Compatible Likelihoods for Flow Matching on Manifolds

Lucas Ng[1], *Georgios Batzolis*[2], *Mark Girolami*[1,2]

[1] University of Cambridge

[2] Devotion AI Labs

IN SHORT

Setting: conditional flow matching on Riemannian manifolds.
Objective: likelihood-based loss functions of the form

$$\mathsf{L}(\theta) = \mathbb{E}_{t,x_0,x_1}\left[-\log q_t^\theta(z \mid x_t)\right].$$

$q_t^\theta(z \mid x_t)$: a chosen family of distributions over latent variable z.
We want *compatibility*: the best loss yields the marginal field,

$$u_t(x_t) = \mathbb{E}_{p_t(x_1|x_t)}\left[u_t^{\mathsf{cond}}(x_t \mid x_1)\right].$$

We prove: $q_t^\theta(z \mid x_t)$ must satisfy a restrictive affine condition.
RG-VFM: sacrifices this to incorporate geometric information.
RFM: satisfies compatibility but puts mass where infeasible.
We propose: a new method that tackles both issues.

COMPATIBILITY THEORY

A loss is *compatible* if, after decoding the network output into a velocity prediction, every population minimizer produces the true marginal velocity. This needs the existence of some z-independent linear map $A_{x,t}^\theta$ and tangent vector $b_{x,t}^\theta$ such that

$$\psi_{x,t}(z) = A_{x,t}^\theta \nabla_\theta \log q_t^\theta(z \mid x_t) + b_{x,t}^\theta, \quad \theta = \theta^*_{\mathsf{unknown}}$$

where ψ maps the likelihood variable z to the conditional velocity. At optima, the error satisfies the locally sharp bound

$$\|\mathbb{E}_p[\psi_{x,t}] - \mathbb{E}_q[\psi_{x,t}]\|_g \le \sqrt{\chi^2(p\|q)}\sqrt{\|\mathbb{E}_q[r_{x,t}\otimes r_{x,t}]\|_{\mathsf{op},g}},$$

where $p = p_t(z \mid x_t)$, $q = q_t^{\theta^*}(z \mid x_t)$, and $r_{x,t}$ is the affine residual.
A similar bound holds for other decoders via triangle inequality.

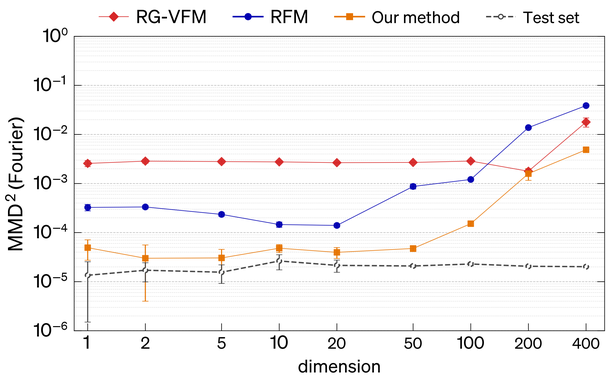


Figure 1. Sample quality across methods on product tori of high dimension.

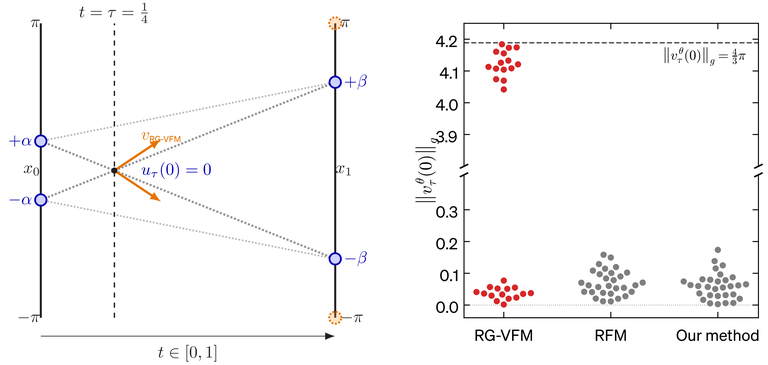


Figure 2. Conditional paths (left) and velocity errors (right) on $\mathbb{T}^1$.

METHODS OLD AND NEW

Riemannian Gaussian variational flow matching (RG-VFM) maximizes a Riemannian Gaussian over endpoints,

$$q_{\text{RG-VFM}}(x_1 \mid x_t) \propto \exp\left\{-\tfrac{1}{2}\sigma^{-2}\,\mathrm{dist}_g(x_1, \mu_t^\theta(x_t))^2\right\}.$$

RG-VFM is not compatible. In Figure 2, both of the global optima push generated samples away from the data distribution. Implicitly, Riemannian flow matching (RFM) uses

$$q_{\text{RFM}}(u \mid x_t) \propto \exp\left\{-\tfrac{1}{2}\|u - v_t^\theta(x_t)\|_g^2\right\},$$

optimizing a Gaussian over velocities in the tangent space. Although compatible, on compact manifolds it assigns density to infeasible velocities. *Our method* uses an exponential family over endpoints with the conditional velocity as sufficient statistic. This marries compatibility with support-awareness:

$$q_t^\theta(x_1 \mid x_t) \propto \exp\left\{\langle \eta, u_t^{\mathsf{cond}}(x_t \mid x_1)\rangle_g - \tfrac{1}{2}\|u_t^{\mathsf{cond}}(x_t \mid x_1)\|_g^2\right\}.$$

Here $\eta = m_{x,t}^{-1}(v_t^\theta(x_t))$, where $m_{x,t}$ is the mean velocity map. It is compatible and supports only the velocities of real endpoints.

PRACTICAL COMPUTATION

With the same gradients as likelihood maximization, training simplifies analytically to a Fisher reweighting of the RFM loss,

$$\mathsf{L}(\theta) = \tfrac{1}{2}\mathbb{E}_{t,x_0,x_1}\left[\left\langle v_t^\theta(x_t) - u_t^{\mathsf{cond}}, \mathsf{sg}(V_{x,t})^{-1}(v_t^\theta(x_t) - u_t^{\mathsf{cond}})\right\rangle_g\right].$$

Here $V_{x,t}$ is the covariance of conditional velocities under q_t^θ, evaluated at the predicted velocity. Symmetry cuts this down to a few scalar functions which can be pretabulated and reused across datasets. On product spaces, factors share one lookup table whose dimension stays fixed as the manifold dimension grows. After compilation, the overhead per step is *negligible*.

Table 1. Geospatial datasets on $\mathbb{S}^2$, test NLL.

	VOLCANO	QUAKE	FLOOD	FIRE
Riemannian CNF	−0.97±0.15	0.19±0.04	0.90±0.03	−0.66±0.05
Moser Flow	−2.02±0.42	−0.09±0.02	0.62±0.04	−1.03±0.03
CNF Matching	−2.38±0.17	−0.38±0.01	**0.25**±0.02	−1.40±0.02
Riemannian Score	−4.92±0.25	−0.19±0.07	0.45±0.17	−1.33±0.06
Riemannian Diffusion	−6.61±0.97	**−0.40**±0.05	0.43±0.07	−1.38±0.05
RFM (Paper results)	−7.93±1.67	−0.28±0.08	0.42±0.05	−1.86±0.11
RFM (Reproduced)	−6.96±1.71	−0.21±0.05	0.48±0.04	−1.77±0.06
Our method	**−9.69**±1.94	−0.28±0.04	0.47±0.04	**−2.11**±0.11

Table 2. Human pose data on SO(3)[19], test metrics.

	NLL	ENERGY	MMD^2
HuProSO3	−137.78	56.76	9.746
RG-VFM	−150.58±0.41	18.60±1.72	3.601±0.308
RFM	**−198.11**±0.81	2.16±0.20	0.379±0.035
Our method	−197.41±0.44	**1.68**±0.04	**0.293**±0.008

When does flow matching on manifolds *actually work?*

NEURAL INFORMATION PROCESSING SYSTEMS

This research is supported through the EPSRC ROSEHIPS grant [EP/W005816/1].