A review on the research paper of Trajectory Inference via Mean-Field Langevin in PathSpace

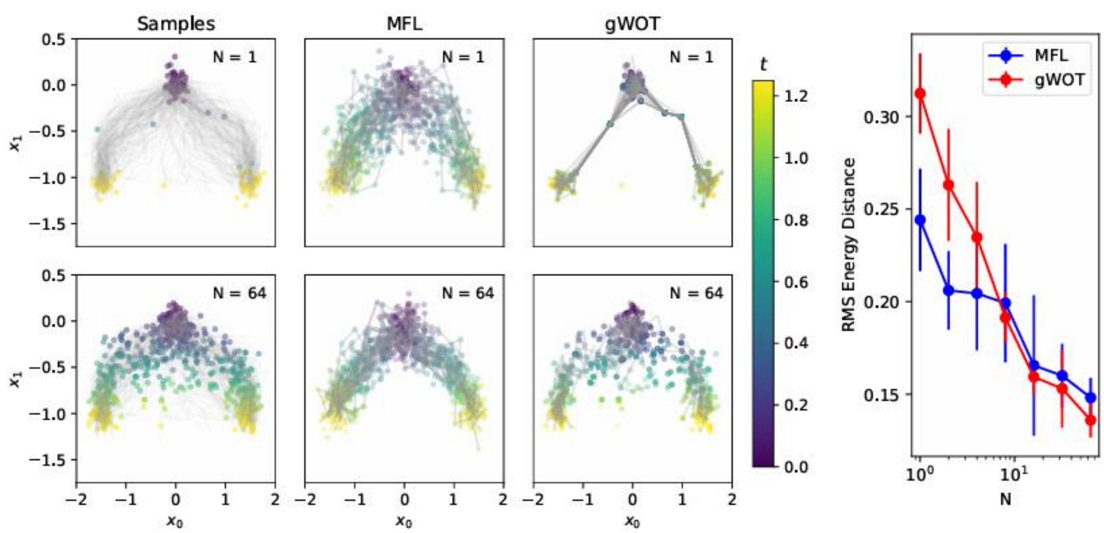
Author: Stephen Zhang (University of Melbourne)   Schiebinger, Matthieu Heitz (UBC, Canada)

This paper is presented in the 36th conference on Neural Information Processing Systems (NeurIPS 2022)

Overall, this research is aimed at improving the model from Lavenant et al(1), which is the first research paper (Mathematical theory towards Trajectory inference).

Problem they are trying to solve:

- In the first research paper, it gave us an estimator relative to the Wiener measure in path space, and they consistently reconstruct the dunamics from dimensional convex optimization problem**.** Still, the problem is they rely on structured grids**,** makes sets of points discrete -> so there are gaps in sample points, and the model (gWOT) would struggle to capture the dynamics.
- MFL, on the other hand, will try to use grid-free, point clouds (one per snap shot) with Schrodinger-bridges.
- This paper want to show the specific structure of Lavenant et al estimator(gWOT) is to make to amenable to the new stochastic methods, MFL.



gWOT works on fixed, discretized suppoer: union of observed sample points -> they perform poorly when the support has missing regions or gaps. Output of MFl has small difference between N = 1 and N = 64. MFL outperform gWOT in smaller N.

How did they suppose to do that?

MFL use particles that move continuosly in the space, instead of fixed grids of points. These particles evolve based on diffusion and deterministic movements towards optimal solutions. In gWOT, sparse data causes gaps in discretized grid -> but in MFL, they fill in the gaps using optimal transport principles.

⇨ In overall, they replace the gWOTs with a more flexible, particle-based representation, they reduced formulation, make the model robust to sparse data.

Numerical Experiments explained:

This experiements focus a lot in cells and biology, where they take the branching in this field as a machine learning experiment.

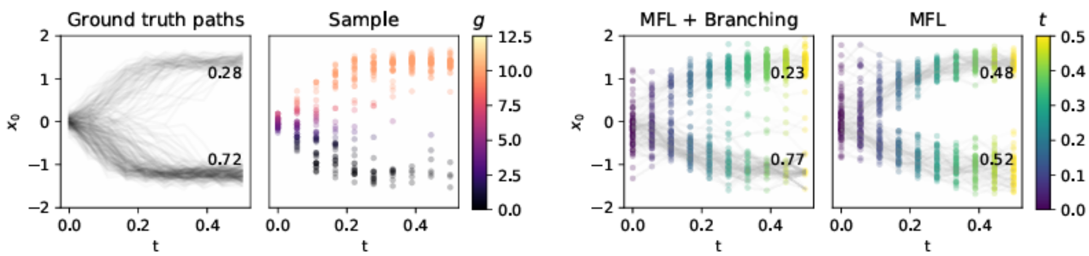
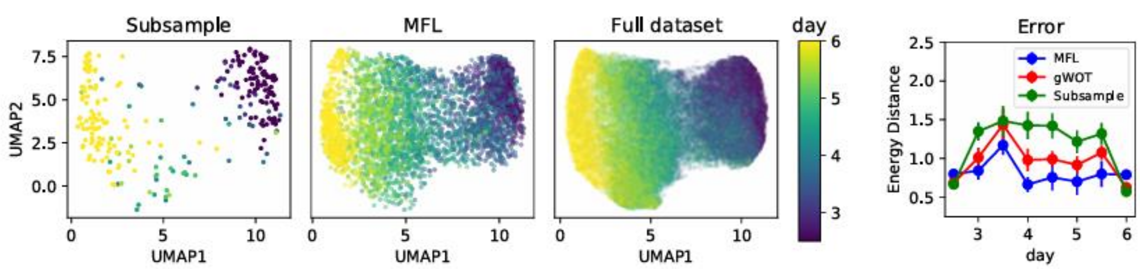


Figure 3: Accounting for branching rates allows for the separation of spatial dynamics from growth. (left) Ground truth paths and sample points. (right) Reconstruction produced by MFL dynamics with and without accounting for branching. Fraction of paths terminating in the upper (resp. lower) branch is annotated in the plot.

They have the ground truth path with 2 branching (upper ones going up and the down ones). If they account for MFL with the explicitly modeling branching, they would get close to ground truth paths with 77% (near 72%) for down branching.



(the experiment starts with 100 cells at first, and with N = 10 (10 cells each timepoint), best performances as measured by Energy Distance to full data sets of gWOT and MFL)

Research advantage and disadvantage for our project:

- This deals a lot with sparse and noisy data, which are much like vessel data.
- MFL would be a great fit since it may adapt well to incomplete or even data, more flexible for continuos particle flow.
- It may help high-demensional data.
- Still have to be careful with reconstructed tracks to make sense.