Propagation of Chaos for Conditional McKean—Vlasov Equations

Manuel Arnese

Columbia University

Problem Statement

- We want a particle system for the coupled 2-dimensional SDE system

$$dX_t = \mathbb{E}[\, a(X_t, Y_t) \mid X_t \,]dt + U(X_t, Y_t)dt + dW_t$$

$$dY_t = \mathbb{E}[\, b(X_t, Y_t) \mid Y_t \,]dt + V(X_t, Y_t)dt + dB_t$$

- Non-Linearity through conditional expectations.

- Applications: SDEs on large graphs, Molecular Dynamics, Finance...

Motivation: Wasserstein Gradient Flow

- Fix probability measures on $\mathbb{R}$

$$e^{-U}, e^{-V}$$

- Fix a cost function $c: \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$

Motivation: Wasserstein Gradient Flow

- Conforti-Lacker-Pal (2025):

$$dX_t = \Big(\mathbb{E}[\,\nabla_x c(X_t, Y_t) \mid X_t\,] - \nabla_x c(X_t, Y_t)\Big)dt - \nabla U(X_t)dt + dW_t$$

$$dY_t = \Big(\mathbb{E}\big[\,\nabla_y c(X_t, Y_t) \mid Y_t\,\big] - \nabla_y c(X_t, Y_t)\Big)\,dt - \nabla V(Y_t)dt + dB_t$$

Motivation: Wasserstein Gradient Flow

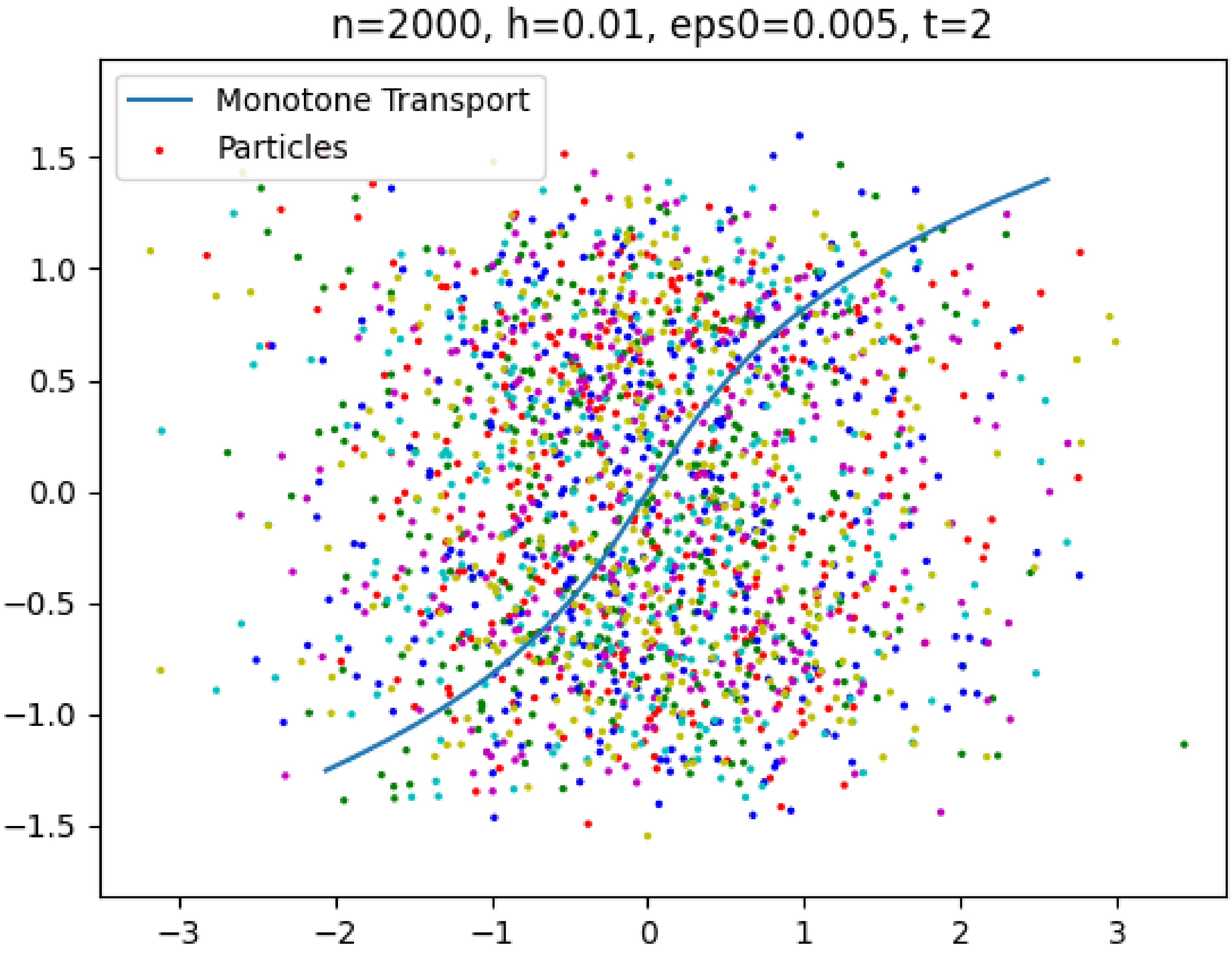
- Conforti-Lacker-Pal (2025):

$$dX_t = \big(\mathbb{E}[\, \nabla_x c(X_t, Y_t) \mid X_t \,] - \nabla_x c(X_t, Y_t)\big) dt - \nabla U(X_t) dt + dW_t$$

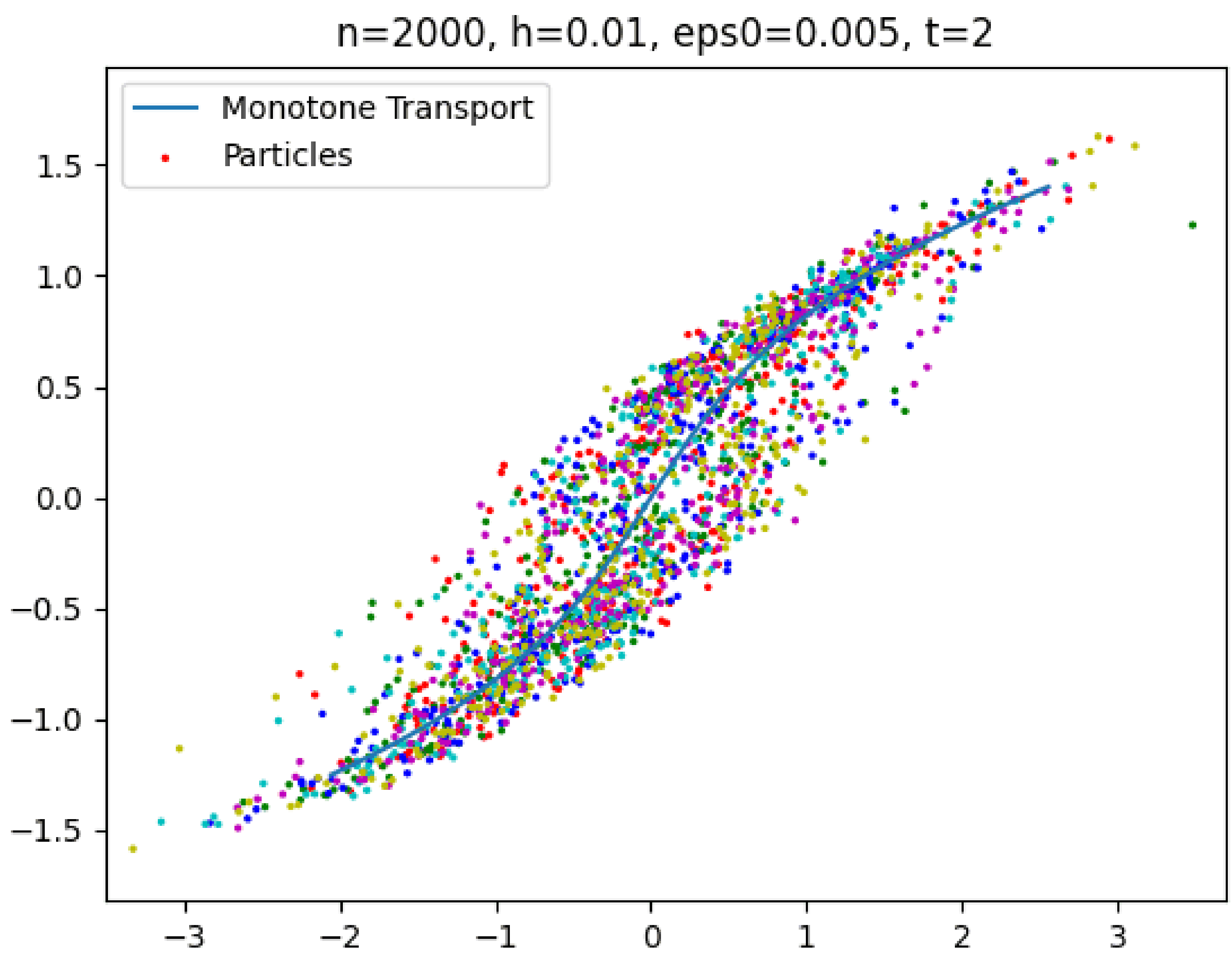
$$dY_t = \Big(\mathbb{E}\big[\, \nabla_y c(X_t, Y_t) \mid Y_t \,\big] - \nabla_y c(X_t, Y_t)\Big) dt - \nabla V(Y_t) dt + dB_t$$

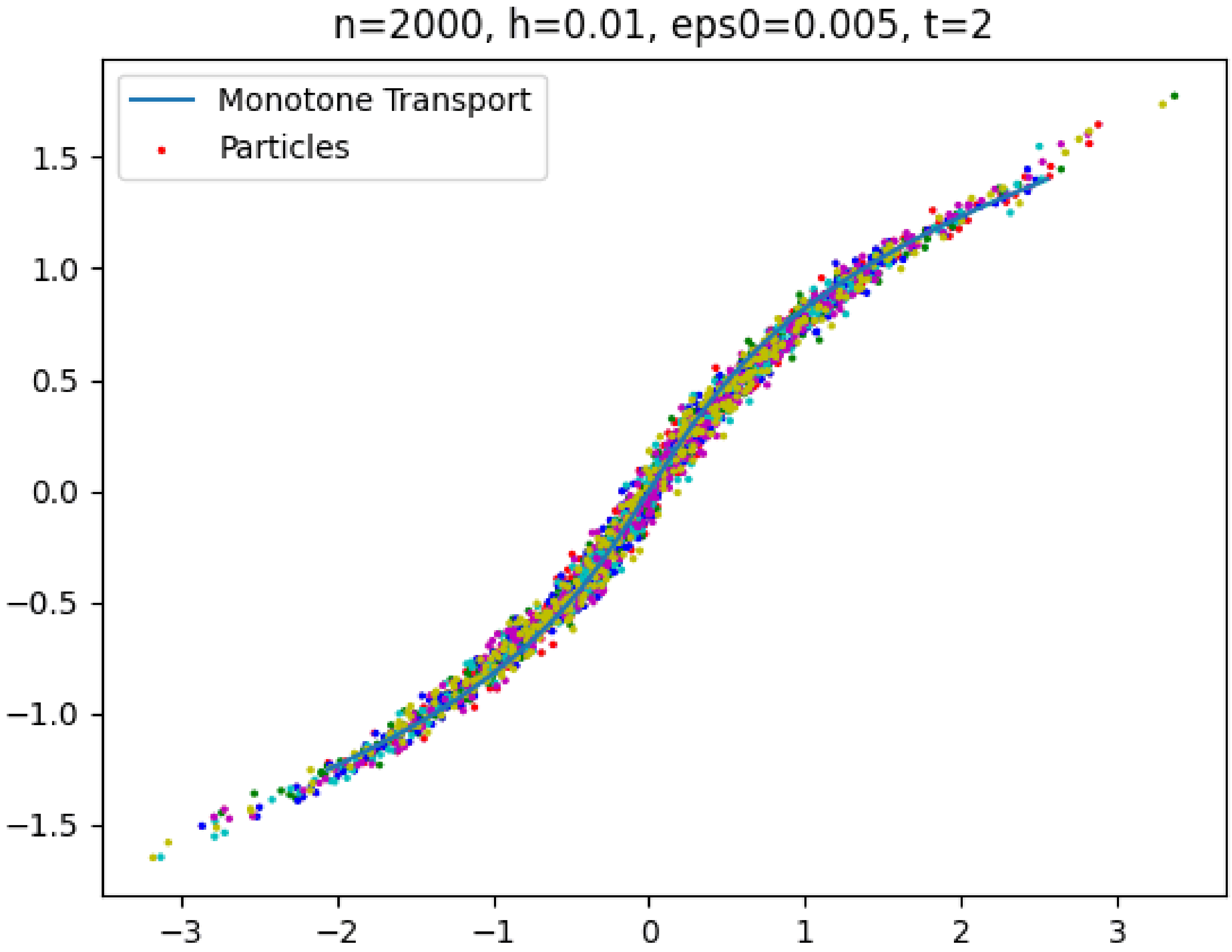
- $P_t = \text{Law}(X_t, Y_t)$ converges to solution of entropic OT:

$$P_t \to \min_{\pi \in \Pi(e^{-U}, e^{-V})} \int c(x, y) d\pi(x, y) + H(\pi | e^{-U} \otimes e^{-V})$$



n=2000, h=0.01, eps0=0.005, t=2
Monotone Transport
Particles
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−3
−2
−1
0
1
2
3


n=2000, h=0.01, eps0=0.005, t=2
Monotone Transport
Particles
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−3
−2
−1
0
1
2
3



n=2000, h=0.01, eps0=0.005, t=2
Monotone Transport
Particles
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−3
−2
−1
0
1
2
3

Particle System

- Consider vectors $\boldsymbol{X} = \left(X^i\right)_{i=1}^n, \boldsymbol{Y} = \left(Y^i\right)_{i=1}^n$;

- Nadaraya-Watson estimator for conditional expectations:

$$\hat{a}(X_t^i|\boldsymbol{X}_t, \boldsymbol{Y}_t) = \frac{\frac{1}{n}\sum_j a\left(X_t^i, Y_t^j\right)K_h(X_t^i - X_t^j)}{\frac{1}{n}\sum_j K_h\left(X_t^i - X_t^j\right) + \varepsilon}$$

Particle System

- Moderately interacting particle system $(\boldsymbol{X}_t, \boldsymbol{Y}_t) = \left(X_t^i, Y_t^i\right)_{i=1}^n$:

$$dX_t^i = \hat{a}(X_t^i \,|\boldsymbol{X}_t, \boldsymbol{Y}_t)dt + U\left(X_t^i, Y_t^i\right)dt + dW_t^i$$

$$dY_t^i = \hat{b}(Y_t^i \,|\boldsymbol{X}_t, \boldsymbol{Y}_t)dt + V\left(X_t^i, Y_t^i\right)dt + dB_t^i$$

- Two free parameters: h and ε

Main Result

- Let π law of the particle system, first k marginals are π^k; μ law of the non-linear limit

- Assume initial condition and coefficients are nice (a, b bounded). Convergence in Total Variation (arbitrary $p < 1$):

$$\left\|\pi_t^k - \mu_t^{\otimes k}\right\|_{TV} \lesssim \frac{e^{\frac{1}{\varepsilon^2 h}}}{\varepsilon\sqrt{h}} \cdot \sqrt{\frac{k}{n}} + \sqrt{k}\left(h + \varepsilon^{\frac{p}{2}}\right)$$

Proof Sketch:

- Take ρ to be the large n limit of π^1 for a fixed $h > 0$.

- Triangle inequality:

$$\left\|\pi_t^k - \mu_t^{\otimes k}\right\|_{TV} \leq \left\|\pi_t^k - \rho_t^{\otimes k}\right\|_{TV} + \left\|\rho_t^{\otimes k} - \mu_t^{\otimes k}\right\|_{TV}$$

Proof Sketch:

- Take ρ to be the large n limit of π^1 for a fixed $h > 0$.

- Triangle inequality:

$$\left\|\pi_t^k - \mu_t^{\otimes k}\right\|_{TV} \le \underbrace{\left\|\pi_t^k - \rho_t^{\otimes k}\right\|_{TV}}_{\text{Variance}} + \underbrace{\left\|\rho_t^{\otimes k} - \mu_t^{\otimes k}\right\|_{TV}}_{\text{Bias}}$$

Proof Sketch:

- The Variance term is (almost) a standard Propagation of chaos problem.

- Bias term: information theoretic inequalities to show for some $p < 1$

$$\|\rho_t - \mu_t\|^2{}_{TV} \lesssim \varepsilon^p + \int_0^t H(\mu_s | \mu_s \star K_h) ds$$

Regularity

- Informal calculation:

$$H(\mu_s|\mu_s \star K_h) \lesssim h^2(1 + \mathbb{E}[|\nabla^2 \log \mu_t(X_t)|^3])$$

- Conditional Expectations a priori rough. Pointwise estimates difficult, integrated control reachable.

$$|\nabla^k \mathbb{E}[a(X_t, Y_t) \mid X_t]| \lesssim \mathbb{E}\left[\left|\nabla_{\mathrm{x}}^k \log \mu_t\,(X_t, Y_t)\right| \mid X_t\right] + \text{lower order terms}$$

- Idea: bootstrap argument and energy estimates.

Thank you!