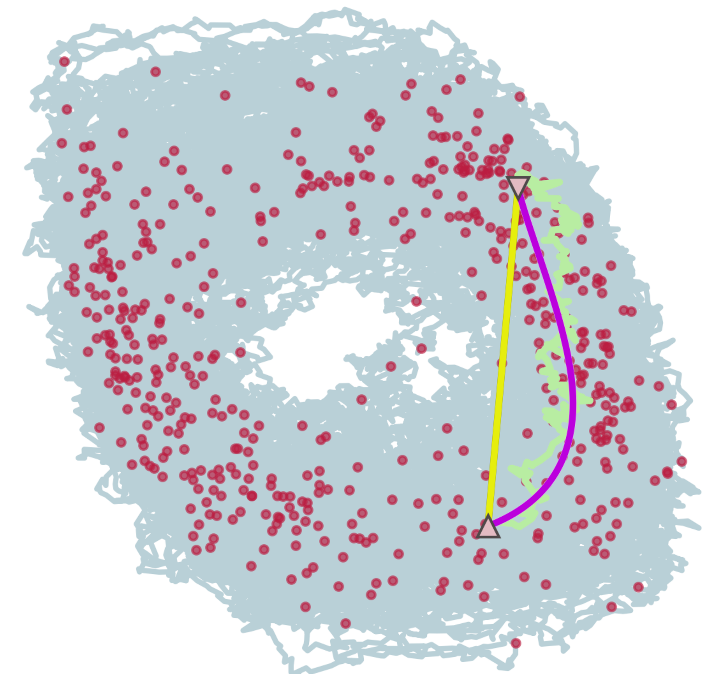
Flow curvature explains failed SDE drift estimation under sparse sampling

ICLR
International Conference On Learning Representations

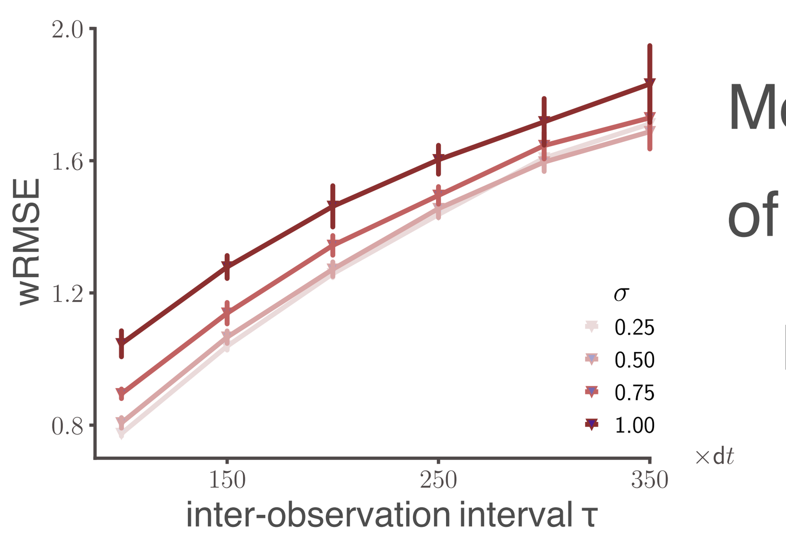
Author
Dimitra Maoutsa
Affiliations
Technical University of Berlin, Berlin, Germany

Workshop on Geometry-grounded Representation Learning and Generative Modeling

We study why **drift estimation** from **sparsely** observed stochastic trajectories fails as the sampling interval grows. We show that standard finite-difference estimators implicitly **underestimate the curvature** of unobserved paths between consecutive observations. This induces a **systematic bias** controlled by the evolution of flow curvature along trajectories, which motivates a **geometric path-augmentation method** based on **geodesic curves on the empirical observation manifold**.



When does drift inference fail?



Most inference methods start from the following discretization of the stochastic dynamics $\mathbf{X}_{t_0+\tau} = \mathbf{X}_{t_0} + \mathbf{f}(\mathbf{X}_{t_0})\,\tau + \sigma(\mathbf{W}_{t_0+\tau} - \mathbf{W}_{t_0}) + \cancel{R_1},$

Even at moderate noise, drift inference degrades for increasing sampling interval

Second-order terms in the Itô–Taylor expansion quantify the bias of the approximation

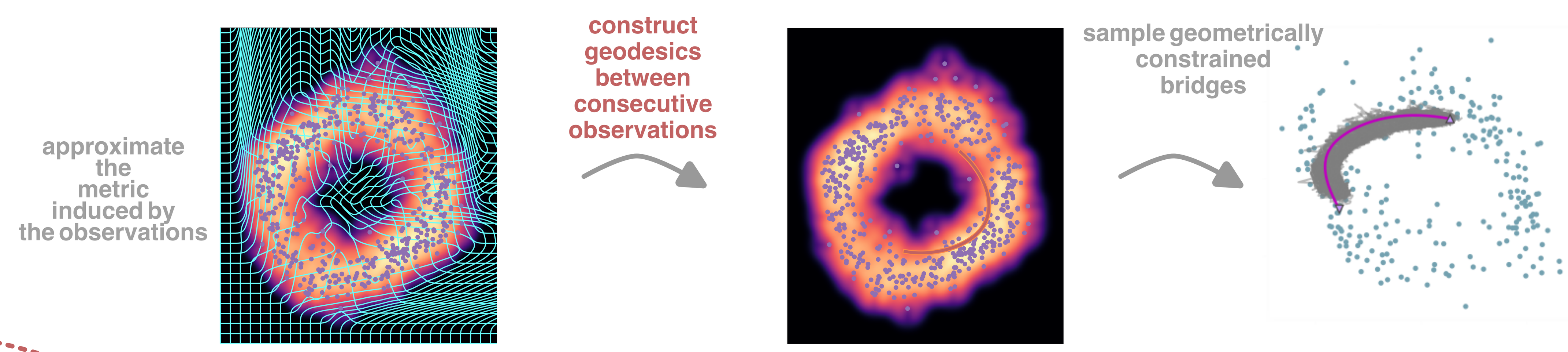
$$\text{bias}(\mathbf{x}) \approx \frac{\tau^2}{2}\left[\underbrace{\nabla(J_{\mathbf{f}}\mathbf{f})\cdot\mathbf{f}}_{\text{flow curvature evolution}} + \nabla\left(\tfrac{1}{2}\Delta_D\mathbf{f}\right)\cdot\mathbf{f}\right]$$

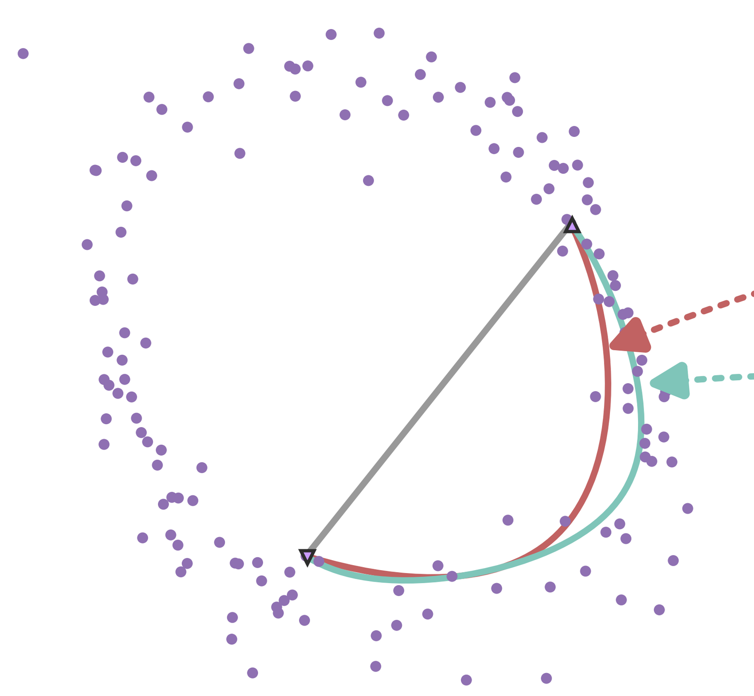
$\nabla(J_{\mathbf{f}}\mathbf{f})\cdot\mathbf{f}$ characterizes how the curvature changes as the state evolves. It is large when the flow bends on short lengthscales.

Why?

We use the empirical manifold as geometric constraint

Link to the paper





Geodesic curves on the empirical manifold provide coarse approximations to the Onsager-Machlup most probale paths

strength of geometric control term

augmentations