The Mean Field Schrödinger problem

Daniela Tonon
an ongoing collaboration with Giovanni Conforti (École Polytechnique) and Richard Kraaij (TU Delft)

CEREMADE, Université Paris Dauphine

Two-days online workshop on Mean Field Games
June 18th 2020

Dauphine UNIVERSITÉ PARIS | PSL CEREMADE UMR 7534

The Schrödinger problem

In his seminal article "Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique" in the Ann. Inst. Henri Poincaré '32 Schrödinger wrote

*Imaginez que vous observez un **système de particules en diffusion**, qui soient en **équilibre thermodynamique.** Admettons qu'à un instant donné* 0 *vous les ayez trouvées en **répartition à peu près uniforme** et qu'à* T *vous ayez trouvé un **écart spontané et considérable par rapport à cette uniformité**. On vous demande de quelle manière cet écart s'est produit. Quelle en est la **manière** la plus probable?*

In plain words, the Schrödinger problem (SP) is the problem of finding the most likely evolution of a cloud of independent Brownian particles conditionally on the observation of their initial and final configuration, i.e. an entropy minimization problem with marginal constraints

The Schrödinger problem

In his seminal article "Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique" in the Ann. Inst. Henri Poincaré '32 Schrödinger wrote

*Imaginez que vous observez un **système de particules en diffusion**, qui soient en **équilibre thermodynamique.** Admettons qu'à un instant donné* 0 *vous les ayez trouvées en **répartition à peu près uniforme** et qu'à T vous ayez trouvé un **écart spontané et considérable par rapport à cette uniformité**. On vous demande de quelle manière cet écart s'est produit. Quelle en est la **manière** la plus probable?*

SP is the object of a very dynamic research activity:

It has powerful connections with the theory of Large Deviations, PDEs, Optimal transport, statistical machine learning and numerical algorithms for PDE related problems

KEY IDEA: SP may be viewed as a (entropic) regularization of the Optimal Transport problem

The Mean Field Schrödinger problem

The Mean Field Schrödinger Problem (MFSP) is obtained by replacing in the previous description the independent particles by interacting ones

Interacting Particle System

$(\Omega, \mathcal{F}_t, \mathcal{F}_T)$ where $\Omega = C([0,T];\mathbb{R}^d)$ with the uniform topology and $\{\mathcal{F}_t\}_{t\in[0,T]}$ the coordinate filtration

Interaction Potential: a symmetric $\mathcal{C}^2$ function $W:\mathbb{R}^d \to \mathbb{R}$ s.t. $\sup_{z,v\in\mathbb{R}^d,|v|=1} v\cdot\nabla^2 W(z)\cdot v < +\infty$

For N large, we consider Brownian particles $(X_t^{i,N})_{t\in[0,T],1\le i\le N}$

$$\begin{cases} \mathrm{d}X_t^{i,N} = & -\frac{1}{N}\sum_{k=1}^N \nabla W(X_t^{i,N}-X_t^{k,N})\mathrm{d}t + \mathrm{d}B_t^i \\ X_0^{i,N} \sim & \mu^{\mathrm{in}} \in \mathcal{P}_2(\mathbb{R}^d) \end{cases}$$

Driving Question: If at time T we observe that the sequence of empirical path measures

$$\frac{1}{N}\sum_{i=1}^N \delta_{X_T^{i,N}} \approx \mu^{\mathrm{fin}} \in \mathcal{P}_2(\mathbb{R}^d),$$

what have done the particles in between?

The Mean Field Schrödinger problem

The Mean Field Schrödinger Problem (MFSP) is obtained by replacing in the previous description the independent particles by interacting ones

Interacting Particle System

$(\Omega, \mathcal{F}_t, \mathcal{F}_T)$ where $\Omega = C([0,T]; \mathbb{R}^d)$ with the uniform topology and $\{\mathcal{F}_t\}_{t\in[0,T]}$ the coordinate filtration

Interaction Potential: a symmetric $\mathcal{C}^2$ function $W : \mathbb{R}^d \to \mathbb{R}$ s.t. $\sup_{z,v\in\mathbb{R}^d, |v|=1} v \cdot \nabla^2 W(z) \cdot v < +\infty$

For N large, we consider Brownian particles $(X_t^{i,N})_{t\in[0,T], 1\le i\le N}$

$$\begin{cases} \mathrm{d}X_t^{i,N} = & -\frac{1}{N}\sum_{k=1}^N \nabla W(X_t^{i,N} - X_t^{k,N})\mathrm{d}t + \mathrm{d}B_t^i \\ X_0^{i,N} \sim & \mu^{\mathrm{in}} \in \mathcal{P}_2(\mathbb{R}^d) \end{cases}$$

Under suitable assumptions, the problem is equivalent to

"minimizing the LDP rate function among all path measures whose marginal at time 0 is μ^{in} and whose marginal at time T is μ^{fin}"

Denote by

$$\Pi(\mu^{\mathrm{in}}, \mu^{\mathrm{fin}}) := \left\{ P \in \mathcal{P}_1(C([0,T];\mathbb{R}^d)) : P_0 = \mu^{\mathrm{in}}, P_T = \mu^{\mathrm{fin}} \right\}$$

and for $P, Q \in \mathcal{P}_1(C([0,T];\mathbb{R}^d))$, let $\mathcal{H}(P|Q)$ denote the relative entropy of P with respect to Q,

$$\mathcal{H}(P|Q) = \begin{cases} \mathbb{E}_P \left[\log\left(\frac{\mathrm{d}P}{\mathrm{d}Q} \right) \right] & P \ll Q \\ +\infty & \text{otherwise} \end{cases}$$

$\frac{\mathrm{d}P}{\mathrm{d}Q}$ denotes the Radon-Nikodym density of P against Q

The mean field Schrödinger problem can be stated as

$$\mathscr{C}_T(\mu^{\mathrm{in}},\mu^{\mathrm{fin}}) := \inf\left\{\mathcal{H}(P|\Gamma(P)) \; : \; P\in\Pi(\mu^{\mathrm{in}},\mu^{\mathrm{fin}})\right\}$$

where $\Gamma(P)$ is the law of the unique solution to

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

Its optimal value is called mean field entropic transportation cost and its optimizers are called mean field Schrödinger bridges (MFSB)

Theorem (Backhoff, Conforti, Gentil, Léonard '19)

Under mild assumptions MFSB exist

Uniqueness is still an open question

Equivalent Formulations (BCGL '19)

Benamou-Brenier Formulation:

It relates to the well known fluid dynamics representation of the Monge Kantorovich distance due to Benamou and Brenier that has been recently extended to the standard entropic transportation cost

$$\inf \frac{1}{2}\int_0^T \int_{\mathbb{R}^d} \left| w_t(z) + \frac{1}{2}\nabla \log \mu_t(z) + \nabla W * \mu_t(z) \right|^2 \mu_t(\mathrm{d}z)\mathrm{d}t$$

over all absolutely continuous curves $(\mu_t)_{t\in[0,T]} \subset \mathcal{P}_2(\mathbb{R}^d)$ s.t.

$$(t,z) \mapsto \nabla \log \mu_t(z) \in L^2\left(\mathrm{d}\mu_t \mathrm{d}t\right) \quad (t,z) \mapsto \nabla W * \mu_t(z) \in L^2\left(\mathrm{d}\mu_t \mathrm{d}t\right)$$

and that are weak solutions of the following continuity equation

$$\partial_t \mu_t + \nabla \cdot (w_t \mu_t) = 0 \quad \mu_0 = \mu^{\mathrm{in}},\ \mu_T = \mu^{\mathrm{fin}}$$

This formulation allows to interpret (MFSP) as a control problem in the Riemannian manifold of optimal transport

Connections with MFG

Theorem (BCGL '19)

Let P be an optimizer for (MFSP). Then there exists a weak gradient field Ψ s.t.

$$\mathrm{d}X_t = (\Psi_t(X_t) - \nabla W * P_t(X_t))\mathrm{d}t + dB_t$$

Now, set $\mu_t = (X_t)_{\#}P$ for all $t \in [0, T]$ and let μ and Ψ be $\mathcal{C}^{1,2}$, $\mu > 0$
Then there exists $\psi : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\Psi_t(x) = \nabla\psi_t(x) \quad \forall t \in [0, T], x \in \mathbb{R}^d$$

and $(\psi(\cdot), \mu(\cdot))$ is a classical solution of the following mean field planning PDE system

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

This type of PDE system has a similar structure to the planning MFG

Connections with MFG

Benamou, Carlier, Di Marino, Nenna '19 proposed an entropy minimization viewpoint on variational MFG of this type

$$\begin{cases} -\partial_t \psi_t - \frac{1}{2}\Delta\psi_t + \frac{1}{2}|\nabla\psi_t|^2 = f[\mu_t] & \text{in } (0,T)\times\mathbb{R}^d \\ \partial_t \mu_t - \frac{1}{2}\Delta\mu_t - \nabla\cdot(\mu_t\nabla\psi_t) = 0 & \text{in } (0,T)\times\mathbb{R}^d \\ \mu|_{t=0} = \mu_0,\ \psi_T = g[\mu_T] \end{cases}$$

and developed a suitable efficient algorithm (using the Sinkhorn algorithm) based on this entropic interpretation

The starting point of their analysis is the equivalence between the classical Schrödinger bridge problem and the optimal control (with kinetic energy as cost) of the Fokker-Planck equation

Connections with MFG

Benamou, Carlier, Di Marino, Nenna '19 proposed an entropy minimization viewpoint on variational MFG of this type

$$\begin{cases} -\partial_t \psi_t - \frac{1}{2}\Delta\psi_t + \frac{1}{2}|\nabla\psi_t|^2 = f[\mu_t] & \text{in } (0,T)\times\mathbb{R}^d \\ \partial_t \mu_t - \frac{1}{2}\Delta\mu_t - \nabla\cdot(\mu_t\nabla\psi_t) = 0 & \text{in } (0,T)\times\mathbb{R}^d \\ \mu|_{t=0} = \mu_0, \;\; \psi_T = g[\mu_T] \end{cases}$$

and developed a suitable efficient algorithm (using the Sinkhorn algorithm) based on this entropic interpretation

IDEA: we control the state variable μ through a vector field $v : (0,T)\times\mathbb{R}^d \to \mathbb{R}^d$ in order to minimize

$$\frac{1}{2}\int_0^T\int_{\mathbb{R}^d} |v_t|^2\mu_t(\mathrm{d}x)\mathrm{d}t + \int_0^T F(\mu_t)\mathrm{d}t + G(\mu_T)$$

when μ solves $\partial_t\mu + \nabla\cdot(\mu v) = 0$ with $\mu|_{t=0} = \mu_0$

The ergodic problem

Assume now that W is convex, then the particles system is rapidly mixing and there is a well defined equilibrium μ_∞

To the coupled HJB-FP systems we can associate the ergodic problem with unknowns (λ, ψ, μ)

$$\begin{cases} \lambda + \frac{1}{2}\Delta\psi(x) + \frac{1}{2}|\nabla\psi(x)|^2 = \int_{\mathbb{R}^d} \nabla W(x-\tilde{x})\cdot(\nabla\psi(x) - \nabla\psi(\tilde{x}))\mu(\mathrm{d}\tilde{x}) \\ -\frac{1}{2}\Delta\mu(x) + \nabla\cdot((-\nabla W * \mu(x) + \nabla\psi(x))\mu(x)) = 0 \end{cases}$$

The equilibrium solution $(0, 0, \mu_\infty)$ is a solution to the above equation

These systems have a broad range of applications:

- in the theory of MFGs they describe Nash equilibria of a large number of players;
- when minimizing the rate function associated with a Large Deviations principle or the objective function of a McKean Vlasov control problem they express necessary optimality conditions

Free energy functional

The free energy (or entropy) functional is defined for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ as

$$\mu \mapsto \tilde{\mathcal{F}}(\mu) := \begin{cases} \int \log \mu(x)\mu(\mathrm{d}x) + \int\int W(x-y)\mu(\mathrm{d}y)\mu(\mathrm{d}x) & \mu \ll \mathcal{L} \\ +\infty & \text{otherwise} \end{cases}$$

$\tilde{\mathcal{F}}$ is a Lyapunov function for the mean-field SDE

Its unique minimizer is the stationary solution μ_∞

BCGL ('19) give an answer to the following questions:

- For T large, how far is the time $T/2$ marginal, $P_{T/2}$, of a MFSB from μ_∞?
- For $t \ll T$, how far is the time t marginal, P_t, of a MFSB from the solution P^{MFSDE} of the mean-field SDE?

P^{MFSDE} is the law of the unique solution to McKean-Vlasov Dynamics SDE

$$\begin{cases} \mathrm{d}X_t = & -\nabla W * \mu_t(X_t)\mathrm{d}t + \mathrm{d}B_t \\ X_0 \sim & \mu^{\mathrm{in}}, \quad \mu_t = Law(X_t) \quad \forall t \in [0,T] \end{cases}$$

The non linear Fisher information functional

The non linear Fisher information functional $\mathcal{I}_{\tilde{\mathcal{F}}}$ is defined for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$

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

where by $\nabla \log \mu \in L^2_\mu$ we mean $\mu \ll \mathcal{L}$ (the Lebesgue measure) and $\log \mu$ is an absolutely continuous function in $\mathbb{R}^d$ whose derivative is in L^2_μ

The non linear Fisher information functional is formally the expected value of the observed information

The non linear Fisher information functional can be seen to be equal to the gradient of the free energy $\tilde{\mathcal{F}}$ along the marginal flow of the McKean Vlasov dynamics

The non linear Fisher information functional is used to state an HWI inequality, a powerful functional inequality relating the relative entropy (H) to the quadratic transport cost (W) and the Fisher information (I)

An HJB equation on the space of probability measures

Mimicking the well-known duality between the Monge-Kantorovich problem and the Hamilton-Jacobi equation, the MFSP can be formally seen as in duality with the solution of an infinite dimensional Hamilton-Jacobi-Bellmann (HJB) equation in $\mathcal{P}_2(\mathbb{R}^d)$

Let us modify the problem adding a penalization at the final time and removing the corresponding marginal constraint

For all $t \in [0, T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ we define

$$u(t, \mu) := \inf\{\mathcal{H}(P|\Gamma(P)) + \mathcal{G}(P_T) : P \in \mathcal{P}_1(\Omega), P_t = \mu\}$$

As we have seen for the classical MFSP, the previous problem can be rewritten equivalently as

$$u(t,\mu) := \inf \frac{1}{2}\int_t^T\int_{\mathbb{R}^d}\left| w_s(z) + \frac{1}{2}\nabla \log \mu_s(z) + \nabla W * \mu_s(z)\right|^2 \mu_s(\mathrm{d}z)\mathrm{d}s + \mathcal{G}(\mu_T)$$

over all absolutely continuous curves $(\mu_s)_{s\in[t,T]} \subset \mathcal{P}_2(\mathbb{R}^d)$ s.t. that are weak solutions of the following continuity equation

$$\partial_s \mu_s + \nabla\cdot(w\mu_s) = 0 \quad \mu_t = \mu,$$

Then the optimal value $u(t,\mu)$ is a candidate solution for an HJB equation on the space of probability measures

Since the articles of Crandall and Lions '84 on infinite dimensional HJB equations, the last years have witnessed a massive scientific production around the study of these equations

Several different strategies:

- Lifting of functions: we associate to any $v : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ a function V defined on $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ by setting for any random variable $X \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$
$$V(X) = v(\mu)$$
where $\mu \in \mathcal{P}_2(\mathbb{R}^2)$ is such that $\mu = Law(X)$
For derivatives we use Lions derivative exploiting the Hilbert space properties of $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$
- use the intrinsic notion of derivative on Wasserstein spaces that comes form optimal transport theory
- when the infinite dimensional HJB equations is associated to controlled gradient flows of a free energy, as in our case, a powerful approach exploiting the geometry of the underlying control problem and the HWI inequality is the one developed by Feng and collaborators
- ...

The HJB equation on the space of probability measures

Formally the HJB equation looks like

$$\begin{cases} -\partial_t u(t,\mu) + Hu(t,\mu) = 0, \\ u(T,\mu) = \mathcal{G}(\mu_T) \end{cases}$$

where the Hamiltonian is written as an operator over functions on $\mathcal{P}_2(\mathbb{R}^d)$

$$Hf(\mu) = \frac{1}{2}\int_{\mathbb{R}^d}\left(\mathrm{grad}^{W_2} f(\mu)\cdot \mathrm{grad}^{W_2}\tilde{\mathcal{F}}(\mu)\right)\mu(\mathrm{d}x) + \frac{1}{2}\int_{\mathbb{R}^d}|\mathrm{grad}^{W_2} f(\mu)|^2\mu(\mathrm{d}x)$$

Note that

$$H\tilde{\mathcal{F}}(\mu) = \mathcal{I}_{\tilde{\mathcal{F}}}(\mu)$$

indeed $\mathcal{I}_{\tilde{\mathcal{F}}}(\mu) = \int_{\mathbb{R}^d}|\mathrm{grad}^{W_2}\tilde{\mathcal{F}}(\mu)|^2\mu(\mathrm{d}x)$

IDEA: the problem can be represented as an infinite dimensional gradient flow for the free energy functional $\tilde{\mathcal{F}}$

$$u(t,\mu) := \inf \frac{1}{2}\int_t^T\int_{\mathbb{R}^d} |v_s(z)|^2 \mu_s(\mathrm{d}z)\mathrm{d}s + \mathcal{G}(\mu_T)$$

over all absolutely continuous curves $(\mu_s)_{s\in[t,T]} \subset \mathcal{P}_2(\mathbb{R}^d)$ s.t. that are weak solutions of the following continuity equation

$$\dot{\mu}_s = -\frac{1}{2}\mathrm{grad}^{W_2}\tilde{\mathcal{F}}(\mu_s) + v(s) \quad \mu_t = \mu$$

for a control $v(s) \in T_{\mu_s}\mathcal{P}_2(\mathbb{R}^d)$

A non-trivial obstruction to the adaptation of Feng's technique to this setup is that the free energy $\tilde{\mathcal{F}}$ does not have compact level sets

IDEA: use of Tataru's distance and Ekeland's optimization principle

Some properties of the gradient flow:

In the following, let us call

$$d := W_2, \quad \mathcal{E} := \frac{1}{2}\tilde{\mathcal{F}} \quad \text{and} \quad \mathcal{I}_{\mathcal{E}} := \|\mathrm{grad}^{W_2}\mathcal{E}\|^2 = \frac{1}{4}\mathcal{I}_{\tilde{\mathcal{F}}}$$

For simplicity, let us consider the stationary HJ equation

$$u - \lambda Hu = h$$

with $Hu = \frac{1}{2}\|\mathrm{grad}^{W_2}u\|^2 - \langle \mathrm{grad}^{W_2}\mathcal{E}, \mathrm{grad}^{W_2}u\rangle$

- The metric d is such that $\forall \rho, \gamma \in \mathcal{P}_2(\mathbb{R}^d)$

$$\|\frac{1}{2}\mathrm{grad}^{W_2}d^2(\rho,\gamma)\|^2 = d^2(\rho,\gamma)$$

- Let $S(t)$ be the semigroup generated by the gradient flow

$$\dot{\mu}_t = -\frac{1}{2}\mathrm{grad}^{W_2}\tilde{\mathcal{F}}(\mu_t)$$

Some properties of the gradient flow:

- Let μ_t be the gradient flow, then for any $0 \le t \le T$

$$\mathcal{E}(\mu_t) - \mathcal{E}(\mu(0)) \le -\int_0^t \mathcal{I}_{\mathcal{E}}(\mu(r))\mathrm{d}r$$

Moreover, there exists $\kappa \in \mathbb{R}$ s.t. for any $\gamma \in \mathcal{P}_2(\mathbb{R}^d)$ and any $t \in [0, T]$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(d^2(\mu(t),\gamma)\right) \le \mathcal{E}(\gamma) - \mathcal{E}(\mu(t)) - \frac{\kappa}{2}d^2(\mu(t),\gamma)$$

- HWI inequality: Let $\forall \mu, \gamma \in \mathcal{P}_2(\mathbb{R}^d)$
If $\mathcal{I}_{\mathcal{E}}(\mu) < \infty$, then

$$\langle -\mathrm{grad}^{W_2}\mathcal{E}(\mu), \frac{1}{2}\mathrm{grad}^{W_2}d^2(\mu,\gamma)\rangle \le \mathcal{E}(\gamma) - \mathcal{E}(\mu) - \frac{\kappa}{2}d^2(\mu,\gamma)$$

Strategy for the comparison principle

We recall the stationary HJ equation

$$u - \lambda Hu = h$$

with $Hu = \frac{1}{2}\|\mathrm{grad}^{W_2} u\|^2 - \langle \mathrm{grad}^{W_2}\mathcal{E}, \mathrm{grad}^{W_2} u\rangle$

Notation: we say that $(f, g) \in H$ if f belongs to the domain of H and $g \in Hf$

Definition: We say that $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ is a (viscosity) subsolution if u is bounded, upper semi-continuous and if for all $(f, g) \in H$ there exists a $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ such that

$$u(\rho_0) - f(\rho_0) = \sup_{\rho} u(\rho) - f(\rho),$$

$$u(\rho_0) - \lambda g(\rho_0) - h(\rho_0) \leq 0$$

Strategy for the comparison principle

We recall the stationary HJ equation

$$u - \lambda H u = h$$

with $Hu = \frac{1}{2}\|\mathrm{grad}^{W_2} u\|^2 - \langle \mathrm{grad}^{W_2}\mathcal{E}, \mathrm{grad}^{W_2} u\rangle$

Notation: we say that $(f, g) \in H$ if f belongs to the domain of H and $g \in Hf$

Definition: We say that $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ is a (viscosity) supersolution if u is bounded, lower semi-continuous and if for all $(f, g) \in H$ there exists a $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ such that

$$u(\rho_0) - f(\rho_0) = \inf_{\rho} u(\rho) - f(\rho),$$
$$u(\rho_0) - \lambda g(\rho_0) - h(\rho_0) \geq 0$$

The comparison principle

Let w be a weak sub-solution to $u - \lambda Hu = h_1$ and let v be a weak super-solution to $u - \lambda Hu = h_2$. Then we have

$$\sup_{\mu} w(\mu) - v(\mu) \leq \sup_{\mu} h_1(\mu) - h_2(\mu)$$

REM: In general, the comparison principle proof relies upon test functions which behave like distance functions

For instance, in the $\mathbb{R}^d$ case, these test functions take the form $\frac{1}{2}|x - y|$

NOTE: In the infinite dimensional case, functions like d^2 are not necessarily included in the domain of the Hamiltonian

IDEA: if $\varphi(\mu) = \frac{1}{2} a d^2(\mu, \gamma)$ for some $a > 0$ and $\gamma \in \mathcal{P}_2(\mathbb{R}^d)$, then formally

$$H\varphi(\mu) = -a \langle \mathrm{grad}^{W_2} \mathcal{E}(\mu), \frac{1}{2} \mathrm{grad}^{W_2} d^2(\mu, \gamma) \rangle + \frac{1}{2} a^2 d^2(\mu, \gamma)$$

Applying HWI we get a proper upper bound

$$H\varphi(\mu) \leq a \left[\mathcal{E}(\gamma) - \mathcal{E}(\mu)\right] - a \frac{\kappa}{2} d^2(\rho, \gamma) + \frac{1}{2} a^2 d^2(\mu, \gamma)$$

This leads to the definition of a new Hamiltonian $H_\dagger$:

$$\mathcal{D}(H_\dagger) := \left\{ \varphi(\mu) = \frac{1}{2} a d^2(\mu, \gamma) \,\middle|\, \forall\, a > 0, \forall\, \gamma : \mathcal{E}(\gamma) < \infty \right\}$$

and for $\varphi(\mu) = \frac{1}{2} a d^2(\mu, \gamma)$ we set

$$H_\dagger \varphi(\mu) = a \left[\mathcal{E}(\gamma) - \mathcal{E}(\mu)\right] - a \frac{\kappa}{2} d^2(\rho, \gamma) + \frac{1}{2} a^2 d^2(\mu, \gamma) \geq H\varphi(\mu)$$

NOTE: for this new Hamiltonian optimizers did not exist in general

Ekeland's perturbed optimization principle claims that, if we add a small perturbation to the test function, we can always attain the extrema

New test functions $\Rightarrow$ new Hamiltonian $\widetilde{H}_\dagger$

IDEA: use of Tataru distance as a penalization function defined as

$$d_T(\mu,\nu) := \inf_{t\geq 0}\left\{t + e^{\hat{\kappa}t} d(\mu, S(t)\nu)\right\}$$

where $\hat{\kappa} = 0 \wedge \kappa$
i.e. we will work with test functions

$$f_0(\rho) = \frac{1}{2} a d^2(\rho,\gamma) + b d_T(\rho,\pi) + c$$

for $a, b > 0$, $c \in \mathbb{R}$, and γ, π such that $\mathcal{E}(\gamma) + \mathcal{E}(\pi) < \infty$ and modify the Hamiltonian in this way

$$\widetilde{H}_\dagger f_0(\rho) = a\left[\mathcal{E}(\gamma) - \mathcal{E}(\rho)\right] - a\frac{\kappa}{2} d^2(\rho,\gamma) + b + \frac{1}{2} a^2 d^2(\rho,\gamma) + abd(\rho,\gamma) + \frac{1}{2} b^2$$

Strategy for the comparison principle

$$
\begin{tikzcd}
 & \mathrm{H}_\dagger \arrow[r, "\mathrm{sub}"] & \widetilde{\mathrm{H}}_\dagger \\
\mathrm{H} \arrow[ur, dotted, "\mathrm{sub}"] \arrow[dr, dotted, "\mathrm{super}"'] & & \\
 & \mathrm{H}_\ddagger \arrow[r, "\mathrm{super}"'] & \widetilde{\mathrm{H}}_\ddagger
\end{tikzcd}
$$

The comparison principle holds for $\widetilde{H}_\dagger$ and $\widetilde{H}_\ddagger$ so that for them we know there is uniqueness of viscosity solutions, however a-priori it is unclear how to show that such solutions exist

It is much easier to construct viscosity solutions for approximations of the operators $H_\dagger$ and $H_\ddagger$ that are in terms of smooth test functions

Perspectives

Our Aims are:

- well-posedness of the Hamilton-Jacobi equation
- Existence of strong solutions for HJB (bounds on the derivatives and regularity results)
- Long time behavior
- Link with FBSDE and MFSP
- study a richer class of equations, possibly including a stochastic component modeling a source of common noise