Neural Schrödinger Bridge with Sinkhorn Losses

Charlie Yan

Department of Electrical and Computer Engineering

University of California Santa Cruz, CA 95064

Advisor: Prof. Abhishek Halder

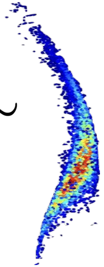
Committee: Profs. Dejan Milutinović, Qi Gong

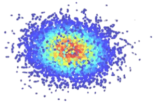
MS Thesis Presentation, May 18, 2023

Problem: Optimal Steering of Stochastic State

Motivational application #1

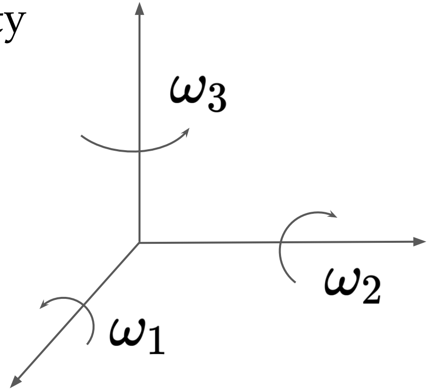
 $\rightsquigarrow$ 

$\boldsymbol{\omega}(t=0) \sim$ 

$\boldsymbol{\omega}(t=T) \sim$ 

Stochastic state:

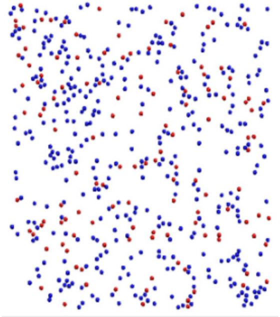
Angular velocity



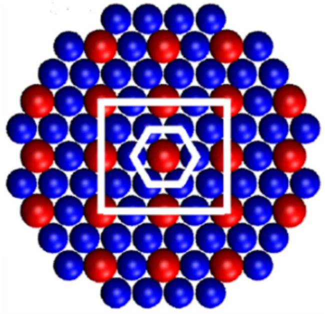
Distributional regulation of stochastic spin over given horizon length T

Problem: Optimal Steering of Stochastic State

Motivational application #2



Dispersed particles



Ordered structure

Stochastic state:

Measure of crystallinity order

$(\langle C_{10} \rangle, \langle C_{12} \rangle)$ Steinhart bond order parameters

Data-driven controlled colloidal self-assembly for precision (sub nm scale) manufacturing

Generalized Optimal Mass Transport (GOMT)

Stochastic Optimal Control Problem:

$$\underset{(\rho^{\boldsymbol{u}},\boldsymbol{u})\in\mathcal{P}_2(\mathbb{R}^n)\times\mathcal{U}}{\arg\inf}\int_0^T\int_{\mathbb{R}^n}\left(q(\boldsymbol{x}^{\boldsymbol{u}})+r(\boldsymbol{u})\right)\ \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t)\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}}\mathrm{d}t$$

$$\boxed{\frac{\partial\rho^{\boldsymbol{u}}}{\partial t}+\nabla_{\boldsymbol{x}^{\boldsymbol{u}}}\cdot\left(\rho^{\boldsymbol{u}}\boldsymbol{f}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\right)=0,}$$

Liouville PDE

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t=0)=\rho_0,\quad \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t=T)=\rho_T.$$

Generalized Optimal Mass Transport (GOMT)

Stochastic Optimal Control Problem:

$$\underset{(\rho^{\boldsymbol{u}},\boldsymbol{u})\in\mathcal{P}_2(\mathbb{R}^n)\times\mathcal{U}}{\arg\inf}\int_0^T\int_{\mathbb{R}^n}\left(q(\boldsymbol{x}^{\boldsymbol{u}})+r(\boldsymbol{u})\right)\,\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t)\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}}\mathrm{d}t$$

$$\boxed{\frac{\partial\rho^{\boldsymbol{u}}}{\partial t}+\nabla_{\boldsymbol{x}^{\boldsymbol{u}}}\cdot\left(\rho^{\boldsymbol{u}}\boldsymbol{f}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\right)=0,}$$

Liouville PDE

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t=0)=\rho_0,\quad \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t=T)=\rho_T.$$

Sample path dynamics associated with the Liouville PDE: $\dot{\boldsymbol{x}}^{\boldsymbol{u}}=\boldsymbol{f}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)$

Classical OMT (Benamou-Brenier, 1999): $q(\cdot)\equiv 0,\quad r(\cdot)\equiv\frac{1}{2}\|\cdot\|_2^2,\quad \boldsymbol{f}\equiv\boldsymbol{u}$

Generalized Schrödinger Bridge Problem (GSBP)

Stochastic Optimal Control Problem:

Fokker-Planck-Kolmogorov (FPK) PDE

$$\inf_{(\rho^{\boldsymbol{u}},\boldsymbol{u})\in\mathcal{P}_2(\mathbb{R}^n)\times\mathcal{U}} \int_0^T \int_{\mathbb{R}^n} \left(q(\boldsymbol{x}^u) + r(\boldsymbol{u})\right) \rho^{\boldsymbol{u}}(\boldsymbol{x},t)\, \mathrm{d}\boldsymbol{x}\mathrm{d}t$$

$$\boxed{\frac{\partial \rho^{\boldsymbol{u}}}{\partial t} + \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \cdot \left(\rho^{\boldsymbol{u}} \boldsymbol{f}\left(t, \boldsymbol{x}^{\boldsymbol{u}}, \boldsymbol{u}\right)\right) = \beta^{-1} \langle \boldsymbol{G}\left(t, \boldsymbol{x}^{\boldsymbol{u}}, \boldsymbol{u}\right), \nabla^2_{\boldsymbol{x}^{\boldsymbol{u}}} \rho^{\boldsymbol{u}} \rangle,}$$

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=0) = \rho_0, \quad \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=T) = \rho_T,$$

Generalized Schrödinger Bridge Problem (GSBP)

Stochastic Optimal Control Problem:

Fokker-Planck-Kolmogorov (FPK) PDE

$$\inf_{(\rho^{\boldsymbol{u}},\boldsymbol{u})\in\mathcal{P}_2(\mathbb{R}^n)\times\mathcal{U}} \int_0^T \int_{\mathbb{R}^n} \left(q(\boldsymbol{x}^u) + r(\boldsymbol{u})\right) \rho^{\boldsymbol{u}}(\boldsymbol{x},t)\, \mathrm{d}\boldsymbol{x}\mathrm{d}t$$

$$\boxed{\frac{\partial \rho^{\boldsymbol{u}}}{\partial t} + \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \cdot \left(\rho^{\boldsymbol{u}} \boldsymbol{f}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\right) = \beta^{-1} \langle \boldsymbol{G}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right), \nabla^2_{\boldsymbol{x}^{\boldsymbol{u}}} \rho^{\boldsymbol{u}} \rangle,}$$

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=0) = \rho_0, \quad \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=T) = \rho_T,$$

Diffusion tensor

$$\left(\boldsymbol{g}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\right)^\top \boldsymbol{g}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)$$

Sample path dynamics associated with the FPK PDE:

$$\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}} = \boldsymbol{f}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\, \mathrm{d}t + \sqrt{2\beta^{-1}}\, \boldsymbol{g}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right) \mathrm{d}\boldsymbol{w}$$

Classical SBP (Schrödinger, 1931-32): $q(\cdot) \equiv 0, \quad r(\cdot) \equiv \frac{1}{2}\|\cdot\|_2^2, \quad \boldsymbol{f} \equiv \boldsymbol{u}, \quad \boldsymbol{g} \equiv \boldsymbol{I}_n$

Solution Structure for Minimum Energy Generalized Schrödinger Bridge Problem (MEGSBP)

GSBP with $r(\cdot) \equiv \frac{1}{2}\|\cdot\|_2^2$

Let $\boldsymbol{G}_\beta := \beta^{-1}\boldsymbol{G}$

Conditions for Optimality: 3 coupled PDEs + endpoint BCs

$$\boxed{\frac{\partial \psi}{\partial t} = q + \frac{1}{2}\|\boldsymbol{u}_{\mathrm{opt}}\|_2^2 - \langle \nabla_{\boldsymbol{x}^{\boldsymbol{u}}}\psi, \boldsymbol{f}\rangle - \langle \boldsymbol{G}_\beta, \mathbf{Hess}(\psi)\rangle}$$

HJB PDE

$$\boxed{\frac{\partial \rho^{\boldsymbol{u}}_{\mathrm{opt}}}{\partial t} = -\nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \cdot \left(\rho^{\boldsymbol{u}}_{\mathrm{opt}}\boldsymbol{f}\right) + \langle \boldsymbol{G}_\beta, \mathbf{Hess}(\rho^{\boldsymbol{u}}_{\mathrm{opt}})\rangle,}$$

FPK PDE

$$\boxed{\boldsymbol{u}_{\mathrm{opt}} = \nabla_{\boldsymbol{u}_{\mathrm{opt}}}\left(\langle \nabla_{\boldsymbol{x}^{\boldsymbol{u}}}\psi, \boldsymbol{f}\rangle + \langle \boldsymbol{G}_\beta, \mathbf{Hess}(\psi)\rangle\right).}$$

Policy PDE

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=0) = \rho_0,$$

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=T) = \rho_T.$$

Existing Literature on Solving MEGSBP

- [Caluya, Halder TAC 2022] Control affine + gradient/mixed conservative-dissipative drift

- [Caluya, Halder ACC 2021] Ditto + hard path constraints

- [Cauya, Halder ACC 2020, Haddad *et. al.* L-CSS 2020] Full state feedback linearizable

- [Nodozi, Halder CDC 2022] Control affine 1st/2nd order nonuniform Kuramoto drift

- [Nodozi *et. al.* ACC 2023] Control non-affine model-based colloidal self-assembly

No existing works on: control non-affine drift and diffusion, model free setting

Contribution of this Work

Neural Schrödinger Bridge: computational framework to learn the solution of MEGSBP

High level idea:

Train PINNs to learn the solution of 3 coupled PDEs + endpoint BCs

Challenges:

PINNs work …

Contribution of this Work

Neural Schrödinger Bridge: computational framework to learn the solution of MEGSBP

High level idea:

Train PINNs to learn the solution of 3 coupled PDEs + endpoint BCs

Challenges:

PINNs work … with carefully set up simulation

Our BCs enforce exact PDF constraints: MSE between PDFs makes less sense

Idea: Regularized Wasserstein Losses for BCs

- Squared Wasserstein metric: $W^2(\mu_0, \mu_1) := \inf\limits_{\pi \in \Pi_2(\mu_0,\mu_1)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \, \mathrm{d}\pi(\boldsymbol{x}, \boldsymbol{y})$

- But PINN training with Wasserstein loss requires differentiating through large scale LP

- Entropy regularized squared Wasserstein distance a.k.a. Sinkhorn divergence:

$$W_\varepsilon^2(\mu_0, \mu_1) := \inf_{\pi \in \Pi_2(\mu_0,\mu_T)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \{\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \varepsilon \log \pi(\boldsymbol{x}, \boldsymbol{y})\} \mathrm{d}\pi(\boldsymbol{x}, \boldsymbol{y})$$

- As $\varepsilon \downarrow 0$, we have $W_\varepsilon^2 \longrightarrow W^2$

Sinkhorn Divergence: Discrete Version

Euclidean distance matrix

$$W_\varepsilon^2(\mu_0, \mu_1) = \min_{\boldsymbol{M}} \left\{ \langle \boldsymbol{C}, \boldsymbol{M} \rangle + \varepsilon \langle \boldsymbol{M}, \log \boldsymbol{M} \rangle \right\}$$

$$\text{subject to} \quad \boldsymbol{M}\mathbf{1} = \mu_0, \quad \boldsymbol{M}^\top \mathbf{1} = \mu_1$$

- Solution via Sinkhorn iteration (iterative matrix scaling):

  Let $\boldsymbol{\Gamma} := \exp(-\boldsymbol{C}/2\varepsilon)$

  Then $\boldsymbol{M}_{\text{opt}} = \operatorname{diag}(\boldsymbol{u})\, \boldsymbol{\Gamma} \operatorname{diag}(\boldsymbol{v})$

  Iterate (guaranteed linear convergence):

$$\boldsymbol{u}^{k+1} = \mu_0 \oslash \left(\boldsymbol{\Gamma} \boldsymbol{v}^k \right)$$
$$\boldsymbol{v}^{k+1} = \mu_1 \oslash \left(\boldsymbol{\Gamma}^\top \boldsymbol{u}^{k+1} \right)$$

Sinkhorn Divergence: Discrete Version

$$W_\varepsilon^2(\mu_0, \mu_1) = \min_{\boldsymbol{M}} \left\{ \langle \underbrace{\boldsymbol{C}}^{\text{Euclidean distance matrix}}, \boldsymbol{M} \rangle + \varepsilon \langle \boldsymbol{M}, \log \boldsymbol{M} \rangle \right\}$$

$$\text{subject to} \quad \boldsymbol{M}\mathbf{1} = \mu_0, \quad \boldsymbol{M}^\top \mathbf{1} = \mu_1$$

- Solution via Sinkhorn iteration (iterative matrix scaling):

  Let $\boldsymbol{\Gamma} := \exp(-\boldsymbol{C}/2\varepsilon)$

  Then $\boldsymbol{M}_{\text{opt}} = \operatorname{diag}(\boldsymbol{u})\, \boldsymbol{\Gamma} \operatorname{diag}(\boldsymbol{v})$

  Iterate (guaranteed linear convergence):

$$\boldsymbol{u}^{k+1} = \mu_0 \oslash \left(\boldsymbol{\Gamma} \boldsymbol{v}^k\right)$$
$$\boldsymbol{v}^{k+1} = \mu_1 \oslash \left(\boldsymbol{\Gamma}^\top \boldsymbol{u}^{k+1}\right)$$

Can compute via autodiff:

$$\nabla_{\boldsymbol{\theta}} W_\varepsilon^2 \left(\mu_0, \mu_0^{\text{epoch index}}(\boldsymbol{\theta})\right)$$

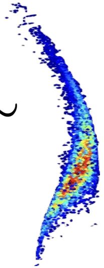
$$\nabla_{\boldsymbol{\theta}} W_\varepsilon^2 \left(\mu_T, \mu_T^{\text{epoch index}}(\boldsymbol{\theta})\right)$$

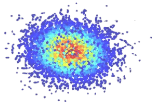
Case Study #1

Optimal Steering of Angular Velocity Distribution

Recap: Optimal Steering of Stochastic State

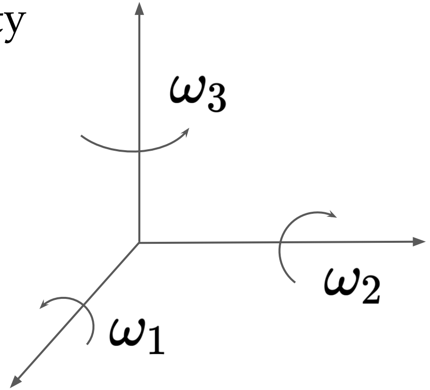
$\rightsquigarrow$

$\boldsymbol{\omega}(t = 0) \sim$ 

$\boldsymbol{\omega}(t = T) \sim$ 

Stochastic state:

Angular velocity



Distributional regulation of stochastic spin over given horizon length T

Euler Equation: Deterministic Controlled Dynamics

$$\boldsymbol{J}\dot{\boldsymbol{\omega}} = -[\boldsymbol{\omega}]^{\times}\boldsymbol{J}\boldsymbol{\omega} + \boldsymbol{\tau}$$

Principal moment of inertia matrix

Control torque

$$[\boldsymbol{\omega}]^{\times} := \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \in \mathfrak{so}(3)$$

Rewrite as:

$$\dot{\boldsymbol{x}}^{\boldsymbol{u}} = \boldsymbol{\alpha} \odot \boldsymbol{f}(\boldsymbol{x}^{\boldsymbol{u}}) + \boldsymbol{\beta} \odot \boldsymbol{u}, i \in [\![3]\!] := \{1, 2, 3\}$$

where

$$\boldsymbol{f}(\boldsymbol{z}) := (z_2 z_3, z_3 z_1, z_1 z_2)^{\top} \text{ for } \boldsymbol{z} \in \mathbb{R}^3$$

$$\alpha_i := (J_{i+1 \mod 3} - J_{i+2 \mod 3})/J_i,\ \beta_i := 1/J_i,\ i \in [\![3]\!]$$

MEGSBP with Euler Drift, without State Cost

$$\underset{(\rho^{\boldsymbol{u}},\boldsymbol{u})\in\mathcal{P}_2(\mathbb{R}^n)\times\mathcal{U}}{\arg\inf} \int_0^T \int_{\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{u}\|_2^2\, \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t)\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}}\mathrm{d}t$$

$$\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}} = (\boldsymbol{\alpha}\odot \boldsymbol{f}(\boldsymbol{x}^{\boldsymbol{u}}) + \boldsymbol{\beta}\odot\boldsymbol{u})\,\mathrm{d}t + \sqrt{2\delta}\,\mathrm{d}\boldsymbol{w},\ \ \delta > 0$$

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=0) = \rho_0, \quad \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t=T) = \rho_T.$$

Thm. The minimizing tuple for this problem exists and is unique

Conditions of Optimality for MEGSBP with Euler Drift

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}\|\boldsymbol{\beta} \odot \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \phi\|_2^2 + \langle \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \phi, \boldsymbol{\alpha} \odot \boldsymbol{f}(\boldsymbol{x}^{\boldsymbol{u}}) \rangle = -\delta \Delta_{\boldsymbol{x}^{\boldsymbol{u}}} \phi,$$

$$\frac{\partial \rho^{\text{opt}}}{\partial t} + \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \cdot \left(\rho^{\text{opt}} \left(\boldsymbol{\alpha} \odot \boldsymbol{f}(\boldsymbol{x}^{\boldsymbol{u}}) + \boldsymbol{\beta}^2 \odot \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \phi \right) \right) = \delta \Delta_{\boldsymbol{x}^{\boldsymbol{u}}} \rho^{\text{opt}},$$

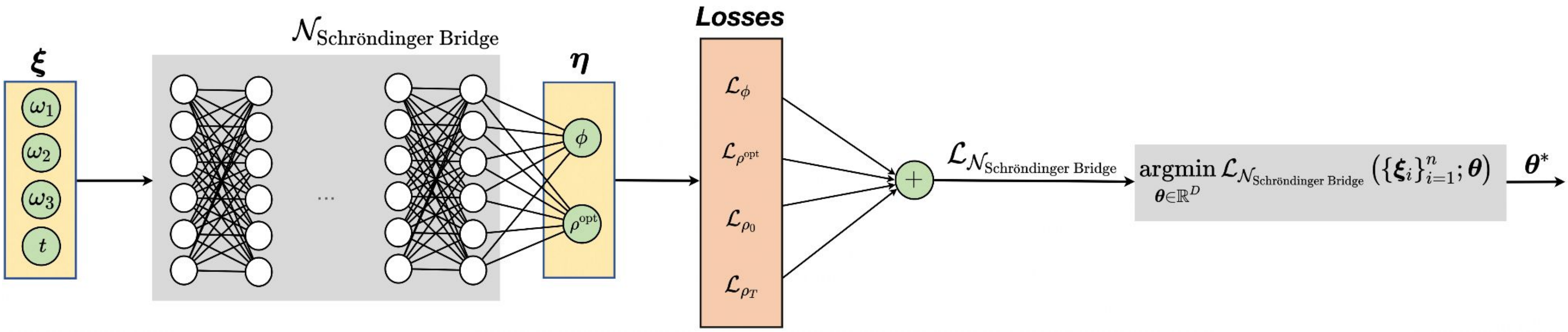
$$\rho^{\text{opt}}(\boldsymbol{x}^{\boldsymbol{u}}, t = 0) = \rho_0, \quad \rho^{\text{opt}}(\boldsymbol{x}^{\boldsymbol{u}}, t = T) = \rho_T,$$

$$\boxed{\boldsymbol{u}^{\text{opt}} = \boldsymbol{\beta} \odot \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \phi,}$$

Optimal policy is explicit in terms of the value function ϕ

Training Architecture

- HJB PDE loss: $\mathcal{L}_{\phi}$
- FPK PDE loss: $\mathcal{L}_{\rho^{opt}}$
- Sinkhorn regularized losses: $\mathcal{L}_{\rho_0}$+ $\mathcal{L}_{\rho_T}$



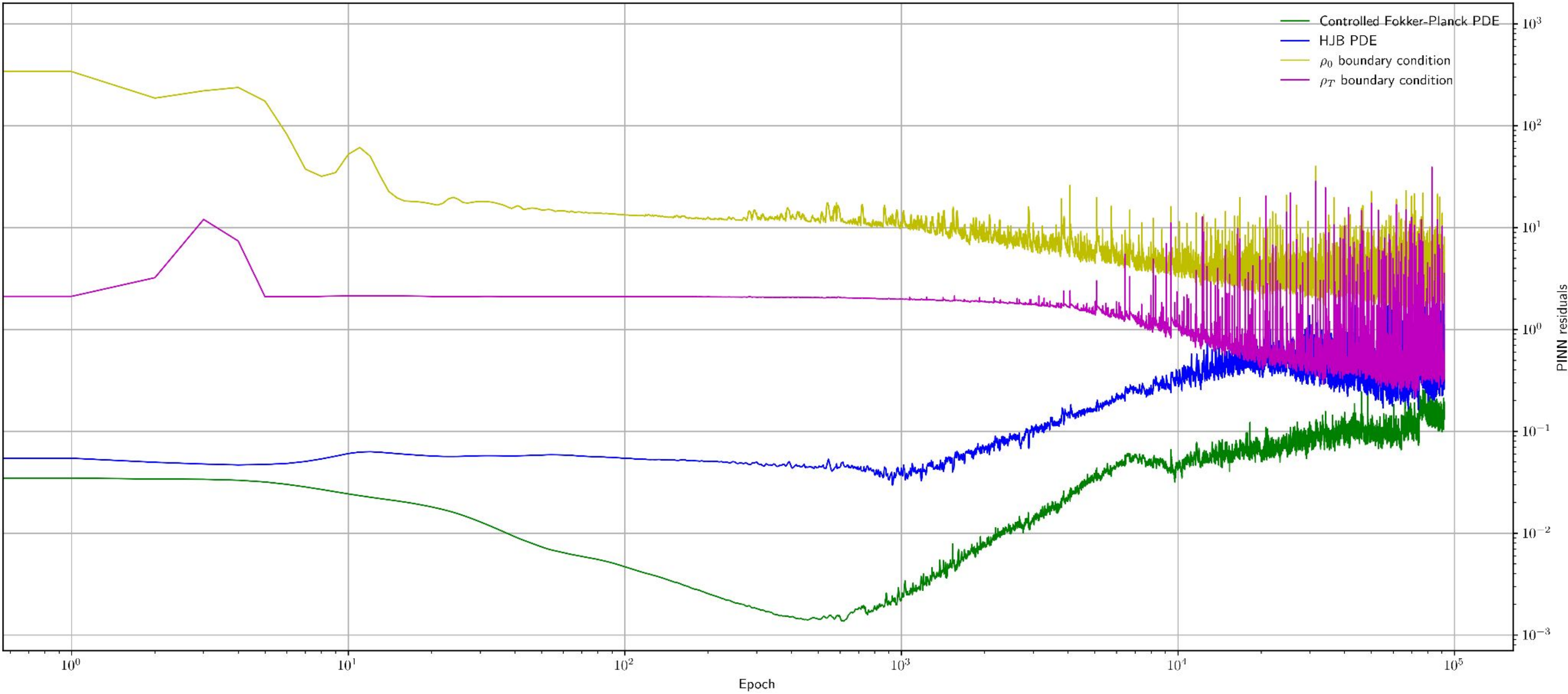
Numerical Simulation

- $\rho_0 = \mathcal{N}((2,2,2), 0.5\boldsymbol{I}_3), \quad \rho_T = \mathcal{N}((0,0,0), 0.5\boldsymbol{I}_3)$
- 3 hidden layers, 70 neurons in each, tanh activation, Glorot normal initialization, adam SGD w/ lr = 10^-3
- 80k epochs, 100k domain samples (mini-batched 35k every 40k epochs) + 1250 boundary condition samples
- Sinkhorn loss entropic regularizer $\varepsilon = 0.1$
- Principal moments of inertia:

$$J_1 = 0.45, J_2 = 0.50, J_3 = 0.55$$

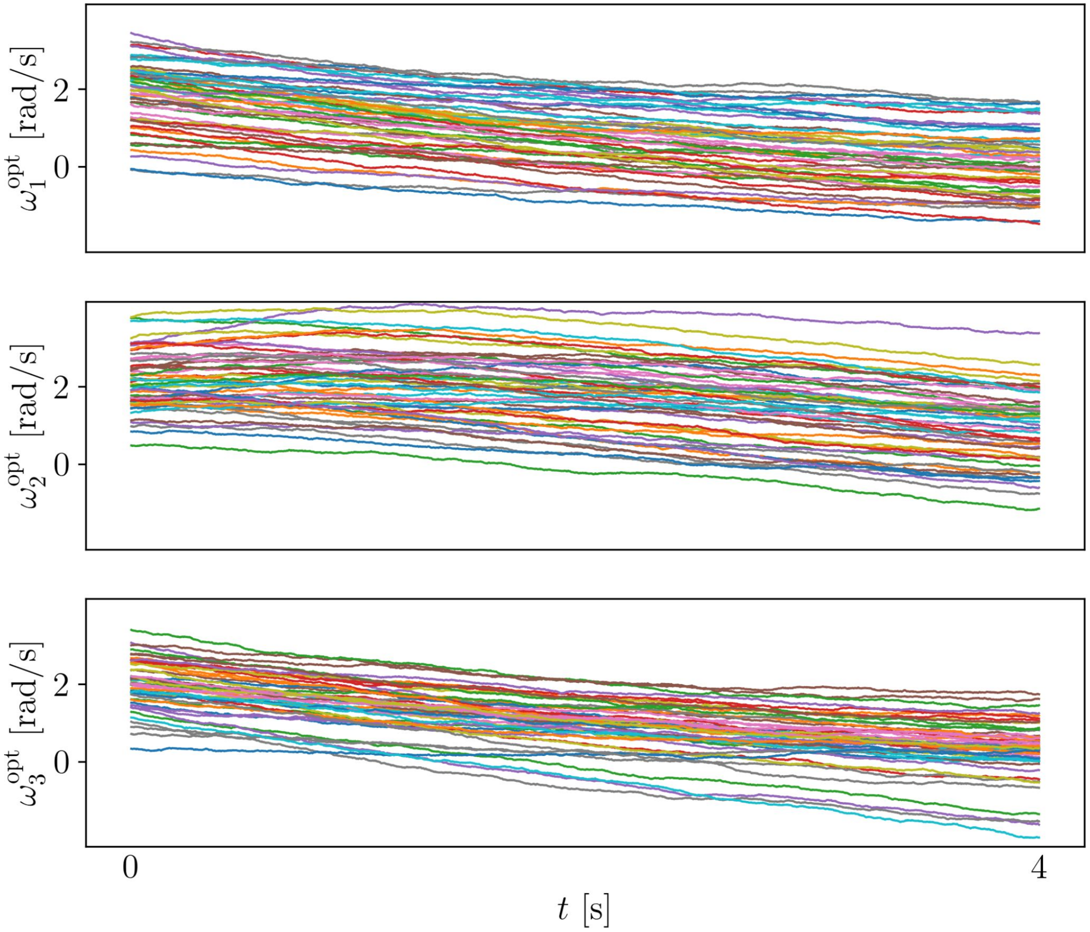
- Final time $T = 4$
- PINN space-time collocation domain: $[-5,5]^3 \times [0,4]$

Training Residuals, 80k Epochs



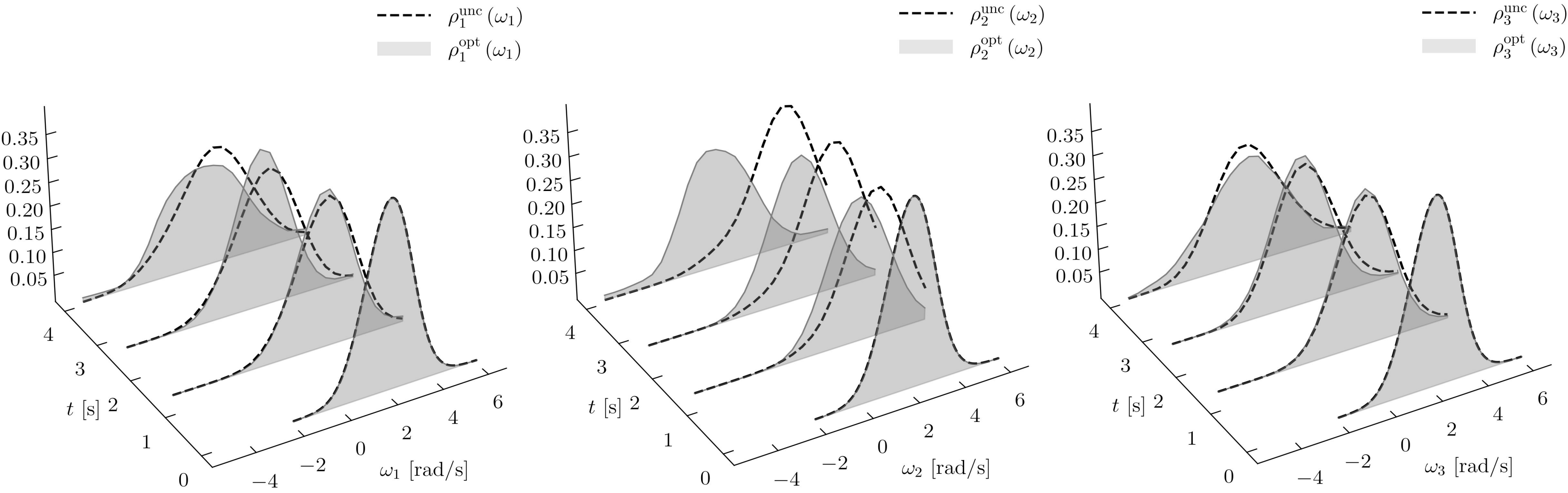
50 Optimally Controlled Closed-loop State Sample Paths

- Euler-maruyama integration
  - Noise strength 0.1



Univariate Marginals of the Optimally Controlled Joint

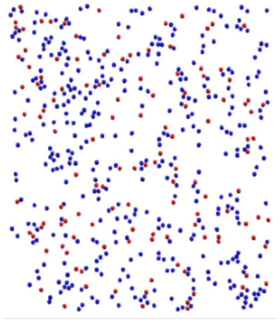
- Four snapshots
- Uncontrolled (--) vs controlled ($\boxed{}$) for $\omega_1, \omega_2, \omega_3$



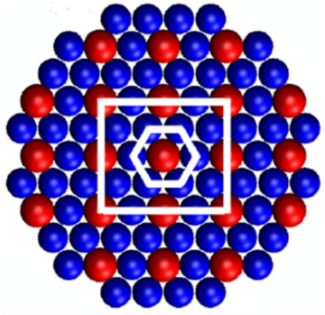
Case Study #2

Data-driven Controlled Colloidal Self-assembly

Recap: Optimal Steering of Stochastic State



Dispersed particles



Ordered structure

Stochastic state:

Measure of crystallinity order

$(\langle C_{10} \rangle, \langle C_{12} \rangle)$ Steinhart bond order parameters

Data-driven controlled colloidal self-assembly for precision (sub nm scale) manufacturing

MEGSBP with Data-driven Drift and Diffusion

$$\underset{(\rho^{\boldsymbol{u}},\boldsymbol{u})\in\mathcal{P}_2(\mathbb{R}^n)\times\mathcal{U}}{\arg\inf} \int_0^T \int_{\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{u}\|_2^2\, \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t)\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}}\mathrm{d}t$$

$$\mathrm{d}\boldsymbol{x}^{\boldsymbol{u}} = \boldsymbol{f}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\boldsymbol{g}\left(t,\boldsymbol{x}^{\boldsymbol{u}},\boldsymbol{u}\right)\mathrm{d}\boldsymbol{w}$$

As two NN approximants from the Molecular Dynamics (MD) simulation data [in collaboration with Prof. Mesbah's group at UC Berkeley]

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t=0)=\rho_0,\quad \rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}},t=T)=\rho_T.$$

Conditions of Optimality for MEGSBP

$$\boxed{\frac{\partial \psi}{\partial t} = q + \frac{1}{2}\|\boldsymbol{u}_{\text{opt}}\|_2^2 - \langle \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \psi, \boldsymbol{f} \rangle - \langle \boldsymbol{G}_{\beta}, \textbf{Hess}(\psi) \rangle}$$

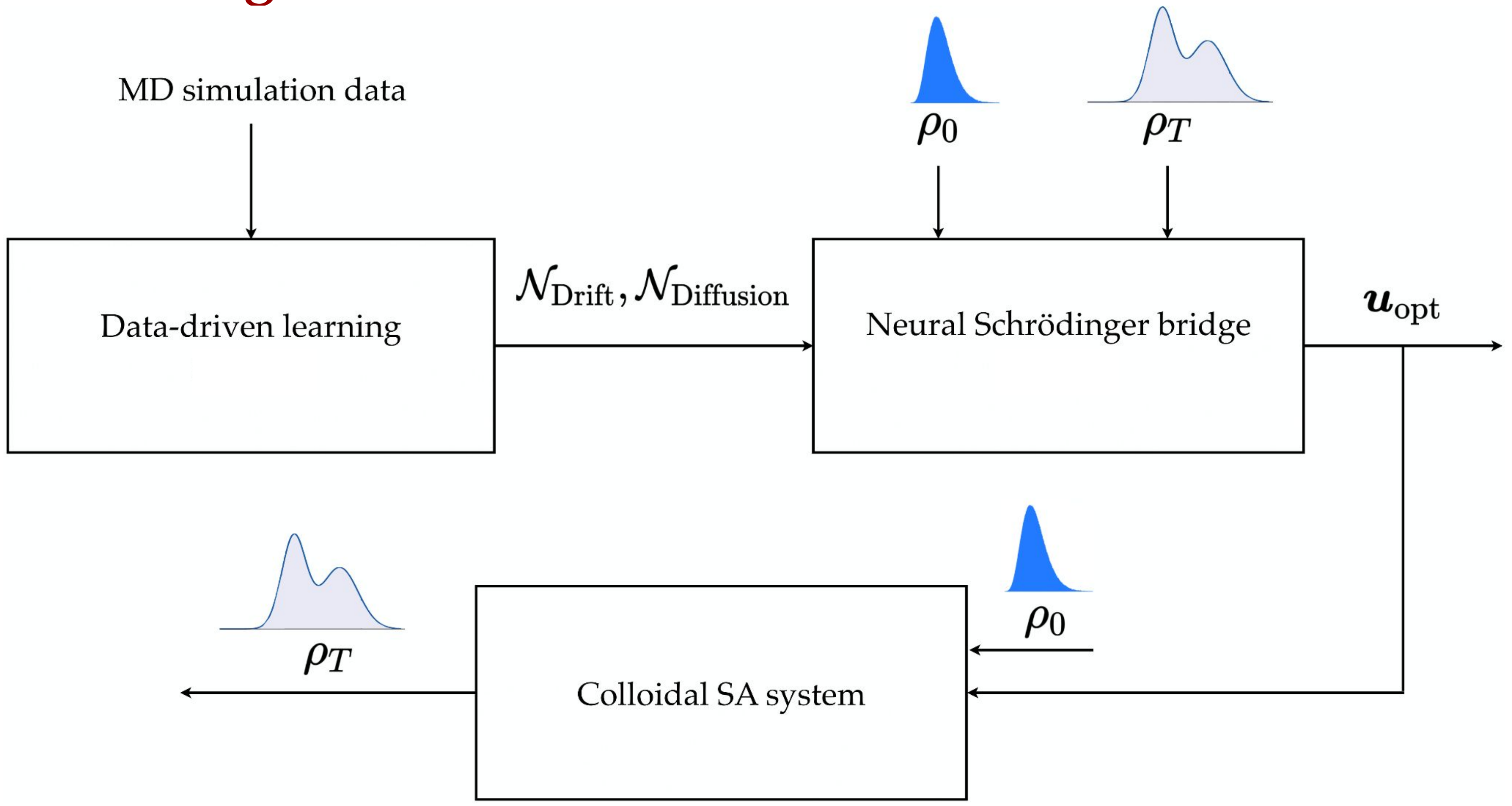
$$\boxed{\frac{\partial \rho^{\boldsymbol{u}}_{\text{opt}}}{\partial t} = -\nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \cdot (\rho^{\boldsymbol{u}}_{\text{opt}} \boldsymbol{f}) + \langle \boldsymbol{G}_{\beta}, \textbf{Hess}(\rho^{\boldsymbol{u}}_{\text{opt}}) \rangle,}$$

$$\boxed{\boldsymbol{u}_{\text{opt}} = \nabla_{\boldsymbol{u}_{\text{opt}}} \left(\langle \nabla_{\boldsymbol{x}^{\boldsymbol{u}}} \psi, \boldsymbol{f} \rangle + \langle \boldsymbol{G}_{\beta}, \textbf{Hess}(\psi) \rangle \right).}$$

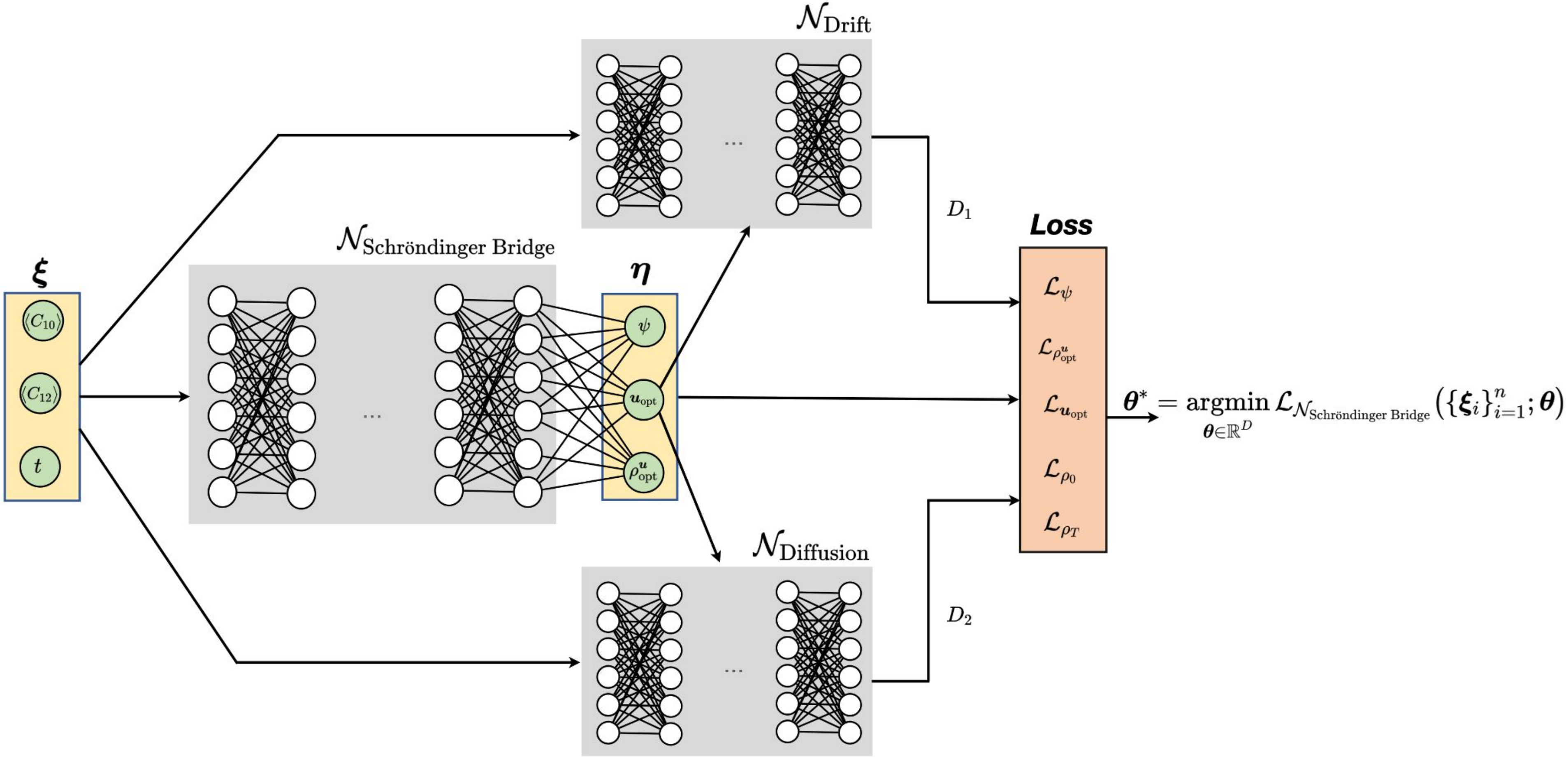
$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t = 0) = \rho_0,$$

$$\rho^{\boldsymbol{u}}(\boldsymbol{x}^{\boldsymbol{u}}, t = T) = \rho_T.$$

Block Diagram



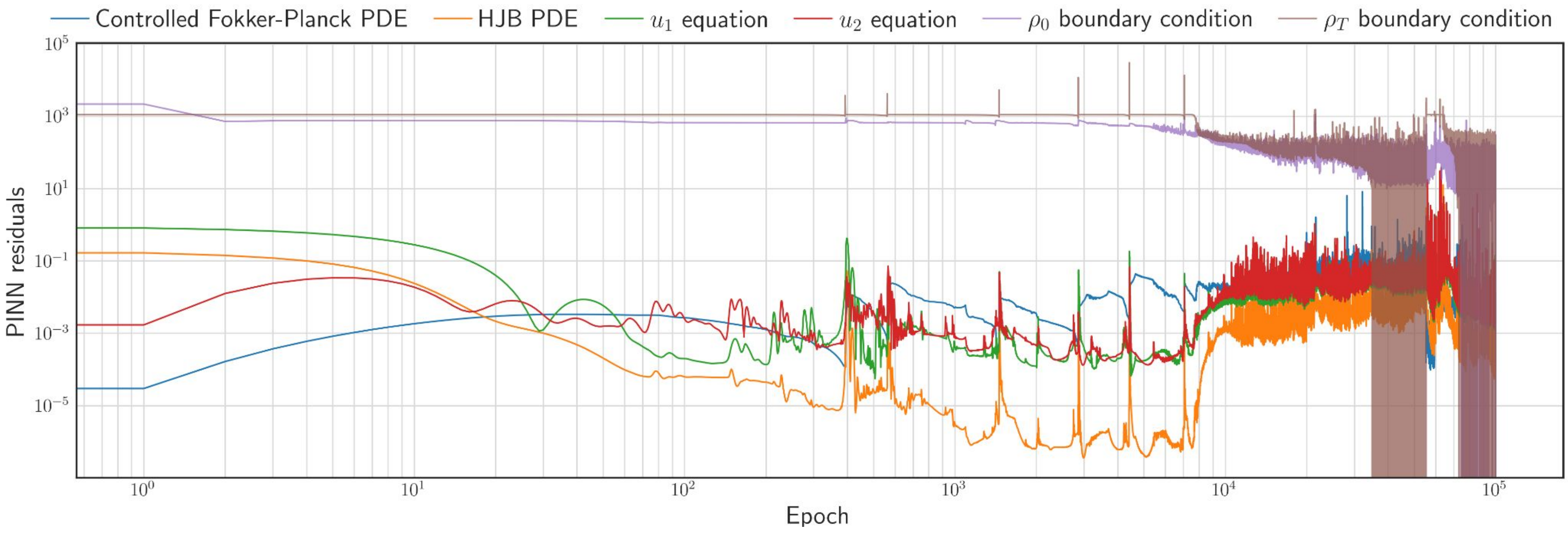
Training Architecture



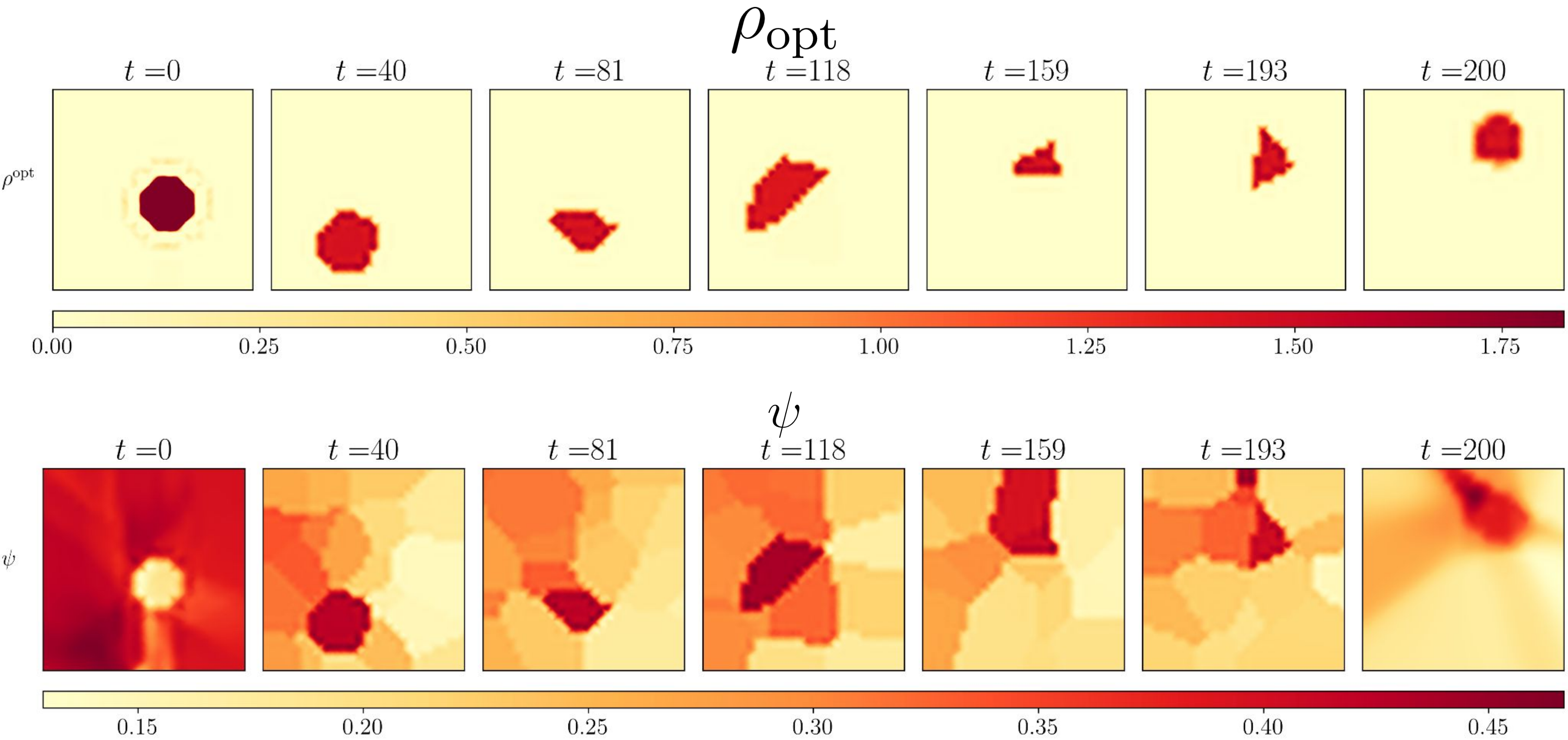
Numerical Simulation

- $m_0 = (0.2, 0.2)^\top, m_T = (0.4, 0.375)^\top, \Sigma_0 = \Sigma_T = 0.1 I_2.$
- 3 hidden layers, 70 neurons in each, tanh activation, Glorot normal initialization, adam SGD w/ lr = 10^-3
- 100k epochs, 100k domain samples (mini-batched 3k every 20k epochs) + 968 boundary condition samples
- Sinkhorn loss on entropic regularization $\epsilon = 0.1$
- Steering Gaussian to Gaussian
- Final time T = 200s

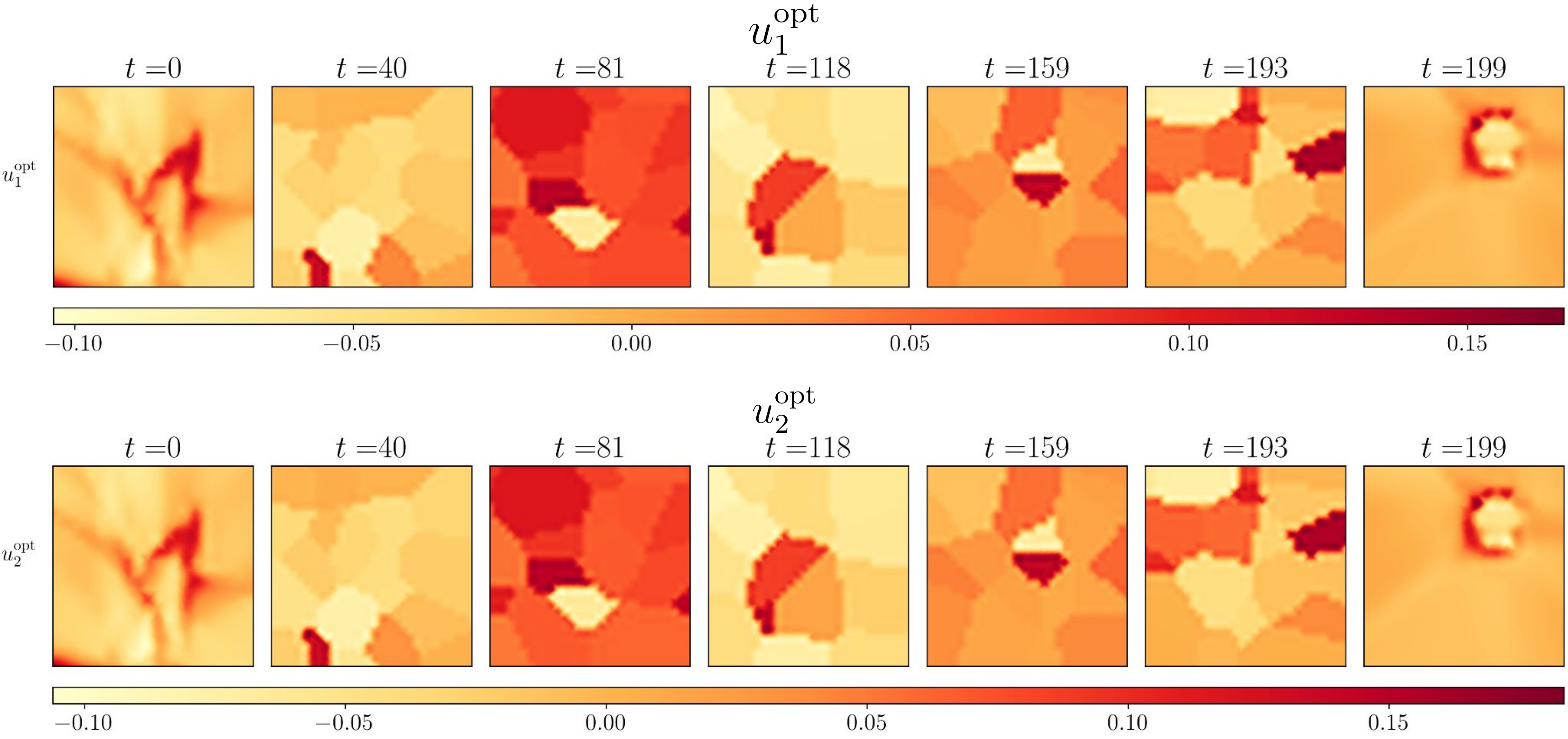
Training Residuals



Optimally Controlled Joint and Value Function



Optimal Control Policy



Necessary Training Tricks

- Since PINN's activations function is (necessarily) $\tanh$, its output tensors could be negative, positive, trivial, take on any distribution shape, and may not be a valid PDF during training.
  - To teach the network in a 'nice / convex' way:
    - Counting is **not** a convex way to ask
    - **p1 = -torch.sum(y_pred[y_pred < 0])**
  - To encourage a NN to prioritize a criteria and get it BEFORE another, weight it more
    - y_pred = torch.where(y_pred < 0, 0, y_pred)
    - dist, _, _ = sinkhorn(C, y_pred.reshape(-1), rho_tensor)
    - 10 * p1 + dist
  - To preserve **numerical stability** for computing Sinkhorn regularized Wasserstein distance when input is **not a valid distribution**, use log-sum-exp(LSE) trick

Numerical Experiment Takeaways

- 3 things in tension:
  - Correctness and completeness, numerical stability etc., minimal time and space complexity / mini-batching requirement
- PINN convergence on residuals alone does NOT imply good control policy
  - Correct training dataset *density* and *distribution pseudorandomness* seems necessary / related to good control
- HJB PDE clamping to a sum or trapz or any other clamping will result in a **trivial control policy**, so we only care about Sinkhorn distance
- Shape of training: PINN solves a **trivial solution** for FPK PDE / HJB PDE, but then to solve the boundary conditions, is forced to find a **nontrivial solution**. PDE ‘W’ shape.

Thank you!