Using Linearized Optimal Transport to Predict the Movement of Stochastic Particle Systems

Nicholas Karris[1], **Evangelos A. Nikitopoulos**[2], **Ioannis Kevrekidis**[3], **Seungjoon Lee**[4], and **Alexander Cloninger**[5],

[1,5]University of California, San Diego, [2]University of Michigan, [3]Johns Hopkins University, [4]California State University, Long Beach

Simulating Particle Systems

General Set Up: Have collection of particles $\{x_i(t)\}_{i=1}^N \subset \mathbb{R}^d$ that form an empirical distribution μ_t^N, and have some dynamics that describe their evolution via

$$\mu_{t+h}^N \approx F_h(\mu_t^N).$$

for some noisy F_h. We can numerically simulate the particle system by successively applying F_h for some small h.

Goal: Want to use μ_t^N and μ_{t+h}^N to approximate μ_{t+H}^N without needing to apply F repeatedly. If $H \gg h$, this drastically reduces computation time for simulating many steps.

Idea: Use an analog of Euler's method:

$$\gamma(t+H) \approx \gamma(t) + H\gamma'(t)$$

for distributions μ_t on the Wasserstein manifold.

Why Optimal Transport?

Motivating Theorem: [1] Let $t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ be absolutely continuous in the Wasserstein-2 distance. Then there is a unique minimal-energy vector field $\mathbf{v}_t : \mathbb{R}^d \to \mathbb{R}^d$ such that the continuity equation

$$\partial_t \mu_t + \nabla \cdot (\mathbf{v}_t \mu_t) = 0$$

holds in the sense of distributions. Moreover, for $\mathcal{L}^1$-a.e. t,

$$\lim_{H\to 0} \frac{W_2(\mu_{t+H}, (\mathrm{id} + H\mathbf{v}_t)_{\#}\mu_t)}{|H|} = 0 \qquad (*)$$

and, assuming $\mu_t \ll \mathcal{L}^d$,

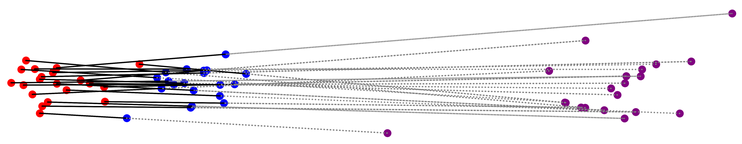
$$\lim_{h\to 0} \frac{1}{h}(T_{\mu_t}^{\mu_{t+h}} - \mathrm{id}) = \mathbf{v}_t \qquad \text{in } L^2(\mu_t). \qquad (\dagger)$$

Approximation Error: As with Euler's method in $\mathbb{R}^d$, the error depends on the "second derivative" of the curve $t \mapsto \mu_t$. Under certain regularity assumptions, the error can be bounded by

$$W_2(\mu_{t+h}, (\mathrm{id} + H\mathbf{v}_t)_{\#}\mu_t) \le \frac{H^2}{2}\left(\int_{\mathbb{R}^d} \max_{r\in[t,t+h]} \left\| \frac{d}{ds}\tilde{\mathbf{v}}_{r+s}(x)\right\|_2^2 d\mu_t(x)\right)^{1/2}.$$

for a slightly modified vector field $\tilde{\mathbf{v}}$.

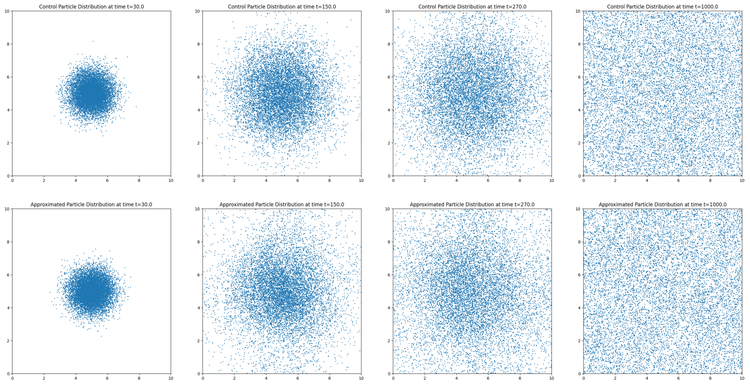
Discrete Algorithm



Key Idea: Take an Euler-step with each particle, where the "velocity" of the particle is given by the optimal-transport approximation of the velocity field.

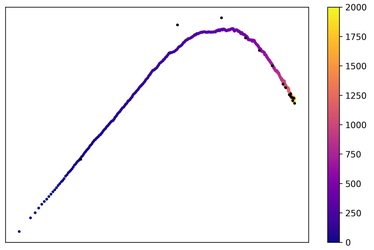
Experimental Results

Gaussian Diffusion: Results of applying the algorithm on a simulation of particles diffusing in two dimensions.

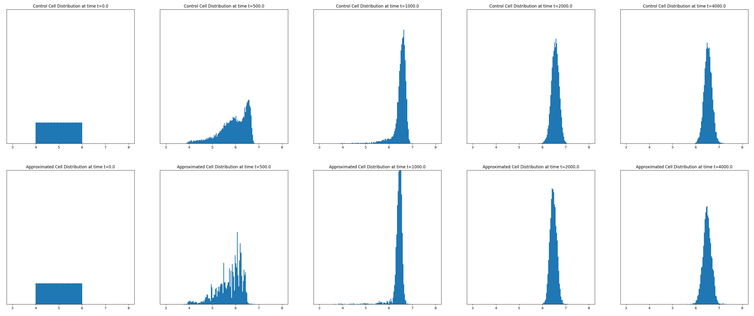


We perform PCA to visualize how well our method approximates the distribution's true evolution.

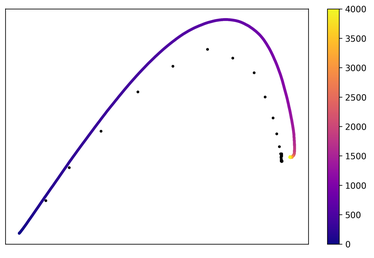
- Color: Control (true) evolution using Euler-Maruyama
- Black: Approximated evolution using discrete approximation algorithm



Cell Chemotaxis: We also apply the algorithm on a cell chemotaxis simulation that models cells responding to a chemoattractant, as described in [3].



PCA plot to compare the approximations with the control cell distributions. The approximations appear slightly behind the true curve because some underlying cell variables take longer to equilibriate after each Euler-step.

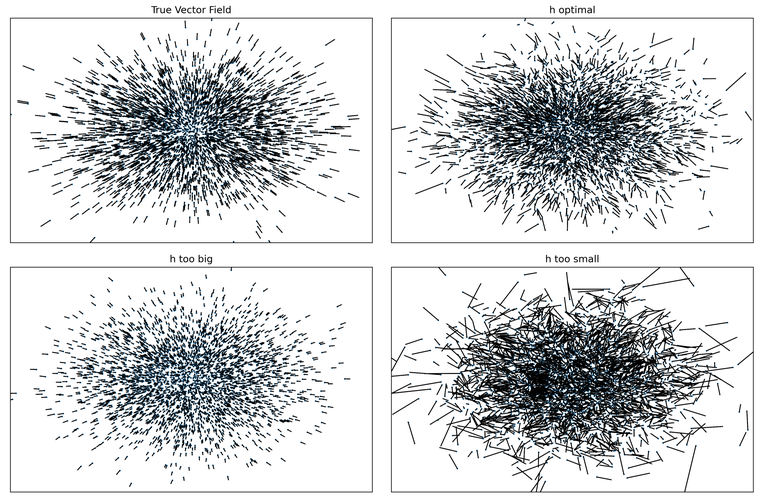


Choosing an Optimal Step Size

We approximate the velocity field

$$\mathbf{v}_t(x_i) \approx \frac{1}{h}(T_{\mu_t}^{\mu_{t+h}}(x_i) - x_i)$$

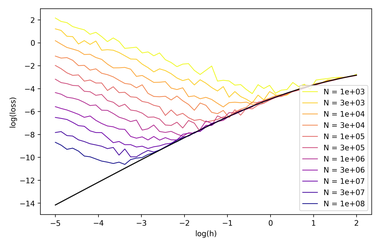
for an appropriate h. Why not take $h \to 0$?



The approximation error

- is large for small h
- approaches the theoretical finite difference error for large h

Empirically, there is an optimal choice for h that decreases as N increases.



Connections to Fast-Slow Systems

Application: We get a "time stepper" that sits between the two time scales of particle systems. We can numerically integrate on the long scale, which reduces the number of short scale steps needed to observe the long-term behavior.

Ongoing Work: (1) How does the vector field error depend on step size h and number of particles N? (2) Can we estimate the error without knowing the underlying (theoretical) vector field? (3) Can we numerically estimate the second-derivative bound based on simulation data?

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, 2. ed edition, 2008. OCLC: 254181287.

[2] Nicholas Karris, Evangelos A. Nikitopoulos, Ioannis Kevrekidis, Seungjoon Lee, and Alexander Cloninger. Using linearized optimal transport to predict the evolution of stochastic particle systems, 2024.

[3] S. Setayeshgar, C. W. Gear, H. G. Othmer, and I. G. Kevrekidis. Application of coarse integration to bacterial chemotaxis. *Multiscale Modeling & Simulation*, 4(1):307–327, 2005.