Path-dependent controlled mean-field FBSDEs. The associated stochastic maximum principle

Juan Li

Shandong University, Weihai & Qingdao, China.

Email: juanli@sdu.edu.cn

Based on a joint work with

Rainer Buckdahn (UBO, France) **Junsong Li** (SDU, Weihai, China)

Chuanzhi Xing (SDU, Qingdao, China)

The 19th International Workshop on Markov Processes and Related Topics (2024/07/23-07/26), Fujian Normal University, Fuzhou, 2024/07/23.

Outline

1. Objective of the talk

Let T be a fixed time horizon, b, σ measurable mappings defined over appropriate spaces. We are interested in the following nonlinear diffusions

Mean Field (McKean-Vlasov) SDE :

$$X_t = X_0 + \int_0^t b(s, X_s, P_{X_s})ds + \int_0^t \sigma(s, X_s, P_{X_s})dB_s, \ t \in [0,T], \quad (1.1)$$

where P is a probability measure with respect to which B is a B.M., and X_0 obeys a given probability law $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$.

Remark: 1. P_{X_s} is the law of X_s w.r.t. P.

2. $\mathcal{P}(\mathbb{R}^d)$: the space of the probability measures over $\mathbb{R}^d$.

3. $\mathcal{P}_p(\mathbb{R}^d) := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^p \mu(dx) < +\infty\}, \ p \geq 1.$

1. Brief state of art — McKean-Vlasov SDEs

The simplest: mean field approximation

Consider a large system of interacting diffusion:

$$dX_t^{N,i} = \sigma(X_t^{N,i}, \frac{1}{N}\sum_{j=1}^{N}\delta_{X_t^{N,j}})dW_t^i, \quad X_0^{N,i} = x_0, \ x_0 \in \mathbb{R}^d, \ 1 \le i \le N,$$

where W^i, $i \geq 1$, are independent B.M., and $\delta_x \in \mathcal{P}(\mathbb{R}^d)$ is the Dirac measure with mass at x.

$\underrightarrow{N \to \infty}$ $\quad dX_t^i = \sigma(X_t^i, P_{X_t^i})dW_t^i, \ X_0^i = x_0, \ x_0 \in \mathbb{R}^d.$

- **Mean Field SDEs**: Mean Field SDEs have been intensively studied for a longer time as limit equ. for systems with a large number of particles (propagation of chaos)(Sznitman (1984, 1991), Kotelenez (1995), Overbeck (1995), Méléard (1996), Talay & Vaillant (2003),......)

1. Brief state of art — Mean Field Games

• **Mean Field Games**: Mean-Field Games and related topics, since 2006-2007 by J.M. Lasry and P.L. Lions, Huang-Caines-Malhamé (Nash certainty equivalence principle) (2006);

....

Mean field game system:

$$\begin{cases} \textbf{i)} & -\partial_t u - v\triangle u + H(x, Du, m) = 0 \ \text{ in } (0,T)\times\mathbb{R}^d \text{ Hamilton-Jacobi-Bellman equation} \\ \textbf{ii)} & \partial_t m - v\triangle m - \text{div}(H_p(x, Du, m)m) = 0 \ \text{ in } (0,T)\times\mathbb{R}^d \text{ Kolmogorov-Fokker-Planck equation} \\ \textbf{iii)} & m(0) = m_0,\ u(x,T) = G(x, m(T)) \quad \text{in } \mathbb{R}^d \end{cases}$$

Master equation:

$$\begin{cases} -\partial_t U(t,x,m) - v\triangle_x U(t,x,m) + H(x, D_x U(t,x,m), m) - v\int_{\mathbb{R}^d} \text{div}_y D_m U(t,x,m,y)m(dy) \\ \qquad + \int_{\mathbb{R}^d} D_m U(t,x,m,y)\cdot D_p H(y, D_x U(t,y,m), m)m(dy) = 0 \quad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2 \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2 \end{cases}$$

1. Brief state of the art — Mean Field BSDEs

- **Mean Field BSDEs (MFBSDEs)**:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, P_{(Y_s,Z_s)})ds - \int_t^T Z_s dW_s,\ t \in [0,T], \quad (1.2)$$

where $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}^2) \to \mathbb{R}$ is Lipschitz, $\xi \in L^2(\mathcal{F}_T)$.

Existence and uniqueness of the solution: (Y, Z).

References:

- Discussion of the MFBSDE as the limit over a sequence of "nearly classical" BSDEs: *Buckdahn, Djehiche, L., Peng. 2009.*
- Discussion of general properties of MFBSDEs; their interpretation as generalized Feynman-Kac formula for associated nonlocal PDEs: *Buckdahn, L., Peng. 2009.*
- Classical solution of non-local PDE related with the general mean-field SDE: *Buckdahn, L., Peng, Rainer (2017 (2014, Arxiv))*

1. Brief state of the art — Controlled MFFBSDEs

- For Pontryagin's maximum principle: L. (2012);

  + with partial observations: Buckdahn, L., Ma (2017);

$\rightsquigarrow$ Acciaio, Backhoff-Veraguas, Carmona (2019):

Controlled mean-field stochastic system:

$dX_t^v = b(t, P_{(X_t^v, v_t)}, X_t^v, v_t)dt + \sigma(t, P_{(X_t^v, v_t)}, X_t^v, v_t)dW_t, \ t \in [0, T]...$

- For Peng's maximum principle: Buckdahn, Djehiche, L. (2011);

$\rightsquigarrow$Buckdahn, L., Ma (2016): Controlled mean-field stochastic system:

$dX_t^v = b(t, P_{X_t^v}, X_t^v, v_t)dt + \sigma(t, P_{X_t^v}, X_t^v, v_t)dW_t, \ t \in [0, T]...$

$\rightsquigarrow$Buckdahn, Chen, L. (2021): Controlled mean-field stochastic system:

$dX_t^v = b(t, P_{(X_t^v, v_t)}, X_t^v, v_t)dt + \sigma(t, P_{(X_t^v, v_t)}, X_t^v, v_t)dW_t, \ t \in [0, T]...$

+ with partial observations: L., Liang, Mi (2023 (2021, Arxiv))

- For Zero-sum stochastic differential games:

  L., Min (2016)

....

1. Objective of the talk

Study of Pontryagin's stochastic maximum principle
for a path-dependent mean-field forward and backward stochastic system.

The novelties in our work:

- The MFFBSDE is fully coupled through the law of the paths of (X, Y) in the coefficients of both the forward and the backward equation. Existence is proved just under the 2-Wasserstein continuity of the coefficients w.r.t. the law of (X, Y).

- The coefficients in both the forward as well as the backward SDEs depend not only on the controlled solution processes (X_t, Y_t, Z_t) at the current time t, but also on the law of the paths (X, Y, u) of the solution process and the control. The cost functional too depends on the law of the paths of (X, Y, u).

- The Hamiltonian as well as the SMP we obtain are quite novel, far beyond the classical ones.

2. Preliminaries

- $(\Omega, \mathcal{F}, P)$ – complete probability space endowed with a d-dim. B.M. B.
- $\mathcal{F}_0 \subset \mathcal{F}$ – a sub-σ-field independent of B and "rich enough", i.e.,
$$\mathcal{P}_2(\mathbb{R}^k) = \{P_\xi, \xi \in L^2(\mathcal{F}_0; \mathbb{R}^k)\},\ k \geq 1.$$
- $T > 0$ – fixed time horizon, $\mathcal{C}_T^k := \mathcal{C}([0,T]; \mathbb{R}^k)$.
- $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ – a copy of the probability space $(\Omega, \mathcal{F}, P)$. For any r.v. ϑ over $(\Omega,\ \mathcal{F}, P)$ we denote by $\widetilde{\vartheta}$ a copy over $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$: $\widetilde{P}_{\widetilde{\vartheta}} = P_\vartheta$.
- $\mathcal{P}(\mathcal{C}_T^k)$ – the space of all probability measures on $(\mathcal{C}_T^k, \mathcal{B}(\mathcal{C}_T^k))$;
- For $p \geq 1$, $\mathcal{P}_p(\mathcal{C}_T^k)$ $(\subseteq \mathcal{P}(\mathcal{C}_T^k))$ the space of all probabilities with finite p-th moment; the p-Wasserstein metric on $\mathcal{P}_p(\mathcal{C}_T^k)$: For all $\mu, \nu \in \mathcal{P}_p(\mathcal{C}_T^k)$,

$$W_p(\mu,\nu) := \inf\{(\int_{\mathcal{C}_T^{2k}} |x-y|^p_{\mathcal{C}_T} \pi(dx,dy))^{\frac{1}{p}} : \pi \in \mathcal{P}_p(\mathcal{C}_T^{2k}) \text{ with marginals } \mu \text{ and } \nu\}.$$

For $p = 1$, we have the Kantorovich-Rubinstein Theorem,

$$W_1(\mu,\nu) = \sup\{|\int_{\mathcal{C}_T^k} h d\mu - \int_{\mathcal{C}_T^k} h d\nu|,\ h \in Lip_1(\mathcal{C}_T^k),\ h(0) = 0.\}.$$

We consider the following coupled mean-field FBSDE

$$\left\{\begin{array}{l} dX_t=\sigma(t,X_t,P_{(X,Y)})dB_t,\ t\in[0,T],\\ dY_t=-f(t,X_t,Y_t,Z_t,P_{(X,Y)})dt+Z_tdB_t,\ t\in[0,T],\\ X_0=x\in\mathbb{R}^d,\ Y_T=\Phi(X_T,P_{(X,Y)}),\end{array}\right. \tag{3.1}$$

where

$$\sigma:[0,T]\times\mathbb{R}\times\mathcal{P}_2(\mathcal{C}_T^2)\to\mathbb{R},\ f:[0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathcal{C}_T^2)\to\mathbb{R},$$
$$\Phi:\mathbb{R}\times\mathcal{P}_2(\mathcal{C}_T^2)\to\mathbb{R},$$

satisfy:

Assumptions

(**H1**) $\sigma,\ \Phi$ and f are measurable, and continuous over $\mathcal{P}_1(\mathcal{C}_T^2)$, with continuity modulus $\rho:\mathbb{R}_+\to\mathbb{R}_+$ with $\rho(0+)=0$, uniformly w.r.t. the other variables.

(**H2**) $\sigma,\ \Phi$ and f are bounded and have bounded derivatives w.r.t. x and to (x,y,z), resp.

Theorem 3.1

We assume (H1) and (H2) hold true. Then mean-field FBSDE (3.1) has an adapted solution (X, Y, Z).

Sketch of Proof. The key for the proof is an application of Schauder's fixed point theorem stating that if V is a Hausdorff topological vector space, $\mathcal{K} \subset V$ is a nonempty convex closed subset, and $\mathcal{T} : \mathcal{K} \to \mathcal{K}$ is a continuous mapping such that $\mathcal{T}(\mathcal{K}) \subset \mathcal{K}$ is contained in a compact subset of $\mathcal{K}$, then there exists $\mu \in \mathcal{K}$ such that $\mathcal{T}(\mu) = \mu$.

Let $\mu \in \mathcal{P}_1(\mathcal{C}_T^2)$, and we consider

$$\begin{cases} X_t^\mu = x + \displaystyle\int_0^t \sigma(s, X_s^\mu, \mu) dB_s, \\ Y_t^\mu = \Phi(X_T^\mu, \mu) + \displaystyle\int_t^T f(s, X_s^\mu, Y_s^\mu, Z_s^\mu, \mu) ds - \int_t^T Z_s^\mu dB_s, \quad t \in [0, T]. \end{cases} \tag{3.2}$$

Then, under the assumptions (H1) and (H2), (3.2) has a unique solution $(X^\mu, Y^\mu, Z^\mu) \in S_{\mathbb{F}}^2 \times S_{\mathbb{F}}^2 \times M_{\mathbb{F}}^2$. Put $\theta_s^\mu := (X_s^\mu, Y_s^\mu, Z_s^\mu)$, $s \in [0, T]$.

Using the Malliavin derivative we show,

- For $p\geq 1$, $E[|\sup_{t\in[r,T]} D_r[X_t^\mu]|^p|\mathcal{F}_r]\leq C_p,\ t\in[0,T],\ 0\leq r\leq\ T,$
- $|Y_t^\mu|\leq C,\ t\in[0,T], P$-a.s., $|Z_t^\mu|\leq C,\ \mathrm{drdP}$-a.e., for some $C\in\mathbb{R}$.

With $\mathscr{C}=(\mathscr{C}_t)$ as coordinate process on $\mathcal{C}_T^2$: $\mathscr{C}_t(\phi)=\phi_t, \phi\in\mathcal{C}_T^2, t\in[0,T]$, we consider

$$\mathcal{K}:=\Big\{\mu\in\mathcal{P}_1(\mathcal{C}_T^2)\Big|\ \int_{\mathcal{C}_T^2}\sup_{t\in[0,T]}|\mathscr{C}_t|^4d\mu\leq C;\ \int_{\mathcal{C}_T^2}|\mathscr{C}_t-\mathscr{C}_s|^4d\mu\leq C|t-s|^2,\ t,s\in[0,T]\Big\}.$$

Step 1. $\mathcal{K}$ is a convex and compact subset of $(\mathcal{P}_1(\mathcal{C}_T^2),W_1(\cdot,\cdot))$.

Step 2. $\mathcal{T}$ defined by $\mathcal{T}(\mu):=P_{(X^\mu,Y^\mu)},\ \mu\in\mathcal{P}_1(\mathcal{C}_T^2)$, maps $\mathcal{K}$ into $\mathcal{K}$.

Step 3. $\mathcal{T}:\mathcal{K}\to\mathcal{K}$ is continuous.

Step 4. Application of Schauder's fixed point theorem.

$\mathcal{T}(\mathcal{K})$ is compact in $\mathcal{K}$, since $\mathcal{K}$ is compact and $\mathcal{T}:\mathcal{K}\to\mathcal{K}$ is continuous. Embed $\mathcal{K}\subset\mathcal{P}_1(\mathcal{C}_T^2)$ into the separable linear Hausdorff space

$$\mathcal{M}_1(\mathcal{C}_T^2) := \Big\{\gamma \text{ signed measure over } (\mathcal{C}_\mathrm{T}^2,\mathcal{B}(\mathcal{C}_\mathrm{T}^2)) \;\Big|\; \int_{\mathcal{C}_\mathrm{T}^2} |\phi|_{\mathcal{C}_\mathrm{T}^2}\cdot|\gamma|(\mathrm{d}\phi) < +\infty\Big\},$$

with the norm: $\|\gamma\|_1 := \sup\{|\int_{\mathcal{C}_T^2} h d\gamma| \mid h\in Lip_1(\mathcal{C}_T^2),\ h(0)=0\}$
(Note: $\|\mu-\nu\|_1 = W_1(\mu,\nu)$, $\mu,\nu\in\mathcal{P}_1(\mathcal{C}_T^2)$).

This allows to apply Schauder's fixed point theorem: There exists $\mu\in\mathcal{K}(\subset\ \mathcal{P}_1(\mathcal{C}_T^2))$ such that $\mu=\mathcal{T}(\mu)=P_{(X^\mu,Y^\mu)}$, where (X^μ,Y^μ,Z^μ) is the solution of FBSDE (3.2). □

Remark 3.1.

Theorem 3.1 uses the $W_1(\cdot,\cdot)$-continuity of the coefficients w.r.t. the measure, with continuity modulus ρ. However, the assumption of $W_1(\cdot,\cdot)$-continuity can be replaced by that of $W_2(\cdot,\cdot)$-continuity. Indeed, assuming w.l.o.g that $\rho:\mathbb{R}_+\to\mathbb{R}_+$ is increasing, we have

$$\rho(W_1(\mu,\mu'))\le\rho(W_2(\mu,\mu'))\le\rho'(W_1(\mu,\mu')),\ \ \mu,\ \ \mu'\in\mathcal{K}, \tag{3.3}$$

for $\rho'(r):=\rho((2C^{\frac{1}{4}})^{\frac{3}{4}}r^{\frac{1}{4}}), r\geqslant 0$, with C from the definition of $\mathcal{K}$.

Let us discuss the uniqueness. For this end we consider the following mean-field FBSDE,

$$\begin{cases} dX_t = \sigma(t, X_t, P_{X_{.\wedge t}})dB_t, \ t \in [0, T], \\ dY_t = -f(t, X_t, Y_t, Z_t, P_{(X, Y_{.\vee t})})dt + Z_t dB_t, \ t \in [0, T], \\ X_0 = x \in \mathbb{R}, \ Y_T = \Phi(X_T, P_{X_{.\wedge T}}), \end{cases} \tag{3.4}$$

where $X_{.\wedge t} = (X_{s\wedge t})_{s\in[0,T]}, Y_{.\vee t} = (Y_{s\vee t})_{s\in[0,T]}$.

It is a special case of MFFBSDE (3.1), and so the existence of a solution is proved.

Let us replace (H1) by the following assumption:

Assumption

(H3) There is a constant $C > 0$ s.t., for all $t \in [0, T]$, (x, y, z), (x', y', z') $\in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d, \mu,\ \mu' \in \mathcal{P}_1(\mathcal{C}_T^2),\ \nu,\ \nu' \in \mathcal{P}_1(\mathcal{C}_T)$,

$$|\sigma(t, x, \nu) - \sigma(t, x', \nu')| + |\Phi(x, \nu) - \Phi(x', \nu')| \leq C(|x - x'| + W_2(\nu, \nu')),$$
$$|f(t, x, y, z, \mu) - f(t, x', y', z', \mu')| \leq C(|x - x'| + |y - y'| + |z - z'| + W_2(\mu, \mu')).$$

Theorem 3.2.

Under assumptions (H2) and (H3), equation (3.4) has a unique solution $(X, Y, Z) \in S_{\mathbb{F}}^2 \times S_{\mathbb{F}}^2 \times M_{\mathbb{F}}^2$.

Examples:

1) $\sigma(t, x, P_{(X,Y)}) := \widetilde{\sigma}(t, x, \int_0^T E[h(X_s, Y_s)]ds)$, where h is a deterministic function;

2) $\sigma(t, x, P_{(X,Y)}) := \widetilde{\sigma}(t, x, E[g(X_{\phi(t)}, Y_{\psi(t)})])$, where $\phi,\ \psi : [0, T] \to [0, T]$ are measureable functions.

......

σ defined in 1) and in 2) satisfies (H1) and (H2).

- $(\mathcal{H}, |\cdot|_{\mathcal{H}})$ - a real separable Banach space ;
- $\mathcal{H}' = \{\ell : \mathcal{H} \to \mathbb{R} \; | \ell$ continuous linear functional $\}$ the dual space of $\mathcal{H}$;
- $\langle l, x \rangle_{\mathcal{H}' \times \mathcal{H}} := l(x)$ the duality product on $\mathcal{H}' \times \mathcal{H}$.
- We are particularly interested in:

  $+ \; \mathcal{H} = \; \mathcal{C}([0,T];\mathbb{R}^n) \times L^2([0,T];\mathbb{R}^d)$,

  $+ \; \mathcal{H}' = \; BV([0,T];\mathbb{R}^n) \times L^2([0,T];\mathbb{R}^d)$;

  $BV([0,T];\mathbb{R}^n)$-space of all $\mathbb{R}^n$-valued bounded variational càdlàg

functions over $[0,T]$. So we have:

$$\langle (h,y), (\varphi, x) \rangle_{\mathcal{H}' \times \mathcal{H}} = \int_{[0,T]} \varphi(t) h(dt) + (y, x)_{L^2([0,T];\mathbb{R}^d)}, \tag{4.1}$$

where $(\varphi, x) \in \mathcal{C}_T^n \times L^2([0,T];\mathbb{R}^d)$, $(h,y) \in BV_T^n \times L^2([0,T];\mathbb{R}^d)$.

Let us come back to a general real separable Banach space $(\mathcal{H}, |\cdot|_{\mathcal{H}})$. Recall that

$$\mathcal{P}_2(\mathcal{H}) = \{m \text{ probability over } (\mathcal{H},\ \mathcal{B}(\mathcal{H})) : \int_{\mathcal{H}} |x|^2_{\mathcal{H}} m(dx) < \infty\}.$$

The notion of differentiability of a function $f : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ (For $\mathcal{H} = \mathbb{R}^d$, see the book by Carmona and Delarue):

Definition 3.2.

We say that $u : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ has the linear functional derivative $\mathcal{D}_m u : \mathcal{P}_2(\mathcal{H}) \times \mathcal{H} \to \mathbb{R}$, if $\mathcal{D}_m u$ is a continuous function over $\mathcal{P}_2(\mathcal{H}) \times \mathcal{H}$ with at most quadratic growth such that, for all $m, m' \in \mathcal{P}_2(\mathcal{H})$,

$$u(m') - u(m) = \int_0^1 \int_{\mathcal{H}} \mathcal{D}_m u(tm' + (1-t)m, x)(m'(dx) - m(dx))dt. \tag{4.2}$$

Let us suppose that for $u : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ the derivative $\mathcal{D}_m u : \mathcal{P}_2(\mathcal{H}) \times \mathcal{H} \to \mathbb{R}$ exists, is continuous and of at most quadratic growth, and that, for all $m \in \mathcal{P}_2(\mathcal{H}), \mathcal{D}_m u(m, \cdot) : \mathcal{H} \to \mathbb{R}$ is differentiable, i.e., there exists $\partial_x(\mathcal{D}_m u)(m, \cdot) : \mathcal{H} \to \mathcal{H}'$ such that, for all $x \in \mathcal{H}$, as $y \in \mathcal{H}$ tends to x

$$\mathcal{D}_m u(m, y) - \mathcal{D}_m u(m, x) = \langle \partial_x(\mathcal{D}_m u)(m, x), y - x \rangle_{\mathcal{H}' \times \mathcal{H}} + o(|y - x|_{\mathcal{H}}). \tag{4.3}$$

Let us make the following assumptions for $\mathcal{D}_m u : \mathcal{P}_2(\mathcal{H}) \times \mathcal{H} \to \mathbb{R}$:

Assumption

(**H4**) The derivative $\partial_x(\mathcal{D}_m u) : \mathcal{P}_2(\mathcal{H}) \times \mathcal{H} \to \mathcal{H}'$ exists, is continuous and of at most linear growth.

However, to simplify our arguments, we suppose also (H4)':

Assumption

(H4') (i) $\partial_x(\mathcal{D}_m u) : \mathcal{P}_2(\mathcal{H}) \times \mathcal{H} \to \mathcal{H}'$ is bounded, and

(ii) There exists a continuity modulus $\rho_u : \mathbb{R}_+ \to \mathbb{R}_+$ continuous and increasing, with $\rho_u(0) = 0$ and $\rho_u^2(\cdot)$ concave s.t.

$$|\partial_x(\mathcal{D}_m u)(m', x') - \partial_x(\mathcal{D}_m u)(m, x)|_{\mathcal{H}'} \leq \rho_u(W_2(m, m') + |x - x'|_{\mathcal{H}}).$$

for all $(m', x'),\ (m, x) \in \mathcal{P}_2(\mathcal{H}) \times \mathcal{H}$.

Proposition 4.1.

Under our assumption (H4') on u and $\mathcal{D}_m u$ we have the following first order Taylor expansion for $u : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ at $m \in \mathcal{P}_2(\mathcal{H})$ holds true:

$$u(m') = u(m) + \int_{\mathcal{H}} \mathcal{D}_m u(m,x)(m'(dx) - m(dx)) + o(W_2(m,m')), \quad \text{as } W_2(m,m') \to 0 \ (m' \in \mathcal{P}_2(\mathcal{H})). \tag{4.4}$$

The proof of this proposition is a slight extension of the corresponding result in Carmona and Delarue (2018). Moreover, also under our above assumptions the following result concerning (Lion's) L-derivative of $u : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ extends easily from $\mathcal{H} = \mathbb{R}^n$ to the general separable Banach space $\mathcal{H}$.

Proposition 4.2.

Given any $\xi \in L^2(\Omega, \mathcal{F}, P; \mathcal{H})$, we have

$$u(P_{\xi+\eta}) = u(P_\xi) + E[\langle \partial_x(\mathcal{D}_m u)(P_\xi, \xi), \eta \rangle_{\mathcal{H}' \times \mathcal{H}}] + R(\xi, \eta), \eta \in L^2(\Omega, \mathcal{F}, P; \mathcal{H}), \tag{4.5}$$

where $R(\xi, \eta) = o((E[|\eta|^2_{\mathcal{H}}])^{\frac{1}{2}})$ as $E[|\eta|^2_{\mathcal{H}}] \to 0$.

Definition 4.2.

We denote the L-derivative by $(\partial_\mu u(P_\xi, x))$:

$$\partial_\mu u(m, x) = \partial_x(\mathcal{D}_m u)(m, x), \ (m, x) \in \mathcal{P}_2(\mathcal{H}) \times \mathcal{H}. \tag{4.6}$$

Moveover, when we speak about the differentiability of $u : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$, we mean the L-differentiability.

Here we are interested in the case $\mathcal{H} = \mathcal{C}_T^\ell \times L^2([0,T];\mathbb{R}^k)$. So, if $\sigma : \mathcal{P}_2(\mathcal{C}_T^\ell \times L^2([0,T];\mathbb{R}^k)) \to \mathbb{R}$ is differentiable,

$$\partial_\mu \sigma : \mathcal{P}_2(\mathcal{C}_T^\ell \times L^2([0,T];\mathbb{R}^k)) \times \big(\mathcal{C}_T^\ell \times L^2([0,T];\mathbb{R}^k)\big) \to BV_T^\ell \times L^2([0,T];\mathbb{R}^k),$$

and $(\partial_\mu \sigma)(m,\varphi) = ((\partial_\mu \sigma)_i(m,\varphi))_{1 \le i \le \ell+k}$.

Moreover, we will write $\langle \cdot,\cdot \rangle_\ell = \langle \cdot,\cdot \rangle_{\mathcal{H}' \times \mathcal{H}}$ for the duality product.

Definition 4.3.

For $l = (l_1, \cdots, l_\ell) \in BV_T^\ell$, $\widehat{v} \in L^2([0,T];\mathbb{R}^k)$, $(f(= (f_1, \cdots, f_\ell)), v) \in \mathcal{C}_T^\ell \times L^2([0,T];\mathbb{R}^k)$, the duality product $\langle \cdot,\cdot \rangle_\ell$ is given by

$$\langle (l,\widehat{v}), (f,v) \rangle_\ell := \sum_{i=1}^{\ell} \int_0^T f_i(t) l_i(dt) + \int_0^T \widehat{v}_s v_s ds.$$

- $\mathcal{U}_T := L^2([0,T];\mathbb{R}^k)$,
- $L^{\infty-}_{\mathbb{F}}(\Omega, L^2([0,T];\mathbb{R}^k))$ - space of all $\mathbb{F}$-adapted processes s.t.
  $E[(\int_0^T |v(t)|^2 dt)^{\frac{p}{2}}] < +\infty,\ p \geq 2.$
- U - convex subset of $\mathbb{R}^k$: our control state space;
- The space of admissible controls:

$$\mathcal{U}_{ad} = \{v \in L^{\infty-}(\Omega, L^2_{\mathbb{F}}([0,T];\mathbb{R}^k)) |\ v_t \in U,\ 0 \leq t \leq T\}.$$

The dynamics of the controlled stochastic system:

$$\begin{cases} dX_t^v = \sigma(t, X_t, P_{(X^v_{\cdot\wedge t}, v)})dB_t, \ t \in [0,T], \\ dY_t^v = -f(t, X_t^v, Y_t^v, Z_t^v, P_{(X^v, Y^v_{t\vee \cdot}, v)})dt + Z_t^v dB_t, \ t \in [0,T], \\ \ X_0^v = x \in \mathbb{R}, \ Y_T^v = \Phi(X_T^v, P_{(X^v, v)}), \end{cases} \quad (5.1)$$

where $v \in \mathcal{U}_{ad}$ and $f: \ [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathcal{C}_T^2 \times \mathcal{U}_T) \to \mathbb{R}, \ \sigma:$ $[0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathcal{C}_T \times \mathcal{U}_T) \to \mathbb{R}, \ \Phi: \ \mathbb{R} \times \mathcal{P}_2(\mathcal{C}_T \times \mathcal{U}_T) \to \mathbb{R}.$

The cost functional:

$$J(v) = E[\int_0^T L(t, X_t^v, Y_t^v, Z_t^v, P_{(X^v, Y^v, v)})dt + \varphi(X_T^v, P_{(X^v, Y^v, v)})],$$

where $L: [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathcal{C}_T^2 \times \mathcal{U}_T) \to \mathbb{R}, \ \ \varphi: \mathbb{R} \times \mathcal{P}_2(\mathcal{C}_T^2 \times$ $\mathcal{U}_T) \to \mathbb{R}.$

Control problem: A control $u \in \mathcal{U}_{ad}$ satisfying

$$J(u) = \inf_{u \in \mathcal{U}_{ad}} J(v)$$

is said to be optimal.

Objective: Suppose there is an optimal control u,

$$J(u) = \inf_{u \in \mathcal{U}_{ad}} J(v)$$

and characterize it with the help of Pontryagin's SMP.

We have the following standard assumptions:

Assumption

(**H5**) The functions $f,\ \sigma,\ L,\ \Phi,$ and φ are bounded and continuously differentiable to $(x,\ y,\ z,\ \mu)$ and the derivatives are bounded.

Recall that $(\partial_\mu\sigma)_1(t,x,\mu;(\phi,\widehat{v}))\in BV_T$. Identifying the elements of BV_T with measures on $([0,T];\mathcal{B}([0,T]))$, we suppose:

Assumption

(H6) (i) $(\partial_\mu\sigma)_1(t,x,\mu;(\phi,\widehat{v}))(dr):=(\partial_\mu\sigma)_1(t,x,\mu;(\phi,\widehat{v}))(r)dr$, where the function at the right-hand side $(\partial_\mu\sigma)_1:[0,T]\times\mathbb{R}\times\mathcal{P}_2(\mathcal{C}_T\times\mathcal{U}_T)\times(\mathcal{C}_T\times\mathcal{U}_T)\to\mathbb{R}$ is Borel measurable, bounded as well as Lipschitz in $(x,\mu,\phi,\widehat{v})$, uniformly w.r.t. $t,r\in[0,T]$.

In the same manner we suppose

$$\begin{aligned}
(\partial_\mu\Phi)_1(t,x,\mu;(\phi,\widehat{v}))(dr) &:=(\partial_\mu\Phi)_1(t,x,\mu;(\phi,\widehat{v}))(r)dr,\\
(\partial_\mu f)_i(t,x,y,z,\mu;(\phi_1,\phi_2,\widehat{v}))(dr) &:=(\partial_\mu f)_i(t,x,y,z,\mu;(\phi_1,\phi_2,\widehat{v}))(r)dr,\\
(\partial_\mu L)_i(t,x,y,z,\mu;(\phi_1,\phi_2,\widehat{v}))(dr) &:=(\partial_\mu L)_i(t,x,y,z,\mu;(\phi_1,\phi_2,\widehat{v}))(r)dr,\\
(\partial_\mu\varphi)_i(t,x,y,z,\mu;(\phi_1,\phi_2,\widehat{v}))(dr) &:=(\partial_\mu\varphi)_i(t,x,y,z,\mu;(\phi_1,\phi_2,\widehat{v}))(r)dr,
\end{aligned}$$

where for the functions at the right-hand side we assume:

Assumption (continued)

(H6) (ii) $(\partial_\mu f)_i, (\partial_\mu L)_i, (\partial_\mu \varphi)_i, i = 1, 2, (\partial_\mu \Phi)_1$, are Borel measurable, bounded, and Lipschitz in the other variables, uniformly w.r.t. the time parameters.

Under the above assumptions, the existence and the uniqueness of the solution of (5.1) can be shown by the arguments used for the proof of the Theorems 3.1 and 3.2.

Proposition 5.1

Suppose $(\mathrm{H5})$ and $(\mathrm{H6})$. Then the controlled MFFBSDE (5.1) has a unique solution.

Convex variational method.

Let u be an optimal control and $(X,Y,Z) := (X^u, Y^u, Z^u)$ be the corresponding optimal solution. Let v be such that $u+v \in \mathcal{U}_{ad}$. Since U is convex, then also for any $0 \le \rho \le 1$, $u^\rho = u + \rho v$ is in $\mathcal{U}_{ad}$.
Formal differentiation w.r.t. to ρ at $\rho = 0$ of $(X^\rho, Y^\rho) := (X^{u+\rho v}, Y^{u+\rho v})$ yields the variational equation:

$$\begin{cases} dX_t^1 = \{\sigma_x(t, X_t, P_{(X_{\cdot\wedge t}, u)})X_t^1 \\ \qquad\qquad + \widetilde{E}[\langle(\partial_\mu \sigma)(t, X_t, P_{(X_{\cdot\wedge t},u)}; (\widetilde{X}_{\cdot\wedge t}, \widetilde{u})),\ (\widetilde{X}^1_{\cdot\wedge t}, \widetilde{v})\rangle_1]\} dB_t, \\ dY_t^1 = -\{f_x(\theta_t)X_t^1 + f_y(\theta_t)Y_t^1 + f_z(\theta_t)Z_t^1 \\ \qquad\qquad + \widetilde{E}[\langle(\partial_\mu f)(\theta_t; (\widetilde{X}, \widetilde{Y}_{\cdot\vee t}, \widetilde{u})),\ (\widetilde{X}^1, \widetilde{Y}^1_{\cdot\vee t}, \widetilde{v})\rangle_2]\} dt + Z_t^1 dB_t, \\ X_0^1 = 0, \\ Y_T^1 = \Phi_x(X_T, P_{(X,u)})X_T^1 + \widetilde{E}[\langle(\partial_\mu \Phi)(X_T, P_{(X,u)}; (\widetilde{X}, \widetilde{u})),\ (\widetilde{X}^1, \widetilde{v})\rangle_1], \end{cases} \tag{5.2}$$

where $\theta_t := (t, X_t, Y_t, Z_t, P_{(X, Y_{\cdot\vee t}, u)})$.

Recall that, e.g.,

$$\langle (\partial_\mu f)(\theta_t; (\widetilde{X}, \widetilde{Y}_{\cdot\vee t}, \widetilde{u})), (\widetilde{X}^1, \widetilde{Y}^1_{\cdot\vee t}, \widetilde{v})\rangle_2$$
$$= \int_0^T (\partial_\mu f)(\theta_t; (\widetilde{X}, \widetilde{Y}_{\cdot\vee t}, \widetilde{u}))(r) \cdot (\widetilde{X}^1_r, \widetilde{Y}^1_{r\vee t}, \widetilde{v}_r)^T dr.$$

Lemma 5.1.

Let the assumptions (H5) and (H6) be satisfied. Then the above linear mean-field FBSDE (5.2) has a unique solution $(X^1, Y^1, Z^1) \in S^2_{\mathbb{F}} \times S^2_{\mathbb{F}} \times M^2_{\mathbb{F}}$.

Recall $(X^\rho,\ Y^\rho,\ Z^\rho)=(X^{u^\rho},\ Y^{u^\rho},\ Z^{u^\rho})$ for $u^\rho=u+\rho v$, and $(X,Y,Z)=(X^u,Y^u,Z^u)$. Then:

Lemma 5.2.

Assume (H5) and (H6). Then, for all $p\geq 2$, there is $C_p\in\mathbb{R}_+$ s.t.

$$E\Big[\sup_{t\in[0,T]}\Big(\big|\frac{1}{\rho}(X_t^\rho-X_t)\big|+\big|\frac{1}{\rho}(Y_t^\rho-Y_t)\big|\Big)^p\Big]$$
$$+E\Big[\Big(\int_0^T\Big(\frac{1}{\rho}|Z_s^\rho-Z_s|\Big)^2ds\Big)^{\frac{p}{2}}\Big]\leq C_pE\Big[\Big(\int_0^T|v_t|^2dt\Big)^{\frac{p}{2}}\Big],$$

for all $\rho\in(0,1)$, $v\in\mathcal{U}_{ad}$,.

Using Lemma 5.2 we prove:

Lemma 5.3.

We assume (H5) and (H6). Then,

$$\frac{1}{\rho}\big((X_t^\rho, Y_t^\rho, Z_t^\rho) - (X_t, Y_t, Z_t)\big) \to (X_t^1, Y_t^1, Z_t^1),\ \rho \searrow 0,$$

with convegence in $S_{\mathbb{F}}^2 \times S_{\mathbb{F}}^2 \times M_{\mathbb{F}}^2$.

Let us now study the variational inequality. As u is an optimal control,

$$\rho^{-1}[J(u(\cdot)+\rho v(\cdot))-J(u(\cdot))]\geq 0.$$

Thus thanks to the Lemmas 5.2 and 5.3 we have

Theorem 5.1.

We suppose (H5) and (H6) hold. Then, the following variational inequality holds true:

$$\begin{aligned}0\leq &E\Big[\int_0^T (L_x(\theta_t)X_t^1+L_y(\theta_t)Y_t^1+L_z(\theta_t)Z_t^1\\&+\widetilde{E}[\langle(\partial_\mu L)(\theta_t;(\widetilde{X},\widetilde{Y},\widetilde{u})),(\widetilde{X}^1,\widetilde{Y}^1,\widetilde{v})\rangle_2])dt\\&+\varphi_x(X_T,P_{X,Y,u})X_T^1\\&+\widetilde{E}[\langle(\partial_\mu\varphi)(X_T,P_{(X,Y,u)};(\widetilde{X},\widetilde{Y},\widetilde{u})),(\widetilde{X}^1,\widetilde{Y}^1,\widetilde{v})\rangle_2]\Big],\end{aligned} \tag{5.3}$$

where $\theta_t=(t,X_t,Y_t,Z_t,P_{(X,Y_{\cdot\vee t},v)})$.

In order to derive the maximum principle, we have to introduce the adjoint equation. For this, we make use of the following adapted processes:

$$\begin{aligned}
&\sigma_x(t) := \sigma_x(t, X_t, P_{(X_{\cdot\wedge t},u)});\\
&(\partial_\mu\sigma)_1^*(t)[k] := E\Big[\widetilde{E}\Big[\int_t^T \{(\partial_\mu\sigma)_1(r, \widetilde{X}_r, P_{(X_{\cdot\wedge r},u)}; (X_{\cdot\wedge r}, u))(t)\widetilde{k}_r\\
&\qquad\qquad + (\partial_\mu\sigma)_1(t, \widetilde{X}_t, P_{(X_{\cdot\wedge t},u)}; (X_{\cdot\wedge t}, u))(r)\widetilde{k}_t\}dr\Big]\Big|\mathcal{F}_t\Big];\\
&(\partial_\mu\sigma)_2^*(t)[k] := E\Big[\widetilde{E}\Big[\int_0^T (\partial_\mu\sigma)_2(r, \widetilde{X}_r, P_{(X_{\cdot\wedge r},u)}; (X_{\cdot\wedge r}, u))(t)\widetilde{k}_r dr\Big]\Big|\mathcal{F}_t\Big];
\end{aligned} \tag{5.4}$$

$$
\begin{aligned}
(\partial_\mu f)^*_j(t)[p] &:= E\Big[\widetilde{E}\Big[\int_0^T (\partial_\mu f)_j(\widetilde{\theta}_r;(X,Y_{\cdot\vee r},u))(t)\widetilde{p}(r)dr\Big]\Big|\mathcal{F}_t\Big],\ j=1,3;\\
(\partial_\mu f)^*_2(t)[p] &:= E\Big[\widetilde{E}\Big[\int_0^t \Big\{(\partial_\mu f)_2(\widetilde{\theta}_r;(X,Y_{\cdot\vee r},u))(t)\widetilde{p}(r)\\
&\qquad + (\partial_\mu f)_2(\widetilde{\theta}_t;(X,Y_{\cdot\vee t},u))(r)\widetilde{p}(t)\Big\}dr\Big]\Big|\mathcal{F}_t\Big];\\
(\partial_\mu L)^*_j(t) &:= E\Big[\widetilde{E}\Big[\int_0^T (\partial_\mu L)_j(\widetilde{\theta}_r;(X,Y,u))(t)dr\Big]\Big|\mathcal{F}_t\Big],\ j=1,2,3;\\
(\partial_\mu \varphi)^*_j(t) &:= E\Big[\widetilde{E}\Big[(\partial_\mu \varphi)_j(\widetilde{X}_T,P_{(X,Y,u)};(X,Y,u))(t)\Big]\Big|\mathcal{F}_t],\ j=1,2,3;\\
(\partial_\mu \Phi)^*_j(t)[p(T)] &:= E\Big[\widetilde{E}\Big[(\partial_\mu \Phi)_j(\widetilde{X}_T,P_{(X,u)};(X,u))(t)\widetilde{p}(T)\Big]\Big|\mathcal{F}_t\Big],\ j=1,2.
\end{aligned}
\tag{5.5}
$$

Using the notations introduced above we consider the following adjoint FBSDE,

$$\begin{cases} dp(t) = \{f_y(\theta_t)p(t) + (\partial_\mu f)_2^*(t)[p] - L_y(\theta_t) - (\partial_\mu L)_2^*(t) - (\partial_\mu \varphi)_2^*(t)\}dt \\ \qquad\qquad + \{f_z(\theta_t)p(t) - L_z(\theta_t)\}dB_t,\ t \in [0,T], \\ p(0) = 0, \end{cases} \tag{5.6}$$

$$\begin{cases} dq(t) = -\{\sigma_x(t)k(t) + (\partial_\mu \sigma)_1^*(t)[k] - f_x(\theta_t)p(t) - (\partial_\mu f)_1^*(t)[p] \\ \qquad\qquad + L_x(\theta_t) + (\partial_\mu L)_1^*(t) + (\partial_\mu \varphi)_1^*(t) - (\partial_\mu \Phi)^*(t)[p(T)]\}dt \\ \qquad\qquad + k(t)dB_t,\ t \in [0,T], \\ q(T) = \varphi_x(X_T, P_{(X,Y,u)}) - \Phi_x(X_T, P_{(X,u)})p(T). \end{cases} \tag{5.7}$$

Lemma 5.4.

Under the assumptions (H5) and (H6) the adjoint equation (5.6)-(5.7) has a unique adapted solution $(p, q, k) \in S_{\mathbb{F}}^2 \times S_{\mathbb{F}}^2 \times M_{\mathbb{F}}^2$.

Proof: Equation (5.6) is an affine mean-field forward stochastic equation with delay (Recall the definition of $(\partial_\mu f)_2^*(t)[p]$), thus the existence and the uniqueness is obvious. After getting the solution of (5.6), equation (5.7) is a mean-field BSDE with anticipation like (5.2) (Recall the definition of $(\partial_\mu \sigma)_1^*(t)[k]$). From Lemma 5.1, we see that there exists a unique solution $(q, k) \in S_{\mathbb{F}}^2 \times M_{\mathbb{F}}^2$. The proof is complete. $\square$

The following lemma studies the duality relation between variational equation and adjoint FBSDE:

Lemma 5.5.

Let p be the solution to the adjoint SDE (5.6), (q,k) the solution to the adjoint BSDE (5.7), and (X^1,Y^1,Z^1) the solution to (5.2). Then we have

$$
\begin{aligned}
& E[X_T^1 q(T) + Y_T^1 p(T)] \\
&= E\Big[\int_0^T X_t^1 \Big\{(\partial_\mu \Phi)_1^*(t)[p(T)] - L_x(\theta_t) - (\partial_\mu L)_1^*(t) - (\partial_\mu \varphi)_1^*(t)\Big\}dt\Big] \\
&\quad - E\Big[\int_0^T Y_t^1 \big\{L_y(\theta_t) + (\partial_\mu L)_2^*(t) + (\partial_\mu \varphi)_2^*(t)\big\}dt\Big] - E\Big[\int_0^T Z_t^1 L_z(\theta_t)dt\Big] \\
&\quad + E\Big[\int_0^T v_t\{(\partial_\mu \sigma)_2^*(t)[k] - (\partial_\mu f)_3^*(t)[p]\}dt\Big]
\end{aligned} \tag{5.8}
$$

The Hamiltonian H associated with our control problem:

The classical Hamiltonian H

$$H(t,x,y,z,p,k,\mu,\nu) = -f(t,x,y,z,\mu)p + \sigma(t,x,\nu)k + L(t,x,y,z,\mu),$$
$$(t,x,y,z,p;k,\mu,\nu)\in [0,T]\times\mathbb{R}^4\times\mathcal{P}_2(C_T^2)\times L^2([0,T]))\times\mathcal{P}_2(C_T)\times L^2([0,T])),$$

cannot be that associated with our control problem.

Indeed, our new Hamiltonian has to consider two effects:

• Let u be an optimal control process and define for the associate dynamics (X,Y,Z) the probability measures $\nu = P_{(X,u)}, \mu = P_{(X,Y,u)}$.

The terminal value $Y_T = \Phi(X_T, P_{(X,u)})$ and also $\varphi(X_T, P_{(X,Y,u)})$ depend on the law of the whole path (X,u) and (X,Y,u), respectively.

This has as consequence that they produce their own time-dependent adapted coefficients $(\partial_\mu\Phi)^*_j(t)[p(T)]$ and $(\partial_\mu\varphi)^*_i(t),\ j=1,2,\ i=1,2,3,$ which we have to take into account in the definition of the Hamiltonian and its derivatives.

• It adds that, like $(\partial_\mu \Phi)^*_j(t)[p(T)]$, the derivatives w.r.t. the measure $(\partial_\mu \sigma)^*_1(t)[k]$, $(\partial_\mu f)^*_i(t)[p]$, $i = 1, 2$, are linear functionals but now of the whole solution process $(p, (q, k))$ of the adjoint forward-backward SDE, and don't depend on p and k only in a multiplicative way.

This makes that our Hamiltonian cannot have the classical form. We define the Hamiltonian just as the following vector function:

$$\begin{aligned} &H(t, x, y, z, \nu, \mu) \\ &:= (-f(t, x, y, z, \mu),\ \sigma(t, x, \nu),\ L(t, x, y, z, \mu), -\Phi(\cdot, \nu), \varphi(\cdot, \mu)). \end{aligned} \tag{5.9}$$

For the derivatives of the Hamiltonian, we introduce the following notations. For $\nu = P_{(X,u)}, \mu = P_{(X,Y,u)}$, $\sigma(t) = \sigma(t, X_t, \nu)$, $\sigma_x(t) = \sigma_x(t, X_t, \nu)$, $f_x(\theta_t) = f_x(t, X_t, Y_t, Z_t, \mu)$, and $L_x(\theta_t) = L_x(t, X_t, Y_t, Z_t, \mu)$, we put

$$(\partial_x H)(t) = (-f_x(\theta_t), \sigma_x(t), L_x(\theta_t), 0, 0),$$

$$H_x(t) = ((\partial_x H)(t), (p(t), k(t), 1, 0, 0))_{\mathbb{R}^5},$$

where $(\cdot, \cdot)_{\mathbb{R}^5}$ denotes the inner product in $\mathbb{R}^5$.

In the same sense we define

$$(\partial_y H)(t) = (-f_y(\theta_t), 0, L_y(\theta_t), 0, 0),$$

$$(\partial_z H)(t) = (-f_z(\theta_t), 0, L_z(\theta_t), 0, 0),$$

and

$$H_y(t) = ((\partial_y H)(t), (p(t), k(t), 1, 0, 0))_{\mathbb{R}^5},$$

$$H_z(t) = ((\partial_z H)(t), (p(t), k(t), 1, 0, 0))_{\mathbb{R}^5}.$$

Concerning the derivatives with respect to the measure, we write

$$\begin{aligned}(\partial_\mu H)_x(t) = &- (\partial_\mu f)_1^*(t)[p] + (\partial_\mu \sigma)_1^*(t)[k] + (\partial_\mu L)_1^*(t)\\ &- (\partial_\mu \Phi)_1^*(t)[p(T)] + (\partial_\mu \varphi)_1^*(t),\\ (\partial_\mu H)_y(t) = &- (\partial_\mu f)_2^*(t)[p] + (\partial_\mu L)_2^*(t) + (\partial_\mu \varphi)_2^*(t),\\ (\partial_\mu H)_v(t) = &- (\partial_\mu f)_3^*(t)[p] + (\partial_\mu \sigma)_2^*(t)[k] + (\partial_\mu L)_3^*(t)\\ &- (\partial_\mu \Phi)_2^*(t)[p(T)] + (\partial_\mu \varphi)_3^*(t).\end{aligned}$$

Putting

$$\begin{aligned}\mathcal{D}_x H(t) :=& H_x(t) + (\partial_\mu H)_x(t),\\ \mathcal{D}_y H(t) :=& H_y(t) + (\partial_\mu H)_y(t),\\ \mathcal{D}_z H(t) :=& H_z(t),\\ \mathcal{D}_v H(t) :=& (\partial_\mu H)_v(t).\end{aligned} \tag{5.10}$$

We can write now the adjoint FBSDE (5.6)-(5.7) in the following way:

$$\begin{cases} dp(t) = -\mathcal{D}_y H(t)dt - \mathcal{D}_z H(t)dB_t,\ t \in [0,T], \\ dq(t) = -\mathcal{D}_x H(t)dt + k(t)dB_t,\ t \in [0,T], \\ p(0) = 0,\ q(T) = \varphi_x(X_T, P_{(X,Y,u)}) - \Phi_x(X_T, P_{(X,u)})p(T). \end{cases} \tag{5.11}$$

Then we will get the following stochastic maximum principle:

Theorem 5.2.

Let u be an optimal control of the mean-field FBSDE control problem.

Then, recalling the definition of $\mathcal{D}_v H(t)$, we have the maximum principle:

$$\mathcal{D}_v H(t)(v - u(t)) \geq 0,\ \text{for all } v \in U,\ dtdP\text{-}a.e. \tag{5.12}$$

where (p, q, k) is the solution of the adjoint equation (5.6) and (5.7).

Last but not least we show that the optimality condition given by Pontryagin's SMP (5.12) is not only necessary, but, combined with a suitable convexity assumption for the Hamiltonian H, it is also sufficient.

Theorem 6.1.

Let us suppose the convexity of the Hamiltonian $(-f(t,x,y,z,\mu)p(t), \sigma(t,x,\nu)k(t), L(t,x,y,z,\mu), -\Phi(x',\nu)p(T), \varphi(x',\mu))$ in $(x,x',y,z,\nu,\mu) \in \mathbb{R}^4 \times \mathcal{P}_2(\mathcal{C}_T \times \mathcal{U}_T) \times \mathcal{P}_2(\mathcal{C}_T^2 \times \mathcal{U}_T)$, where (p,q,k) is the solution of the adjoint equation (5.6)-(5.7). Furthermore, we continue to suppose the standard assumptions (H5)-(H6) of the preceding section. Then, if an admissible control $u \in \mathcal{U}_{ad}$ satisfies (5.12), it is optimal:

$$J(w) \geq J(u), \text{ for all } w \in \mathcal{U}_{ad} \text{ such that } v := w - u \in \mathcal{U}_{ad}.$$

Remark 6.1.

The convexity of $(x, x', y, z, \mu, \nu) \to (-f(t, x, y, z, \mu)p(t), -\Phi(x', \nu)p(T))$ can be got by supposing, for instance, the convexity of $-f(t, \cdot, \cdot, \cdot, \cdot)$ and $-\Phi$, and by taking assumptions on the coefficients of the adjoint forward SDE (5.6) which guarantee that $p(t) \geq 0$, $t \in [0, T]$.

Thank you very much
for your attention!