AIMS *Mathematics*

Research article

Mean-field-type games

Hamidou Tembine[*]

Learning and Game Theory Laboratory, New York University Abu Dhabi

* Correspondence: tembine@nyu.edu

Abstract: This article examines games in which the payoffs and the state dynamics depend not only on the state-action profile of the decision-makers but also on a measure of the state-action pair. These game situations, also referred to as mean-field-type games, involve novel equilibrium systems to be solved. Three solution approaches are presented: (i) dynamic programming principle, (ii) stochastic maximum principle, (iii) Wiener chaos expansion. Relationship between dynamic programming and stochastic maximum principle are established using sub/super weak differentials. In the non-convex control action spaces, connections between the second order weaker differentials of the dual function and second order adjoint processes are provided. Multi-index Wiener chaos expansions are used to transform the non-standard game problems into standard ones with ordinary differential equations. Aggregative and moment-based mean-field-type games are discussed.

Keywords: game theory; coalition; mean-field; Wiener chaos; dynamic programming; maximum principle
Mathematics Subject Classification: 91A15, 91A25, 49L20, 49L25

1. Introduction

The term "mean-field" has been referred to a physics concept that attempts to describe the effect of an infinite number of particles on the motion of a single particle. Researchers began to apply the concept to social sciences in the early 1960s to study how an infinite number of factors affect individual decisions. However, the key ingredient in a game-theoretic context is the influence of the distribution of states and or control actions into the payoffs of the decision-makers who may have different preferences and characters, and are not necessarily exchangeable per class (or indistinguishable per class/type). A mean-field-type game is a game in which the payoffs and/or the state dynamics coefficient functions involve not only the state and actions profiles but also the distributions of state-action process (or its marginal distributions). In contrast to mean-field games [2, 3] in which a single player does not influence of the mean-field terms, here, it is totally different. In mean-field-type games, a single player

may have a strong impact on the mean-field terms. This paper presents the key ingredients and recent development of mean-field-type game theory.

1.1. Related works

Stochastic games is a model for dynamic interactions in which the state evolves in a way that depends on the actions of the decision-makers. The model was introduced in [1], who proved that two-player zero-sum discounted games with finite states under perfect state observation have a value and both players have optimal stationary strategies. In Shapley's initial model of stochastic games, the state transitions and instant payoff were given by the state-action profiles. However, explicit dependence on the distribution of states or distribution of actions were not examined.

1.1.1. Mean-field-type games with one player

A stochastic optimization problem in which the first moments (expected values of the state) influences the state dynamics and the performance criterion were examined in [10]. This corresponds to a first moment-based mean-field-type game with a single decision-maker. The model was extended by [11] to include aggregative structures of the distribution of states and a new stochastic maximum principle was established in the risk-neutral setting. The resulting system is not a simple augmented state space approach because the expected value of the Hamiltonian is involved in it. The work of [17] extends that result to performance criteria that include the entire probability measure of the states. The work of [12, 13] consider both expected value of states and expected of control-actions in the performance criteria. We refer the reader to [4, 14, 15, 16] for stochastic maximum principle of mean-field type.

1.1.2. Mean-field-type games with several players

Mean-field-type games with two or more players can be seen as the multi-agent generalization of the single agent mean-field-type control problem. In [18, 19] it is shown that the methodology can be used for risk-sensitive mean-field-type games where weakened conditions on the drifts are provided. State measurement noise and partial observation studies were conducted in [20, 21]. See also [27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 46, 47, 48, 49, 50].

Wiener chaos expansion has been proposed in [30] to transform the mean-field-type games into equivalent deterministic game problems where the state process is replaced by its polynomial chaos or Wiener chaos expansion counterpart. The decomposition into decoupling series expansion is similar to the one known in the theory of stochastic processes, as the Kosambi-Karhunen-Loeve theorem, is a representation of a stochastic process as an infinite linear combination of orthogonal functions, analogous to a Fourier series representation of a function on a bounded domain. Cooperative mean-field-type games were introduced in [5, 41, 42, 43, 44, 45]. In [40] several applications of mean-field-type games in engineering are provided.

1.2. Structure

The rest of the paper is structured as follows. Section 2 presents a generic model. Section 3 focuses on solution methods. Classes of mean-field-type games are provided in Section 4. Section 5 examines

the partial state observation case. Section 6 presents some recent development and or extensions of mean-field-type games. Section 7 concludes the paper.

Table 1 summarizes the notations used in the article.

Table 1. Table of Notations

$\mathcal{I}$	$\triangleq$	set of decision-makers
T	$\triangleq$	Length of the horizon
$[0, T]$	$\triangleq$	horizon of the mean-field-type game
t	$\triangleq$	time index.
$\mathcal{S}$	$\triangleq$	state space
$s(t)$	$\triangleq$	state at time t
$\Delta(\mathcal{S})$	$\triangleq$	set of probability measures on $\mathcal{S}$
$m(t, .)$	$\triangleq$	probability measure of the state at time t
A_i	$\triangleq$	control action set of decision-maker $i \in \mathcal{I}$
$a_i(.)$	$\triangleq$	strategy of decision-maker $i \in \mathcal{I}$
$a = (a_i)_{i \in \mathcal{I}}$	$\triangleq$	strategy profile
$b(t, s, m, a)$	$\triangleq$	drift coefficient function
$\sigma(t, s, m, a)$	$\triangleq$	diffusion coefficient function
$r_i(t, s, m, a)$	$\triangleq$	instant payoff of decision-maker $i \in \mathcal{I}$
$g_i(s, m)$	$\triangleq$	terminal payoff of decision-maker $i \in \mathcal{I}$
$\mathcal{R}_{i,T}(m_0, a)$	$\triangleq$	cumulative payoff of i
$V_i(t, m)$	$\triangleq$	equilibrium payoff of $i \in \mathcal{I}$
(p_i, q_i)	$\triangleq$	first order adjoint process of $i \in \mathcal{I}$
(P_i, Q_i)	$\triangleq$	second order adjoint process of $i \in \mathcal{I}$
$v_i^*(t, s)$	$\triangleq$	dual function of $i \in \mathcal{I}$
H_i	$\triangleq$	$r_i + bp_i + \sigma q_i$ of $i \in \mathcal{I}$
$H_{i,s}$	$\triangleq$	$r_{i,s} + b_s p_i + \sigma_s q_i$
$H_{i,ss}$	$\triangleq$	$r_{i,ss} + b_{ss} p_i + \sigma_{ss} q_i$

2. Mean-field-type games

2.1. The interaction model

A basic mean-field-type game in continuous time is described by:

- The set of decision-makers $\mathcal{I} = \{1, 2, \ldots, I\}$, where $I \in \mathbb{N}$.
- The horizon of the interaction is the interval $[0, T]$, $T > 0$.
- There is a non-empty state space $\mathcal{S}$. The set of probability measures on $\mathcal{S}$ is denoted by $\Delta(\mathcal{S})$.
- For each decision-maker i, a non-empty control action set A_i is available. The set A_i is not necessarily convex. The set of all control actions of all the decision-makers is $A = \prod_{i \in \mathcal{I}} A_i$.
- An instant payoff of decision-maker i is $r_i : \mathcal{S} \times \Delta(\mathcal{S}) \times A \to \mathbb{R}$
- The state evolution is explicitly given by a controlled Itô's stochastic differential equation of

mean-field type, called controlled McKean-Vlasov equation:

$$s(t) = s_0 + \int_0^t b \, dt' + \int_0^t \sigma \, dB(t'), \; t > 0.$$

with $s(0) = s_0 \in S$, and $b, \sigma : \; S \times \Delta(S) \times A \to \mathbb{R}$, b is the drift coefficient functional and σ is the diffusion coefficient functional, B is a one-dimensional standard Brownian motion on a given probability space $(\Omega, \mathbb{P}, \mathbb{F})$.

Given a initial state s_0 which is drawn from the initial distribution m_0, the game $\mathcal{G}(s_0)$ proceeds as follows. At each instant, each decision-maker observes the state (perfect monitoring, perfect state observation), chooses a control action according her strategy (defined below) and observes/measures her payoff.

For this game, we adopt the following notions: An admissible control of decision-maker $i \in \mathcal{I}$ is progressively measurable process with respect to the filtration $\mathcal{F}_t$, taking values in A_i. $L_{\mathcal{F}}^2([0,T], A_j)$ denotes the set of $\mathcal{F}$-adapted, A_j-valued processes within $[0,T]$, and $C([0,T], \mathcal{X})$ denotes the set of continuous (hence measurable) $[0,T] \to \mathcal{X}$. The set of admissible controls of decision-maker i is denoted by $\mathcal{A}_i$.

We identify two processes a_i and $\tilde{a}_i$ in $\mathcal{A}_i$ if

$$\mathbb{P}(a_i = \tilde{a}_i, \; a.e. \; on \; [0,T]) = 1.$$

A strategy for decision-maker i starting at time 0 is a measurable map (renamed again by) $a_i : \; [0,T] \times C([0,T], \mathcal{X}) \times \prod_{j \neq i} L_{\mathcal{F}}^2([0,T], A_j) \to A_i$ for which there exists $\epsilon > 0$ such that for any

$$(t', f_1, f_2) \in (0,T] \times \{C([0,T], \mathcal{X}) \times \prod_{j \neq i} L_{\mathcal{F}}^2([0,T], A_j)\}^2,$$

if $f_1 = f_2$ on $[0,t']$ then $a_i(., f_1) = a_i(., f_2)$, on $[0, t' + \epsilon]$. To each strategy profile one can associate a control process profile. This will allow us to work with both open-loop and feedback form of strategies by considering $a_i(., s^{s_0,a}, m)$. For open-loop strategies the information structure is limited to $\{t\}$, a_i is simply a measurable function of time (and the initial data). The stochastic maximum principle can be used as a methodology for finding optimal open-loop strategies. The information structure for feedback strategies is s and its distribution (and the initial data). The dual adjoint functions which are obtained from the Bellman functions can be used for finding feedback strategies where here the feedback is a state-and-mean-field feedback.

The cumulative payoff of decision-maker i is

$$\mathcal{R}_{i,T}(m_0, a) := g_i(s(T), m(T, .)) + \int_0^T r_i(s(t), m(t, .), a(t)) \, dt,$$

where $g_i(s(T), m(T, .))$ is the terminal payoff of i. The risk-neutral payoff of i is $\mathbb{E}[\mathcal{R}_{i,T}(m_0, a)]$. The risk-neutral mean-field-type game $\mathcal{G}_{0,T}(m_0)$ is the normal-form game $(\mathcal{I}, (\mathcal{A}_i, \mathbb{E}\mathcal{R}_{i,T})_{i \in \mathcal{I}})$.

2.2. Solution concept

We now introduce the key problems addressed in this paper and some solution concepts.

2.2.1. Best-response correspondence

Given $a_{-i} := (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots)$, the best-response value problem associated with decision-maker i, is

$$V_i(0, m_0) = \begin{cases} \sup_{a_i \in \mathcal{A}_i} \mathbb{E}[\mathcal{R}_{i,T}(m_0, a)] \\ s(t) = s^{s_0, a}(t), \ s(0) = s_0 \sim m_0 \\ m(t, .) = m^{m_0, a}(t, .) = \mathbb{P}_{s^{s_0, a}(t)}, \end{cases} \tag{1}$$

The strategies a_i of i that achieve the above best-response value $V(0, m_0)$ are called best-response strategies of i to $a_{-i} := (a_j)_{j \neq i}$. The set of such strategies is denoted by $BR_i(a_{-i})$. We are interested in characterizing best response strategies of every decision-maker.

2.2.2. Mean-field-type equilibria

An equilibrium point is a strategy profile a such that for every decision-maker i, $a_i \in BR_i(a_{-i})$ and the resulting mean-field is $m(t, .) = \mathbb{P}_{s^{s_0, a}(t)}$. We are interested in characterizing equilibrium strategies of every decision-maker.

2.2.3. Well-posedness of the state equation

As the state dynamics is a stochastic differential equation of mean-field type, the existence and uniqueness of solution to the state equation is not always guaranteed. Below, we provide sufficiency conditions for existence and uniqueness of a solution.

Lemma 1 (Existence). *If the coefficient functions b, σ are continuously differentiable with respect to s, m and the Gateaux-derivatives are bounded, then for each strategy profile $(a_i)_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} \mathcal{A}_i$, there is a unique solution to the state dynamics which we denote by $s(t) := s^{s_0, a}(t)$.*

A proof of this Lemma can be directly obtained from [7] by choosing a strategy profile $(a_i)_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} \mathcal{A}_i$.

Lemma 2. *The probability law of the state solves (in the weak sense) the following Fokker-Planck-Kolmogorov equation of mean-field type:*

$$m_t + (bm)_s - \frac{1}{2}(\sigma^2 m)_{ss} = 0, \ \ m(0, .) = m_0(.). \tag{2}$$

The proof is by now standard using integration by parts in the sense of distribution.

3. Solution approaches

In this section we present three different solution approaches for solving mean-field-type game problems. We also provide some relationships between these methods. We introduce a notion of Gâteaux differentiability with respect to a measure.

Let $\mathcal{S}$ be a vector space, $b \in L^{\alpha}(\Omega \times [0, T] \times \mathcal{S} \times \Delta(\mathcal{S}), \mathbb{R})$, $O \subset L^{\alpha^*}$ is an open set.

Definition 1 (Gâteaux derivative). *The Gâteaux differential $b_m[\tilde{m}]$ of the functional b at $m \in O$ in the direction $\tilde{m}$ is defined as*

$$\lim_{\epsilon \to 0_+} \frac{d}{d\epsilon} b(., t, s, m + \epsilon\tilde{m}) = \int b_m(., t, s, m)(\xi)\, \tilde{m}(d\xi).$$

If the limit exists for all direction $\tilde{m}$, then one says that b is Gâteaux differentiable at m.

If $m \in L^\alpha$ and $b_m \in L^{\alpha^*}$, the limit above is finite. The limit appearing in Definition 1 is taken relative to the topology of L^{α^*}.

Note that one can connect the derivative with the respect to s of the Gateaux derivative b_m to the notion of functional derivative to respect to measure and also to the Wasserstein gradient.

We provide Gâteaux differentiation of $\|s\|_\alpha$−based functions:

- α−norm: Let $g(m) = (\int |s|^\alpha m(t, ds))^{1/\alpha} =: m_\alpha^{1/\alpha}$, and $m \in L^\alpha$, then

$$\lim_{\epsilon \to 0_+} \frac{d}{d\epsilon} g(m + \epsilon\tilde{m}) = \frac{1}{\alpha} \left[\int |\xi|^\alpha\, m(d\xi) \right]^{\frac{1}{\alpha}-1} \left[\int |\xi|^\alpha\, \tilde{m}(d\xi) \right].$$

Thus, $g_m(m)(s) = \frac{s^\alpha}{\alpha m_\alpha^{\alpha-1}}$, and $\partial_s g_m(m)(s) = \frac{s^{\alpha-1}}{m_\alpha^{\alpha-1}} \neq 0 = \partial_m[g_s]$.

- L^α−normed drift: Let

$$\bar{b} = \left(\int_{y \in \mathcal{S}} |b|^\alpha(., t, s^a(t), y, a(t)) m^a(t, dy) \right)^{\frac{1}{\alpha}},$$

i.e., the L^α−norm of b with respect to the measure $m^a(t, .)$. We compute the Gâteaux-derivative of the L^α−normed drifts: $\bar{b}_m(., t, s, m)(\xi) := \frac{b^\alpha(., t, s, \xi)}{\alpha \bar{b}^{\alpha-1}}$. By changing variables, one has $\bar{b}_m(., t, \xi, m)(s) = \frac{b^\alpha(., t, \xi, s)}{\alpha \bar{b}^{\alpha-1}(t, \xi, m)}$. We differentiate with respect to the state s to get:

$$\partial_s \bar{b}_m(., t, \xi, m)(s) = \frac{b^{\alpha-1}(., t, \xi, s) b_y(., t, \xi, s)}{\bar{b}^{\alpha-1}(t, \xi, m)}. \tag{3}$$

$E[L \partial_s \bar{b}_m(., t, S, m)(s)] = E[L \frac{b^{\alpha-1}(., t, S, s) b_y(., t, S, s)}{\bar{b}^{\alpha-1}(t, S, m)}] := \tilde{E}[\tilde{L} \frac{b^{\alpha-1}(., t, \tilde{S}, s) b_y(., t, \tilde{S}, s)}{\bar{b}^{\alpha-1}(t, \tilde{S}, m)}]$, where the notation $\tilde{E}$ denotes the expectation with respect to the variables with $\tilde{S}$ which is a copy of S. We now replace the argument s by S to get $\tilde{E}[\tilde{L} \partial_s \bar{b}_m(., t, \tilde{S}, m)(S)] = \tilde{E}[\tilde{L} \frac{b^{\alpha-1}(., t, \tilde{S}, S) b_y(., t, \tilde{S}, S)}{\bar{b}^{\alpha-1}(t, \tilde{S}, m)}]$.

3.1. Dynamic programming principle

In contrast to classical differential games or classical mean-field games in which one can directly use the standard Dynamic Programming Principle (DPP) on $s(t)$, here, the presence of the term $m^a(t, .)$ in the functions r_i, b, σ may create a time-inconsistency, and $s(t)$ is not the appropriate state for DPP. The next Proposition identifies an appropriate state space under which DPP can be obtained and a Hamilton-Jacobi-Bellman (HJB) equation can be derived.

Proposition 1. *If there exists V such that*

$$\begin{cases} i \in \mathcal{I}, \\ V_{i,t} + \int_s [\sup_{a_i}\ r_i + b V_{i,sm} + \frac{\sigma^2}{2} V_{i,ssm}] m(t, ds) = 0, \\ V_i(T, m) = \int_s m(T, ds) g_i(s, m(T, .)), \\ m(t, .) = \mathbb{P}_{s^a(t)} \end{cases} \tag{4}$$

then $V_i(0, m_0)$ is an equilibrium payoff and the optimal strategy a_i of i maximizes $r_i + bV_{i,sm} + \frac{\sigma^2}{2}V_{i,ssm}$.

Proof. In order to use a dynamic programming principle we look for an augmented state such that one gets a Markovian system. By rewriting the payoff functional $\mathbb{E}\mathcal{R}_{i,T}$ as a function of the measures $m(t, .)$ and the strategies $a_i(.)$, we obtain a classical setup. Hence, we only need the evolution of the measures $m(t, .)$ which is given by (2). Then, one obtains a deterministic differential game problem with m as state (in infinite dimensions). The best response of i is

$$V_i(0, m_0) = \begin{cases} \sup_{a_i \in \mathcal{A}_i} \mathbb{E}[\mathcal{R}_{i,T}(m_0, a)] \\ m(t, .) \text{ solves (2).} \end{cases}$$

Applying the classical DPP to the deterministic problem yields

$$V_{i,t} + \sup_{a_i}[\int_s r_i m(t, ds) + \langle \tilde{b}, V_{i,m} \rangle] = 0,$$

where $\tilde{b}(m)(t, s) = -(bm)_s + \frac{1}{2}(\sigma^2 m)_{ss}$. We apply the theory of distributions to establish the following equalities:

$$\langle \tilde{b}, V_{i,m} \rangle = \int_s \tilde{b}(m) V_{i,m} \, ds \tag{5}$$

$$= \int_s -(bm)_s V_{i,m} + \frac{1}{2}(\sigma^2 m)_{ss} V_{i,m} \tag{6}$$

$$= \int_s [bV_{i,sm} + \frac{\sigma^2}{2} V_{i,ssm}] m(t, ds), \tag{7}$$

which is obtained by integration by part (in the sense of distribution). This completes the proof. $\square$

Example 1 (Mean-field-free equations). *If r_i, g_i, b, σ are all independent of m then $V_i(t, m) = \int v_i(t, s) \, m(t, ds) = \mathbb{E}_{m(t,.)} v_i(t, s(t))$ where the function $(v_i(t, s))_i$ solves the classical Bellman system*

$$\begin{cases} i \in \mathcal{I}, \\ v_{i,t} + H_i = 0, \\ v_i(T, s) = g_i(s, m(T, .)), \end{cases} \tag{8}$$

Moreover

$$v_i(0, s_0) = \begin{cases} \sup_{a_i \in \mathcal{A}_i} \mathbb{E}[\mathcal{R}_{i,T}(s_0, a)] \\ s(t) = s^{s_0,a}(t), \ s(0) = s_0 \end{cases} \tag{9}$$

Proposition 1 provides a sufficiency condition for equilibria. However, the existence of classical solution to the infinite-dimensional HJB equation is not guaranteed and the value function may not be differentiable as in the classical case. Weaker notion of solution such as viscosity solution using weak sub/super-differential set will be introduced in the next section.

3.2. Dual functions

Since the DPP equilibrium system (4) is in infinite dimensions, the solvability of such an integro-partial differential equation poses some technical difficulties. However, when the value function V is weakly differentiable with respect to (s, m) then a finite dimensional quantity can be obtained as it is in the dual space.

Definition 2. *The function $v^*(t, s) = V_m(t, s)$ is the dual function associated to V.*

Example 2. *In the mean-field-free setting, i.e., when r_i, g_i, b, σ are independent of m, the dual function $v^*(t, s) = V_m(t, s)$ coincides with the function $v(t, s)$. However, the two functions do not coincide in the mean-field-dependent setting.*

Proposition 2. *If V_m, V_{sm}, V_{ssm} exist and are continuous then the equilibrium dual function $v^* = V_m$ solves the following system of partial differential equations*

$$
\begin{cases}
i \in \mathcal{I}, \\
v_{i,t}^* + H_i + \mathbb{E}H_{i,m} = 0, \\
v_i^*(T, s) = g_i(s, m(T, .)) + \mathbb{E}g_{i,m}, \\
m(t, .) = \mathbb{P}_{s^a(t)}, \ m(0, .) = m_0(.)
\end{cases}
\tag{10}
$$

where $H_i(t, s, m, v_{i,s}^*, v_{i,ss}^*) = \sup_{a_i} [r_i + bv_{i,s}^* + \frac{\sigma^2}{2}v_{i,ss}^*].$

Proof. We differentiate the HJB system (4) in the direction $\tilde{m} = m$,

$$
\begin{cases}
i \in \mathcal{I}, \\
(V_{i,t})_m + [\sup_{a_i} \ r_i + bV_{i,sm} + \frac{\sigma^2}{2}V_{i,ssm}] \\
+ \int_{\tilde{s}}[\sup_{a_i} \ r_i + bV_{i,sm} + \frac{\sigma^2}{2}V_{i,ssm}]_m(t, \tilde{s}, m)(s)m(t, d\tilde{s}) = 0, \\
V_{i,m}(T, s) = g_i(s, m(T, .)) + \int_s m(T, ds)g_i(s, m(T, .)), \\
m(t, .) = \mathbb{P}_{s^a(t)}
\end{cases}
\tag{11}
$$

Noting that $(V_{i,t})_m = v_{i,t}^*$, $V_{i,sm} = v_{i,s}^*$ and $V_{i,ssm} = v_{i,ss}^*$, the latter system (11) becomes the announced one. This completes the proof.

$\square$

As pointed out by [17], the dual function $v_i^*(0, s_0)$ is not the equilibrium payoff of the decision-maker i. The dual function plays an important role in establishing a stochastic maximum principle as we will see in the next subsection.

3.3. Stochastic maximum principle

We now present first order and second-order adjoint processes that are useful in establishing necessary conditions for mean-field-type equilibria. The advantage of the adjoint processes is that they solve a class of linear backward SDEs of mean-field type. The following result provides existence and uniqueness conditions.

Lemma 3 ([9, 11]). *Consider the following mean-field backward SDE*

$$
p(t) = p(T) + \int_t^T \tilde{E}[\hat{f}(t', \tilde{p}(t'), \tilde{q}(t'), p(t'), q(t'))] \, dt' - \int_t^T q(t')dB(t'),
$$

where $p(T)$ is a progressively measurable, square integrable random variable. Let $\hat{f}(t, ., ., ., .)$ be Lipschitz for all time $t \in [0, T]$ and $t \mapsto \hat{f}(t, 0, 0, 0, 0)$ be square integrable over $[0, T]$. Then, the mean-field backward SDE has a unique adapted solution satisfying

$$
E\left[\sup_{t\in[0,T]} |p(t)|^2 + \int_0^T |q(t)|^2 dt\right] < \infty.
\tag{12}
$$

Consider the first-order adjoint processes $(p_i, q_i)_{i \in \mathcal{I}}$

$$\begin{cases} dp_i = -[H_{i,s} + \tilde{\mathbb{E}}\partial_s H_{i,m}]dt + q_i dB, \\ p_i(T) = g_{i,s}(T) + \tilde{\mathbb{E}}\{\partial_s g_{i,m}(T)\}, \\ i \in \mathcal{I}. \end{cases} \qquad (13)$$

Lemma 4. *If the functions b, σ, r_i, g_i are continuously differentiable with respect to (s, m), all their first-order derivatives with respect to (s, m) are continuous in (s, m, a), and bounded. Then, the first-order adjoint system is a linear SDE with almost surely bounded coefficient functions. There is a unique $\mathcal{F}$−adapted solution such that*

$$\mathbb{E}\left[\sup_{t \in [0,T]} |p_i(t)|^2 + \int_0^T |q_i(t)|^2 dt \right] < +\infty.$$

These strong smoothness conditions on b, σ, r_i, g_i can be considerably weakened using representations of weak sub/super-differential sets. We refer the reader to [8, 9] for more details on existence and uniqueness of solutions to backward SDE of mean-field type.

Proof. By choosing $\hat{f}(t, \tilde{p}(t), \tilde{q}(t), p(t), q(t)) = \alpha_0(t, .) + \alpha_1(t, .)\tilde{p}(t) + \alpha_2(t, .)\tilde{q}(t) + \alpha_3(t, .)p(t) + \alpha_4(t, .)q(t)$ where $\alpha_i(t, .)$ are measurable bounded coefficient functions, one gets a backward equation in the form of the adjoint equations. Under the assumptions imposed on b, σ, r_i, g_i, the coefficient α_i fulfill the requirement and result follows as a direct application. □

Note that for linear-quadratic mean-field-type games, the payoff function $r_i = -s^2 - \int y^2 m(t, dy) - a_i^2 - \int_{a_i'} (a_i')^2 \mathbb{P}_{a_i}(da_i')$ does not satisfy the above assumptions. In that case, one can relax the boundedness assumption and replace it by L^2−estimates if the second moment is finite.

Consider the second-order adjoint processes $(P_i, Q_i)_{i \in \mathcal{I}}$

$$\begin{cases} dP_i = -\{H_{i,ss} + \tilde{\mathbb{E}}\partial_{ss} H_{i,m} + (2b_s + \sigma_s^2)P_i + 2\sigma_s Q_i\}dt + Q_i dB, \\ P_i(T) = g_{i,ss}(T) + \tilde{\mathbb{E}}\{\partial_{ss} g_{i,m}(T)\}, \\ i \in \mathcal{I}. \end{cases} \qquad (14)$$

Lemma 5. *If the functions b, σ, r_i, g_i are twice continuously differentiable with respect to (s, m), all their derivatives up to second order with respect to (s, m) are continuous in (s, m, a), and bounded. Then, the second order adjoint system is a linear backward SDE with almost surely bounded coefficient functions. There is a unique $\mathcal{F}$−adapted solution such that*

$$\mathbb{E}\left[\sup_{t \in [0,T]} |P_i(t)|^2 + \int_0^T |Q_i(t)|^2 dt \right] < +\infty.$$

Proof. Follows immediately from Lemma 3. □

Proposition 3 (Stochastic Maximum Principle). *If (s, a) is a pair of equilibrium state and equilibrium strategy, then there is vector of $\mathbb{F}$−adapted processes $(p_i, q_i, P_i, Q_i)_{i \in \mathcal{I}}$ such that $H_i(s, m, a_i', a_{-i}) - H_i(s, m, a) + \frac{1}{2}P_i(\sigma(s, m, a_i', a_{-i}) - \sigma(s, m, a))^2 \leq 0$ for all $a_i' \in \mathcal{A}_i$, almost every time $t \in [0, T]$ and $\mathbb{P}$−almost surely. In particular, if $\sigma(t, s, m, a) = \sigma(t, s, m)$ is independent of a_i then $\sup_{a_i' \in \mathcal{A}_i} H_i(s, m, a_i', a_{-i}) \leq H_i(s, m, a)$.*

Proof. The proof follows similar steps as in [10] by replacing the derivative with the respect to first moment component by the Gateaux-derivative with respect to m, for each decision-maker i when fixing the strategies of the other decision-makers. □

3.4. Relationships between DPP and SMP

In order to provide weaker conditions we introduce super- and subdifferentials of the viscosity solution V. Define the first order super-differential of the dual function as

$D_{t,w}^{1,+} V_m(t,w) = \{(d_1, d_2) \in \mathbb{R} \times T(\mathcal{S}) \mid \limsup_{t',w' \to t,w} \frac{V_m(t',w') - V_m(t,w) - d_1(t'-t) - d_2(w'-w)}{|t'-t| + \|w'-w\|} \leq 0\}$.

Similarly, the first order sub-differential is

$D_{t,w}^{1,-} V_m(t,w) = \{(d_1, d_2) \in \mathbb{R} \times T(\mathcal{S}) \mid \liminf_{t',w' \to t,w} \frac{V_m(t',w') - V_m(t,w) - d_1(t'-t) - d_2(w'-w)}{|t'-t| + \|w'-w\|} \geq 0\}$.

Using these weak derivatives, one has a relationship between stochastic maximum principle and dynamic programming principle in terms of inclusion: $p(t) \in D_s^{1,+} V_m(t,w)$. In particular, if V is Gâteaux-differentiable with respect to m and s then

$$\begin{cases} p_i(t) = V_{i,sm}(s(t), m(t)) = d_2, \\ q_i(t) = \sigma V_{i,ssm}(s(t), m(t)). \end{cases}$$

Proposition 4. *If the first and second order weak derivatives with respect to s of V_m exist then the process $p_i(t) = V_{i,sm}(s(t))$ solves the backward SDE:*

$$dp_i = -\alpha_i dt + \beta_i dB, \quad p_i(T) = g_{i,s}(s(T), m(T)) + \int_s \partial_s g_{i,m} m(T, ds),$$

where $\alpha_i = H_{i,s} + \tilde{\mathbb{E}}[\partial_s H_{i,m}]$, $H_{i,s}$ is a notation for $r_{i,s} + b_s p_i + \sigma_s q_i$ and should not be confused with derivative of the functional H_i. $\beta_i = \sigma V_{i,ssm} = q_i$.

Proof. Let $\tilde{f}_i(t, s) = V_{i,sm}(t, s) = v_{i,s}^*(t, s)$ Applying Ito's formula applied to $\tilde{p}_i(t) = \tilde{f}_i(t, s(t))$ yields

$$d\tilde{p}_i = [(\tilde{f}_i)_t + b(\tilde{f}_i)_s + \frac{\sigma^2}{2}(\tilde{f}_i)_{ss}]dt + \sigma(\tilde{f}_i)_s dB ,$$

We identify the coefficient processes $\tilde{q}_i = \sigma(\tilde{f}_i)_s = \sigma v_{i,ss}^* = \sigma V_{i,ssm}$. It remains to find the drift coefficient of $\tilde{p}_i$.

$$\begin{cases} \tilde{p}_i = v_{i,s}^*(t, s) = V_{i,sm}(t, s), \\ (\tilde{f}_i)_t = v_{i,ts}^*, \\ b(\tilde{f}_i)_s = b v_{i,ss}^*, \\ \frac{\sigma^2}{2}(\tilde{f}_i)_{ss} = \frac{\sigma^2}{2} v_{i,sss}^* \end{cases}$$

Summing together one obtains

$$\left\{ (\tilde{f}_i)_t + b(\tilde{f}_i)_s + \frac{\sigma^2}{2}(\tilde{f}_i)_{ss} = v_{i,ts}^* + b v_{i,ss}^* + \frac{\sigma^2}{2} v_{i,sss}^* \right. \tag{15}$$

In order to identify the latter term, one differentiates with respect to s the equation (10) satisfied by the dual function v_i^*.

$$\begin{cases} i \in \mathcal{I}, \\ v_{i,st}^* + r_{i,s} + (b v_{i,s}^*)_s + (\frac{\sigma^2}{2} v_{i,ss})_s + \tilde{\mathbb{E}} H_{i,sm} = 0, \\ v_{i,s}^*(T, s) = g_{i,s}(s, m(T, .)) + \tilde{\mathbb{E}} g_{i,sm}, \\ m(t, .) = \mathbb{P}_{s^a(t)}, \ m(0, .) = m_0(.) \end{cases} \tag{16}$$

which is expanded as

$$\begin{cases} i \in \mathcal{I}, \\ (v_{i,s}^*)_t + b v_{i,ss}^* + \frac{\sigma^2}{2} v_{i,sss}^* \\ + [r_{i,s} + b_s v_{i,s}^* + \sigma_s(\sigma v_{i,ss}^*) + \tilde{\mathbb{E}} H_{i,sm}] = 0, \\ v_{i,s}^*(T, s) = g_{i,s}(s, m(T, .)) + \tilde{\mathbb{E}} g_{i,sm}, \\ m(t, .) = \mathbb{P}_{s^a(t)}, \ m(0, .) = m_0(.) \end{cases} \tag{17}$$

It follows that $(v^*_{i,s})_t + bv^*_{i,ss} + \frac{\sigma^2}{2}v^*_{i,sss} = -[r_{i,s} + b_s v^*_{i,s} + \sigma_s(\sigma v^*_{i,ss}) + \tilde{\mathbb{E}}H_{i,sm}]$, and

$$dp̃_i = -[r_{i,s} + b_s v^*_{i,s} + \sigma_s(\sigma v^*_{i,ss}) + \tilde{\mathbb{E}}H_{i,sm}]dt + \sigma v^*_{i,ss}dB$$
$$= -[H_{i,s} + \tilde{\mathbb{E}}H_{i,sm}]dt + q_i dB \quad,$$

The two pair of processes (p_i, q_i) and $(p̃_i = v^*_{i,s}, q̃_i = \sigma v^*_{i,ss})$ solves the same backward SDE with the same terminal condition. Under the assumptions above, these two processes are identical. This completes the proof. $\quad\square$

Define the second order super-differential is $D^{1,2,+}_{t,w} V_m(t,w) = \{(d_1, d_2, d_3) \in \mathbb{R} \times T(\mathcal{S}) \times S^n \mid$
$\limsup_{t',w' \to t,w} \frac{V_m(t',w') - V_m(t,w) - d_1(t'-t) - d_2(w'-w) - \frac{1}{2}(w'-w)'d_3(w'-w)}{|t'-t| + \|w'-w\|^2} \leq 0\}$
Similarly, the second order sub-differential is $D^{1,2,-}_{t,w} V_m(t,w) = \{(d_1, d_2, d_3) \in \mathbb{R} \times T(\mathcal{S}) \times \mathbb{R}^n \times S^n \mid$
$\liminf_{t',w' \to t,w} \frac{V_m(t',w') - V_m(t,w) - d_1(t'-t) - d_2(w'-w) - \frac{1}{2}(w'-w)'d_3(w'-w)}{|t'-t| + \|w'-w\|^2} \geq 0\}$
V_i is a viscosity solution if and only if $V_i(T, m) = \int_s g_i m(T, ds)$ for any measure m and for all $(t, m) \in [0, T) \times \Delta(\mathcal{S})$,

$$\begin{cases} d_1 + \int \sup_{a_i(.)\in\mathcal{A}_i} H_i(t, s, m, a, d_2, d_3) \leq 0, \ \forall \ d \in D^{1,2,+}_{t,w} V_{i,m}(t,w), \\ d_1 + \int \sup_{a_i(.)\in\mathcal{A}_i} H_i \geq 0, \ \forall \ d \in D^{1,2,-}_{t,w} V_{i,m}(t,w). \end{cases}$$

In particular, if $V_{i,m} \in C^{1,3}$ then
$$\begin{cases} P_i(t) = V_{i,ssm}(t, s(t)) = d_3, \\ Q_i(t) = \sigma V_{i,sssm}(t, s(t)). \end{cases}$$

Proposition 5. *If the first and second order weak derivatives with respect to s of V_m exist then the process $P_i(t) = V_{i,ssm}(t, s(t))$ solves the backward SDE:*

$$dP_i = -\gamma_i dt + \kappa_i dB, \quad P_i(T) = g_{i,ss}(s(T), m(T)) + \int_s \partial_{ss} g_{i,m} m(T, ds),$$

where $\gamma_i = H_{i,ss} + \tilde{\mathbb{E}}\partial_{ss} H_{i,m} + (2b_s + \sigma^2_s)P_i + 2\sigma_s Q_i, \ \kappa_i = \sigma V_{i,sssm} = Q_i$

Proof. We compute explicitly the second order terms. Let $g̃_i(t, s) = V_{i,ssm}(t, s) = v^*_{i,ss}(t, s)$. Applying Ito's formula to $P̃_i(t) = g̃_i(t, s(t))$ yields

$$dP̃_i = [(g̃_i)_t + b(g̃_i)_s + \frac{\sigma^2}{2}(g̃_i)_{ss}]dt + \sigma(g̃_i)_s dB \quad,$$

We identify the coefficient processes $Q̃_i = \sigma(g̃_i)_s = \sigma v^*_{i,sss} = \sigma V_{i,sssm}$. It remains to find the drift coefficient of $P̃_i$.

$$\begin{cases} P̃_i = v^*_{i,ss}(t, s) = V_{i,ssm}(t, s), \\ (g̃_i)_t = v^*_{i,tss}, \\ b(g̃_i)_s = bv^*_{i,sss}, \\ \frac{\sigma^2}{2}(g̃_i)_{ss} = \frac{\sigma^2}{2}v^*_{i,ssss} \end{cases}$$

Summing together one obtains

$$\begin{cases} (g̃_i)_t + b(g̃_i)_s + \frac{\sigma^2}{2}(g̃_i)_{ss} = v^*_{i,tss} + bv^*_{i,sss} + \frac{\sigma^2}{2}v^*_{i,ssss} \end{cases} \tag{18}$$

In order to identify the latter term, one can differentiate twice with respect to s the equation (10) satisfied by the dual function v_i^*.

$$\begin{cases} i \in \mathcal{I}, \\ v_{i,sst}^* + r_{i,ss} + (bv_{i,s}^*)_{ss} + (\frac{\sigma^2}{2}v_{i,ss}^*)_{ss} + \tilde{\mathbb{E}}H_{i,ssm} = 0, \\ v_{i,ss}^*(T,m) = g_{i,ss}(s, m(T,.)) + \tilde{\mathbb{E}}g_{i,ssm}, \\ m(t,.) = \mathbb{P}_{s^a(t)}, \ m(0,.) = m_0(.) \end{cases} \tag{19}$$

$$\begin{cases} (bv_{i,s}^*)_{ss} = (b_s v_{i,s}^* + bv_{i,ss}^*)_s = b_{ss}v_{i,s}^* + 2b_s v_{i,ss}^* + bv_{i,sss}^*, \\ (\frac{\sigma^2}{2}v_{i,ss}^*)_{ss} = [\sigma_s \sigma v_{i,ss}^* + \frac{\sigma^2}{2}v_{i,sss}^*]_s = (\sigma_{ss}\sigma + \sigma_s^2)v_{i,ss}^* + 2\sigma_s \sigma v_{i,sss}^* + \frac{\sigma^2}{2}v_{i,ssss}^* \end{cases}$$

By substitution we obtain

$$\begin{cases} 0 = v_{i,sst}^* + r_{i,ss} + (bv_{i,s}^*)_{ss} + (\frac{\sigma^2}{2}v_{i,ss}^*)_{ss} + \tilde{\mathbb{E}}H_{i,ssm} \\ = v_{i,tss}^* + bv_{i,sss}^* + \frac{\sigma^2}{2}v_{i,ssss}^* \\ + r_{i,ss} + b_{ss}v_{i,s}^* + \sigma_{ss}(\sigma v_{i,ss}^*) + \\ + (2b_s + \sigma_s^2)v_{i,ss}^* + 2\sigma_s(\sigma v_{i,sss}^*) + \tilde{\mathbb{E}}H_{i,ssm} \end{cases}$$

$$\begin{cases} -[v_{i,tss}^* + bv_{i,sss}^* + \frac{\sigma^2}{2}v_{i,ssss}^*] \\ = r_{i,ss} + b_{ss}v_{i,s}^* + \sigma_{ss}\sigma v_{i,ss}^* + \\ + (2b_s + \sigma_s^2)v_{i,ss}^* + 2\sigma_s(\sigma v_{i,sss}^*) + \tilde{\mathbb{E}}H_{i,ssm} \\ = H_{i,ss} + (2b_s + \sigma_s^2)\tilde{P}_i + 2\sigma_s \tilde{Q}_i + \tilde{\mathbb{E}}H_{i,ssm} \end{cases}$$

$$\begin{aligned} d\tilde{P}_i &= [(\tilde{g}_i)_t + b(\tilde{g}_i)_s + \frac{\sigma^2}{2}(\tilde{g}_i)_{ss}]dt + \sigma(\tilde{g}_i)_s dB \\ &= [v_{i,tss}^* + bv_{i,sss}^* + \frac{\sigma^2}{2}v_{i,ssss}^*]dt + (\sigma v_{i,sss}^*)dB \\ &= -[H_{i,ss} + \tilde{\mathbb{E}}H_{i,ssm} + (2b_s + \sigma_s^2)\tilde{P}_i + 2\sigma_s \tilde{Q}_i]dt + \tilde{Q}_i dB \end{aligned} ,$$

where $H_{i,ss} = r_{i,ss} + b_{ss}p_i + \sigma_{ss}q_i$. This completes the proof. $\qquad\square$

3.5. Wiener chaos expansion

The space of square integrable real-valued deterministic functions over $[0, T]$, is denoted as $L^2([0, T], \mathbb{R})$. From Riesz and Fischer's theorem, the space $L^2([0, T], \mathbb{R})$ of square Lebesgue-integrable functions over $[0, T]$ is an (infinite-dimensional) complete metric space. Thus, $L^2([0, T], \mathbb{R})$ is an Hilbert space (with the inner product $\langle f, g \rangle = \int_0^T fg \, dt$).

Lemma 6 ([22], prop. 5.14). *Every Hilbert space, that is not reduced to singleton $\{0\}$, has an orthonormal basis. In particular, $L^2([0, T], \mathbb{R})$ has an orthonormal basis. Let $\{\hat{m}_k(.), k \in \mathbb{Z}_+\}$ be an orthonormal basis in the Hilbert space $L^2([0, T], \mathbb{R})$. Then, the following are equivalent:*

- *$\{\hat{m}_k(.), k \in \mathbb{Z}_+\}$ is an orthonormal basis in the Hilbert space $L^2([0, T], \mathbb{R})$.*
- *$\forall f \in L^2([0, T], \mathbb{R})$, $f(t) = \sum_{k \in \mathbb{Z}_+} \langle f, \hat{m}_k \rangle \hat{m}_k(t) = \sum_{k \in \mathbb{Z}_+} \left(\int_0^T f(t')\hat{m}_k(t')dt' \right) \hat{m}_k(t)$.*
- *One has $\|f\|_{L^2([0,T],\mathbb{R})}^2 = \sum_{k \in \mathbb{Z}_+} |\langle f, \hat{m}_k \rangle|^2 = \sum_{k \in \mathbb{Z}_+} \left(\int_0^T f(t)\hat{m}_k(t) \, dt \right)^2$ for all $f \in L^2([0, T], \mathbb{R})$.*
- *The system $\langle f, \hat{m}_k \rangle = 0 \ \forall k \in \mathbb{Z}_+$ implies that $f = 0$ (identically null function over $[0, T]$).*

See Robinson ([22], prop. 5.14) for a proof of this result.

Example 3. *Orthonormal Basis of $L^2([0,T],\mathbb{R})$:*

- *An important orthogonal basis of $L^2[0,T]$ is the set $\{1, \cos(\frac{2\pi}{T}kt), \sin(\frac{2\pi}{T}kt)\}_{k\geq 1}$. Normalizing by the respective norms in $L^2([0,T],\mathbb{R})$, one gets an orthonormal basis.*
- *One can use the Gram-Schmidt algorithm to construct new bases from more-or-less arbitrary collections of vectors. This is an inductive process which any basis of square integrable functions over $[0,T]$, $\{f_k \mid k \in \mathbb{Z}_+\}$. (i) Since $f_0 \neq 0$, we set $\hat{m}_0 = \frac{f_0}{\|f_0\|_{L^2([0,T],\mathbb{R})}}$ and (ii) then inductively*

$$\hat{f}_{k+1}(t) = f_{k+1}(t) - \sum_{i=0}^{k} \langle f_{k+1}, \hat{m}_i \rangle \hat{m}_i(t),$$

$\hat{f}_{k+1} \neq 0$ because otherwise it can be written as a linear combination of the vector $\hat{m}_0(), \ldots, \hat{m}_k()$. Thus, set $\hat{m}_{k+1} = \frac{\hat{f}_{k+1}}{\|\hat{f}_{k+1}\|_{L^2([0,T],\mathbb{R})}}$ to get a unit vector. (iii) Then the set $\{\hat{m}_k(.) \mid k \in \mathbb{Z}_+\}$ is an orthonormal basis of $L^2([0,T],\mathbb{R})$.

Lemma 7. *From any orthonormal basis $\{\hat{m}_k\}_k$ of $L^2([0,T],\mathbb{R})$, we set $\xi_k = \int_0^T \hat{m}_k(t)dB(t)$ be the stochastic Itô's integral. Then, the random variables $\{\xi_k\}_k$ are identically distributed and Gaussian with zero mean and variance equals to*

$$\int_0^T \hat{m}_k^2(t)dt = \|\hat{m}_k\|_{L^2([0,T],\mathbb{R})}^2 = 1.$$

Furthermore, the standard Brownian motion can be decomposed as follows:

$$
\begin{aligned}
B(t) &= \int_0^t dB(t') = \int_0^T \mathbb{1}_{[0,t]}(t')dB(t') \\
&= \int_0^T \left(\sum_{k \geq 0} \hat{m}_k(t') \langle \mathbb{1}_{[0,t]}(.), \hat{m}_k(.) \rangle \right) dB(t') \\
&= \sum_{k \geq 0} \langle \mathbb{1}_{[0,t]}(.), \hat{m}_k(.) \rangle \int_0^T \hat{m}_k(t')dB(t') \\
&= \sum_{k \geq 0} \left(\int_0^t \hat{m}_k(t')dt' \right) \xi_k.
\end{aligned}
\tag{20}
$$

The expansion

$$B(t) = \sum_{k \geq 0} \xi_k \int_0^t \hat{m}_k(t')dt'$$

converges in the mean-square sense:

$$\mathbb{E}\left[\|B(t) - \sum_{k=0}^{K} \xi_k \int_0^t \hat{m}_k(t')dt'\|^2 \right] \to 0$$

as K goes to infinity, for $t \leq T$.

The mean-square error is

$$\mathbb{E}\left[\| \sum_{k=K+1}^{+\infty} \xi_k \int_0^t \hat{m}_k(t')dt'\|^2 \right] = \sum_{k=K+1}^{+\infty} \left(\int_0^t \hat{m}_k(t')dt' \right)^2 = O(\frac{t}{K}).$$

As a result of the above Lemma, we can view the Ito's process

$$s(t) = s(0) + \int_0^t b(t', s(t'), m(t'), a(t'))dt' + \int_0^t \sigma(t', s(t'), m(t'), a(t'))dB(t'),$$

as a function of t, $s(0)$, w and the set of random variables $\xi = (\xi_k)_k$. We denote the solution of the state equation as $s(t) := s^{s_0, a}(t)$.

Definition 3. *With the one-dimensional Hermite polynomials*

$$\hat{H}_k(t) = \frac{(-1)^k}{k!} e^{\frac{t^2}{2}} \frac{d^k}{dt^k} \left[e^{-\frac{t^2}{2}} \right],$$

the basis polynomials of the Wiener chaos space are defined by

$$\chi_\alpha(\xi) = \sqrt{\alpha!} \prod_i \hat{H}_{\alpha_i}(\xi_i)$$

where $\alpha!$ is the product of all the components' factorials, α denotes the multi-index from the set

$$\{\alpha = (\alpha_i)_i \mid \alpha_i \geq 0, \ \mid |\alpha| = \sum_{i=0}^{+\infty} \alpha_i < +\infty\}.$$

Note that, for a multi-index α such that $|\alpha| < +\infty$, the number of the terms in the product $\prod_i \hat{H}_{\alpha_i}$ is finite. From Definition 3 and the properties of the Hermite polynomials one directly shows the orthogonality of the above polynomial basis, which are often referred to as Wick polynomials. If the controlled state process $s \in L^2(\Omega \times [0, T], \mathbb{R})$ then one can decompose it as

$$s(t) = \sum_\alpha \langle s(t), \chi_\alpha(\xi) \rangle \chi_\alpha(\xi) = \sum_\alpha E[s(t)\chi_\alpha(\xi)]\chi_\alpha(\xi),$$

where $\langle x, y \rangle = E[xy]$, when x and y are random variables. Particularly the coefficient function of order zero, $s_0(t) = E[s(t)\chi_0(\xi)]$, obtains a special meaning. It coincides with the expectation of the process $x(t)$, as the basis polynomial of order zero is identically one.

This is summarized in the following well-known result:

Theorem 3.1 (Cameron and Martin [23]). *Assume that the state process $s(t)$ is adapted to the filtration generated by the Wiener chaos $\{\chi_\alpha(\xi)\}_\alpha$, and is in $L^2(\Omega \times [0, T], \mathbb{R})$ i.e., it satisfies the integrability condition $E[\int_0^T |s(t)|^2 dt] < +\infty$. Then, $s(t)$ can be expanded in $[0, T]$ as*

$$s(t, \xi) = \sum_\alpha s_\alpha(t)\chi_\alpha(\xi),$$

where the deterministic coefficient functions $s_\alpha(t) = E[x(t)\chi_\alpha(\xi)]$ can be interpreted as projections of the process $s(t)$ onto the corresponding chaos basis.

Note that the variance starts with index 1 and not from 0. The polynomial chaos (PC) or Wiener chaos framework was developed by Norbert Wiener [24, 25] and generalized by Cameron and Martin later on [23]. The Fourier-Hermite series of $s(t)$ is often called the Wiener chaos expansion (WCE).

Theorem 3.2 (Error estimates, Theorem 2.1 in [26]). *If there exists k such that $\|\partial_\xi^k s(t, \xi)\|^2 < +\infty$, then the following error estimates holds:*

$$\|s(t) - \sum_{|\alpha| \leq K} s_\alpha(t) \chi_\alpha(\xi)\|^2 \leq \frac{\|\partial_\xi^k s(t, \xi)\|^2}{\prod_{i=0}^{k-1}(K - i + 1)} =: \epsilon_K.$$

It is important to notice that this error ϵ_K is much better than Monte-Carlo sampling (in the order of $O\left(\frac{1}{\sqrt{K}}\right)$ for K samples) when the random processes of the basis capture the state process s. This provides a more efficient way to solve mean-field-type game problem. However, the choice of a basis is crucial and the chosen elements in the basis need to be sparse enough in order to reduce the curse of dimensionality.

We now reformulate the mean-field-type game problems using the PC framework above. We replace the state process s and the control process by their respective PC expansions. Therefore, s is determined by its coefficient functions $(s_\alpha)_\alpha$. We can compute the cost as

$$\mathbb{E}g_i(s(T), m(T)) = \hat{g}_i((s_\alpha(T))_\alpha) \tag{21}$$

$$\mathbb{E}r_i(t, s(t), m(t), a(t)) = \hat{r}_i(t, (s_\alpha(t))_\alpha, (a_\alpha(t))_\alpha).$$

and the state dynamics coefficients are

$$s_\alpha(t) = s_\alpha(0)\mathbb{1}_{\alpha=0} + \langle \chi_\alpha(.), \int_0^t b\,dt' \rangle + \langle \chi_\alpha(.), \int_0^t \sigma\,dB(t') \rangle.$$

Since the Lebesgue integral term is $\langle \chi_\alpha(.), \int_0^t b\,dt' \rangle = \int_0^t b_\alpha\,dt'$, the Ito's integral term becomes

$$\langle \chi_\alpha(.), \int_0^t \sigma\,dB(t') \rangle = \int_0^t \sigma \partial_{t'}[\chi_\alpha]\,dt'.$$

Using Hermite polynomials, we know that

$$\partial_{t'}[\chi_\alpha] = \sum_{k=1}^\infty \sqrt{\alpha_k} m_k(t') \chi_{\hat{\alpha}^k}(\xi)$$

where $\hat{\alpha}^k = (\alpha_1, \ldots, \alpha_{k-1}, \alpha_k - 1, \alpha_{k+1}, \ldots)$, i.e., $\hat{\alpha}_i^k = (\alpha_k - 1)\mathbb{1}_{i=k} + \alpha_i \mathbb{1}_{i \neq k}$. It turns out that s_α solves the ordinary differential system:

$$\dot{s}_\alpha(t) = b_\alpha(t) + \sum_{k=1}^\infty \sqrt{\alpha_k} m_k(t) E[\sigma(t, .) \chi_{\hat{\alpha}^k}(\xi)]$$

with initial condition $s(0)\mathbb{1}_{\alpha=0}$.

Proposition 6. *The best response problem (1) becomes the following: For each decision-maker i, given the coefficient strategies of the others $(\{a_{i,\alpha}\}_\alpha, \ j \neq i)$ the best response solves*

$$\begin{cases} \sup_{(a_{i,\alpha}(\cdot))_\alpha} \mathbb{E}\hat{\mathcal{R}}_{i,T}, \\ subject\ to \\ \dot{s}_\alpha(t) = b_\alpha(t) + \sum_{k=1}^\infty \sqrt{\alpha_k}\hat{m}_k(t) E[\sigma(t, .) \chi_{\hat{\alpha}^k}] \\ s_\alpha(0) = s(0)\mathbb{1}_{\alpha=0}, \end{cases} \tag{22}$$

where $\hat{\mathcal{R}}_{i,T} = \hat{g}_i + \int_0^T \hat{r}_i\,dt.$

Proof. We first rewrite the instantaneous payoff functions in terms Wiener chaos. We replace the state process $s(t)$ by its decomposition $\sum_\alpha s_\alpha(t)\chi_\alpha(\xi)$.

$$\mathbb{E}r_i(t, s(t), m(t), a(t)) =: \hat{r}_i(t, \{s_\alpha(t)\}_\alpha, \{a_\alpha(t)\}_\alpha) \tag{23}$$

In addition, the coefficients dynamics $s_\alpha(t)$ can be easily obtained from the state dynamics and s_α solves a standard differential system:

$$\dot{s}_\alpha(t) = b_\alpha(t) + \sum_{k=1}^\infty \sqrt{\alpha_k}\hat{m}_k(t)\mathbb{E}[\sigma(t,.)\chi_{\hat{a}^k}(\xi)].$$

with initial condition $s(0)\mathbb{1}_{\alpha=0}$. Collecting together, we obtain a standard dynamic optimization problem subject to coefficient dynamics with multiple index. This completes the proof. $\quad\square$

As we can observe, problem (22) is now a standard differential game with standard deterministic state dynamics. Note that, in order to get these transformations, the underlying problem must be included in the space of square integrable random process Hilbert space. This implies, that the variance of the problem must be finite.

4. Aggregative mean-field-type games

In this section, equilibrium system for mean-field-type games with aggregative structures are presented.

4.1. Aggregative structures

In this subsection we take b, σ, r_i, g_i as functions of m^a only through aggregative terms $\int \phi_.(s)m^a(t, ds) = \mathbb{E}[\phi_.(s^a(t))]$, and $\int \phi_.(a')\mathbb{P}_a(t, da') = \mathbb{E}[\phi_.(a(t))]$. In that case the Gâteaux differentiation with respect to m can be reduced to a finite-dimensional differentiation.

$$\begin{cases} \mathcal{R}_{i,T} = \left[\int_0^T r_i(t, s^a(t), \mathbb{E}[\phi_{r_i}(s^a(t))], a(t), \mathbb{E}[\phi_{r_i}(a(t))])\, dt + g_i(s^a(T), \mathbb{E}[\phi_{g_i}(s^a(t))])\right], \\ \sup_{a_i \in \mathcal{U}_i} \mathbb{E}\mathcal{R}_{i,T} \text{ subject to} \\ ds^a(t) = b(t, s^a(t), \mathbb{E}[\phi_b(s^a(t))], a(t), \mathbb{E}[\phi_b(a(t))])\, dt + \sigma(t, s^a(t), \mathbb{E}[\phi_\sigma(s^a(t))], a(t), \mathbb{E}[\phi_\sigma(a(t))])dB(t), \\ s^a(0) = s_0. \end{cases} \tag{24}$$

If b, σ, r_i are only functions of $(t, s^a, \mathbb{E}[\phi_.(s^a(t))], a, \mathbb{E}[\phi_.(s(t))])$, then the partial derivatives of the dual functions $p_i = V_{sm}(t, s^a(t))$ and $q_i = \sigma V_{ssm}(t, s^a(t))$ solve backward SDE system:

$$\begin{cases} dp_i = \{-H_{i,s} - \phi_{r_i,s}\mathbb{E}[r_{i,y}] - \phi_{b,s}\mathbb{E}[b_y p_i] - \phi_{\sigma,s}\mathbb{E}[\sigma_y q_i]\} + q_i dB, \\ p_i(T) = g_{i,s}(s^a(T), \mathbb{E}[\phi_{g_i}(s^a(T))]) + \phi_{g_i,s}(T)\mathbb{E}[g_{i,y}(s^a(T), \mathbb{E}[\phi_{g_i}(s^a(T))])] \end{cases} \tag{25}$$

where $H_{i,s} = r_{i,s} + b_s.p_i + \sigma_s q_i$.

Note that the aggregative structure in the form $\tilde{E}[\phi(S(t), \tilde{S}(t))] = \int_w \phi(S, w)m(t, dw)$ which is a random variable, is already included in the cases $\phi(t, s, m)$ discussed in the previous section.

4.2. Moment-based mean-field-type games

We take b, σ, r_i as functions of the $k-$th moment of the state $\int s^k m^a(t, ds) = \mathbb{E}[(s^a(t))^k]$, and the $l-$moment of the control action of decision-maker i $\int (a'_i)^l \mathbb{P}_{a_i}(t, da'_i) = \mathbb{E}[a^l_i(t)]$. Thus, aggregative function is $\phi.(y) = y^k$. Then, the derivative of ϕ is $\phi_{.,y} = ky^{k-1}$. For $k > 1$, the dynamics yields

$$\begin{cases} dp_i = -\{H_{i,s} + ks^{k-1}\mathbb{E}H_{i,y}\}dt + q_i dB, \\ p_i(T) = g_{i,s}(s^a(T), \mathbb{E}[(s^a(T))^k]) + ks^{k-1}(T)\mathbb{E}[g_{i,y}(s^a(T), \mathbb{E}[(s^a(T))^k])] \end{cases} \tag{26}$$

The existence of solution is not immediate. It requires the $(k-1)-$moment estimates of the state for $k > 1$.

4.3. First moment mean-field-type games

We take b, σ, r_i as functions of the first moments $\int s m^a(t, ds) = \mathbb{E}[s^a(t)]$, and $\int a' \mathbb{P}_a(t, da') = \mathbb{E}[a(t)]$. Then the aggregative functions are reduced to the identity function $\phi.(y) = y$. Then, the derivative of ϕ is $\phi_{.,y} = 1$. Hence, the first order risk-neutral adjoint processes $p_i(t) = V_{i,sm}(t, s^a(t))$ and $q_i(t) = \sigma V_{i,ssm}(t, s^a(t))$ solve a simpler backward SDE system and (25) reduces to

$$dp_i = \{-H_{i,s} - \mathbb{E}[H_{i,y}(t, s^a, \mathbb{E}[s^a], p_i, q_i)]\} + q_i dB$$
$$p_i(T) = g_{i,s}(s^a(T), \mathbb{E}[s^a(T)]) + \mathbb{E}[g_{i,y}(s^a(T), \mathbb{E}[s^a(T)])].$$

However, the optimal control strategy equation is modified because of the presence of $E[a(t)]$ In the convex control set case, one obtains the variational inequality $(H_{i,a} + \mathbb{E}H_{i,\bar{a}})(a_i - a'_i) \geq 0$ where $\mathbb{H}_{i,\bar{a}_i}$ denotes the Hamiltonian derivative with respect to the component $\bar{a}_i = \mathbb{E}[a_i(t)]$.

5. Partial observation

This section examines situations in which the state is partially observed.

$$\begin{cases} \mathcal{R}_{i,T}(m_0, a) = \int_0^T r_i(t, s^{s_0,a}(t), m^{m_0,a}(t), a(t)) \, dt + g_i(s^{s_0,a}(T), m_1^{m_0,a}(T)), \\ \sup_{a_i \in \mathcal{A}_i} \mathbb{E}^a \mathcal{R}_{i,T}(m_0, a) \text{ subject to} \\ ds^{s_0,a}(t) = b(t, s^{s_0,a}(t), m^{m_0,a}(t), a(t)) \, dt \\ \qquad + \sigma(t, s^{s_0,a}(t), m^{m_0,a}(t), a(t)) dB(t) \\ \qquad + \sigma_o(t, s^{s_0,a}(t), m^{m_0,a}(t), a(t)) dB_o(t), \\ s^a(0) = s_0, \\ dy^a(t) = b_o(t, s^{s_0,a}(t), m^{m_0,a}(t), a(t)) \, dt + dB_o(t), \; y^a(0) = 0, \\ m^{m_0,a}(t, .) := \mathbb{P}_{s^{s_0,a}(t)}. \end{cases} \tag{27}$$

where $\mathbb{E}^a$ denotes the the expectation with respect to the probability space $(\Omega, \mathbb{F}, \{\mathcal{F}_t\}_t, \mathbb{P}^a)$ where $\mathbb{P}^a$ is defined below. Under partial state observation an admissible strategy a_i is a random process $a_i : [0, T] \times \Omega \to A_i$ that is adapted to $\mathcal{F}_t^y = \sigma(y(t'), \; t' \leq t)$ and $\mathbb{E} \int_0^T |a_i(t, .)|^\alpha dt < \infty$. The set of admissible strategies is denoted by $\hat{\mathcal{A}}_i$. Introduce the density process $\rho^a(t) = e^{\int_0^t b_o(t')dy^a(t') - \frac{1}{2}\int_0^t |b_o(t')|^2 dt'}$. Then, $\rho^a(t)$ solves the forward SDE $d\rho^a = \rho^a b_o dy^a$, $\rho^a(0) = 1$. By Girsanov transform, $d\mathbb{P}^a = \rho^a d\mathbb{P}$, the partial observation problem is transformed into a full observation problem with respect to $\mathbb{P}$ and with a new

state $(\hat{s}^a, \rho^a)$.

$$\begin{cases} \sup_{a_i \in \hat{\mathcal{A}}_i} \quad \mathbb{E}[\int_0^T \rho^a(t) r_i(t, \hat{s}^{s_0,a}(t), m^{m_0,a}(t), a(t)) \, dt + \rho^a(T) g_i(\hat{s}^{s_0,a}(T), m^{m_0,a}(T))] \\ \text{subject to} \\ d\hat{s}^a(t) = b \, dt + \sigma dB(t) + \sigma_o[dy^a(t) - b_o \, dt], \\ = [b - \sigma_o b_o] \, dt + \sigma dB + \sigma_o dy^a(t), \\ \hat{s}^a(0) = s_0, \\ d\rho^a = \rho^a b_o dy^a, \ \rho^a(0) = 1. \end{cases} \tag{28}$$

This problem is similar to (1) but with a new state $(\hat{s}^{s_0,a}, \rho^a)$. One obtains an augmented state $(\hat{s}, \rho)$ and the infinite-dimensional DPP can be directly applied by considering the probability measure $m(t, d\hat{s}d\rho) = \mathbb{P}_{\hat{s}^a(t), \rho^a(t)}$. Similarly the stochastic maximum principle can be applied with the state $(\hat{s}, \rho)$.

$$d \begin{pmatrix} \hat{s}^a \\ \rho \end{pmatrix} = \begin{pmatrix} b - \sigma_o b_o \\ 0 \end{pmatrix} dt + \begin{pmatrix} \sigma & \sigma_o \\ 0 & \rho^a b_o \end{pmatrix} \begin{pmatrix} dB \\ dy^a \end{pmatrix}.$$

The new Hamiltonian for decision-maker i is

$$H_i(t, s, \rho, m, (p_1, p_2), q) = \rho r_i + [b - \sigma_o b_o] p_1 + trace(\Gamma' q),$$

where $\Gamma := \begin{pmatrix} \sigma & \sigma_o \\ 0 & \rho^a b_o \end{pmatrix}$, and Γ' is the transpose of the matrix Γ, and $q := \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$. For the optimal control action, however, it should be conditioned on the observation filtration $\mathcal{F}_t^y$. A necessary condition in the smooth case yields

$$\mathbb{E}\left[\delta H_i + \frac{1}{2} P_i [\delta\sigma]^2 \mid \mathcal{F}_t^y\right] \leq 0.$$

Note that the methodology extends to the case of individual observation per decision-maker i using the filtration $\mathcal{F}_{i,t}^{y_i}$.

6. Extensions

In this section focuses on recent development and or extensions of mean-field-type games.

6.1. Discrete time games

Consider a mean-field-type game setup with the following data:

$$\begin{cases} \text{Time step:} & t \in \{0, 1, \ldots, T-1\} \\ \text{Set of decision-makers:} & \mathcal{I} \\ \text{Initial state :} & s_0 \sim m_0^s \\ \text{Stochastic state dynamics:} & s(t+1) \sim q_{t+1}(. \mid s(t), m^s(t), m^a(t), a(t)) \\ \text{Instant payoff of } i : & r_i(t, s(t), m^s(t), m^a(t), a(t)) \\ \text{Terminal payoff of } i : & g_i(s(T), m^s(T)) \end{cases}$$

Here the time space is a discrete set and denoted by $\{0, 1, \ldots, T-1\}$, $T \geq 1$ is the length of the horizon, t denotes a time step, s is the state process of the system, m_0^s is the initial distribution of states.

A decision-maker is denoted by i, has an action space A_i. The system state $s(t)$ is stochastic and its probability transition from t to $t + 1$ is given by $q_{t+1} : \mathcal{S} \times \Delta(\mathcal{S}) \times \prod_i \Delta(A_i) \times \prod_i A_i \to \Delta(\mathcal{S})$. Denote by $m^s(t)$ the distribution of state and by $m^a(t)$ the distribution of actions at time t. Decision-maker i's cumulative payoff is

$$\mathcal{R}_{i,T}(m_0^s, a) = \sum_{t=0}^{T-1} r_i(t, s(t), m^s(t), m^a(t), a(t)) + g_i(s(T), m^s(T)),$$

The risk-neutral best response problem of i is

$$\begin{cases} \sup_{a_i} \mathbb{E}\mathcal{R}_{i,T}(m_0^s, a) \text{ subject to} \\ s(t + 1) \sim q_{t+1}(.|\, s(t), m^s(t), m^a(t), a(t)), \\ m^s(t, .) = \mathbb{P}_{s(t)}, \ s_0 \sim m_0^s, \end{cases} \tag{29}$$

Let express the expected payoff in terms of the measure $m(t, .)$.

$$Er_i(t, s(t), m^s(t), m^a(t), a(t)) = \int r_i(t, \bar{s}, m^s(t), m^a(t), a(t)) m^s(t, d\bar{s}) = \hat{r}_i(t, m(t), a(t))$$

where $\hat{r}_i$ depends only on the measure $m(t)$ and the strategy profile $a(t)$. Similarly one can rewrite the expected value of the terminal payoff as

$$Eg_i(s(T), m^s(T)) = \int g_i(\bar{s}, m^s(T)) m^s(T, d\bar{s}) = \hat{g}_i(m^s(T)).$$

Proposition 7. *On the space of measures, one has a deterministic dynamic game problem over multiple stages. Therefore a dynamic programming principle (DPP) holds:*

$$\begin{cases} V_i(t, m^s(t)) = \sup_{a_i'} \left\{ \hat{r}_i(t, m^s(t), a_i'(t), a_{-i}(t)) + V_i(t + 1, m^s(t + 1)) \right\} \\ m^s(t + 1, ds') = \int_s q_{t+1}(ds'|\, s, m^s(t), m^a(t), a(t)) m^s(t, ds) \end{cases} \tag{30}$$

As we can see the best-response strategy may be dependent on the state, the mean-field m^s, which is referred to as (state-and-mean-field) feedback strategy. Therefore, $m^{a(t)}$ can be expressed as a function of $(s(t), m^{s(t)}(t, .))$. Thus, the payoff $\hat{r}_i(t, .)$ can be expressed as a function $(m(t, .), a(t))$.

Proposition 8. *Suppose a sequence of real-valued function $V_i(t, .)$, $t \leq T$ defined on the set of probability measures over $\mathcal{S}$ is satisfying the DPP relation above. Then $V_i(t, m)$ is the value function on $\Delta(\mathcal{S})$ starting from $m(t) = m$. Moreover if the supremum is achieved for some action $a_i^*(., m)$, then the best response strategy is in (state-and-mean-field) feedback form, and the payoff value is*

$$R_i(a^*) = V_i(0, m_0).$$

Proposition 8 provides a sufficiency condition for best-response strategies in terms of $(s, m^s(t))$. The proof is immediate and follows from the verification theorem of DPP in deterministic dynamic games.

Particular cases of interest

- Finite state space: Suppose that the state space and the action spaces are nonempty and finite. Let the state transition be

$$\mathbb{P}(s(t+1) = s' \mid s(t), m^s(t), m^a(t), a(t)) = q_{t+1}(s' \mid s(t), m^s(t), m^a(t), a(t)),$$

DPP becomes

$$\begin{cases} V_i(t, m^s(t)) = \sup_{a'_i} \left\{ \hat{r}_i(t, m(t), a'_{it}, a_{-i,t}) + V_i(t+1, m^s(t+1)) \right\} \\ m^s(t+1, s') = \sum_{s \in S} q_{t+1}(s' \mid s, m^s(t), m^a(t), a(t)) m^s(t, s) \end{cases}$$

As for classical polymatrix games, a pure mean-field equilibrium may not exist in general in mean-field-type games with finite actions. However mixed extension can be adopted. By extending the action space to the set of probability measures on A_i and the functions $\hat{r}_{it}, \hat{g}_{iT} q_{t+1}$ to the corresponding mixed extensions, one gets the existence of mean-field equilibria in behavioral (mixed) strategies for payoffs that are independent of m^a.

- Continuous state space: Consider the interactive state dynamics in discrete time

$$s(t+1) = s(t) + b(t+1, s(t), m^s(t), m^a(t), a(t), \eta_{t+1})$$

where η is random process. The transition kernel of $s(t+1)$ given $s(t), m^s(t), m^a(t), a(t)$ is

$$q_{t+1}(ds' \mid s(t), m^s(t), m^a(t), a(t)) = \int_\eta P(ds' \ni s(t) + b_{t+1}(s(t), m^s(t), m^a(t), a(t), \eta)) \mathbb{P}_{\eta_{t+1}}(d\eta)$$

where $\mathbb{P}_{\eta_{t+1}}(d\eta)$ denotes the probability distribution of η_{t+1}. The Bellman equation 7 applies this case.

- Mean-field free case: If $r_i(s, a, m^s, m^a) = r_i(s, a)$ and $g_i(s, m^s) = g_i(s)$ for every decision-maker i then

$$\hat{r}_i(m, a) = \int_s r_i(s, a) m^s(t, ds).$$

There exists a function v_i such that

$$V_i(t, m^s(t)) = \langle v_i(t, .), m^s(t, .) \rangle = \int_s v_i(t, s) m^s(t, ds),$$

$v_i(t, s)$ is a mean-field free equilibrium payoff function of i. In that case, the dynamic programming reduces to $v_i(t, s) = \sup_{a'_{it}} H_i(t, s, a'_{it}, a_{-i,t})$,

$$H_i = r_i(t, s, a'_{it}, a_{-i,t}) + \int_{s'} v_i(t+1, s') q_{t+1}(ds' \mid s(t), a'_i(t), a_{-i}(t)),$$

which is the classical Bellman-Shapley equilibrium system. Note that in this case $v_i(t, s) = \partial_{m(t,s)} V(t, m)$.

Finite state continuous time games

Discrete state (countable or finite) games of mean-field type can be analyzed using the above methodologies. State process becomes a continuous time interactive Markovian decision process in which mean-field terms such as distribution of states and distribution of control actions are involved in it.

6.2. Risk-sensitive games of mean-field type

Each decision-maker has a risk-sensitivity index $\theta_i \in \mathbb{R}$. When θ_i vanishes, one gets the risk-neutral setup discussed above. Consider the risk-sensitive best response of i as

$$
\begin{cases}
\sup_{a_i \in \mathcal{A}_i} \frac{1}{\theta_i} \log \left(\mathbb{E} e^{\theta_i \mathcal{R}_{i,T}} \right) \text{ subject to} \\
s(t) = s_0 + \int_0^t b \, dt' + \int_0^t \sigma \, dB(t'), \ t > 0 \\
s(t) = s^{s_0,a}(t), \ s(0) = s_0 \sim m_0 \\
m(t,.) = m^{m_0,a}(t,.) = \mathbb{P}_{s^{s_0,a}(t)}
\end{cases}
\tag{31}
$$

Introduce the following augmented state (s,z) such that $z_i(0) = 0$, and $dz_i(t) = r_i(t, s^a(t), m^a(t), a(t)) \, dt$. Then,

$$
e^{\theta_i \mathcal{R}_{i,T}} = e^{\theta_i z_i(T) + \theta_i g_i(s^a(T), m^a(T))}.
$$

Thus, we obtain the following mean-field-type control problem for decision-maker i:

$$
\begin{cases}
\sup_{a_i} \mathbb{E} e^{\theta_i[z_i(T) + g_i(s^a(T), m^a(T))]} \\
z_i^a(t) = \int_0^t r_i \, dt', \\
z_i^a(0) = 0, \\
s(t) = s_0 + \int_0^t b \, dt' + \int_0^t \sigma \, dB(t'), \ t > 0 \\
s^a(0) = s_0.
\end{cases}
\tag{32}
$$

Let us introduce $\mu(t, ds\,dz) = \mathbb{P}_{s^a(t), z^a(t)}$, the probability measure associated with the joint process $(s^a(t), z^a(t))$. Generically, the measure μ solves the Fokker-Planck-Kolmogorov forward equation in the distributional (weak) sense. The term $\mathbb{E} e^{\theta_i \mathcal{R}_{i,T}}$ can be rewritten in a deterministic manner as

$$
\int \mu(T, ds\,dz) \, e^{\theta_i(z_i + g_i(s, \int_{\bar{z}} \mu(T,.,d\bar{z})))} \,.
$$

This is a terminal cost in the sense that it is evaluated only at $\mu(T,.)$. μ is an infinite dimensional quantity that serves as a state in the problem. The advantage now is that $\mu(.)$ is a deterministic quantity. Since there is no running cost, one can write directly the HJB equation using classical calculus of variations for

$$
\hat{V}_i^\theta(t,\mu) = \sup_{a_i \in \mathcal{A}_i} \int \mu(T, ds\,dz) \, e^{\theta_i(z_i + g_i(s, \int_{\bar{z}} \mu(T,.,d\bar{z})))},
$$

starting from $\mu(t,.) = \mu$ at time t : The risk-sensitive best-value is $V_i^\theta(t,\mu) := \frac{1}{\theta_i} \log \hat{V}_i^\theta(t,\mu)$.

$$
\mu_t = -\partial_s[b\mu] - \text{div}_z(f\mu) + \frac{1}{2}\partial_{ss}(\sigma'\sigma\mu),
\tag{33}
$$

with initial distribution $\mu(0, dx, dz) = m_0(dx)\delta_0(dz)$ and σ' is the transpose of σ.

The risk-sensitive HJB system yields

$$
\begin{aligned}
0 = {} & \hat{V}_{i,t}^\theta(t,\mu) + \int H_i^\theta \mu(t, ds\,dz), \\
& \hat{V}_i^\theta(T,\mu) = \int \mu(T, ds\,dz) \, e^{\theta_i(z_i + g_i(s, \int_{\bar{z}} \mu(T,.,d\bar{z})))}, \\
& i \in \mathcal{I}
\end{aligned}
\tag{34}
$$

where $H_i^\theta(t, s, z, \mu, \hat{V}_{i,z\mu}, \hat{V}_{i,ss\mu}) = \sup_{a_i} b\partial_s \hat{V}_{i,\mu} + \langle r, \partial_z \hat{V}_{i,\mu}\rangle + \frac{\sigma^2}{2}\partial_{ss}\hat{V}_{i,\mu}$. Note that

$$
\begin{aligned}
\hat{V}_i^\theta(T, \mu) &= \int \mu(T, dsdz)\, e^{\theta_i(z_i + g_i(s, \int_{\bar{z}} \mu(T,.,d\bar{z})))} \\
&= \int_{s,z_i} e^{\theta_i(z_i + g_i(s, \int_{\bar{z}_i} \int_{\bar{z}_{-i}} \mu(T,.,d\bar{z})))} \int_{z_{-i}} \mu(T, dsdz) \\
&= \int_{s,z_i} e^{\theta_i(z_i + g_i(s, \int_{\bar{z}_i} \mu_i(T,.,d\bar{z}_i)))} \mu_i(T, dsdz_i) = \hat{V}_i^\theta(T, \mu_i),
\end{aligned}
\tag{35}
$$

where $\mu_i(t, dsdz_i) = \int_{z_{-i}} \mu(t, dsdz_i dz_{-i})$, is the marginal of μ with respect to (s, z_i), and $\hat{V}_i^\theta(t, \mu) = \tilde{V}_i(t, \mu_i)$, The above calculation shows that the value function $\frac{1}{\theta_i}\log \hat{V}_i^\theta$ depends only on the measure μ_i

Proposition 9. *If there exists a function $\tilde{V}$ with*

$$
\tilde{V}_{i,t}, \tilde{V}_{i,s\mu_i}, \tilde{V}_{i,ss\mu_i}, \tilde{V}_{i,zz\mu_i},
$$

satisfying

$$
\begin{aligned}
0 &= \tilde{V}_{i,t} + \int H_i^\theta \mu_i(t, dsdz_i), \\
\tilde{V}_i(T, \mu_i) &= \int \mu_i(T, dsdz_i)\, e^{\theta_i(z_i + g_i(s, \int_{\bar{z}} \mu_i(T,.,d\bar{z})))}, \\
\mu_i(t, dsdz_i) &= \int_{z_{-i}} \mu(t, dsdz_i dz_{-i}), \\
\mu_t &= -\partial_s[b\mu] - div_z(r\mu) + \frac{1}{2}\partial_{ss}(\sigma^2\mu), \\
\mu(0, dsdz) &= m_0(ds)\delta_0(dz), \\
i &\in \mathcal{I},
\end{aligned}
\tag{36}
$$

where the integrand Hamiltonian is

$$
\begin{aligned}
H_i^\theta(t, s, z, \mu_i, \tilde{V}_{i,s\mu_i}, \tilde{V}_{i,z_i\mu_i}, \tilde{V}_{i,ss\mu_i}) &= \sup_{a_i}[b\tilde{V}_{i,s\mu_i} + r_i\tilde{V}_{i,z_i\mu_i} + \frac{\sigma^2}{2}\tilde{V}_{i,ss\mu_i}] \\
= \tilde{V}_{i,z_i\mu_i}\sup_{a_i}[r_i + b\frac{\tilde{V}_{i,s\mu_i}}{\tilde{V}_{i,z_i\mu_i}} + \frac{\sigma^2}{2}\frac{\tilde{V}_{i,ss\mu_i}}{\tilde{V}_{i,z_i\mu_i}}] &= \tilde{V}_{i,z_i\mu_i}H_i(t, s, z, \mu_i, \frac{\tilde{V}_{i,s\mu_i}}{\tilde{V}_{i,z_i\mu_i}}, \frac{\tilde{V}_{i,ss\mu_i}}{\tilde{V}_{i,z_i\mu_i}})
\end{aligned}
$$

then $\frac{1}{\theta_i}\log \tilde{V}_i(0, \mu)$ is an equilibrium payoff for decision-maker i and the best-response strategy a_i minimizes H_i^θ given the other decision-makers' strategies a_{-i}.

Proof. Apply the recipe of Proposition 1 with the augmented state (s, z) and the measure μ $\qquad\square$

From the relation

$$
V_i^\theta(t, \mu) := \frac{1}{\theta_i}\log \hat{V}_i^\theta(t, \mu) \implies e^{\theta_i V_i^\theta(t,\mu)} = \hat{V}_i^\theta(t, \mu),
\tag{37}
$$

It follows that

$$\hat{V}^{\theta}_{i,\mu}(t,\mu) = \theta_i V^{\theta}_{i,\mu}(t,\mu) e^{\theta_i V^{\theta}_i(t,\mu)} = \theta_i V^{\theta}_{i,\mu}(t,\mu)\hat{V}^{\theta}_i(t,\mu),$$
$$\hat{V}^{\theta}_{i,s\mu}(t,\mu) = \theta_i V^{\theta}_{i,s\mu}(t,\mu)\hat{V}^{\theta}_i(t,\mu),$$
$$\hat{V}^{\theta}_{i,z\mu}(t,\mu) = \theta_i V^{\theta}_{i,z\mu}(t,\mu)\hat{V}^{\theta}_i(t,\mu),$$
$$\hat{V}^{\theta}_{i,ss\mu}(t,\mu) = \theta_i V^{\theta}_{i,ss\mu}(t,\mu)\hat{V}^{\theta}_i(t,\mu),$$
$$\hat{V}^{\theta}_{i,t}(t,\mu) = \theta_i V^{\theta}_{i,t}(t,\mu)\hat{V}^{\theta}_i(t,\mu).$$

Thus,

$$H^{\theta}_i(t,s,z,\mu_i,\hat{V}_{i,s\mu_i},\hat{V}_{i,z_i\mu_i},\hat{V}_{i,ss\mu_i}) = \hat{V}_{i,z_i\mu_i} H_i(t,s,z,\mu_i,\frac{\hat{V}_{i,s\mu_i}}{\hat{V}_{i,z_i\mu_i}},\frac{\hat{V}_{i,ss\mu_i}}{\hat{V}_{i,z_i\mu_i}})$$
$$H_i(t,s,z,\mu_i,\frac{V^{\theta}_{i,s\mu_i}}{V^{\theta}_{i,z_i\mu_i}},\frac{V^{\theta}_{i,ss\mu_i}}{V^{\theta}_{i,z_i\mu_i}})$$

$$0 = \theta_i V^{\theta}_{i,t}(t,\mu)\hat{V}^{\theta}_i(t,\mu) + \int \theta_i V^{\theta}_{i,z\mu}\hat{V}^{\theta}_i(t,\mu).H_i(t,s,z,\mu_i,\frac{V^{\theta}_{i,s\mu_i}}{V^{\theta}_{i,z_i\mu_i}},\frac{V^{\theta}_{i,ss\mu_i}}{V^{\theta}_{i,z_i\mu_i}})\mu_i(t,dsdz_i),$$

$$V^{\theta}_i(T,\mu_i) = \frac{1}{\theta_i}\log\left[\int \mu_i(T,dsdz_i)\, e^{\theta_i(z_i+g_i(s,\int_{\tilde{z}}\mu_i(T,.,d\tilde{z})))}\right],$$

By symplifying by $\theta_i\hat{V}^{\theta}_i(t,\mu) \neq 0$, we obtain that the risk-sensitive best-response payoff $V^{\theta}_i(t,\mu)$ solves the functional PDE given by

$$0 = V^{\theta}_{i,t}(t,\mu) + \int V^{\theta}_{i,z\mu}.H_i(t,s,z,\mu_i,\frac{V^{\theta}_{i,s\mu_i}}{V^{\theta}_{i,z_i\mu_i}},\frac{V^{\theta}_{i,ss\mu_i}}{V^{\theta}_{i,z_i\mu_i}})\mu_i(t,dsdz_i),$$

$$V^{\theta}_i(T,\mu_i) = \frac{1}{\theta_i}\log\left[\int \mu_i(T,dsdz_i)\, e^{\theta_i(z_i+g_i(s,\int_{\tilde{z}}\mu_i(T,.,d\tilde{z})))}\right],$$

$$\mu_i(t,dsdz_i) = \int_{z_{-i}}\mu(t,dsdz_idz_{-i}),$$

$$\mu_t = -\partial_s[b\mu] - \text{div}_z(r\mu) + \frac{1}{2}\partial_{ss}(\sigma^2\mu),$$

$$\mu(0,dsdz) = m_0(ds)\delta_0(dz),$$

$$i \in \mathcal{I},\tag{38}$$

Let $\tilde{v}^*_i(t,s,z_i) := \tilde{V}_{i,\mu_i}(\mu_i)(t,s,z_i)$. It follows that $\tilde{v}^*_i(t,s,z_i)$ solves the following PDE system:

$$0 = \tilde{v}^*_{i,t} + \tilde{v}^*_{i,z_i} H_i(t,s,z_i,\mu_i,\frac{\tilde{v}^*_{i,s}}{\tilde{v}^*_{i,z_i}},\frac{\tilde{v}^*_{i,ss}}{\tilde{v}^*_{i,z_i}}) + \int \tilde{v}^*_{i,\tilde{z}_i} H_{i,\mu_i}\mu_i(t,d\tilde{s},d\tilde{z}_i).$$

The terminal condition for this PDE is

$$\tilde{v}^*_i(T,s,z_i) = e^{\theta_i(z_i+g_i(s,m(T)))} + \theta_i \int g_{i,m}(\tilde{s},m(T))(s)e^{\theta_i(\tilde{z}_i+g_i(\tilde{s},m(T)))}\,\mu_i(T,d\tilde{s},d\tilde{z}_i).$$

From the equality $\tilde{V}^{\theta}_{i,\mu}(t,\mu) = \theta_i V^{\theta}_{i,\mu}(t,\mu)\tilde{V}^{\theta}_i(t,\mu)$, we deduce that the function $v^*_i(t,s,z_i) := V^{\theta}_{i,\mu_i}(\mu_i)(t,s,z_i) = \frac{\tilde{V}^{\theta}_{i,\mu}}{\theta_i\tilde{V}^{\theta}_i} = \frac{\tilde{v}^*_i(t,s,z_i)}{\theta_i\tilde{V}^{\theta}_i}$ is the dual function associated with the risk-sensitive best response value of decision-maker i.

Note that, as in the risk-neutral case, in general, the risk-sensitive dual function $v^*(t, s, z_i)$ is different than the risk-sensitive payoff function $V^\theta(t, \mu_i)$ or $V^\theta(t, \mu)$.

Itô's formula applied to

$$p_i^\theta(t) = \frac{\tilde{v}^*_{i,s}(t, s(t), z_i(t))}{\tilde{v}^*_{i,z_i}(t, s(t), z_i(t))} = \frac{\tilde{V}_{i,s\mu_i}(t, s(t), z_i(t))}{\tilde{V}_{i,z_i\mu_i}(t, s(t), z_i(t))},$$

provides a risk-sensitive first-order adjoint equation with the risk-sensitive Hamiltonian $H_i^\theta = H_i(t, s, m, p_i^\theta, q_i^\theta + p_i^\theta l_i)$.

Proposition 10. *For $\theta_i \neq 0$, the risk-sensitive first order adjoint process $(p_i^\theta, q_i^\theta, l_i, \eta_i)_{i \in I}$ solves*

$$dp_i^\theta = -\{H_{i,s} + \tfrac{1}{\eta_i}\mathbb{E}[\eta_i H_{i,s\mu_i}]\}dt + q_i^\theta[-l_i dt + dB],$$
$$p_i^\theta(T) = g_{i,s} + \tfrac{1}{\eta_i(T)}\mathbb{E}\{\eta_i(T)\partial_s g_{i,m}\},$$
$$d\eta_i = \eta_i l_i dB,$$
$$\eta_i(T) = \theta_i e^{\theta_i[z_i(T) + g_i(s(T), m(T, .))]}$$

Proposition 10 provides a risk-sensitive first order system for stochastic maximum principle of mean-field type. By identification of processes, one obtains

$$p_i^\theta(t) = \frac{\tilde{v}^*_{i,s}(t, s(t), z_i(t))}{\tilde{v}^*_{i,z_i}(t, s(t), z_i(t))} = \frac{\tilde{V}_{i,s\mu_i}(t, s(t), z_i(t))}{\tilde{V}_{i,z_i\mu_i}(t, s(t), z_i(t))},$$
$$q_i^\theta = \sigma \frac{\tilde{v}^*_{i,ss}}{\eta_i} - p_i^\theta l_i,$$
$$\eta_i(t) = \tilde{v}^*_{i,z_i}(t, s(t), z_i(t)),$$
$$l_i = \sigma \frac{\eta_{i,s}}{\eta_i} = \sigma \partial_s[\log \eta_i]$$

Proof. A proof can be obtained by following similar steps as for the proof of Proposition 4. □

6.3. Coalitional mean-field-type games

One of the fundamental element in the theory of cooperative games is the formulation of the optimal behavior for the decision-makers. Decision-maker behavior (control action and imputations) satisfying specific optimality behaviors then constitutes a solution of the game. In other words, a solution concept of a dynamic cooperative game is produced by a set of optimality principles such as dynamic bargaining solution and payoff allocation procedure. Altruism and cooperation are fascinating research areas. Intuitively, one has attempted to claim that the decision-makers are better off when they all work cooperatively. However, we are often observing very strange behaviors that are far from cooperation. So, if cooperation is answer, what is the question and why these strange behaviors?

Let us consider a simple example cooperative mean-field-type game [42] with two decision-makers. Assume that if they work together (jointly) they will be able to get $V(\{12\}, m_0, [0, T])$. DM 1 gets $V(\{1\}, m_0, [0, T])$ if he or she works alone and DM 2 gets $V(\{2\}, m_0, [0, T])$. From these three values, it is not clear why these decision-makers should work together. In order to formalize it in terms of their interest, we introduce a cost of making a coalition, $C(\{12\}, m_0, [0, T]) \geq 0$ which is the cost incurred when both decision-makers pool their effort (it includes information exchanging cost, coalition creation cost, etc). While this cost is often neglected in the literature, it may be important in many setups. Thus, a necessary condition for possible cooperation between the decision-makers is

$$V(\{12\}, m_0, [0, T]) - C(\{12\}, m_0, [0, T]) > V(\{1\}, m_0, [0, T]) + V(\{2\}, m_0, [0, T]).$$

Then, the next question is: what will be their payoff if they cooperate? To answer to this question, we need to know how to share the outcome of the cooperation. It is clear that allocating the equal share $\frac{1}{2}[V(\{12\}, m_0, [0, T]) - C(\{12\}, m_0, [0, T])]$ to each decision-maker is not necessary appropriate since it can be less than $\min_i(V(\{i\}, m_0, [0, T]))$. Thus, the allocation has to be done in a more clever way.

Cooperative game-theoretic solutions such as Bargaining solution, Core, Shapley value, Nucleolus dealt with such problems. When stochasticity and time-dependence are involved, the solution concepts require careful adaptations. In addition, if the payoff function and the state dynamics are of mean-field type, the optimality equations need to be established. In terms of equilibrium payoffs, a mean-field-type Nash equilibrium could plays the role of a benchmark in a cooperative game, i.e., gives what decision-makers could secure for themselves if there is no agreement, i.e., $V(\{i\}, m_0, [0, T])$.

Let $R_{c,i}$ be the total before-side-payment cooperative payoff of decision-maker i i.e.,

$$\mathbb{E}[g_i(s_c(T), m_c(T)) + \int_0^T r_i(t, s_c(t), m_c(t), a_c(t))dt]$$

s_c is the optimal state under cooperative (joint) decision-making scenario. One has $\sum_{j \in \mathcal{J}} R_{c,j} = V(\mathcal{J}, m_0, [0, T])$. However, one needs to find a better way to share the payoff $V(\mathcal{J}, m_0, [0, T]) - C(\mathcal{J}, m_0, [0, T])$. This leads to the introduction of the notion of imputation, i.e., a vector profile $(\gamma_j)_{j \in \mathcal{J}}$ such that $\sum_{j \in \mathcal{J}'} \gamma_j \geq V(\mathcal{J}', m_0, [0, T]) - C(\mathcal{J}', m_0, [0, T])$ for any $\mathcal{J}' \subset \mathcal{J}$, and $\sum_{j \in \mathcal{J}} \gamma_j = V(\mathcal{J}, m_0, [0, T]) - C(\mathcal{J}, m_0, [0, T])$

By virtue of mean-field type joint optimization, the sum of individual payoffs under cooperation is greater or equal to its noncooperative counterpart, i.e.,

$$\sum_{j \in \mathcal{J}} R_{c,j} \geq \sum_{j \in \mathcal{J}} V(\{j\}, m_0, [0, T]).$$

Thus, the dividend of cooperation (without the coalition making cost) is

$$DC = \sum_{j \in \mathcal{J}} R_{c,j} - \sum_{j \in \mathcal{J}} V(\{j\}, m_0, [0, T]) \geq 0.$$

Thus, the dividend of cooperation (with the coalition making cost) to be distributed among the decision-makers is $DC - C(\mathcal{J}, m_0, [0, T])$.

As a first consequence, it is clear if the coalition making cost is too high (compared to the game coalition value) then there is no reason for the decision-makers to form coalition.

Therefore for cooperation purpose we require the positivity of $DC - C(\mathcal{J}, m_0, [0, T])$. Using a cooperative game approach yields individual payoffs for the whole interval $[0, T]$. The selected imputation has, by definition, the property that each decision-maker's payoff in the cooperative game is higher or equal to what she would get in a noncooperative game played on the same time interval.

Let $\gamma_i(t) = \gamma(\{i\}, \mathcal{J}, m_0, [t, T])$ be the cooperative payoff-to-go after side payment for decision-maker i at position $[t, T], 0 < t < T$ of the game. This is the amount of individual payoff that decision-maker i will actually get. One way sharing the payoff is to use a dynamical Shapley value. The allocated payoff to decision-maker i under Shaley value is

$$\gamma_i = \sum_{\mathcal{J}' \subset \mathcal{J}, i \notin \mathcal{J}'} \frac{|\mathcal{J}'|!(|\mathcal{J}| - |\mathcal{J}'| - 1)!}{|\mathcal{J}|!} [(V - C)(\mathcal{J}' \cup \{i\}, [t, T]) - (V - C)(\mathcal{J}', [t, T])]. \quad (39)$$

For two decision-makers case, the payoffs of the decision-makers are

$$\gamma_1 =$$
$$V(\{1\}, m_0, [0, T]) + \frac{1}{2} \left[V(\{12\}, m_0, [0, T]) - C(\{12\}, m_0, [0, T]) - V(\{1\}, m_0, [0, T]) - V(\{2\}, m_0, [0, T]) \right],$$
$$\gamma_2 =$$
$$V(\{2\}, m_0, [0, T]) + \frac{1}{2} \left[V(\{12\}, m_0, [0, T]) - C(\{12\}, m_0, [0, T]) - V(\{1\}, m_0, [0, T]) - V(\{2\}, m_0, [0, T]) \right].$$

Definition 4. *A cooperative solution is time consistent if, at any position $t \in [0, T]$ the cooperation solution payoff to go $\gamma(\{i\}, m(t), [t, T]) \geq V(\{i\}, m(t), [t, T])$ where the deviating payoff $V(\{i\}, m(t), [t, T])$ is computed along the cooperative state trajectory $s_c(t)$.*

This notion of time consistency and its implementation in cooperative differential games was introduced in [6]. A stronger notion of time consistency is that the cooperative payoff-to-go dominates the noncooperative payoff-to-go for any state $s(t)$, $t \in [[0, T]$. This is called sub-game consistent solution.

Though one of the most commonly used allocation principles is the dynamical Shapley value, however, in the case when decision-makers may be asymmetric in their powers and sizes of payoffs, equal imputation of cooperative gains may not be totally agreeable to asymmetric decision-maker. To overcome this, one can suggest the allocation principle in which the decision-makers, shares of the gain from cooperation are proportional to the relative sizes of their expected deviating payoffs. Thus, a proportional time-consistent solution is given by

$$\gamma_i(s) = \frac{V(\{i\}, [t, T])}{\sum_i V(\{i\}, [t, T])} [V(\mathcal{J}, [t, T]) - C(\mathcal{J}, [t, T])]. \tag{40}$$

If these quantities are positive, one gets $\gamma(\{i\}, \mathcal{I}, [t, T]) \geq V(\{i\}, m(t), [t, T]), \forall t$, (individual rationality) and $\sum_{i \in \mathcal{I}} \gamma(\{i\}, \mathcal{I}, [t, T]) = V(\mathcal{I}, m(t), [t, T]) - C(\mathcal{I}, [s, T])$ (efficiency at any time).

For dynamic games, an additional and stringent condition on the solutions is required: The specific optimality principle must remain optimal at any instant of time throughout the game duration along the optimal state trajectory. This condition is known as dynamic stability or time consistency.

In the context of mean-field type games, the notion of time consistency is crucial since the initial distribution of states and starting time influences naturally the Kolmogorov forward equation. A cooperative solution is sub-game consistent if an extension of the cooperative strategy to a situation with a later starting time and to any possible state brought about by the prior optimal behavior of the decision-makers remains optimal. Sub-game consistent is a stronger notion of time consistency. In the presence of stochastic elements, sub-game consistency is required in a credible cooperative solution. In the field of cooperative mean-field type games, little research has been published to date on sub-game consistent solutions.

If the set $\mathcal{A}$ is a non-convex but a general separable complete metric space (Polish space), Pontryagin's approach suggests the following perturbation method called spike variation. The approach is well-adapted to sub-game perfection in games. Fix $(t, s) \in [0, T] \times \mathcal{S}$ and define the control law a_ϵ as the spike variation of $\hat{a}$ over the set $[t, t + \epsilon]$, $\epsilon > 0$ i.e.,

$$a_\epsilon(t') = a(t') \mathbb{1}_{[t, t+\epsilon]}(t') + \hat{a}(t') \mathbb{1}_{[0, T] \setminus [t, t+\epsilon]}(t'),$$

where a is an arbitrary admissible control and $\mathbb{1}_{[t, t+\epsilon]}$ is the indicator function over the set $[t, t + \epsilon]$.

Definition 5. *Let R_J be objective to be maximized in a. We say that $\hat{a}$ is a sub-game perfect cooperative strategy under spike variation if for any t_0, s_0, a,*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} [R_J([t_0, T], s_0, \hat{a}) - R_J([t_0, T], s_0, a_\epsilon)] \geq 0.$$

Note that a sub-game perfect cooperative strategy under spike variation is in particular a time consistent solution.

The key difference here is that the solution that we are looking for, should not depend on the initial data (when and where we started).

Let H^{t_0, s_0} be the Pontryagin function associated with the random variable s that starts from s_0 at $t_0 \in [0, T]$. $H^{t_0, s_0}(t, s, m, a, p, q) = bp^{t_0, s_0} + \sigma q^{t_0, s_0} + r_J^{t_0, s_0}$, where the notation r^{t_0, s_0} is obtained from r when the aggregate term is $m = m^{t_0, s_0, a}$. The first order adjoint equation process $(p^{t_0, s_0}, q^{t_0, s_0})$ under optimal cooperative control law becomes

$$dp^{t_0, s_0} = -[H_s^{t_0, s_0} + \mathbb{E} \partial_s H_m^{t_0, s_0}]dt + q^{t_0, s_0} dB(t), \ t > t_0,$$
$$p^{t_0, x_0}(T) = g_s^{t_0, s_0}(T) + \mathbb{E} \partial_s g_m^{t_0, s_0}(T). \tag{41}$$

The second order adjoint equation $(P^{t_0, s_0}, Q^{t_0, s_0})$ is defined following similar steps as in (14) but with conditioning on t_0, s_0.

Proposition 11. *Let the assumptions of Lemma 5 hold. If $(\hat{x}, \hat{a})$ is an optimal solution of the cooperative mean-field type game then there are two pairs of processes $(p, q), (P, Q)$ that satisfy the first order and the second order adjoint equations, such that*

$$H^{t, \hat{x}(t)}(t, \hat{x}, \hat{m}, \hat{a}, p^{t, \hat{x}(t)}, q^{t, \hat{x}(t)}) - H^{t, \hat{x}(t)}(t, \hat{x}, \hat{m}, a, p^{t, \hat{x}(t)}, q^{t, \hat{x}(t)})$$

$$+ \frac{1}{2} P^{t, \hat{x}(t)}(t) (\sigma(t, \hat{x}, \hat{m}, \hat{a}) - \sigma(t, \hat{x}, \hat{m}, a))^2 \geq 0,$$

for all $a(.) \in \mathcal{A}$, almost every t and $\mathbb{P}-$almost surely.

7. Concluding remarks

In this article, basic results on mean-field-type games were presented. Most of the results presented here extend to the multi-dimensional state case (vector or matrix), and the state-of-the-art stochastic maximum principle can carried out random coefficient functions, common noise as well as time delays and backward-forward stochastic integro-differential equation. The state-of-art dynamic programming principle works on infinite dimension, in the space of measures. By introducing dual functions which are weak Gateaux-derivatives, relationships between stochastic maximum principle and dynamic programming were established. The methodology was shown to be flexible enough to carry out partial state observation and imperfect state measurement in the non-degenerate case using Girsanov transform. Wiener chaos expansion of the underlying processes were proposed to solve the game problems and truncature error bounds were derived using Kosambi-Karhunen-Loeve's approach. This allows one to solve efficiently of the mean-field-type problem much faster than standard methods as multi-level Monte-Carlo sampling or stochastic collocation methods. The choice of a basis is crucial as the number of elements influence the curse of dimensionality of the problem. Sparse representation and non-intrusive proper generalized decomposition of the processes may be needed in order to significantly reduce the complexity.

Acknowledgements

This research work is supported by U.S. Air Force Office of Scientific Research under grant number FA9550-17-1-0259. The author is grateful to Prof. Boualem Djehiche for useful comments on the Wiener Chaos Expansions in mean-field-type games. The author is grateful to the Editor and the reviewers' valuable comments that improved the manuscript.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

References

1. L. S. Shapley, *Stochastic games*, PNAS, P. Natl Acad. Sci. USA, **39** (1953), 1095-1100.

2. B. Jovanovic, R. W. Rosenthal, *Anonymous sequential games*, J. Math. Econ. Elsevier, **17** (1988), 77-87.

3. J. M. Lasry, P. L. Lions, *Mean field games*, Jpn. J. Math. **2** (2007), 229-260.

4. A. Bensoussan, B. Djehiche, H. Tembine, et al. *Risk-sensitive mean-field-type control*, IEEE CDC, 2017.

5. B. Djehiche, S. A. Tcheukam and H. Tembine, *A Mean-Field Game of Evacuation in Multi-Level Building*, IEEE T. Automat. Contr. **62** (2017), 5154-5169.

6. L. A. Petrosjan, *Stability of the solutions in differential games with several players*, Vestnik Leningrad. univ. **19** (1977), 46-52.

7. B. Jourdain, S. Méléard and W. Woyczynski, *Nonlinear SDEs driven by Lévy processes and related PDEs*, Alea, **4** (2008), 1-29.

8. R. Buckdahn, B. Djehiche, J. Li, et al. *Mean-field backward stochastic differential equations: a limit approach*, Ann. Probab. **37** (2009), 1524-1565.

9. R. Buckdahn, J. Li and S. Peng, *Mean-field backward stochastic differential equations and related partial differential equations*, Stoch. Proc. Appl. **119** (2009), 3133-3154.

10. D. Andersson, B. Djehiche, *A maximum principle for SDEs of Mean-Field Type*, Appl. Math. Opt. **63** (2011), 341-356.

11. R. Buckdahn, B. Djehiche and J. Li, *A General Stochastic Maximum Principle for SDEs of Mean-Field Type*, Appl. Math. Opt. **64** (2011), 197-216.

12. J. Hosking, *A stochastic maximum principle for a stochastic differential game of a mean-field type*, Appl. Math. Opt. **66** (2012), 415-454.

13. J. Li, *Stochastic maximum principle in the mean-field controls*, Automatica, **48** (2012), 366-373.

14. M. Hafayed, S. Abbas, A. Abba, *On mean-field partial information maximum principle of optimal control for stochastic systems with Lévy processes*, J. Optimiz. Theory App. **167** (2015), 1051-1069.

15. R. Elliott, X. Li and Y. H. Ni, *Discrete time mean-field stochastic linear-quadratic optimal control problems*, Automatica, **49** (2013), 3222-3233.

16. M. Hafayed, *A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes*, International Journal of Dynamics and Control, **1** (2013), 300-315.

17. A. Bensoussan, J. Frehse and S. C. P. Yam, *Mean-field games and mean-field-type control theory*, SpringerBriefs in mathematics, Springer, 2013.

18. B. Djehiche, H. Tembine and R. Tempone, *A Stochastic Maximum Principle for Risk-Sensitive Mean-Field-Type Control*, IEEE T. Automat. Contr. **60** (2015), 2640-2649.

19. H. Tembine, *Risk-sensitive mean-field-type games with Lp-norm drifts*, Automatica, **59** (2015), 224-237.

20. B. Djehiche, H. Tembine, *Risk-Sensitive Mean-Field-Type Control under Partial Observation*, In Stochastics in environmental and financial economics, Springer Proceedings in Mathematics and Statistics, 2016.

21. K. Law, H. Tembine and R. Tempone, *Deterministic Mean-Field Ensemble Kalman Filtering*, SIAM J. Sci. Comput. **38** (2016), 1251-1279.

22. J. Robinson, *Applied Analysis lecture notes*, University of Warwick, 2006.

23. R. H. Cameron, W. T. Martin, *The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals*, Ann. Math. **48** (1947), 385-392.

24. N. Wiener, *The homogeneous chaos*, Am. J. Math. **60** (1938), 897-936.

25. N. Wiener, *Nonlinear Problems in Random Theory*, Technology Press of the Massachusetts Institute of Technology/Wiley, New York, 1958.

26. F. Augustin, A. Gilg, M. Paffrath, et al. *Polynomial chaos for the approximation of uncertainties: Chances and limits*, Eur. J. Appl. Math. **19** (2008), 149-190.

27. G. Wang, C. Zhang and W. Zhang, *Stochastic maximum principle for mean-field type optimal control under partial information*, IEEE T. Automat. Contr. **59** (2014), 522-528.

28. Z. Wu, *A maximum principle for partially observed optimal control of forward-backward stochastic control systems*, Science China information sciences, **53** (2010), 2205-2214.

29. J. Yong, *Linear-quadratic optimal control problems for mean-field stochastic differential equations*, SIAM J. Control Optim. **51** (2013), 2809-2838.

30. H. Tembine, *Uncertainty quantification in mean-field-type teams and games*, In Proceedings of 54th IEEE Conference on Decision and Control (CDC), 2015, 4418-4423,

31. S. Tang, *The maximum principle for partially observed optimal control of stochastic differential equations*, SIAM J. Control Optim. **36** (1998), 1596-1617.

32. H. Ma, B. Liu, *Linear Quadratic Optimal Control Problem for Partially Observed Forward Backward Stochastic Differential Equations of Mean-Field Type*, Asian J. Control, 2017.

33. G. Wang, Z. Wu and J. Xiong, *A linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information*, IEEE T. Automat. Contr. **60** (2015), 2904-2916.

34. Q. Meng, Y. Shen, *Optimal control of mean-field jump-diffusion systems with delay: A stochastic maximum principle approach*, J. Comput. Appl. Math. **279** (2015), 13-30.

35. T. Meyer-Brandis, B. Oksendal and X. Y. Zhou, *A mean-field stochastic maximum principle via Malliavin calculus*, Stochastics, **84** (2012), 643-666.

36. Y. Shen, Q. Meng and P. Shi, *Maximum principle for mean-field jump-diffusion stochastic delay differential equations and its application to finance*, Automatica, **50** (2014), 1565-1579.

37. Y. Shen, T. K. Siu, *The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem*, Nonlinear Anal-Theor, **86** (2013), 58-73.

38. G. Wang, Z. Wu, *The maximum principles for stochastic recursive optimal control problems under partial information*, IEEE T. Automat. Contr. **54** (2009), 1230-1242.

39. G. Wang, Z. Wu and J. Xiong, *Maximum principles for forward-backward stochastic control systems with correlated state and observation noises*, SIAM J. Control Optim. **51** (2013), 491-524.

40. B. Djehiche, A. Tcheukam and H. Tembine, *Mean-Field-Type Games in Engineering*, AIMS Electronics and Electrical Engineering, **1** (2017), 18-73.

41. J. Gao, H. Tembine, *Distributed Mean-Field-Type Filters for Big Data Assimilation*, IEEE International Conference on Data Science Systems (DSS), Sydney, Australia, 12-14 Dec. 2016.

42. A. K. Cisse, H. Tembine, *Cooperative Mean-Field Type Games*, 19th World Congress of the International Federation of Automatic Control (IFAC), Cape Town, South Africa, 24-29 August 2014, 8995-9000.

43. H. Tembine, *Tutorial on Mean-Field-Type Games*, 19th World Congress of the International Federation of Automatic Control (IFAC), Cape Town, South Africa, 24-29 August 2014.

44. G. Rossi, A. S. Tcheukam and H. Tembine, *How Much Does Users' Psychology Matter in Engineering Mean-Field-Type Games*, Workshop on Game Theory and Experimental Methods, June 6-7, Second University of Naples, Italy, 2016.

45. A. S. Tcheukam, H. Tembine, *On the Distributed Mean-Variance Paradigm*, 13th International Multi-Conference on Systems, Signals & Devices. Conference on Systems, Automation & Control. March 21-24, 2016 - Leipzig, Germany, 604-609.

46. T. E. Duncan, H. Tembine, *Linear-quadratic mean-field-type games: A direct method*, Preprint, 2017.

47. T. E. Duncan, H. Tembine, *Linear-quadratic mean-field-type games with common noise: A direct method*, Preprint, 2017.

48. T. E. Duncan, H. Tembine, *Other-regarding payoffs in linear-quadratic mean-field-type games with common noise: A direct method*, Preprint, 2017.

49. T. E. Duncan, H. Tembine, *Nash bargaining solution in linear-quadratic mean-field-type games: A direct method*, Preprint, 2017.

50. H. Tembine, B. Djehiche, P. Yam, et al. *Mean-field-type games with jumps and regime switching*, Preprint, 2017.

AIMS Press