

Policy Gradient Descent for Mean-field Stochastic Control

Zhenjie Ren

LaMME, Université Evry Paris-Saclay
joint work with Yupeng Bai, Louis-Pierre Chaintron, Songbo Wang

International Seminar on SDEs and Related Topics
October 16, 2026



Table of Contents

- 1 Policy Gradient Descent
- 2 Normal Policy Gradient Flow and Mirror Flow
- 3 Mirror Flow on Wiener Space
 - Bregman divergence for stochastic control
 - Mirror flow and its dual
 - Trinity: PGD, Mirror Flow, Dual Coordinate Flow
- 4 PGD for Mean-Field Stochastic Control
 - Extension to MFC
 - Long-time behavior

Stochastic Control

A simple **controlled** diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T].$$

Consider the cost functional (finance, fine tuning, RL...)

$$V^\beta = \mathbb{E} \left[\int_0^T \ell(X_t, \beta_t) dt + g(X_T) \right], \quad \nabla_\beta^2 \ell \geq \lambda I_d.$$

Markovian case

The control is taken as a deterministic state-feedback function

$$\beta_t = \beta(t, X_t), \quad \beta : [0, T] \times \mathcal{X} \rightarrow \mathbb{R}^d.$$

Thus, the marginal law of X_t is described by a Fokker–Planck PDE and the value function by HJB PDE.

Policy Gradient Descent in Stochastic Control

Policy evaluation: linear HJB (heat) equation

$$\begin{cases} -\partial_t V^\beta - \frac{1}{2} \Delta V^\beta - \langle \nabla V^\beta, \beta \rangle - \ell(x, \beta) = 0, \\ V^\beta(T, x) = g(x). \end{cases}$$

Policy Gradient Descent in Stochastic Control

Policy evaluation: linear HJB (heat) equation

$$\begin{cases} -\partial_t V^\beta - \frac{1}{2} \Delta V^\beta - \langle \nabla V^\beta, \beta \rangle - \ell(x, \beta) = 0, \\ V^\beta(T, x) = g(x). \end{cases}$$

Let the control evolves as β^τ . The calculus of variation gives

$$\partial_\tau V^{\beta^\tau} = \int_0^T \mathbb{E}^{\beta^\tau} \left[\left(\nabla_x V^{\beta^\tau}(t, X_t) + \nabla_\beta \ell(X_t, \beta^\tau(t, X_t)) \right) \cdot \frac{d\beta^\tau}{d\tau} \right] dt.$$

Policy Gradient Descent in Stochastic Control

Policy evaluation: linear HJB (heat) equation

$$\begin{cases} -\partial_t V^\beta - \frac{1}{2} \Delta V^\beta - \langle \nabla V^\beta, \beta \rangle - \ell(x, \beta) = 0, \\ V^\beta(T, x) = g(x). \end{cases}$$

Let the control evolves as β^τ . The calculus of variation gives

$$\partial_\tau V^{\beta^\tau} = \int_0^T \mathbb{E}^{\beta^\tau} \left[\left(\nabla_x V^{\beta^\tau}(t, X_t) + \nabla_\beta \ell(X_t, \beta^\tau(t, X_t)) \right) \cdot \frac{d\beta^\tau}{d\tau} \right] dt.$$

Therefore, it is natural to define the policy gradient flow:

$$\partial_\tau \beta^\tau(t, x) := - \left(\nabla_x V^{\beta^\tau}(t, x) + \nabla_\beta \ell(x, \beta^\tau(t, x)) \right)$$

Policy Gradient Descent in Stochastic Control

Policy evaluation: linear HJB (heat) equation

$$\begin{cases} -\partial_t V^\beta - \frac{1}{2} \Delta V^\beta - \langle \nabla V^\beta, \beta \rangle - \ell(x, \beta) = 0, \\ V^\beta(T, x) = g(x). \end{cases}$$

Let the control evolves as β^τ . The calculus of variation gives

$$\partial_\tau V^{\beta^\tau} = \int_0^T \mathbb{E}^{\beta^\tau} \left[\left(\nabla_x V^{\beta^\tau}(t, X_t) + \nabla_\beta \ell(X_t, \beta^\tau(t, X_t)) \right) \cdot \frac{d\beta^\tau}{d\tau} \right] dt.$$

Therefore, it is natural to define the policy gradient flow:

$$\partial_\tau \beta^\tau(t, x) := - \left(\nabla_x V^{\beta^\tau}(t, x) + \nabla_\beta \ell(x, \beta^\tau(t, x)) \right)$$

Value dissipation along the PGD flow

$$\partial_\tau V^{\beta^\tau} = - \int_0^T \mathbb{E}^{\beta^\tau} \left[\left| \nabla_x V^{\beta^\tau}(t, X_t) + \nabla_\beta \ell(X_t, \beta^\tau(t, X_t)) \right|^2 \right] dt \leq 0.$$

Convergence rate?

Global convergence and rate (Zhou-Lu, SICON '25)

Under uniform strong concavity of the Hamiltonian in β , and **additional assumptions**, the value converges exponentially quickly along the PGD:

$$V^{\beta^\tau} - V^{\beta^*} \leq e^{-\kappa\tau} (V^{\beta^0} - V^{\beta^*}).$$

Important limitation

The rate proof relies on a uniform-in- τ regularity assumption along the flow:

$$\beta^\tau \text{ remains smooth and } \|\beta^\tau\|_{C^{2,4}([0,T] \times \mathcal{X})} \leq K \text{ for all } \tau.$$

Together with Schauder-type PDE estimates, this provides the regularity needed in the convergence proof. The final convergence rate κ depends on the regularity estimates.

Table of Contents

- 1 Policy Gradient Descent
- 2 Normal Policy Gradient Flow and Mirror Flow**
- 3 Mirror Flow on Wiener Space
 - Bregman divergence for stochastic control
 - Mirror flow and its dual
 - Trinity: PGD, Mirror Flow, Dual Coordinate Flow
- 4 PGD for Mean-Field Stochastic Control
 - Extension to MFC
 - Long-time behavior

Normal gradient flow in Euclidean space

We will study the PGD through the lens of mirror flow.

Why are GD and MF connected?

Normal gradient flow in Euclidean space

We will study the PGD through the lens of mirror flow.

Why are GD and MF connected?

For an objective $F(\theta)$, the usual gradient flow is

$$\dot{\theta}_t = -\nabla F(\theta_t),$$

the steepest descent dynamics when distance is measured by $\|\theta - \theta_t\|$.

Normal gradient flow in Euclidean space

We will study the **PGD** through the lens of **mirror flow**.

Why are **GD** and **MF** connected?

For an objective $F(\theta)$, the usual gradient flow is

$$\dot{\theta}_t = -\nabla F(\theta_t),$$

the steepest descent dynamics when distance is measured by $\|\theta - \theta_t\|$.

Normal/preconditioned gradient flow

Choose a strictly convex potential Φ . The preconditioned gradient flow is

$$\frac{d}{dt} \nabla \Phi(\theta_t) = -\nabla F(\theta_t), \quad \text{equivalently} \quad \nabla^2 \Phi(\theta_t) \dot{\theta}_t = -\nabla F(\theta_t).$$

In particular, if $\Phi := F$, we call $\frac{d}{dt} \nabla F(\theta_t) = -\nabla F(\theta_t)$ the normal gradient flow.

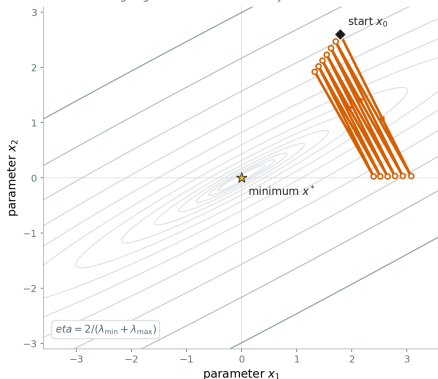
Comparison: GD and Normal GD

Ill-conditioning bends the geometry seen by gradient descent

Same quadratic $f(x) = \frac{1}{2}x^\top Hx$, same start, 12 iterations • condition number $\kappa(H) = 80$

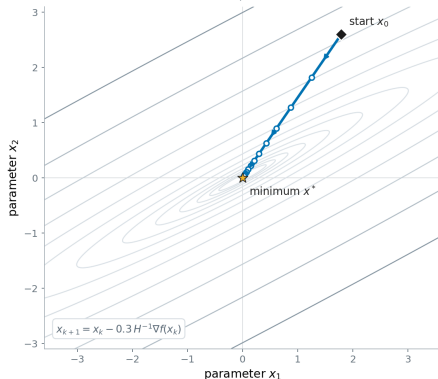
Gradient descent

The iterates zig-zag across the narrow valley



Hessian-preconditioned descent

H^{-1} rescales the curvature: a direct path



For this quadratic, a full Newton step ($\eta = 1$) reaches x^* in one iteration.

Preconditioned gradient flow = mirror flow

The Bregman divergence generated by a convex potential Φ is

$$D_{\Phi}(\theta, \bar{\theta}) = \Phi(\theta) - \Phi(\bar{\theta}) - \langle \nabla \Phi(\bar{\theta}), \theta - \bar{\theta} \rangle.$$

Mirror descent uses it as the local distance to define the JKO variation:

$$\theta_{k+1} = \arg \min_{\theta} \left\{ F(\theta) + \frac{1}{\eta} D_{\Phi}(\theta, \theta_k) \right\}.$$

Preconditioned gradient flow = mirror flow

The Bregman divergence generated by a convex potential Φ is

$$D_{\Phi}(\theta, \bar{\theta}) = \Phi(\theta) - \Phi(\bar{\theta}) - \langle \nabla \Phi(\bar{\theta}), \theta - \bar{\theta} \rangle.$$

Mirror descent uses it as the local distance to define the JKO variation:

$$\theta_{k+1} = \arg \min_{\theta} \left\{ F(\theta) + \frac{1}{\eta} D_{\Phi}(\theta, \theta_k) \right\}.$$

Local intuition: Bregman divergence as a local Hessian metric

For a small displacement v , Taylor expansion gives

$$D_{\Phi}(\theta_k + v, \theta_k) = \frac{1}{2} \langle v, \nabla^2 \Phi(\theta_k) v \rangle + o(\|v\|^2).$$

So Bregman geometry is locally the Hessian metric $\nabla^2 \Phi$. Therefore,

$$\theta_{k+1} \approx \arg \min_{\theta} \left\{ \langle \nabla F(\theta_k), \theta - \theta_k \rangle + \frac{1}{2\eta} \langle \theta - \theta_k, \nabla^2 \Phi(\theta_k) (\theta - \theta_k) \rangle \right\}.$$

Preconditioned gradient flow = mirror flow

The Bregman divergence generated by a convex potential Φ is

$$D_{\Phi}(\theta, \bar{\theta}) = \Phi(\theta) - \Phi(\bar{\theta}) - \langle \nabla \Phi(\bar{\theta}), \theta - \bar{\theta} \rangle.$$

Mirror descent uses it as the local distance to define the JKO variation:

$$\theta_{k+1} = \arg \min_{\theta} \left\{ F(\theta) + \frac{1}{\eta} D_{\Phi}(\theta, \theta_k) \right\}.$$

Local intuition: Bregman divergence as a local Hessian metric

For a small displacement v , Taylor expansion gives

$$D_{\Phi}(\theta_k + v, \theta_k) = \frac{1}{2} \langle v, \nabla^2 \Phi(\theta_k) v \rangle + o(\|v\|^2).$$

So Bregman geometry is locally the Hessian metric $\nabla^2 \Phi$. Therefore,

$$\theta_{k+1} \approx \theta_k - \eta \nabla^2 \Phi(\theta_k)^{-1} \nabla F(\theta_k), \quad \text{(Preconditioned Gradient Flow).}$$

Table of Contents

- 1 Policy Gradient Descent
- 2 Normal Policy Gradient Flow and Mirror Flow
- 3 Mirror Flow on Wiener Space**
 - Bregman divergence for stochastic control
 - Mirror flow and its dual
 - Trinity: PGD, Mirror Flow, Dual Coordinate Flow
- 4 PGD for Mean-Field Stochastic Control
 - Extension to MFC
 - Long-time behavior

Table of Contents

3 Mirror Flow on Wiener Space

- Bregman divergence for stochastic control
- Mirror flow and its dual
- Trinity: PGD, Mirror Flow, Dual Coordinate Flow

Our main contributions

Study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with non-Markovian cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt + \xi(X_{T\wedge\cdot}) \right] \quad \text{with } \nabla_\beta^2 \ell \geq cI_d, \ell \geq 0, \xi \text{ bounded.}$$

Main contributions

- Study the PGD for **non-Markovian**, **drift control**, **MFC**;

Our main contributions

Study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with non-Markovian cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt + \xi(X_{T\wedge\cdot}) \right] \quad \text{with } \nabla_{\beta}^2 \ell \geq cI_d, \ell \geq 0, \xi \text{ bounded.}$$

Main contributions

- Study the PGD for **non-Markovian**, **drift control**, **MFC**;
- Construct a Bregman divergence on Wiener space such that
 - Discrete Mirror descent converges to **continuous-time Mirror flow**;
 - **Preconditioned PGD = Continuous-time Mirror flow**;

Our main contributions

Study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with non-Markovian cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt + \xi(X_{T\wedge\cdot}) \right] \quad \text{with } \nabla_\beta^2 \ell \geq cI_d, \ell \geq 0, \xi \text{ bounded.}$$

Main contributions

- Study the PGD for **non-Markovian**, **drift control**, **MFC**;
- Construct a Bregman divergence on Wiener space such that
 - Discrete Mirror descent converges to **continuous-time Mirror flow**;
 - **Preconditioned PGD = Continuous-time Mirror flow**;
- Using the Mirror flow to obtain the **long-time convergence** rate of Preconditioned PGD, with an **intrinsic rate**.

Entropy and relative entropy on the Wiener space

Let $\mathbb{P}$ be the Wiener measure, and $\mathbb{Q}^\beta$ such that

$$\mathbb{Q}^\beta \sim \mathbb{P}, \quad dX_t = \beta_t dt + dW_t, \quad \mathbb{Q}^\beta\text{-a.s.}$$

By the Girsanov theorem, we have

$$\frac{d\mathbb{Q}^\beta}{d\mathbb{P}} = \exp\left(\int_0^T \beta_t dX_t - \frac{1}{2} \int_0^T |\beta_t|^2 dt\right).$$

Define the entropy:

$$H(\mathbb{Q}^\beta) := \int \ln(d\mathbb{Q}^\beta/d\mathbb{P}) d\mathbb{Q}^\beta = \mathbb{E}^\beta \left[\frac{1}{2} \int_0^T |\beta_t|^2 dt \right].$$

Entropy and relative entropy on the Wiener space

Let $\mathbb{P}$ be the Wiener measure, and $\mathbb{Q}^\beta$ such that

$$\mathbb{Q}^\beta \sim \mathbb{P}, \quad dX_t = \beta_t dt + dW_t, \quad \mathbb{Q}^\beta\text{-a.s.}$$

Define the entropy:

$$H(\mathbb{Q}^\beta) := \int \ln(d\mathbb{Q}^\beta/d\mathbb{P}) d\mathbb{Q}^\beta = \mathbb{E}^\beta \left[\frac{1}{2} \int_0^T |\beta_t|^2 dt \right].$$

Define the first variation

$$\begin{aligned} \delta H(\mathbb{Q}^\beta, X) &:= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(H(\mathbb{Q}^\beta + \varepsilon \cdot) - H(\mathbb{Q}^\beta) \right) \\ &= \ln(d\mathbb{Q}^\beta/d\mathbb{P}) = \int_0^T \beta_t dX_t - \frac{1}{2} \int_0^T |\beta_t|^2 dt. \end{aligned}$$

Entropy and relative entropy on the Wiener space

Let $\mathbb{P}$ be the Wiener measure, and $\mathbb{Q}^\beta$ such that

$$\mathbb{Q}^\beta \sim \mathbb{P}, \quad dX_t = \beta_t dt + dW_t, \quad \mathbb{Q}^\beta\text{-a.s.}$$

Define the entropy:

$$H(\mathbb{Q}^\beta) := \int \ln \left(d\mathbb{Q}^\beta / d\mathbb{P} \right) d\mathbb{Q}^\beta = \mathbb{E}^\beta \left[\frac{1}{2} \int_0^T |\beta_t|^2 dt \right].$$

Define the first variation

$$\delta H(\mathbb{Q}^\beta, X) := \int_0^T \beta_t dX_t - \frac{1}{2} \int_0^T |\beta_t|^2 dt.$$

Entropy and relative entropy on the Wiener space

Let $\mathbb{P}$ be the Wiener measure, and $\mathbb{Q}^\beta$ such that

$$\mathbb{Q}^\beta \sim \mathbb{P}, \quad dX_t = \beta_t dt + dW_t, \quad \mathbb{Q}^\beta\text{-a.s.}$$

Define the entropy:

$$H(\mathbb{Q}^\beta) := \int \ln(d\mathbb{Q}^\beta/d\mathbb{P}) d\mathbb{Q}^\beta = \mathbb{E}^\beta \left[\frac{1}{2} \int_0^T |\beta_t|^2 dt \right].$$

Define the first variation

$$\delta H(\mathbb{Q}^\beta, X) := \int_0^T \beta_t dX_t - \frac{1}{2} \int_0^T |\beta_t|^2 dt.$$

Finally, the relative entropy is a Bregman divergence:

$$\begin{aligned} H(\mathbb{Q}^{\beta'} | \mathbb{Q}^\beta) &:= H(\mathbb{Q}^{\beta'}) - H(\mathbb{Q}^\beta) - \int \delta H(\mathbb{Q}^\beta, \cdot) d(\mathbb{Q}^{\beta'} - \mathbb{Q}^\beta) \\ &= \mathbb{E}^{\beta'} \left[\frac{1}{2} |\beta'_t - \beta_t|^2 dt \right]. \end{aligned}$$

ℓ -entropy and ℓ -divergence for stochastic control

Let $\mathbb{P}$ be the Wiener measure. Consider the set of pairs

$$\mathcal{A} := \{(\mathbb{Q}^\beta, \beta) : \mathbb{Q}^\beta \sim \mathbb{P}, \, dX_t = \beta_t dt + dW_t, \, \mathbb{Q}^\beta\text{-a.s.}\}.$$

ℓ -entropy and ℓ -divergence for stochastic control

Let $\mathbb{P}$ be the Wiener measure. Consider the set of pairs

$$\mathcal{A} := \{(\mathbb{Q}^\beta, \beta) : \mathbb{Q}^\beta \sim \mathbb{P}, dX_t = \beta_t dt + dW_t, \mathbb{Q}^\beta\text{-a.s.}\}.$$

For $(\mathbb{Q}^\beta, \beta) \in \mathcal{A}$, we define the ℓ -entropy:

$$H_\ell(\mathbb{Q}^\beta) := \mathbb{E}^\beta \left[\int_0^T \ell(X_{s\wedge\cdot}, \beta_s) ds \right].$$

On $\mathcal{A}$, $\mathbb{Q}^\beta \mapsto H_\ell(\mathbb{Q}^\beta)$ is strictly flat-convex and weakly l.s.c..

ℓ -entropy and ℓ -divergence for stochastic control

Let $\mathbb{P}$ be the Wiener measure. Consider the set of pairs

$$\mathcal{A} := \{(\mathbb{Q}^\beta, \beta) : \mathbb{Q}^\beta \sim \mathbb{P}, dX_t = \beta_t dt + dW_t, \mathbb{Q}^\beta\text{-a.s.}\}.$$

For $(\mathbb{Q}^\beta, \beta) \in \mathcal{A}$, we define the ℓ -entropy:

$$H_\ell(\mathbb{Q}^\beta) := \mathbb{E}^\beta \left[\int_0^T \ell(X_{s\wedge\cdot}, \beta_s) ds \right].$$

On $\mathcal{A}$, $\mathbb{Q}^\beta \mapsto H_\ell(\mathbb{Q}^\beta)$ is **strictly flat-convex** and **weakly l.s.c.**. Define

$$\begin{aligned} \delta H_\ell(\mathbb{Q}^\beta) &:= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(H_\ell(\mathbb{Q}^\beta + \varepsilon \cdot) - H_\ell(\mathbb{Q}^\beta) \right) \\ &= \int \ell(\beta_t) dt + \int \nabla \ell(\beta_t) (dX_t - \beta_t dt). \end{aligned}$$

ℓ -entropy and ℓ -divergence for stochastic control

Let $\mathbb{P}$ be the Wiener measure. Consider the set of pairs

$$\mathcal{A} := \{(\mathbb{Q}^\beta, \beta) : \mathbb{Q}^\beta \sim \mathbb{P}, dX_t = \beta_t dt + dW_t, \mathbb{Q}^\beta\text{-a.s.}\}.$$

For $(\mathbb{Q}^\beta, \beta) \in \mathcal{A}$, we define the ℓ -entropy:

$$H_\ell(\mathbb{Q}^\beta) := \mathbb{E}^\beta \left[\int_0^T \ell(X_{s \wedge \cdot}, \beta_s) ds \right].$$

On $\mathcal{A}$, $\mathbb{Q}^\beta \mapsto H_\ell(\mathbb{Q}^\beta)$ is **strictly flat-convex** and **weakly l.s.c.**. Define

$$\delta H_\ell(\mathbb{Q}^\beta) := \int \ell(\beta_t) dt + \int \nabla \ell(\beta_t) (dX_t - \beta_t dt).$$

ℓ -entropy and ℓ -divergence for stochastic control

Let $\mathbb{P}$ be the Wiener measure. Consider the set of pairs

$$\mathcal{A} := \{(\mathbb{Q}^\beta, \beta) : \mathbb{Q}^\beta \sim \mathbb{P}, dX_t = \beta_t dt + dW_t, \mathbb{Q}^\beta\text{-a.s.}\}.$$

For $(\mathbb{Q}^\beta, \beta) \in \mathcal{A}$, we define the ℓ -entropy:

$$H_\ell(\mathbb{Q}^\beta) := \mathbb{E}^\beta \left[\int_0^T \ell(X_{s\wedge\cdot}, \beta_s) ds \right].$$

On $\mathcal{A}$, $\mathbb{Q}^\beta \mapsto H_\ell(\mathbb{Q}^\beta)$ is strictly flat-convex and weakly l.s.c.. Define

$$\delta H_\ell(\mathbb{Q}^\beta) := \int \ell(\beta_t) dt + \int \nabla \ell(\beta_t) (dX_t - \beta_t dt).$$

Finally define the ℓ -divergence:

$$\begin{aligned} H_\ell(\mathbb{Q}^{\beta'} | \mathbb{Q}^\beta) &= H_\ell(\mathbb{Q}^{\beta'}) - H_\ell(\mathbb{Q}^\beta) - \int \delta H_\ell(\mathbb{Q}^\beta, \omega) (\mathbb{Q}^{\beta'} - \mathbb{Q}^\beta)(d\omega) \\ &= \mathbb{E}^{\beta'} \left[\int_0^T \left(\ell(\beta'_t) - \ell(\beta_t) - \nabla_\beta \ell(\beta_t) (\beta'_t - \beta_t) \right) dt \right] \geq 0. \end{aligned}$$

Table of Contents

3 Mirror Flow on Wiener Space

- Bregman divergence for stochastic control
- **Mirror flow and its dual**
- Trinity: PGD, Mirror Flow, Dual Coordinate Flow

Discrete mirror descent and its dual

Recall that

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt + \xi(X_{T\wedge\cdot}) \right] = H_\ell(\mathbb{Q}^\beta) + \mathbb{E}^\beta[\xi].$$

Using the ℓ -divergence, we can define the **discrete mirror descent** by the JKO variational formula:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

Discrete mirror descent and its dual

Recall that

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt + \xi(X_{T\wedge\cdot}) \right] = H_\ell(\mathbb{Q}^\beta) + \mathbb{E}^\beta[\xi].$$

Using the ℓ -divergence, we can define the **discrete mirror descent** by the JKO variational formula:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

The first order condition (**under technical conditions**) reads

$$\delta H_\ell(\mathbb{Q}_{k+1}^h) + \xi + \frac{1}{h} (\delta H_\ell(\mathbb{Q}_{k+1}^h) - \delta H_\ell(\mathbb{Q}_k^h)) = 0,$$

which gives the **discrete dynamics of the dual coordinate** $\rho := \delta H_\ell(\mathbb{Q})$.

Discrete mirror descent and its dual

Recall that

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt + \xi(X_{T\wedge\cdot}) \right] = H_\ell(\mathbb{Q}^\beta) + \mathbb{E}^\beta[\xi].$$

Using the ℓ -divergence, we can define the **discrete mirror descent** by the JKO variational formula:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

The first order condition (**under technical conditions**) reads

$$\rho_{k+1}^h + \xi + \frac{1}{h} (\rho_{k+1}^h - \rho_k^h) = 0,$$

which gives the **discrete dynamics of the dual coordinate** $\rho := \delta H_\ell(\mathbb{Q})$.

Link to discrete Normal Policy Gradient descent

Recall the discrete mirror descent:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_{\ell}(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_{\ell}(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

The minimization problem can be viewed as a **stochastic control problem**,

Link to discrete Normal Policy Gradient descent

Recall the discrete mirror descent:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_{\ell}(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_{\ell}(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

The minimization problem can be viewed as a **stochastic control problem**, which can be characterized by a backward SDE (**BSDE**):

$$dY_t = -G(Z_t)dt + Z_t dW_t, \quad Y_T = \xi,$$

$$G(z) = \inf_{\beta} \left\{ \beta z + \ell(\beta) + \frac{1}{h} (\ell(\beta) - \ell(\beta_k^h) - \nabla \ell(\beta_k^h)(\beta - \beta_k^h)) \right\}.$$

Link to discrete Normal Policy Gradient descent

Recall the discrete mirror descent:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_{\ell}(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_{\ell}(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

The minimization problem can be viewed as a **stochastic control problem**, which can be characterized by a backward SDE (**BSDE**):

$$\begin{aligned} dY_t &= -G(Z_t)dt + Z_t dW_t, \quad Y_T = \xi, \\ G(z) &= \inf_{\beta} \left\{ \beta z + \ell(\beta) + \frac{1}{h} (\ell(\beta) - \ell(\beta_k^h) - \nabla \ell(\beta_k^h)(\beta - \beta_k^h)) \right\}. \end{aligned}$$

The optimal control satisfies

$$\nabla \ell(\beta_{k+1}^h) - \nabla \ell(\beta_k^h) = -h(Z_{k+1}^h + \nabla \ell(\beta_{k+1}^h)).$$

Note that in **Markovian case** $Z_{k+1}^h \approx \nabla V^{\beta_{k+1}^h}(X)$.

Link to discrete Normal Policy Gradient descent

Recall the discrete mirror descent:

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_{\ell}(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\xi] + \frac{1}{h} H_{\ell}(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

The minimization problem can be viewed as a **stochastic control problem**, which can be characterized by a backward SDE (**BSDE**):

$$\begin{aligned} dY_t &= -G(Z_t)dt + Z_t dW_t, \quad Y_T = \xi, \\ G(z) &= \inf_{\beta} \left\{ \beta z + \ell(\beta) + \frac{1}{h} (\ell(\beta) - \ell(\beta_k^h) - \nabla \ell(\beta_k^h)(\beta - \beta_k^h)) \right\}. \end{aligned}$$

The optimal control satisfies

$$\nabla \ell(\beta_{k+1}^h) - \nabla \ell(\beta_k^h) = -h(\mathbf{Z}_{k+1}^h + \nabla \ell(\beta_{k+1}^h)).$$

It coincides with **implicit Euler scheme for Normal Policy Gradient Flow**.

Table of Contents

3 Mirror Flow on Wiener Space

- Bregman divergence for stochastic control
- Mirror flow and its dual
- Trinity: PGD, Mirror Flow, Dual Coordinate Flow

Continuous-time limit

As the time step $h \rightarrow 0$, we prove that all three flows converge towards their **unique** continuous-time limits:

- $\mathbb{Q}_k^h \rightarrow \mathbb{Q}^\tau$;
- $\rho_{k+1}^h + \xi + \frac{1}{h}(\rho_{k+1}^h - \rho_k^h) = 0$ converges to $\dot{\rho} = -(\rho + \xi)$;
- $\nabla \ell(\beta_{k+1}^h) - \nabla \ell(\beta_k^h) = -h(Z_{k+1}^h + \nabla \ell(\beta_{k+1}^h))$ converges to

$$d\nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau)). \quad (\text{NPGD})$$

Continuous-time limit

As the time step $h \rightarrow 0$, we prove that all three flows converge towards their **unique** continuous-time limits:

- $\mathbb{Q}_k^h \rightarrow \mathbb{Q}^\tau$;
- $\rho_{k+1}^h + \xi + \frac{1}{h}(\rho_{k+1}^h - \rho_k^h) = 0$ converges to $\dot{\rho} = -(\rho + \xi)$;
- $\nabla \ell(\beta_{k+1}^h) - \nabla \ell(\beta_k^h) = -h(Z_{k+1}^h + \nabla \ell(\beta_{k+1}^h))$ converges to

$$d\nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau)). \quad (\text{NPGD})$$

Furthermore, the **trinity** of the triple $(\mathbb{Q}^h, \beta^h, \rho^h)$ is **preserved** in the limit:

- $dX_t = \beta_t^\tau dt + dW_t, \quad \mathbb{Q}^\tau\text{-a.s.};$
- $\delta H_\ell(\mathbb{Q}^\tau) = \rho_\tau.$

Analog from Euclidean to Wiener Space

Table: Notions involved in GD and MF

	Euclidean	Wiener
(P)GD	$\dot{x}_t = -\nabla F(x_t)$	$\dot{\beta} = -(\nabla V^\beta + \nabla \ell(\beta))$
Bregman	F	H_ℓ
Dual flow	$d\nabla F(x_t) = -\nabla F(x_t)dt$	$\rho := \delta H_\ell(\mathbb{Q}^\beta), \quad \dot{\rho} = -(\rho + \xi)$
N(P)GD	$d\nabla F(x_t) = -\nabla F(x_t)dt$	$d\nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau))d\tau$

Table of Contents

- 1 Policy Gradient Descent
- 2 Normal Policy Gradient Flow and Mirror Flow
- 3 Mirror Flow on Wiener Space
 - Bregman divergence for stochastic control
 - Mirror flow and its dual
 - Trinity: PGD, Mirror Flow, Dual Coordinate Flow
- 4 PGD for Mean-Field Stochastic Control**
 - Extension to MFC
 - Long-time behavior

Table of Contents

- 4 PGD for Mean-Field Stochastic Control
 - Extension to MFC
 - Long-time behavior

Mirror descent for Mean-field Control

Again, study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with **flat-convex** mean-field cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt \right] + F(\mathbb{Q}^\beta) \quad \text{with} \quad cI_d \leq \nabla_{\beta}^2 \ell \leq CI_d, \quad \ell \geq 0.$$

Mirror descent for Mean-field Control

Again, study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with **flat-convex** mean-field cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt \right] + F(\mathbb{Q}^\beta) \quad \text{with} \quad c\mathbf{I}_d \leq \nabla_\beta^2 \ell \leq C\mathbf{I}_d, \quad \ell \geq 0.$$

Then the mirror descent reads

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + F(\mathbb{Q}) + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\}.$$

Mirror descent for Mean-field Control

Again, study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with **flat-convex** mean-field cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt \right] + F(\mathbb{Q}^\beta) \quad \text{with} \quad c\mathbf{I}_d \leq \nabla_\beta^2 \ell \leq C\mathbf{I}_d, \quad \ell \geq 0.$$

Then the mirror descent reads

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\delta F(\mathbb{Q}_{k+1}^h)] + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\},$$

Mirror descent for Mean-field Control

Again, study a simple controlled diffusion:

$$dX_t = \beta_t dt + dW_t, \quad t \in [0, T], \quad \mathbb{Q}^\beta\text{-a.s.}$$

with **flat-convex** mean-field cost functional

$$V^\beta = \mathbb{E}^\beta \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt \right] + F(\mathbb{Q}^\beta) \quad \text{with} \quad c\mathbf{I}_d \leq \nabla_\beta^2 \ell \leq C\mathbf{I}_d, \quad \ell \geq 0.$$

Then the mirror descent reads

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\delta F(\mathbb{Q}_{k+1}^h)] + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\},$$

characterized by the **fixed point** of the mean-field BSDE:

$$\begin{aligned} dY_t &= -G(Z_t)dt + Z_t dW_t, \quad Y_T = \delta F(\mathbb{Q}_{k+1}^h), \\ G(z) &= \inf_{\beta} \left\{ \beta z + \ell(\beta) + \frac{1}{h} (\ell(\beta) - \ell(\beta_k^h) - \nabla \ell(\beta_k^h)(\beta - \beta_k^h)) \right\}. \end{aligned}$$

Normal Policy Gradient Descent for MFC

Following from the mean-field BSDE, the normal PGD is naturally defined as

$$\frac{d}{d\tau} \nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau))$$

where Z^τ is a part of the **fixed point** solution to the linear BSDE:

$$dY_t = -(\beta_t^\tau Z_t + \ell(\beta_t^\tau))dt + Z_t dW_t, \quad Y_T = \delta F(\mathbb{Q}^\tau).$$

Normal Policy Gradient Descent for MFC

Following from the mean-field BSDE, the normal PGD is naturally defined as

$$\frac{d}{d\tau} \nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau))$$

where Z^τ is a part of the **fixed point** solution to the linear BSDE:

$$dY_t = -(\beta_t^\tau Z_t + \ell(\beta_t^\tau))dt + Z_t dW_t, \quad Y_T = \delta F(\mathbb{Q}^\tau).$$

On the other hand, the mirror descent

$$\mathbb{Q}_{k+1}^h := \arg \min_{\mathbb{Q}} \left\{ H_\ell(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}}[\delta F(\mathbb{Q}_{k+1}^h)] + \frac{1}{h} H_\ell(\mathbb{Q} | \mathbb{Q}_k^h) \right\}$$

Normal Policy Gradient Descent for MFC

Following from the mean-field BSDE, the normal PGD is naturally defined as

$$\frac{d}{d\tau} \nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau))$$

where Z^τ is a part of the **fixed point** solution to the linear BSDE:

$$dY_t = -(\beta_t^\tau Z_t + \ell(\beta_t^\tau))dt + Z_t dW_t, \quad Y_T = \delta F(\mathbb{Q}^\tau).$$

On the other hand, the mirror descent gives the dual coordinate flow

$$\delta H_\ell(\mathbb{Q}_{k+1}^h) + \delta F(\mathbb{Q}_{l+1}^h) + \frac{1}{h} (\delta H_\ell(\mathbb{Q}_{k+1}^h) - \delta H_\ell(\mathbb{Q}_k^h)) = 0.$$

Normal Policy Gradient Descent for MFC

Following from the mean-field BSDE, the normal PGD is naturally defined as

$$\frac{d}{d\tau} \nabla \ell(\beta^\tau) = -(Z^\tau + \nabla \ell(\beta^\tau))$$

where Z^τ is a part of the **fixed point** solution to the linear BSDE:

$$dY_t = -(\beta_t^\tau Z_t + \ell(\beta_t^\tau))dt + Z_t dW_t, \quad Y_T = \delta F(\mathbb{Q}^\tau).$$

On the other hand, the mirror descent gives the dual coordinate flow

$$\delta H_\ell(\mathbb{Q}_{k+1}^h) + \delta F(\mathbb{Q}_{k+1}^h) + \frac{1}{h} (\delta H_\ell(\mathbb{Q}_{k+1}^h) - \delta H_\ell(\mathbb{Q}_k^h)) = 0.$$

As before, we can prove that the continuous-time limit of the **Mirror flow**, the **dual coordinate flow** and the **normal PGD** all coincide.

Table of Contents

- 4 PGD for Mean-Field Stochastic Control
 - Extension to MFC
 - Long-time behavior

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Then,

$$\begin{aligned} \frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) &= - \int \rho_\tau \dot{\mathbb{Q}}_\tau - \int \dot{\rho}_\tau (\mathbb{Q}_* - \mathbb{Q}_\tau) + \int \rho_\tau \dot{\mathbb{Q}}_\tau \\ &= \int \delta V(\mathbb{Q}_\tau)(\mathbb{Q}_* - \mathbb{Q}_\tau) \\ &= \int (\delta V(\mathbb{Q}_\tau) - \delta V(\mathbb{Q}_*))(\mathbb{Q}_* - \mathbb{Q}_\tau) \\ &\leq \int (\delta H_\ell(\mathbb{Q}_\tau) - \delta H_\ell(\mathbb{Q}_*))(\mathbb{Q}_* - \mathbb{Q}_\tau) \\ &= -(H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) + H_\ell(\mathbb{Q}_\tau|\mathbb{Q}_*)) \end{aligned}$$

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Then,

$$\frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) = \int \delta V(\mathbb{Q}_\tau)(\mathbb{Q}_* - \mathbb{Q}_\tau) \leq -H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau)$$

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Then,

$$\frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) = \int \delta V(\mathbb{Q}_\tau)(\mathbb{Q}_* - \mathbb{Q}_\tau) \leq -H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau)$$

To obtain convergence of the value, we first note that

$$V(\mathbb{Q}') - V(\mathbb{Q}) \geq \delta V(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}) + H_\ell(\mathbb{Q}'|\mathbb{Q}).$$

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Then,

$$\frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) = \int \delta V(\mathbb{Q}_\tau)(\mathbb{Q}_* - \mathbb{Q}_\tau) \leq -H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau)$$

To obtain convergence of the value, we first note that

$$V(\mathbb{Q}') - V(\mathbb{Q}) \geq \delta V(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}) + H_\ell(\mathbb{Q}'|\mathbb{Q}).$$

Therefore,

$$\frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) \leq -H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) + V(\mathbb{Q}^*) - V(\mathbb{Q}_\tau)$$

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Then,

$$\frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) = \int \delta V(\mathbb{Q}_\tau)(\mathbb{Q}_* - \mathbb{Q}_\tau) \leq -H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau)$$

To obtain convergence of the value, we first note that

$$V(\mathbb{Q}') - V(\mathbb{Q}) \geq \delta V(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}) + H_\ell(\mathbb{Q}'|\mathbb{Q}).$$

Therefore,

$$\frac{d}{d\tau} e^\tau H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) \leq e^\tau (V(\mathbb{Q}_*) - V(\mathbb{Q}_\tau))$$

Magic of Mirror flow: Energy dissipation

Denote by $\rho_\tau := \delta H_\ell(\mathbb{Q}_\tau)$. The continuous-time dual coordinate flow reads: $\dot{\rho}_\tau = -(\rho_\tau + \delta F(\mathbb{Q}_\tau)) = -\delta V(\mathbb{Q}_\tau)$, where the value of the control problem $V := H_\ell + F$. Recall that

$$H_\ell(\mathbb{Q}'|\mathbb{Q}) = H_\ell(\mathbb{Q}') - H_\ell(\mathbb{Q}) - \int \delta H_\ell(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}).$$

Then,

$$\frac{d}{d\tau} H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau) = \int \delta V(\mathbb{Q}_\tau)(\mathbb{Q}_* - \mathbb{Q}_\tau) \leq -H_\ell(\mathbb{Q}_*|\mathbb{Q}_\tau)$$

To obtain convergence of the value, we first note that

$$V(\mathbb{Q}') - V(\mathbb{Q}) \geq \delta V(\mathbb{Q})(\mathbb{Q}' - \mathbb{Q}) + H_\ell(\mathbb{Q}'|\mathbb{Q}).$$

Therefore,

$$(e^\tau - 1)(V(\mathbb{Q}_\tau) - V(\mathbb{Q}_*)) \leq \int_0^\tau e^s (V(\mathbb{Q}_s) - V(\mathbb{Q}_*)) ds \leq H_\ell(\mathbb{Q}_*|\mathbb{Q}_0).$$

Particle system of MFC

Recall the MFC:

$$\inf_{\beta} \mathbb{E}^{\beta} \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt \right] + F(\mathbb{Q}^{\beta}).$$

Now introduce the N-particel control problem:

$$\inf_{\beta} \mathbb{E}^{\beta} \left[\int_0^T \sum_{i=1}^N \ell(X_{t\wedge\cdot}^i, \beta_t^i) dt \right] + \mathbb{E}^{\beta} [NF(\mu)], \quad \mu := \frac{1}{N} \sum_{i=1}^N \delta_{X_{t\wedge\cdot}^i}.$$

Particle system of MFC

Recall the MFC:

$$\inf_{\beta} \mathbb{E}^{\beta} \left[\int_0^T \ell(X_{t\wedge\cdot}, \beta_t) dt \right] + F(\mathbb{Q}^{\beta}).$$

Now introduce the N -particel control problem:

$$\inf_{\beta} \mathbb{E}^{\beta} \left[\int_0^T \sum_{i=1}^N \ell(X_{t\wedge\cdot}^i, \beta_t^i) dt \right] + \mathbb{E}^{\beta} [NF(\mu)], \quad \mu := \frac{1}{N} \sum_{i=1}^N \delta_{X_{t\wedge\cdot}^i}.$$

For the N -particle control, the **dual coordinate flow** reads:

$$\frac{d}{d\tau} \delta H_{\ell}^N(\mathbb{Q}_{\tau}^N) = -\delta H_{\ell}^N(\mathbb{Q}_{\tau}^N) - NF(\mu).$$

Long time behavior

We study the defective energy dissipation:

$$\frac{d}{d\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N).$$

Long time behavior

We study the defective energy dissipation:

$$\begin{aligned}\frac{d}{d\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) &= - \int \frac{d}{d\tau} \delta H^N(\mathbb{Q}_\tau^N) (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_\tau^N) \\ &= \int \left(\delta H_\ell^N(\mathbb{Q}_\tau^N) + NF(\mu) \right) (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_\tau^N).\end{aligned}$$

Long time behavior

We study the defective energy dissipation:

$$\frac{d}{d\tau} H_{\ell}^N(Q_*^{\otimes N} | Q_{\tau}^N) = \int \left(\delta H_{\ell}^N(Q_{\tau}^N) + NF(\mu) - \delta H_{\ell}^{\otimes N}(Q_*) - \delta F(Q_*) \right) (Q_*^{\otimes N} - Q_{\tau}^N).$$

Long time behavior

We study the defective energy dissipation:

$$\frac{d}{d\tau} H_{\ell}^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_{\tau}^N) = \int \left(\delta H_{\ell}^N(\mathbb{Q}_{\tau}^N) - \delta H_{\ell}^{\otimes N}(\mathbb{Q}_*) + NF(\mu) - \delta F^{\otimes N}(\mathbb{Q}_*) \right) (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_{\tau}^N)$$

.

Long time behavior

We study the defective energy dissipation:

$$\begin{aligned}
 \frac{d}{d\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) &= \int \left(\delta H_\ell^N(\mathbb{Q}_\tau^N) - \delta H_\ell^{\otimes N}(\mathbb{Q}_*) + NF(\mu) - \delta F^{\otimes N}(\mathbb{Q}_*) \right) \\
 &\quad (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_\tau^N) \\
 &= -H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) - H_\ell^N(\mathbb{Q}_\tau^N | \mathbb{Q}_*^{\otimes N}) \\
 &\quad + \int \left(NF(\mu) - \delta F^{\otimes N}(\mathbb{Q}_*) \right) (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_\tau^N).
 \end{aligned}$$

Long time behavior

We study the defective energy dissipation:

$$\begin{aligned} \frac{d}{d\tau} H_{\ell}^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_{\tau}^N) &= -H_{\ell}^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_{\tau}^N) - H_{\ell}^N(\mathbb{Q}_{\tau}^N | \mathbb{Q}_*^{\otimes N}) \\ &\quad + \int \left(NF(\mu) - \delta F^{\otimes N}(\mathbb{Q}_*) \right) (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_{\tau}^N) \end{aligned}$$

.

Long time behavior

We study the defective energy dissipation:

$$\begin{aligned}
 \frac{d}{d\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) &= -H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) - H_\ell^N(\mathbb{Q}_\tau^N | \mathbb{Q}_*^{\otimes N}) \\
 &\quad + \int \left(NF(\mu) - \delta F^{\otimes N}(\mathbb{Q}_*) \right) (\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_\tau^N) \\
 &\leq -H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) - H_\ell^N(\mathbb{Q}_\tau^N | \mathbb{Q}_*^{\otimes N}) \\
 &\quad + \int NF(\mu)(\mathbb{Q}_*^{\otimes N} - \mathbb{Q}_\tau^N) + \int NF(\mu)\mathbb{Q}_\tau^N - NF(\mathbb{Q}_*)
 \end{aligned}$$

Long time behavior

We study the defective energy dissipation:

$$\begin{aligned} \frac{d}{d\tau} H_\ell^N(Q_*^{\otimes N} | Q_\tau^N) &\leq -H_\ell^N(Q_*^{\otimes N} | Q_\tau^N) - H_\ell^N(Q_\tau^N | Q_*^{\otimes N}) \\ &\quad + \int NF(\mu) Q_*^{\otimes N} - NF(Q_*). \end{aligned}$$

Long time behavior

We study the defective energy dissipation:

$$\frac{d}{d\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) \leq -H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) - H_\ell^N(\mathbb{Q}_\tau^N | \mathbb{Q}_*^{\otimes N}) + O(1).$$

Long time behavior

We study the defective energy dissipation:

$$\frac{d}{d\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) \leq -H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) - H_\ell^N(\mathbb{Q}_\tau^N | \mathbb{Q}_*^{\otimes N}) + O(1).$$

Therefore, we obtain the **long time error** of the particle system:

$$H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_\tau^N) \leq e^{-\tau} H_\ell^N(\mathbb{Q}_*^{\otimes N} | \mathbb{Q}_0^N) + O(1).$$

Thank you for your attention!