Mean-Field Schrödinger Problem

Gianluca Covini

Presentation for the exam of *Stochastic Processes II*

Forward Outline

Goal

Introduce the **Mean-field Schrödinger problem**, the study of the most likely evolution of a cloud of interacting Brownian particles conditional on observing initial and final empirical distributions, μ^{in} and μ^{fin} and interpret it as a way to **interpolate between μ^{in} and μ^{fin}** $\in \mathcal{P}_2(\mathbb{R}^d)$.

Outline

- Classical Schrödinger Problem
- Mean-field Schrödinger Problem
  - Definition and Properties
  - Stochastic Control Formulation
  - A Geometric Perspective
- Conclusion

Schrödinger Problem

Schrödinger Problem: Statistical Physics Motivation

In 1931, E. Schrödinger proposed in the article *"Über die Umkehrung der Naturgesetze"*, the following problem (expressed in modern terms)

Statistical Physics Schrödinger Problem

Consider $N \gg 1$ **independent** Brownian particles $(X_t^i)_{t\in[0,1], i=1,\dots,N}$. The law of the system at time t can be described by

$$\hat{\mu}_t := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^i}$$

Given that we observe $\hat{\mu}_0 \approx \mu^{\text{in}}$ and $\hat{\mu}_1 \approx \mu^{\text{fin}}$,

- **What is the most likely evolution of the system in** $[0,1]$?

Schrödinger Problem: a Way to Interpolate Between Measures

Schrödinger Problem can be seen as a way to interpolate between the initial measure μ^{in} and the final measure μ^{fin}.

In probabilistic terms,

- We consider $\Omega := C([0,1], \mathbb{R}^d)$ the space of continuous paths in $\mathbb{R}^d$;
- $\mathcal{P}(\Omega)$ is the space of probabilities on the paths in Ω.

We study

→ $P \in \mathcal{P}(\Omega)$, a probability on paths in $\mathbb{R}^d$, to express $P_0 = \mu^{\text{in}}$ and $P_1 = \mu^{\text{fin}}$;

→ $X = (X_t)_{t \in [0,1]} \sim P$ to express the evolution of the system.

Schrödinger Problem: a Way to Interpolate Between Measures

→ Taking $P \in \mathcal{P}(\Omega)$ as law of a path $X = (X_t)_{t\in[0,1]}$;

→ P_t the t-th marginal of P, i.e. $P_t(B) := \mathbb{P}_{X\sim P}[X_t \in B]$ for $B \in \mathcal{B}(\mathbb{R}^d)$.

Schrödinger Problem

In probabilistic terms, it corresponds to the problem

$$\begin{aligned} &\min \mathrm{KL}(P \mid R) \\ &\text{s.t. } P \in \mathcal{P}(\Omega), P_0 = \mu^{\text{in}}, P_1 = \mu^{\text{fin}} \end{aligned}$$

where R is the law of Brownian motion with starting measure the Lebesgue measure, i.e.,

$$R(\cdot) = \int_{\mathbb{R}^d} \mathbb{P}[B_t + x \in \cdot]\mathrm{d}x$$

Schrödinger Problem: a Way to Interpolate Between Measures

Entropic Optimal Transport [C. Léonard, *A Survey on the Schrödinger Problem and some of its Connections with Optimal Transport*, 2013]

$$(SP) \quad \min_{\substack{P \in \mathcal{P}(\Omega) \\ P_0 = \mu^{\text{in}}, P_1 = \mu^{\text{fin}}}} \mathsf{KL}(P \mid R) \quad \rightsquigarrow \quad (SP_{\text{static}}) \quad \min_{\pi \in \Pi(\mu^{\text{in}}, \mu^{\text{fin}})} \mathsf{KL}(\pi \mid R_{01})$$

with $R_{01} = (X_0, X_1)_{\#} R$, projection of R on $\mathbb{R}^{2d}$.

(SP_{static}) is equivalent (up to rescaling) to the **entropic optimal transport** problem

$$\min_{\pi \in \Pi(\mu^{\text{in}}, \mu^{\text{fin}})} \left\{ \int_{\mathbb{R}^{2d}} \frac{\|x - y\|^2}{2} \mathrm{d}\pi(x, y) + \varepsilon \mathsf{KL}(\pi \mid \mu^{\text{in}} \otimes \mu^{\text{fin}}) \right\}$$

→ **Idea:** $R_{01}(\mathrm{d}x, \mathrm{d}y) \propto e^{-\|x-y\|^2/2} \mathrm{d}x \mathrm{d}y$

→ Computationally tractable approximation of optimal transport.

→ Schrödinger Problem gives a convex equivalent useful to derive many properties.

Mean-Field Schrödinger Problem: definition and properties

Adding Interaction

$\rightarrow$ *"The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities"* (2020), Backhoff, Conforti, Gentil and Léonard: generalization of Schrödinger problem including **mean-field interactions** between the Brownian particles.

Adding Interaction

→ *"The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities"* (2020), Backhoff, Conforti, Gentil and Léonard: generalization of Schrödinger problem including **mean-field interactions** between the Brownian particles.

Statistical Physics Problem

Let $(X_t^i)_{t\in[0,T],i=1,\dots,N}$ be Brownian particles interacting through a potential W according to the system of SDEs:

$$\begin{cases} \mathrm{d}X_t^i = -\frac{1}{N}\sum_{k=1}^N \nabla W(X_t^i - X_t^k)\mathrm{d}t + \mathrm{d}B_t^i \\ X_0^i \sim \mu^{\text{in}} \end{cases}$$

Again, we observe the law of the system at time T: $\hat{\mu}_T = \frac{1}{N}\sum_{i=1}^N \delta_{X_T^i} \approx \mu^{\text{fin}}$.

- **What is the most likely evolution of the system in $[0, T]$?**

Probabilistic Problem

Mean-Field Schrödinger Problem: Probabilistic formulation

$$\min \mathrm{KL}(P \mid \Gamma(P)) \quad \text{s.t. } P \in \mathcal{P}_1(\Omega), P_0 = \mu^{\mathrm{in}}, P_T = \mu^{\mathrm{fin}}$$

and $\Gamma(P)$ is the unique strong solution of

$$\begin{cases} \mathrm{d}X_t = -(\nabla W * P_t)(X_t)\mathrm{d}t + \mathrm{d}B_t = \mathrm{d}X_t = -\left(\int_{\mathbb{R}^d} \nabla W(X_t - y)\, \mathrm{d}P_t(y)\right) \mathrm{d}t + \mathrm{d}B_t \\ X_0 \sim \mu^{\mathrm{in}} \end{cases}$$

Probabilistic Problem

Mean-Field Schrödinger Problem: Probabilistic formulation

$$\min \mathrm{KL}(P \mid \Gamma(P)) \quad \text{s.t. } P \in \mathcal{P}_1(\Omega), P_0 = \mu^{\mathrm{in}}, P_T = \mu^{\mathrm{fin}}$$

and $\Gamma(P)$ is the unique strong solution of

$$\begin{cases} \mathrm{d}X_t = -(\nabla W * P_t)(X_t)\mathrm{d}t + \mathrm{d}B_t = \mathrm{d}X_t = -\left(\int_{\mathbb{R}^d} \nabla W(X_t - y)\, \mathrm{d}P_t(y)\right) \mathrm{d}t + \mathrm{d}B_t \\ X_0 \sim \mu^{\mathrm{in}} \end{cases}$$

Remarks

→ P imposes the satisfaction of initial and final distributions μ^{in} and μ^{fin}.

→ $\Gamma(P)$ describes evolution given initial status μ^{in}, Brownian evolution and mean-field interactions;

→ $\Gamma(P)$ depends on P only through the time marginals $(P_t)_{t\in[0,T]}$.

Probabilistic Problem

Mean-Field Schrödinger Problem: Probabilistic formulation

$$\min \mathsf{KL}(P \mid \Gamma(P)) \quad \text{s.t. } P \in \mathcal{P}_1(\Omega), P_0 = \mu^{\mathsf{in}}, P_T = \mu^{\mathsf{fin}}$$

and $\Gamma(P)$ is the unique strong solution of

$$\begin{cases} \mathrm{d}X_t = -(\nabla W * P_t)(X_t)\mathrm{d}t + \mathrm{d}B_t \\ X_0 \sim \mu^{\mathsf{in}} \end{cases}$$

Motivations

- Statistical Physics.
- New way to interpolate between measures μ^{in} and μ^{fin}.
- (Formal) way to interpret second order ODEs in the Wasserstein space.
- Functional inequalities.

Probabilistic Problem

Mean-Field Schrödinger Problem: Probabilistic formulation

$$\min \mathrm{KL}(P \mid \Gamma(P)) \quad \text{s.t. } P \in \mathcal{P}_1(\Omega), P_0 = \mu^{\mathrm{in}}, P_T = \mu^{\mathrm{fin}}$$

and $\Gamma(P)$ is the unique strong solution of

$$\begin{cases} \mathrm{d}X_t = -(\nabla W * P_t)(X_t)\mathrm{d}t + \mathrm{d}B_t \\ X_0 \sim \mu^{\mathrm{in}} \end{cases}$$

Existence of a solution

Assuming

- (H1) $W \in C^2(\mathbb{R}^d; \mathbb{R})$, symmetric, $\sup_{\substack{z,v \in \mathbb{R}^d \\ |v|=1}} v^\top \cdot \nabla^2 W(z) \cdot v < +\infty$
- (H2) $\mu^{\mathrm{in}}, \mu^{\mathrm{fin}} \in \mathcal{P}_2(\mathbb{R}^d)$ and with finite free-energy,

the mean-field Schrödinger problem has a solution.

Large Deviations Principle

The interpretation of most likely evolution of particles is justified by:

Large Deviations Principle

Under (H1), (H2) and $\int_{\mathbb{R}^d} \exp(r|x|)\mathrm{d}\mu^{\mathrm{in}}(x) < +\infty \quad \forall r > 0$, for mean-field interacting Brownian particles $(X^i)_{i=1}^N$,

$$\left\{ \hat{\mu}^{(N)} := \frac{1}{N}\sum_{i=1}^{N} \delta_{X^i} \mid N \in \mathbb{N} \right\}$$

satisfies a **large deviations principle** with rate function $\mathcal{P}_1(\Omega) \ni P \mapsto \mathsf{KL}(P \mid \Gamma(P))$, namely, with respect to the 1-Wasserstein topology, for every $A \subseteq \mathcal{P}_1(\Omega)$ measurable

$$\begin{aligned} - \inf_{P \in A^\circ} \{\mathsf{KL}(P \mid \Gamma(P))\} &\leq \liminf_{N \uparrow +\infty} \frac{1}{N} \log \mathbb{P}[\hat{\mu}^{(N)} \in A] \leq \\ &\leq \limsup_{N \uparrow +\infty} \frac{1}{N} \log \mathbb{P}[\hat{\mu}^{(N)} \in A] \leq - \inf_{P \in \bar{A}} \{\mathsf{KL}(P \mid \Gamma(P))\} \end{aligned}$$

Large Deviations Principle

Large Deviations Principle: $\hat{\mu}^{(N)}$ empirical path measure

$$-\inf_{P\in A^\circ}\{\mathsf{KL}(P \mid \Gamma(P))\} \le \liminf_{N\uparrow+\infty}\frac{1}{N}\log\mathbb{P}[\hat{\mu}^{(N)}\in A] \le$$
$$\le \limsup_{N\uparrow+\infty}\frac{1}{N}\log\mathbb{P}[\hat{\mu}^{(N)}\in A] \le -\inf_{P\in\bar{A}}\{\mathsf{KL}(P \mid \Gamma(P))\}$$

Interpretation

For N big enough,

$$\mathbb{P}\left[\hat{\mu}^{(N)} \approx Q\right] \approx \exp(-N\mathsf{KL}(Q \mid \Gamma(Q))) \text{ for candidate minimizer of MFSP } Q\in\mathcal{P}_1(\Omega)$$

Mean-field Schrödinger Problem: a stochastic optimal control problem

Stochastic Optimal Control Interpretation

$\Gamma(Q)$ solves the particle dynamics $\mathrm{d}X_t = -(\nabla W * Q_t)(X_t)\mathrm{d}t + \mathrm{d}B_t$, $X_0 \sim \mu^{\text{in}}$.

$\rightarrow$ **Idea:** Choosing $P \in \arg\min_Q \mathrm{KL}(Q \mid \Gamma(Q))$, means adding the minimum drift such that the final condition $P_T = \mu^{\text{fin}}$ is satisfied.

Stochastic Optimal Control Interpretation

$\Gamma(Q)$ solves the particle dynamics $\mathrm{d}X_t = -(\nabla W * Q_t)(X_t)\mathrm{d}t + \mathrm{d}B_t$, $X_0 \sim \mu^{\mathrm{in}}$.

→ **Idea:** Choosing $P \in \arg\min_Q \mathrm{KL}(Q \mid \Gamma(Q))$, means adding the minimum drift such that the final condition $P_T = \mu^{\mathrm{fin}}$ is satisfied.

Stochastic Control Interpretation

Under (H1),(H2), the mean-field Schrödinger problem is equivalent to

$$\begin{aligned} \inf_P \quad & \frac{1}{2}\mathbb{E}_P\left[\int_0^T |\alpha_t^P|^2\, dt\right] \\ \text{s.t.} \quad & P \in \mathcal{P}_1(\Omega), \qquad P_0 = \mu^{\mathrm{in}}, \quad P_T = \mu^{\mathrm{fin}}, \\ & \left(X - \int_0^\cdot \left[-(\nabla W * P_s)(X_s) + \alpha_s^P\right] \mathrm{d}s\right)_{\#} P = R^{\mu^{\mathrm{in}}} \end{aligned}$$

where $R^{\mu^{\mathrm{in}}}$ is the Wiener measure with starting distribution μ^{in} and α^P a predictable stochastic process (well-defined given P).

Stochastic Optimal Control Interpretation

$\Gamma(Q)$ solves the particle dynamics $\mathrm{d}X_t = -(\nabla W * Q_t)(X_t)\mathrm{d}t + \mathrm{d}B_t$, $X_0 \sim \mu^{\text{in}}$.

→ **Idea:** Choosing $P \in \arg\min_Q \mathrm{KL}(Q \mid \Gamma(Q))$, means adding the minimum drift such that the final condition $P_T = \mu^{\text{fin}}$ is satisfied.

Stochastic Control Interpretation

Under (H1),(H2), the mean-field Schrödinger problem is equivalent to

$$\begin{aligned} \inf_P \quad & \frac{1}{2}\mathbb{E}_P\left[\int_0^T |\alpha_t^P|^2 \mathrm{d}t\right] \\ \text{s.t.} \quad & P \in \mathcal{P}_1(\Omega), \qquad P_0 = \mu^{\text{in}}, \quad P_T = \mu^{\text{fin}}, \\ & \begin{cases} dX_t = \left[-(\nabla W * P_t)(X_t) + \alpha_t^P\right]\mathrm{d}t + \mathrm{d}B_t, \\ X_0 \sim \mu^{\text{in}} \end{cases} \end{aligned}$$

Here B is a Brownian motion under P, $(X_t)_t$ canonical process under P and α^P a predictable stochastic process (well-defined given P).

Stochastic Control Properties

Stochastic control formulation

The stochastic control formulation is a consequence of Girsanov theorem as presented by C. Léonard's paper in *Girsanov theorem under a finite entropy condition* (2011).

Stochastic Control Properties

Stochastic control formulation

The stochastic control formulation is a consequence of Girsanov theorem as presented by C. Léonard's paper in *Girsanov theorem under a finite entropy condition* (2011).

Optimality Conditions

Optimality conditions of the MFSP entail $\exists \Psi \in \overline{\{\nabla\psi \mid \psi \in C_c^\infty([0,T] \times \mathbb{R}^d)\}}^{L^2((P_t)_t)}$, *corrector* of P,

- $\Psi_t(X_t) = \alpha_t^P \quad \mathrm{d}t \otimes \mathrm{d}P$-a.s.
- and

$$M_t := \Psi_t(X_t) - \int_0^t \widetilde{\mathbb{E}}\left[\nabla^2 W(X_s - \widetilde{X}_s) \cdot (\Psi_s(X_s) - \Psi_s(\widetilde{X}_s))\right] \mathrm{d}s$$

  is a continuous martingale under P on $[0,T)$, with $(\widetilde{X}_t, \widetilde{Y}_t)$ independent copies of (X_t, Y_t).

Stochastic Control Properties

Stochastic control formulation

The stochastic control formulation is a consequence of Girsanov theorem as presented by C. Léonard's paper in *Girsanov theorem under a finite entropy condition* (2011).

Optimality Conditions

Optimality conditions of the MFSP entail $\exists \Psi \in \overline{\{\nabla \psi \mid \psi \in C_c^\infty([0,T] \times \mathbb{R}^d)\}}^{L^2((P_t)_t)}$, *corrector* of P,

- $\Psi_t(X_t) = \alpha_t^P \quad \mathrm{d}t \otimes \mathrm{d}P$-a.s.
- formally, setting $Y_t := \Psi_t(X_t)$ and using decomposition of the martingale $(M_t)_t$

$$\mathrm{d}Y_t = \widetilde{\mathbb{E}}\left[\nabla^2 W(X_t - \widetilde{X}_t) \cdot (Y_t - \widetilde{Y}_t)\right] \mathrm{d}t + Z_t \cdot \mathrm{d}B_t$$

  with $(\widetilde{X}_t, \widetilde{Y}_t)$ independent copies of (X_t, Y_t).

FBSDE Dynamics

Forward-backward SDEs

Let P solve MFSP, $X \sim P$ and $Y_t := \Psi_t(X_t)$. Then we have the system follows the dynamic

$$\begin{cases} \mathrm{d}X_t = -\widetilde{\mathbb{E}}\left[\nabla W(X_t - \widetilde{X}_t)\right]\mathrm{d}t + Y_t \mathrm{d}t + \mathrm{d}B_t \\ \mathrm{d}Y_t = \widetilde{\mathbb{E}}\left[\nabla^2 W(X_t - \widetilde{X}_t) \cdot (Y_t - \widetilde{Y}_t)\right]\mathrm{d}t + Z_t \cdot \mathrm{d}B_t \\ X_0 \sim \mu^{\mathrm{in}}, X_T \sim \mu^{\mathrm{fin}} \end{cases}$$

$\to$ Under regularity assumptions, if $\Psi_t(X_t) = \nabla\psi_t(X_t)$ and $\mu_t := (X_t)_{\#}P$, then (ψ_t, μ_t) solve a coupled PDE dynamic of the form of planning mean-field games.

Mean-field Schrödinger Problem: a geometric perspective

Interpolating between Measures

- → **Optimal transport** and **classical Schrödinger problem** (entropic optimal transport) give a way to interpolate between measures in the space $\mathcal{P}_2(\mathbb{R}^d)$.
- → Mean-field Schrödinger problem enters this line of research, being interpreted as an interpolation dynamic between μ^{in} and $\mu^{\text{fin}} \in \mathcal{P}_2(\mathbb{R}^d)$.

A Benamou-Brenier perspective

$$\mathcal{C}_T(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}}) := \inf_{\substack{P \in \mathcal{P}_1(\Omega) \\ P_0 = \mu^{\mathsf{in}}, P_T = \mu^{\mathsf{fin}}}} \mathsf{KL}(P \mid \Gamma(P))$$

Given $\mathcal{A} := \{(\mu_t)_t \subseteq \mathcal{P}_2(\mathbb{R}^d) \mid$ absolutely continuous and s.t. $\mu_0 = \mu^{\mathsf{in}}, \mu_T = \mu^{\mathsf{fin}}\}$,

Benamou-Brenier Formula for W_2

$$W_2^2(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}}) = \inf_{\substack{(\mu_t)_{t \in [0,T]} \in \mathcal{A}, (v_t)_t \\ \partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 0}} \int_0^T \int_{\mathbb{R}^d} |v_t(z)|^2 \, \mathrm{d}\mu_t(z) \mathrm{d}t$$

Benamou-Brenier Interpretation of MFSP

$$\mathcal{C}_T(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}}) = \inf_{\substack{(\mu_t)_{t \in [0,T]} \in \mathcal{A}, (v_t)_t \\ \partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 0}} \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} \left| v_t(z) + \frac{1}{2} \nabla \log \mu_t(z) + (\nabla W * \mu_t)(z) \right|^2 \mathrm{d}\mu_t(z) \mathrm{d}t$$

Accelerated Dynamics

Benamou-Brenier Formula for W_2

$$
\begin{aligned}
W_2^2(\mu^{\text{in}}, \mu^{\text{fin}}) = \inf \quad & \int_0^T \int_{\mathbb{R}^d} |v_t(z)|^2 \, \mathrm{d}\mu_t(z)\mathrm{d}t \\
\text{s.t.} \quad & (\mu_t)_{t\in[0,T]} \in \mathcal{A}, (v_t)_t \\
& \partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 0
\end{aligned}
$$

Benamou-Brenier Interpretation of MFSP

$$
\begin{aligned}
\mathcal{C}_T(\mu^{\text{in}}, \mu^{\text{fin}}) = \inf \quad & \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |w_t(z)|^2 \, \mathrm{d}\mu_t(z)\mathrm{d}t \\
\text{s.t.} \quad & (\mu_t)_{t\in[0,T]} \in \mathcal{A}, (w_t)_t \\
& \partial_t \mu_t + \nabla \cdot ((w_t - \frac{1}{2}\nabla \log \mu_t - \nabla W * \mu_t)\mu_t) = 0
\end{aligned}
$$

A meaningful way to interpolate between probabilities: turnpike

→ Under (H3): $\nabla^2 W(z) \geq \kappa I_d$, $\kappa > 0, \forall z \in \mathbb{R}^d$, those dynamics are stable for $T \uparrow +\infty$.

→ ∃ an equilibrium configuration $\mu_\infty \in \arg\min_{\mu \in \mathcal{P}_2(\Omega)} \widetilde{\mathcal{F}}(\mu)$

***Turnpike* property: optimal trajectories spend most of their time around the equilibrium configuration μ_∞**

Given the free energy functional

$$\widetilde{\mathcal{F}}(\mu) = \int_{\mathbb{R}^d} \log \mu(x) \mathrm{d}\mu(x) + \int_{\mathbb{R}^d} (W * \mu)(x) \mathrm{d}\mu(x) \text{ if } \mu \ll Leb; +\infty \text{ otherwise}$$

Then, for every $\gamma \in (0,1)$, P being optimal for MFSP,

$$\widetilde{\mathcal{F}}(P_{\gamma T}) - \widetilde{\mathcal{F}}(\mu_\infty) \to 0 \text{ exponentially fast as } T \to \infty.$$

A meaningful way to interpolate between probabilities: functional inequalities

The MFSP defines a new interpolation between measures, so one can study the evolution of entropy-like functionals along this interpolation, as in optimal transport.

$$\rightarrow \qquad \mathcal{C}_T(\mu^{\mathrm{in}}, \mu^{\mathrm{fin}}) := \inf_{\substack{P \in \mathcal{P}_1(\Omega) \\ P_0 = \mu^{\mathrm{in}}, P_T = \mu^{\mathrm{fin}}}} \mathrm{KL}(P \mid \Gamma(P))$$

A Talagrand inequality

$$\mathcal{C}_T(\mu^{\mathrm{in}}, \mu_\infty) \leq \frac{1}{e^{2\kappa T} - 1} (\widetilde{\mathcal{F}}(\mu_{\mathrm{in}}) - \widetilde{\mathcal{F}}(\mu_\infty))$$

- $\mathcal{C}_T$ plays the role of an entropic transportation cost.
- For large T, reaching equilibrium is exponentially cheap.
- The paper also obtains an HWI inequality and the energy–transport inequality:

$$|\mathcal{E}_P(\mu^{\mathrm{in}}, \mu^{\mathrm{fin}})| \leq \frac{4\kappa}{\exp(\kappa T) - 1} \Big(\mathcal{C}_T(\mu^{\mathrm{in}}, \mu^{\mathrm{fin}}) \mathcal{C}_T(\mu^{\mathrm{fin}}, \mu^{\mathrm{in}}) \Big)^{1/2}$$

Conclusion

Backward outline

- We started from evolving particles $(X_t)_{t\in[0,T]}$ with $X_0 \sim \mu^{\mathrm{in}}$, $X_T \sim \mu^{\mathrm{fin}}$
  - under independent Brownian paths $\rightsquigarrow$ classical Schrödinger problem.
  - under mean-field interactions $\rightsquigarrow$ mean-field Schrödinger problem.

$\to$ We looked for most likely evolution.

Backward outline

- We started from evolving particles $(X_t)_{t\in[0,T]}$ with $X_0 \sim \mu^{\mathsf{in}}$, $X_T \sim \mu^{\mathsf{fin}}$
  - under independent Brownian paths $\rightsquigarrow$ classical Schrödinger problem.
  - under mean-field interactions $\rightsquigarrow$ mean-field Schrödinger problem.

$\rightarrow$ We looked for most likely evolution.

- MFSP had a stochastic control interpretation, as an evolution under controlled drift.

Backward outline

- We started from evolving particles $(X_t)_{t\in[0,T]}$ with $X_0 \sim \mu^{\text{in}}$, $X_T \sim \mu^{\text{fin}}$
  - under independent Brownian paths $\rightsquigarrow$ classical Schrödinger problem.
  - under mean-field interactions $\rightsquigarrow$ mean-field Schrödinger problem.

$\rightarrow$ We looked for most likely evolution.

- MFSP had a stochastic control interpretation, as an evolution under controlled drift.
- MFSP describes a way to interpolate between measures μ^{in} and μ^{fin} through their path law.
- The geometry of MFSP, though formally, induced dynamics
  - connected to accelerated dynamics between measures.
  - with a turnpike property.
  - with new functional inequalities.

Backward outline

- We started from evolving particles $(X_t)_{t\in[0,T]}$ with $X_0 \sim \mu^{\mathsf{in}}$, $X_T \sim \mu^{\mathsf{fin}}$
  - under independent Brownian paths $\rightsquigarrow$ classical Schrödinger problem.
  - under mean-field interactions $\rightsquigarrow$ mean-field Schrödinger problem.

$\rightarrow$ We looked for most likely evolution.

- MFSP had a stochastic control interpretation, as an evolution under controlled drift.
- MFSP describes a way to interpolate between measures μ^{in} and μ^{fin} through their path law.
- The geometry of MFSP, though formally, induced dynamics
  - connected to accelerated dynamics between measures.
  - with a turnpike property.
  - with new functional inequalities.

Thank you for your attention!

An useful result

Result

In the Brownian filtration, a square-integrable continuous martingale can be represented as

$$M_t = M_0 + \int_0^t Z_s \mathrm{d}B_s$$

In our case, if $\Psi_t(X_t) = \nabla\psi_t(X_t)$, then $Z_t = \nabla^2\psi_t(X_t)$.

PDE Dynamics

Planning Mean Field Games System

Let P solution of MFSP, assume $\Psi = \nabla\psi$ the corrector.

Under regularity assumptions, it holds that $(\mu_t = (X_t)_\# P, \psi_t)$ formally solves

$$\begin{cases} \partial_t \psi_t(x) + \frac{1}{2}\Delta\psi_t(x) + \frac{1}{2}|\nabla\psi_t(x)|^2 = \int \nabla W(x-\tilde{x}) \cdot (\nabla\psi_t(x) - \nabla\psi_t(\tilde{x}))\mathrm{d}\mu_t(\tilde{x}) \\ \partial_t \mu_t(x) - \frac{1}{2}\Delta\mu_t(x) + \nabla \cdot ((-(\nabla W * \mu_t)(x) + \nabla\psi_t(x))\mu_t(x)) = 0 \\ \mu_0(x) = \mu^{\mathrm{in}}(x), \mu_T(x) = \mu^{\mathrm{fin}}(x) \end{cases}$$

$\rightarrow$ Hamilton-Jacobi-Bellman PDE is related to microscopic dynamic of the particles;

$\rightarrow$ Fokker-Planck PDE describes the macroscopic dynamic of the system.

Functional inequalities

This new geometric perspective to study interpolation between measures, allows to study how entropy-like functionals evolve along this interpolation as done in optimal transport.

Let $\mathcal{C}_T(\mu^{\text{in}}, \mu^{\text{fin}}) := \inf_{\substack{P \in \mathcal{P}_1(\Omega) \\ P_0 = \mu^{\text{in}}, P_1 = \mu^{\text{fin}}}} \mathrm{KL}(P \mid \Gamma(P))$, under (H1), (H2) and (H3),

A Talagrand inequality

$$\mathcal{C}_T(\mu^{\text{in}}, \mu_\infty) \leq \frac{1}{\exp(2kT) - 1} \mathcal{F}(\mu^{\text{in}})$$

→ $\mathcal{F}(\mu^{\text{in}})$ controls the cost of steering the system from μ^{in} to equilibrium.

→ if we give the system a long time, reaching equilibrium is exponentially cheap.

Functional inequalities

Let $\quad \mathcal{C}_T(\mu^{\text{in}}, \mu^{\text{fin}}) := \inf_{\substack{P \in \mathcal{P}_1(\Omega) \\ P_0 = \mu^{\text{in}}, P_1 = \mu^{\text{fin}}}}$ KL$(P \mid \Gamma(P))$, under (H1), (H2), (H3):

An HWI inequality

$$\mathcal{F}(\mu^{\text{in}}) \leq \frac{1 - \exp(-2kT)}{2k} \left(\mathcal{I}_{\mathcal{F}}(\mu^{\text{in}}) \left(\frac{1}{4} \mathcal{I}_F(\mu^{\text{in}}) - \mathcal{E}_P(\mu^{\text{in}}, \mu_\infty) \right) \right)^{1/2} - (1 - \exp(-2kT)) \mathcal{C}_T(\mu^{\text{in}}, \mu_\infty)$$

with $\mathcal{I}_F$ nonlinear Fisher information.

- $\rightarrow$ Intuitively, $\mathcal{F}(\mu^{\text{in}}) \leq$ Fisher information - cost
- $\rightarrow$ $\mathcal{I}_{\mathcal{F}}$ measures the infinitesimal tendency of the particle flow to dissipate free energy; $\mathcal{C}_T$ is the effort needed to realize the interpolation.

Functional inequalities

$$\mathcal{C}_T(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}}) := \inf_{\substack{P \in \mathcal{P}_1(\Omega) \\ P_0 = \mu^{\mathsf{in}}, P_1 = \mu^{\mathsf{fin}}}} \mathsf{KL}(P \mid \Gamma(P))$$

$$\mathcal{E}_P(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}}) := \mathbb{E}_P[\Psi_t(X_t) \cdot \hat{\Psi}_{T-t}(\hat{X}_{T-t})]$$

(constant in t, conserved energy along the interpolation)

Energy-transport inequality

$$|\mathcal{E}_P(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}})| \leq \frac{4k}{\exp(kT) - 1} \Big(\mathcal{C}_T(\mu^{\mathsf{in}}, \mu^{\mathsf{fin}}) \mathcal{C}_T(\mu^{\mathsf{fin}}, \mu^{\mathsf{in}}) \Big)^{1/2}$$

$\rightarrow$ The conserved energy of the bridge is controlled by the forward and backward entropic transportation costs.

$\rightarrow$ For T large, the energy conserved is exponentially small (turnpike).

Fisher functionals

Definition

We define the nonlinear Fisher functional $\mathcal{I}_{\mathcal{F}}$ as

$$\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto \mathcal{I}_{\mathcal{F}}(\mu) := \begin{cases} \int_{\mathbb{R}^d} |\nabla \log \mu + 2(\nabla W * \mu)(x)|^2 \, \mathrm{d}\mu(x), & \text{if } \nabla \log \mu \in L^2_\mu \\ +\infty, & \text{otherwise} \end{cases}$$

It can be seen as the derivative of the free energy $\tilde{\mathcal{F}}$ along the marginal flow.

Wasserstein gradient of Fisher functional

We have that the Wasserstein gradient of Fisher information is

$$\nabla_{\mathcal{W}_2} \mathcal{I}(\mu) = -2\nabla\Delta \log \mu - \nabla|\nabla \log \mu|^2$$

Definitions from Riemannian geometry

Covariant derivative

Given a curve $(\mu_t)_t$, with velocity field $(v_t)_t$, we define the *covariant derivative* of v_t as

$$\frac{\mathbf{D}}{dt} v_t = \partial_t v_t + \frac{1}{2}\nabla |v_t|^2$$

Dynamics between measures

- $\rightarrow$ OT moves along constant-speed geodesics; mean-field Schrödinger problem interpolates along accelerated trajectories.

$\rightarrow$ Jordan, Kinderlehrer and Otto, *The Variational Formulation of the Fokker-Planck Equation* (1998):

Langevin diffusion $\leftrightarrow$ Fokker-Planck PDE $\leftrightarrow$ I-order ODE in the space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

Dynamics between measures

→ OT moves along constant-speed geodesics; mean-field Schrödinger problem interpolates along accelerated trajectories.

→ Jordan, Kinderlehrer and Otto, *The Variational Formulation of the Fokker-Planck Equation* (1998):

Langevin diffusion ↔ Fokker-Planck PDE ↔ I-order ODE in the space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

Newton's Law

Newton's law on $\mathcal{P}_2(\mathbb{R}^d)$ with Otto calculus (formal) Riemannian structure, taking v_t velocity field from Benamou-Brenier formula:

Optimal transport:

$$\begin{cases} \frac{\mathbf{D}}{\mathrm{d}t} v_t = 0 \\ \mu_0 = \mu^{\text{in}}, \quad \mu_1 = \mu^{\text{fin}} \end{cases}$$

MFSP:

$$\begin{cases} \frac{\mathbf{D}}{\mathrm{d}t} v_t = \frac{1}{8} \nabla_{\mathcal{W}_2} \mathcal{I}_F(\mu_t) \\ \mu_0 = \mu^{\text{in}}, \quad \mu_1 = \mu^{\text{fin}} \end{cases}$$

with $\mathcal{I}_F$ is the nonlinear Fisher information functional.

Proof sketch: Large deviations principle

- One starts from the fact that for the classic Schrödinger problem $\left\{\frac{1}{N}\sum_{i=1}^{N}\delta_{B^i} \mid N \in \mathbb{N}\right\}$ satisfy a large deviations principle with rate function $\mathrm{KL}(\cdot \mid R^{\mu^{\mathrm{in}}})$.
- Then, one defines a function $\Theta : (\mathcal{P}_1(\Omega), W_1) \to (\mathcal{P}_1(\Omega), W_1)$ such that $\Theta(Q) = Y^Q_{\#}Q$ where Y^Q is a stochastic process which encodes an interaction term to Q.
- Using Lipschitzianity of Θ, one proves a contraction principle such that Q satisfying LDP implies $\Theta(Q)$ satisfying LDP.
- By choosing $Q = \frac{1}{N}\sum_{i=1}^{N}\delta_{B^i}$, one finds the LDP for MFSP with desired rate function.

$\square$