Executive summary: Plug-in estimation of Schrodinger bridges

Viresh Pati

January 2025

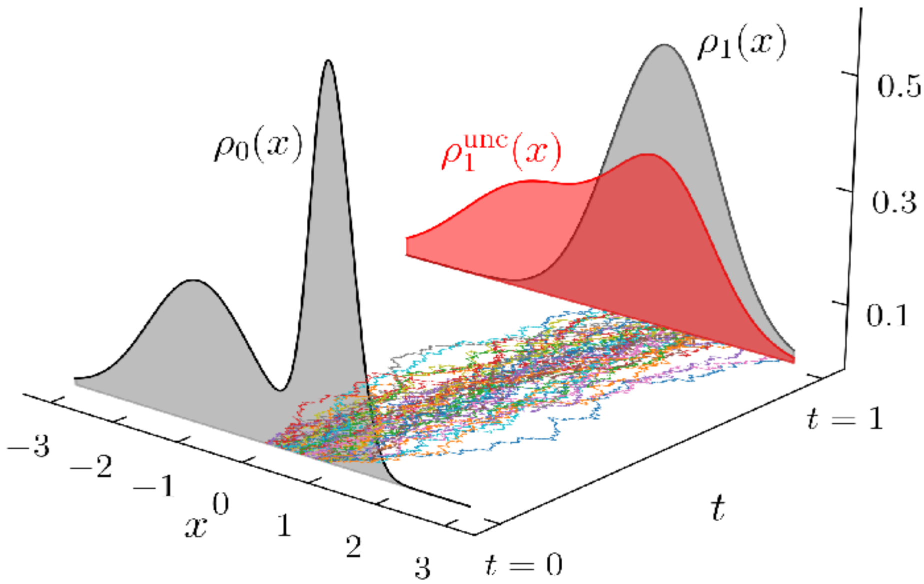
1 Paper Details

Aram-Alexandre Pooladian and Jonathan Niles-Weed. "Plug-in estimation of Schrödinger bridges" *arXiv*.2024. [arxiv]

2 Repository

The authors provide code at [this github] which runs on some benchmarks. We should be able to adapt for NSF challenge.

3 High-level problem



When we have two probability distributions μ and ν , the Schrodinger Bridge gives a most likely stochastic process from source distribution μ to target distribution ν. That is,

- We want to transform samples from μ into samples from ν by simulating a diffusion over time.
- The Schrodinger Bridge is the process that minimizes relative entropy.

Existing methods to estimate that time-dependent diffusion require heavy iterative SDE simulation or neural networks. The authors propose a simple plug-in estimator:

- Solve the static entropic optimal transport between μ and ν once.
- Use these entropic OT potentials to directly plug in a drift formula that defines the Schrodinger bridge over time.

4 Intuition behind the technique

- Entropic OT (via Sinkhorn's Algortim) is already a soft coupling between μ and ν.
- Schrodinger bridge is also an entropic construction in that it is the minimal-entropy stochastic process that connects μ and ν.

- This work shows that the "entropic penalty" that arises in the classical entropic OT problem is the same in the Schrodinger bridge problem, so the dual potentials from Sinkhorn effectively store all the needed information to diffuse between distributions.

5 Results

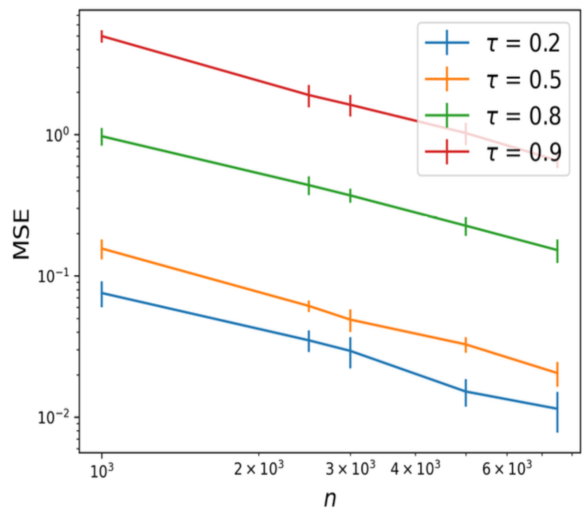


Figure 2: MSE for estimating the Gaussian drift as (n, τ) vary, averaged over 10 trials.

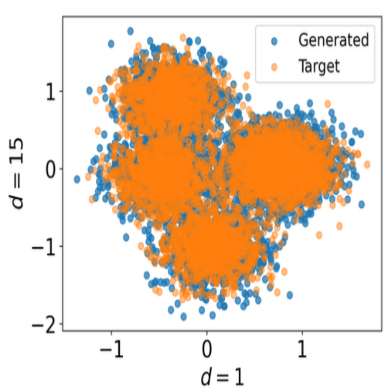


Figure 3: Plotting generated and re-sampled target data in $d = 64$.

Method	BW-UVP
Ours	**0.41 ± 0.03**
MLE-SB	0.56
EgNOT	0.85
FB-SDE-A	0.65

Table 1: Comparison to neural network approaches in BW-UVP for $d = 64$.

6 Some more background and detail

6.1 Optimal Transport

In classical optimal transport, we seek a joint distribution π with marginals μ and ν that minimizes the average cost $|x-y|^2$ of pairing $x \in \mu$ and $y \in \nu$. Entropic OT adds a small entropic penalty that makes π spread out and keeps the problem strictly convex, solvable by Sinkhorn's algorithm.

$$\int |x-y|^2 \pi(dx, dy) + \epsilon \mathrm{KL}(\pi || \mu \otimes v)$$

Note:$\pi(x, y)$ is a coupling, we'll say "how much mass from $x \in \mu$ is shipped to point $y \in \nu$ "

Instead of directly finding joint π we can use a dual-formulation to find two scalar function f that lives on the support of μ and g that lives on the support of ν . We will omit detail for brevity, but Sinkhorn's algorithm updates f and g based on each other until convergence, and you end with approximations for each sample in the distributions.

6.2 From potentials to Schrodinger Bridge

It turns out the potentials from entropic OT encode the correction terms needed to steer a Brownian motion from μ to ν. That is, reading off potentials from the Sinkhorn solution and plugging into a known formula for drift (deterministic part) of an SDE between μ and ν is precisely the Schrodinger Bridge.

7 Application to vessel trajectories

We can treat each data snapshot as a discrete empirical distribution, run EOT and plug dual potentials to get a drift that flows one vessel shape into the other in time. We may want to parameterize multiple time-steps in a neural network, and/or do a latent encoding of raw vessel data to obtain μ and ν.

On runtime: naive Sinkhorn is $O(kn^2)$ where k is the number of iterations to convergence (there are some small optimizations possible and the paper also divides out some constraint tolerances so this is a loose bound). And then there is usually negligible $O(nt)$ drift where t is the discrete number of steps. Sinkhorn's is also highly parallelizable.