We consider a moral hazard problem with multiple principals in a continuous-time model. The agent can only work exclusively for one principal at a given time, so faces an optimal switching problem. Using a randomized formulation and techniques from the theory of backward SDEs, we manage to represent the agent’s value function and his optimal effort by an Itô process. This representation further helps to solve the principals’ problem in case we have infinite number of principals in the sense of mean field game. Finally, to justify the mean field formulation, we develop the so-called backward propagation of chaos, which may carry independent interest itself.

1. Introduction

The principal-agent problem is a study of optimizing the incentives, so central in economics. In particular, the optimal contracting between the two parties, principal and agent(s), is called moral hazard, when the agent’s effort is not observable by the principal. It has been widely applied in many areas of economics and finance, for example in corporate finance, see [10] and the recent works by Cvitanić, Possamaï and Touzi [19], El Euch, Mastrolia, Rosenbaum and Touzi [29], Cvitanić and Xing [21], Baldacci, Manziuk, Mastrolia and Rosenbaum [5]. More recently, we also witness the works Aïd, Possamaï and Touzi [1], Alasseur, Ekeland, Élie, Hernández and Possamai [2], Élie, Hubert, Mastrolia and Possamai [25], Alasseur, Farhat and Saguan [3], using the principal-agent formulation to study how to design an optimal electricity contract and how to encourage people to embrace the energy transition. We would also like to mention the very recent applications related to epidemic control by Aurell, Carmona, Dayanikli and Laurière [4], Hubert, Mastrolia, Possamai and Warin [38].

While the research on the discrete-time model can be dated back further, the first paper on continuous-time principal-agent problem is the seminal work by Holmström and Milgrom [36], who study a simple continuous-time model in which the agent gets paid at the end of a finite time interval. They show that optimal contracts are linear in aggregate output when the agent has exponential utility with a monetary cost of effort. The advantage...
of the continuous-time model is further explored by Sannikov [49]. Not only he considers a new model which allows the agent to retire, but also (and more importantly in the mathematical perspective) he introduces new dynamic insights to the principal-agent problem, and it leads to simple computational procedure to find the optimal contract by solving an ordinary differential equation in his case.

Later, the idea of Sannikov is interpreted and introduced to the mathematical finance community by Cvitanić, Possamaï and Touzi [20]. Let us illustrate their contribution with a toy model. Denote by $\xi$ the contract paid at the end of a finite time interval $[0, T]$. Assume the agent faces the following optimization:

$$
\max_{\alpha} \mathbb{E} \left[ \xi(X^\alpha) - \int_0^T c(\alpha_t) \, dt \right], \quad \text{where } dX_t^\alpha = dW_t + \alpha_t \, dt.
$$

The crucial observation in [20] is that both the contract $\xi$ and the Agent’s best response $\alpha^*[\xi]$ can be characterized by the following backward stochastic differential equation (in short, BSDE, for readers not familiar with BSDE we refer to [24,46], and in particular to [22] for the applications on the contract theory):

$$
dY_t = -c^*(Z_t) \, dt + Z_t \, dW_t, \quad Y_T = \xi, \quad \text{where } c^*(z) = \max_{a} \{az - c(a)\},
$$

namely, $\xi = Y_T$ and $\alpha^*_t[Z] = \arg\max_{a} \{aZ_t - c(a)\}$ for all $t \in [0, T]$. This induces a natural (forward) representation of the couple $(\xi, \alpha^*_\xi)$:

$$
\begin{cases}
\xi = Y_{T}^{Y_0,Z} := Y_0 - \int_0^T c^*(Z_t) \, dt + \int_0^T Z_t \, dW_t, \\
\alpha^*_t[Z] = \alpha^*_t[Y_0,Z] = \arg\max_{a} \{aZ_t - c(a)\},
\end{cases}
$$

for some $(Y_0, Z)$, and this neat representation transforms the once puzzling principal’s problem into a classical control problem, namely,

$$
\max_{\xi} \mathbb{E} \left[ U \left( X_T^{\alpha^*_\xi} - \xi \right) \right] = \max_{Y_0,Z} \mathbb{E} \left[ U \left( X_T^{\alpha^*_\xi[Z]} - Y_T^{Y_0,Z} \right) \right].
$$

Cvitanić, Possamaï and Touzi [20] provide also an extension of Sannikov [49] allowing to address a wide spectrum of principal-agent problems, in particular problems with volatility control, which required the use of second-order BSDEs. We also mention that the work of Sannikov [49] has also been recently revisited by Possamaï and Touzi [48]. This idea of representation is further applied to study the case where the principal can hire multiple agents. See Élie and Possamaï [28], Koo, Shim and Wang [40], Baldacci, Possamaï and Rosenbaum [6]. Eventually, in Élie, Mastrolia and Possamaï [27] the authors follow the same machinery to study the model where the principal hires infinite number of agents using the formulation of mean field games (as for the mean field game we refer to the seminal paper [42] and the recent books [14] and [15]). See also the works by Carmona and Wang [17], Élie, Hubert, Mastrolia and Possamaï [25] for a continuum of agents with mean-field interactions.

There are fewer existing literature on the model concerning one agent facing multiple principals. Meanwhile, along the emerging services, such as private teachings, household services, Uber-like taxis, and so on, more and more people are involved in careers during
which they can switch from one employer (typically an online platform) to another more freely than ever before. This new trend evokes our curiosity to have a theoretical insight on the incentive-mechanism of such job markets.

In the 1980s and 1990s, the economists investigated the common agency problem in the discrete time case, i.e. the agent simultaneously works on different projects for different principals. See [7–9,12,23]. Recently, in [45] the authors the common agency problem in the continuous time case. However, to the best of our knowledge, no one has yet considered a \( n \)-principal/1-agent model where the agent can only exclusively work for one principal at a given time. In such a model, the agent is facing an optimal switching (among the principals) problem, i.e. the agent is looking for optimal stopping times to switch and optimal regimes. According to the classic literature of optimal switching, see e.g. [18,32,37,47], the counterpart of the BSDE characterization (1) for the agent’s problem would be a system of reflected BSDE in the form:

\[
\begin{align*}
    dY_t &= -f_t(Y_t, Z_t)\, dt + Z_t\, dW_t - \, dK_t, \\
    Y_T &= \xi,
\end{align*}
\]

where \( K \) is an increasing process satisfying some (backward) Skorokhod condition. The presence of the process \( K \) and its constraint make it difficult to find a representation of the couple \( (\xi, \alpha^*[\xi]) \) as in [20]. In this paper, we propose an alternative facing this difficulty. Instead of letting the agent choose stopping times to change his employers, we study a randomized optimal switching problem where the switching time is modelled by a random time characterized by a Poisson point process and the agent influences the random time by controlling the intensity of the Poisson point process. Here, the intensity describes the hesitation of changing employer. It is fair to note that similar randomized formulations of switching have been discussed in the literature, see e.g. [11,26]. In such framework, we may characterize \( (\xi, \alpha^*[\xi]) \) by the solution to a system of BSDEs (without reflection).

Unfortunately, this characterization of the agent’s problem and the corresponding representation do not help us to solve simply the principals’ problem. We observe that the optimizations the principals face are time-inconsistent, that is, these cannot be solved by dynamic programming. In order to get around this difficulty, we suggest the following two approaches:

- We propose and study a ‘suboptimal’ optimization for the principals which turns out to be time-consistent. In this case, we find the PDE system characterizing the Nash equilibrium and solve it numerically.
- We also note that this time-inconsistency disappears once the number of principals tends to infinity. In the setting of infinite number of principals, it is natural to adopt the formulation of mean field game. We prove the existence of mean field equilibrium largely based on the recipes in Lacker [41]. Further, in order to justify our mean field game formulation, we introduce a machinery named ‘backward propagation of chaos’ which may carry independent interest itself.

The rest of the paper is organized as follows. In Section 2, we state our moral hazard problem with \( n \)-principal and one agent, we shall solve the agent’s problem under the previously mentioned randomized optimal switching formulation, and observe the time-inconsistency of the \( n \)-principal problem. In Section 3 we study the time-consistent ‘suboptimal’ optimization for the principals. Finally in Section 4, we shall derive the mean
field game among the principals, prove its existence, and justify it using the 'backward propagation of chaos' technique.

2. Moral hazard problem: \( n \) principals and one agent

In this paper we consider the principal-agent problem on the a finite time horizon \([0, T]\) for some \( T > 0 \). The main novelty is to introduce the new model and method to allow the agent to choose working for different employers. In this section, we set up the model in which one agent switches among \( n \) different principals.

For the agent, the set of possible regimes is \( I_n := \{1, 2, \ldots, n\} \). Denote by \( C([0, T]; \mathbb{R}^n) \) the set of continuous functions from \([0, T]\) to \( \mathbb{R}^n \), endowed with the supremum norm \( \| \cdot \|_T \), where

\[
\|x\|_t := \sup_{0 \leq s \leq t} |x_s|, \quad t \in [0, T], \quad x \in C([0, T]; \mathbb{R}^n).
\]

Denote by \( D([0, T], I_n) \) the set of càdlàg functions from \([0, T]\) to \( I_n \). We introduce the canonical space \( \Omega_1 := C([0, T]; \mathbb{R}^n) \times D([0, T]; I_n) \), denote the canonical process by \((W, I)\), and the canonical filtration by \( \mathbb{F} = \{\mathcal{F}_t\}_{t \in [0, T]} \). The process \( X_t \) represents the outputs of the principals and \( I_t \) records for which principal the agent is working at time \( t \). Denote by \( \mathbb{F}^W \) the filtration generated by the process \( W \) alone.

In our model, for simplicity, the output process follows the dynamic

\[
X_t = \int_0^t e_{I_s} \, ds + W_t, \quad \text{where} \quad e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T \in \mathbb{R}^n, \tag{2}
\]

where \( W \) is a \( n \)-dimensional Brownian motion. That is, the principal employing the agent has a constant-1 drift while outputs of the others just follow the Brownian motion. The agent controls the process \( I \), that is, he switches from one employer to another so as to maximize his utility:

\[
V^A_0(\xi, w) = \sup I \left\{ \sum_{i=1}^n \left( \xi^i 1_{I_t = i} + \int_0^T \theta^i_t 1_{I_t = i} \, du \right) - C_T(I) \right\}, \tag{3}
\]

where \( \xi^i \) is the reward at the terminal time \( T, \theta^i \) is the wage payed by principal \( i \), and \( C_T(I) \) denotes the cost of switching \( I \) up to time \( T \) which will be defined later when we analyse the agent’s problem in detail. Naturally, the payment \( \xi^i \) should be an \( \mathcal{F}_T^W \)-measurable random variable and \( \theta^i \) should be an \( \mathcal{F}_T^W \)-adapted process.

Providing the contracts \( \{(\xi^i, \theta^i)\}_{i \in I_n} \), respectively, the \( n \) principals search among them a Nash equilibrium so as to maximize their profits:

\[
V^P_{0,i} = \max_{(\xi^i, \theta^i)} \mathbb{E} \left[ X_T^{i,\xi^i} - U^i(\xi^i) 1_{I_T = i} - \int_0^T U^i(\theta^i_t) 1_{I_t = i} \, dt \right], \tag{4}
\]

where \( X_T^{i,\xi^i} \) denotes the optimal output process controlled by the agent. Here, we use \( U^i \) to describe the principals’ preference. We assume that \( U^i \) is convex and increasing. We also note that \( x \mapsto -U^i(-x) \) is concave and increasing, i.e. the usual definition of a utility function. See Section 3.2 for concretes examples.
Remark 2.1: In the agent problem (4), the function $U^i$ applies only to the final payment $\xi^i$ and the wage $\theta^i$. We may also have another function $U_0^i$ applying to $X^{i,k^i}_T$ as well, i.e.

\[
V_0^{P,i} = \max_{(\xi^i, \theta^i)} \mathbb{E} \left[ -U_0^i(-X^{i,k^i}_T) - U^i(\xi^i) 1_{\{T=i\}} - \int_0^T U^i(\theta^i) 1_{\{T=i\}} \, dt \right].
\]

(5)

For the sake of simplicity, we assume that $U_0^i(x) = x$, for which we have (4). See Remark (2.11) for more details.

As in [20,44,49], we are going to provide a representation of the value function of the agent’s problem so as to solve the principals’ problem by dynamic programming.

2.1. Agent’s problem

In our model, instead of allowing the agent to control the process $I$ in a ‘singular’ way, we assume that $I$ is driven by a Poisson random measure (p.r.m.) and allow the agent to control its intensity. More precisely, in the weak formulation of stochastic control, the agent aims at choosing his optimal control among:

\[
\{ \mathbb{P}^{\alpha} \in \mathcal{P}(\Omega) : \mathbb{P}^{\alpha} \text{-a.s. } W \text{ is a Brownian motion, } dI_t = \int_{\mathbb{N}} (k - I_{t^-}) \mu(dt, dk), \}
\]

where $\mu$ is a p.r.m. with intensity $\alpha(dt, \cdot) = \sum_{k \in \mathbb{N}} \alpha^k_t \delta_k dt$,

for some non-negative $\mathbb{F}$-adapted $(\alpha^k)$, $W$ and $\mu$ are independent,

where $\mathcal{P}(\Omega)$ is the set of all probability measures on $\Omega$ and $\delta_k$ is the Dirac measure.

Remark 2.2: Define the first jump time of the process $I$:

\[
\tau_I := \inf\{s \geq t : I_s \neq I_t\}.
\]

(6)

It follows from the Girsanov theorem for multivariate point processes (see e.g. [39]) that for any $\mathbb{F}^W$-adapted processes $\{\gamma(i)\}_{i \in \mathbb{N}}$, we have

\[
\mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \gamma_{\tau_I}(I_{\tau_I}) 1_{\{\tau_I \leq s, I_{\tau_I} = j\}} \right] = \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \int_t^s \beta^{i,\alpha}(t, u) \alpha^j_u \gamma_u(j) \, du \right], \quad \text{for all } j \in \mathbb{N} \text{ and } j \neq I_t,
\]

as well as $\mathbb{P}^{\alpha}[\tau_I > s | \mathcal{F}_s] = \beta^{i,\alpha}(t, s)$, where $\beta^{i,\alpha}(t, s) := \exp\left( -\int_t^s \sum_{j \neq i} \alpha^j_u \, du \right)$.

These results will be useful in the upcoming calculus.

With the intensity $\alpha$ introduced as above, we can now make it precise how to define the cost of switching in (3). Given $n$ contracts $\{(\xi^i, \theta^i)\}_{i \in \mathbb{N}}$, the agent solves the following
optimal switching problem

\[ V_0^A = \sup_{\alpha} \mathbb{E}^{\mathbb{P}^\alpha} \left[ \sum_{i=1}^{n} \xi^i 1_{\{I_T = i\}} + \int_0^T \left( \sum_{i=1}^{n} \theta^i_u 1_{\{I_u = i\}} - \frac{1}{n-1} \sum_{i \neq I_u} c((n-1)\alpha^i_u) \right) \, du \right]. \]  

(7)

Here, \( c \) is a cost function, i.e. a non-negative convex function. Intuitively, the intensity process \( \alpha \) describes the hesitation of changing employer. The bigger \( \alpha \) is, the less hesitation the agent has to change his employer. The agent, when working for the principal \( i \), can choose some effort \( \alpha^j \) in order to increase the probability of switching from principal \( i \) to principal \( j \). The normalization \( \frac{1}{n-1} c((n-1)\cdot) \) is for later use, as \( n \to \infty \).

**Remark 2.3:** Unlike the moral hazard problems proposed by Holmström and Milgrom [36], Sannikov [49] and Cvitanić, Possamaï and Touzi [20], in our setting since the drift of the output diffusion \( X \) (2) is constant, that is, the principals cannot incite the agent to make more ‘effort’ during his work. Instead, the principals make contracts based on the information of \( W = (W^1, \cdots, W^n) \) to give the agent the incentive to switch among them (see Remark 2.7). Note that the principals still do not have insight to the intensity that the agent chooses to switch among the different job opportunities. Therefore, the principals encounter a new type of moral hazard problem, and need to provide incentives so as to obtain the most advantageous feedback from the agent.

As in the classical literature of the optimal switching problems, we shall use the dynamic programming principle to obtain the system of equations characterizing the value function. First, define the dynamic version of (7):

\[ V_t^A = \text{ess} \sup_\alpha \mathbb{E}^{\mathbb{P}^\alpha} \left[ \sum_{i=1}^{n} \xi^i 1_{\{I_T = i\}} + \int_t^T \left( \sum_{i=1}^{n} \theta^i_u 1_{\{I_u = i\}} - \frac{1}{n-1} \sum_{i \neq I_u} c((n-1)\alpha^i_u) \right) \, du \right]. \]

Recall \( \tau_t \) defined in (6). By the dynamic programming, see, e.g. [30, Theorem 5.2], we have

\[ V_t^A = \text{ess} \sup_\alpha \mathbb{E}^{\mathbb{P}^\alpha} \left[ \sum_{i=1}^{n} \xi^i 1_{\{I_{\tau_t} = i, \, \tau_t > T\}} + V_{\tau_t}^A 1_{\{\tau_t \leq T\}} \right] 
\]

\[ + \int_t^{\tau_t \wedge T} \left( \sum_{i=1}^{n} \theta^i_u 1_{\{I_u = i\}} - \frac{1}{n-1} \sum_{i \neq I_u} c((n-1)\alpha^i_u) \right) \, du \].

Further, by defining \( V_{t, i}^{A,j} := V_t^A |_{I_t = i} \), the value of the control problem at time \( t \) given \( I_t = i \), we obtain

\[ V_{t, i}^{A,j} = \text{ess} \sup_\alpha \mathbb{E}^{\mathbb{P}^\alpha} \left[ \xi^i 1_{\{\tau_t > T\}} \right] 
\]

\[ + \sum_{j \neq i} \left( V_{\tau_t}^{A,j} 1_{\{\tau_t \leq T, \, I_{\tau_t} = j\}} + \int_t^{\tau_t \wedge T} \frac{\theta^j_u - c((n-1)\alpha^j_u)}{n-1} \, du \right) \].
\[
= \text{ess sup}_{\alpha} \mathbb{E}_{t}^{P_{0}} \left[ \xi^{i} \beta^{i \alpha}(t, T) 
+ \int_{t}^{T} \beta^{i \alpha}(t, u) \sum_{j \neq i} \left( \alpha^{j}_{u} V^{A_{j}}_{u} + \theta^{j}_{u} - c((n-1)\alpha^{j}_{u}) \right) du \right].
\] (8)

The second equality above is due to the results in Remark 2.2. In view of (8), it becomes a classical stochastic control problem. In particular, the value function of this control problem can be characterized by the BSDE.

**Assumption 2.4:** Assume that the cost function \( c \) is convex, and \( c \) takes value of \(+\infty\) out of a compact set \( K \). We also assume that there exists a unique maximizer

\[
a^{*}(y) := \arg \max_{a \geq 0} \{ ay - c(a) \} \in K, \quad \text{for all } y \in \mathbb{R}.
\]

Define the convex conjugate of \( c \)

\[
c^{*}(y) := \sup_{a \geq 0} \{ ay - c(a) \}, \quad \text{for all } y \in \mathbb{R}.
\]

Further assume that \( c^{*} \) is Lipschitz continuous.

**Example 2.5:** As an example, consider \( c(a) := \frac{1}{2} a^{2} \) on \( K := [0, \overline{K}] \) for some \( \overline{K} > 0 \). Then,

\[
a^{*}(y) = \begin{cases} 
0, & y \leq 0, \\
y, & 0 \leq y \leq \overline{K}, \\
\overline{K}, & y \geq \overline{K},
\end{cases}
\]

and \( c^{*}(y) = \begin{cases} 
0, & y \leq 0, \\
\frac{1}{2} y^{2}, & 0 \leq y \leq \overline{K}, \\
\overline{K} y - \frac{1}{2} \overline{K}^{2}, & y \geq \overline{K}.
\end{cases} \) (9)

Obviously, the function \( c^{*} \) is Lipschitz continuous with constant \( \overline{K} \).

**Proposition 2.6:** Under Assumption 2.4, given \( \xi^{i} \in L^{2}(\mathbb{P}_{0}), \theta^{i} \in H^{2}(\mathbb{P}_{0}) \) for all \( i \in \mathbb{I}_{n} \), the following system of BSDEs has a unique solution \( (Y^{i}, Z^{i})_{i \in \mathbb{I}_{n}} \):

\[
Y^{i}_{t} = \xi^{i} + \int_{t}^{T} \left( \frac{1}{n-1} \sum_{j \neq i} c^{*}(Y^{j}_{u}^{i} - Y^{i}_{u}^{i}) + \theta^{i}_{u} \right) du - \int_{t}^{T} Z^{i}_{u} \cdot dW_{u}, \quad i \in \mathbb{I}_{n}, \quad \mathbb{P}_{0}\text{-a.s.}
\]

Moreover, we have \( Y^{i}_{t} = V^{A,i}_{t}, \mathbb{P}_{0}\text{-a.s.} \), i.e. the solution \( Y^{i}_{t} \) of the BSDE represents the continuation utility of the agent when working for the principal \( i \) at time \( t \). In particular, the optimal
intensity satisfies

\[ a_t^{j^*} = \frac{1}{n-1} a^*(Y_t^j - Y_t^{I_t}), \quad j \neq I_t, \quad \text{for all } t \in [0, T], \quad \mathbb{P}_0\text{-a.s.} \quad (11) \]

**Proof:** In view of the control problem (8), following the result of [24,31], see also [50], the corresponding BSDE reads

\[
Y_i^t = \xi^i + \int_t^T \left( \frac{1}{n-1} \sum_{j \neq i} \sup_{a^j \geq 0} \left\{ (n-1)a^j(Y_u^j - Y_u^i) - c((n-1)a^j) + \theta_u^i \right\} \right) du \\
- \int_t^T Z_u^i \cdot dW_u.
\]

Then, (10) follows from the definition of \( c^* \). Since all the coefficients are Lipschitz continuous, the wellposedness of the BSDE system and the verification for the control problem is classical, see e.g. [46].

**Remark 2.7:** We recall that \( Y^i \) is the continuation value of the agent when working for the principal \( i \) at time \( t \). The switching of the agent, determined by the optimal intensity

\[ \frac{1}{n-1} a^*(Y_t^j - Y_t^{I_t}), \]

is influenced by principals through their contracts.

**Remark 2.8:** It is noteworthy that the agent problem can be easily solved and enjoys a BSDE representation, thanks to the assumption that the contracts \( (\xi^i) \), are \( \mathcal{F}_T^W \)-measurable. It is crucial that the contracts do not depend on the process \( I \), in other words, the principals have no right to design contracts based on the history of employment of the agent. It may contradict the reality in some applications, but is important for our mathematical reasoning.

**Remark 2.9:** As in [44], the results could also be extended to a more general case, e.g. with discount factors or risk-aversion in the agent’s utility function. With the presence of discount factors, we have another \(-rY^i\) term in (10). Given an invertible utility function \( U_A \), substitute \( \xi^i \) by \( U_A^{-1}(\xi) \) in (10). We need to make the integrability assumptions on \( U_A(\xi^i) \) instead of \( \xi^i \).

Before we continue, let us summarize our model:

- The \( i \)-th principal, \( i \in \{1, \ldots, n\} \), observes only \( W \) and \( X^i \), and in particular, she does not observe for which principal the agent is working, if he is not working for her.
- During the contracting period, the agent can only work for one principal at a time, but can make an effort to increase the intensity of switching to another principal.
- At any time \( t \in [0, T] \), the agent receives a continuous payment from the principal he works for, i.e. \( \theta^i \) if he works for the \( i \)-th principal.
- At the time \( T \), the agent receives a terminal payment from the principal he works for at that time, i.e. \( \xi^i \) if he ends up working for the \( i \)-th principal.
- Each principal \( i \) designs a contract, i.e. a couple \( (\xi^i, \theta^i) \), in order to compensate the agent for his work but most importantly to incentivize the agent to continue working for her, or to come to work for her if he is actually working for another principal \( j \neq i \).
2.2. Principals’ problem: time inconsistency

In the previous section, we managed to represent the value function of the agent by an Itô process (Proposition 2.6). As in [20, 49], we expect that this representation would help us to solve the principals’ problem by the dynamic programming approach. However, in this model, this approach does not work. We shall explain in the case \( n = 2 \) for the simplification of notation.

Consider the set of all contracts

\[
\Xi := \left\{ (\xi^i, \theta^i) \in L^2(\mathbb{P}_0), \quad \theta^i \in H^2(\mathbb{P}_0) \quad \text{and} \quad V^A(\xi, \theta) \geq R \right\},
\]

where \( R \) is the reservation value of the agent for whom only the contracts such that \( V^A(\xi, \theta) \geq R \) are acceptable. Now define

\[
\mathcal{V} := \left\{ (Y^i_0, Z^i) \in L^1(\mathbb{P}_0), \quad Z^i \in H^2(\mathbb{P}_0), \quad i = 1, 2 \right\}.
\]

It follows from Proposition 2.6 that

\[
\Xi = \left\{ (\xi^i, \theta^i) \in H^2(\mathbb{P}_0) \quad \text{and} \quad \xi^i = Y^i T_0, Z^i \theta^i \right\}, \quad \text{where} \quad Y^i T_0, Z^i \theta^i \text{ satisfies}
\]

\[
Y^i T_0 = Y^i 0 - \int_0^T \left( e^i (Y^j t - Y^j t) + \theta^i j \right) dt + \int_0^T Z^i j \cdot dW t, \quad \mathbb{P}_0 - \text{a.s.,}
\]

with \( j \neq i \) and \( (Y^i_0, Z^i) \in \mathcal{V} \). \( \tag{12} \)

For simplicity of notation, if there is no ambiguity, we write simply \( Y^i \) instead of \( Y^i T_0, Z^i \theta^i \). But just keep in mind that the process \( Y^i \) is controlled by \( Z^i \) and \( \theta^i \) with initial value \( Y^i 0 \). The corresponding optimal intensity reads \( \alpha^*_i = (a^*(Y^i t - Y^i t))_{i (i = 2)}, a^*(Y^i t - Y^i t))_{i (i = 1)} \).

Therefore, the principals are searching for a Nash equilibrium so as to maximize:

\[
V^P_0 = \sup_{(\xi^i, \theta^i)} \mathbb{E}^{\mathbb{P}_0} \left[ X^i T - U^i(\xi^i) 1_{\{i = i\}} - \int_0^T U^i(\theta^i) 1_{\{i = i\}} dt \right]
\]

\[
= \sup_{(Y^i_0, Z^i \theta^i)} \mathbb{E}^{\mathbb{P}_0} \left[ -U^i(Y^i t) 1_{\{i = i\}} - \int_0^T \left( U^i(\theta^i) - 1 \right) 1_{\{i = i\}} dt \right] = \sup_{Y^i 0 \geq R} J^i_0(Y^i 0),
\]

where

\[
J^i_0(Y^i 0) := \sup_{(Z^i, \theta^i)} \mathbb{E}^{\mathbb{P}_0} \left[ -U^i(Y^i t) 1_{\{i = i\}} - \int_0^T \left( U^i(\theta^i) - 1 \right) 1_{\{i = i\}} dt \right].
\]

**Remark 2.10:** Regardless the running payment \( \theta^i \), the principal \( i \) can only modify the contract by choosing \( Y^i 0 \) and \( Z^i \). Note that \( Y^i 0 \) represents the ‘expectation’ of the contract, while \( Z^i \) describes its variance (in other word, the risk).
Remark 2.11: If we consider the principal problem (5) with $U_0^i$, we obtain by Itô's formula that

$$ V_0^{P,i} = \sup_{(Y_0^i, Z_i^i, \theta_i^i)} \mathbb{E}^{P_a,*} \left[ -U^i(Y_T^i) 1_{[T = i]} - \int_0^T \left( (U^i(\theta_T^i) - (U_0^i)'(-X_T^i)) 1_{[T = i]} + \frac{1}{2} (U_0^i)''(-X_T^i) \right) dt \right]. $$

Therefore, we have the additional $X$-dependency in the formula. For the sake of simplicity, we only consider the case $U_0(x) = x$.

Up to now, we are applying the same strategy as in [20]. Assume the control problem is time-consistent, that is, admits dynamic programming. Fix $Y_0$ and define the dynamic version of $J_t^i$:

$$ J_t^i(Y_t^i) := \text{ess sup}_{(Z_t^i, \theta_t^i)} \mathbb{E}^{P_a,*}_t \left[ -U^i(Y_T^i) 1_{[T = i]} - \int_t^T (U^i(\theta_s^i) - 1) 1_{[s = i]} ds \right]. $$

Defining $J_t^{i,1} := J_t^i(Y_t^i)|_{l_t = 1}$ and $J_t^{i,2} = J_t^i(Y_t^i)|_{l_t = 2}$ according to the different regimes, we expect to have for $i = 1$

$$ J_t^{1,1} = \text{ess sup}_{(Z_t^i, \theta_t^i)} \mathbb{E}^{P_a,*}_t \left[ -U^i(Y_T^i) 1_{[T > T]} + J_t^{1,2} 1_{[T \leq T]} - \int_t^{T \wedge T} (U^i(\theta_s^i) - 1) ds \right] $$

$$ = \text{ess sup}_{(Z_t^i, \theta_t^i)} \mathbb{E}^{P_a,*}_t \left[ -\beta^1(t, T) U^i(Y_T^i) - \int_t^T \beta^1(t, u)(a^*(Y_u^1 - Y_u^1)) J_t^{1,1} + 1 - U^i(\theta_u^i)) du \right], \quad (13) $$

where

$$ \beta^1(t, s) := \exp \left( -\int_t^s a^*(Y_u^2 - Y_u^1) du \right), $$

and similarly

$$ J_t^{1,2} = \text{ess sup}_{(Z_t^i, \theta_t^i)} \mathbb{E}^{P_a,*}_t \left[ J_t^{1,1} 1_{[T \leq T]} \right] $$

$$ = \text{ess sup}_{(Z_t^i, \theta_t^i)} \mathbb{E}^{P_a,*}_t \left[ \int_t^T \beta^2(t, u)a^*(Y_u^1 - Y_u^2) J_t^{1,1} du \right], \quad (14) $$

where

$$ \beta^2(t, s) := \exp \left( -\int_t^s a^*(Y_u^1 - Y_u^2) du \right). $$

Although it seems promising to solve the system of value functions $(J^{i,1}, J^{i,2})$ as in the agent problem, note that the admissible controls $(Z^i, \theta^i)$ are constrained to be $\mathbb{F}^{X,\alpha}_t$-adapted, in particular, $(Z^i, \theta^i)$ cannot depend on the regimes. Therefore, if the control problem
is time-consistent, then we expect to find the same process \((Z^1, \theta^1)\) solve both the optimizations (13) and (14), which is absurd. It is a contradiction to the time-consistency assumption.

**Remark 2.12:** Besides, the approach for solving the \(n\)-principal problem has another drawback. Note that each contract \(\xi^i\) for \(i \in \mathbb{I}_n\) is a function of the entire vector \(W\), that is, the principal needs to know the other companies’ noise in order to design his own contract. This problem can be avoided once we consider an infinite number of principals (see Section 4), where \(\xi^i\) will be a function of \(W^i\).

In the rest of the paper, we propose two ways to address the time-inconsistency above:

- introduce a time-consistent ‘suboptimal’ solution for the principals;
- consider the case with infinite number of principals using mean-field formulation.

We refer to the recent work of Hernández and Possamaï [35]. They developed a dynamic programming principle for a time-inconsistent agent, and their result is applied to principal-agent model with moral hazard in the thesis of Hernández [34].

### 3. A time-consistent ‘suboptimal’ solution for the principals

#### 3.1. PDE characterization

In this section we assume that the principals only care about maximizing their profits during the period when the agent is hired, i.e. the principal \(i\), instead of solving both optimization problems (13) and (14), only focus on the former. More precisely, we consider the admissible contracts in (12) represented by \(Z^i := (Y^i_0, Z^i, \theta^i)\). The principals now aim at searching for \(\hat{Z}^i := (\hat{Y}^i_0, \hat{Z}^i, \hat{\theta}^i)\) such that

\[
\hat{Z}^i \in \arg\max_{Z^i \in \mathbb{I}} J^i_0 (Z^i; J^i_1, (\hat{Z}^i, Z^j); Z^j)
\]  

(15)

where for \(i \neq j\)

\[
\begin{align*}
J^i_t (Z^i; J^i; Z^j) &= \mathbb{E}^{\mathbb{P}_{\alpha^*}}_t \left[-U^i(Y^i_T)1_{\{\tau_t > T\}} + J^i_\tau 1_{\{\tau_t \leq T\}} - \int_t^{\tau_t \wedge T} \left(U^i(\theta^i_t) - 1\right) \, du\right] \\
&= \mathbb{E}^{\mathbb{P}_{\alpha^*}}_t \left[-\beta^{i,*}(t, T) U^i(Y^i_T) + \int_t^{T} \beta^{i,*}(t, u) \left(a^*(Y^j_u - Y^i_u)J^i_{u} + 1 - U^i(\theta^i_u)\right) \, du\right] \\
&= \mathbb{E}^{\mathbb{P}_{\alpha^*}}_t \left[J^i_{\tau_t 1_{\{\tau_t \leq T\}}} + \int_t^{\tau_t \wedge T} \beta^{i,*}(t, u) a^*(Y^j_u - Y^i_u)J^i_u \, du\right] \quad (16)
\end{align*}
\]

and \(J^i_t (Z^i; J^i; Z^j)\) is the solution to the linear system (16), (17).

The Nash equilibrium of the modified game can be characterized by a highly coupled PDE system. In order to weaken the coupling among the two principals and facilitate the computation of equilibrium, we require the further constraint:

\[
Z^i = (Z^i_1, Z^i_2), \quad \text{with} \quad Z^i_j = 0 \quad \text{for} \ i \neq j, \quad \text{and} |Z^i_i| \leq M \quad \text{for some} \ M > 0. \quad (18)
\]
In order to characterize the Nash equilibrium, we consider the following PDE system:

\[
\begin{align*}
&\partial_t v^{i,j} - \min_{\theta^i} \left\{ \theta^i \partial_t v^{i,j} + U^i(\theta^i) \right\} + \frac{1}{2} M^2 (\partial_{ii}^2 v^{i,j})^+ - c^*(y^j - y^i) \partial_t v^{i,j} \\
&- \left( c^*(y^j - y^i) + \theta^i \right) \partial_j v^{i,j} + \frac{1}{2} |\xi| |z^i|^2 \partial_{jj} v^{i,j} + \alpha^*(y^j - y^i)(v^{i,j} - v^{j,i}) + 1 = 0 \\
&\partial_t v^{j,i} - \left( c^*(y^j - y^i) + \theta^i \right) \partial_j v^{j,i} - \left( c^*(y^j - y^i) + \theta^i \right) \partial_j v^{i,j} \\
&+ \frac{1}{2} |\xi| |z^i|^2 \partial_{jj} v^{j,i} + \frac{1}{2} |\xi| |z^i|^2 \partial_{jj} v^{i,j} + \alpha^*(y^j - y^i)(v^{i,j} - v^{j,i}) = 0 \\
v^{i,j}_T = -U^i(y^j), \quad v^{j,i}_T = 0,
\end{align*}
\]  

(19)

where \( \theta^{i,*} = \arg\min_{\theta^i} \left\{ \theta^i \partial_t v^{i,j} + U^i(\theta^i) \right\} \) and \( z^{i,*} = \arg\max_{|\xi| \leq M} |\xi|^2 \partial_{ii} v^{i,j} = M1_{|\partial_{ii} v^{i,j} \geq 0} \). Further the Nash equilibrium \((\hat{Z}^i)_{i=1,...,n}\) in (15) should satisfy

\[
\hat{\theta}_i^t = \theta^i(t, Y_t), \quad \hat{Z}_i^t = z^i(t, Y_t), \quad \text{and} \quad Y_0^i \in \arg\max_{y} v^i(0, y).
\]

Remark 3.1: Due to the observation in the following Section 4.1 that given a large number of principals, there is a tiny possibility for the agent to return to the same principal after leaving her company. The alternative equilibrium (15) would make more sense provided that there are many principals in the game. Also one may find the reasoning for the constraint \( Z^j_i = 0 \) for \( i \neq j \) in the mean-field analysis in Section 4.3.1.

It is noteworthy that even for this simplified ‘suboptimal’ equilibrium, the corresponding PDE system similar to (19) is non-trivial to solve. In general, for the case with \( n \) principals, one need to solve the PDE system with \( n^2 \) coupled PDE’s and of \( n \) variables. Apparently, the conventional PDE solvers cannot handle the task when \( n \) is large. In the following section we will numerically solve it in case there are only two principals using the explicit Euler scheme.

### 3.2. Numerical tests

In this section, we numerically solve the PDE system (19) for the game involving two principals, using the explicit Euler scheme. In order to further simplify the equations, we assume that the principals do not pay the salaries \( (\theta_i) \) but only the final compensation \( \xi = Y_T \).

In aspect of the coefficients, we adopt the quadratic cost for the hesitation in Example 2.5, namely,

\[ c(a) = \frac{1}{2} a^2 \quad \text{on} \ [0, \overline{K}] \]

and set the constant \( \overline{K} = 1 \), choose the bound of the volatility in (18) to be \( M = 1 \), and consider the time horizon \( T = 0.5 \). As for the utility functions of the principals, we set

\[ U^1(y) = e^y - 1 \quad \text{and} \quad U^2(y) = (e^{0.1y} - 1)/0.1, \]

(20)

so that Principal 1 is more risk-averse than Principal 2.

Recall the function \( c^* \) and \( a^* \) in (9). By solving the PDE system (19), we obtain the (sub-)optimal contract

\[
\xi^i = Y^i_0 - \int_0^T c^*(Y^j_t - Y^i_t) \, dt + \int_0^T M1_{|\partial_{ii} v^{i,j}(t, Y^i_t, Y^j_t) \geq 0} \, dW_t, \quad \text{for} \ i \neq j \in \{1, 2\}.
\]

(21)
Meanwhile, the optimal intensity at time $t$ for the agent to switch from Principal $i$ to Principal $j$ is equal to

$$a^*(Y^j_t - Y^i_t).$$

As a benchmark, we shall look into the situation where there is only one principal in the market. Following the same analysis as before, the principal chooses among the contracts of the form

$$\xi = Y_0 + \int_0^T Z_t \, dW_t.$$  

Since the principal is assumed to be risk-averse, the optimal choice is clearly $Z \equiv 0$, and the value function of the principal reads

$$v(y) = -U(y) + 1 \ast T,$$

where we recall 1 is the drift of the output diffusion $X$. Moreover, since $U$ is strictly increasing, the principal would choose $Y_0$ equal to the reservation value $R$ of the agent in order to maximize her profit.

In Figure 1 we can view the numerical solution to the value functions of the two principals at time 0. The four graphs show the values of functions $v^{1,1}(0, \cdot)$, $v^{1,2}(0, \cdot)$, $v^{2,2}(0, \cdot)$, $v^{2,1}(0, \cdot)$, respectively. Already we can see that the value functions $y^1 \mapsto v^{1,1}(0, y^1, \cdot)$, $y^2 \mapsto v^{2,2}(0, \cdot, y^2)$ are no longer concave. As a result the optimal choice of $Z^i\ast = M^1\{\partial_y^2 v^{ij} \geq 0\}$ is no longer 0 (the optimal choice for a single principal), in other word, the competition does change the strategies of the principals. In order to make this point more clear, we simulate the continuation utility values $(Y^1_t)_{t \in [0,T]}$ and $(Y^2_t)_{t \in [0,T]}$ of the agent under the contacts (21), as well as his decision of switching. By setting $Y^1_0 = 0.3$, $Y^2_0 = -0.5$ and letting the agent start being hired by Principal 2, we obtain the simulation result in Figure 2. As we see, at the beginning Principal 2 sets the volatility in the contract $Z^2\ast$ to be non-zero as an attempt to possibly close the gap between $Y^1$ and $Y^2$ so as to keep the agent working for her and later she sets the volatility $Z^2\ast = 0$ because she would like to allow the agent to leave in order to avoid the terminal payment. Meanwhile since $Y^1$ is always bigger than $Y^2$, the agent is always interested in switching to Principal 1 (in other word, the intensity $a^*(Y^1_t - Y^2_t) > 0$) and eventually he manages it in this simulation.

Further in Figure 3, we observe how the position $y^j$ influences the value function $v^{ij}$ for $i \neq j$. We draw in different colours the value functions given different values of $y^j$, as well as in black the value function of the single principal without competition. Here are some discoveries:

- Most significantly, the value function $v^{ij}$ coincides with the value function without competition, once the value of $y^j$ surpasses that of $y^i$;
- The value function $y^j \mapsto v^{ij}(0, y^j, \cdot)$ is convex only when $y^j < y^i$ and the spread is not large, in particular, it is the scenario where the principals are willing to set the volatility to be non-zero as an attempt to keep the agent working for them, and the more risk-averse principal appears to be more reluctant to submit this risk;
Recall that the principals also choose \((Y^i_0)_{i=1,2}\) to reach an equilibrium such that

\[
Y^1_0 = \arg\max_{y^1} v^{1,1}(0, y^1, Y^2_0) \quad \text{and} \quad Y^2_0 = \arg\max_{y^2} v^{2,2}(0, Y^1_0, y^2).
\]
As shown in Figure 4, we find the equilibria \((Y^i_0)_{i=1,2}\) on the domain \([-1.5, 1.5] \times [-1.5, 1.5]\). Curiously, in this example the equilibria are all on the diagonal \(Y^1_0 = Y^2_0\). Taking into account that the value functions \(v^{1,1}, v^{2,2}\) seem locally concave on the diagonal (see Figure 1), we shall have \(Z^{i,*} = Z^{2,*} = 0\) there. Since in this example \(c^*(0) = 0\), eventually we should have

\[
Y^1_T = Y^1_0 - \int_0^T c^*(Y^2_t - Y^1_t) \, dt + \int_0^T Z^{1,*}_t \cdot dW_t = Y^1_0,
\]

\[
Y^2_T = Y^2_0 - \int_0^T c^*(Y^1_t - Y^2_t) \, dt + \int_0^T Z^{2,*}_t \cdot dW_t = Y^2_0,
\]
that is, both principals would offer constant rewards as in the single principal case. However, the constant reward \( Y^2_0 = Y^2_0 \) may be still higher than the reservation value of the agent.

4. Infinite number of principals: mean field approach

4.1. Heuristic analysis

Heuristically, as \( n \to \infty \) the equation (10), which characterizes the agent’s value function, converges to

\[
dY^i_t = - \left( \int c^*(y - Y^i_t) p_t(dy) + \theta^i_t \right) dt + Z^i_t dW^i_t, \quad p_t = \mathcal{L}(Y^i_t). \tag{22}\]

Besides, it follows from (11) and the definition of the discount factor \( \beta \) that the optimal \( \alpha^* \) and \( \beta^* \) converge to

\[
\alpha^*_j \to 0 \text{ for } j \neq i, \quad \sum_{j \neq i} \alpha^*_j \to \alpha^*_i := \int a^*(y - Y^i_t) p_t(dy), \tag{23}\]

and \( \beta^{n,*}_t \to \beta^*_t := e^{-\int_0^t \alpha^*_s ds} \). In particular, at time \( t \) the agent has \( \alpha^*_i := \int a^*(y - Y^i_t) p_t(dy) \) as intensity to leave the current principal, so his decision of switching is clearly influenced by the continuation value of the contract of the current principal \( Y_t \), and by the distribution of the continuation values of the contracts of all principals \( p_t \).

Remark 4.1: • To be fair, the form of the limit equation (22) does not entirely follow the intuition. Note that we removed the \( e^i \) in front of the term \( Z^i_t dt \) and replaced the stochastic integrator \( dW_t \) by \( dW^i_t \). At this stage, let us admit that once in the drift of the BSDE there is no longer dependence on other \( Y^j (j \neq i) \), the system would be decoupled, i.e. the equation of \( Y^i \) no longer needs the information of \( W^j (j \neq i) \), and it leads to the limit form (22). We will justify the mean field formulation in Section 4.3.

• The first observation in (23) implies that in the limit case once the agent leaves the company \( i \), he will have null probability to come back due to the fact that we have a continuum of players. Therefore, the principal should be only interested in the optimization for the regime where the agent works for her. Remember that in the \( n \)-principal problem, the time-inconsistency is due to the optimizations in the different regimes. Therefore, it is reasonable that the mean field formulation would bypass this difficulty.

Recall the principal’s problem (4). It follows from the observation in Remark 4.1 that

\[
V^P_{0,i} = \begin{cases} 
\mathbb{E}^{\mathbb{P}^{\alpha^*}} [X^i_0 + W_T] = X^i_0, & \text{as } I_0 \neq i, \\
\max_{(\xi^i, \theta^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X^i_{\tau \wedge T} + (X^i_T - U(\xi^i)) \mathbf{1}_{\{\tau > T\}} - \int_0^{\tau \wedge T} U(\theta^i_u) \, du \right], & \text{as } I_0 = i.
\end{cases}
\]

Therefore, in the upcoming mean field game, we should only keep the non-trivial regime \( (I_0 = i) \), i.e. denote

\[
V^P_{0,i} := \max_{(\xi^i, \theta^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ W^i_{\tau \wedge T} + (W^i_T - U(\xi^i)) \mathbf{1}_{\{\tau > T\}} - \int_0^{\tau \wedge T} (U(\theta^i_u) - 1) \, du \right]
\]
It follows from (24) that each principal faces the optimization:

$$\xi \in \Xi(p) = \left\{ \bar{Y}_T : \bar{Y}_T = \bar{Y}_0 - \int_0^T \left( \int c^*(y - \bar{Y}_u) p_u(dy) + \theta_u \right) du + \int_0^T \bar{Z}_u d\bar{W}_u \right\}. $$

It follows from (24) that each principal faces the optimization:

$$V^p_0(p) = \sup_{(\bar{Y}_0, \bar{Z}, \theta)} \mathbb{E}^{\mathbb{P}^\mathbb{F}, \Xi} \left[ \int_0^T \beta^*_u(\alpha^*_u \bar{W}_u - U(\theta_u) + 1) du + \beta^*_T(\bar{W}_T - U(\bar{Y}_T)) \right]. $$

The maximizer $$(\bar{Y}^*_0, \bar{Z}^*, \theta^*)$$ would define a process $\bar{Y}^*$. The law $(p_t)_{t \in [0,T]}$ is a mean field equilibrium, if $p_t = \mathcal{L}(\bar{Y}^*_t)$ for all $t \in [0,T]$.

### 4.2. Mean field game: existence

Our rigorous definition of the mean field game and the argument to prove its existence rely largely on the framework in Lacker [41].

In this part of the paper, we denote the canonical space by $\overline{\Omega} = C([0, T], \mathbb{R}^2)$, the canonical process by $\Theta := (\bar{W}, \bar{Y})$ and the canonical filtration by $\mathbb{F}$. Given $p \in \mathcal{P}(\overline{\Omega})$, define $p_t := p \circ \bar{Y}^{-1}_t \in \mathcal{P}(\mathbb{R})$ and

$$\mathcal{W}(p) := \left\{ \mathbb{P}^{\lambda, \eta} \in \mathcal{P}(\overline{\Omega}) : \bar{W} \text{ is a } \mathbb{P}^{\lambda, \eta}\text{-Brownian motion, and } \mathbb{P}^{\lambda, \eta}\text{-a.s.} \right\}$$

$$d\bar{Y}_t = -\int (c^*(y - \bar{Y}_t)p_t(dy) + \theta_t) dt + \eta_t d\bar{W}_t, \quad \mathcal{L}(\bar{Y}_0) = \lambda,$$

for some $\lambda \in \mathcal{I}$, $\eta \in \mathcal{U}$, $\theta \in \mathbb{H}^2(\mathbb{P}^{\lambda, \eta})$

where, for technical reasons, we define

$$\mathcal{I} := \{ \lambda \in \mathcal{P}(\mathbb{R}) : \lambda \text{ with a compact support } K \text{ in } [R, \infty) \}$$

$$\mathcal{U} := \{ \eta \mathbb{F}\text{-adapted : } \eta \text{ takes values in a compact set } \Sigma \text{ in } \mathbb{R} \}.$$
Assumption 4.2: We assume that $a^* : \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz continuous. $U$ is convex and of $q$-polynomial growth, i.e. there are constants $C < C'$ and $q > 1$ such that

$$C(|\theta|^q - 1) \leq U(\theta) \leq C'(|\theta|^q + 1).$$

Theorem 4.3: Under Assumption 2.4 and 4.2, there exists $p \in \mathcal{P}^c(L_\infty)$, $(\hat{\lambda}, \hat{\eta}) \in \mathcal{I} \times \mathcal{U}$ and an $\hat{\mathbb{F}}$-adapted process $\theta$ such that

$$(\hat{\mathbb{P}}^{\lambda, \eta}, \hat{\theta}) \in \arg\min_{P \in \mathcal{W}(p), \theta} J(\hat{\mathbb{P}}, \theta; p) \quad \text{and} \quad p = \hat{\mathbb{P}}^{\lambda, \eta}.$$

Remark 4.4: • This definition of mean field game via the weak formulation of stochastic control follows the same spirit as those in Carmona and Lacker [16] and Lacker [41]. However, here we also include the control of the initial distribution $\hat{\lambda}$ of $\hat{Y}$.
• It is noteworthy that among the triple of control $(\lambda, \eta, \theta)$ of this mean field game, $\lambda$ takes values of measures, that is, the principals are allowed to play a mixed strategy.
• If we constrain $\theta$ to take values in a compact set in $\mathbb{R}$, then given a bounded function $U$ which is convex on this compact set, the mean field game exists.

We realize that Theorem 4.3 can be proved by essentially the same argument as in [41]. In the rest of the section, we shall outline the strategy of the proof and refer the readers for details to [41].

First, we shall linearize the functional $J$ using the so-called relaxed control. Denote by $\mathcal{D}$ the set of measures $q$ on $[0, T] \times \Sigma \times \mathbb{R}$. Instead of controlling via the processes $\eta$ (taking values in $\Sigma$) and $\theta$ (taking values in $\mathbb{R}$), we shall control through the measure $q$ in the relaxed formulation. The canonical space for the relaxed control becomes $\hat{\Omega} := \Omega \times \mathcal{D}$. Denote the canonical process by $(\hat{W}, \hat{Y}, \hat{\Lambda})$, and the canonical filtration by $\hat{\mathbb{F}}$. Note that one may define a $\mathcal{P}(\Sigma \times \mathbb{R})$-valued $\hat{\mathbb{F}}$-predictable process $(\hat{\Lambda}_t)_{0 \leq t \leq T}$ such that $\hat{\Lambda}(dt, d\eta, d\theta) = \hat{\Lambda}_t(d\eta, d\theta) dt$. Here we abuse the notations using $\eta, \theta$ to represent points in $(\Sigma, \mathbb{R})$.

Denote by $C^\infty_0(\mathbb{R}^2)$ denote the set of infinitely differentiable functions $\phi : \mathbb{R}^2 \mapsto \mathbb{R}$ with compact support. Define the generator $L$ on $C^\infty_0(\mathbb{R}^2)$ by

$$L^{\eta, \phi}(t, x, y) = \left(1, -\int c^\ast(y' - y)p_t(dy') - \theta_t\right) \nabla \phi + \frac{1}{2}\left(\partial_{xx}\phi + \eta^2\partial_{yy}\phi + 2\eta\partial_{xy}\phi\right),$$

for $(t, x, y, p, \eta) \in [0, T] \times \mathbb{R}^2 \times \mathcal{P}(\hat{\Omega}) \times \Sigma$. Further define

$$M^{\rho, \phi}_t := \phi(\bar{X}_t, \bar{Y}_t) - \int_0^t \int_{\Sigma \times \mathbb{R}} L^{\rho, \phi}(s, \bar{X}_s, \bar{Y}_s) \Lambda_s(d\eta)ds.$$

Definition 4.5: Given $p \in \mathcal{P}(\hat{\Omega})$, define the set of the controlled martingale problems:

$$\mathcal{R}(p) = \left\{\hat{\mathbb{P}} \in \mathcal{P}(\hat{\Omega}) : \hat{\mathbb{P}}\text{-a.s. } \bar{Y}_0 \sim \lambda, \text{ for some } \lambda \in \mathcal{I}, \right.$$ 

$$\left.\mathbb{E}^{\hat{\mathbb{P}}_\lambda}\left[\int_0^T |\Lambda_t|^q dt\right] < \infty, \right\}$$

$$M^{\rho, \phi}$$ is a $\hat{\mathbb{P}}$-martingale for each $\phi \in C^\infty_0(\mathbb{R}^2)$. (26)
where \(|\Lambda_t|^q := \int_{\Sigma \times \mathbb{R}} |(\eta, \theta)|^q \Lambda_t(d\eta, d\theta)|.

Further, in the relaxed formulation the object function of the principals reads:
\[
\hat{j}(\mathbb{P}; p) := \mathbb{E}^{\mathbb{P}} \left[ \int_0^T \int_{\mathbb{R}} \beta^*_u (\alpha^*_u \mathbb{W}_u - U(\theta) + 1) \Lambda_u(d\theta) \, du + \beta^*_T (\mathbb{W}_T - U(\mathbb{Y}_T)) \right],
\]
and define the set of the optimal control:
\[
\mathcal{R}^*(p) := \arg \max_{\mathbb{P} \in \mathcal{R}(p)} \hat{j}(\mathbb{P}; p).
\]
We say \(\mathbb{P} \in \mathcal{P}(\hat{\Omega})\) is a relaxed mean field game if \(\mathbb{P} \in \mathcal{R}^*(\mathbb{P} \circ (X, Y)^{-1})\).

Theorem 4.6: Under Assumption 2.4 and 4.2, there exists a relaxed mean field game.

Proof: Our setting slightly distinguishes from the one in [41], because we allow to control the initial law \(\lambda\) of \(\mathbb{Y}\). However, since we constrain the choice of \(\lambda\) among the distribution in \(\mathcal{P}(K)\) where \(K\) is a compact set in \(\mathbb{R}\), we are still able to prove the tightness of \(\mathcal{R}(p)\) (in the case \(\Lambda\) is truncated), and the rest of the argument would follow the same lines in [41, Section 4 and 5].

It remains to construct a strict mean field game (as in Theorem 4.3) based on a relaxed one. Again we can follow the classical argument. We shall only provide the sketch of the proof, for more details we refer the readers to [41, Proof of Theorem 3.7].

Proof of Theorem 4.3: Let \(\hat{\mathbb{P}}\) be a relaxed mean field game. First, by the same argument as in the proof of [41, Proof of Theorem 3.7], we may find an \(\hat{\mathbb{P}}\)-adapted process \(\hat{q} : [0, T] \times \hat{\Omega} \rightarrow \mathcal{P}(\Sigma \times \mathbb{R})\) such that
\[
\hat{q}(t, \mathbb{W}, \mathbb{Y}) = \mathbb{E}^{\hat{\mathbb{P}}} [\Lambda_t | \hat{\mathcal{F}}_t], \quad \hat{\mathbb{P}}\text{-a.s. } t \in [0, T].
\]
Further, for each \((t, w, y, p) \in [0, T] \times \hat{\Omega} \times \mathcal{P}(\hat{\Omega})\), since \(\alpha^*, \beta^*\) are always non-negative and function \(U\) is convex, the following subset
\[
K(t, \theta, y, p) := \{ (\eta^2, \theta, l) : \eta \in \Sigma, l \leq f(t, x, y, p, \theta) := \beta^*_t(y, p)(\alpha^*_t(y, p)\theta_t - U(\theta)) \}
\]
is convex. This verifies the ‘convexity’ assumption in [41, Assumption (Convex)]. Therefore, using the measurable selection result in [33, Lemma 3.1], there exist \(\hat{\mathbb{P}}\)-adapted processes \(\hat{\eta}, \hat{\theta}\) and \(\hat{l} \geq 0\) such that
\[
\int \left( \eta^2, f(t, x, y, p, \theta) \right) \hat{q}(t, x, y)(d\eta, d\theta) = \left( \hat{\eta}^2(t, x, y), f(t, x, y, p, \hat{\theta}(t, x, y)) - \hat{l}(t, x, y) \right).
\]
Further, it would be easy to verify that \(p := \hat{\mathbb{P}} \circ (\mathbb{W}, \mathbb{Y})^{-1} = \mathbb{P}^{\lambda, \hat{\eta}}\) as well as \(J(\mathbb{P}^{\lambda, \hat{\eta}}, \hat{\theta}; p) \geq \hat{j}(\mathbb{P}; p)\). Therefore, we find a mean field game in sense of Theorem 4.3.

From the sketch of proof, we may observe the following.
Corollary 4.7: Under Assumption 2.4 and 4.2, there exists a mean field game such that

\[ \mathbb{E}^\mathbb{P}_{\hat{\lambda}, \hat{\eta}} \left[ \int_0^T |\hat{\theta}_u|^q \, du \right] < \infty. \]

Proof: Note that in the proof of Theorem 4.3, the \( \hat{\theta} \) we constructed satisfies

\[ C(|\hat{\theta}|^q - 1) \leq U(\hat{\theta}) \leq \int_{\mathbb{R}} U(\theta) \hat{q}_t(\theta) \, d\theta, \]

where \( \hat{q} \) is defined in (27). Therefore,

\[ C \mathbb{E}^\mathbb{P}_{\hat{\lambda}, \hat{\eta}} \left[ \int_0^T |\hat{\theta}_u|^q \, du \right] \leq C + \mathbb{E}^\mathbb{P} \left[ \int_0^T \int_{\mathbb{R}} U(\theta) \Lambda_u(\theta) \, d\theta \, du \right] \]

\[ \leq C + \mathbb{E}^\mathbb{P} \left[ \int_0^T \int_{\mathbb{R}} C'(|\theta|^q + 1) \Lambda_u(\theta) \, d\theta \, du \right]. \]

Then, the desired result follows from (26).

4.3. Mean field game: approximation

In this section, we shall justify our mean field formulation, that is, answer the question why the principals would apply the mean field game we studied in the previous section. Unlike the classical cases, recall that in our \( n \)-player problem the stochastic control problem is time-inconsistent. Therefore, it would be difficult to verify whether a contract provided by the mean field game would be \( \varepsilon \)-optimal for the \( n \)-player game. Instead, we shall verify that using a contract provided by the mean field game, the principals would receive, in a \( n \)-player game, a utility \( \varepsilon \)-close to the value function calculated in the mean field game.

4.3.1. Agent problem: backward propagation of chaos

Here we shall analyse the agent’s behaviour, once he is given \( n \) contracts provided by the mean field game. Note that the mean field contract has the form:

\[ \xi = \bar{Y}_0 - \int_0^T \left( \int_{\mathbb{R}} c^*(y - \bar{Y}_u) p_u^*(dy) + \theta_u \right) \, du + \int_0^T \eta_u \, d\bar{W}_u. \] \hspace{1cm} (28)

In this section, we shall assume that \( \xi \) is \( \mathcal{F}_{\mathcal{T}, \mathcal{T}}^\mathcal{W}, \mathcal{Y}_0 \)-measurable and \( \theta \) is \( \mathcal{F}_{\mathcal{T}, \mathcal{T}}^\mathcal{W}, \mathcal{Y}_0 \)-adapted. In this case, we may write \( \xi(\bar{W}, \mathcal{Y}_0) \) and \( \theta(\bar{W}, \mathcal{Y}_0) \). Now, recall the canonical space \( \Omega \) in the setting of \( n \)-principal problem, and the canonical process \( W = \{ W^i \}_{i \in \mathbb{N}} \) representing the outputs of all principals. Further, let \( \{(Y^*_t, \xi^t, \theta^t, \eta^t)\}_{i \in \mathbb{N}} \) be the \( n \) independent copies of \( (\bar{Y}, \xi, \theta, \eta) \) such that

\[ Y^*_{t,i} = \xi^i + \int_t^T \left( \int_{\mathbb{R}} c^*(y - Y^*_{u,i}) p_u^*(dy) + \theta^i_u \right) \, du - \int_t^T \eta^i_u \, dW^i_u \]

\[ = \xi^i + \int_t^T \left( \int_{\mathbb{R}} c^*(y - Y^*_{u,i}) p_u^*(dy) + \theta^i_u \right) \, du - \int_t^T Z^*_{u,i} \cdot dW_u. \] \hspace{1cm} (29)
with \( Z^{*i} := c_{H}^{i} = (0, \ldots, 0, \eta^{i}, 0, \ldots, 0)^{\top} \).

In this section, we shall simply use the notation \( \mathbb{E} \) instead of \( \mathbb{E}_{\mathbb{P}^{\xi, \eta}} \). The following estimate follows directly from Corollary 4.7.

**Lemma 4.8:** Assume \( q \geq 2 \) in Assumption 4.2. Then we have

\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_{t}^{*i}| \right] < \infty.
\]

In particular \( \mathbb{E}[|\xi^{i}|^2] < \infty \).

Given such contracts \( \{(\xi^{i}, \theta^{i})\}_{i \in \mathbb{I}_{n}} \), the agent would solve the system of BSDE (10), namely,

\[
Y_{t}^{ni} = \xi^{i} + \int_{t}^{T} \left( \frac{1}{n-1} \sum_{j=1}^{n} c^{*}(Y_{u}^{nj} - Y_{u}^{ni}) + \theta^{i}_{u} \right) du - \int_{t}^{T} Z_{u}^{ni} \cdot dW_{u}
\]

\[
= \xi^{i} + \int_{t}^{T} \left( \frac{n}{n-1} \int_{\mathbb{R}} c^{*}(y - Y_{u}^{ni}) p_{u}^{n}(dy) - \frac{c^{*}(0)}{n-1} + \theta^{i}_{u} \right) du - \int_{t}^{T} Z_{u}^{ni} \cdot dW_{u},
\]

where \( p^{n} \) is the empirical measure \( \frac{1}{n} \sum_{j=1}^{n} \delta_{Y_{u}^{nj}} \) and \( p_{u}^{n} := \frac{1}{n} \sum_{j=1}^{n} \delta_{Y_{u}^{nj}} \). Define \( \Delta Y^{i} := Y_{u}^{ni} - Y^{*i}, \quad \Delta Z^{i} := Z_{u}^{ni} - Z^{*i} \). Let \( C := C([0, T]; \mathbb{R}) \). Denote the square of the Wasserstein-2 distance on \( \mathcal{P}^{2}(C) \) by

\[
d_{2}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{C \times C} \sup_{t \leq u \leq T} |x_{u} - y_{u}|^2 \pi(dx, dy).
\]

Here is the main result concerning the agent’s problem.

**Proposition 4.9:** Assume that Assumption 2.4 and 4.2 hold true for some \( q \geq 2 \). Then we have

\[
\lim_{n \to \infty} \mathbb{E} \left[ d_{2}^{2} (p^{n}, p^{*}) \right] = 0.
\]

**Remark 4.10:** There is a certain similarity between the previous result and the one in Buckdahn, Djehiche, Li, and Peng [13] where the authors also study the convergence from a \( n \)-player BSDE to the mean field limit equation in the form of (29). To our understanding of their paper, the solution to the \( n \)-player BSDE there is a fixed point of the following map:

\[
p \mapsto Y(p) \mapsto \frac{1}{n} \sum_{i=1}^{n} \delta_{Y^{i}(p)},
\]

where \( Y(p) \) is the solution of the BSDE

\[
Y_{t} = \xi + \int_{t}^{T} \left( \int_{\mathbb{R}} c^{*}(y - Y_{u}) p_{u}(dy) + \theta_{u} \right) du - \int_{t}^{T} Z_{u} dW_{u},
\]

(32)
and \( \{Y^i(p)\}_{i \in I_n} \) are the independent copies of \( Y(p) \). In other words, their formulation of the \( n \)-player problem is via the open loop while ours is via the closed loop.

- As we will show, the technique involved to prove Proposition 4.9 is a combination of the BSDE estimates and the argument for proving the propagation of chaos. That is why we would name this section the backward propagation of chaos. When we revise the paper, we note that this topic has already attracted some following research, see [43].

- Indeed the result of Proposition 4.9 and the upcoming analysis hold valid for the broader class of BSDE systems in the form of

\[
Y^n_i = \xi^i + \int_t^T F(t, Y^n_i, Z^n_i, p^n_t) \, dt - Z^n_i \, dW_t, \quad i \in \mathbb{I}_n,
\]

where \( F \) is Lipschitz continuous in \((y, z, p)\). Apparently, this result could have an independent interest, once one wants to look into the asymptotic behaviour of solutions to such BSDE systems.

Before proving this main result, we first obtain the following estimates through some classical BSDE arguments.

**Lemma 4.11**: We have

\[
\mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y^j_s|^2 \right] \leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) \, du \right] + \frac{1}{n^2} \right),
\]

(33)

\[
\mathbb{E} \left[ \int_t^T |\Delta Z^*_u|^2 \, du \right] \leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) \, du \right] + \frac{1}{n^2} \right).
\]

(34)

**Proof**: Comparing (29) and (30) we obtain

\[
|\Delta Y^i_t| \leq \frac{n}{n-1} \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y^i_u) p^n_u(dy) - \int_{\mathbb{R}} c^*(y - Y^i_u) p^*_u(dy) \right| \, du + \frac{|c^*(0)|}{n-1} T
\]

\[
+ \frac{1}{n-1} \int_t^T \int_{\mathbb{R}} c^*(y - Y^i_u) p^*_u(dy) \, du + \int_t^T \Delta Z^*_u \cdot dW_u |.
\]

(35)

Using the Kontorovich duality and Lipschitz continuity of \( c^* \), we have

\[
\left| \int_{\mathbb{R}} c^*(y - Y^i_u) p^n_u(dy) - \int_{\mathbb{R}} c^*(y - Y^i_u) p^*_u(dy) \right|
\]

\[
\leq \left| \int_{\mathbb{R}} c^*(y - Y^i_u) p^n_u(dy) - \int_{\mathbb{R}} c^*(y - Y^i_u) p^*_u(dy) \right|
\]

\[
+ \left| \int_{\mathbb{R}} c^*(y - Y^i_u) p^*_u(dy) - \int_{\mathbb{R}} c^*(y - Y^*_u) p^*_u(dy) \right|
\]

\[
\leq \mathbb{W}_1(p^n_u, p^*_u) + L|\Delta Y^i_t|
\]

\[
\leq \mathbb{W}_2(p^n_u, p^*_u) + L|\Delta Y^i_u|.
\]

(36)
Squaring and taking supremum and expectation on both sides of (35), and using Jensen and BDG-inequality, we obtain

\[
\mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y^i_s|^2 \right] \leq C_T \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2 (p^n_{u}, p^*_u) \, du \right] + C_{L,T} \mathbb{E} \left[ \int_t^T |\Delta Y^i_u|^2 \, du \right] + \frac{C_T}{n^2} |c^*(0)|^2
\]

\[+ C_T \mathbb{E} \left[ \sup_{t \leq s \leq T} \left| \int_t^s \Delta Z^i_u \cdot dW_u \right|^2 \right]
\]

\[+ \frac{C_T}{n^2} \mathbb{E} \left[ \int_t^T \left| \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) \right|^2 \, du \right]
\]

\[\leq C_T \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2 (p^n_{u}, p^*_u) \, du \right] + C_{L,T} \mathbb{E} \left[ \int_t^T |\Delta Y^i_u|^2 \, du \right] + \frac{C_T}{n^2} |c^*(0)|^2
\]

\[+ C_{T,BDG} \mathbb{E} \left[ \int_t^T |\Delta Z^i_u|^2 \, du \right]
\]

\[+ \frac{C_T}{n^2} \mathbb{E} \left[ \int_t^T \left| \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) \right|^2 \, du \right]. \tag{37}
\]

Further we shall estimate \(\mathbb{E} \left[ \int_t^T |\Delta Z^i_u|^2 \, du \right] \). By Itô's formula,

\[
|\Delta Y^i_t|^2 + \int_t^T |\Delta Z^i_u|^2 \, du
\]

\[= \frac{2n}{n - 1} \int_t^T \Delta Y^i_u \left( \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) - \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) \right) \, du
\]

\[+ \frac{2}{n - 1} \int_t^T \Delta Y^i_u \, du - 2 \int_t^T \Delta Y^i_u \Delta Z^i_u \cdot dW_u
\]

\[+ \frac{2}{n - 1} \int_t^T \Delta Y^i_u \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) \, du.
\]

Together with (36) and Young's inequality, we obtain

\[
\mathbb{E} \left[ \int_t^T |\Delta Z^i_u|^2 \, du \right]
\]

\[\leq 2 \mathbb{E} \left[ \int_t^T |\Delta Y^i_u| \left( \mathcal{W}_2 (p^n_{u}, p^*_u) + C_L |\Delta Y^i_u| \right) \right] + \frac{|c^*(0)|^2}{(n - 1)^2} T
\]

\[+ 2 \mathbb{E} \left[ \int_t^T |\Delta Y^i_u|^2 \, du \right] + \frac{1}{(n - 1)^2} \mathbb{E} \left[ \int_t^T \left| \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) \right|^2 \, du \right]
\]

\[\leq (3 + 2C_L) \mathbb{E} \left[ \int_t^T |\Delta Y^i_u|^2 \, du \right] + \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2 (p^n_{u}, p^*_u) \, du \right]
\]

\[+ \frac{|c^*(0)|^2}{(n - 1)^2} T + \frac{1}{(n - 1)^2} \mathbb{E} \left[ \int_t^T \left| \int \frac{c^*(y - Y^{*,i})}{\Delta 1} p^*_u(dy) \right|^2 \, du \right]. \tag{38}
\]
We now estimate the last term above,

\[
\mathbb{E} \left[ \int_t^T \left( \int_\mathbb{R} c^* (y - Y^*_{u,i} p^*_u(dy) \right)^2 \, du \right]
\]

\[
\leq \mathbb{E} \left[ \int_t^T \left( \int_\mathbb{R} (L|y| + L|Y^*_{u,i}| + |c^*(0)|) p^*_u(dy) \right)^2 \, du \right]
\]

\[
\leq \mathbb{E} \left[ \int_t^T (\mathbb{E}[|Y^*_{u,i}|] + L|Y^*_{u,i}| + |c^*(0)|)^2 \, du \right]
\]

\[
\leq 6L^2 T \mathbb{E} \left[ \sup_{0 \leq u \leq T} |Y^*_{u,i}|^2 \right] + 3T|c^*(0)|^2,
\]

which is bounded from above by a constant \(C_0\), independent of \(t\) and \(i\), by the a priori estimate for \(Y^*\) in Lemma 4.8. Together with (37) we obtain

\[
\mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y_s^i|^2 \right] \leq (C_T + C_{T,\text{BDG}}) \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p^n_u, p^*_u) \, du \right]
\]

\[
+ \left( C_L T + (2C_L + 3)C_{T,\text{BDG}} \right) \int_t^T \mathbb{E} \left[ \sup_{u \leq s \leq T} |\Delta Y_s^i|^2 \right] \, du
\]

\[
+ \frac{1}{n^2} (C_T + 4C_{T,\text{BDG}} T) c^*(0)^2 + \frac{1}{n^2} (C_T + 4C_{T,\text{BDG}}) C_0.
\]

Applying Grönwall inequality, we get

\[
\mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y_s^i|^2 \right] \leq C \left( \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p^n_u, p^*_u) \, du \right] + \frac{1}{n^2} \right)
\]

\[
\leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) \, du \right] + \frac{1}{n^2} \right),
\]

that is, the estimate (33), for some constant depending on \(T, L, C_0\) and the constant from BDG inequality. Finally, the estimate (34) follows from (38) and (33). 

\[\blacksquare\]

**Proof of Proposition 4.9:** Define the empirical measure,

\[\nu^n := \frac{1}{n} \sum_{i=1}^n \delta_{Y^*_{n,i}}.\]

The empirical measure \(\frac{1}{n} \sum_{i=1}^n \delta_{(Y^*_{n,i}, Y^*_{n,i})}\) is a coupling of the empirical measures \(p^n\) and \(\nu^n\), so

\[
d^2_t(p^n, \nu^n) \leq \frac{1}{n} \sum_{i=1}^n \sup_{1 \leq u \leq T} |\Delta Y^i_u|, \quad a.s.
\]

Together with (33), we obtain that

\[
\mathbb{E} \left[ d^2_t(p^n, \nu^n) \right] \leq C \left( \mathbb{E} \left[ \int_t^T d^2_u(p^n, p^*) \, du \right] + \frac{1}{n^2} \right).
\]
Apply the triangle inequality and the previous inequality to obtain
\[
\mathbb{E}[d_t^2(p^n, p^*)] \leq 2\mathbb{E}[d_t^2(p^n, v^n)] + 2\mathbb{E}[d_t^2(v^n, p^*)] \\
\leq 2C\mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) \, du \right] + 2\mathbb{E}[d_0^2(v^n, p^*)] + \frac{2C}{n^2}.
\]
Using Grönwall’s inequality we obtain
\[
\mathbb{E}[d_0^2(p^n, p^*)] \leq 2e^{2CT} \left( \mathbb{E}[d_0^2(v^n, p^*)] + \frac{C}{n^2} \right).
\]
Each \(v^n\) is the empirical measures of i.i.d. samples from the law \(p^*\), (31) follows from the law of large numbers. \(\blacksquare\)

### 4.3.2. Principals’ problem

For technical reasons, in this section we would consider the case where \(\theta\) takes values in a compact set in \(\mathbb{R}\) and \(U\) is convex and bounded on this set. Recall the third point in Remark 4.4, under the above setting, we still have existence of the mean field games.

**Proposition 4.12:** Let all the \(n\) principals offer the contract provided by a mean field game, for principal \(i\) the reward becomes
\[
V_{n,i}^0 = \int_K \lambda(dy_0) \mathbb{E}^{\mathbb{P}_r^*}\left[ X_T^i - U(\xi^i)1_{\{\tau_T^i = t\}} - \int_0^T U(\theta^i_u)1_{\{I_u = i\}} \, du \right], \tag{39}
\]
where \(\alpha^*\) is the optimal intensity of the agent satisfying (11). Then, as \(n \to \infty\), \(V_{n,i}^0\) converges to the value of the mean field game.

**Proof:** The dynamic version of (39) reads
\[
V_{n,i}^t = \mathbb{E}^{\mathbb{P}_r^*}\left[ X_T^i - U(\xi^i)1_{\{\tau_T^i = t\}} - \int_t^T U(\theta^i_u)1_{\{I_u = i\}} \, du \right], \quad \text{for } 0 < t \leq T.
\]
Since \(U\) is bounded
\[
\mathbb{E}^{\mathbb{P}_r^*}\left[ \sup_{0 \leq t \leq T} |V_{n,i}^t|^2 \right] < \infty. \tag{40}
\]
We denote \(V_{n,i,0} := V_{n,i}^t|_{\tau_T^i \neq t}\) and \(V_{n,i,1} := V_{n,i}^t|_{\tau_T^i = t}\) on different regimes. Recall \(\tau_T^i = \inf\{s \geq t : I_s \neq I_t\}\) and define \(\bar{\tau}_T^i := \inf\{s \geq t : I_s = I_t\}\). By the tower property of the conditional expectation, we have
\[
\begin{align*}
V_{n,i,1}^t &= \mathbb{E}^{\mathbb{P}_r^*}_{\tau_T^i} \left[ V_{n,i,0}^t 1_{\{\tau_T^i \leq T\}} + (X_T^i - U(\xi^i))1_{\{\tau_T^i > T\}} - \int_t^{\tau_T^i \wedge T} U(\theta^i_u) \, du \right] \\
V_{n,i,0}^t &= \mathbb{E}^{\mathbb{P}_r^*}_{\tau_T^i} \left[ V_{n,i,1}^t 1_{\{\tau_T^i \leq T\}} + X_T^i 1_{\{\tau_T^i > T\}} \right] = \mathbb{E}^{\mathbb{P}_r^*}_{\tau_T^i} \left[ (V_{n,i,1}^t - X_T^i)1_{\{\tau_T^i > T\}} \right] + X_T^i.
\end{align*}
\] (41)
It follows from (11) that \( \lim_{n \to \infty} \mathbb{P}^{\pi_n} [ \tau^n_t \leq T ] = 0 \). Together with (40), we obtain
\[
\lim_{n \to \infty} \mathbb{E}^{\pi_n} \left[ | V^n_{t,0} - X_t |^2 \right] = 0. \tag{42}
\]

Define
\[
\tilde{V}^{n,i,1}_0 = \int_K \lambda(\mathrm{d}y_0) \mathbb{E}^{\pi_n} \left[ \int_0^T \beta^n_u \left( \alpha^n_u W^n_i - U(\theta^n_u) + 1 \right) \mathrm{d}u + \beta^n_T (W^n_T - U(\tilde{\xi}^i)) \right],
\]
where \( \alpha^n_t = \frac{1}{n} \sum_{j \neq i} \alpha^*(Y^n_{t,j} - Y^n_{t,i}) \) and \( \beta^n_t := \exp(- \int_0^t \alpha^n_u \mathrm{d}u) \). Note that \( \alpha^*, \beta^* \) are bounded. By (42), we have
\[
\lim_{n \to \infty} | \tilde{V}^{n,i,1}_0 - V^n_{0,i,1} | = 0.
\]

Further, on the regime \( I = i \), the expectation \( \int_K \lambda(\mathrm{d}y_0) \mathbb{E}^{\pi_n} [ \cdot ] \) coincides with \( \mathbb{E}( = \mathbb{E}^{\tilde{\lambda}, \tilde{\eta}}) \), so
\[
\tilde{V}^{n,i,1}_0 = \mathbb{E} \left[ \int_0^T \beta^n_u \left( \alpha^n_u W^n_i - U(\theta^n_u) + 1 \right) \mathrm{d}u + \beta^n_T (W^n_T - U(\tilde{\xi}^i)) \right].
\]
Since the function \( a^* \) is bounded and Lipschitz continuous, we have
\[
| \beta^n_t - \beta^*_t |^2 + | \alpha^n_t \beta^n_t - \alpha^*_t \beta^*_t |^2 \leq C \left( d^2_0(p^n, p^*) + | Y^n_{t,i} - Y^*_n_{t,i} |^2 \right).
\]

Finally, the convergence result (31) and the estimate (33) imply that
\[
\lim_{n \to \infty} \tilde{V}^{n,i,1}_0 = \mathbb{E} \left[ \int_0^T \beta^*_u \left( \alpha^*_u W^n_i - U(\theta^*_u) + 1 \right) \mathrm{d}u + \beta^*_T (W^n_T - U(Y^*_n_T)) \right],
\]
and the latter is the value of the mean field game. \( \blacksquare \)

**Note**

1. We denote by \( \mathbb{H}^2(\mathbb{P}_0) \) the space of adapted processes \( w \) such that
\[
\| \theta \|_{\mathbb{H}^2(\mathbb{P}_0)}^2 := \mathbb{E}^{\mathbb{P}_0} \left[ \left( \int_0^T | \theta_u |^2 \mathrm{d}u \right)^{\frac{1}{2}} \right] < \infty.
\]

**Acknowledgments**

The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Programme.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).
References

[1] R. Aid, D. Possamai, and N. Touzi, Optimal electricity demand response contracting with responsiveness incentives, to appear in Mathematics of Operations Research (2022).
[2] C. Alasseur, I. Ekeland, R. Elie, N. Hernández Santibáñez, and D. Possamai, An adverse selection approach to power pricing, SIAM J. Control Optim. 58(2) (2020), pp. 686–713.
[3] C. Alasseur, H. Farhat, and M. Sagué, A principal-agent approach to capacity remuneration mechanisms, Int. J. Theor. Appl. Finance 23(8) (2020), pp. 1–64.
[4] A. Aurell, R. Carmona, G. Dayanikli, and M. Lauriere, Optimal incentives to mitigate epidemics: A Stackelberg mean field game approach. SIAM Journal on Control and Optimization 60 (2022), pp. S294–S322.
[5] B. Baldacci, I. Manziuk, T. Mastrolia, and M. Rosenbaum, Market making and incentives design in the presence of a dark pool: A deep reinforcement learning approach. Preprint (2019).
[6] B. Baldacci, D. Possamai, and M. Rosenbaum, Optimal make-take fees in a multi market-maker environment, SIAM J. Financial Math. 12(1) (2021), pp. 446–486.
[7] D. Baron, Noncooperative regulation of a nonlocalized externality, Rand. J. Econ. 16(4) (1985), pp. 553–568.
[8] B.D. Bernheim and M.D. Whinston, Common agency, Econometrica 54 (1986), pp. 923–942.
[9] B.D. Bernheim and M.D. Whinston, Common marketing agency as a device for facilitating collusion, RAND. J. Econ. 16 (1986), pp. 269–281.
[10] P. Bolton and M. Dewatripont, Contract Theory, MIT Press, 2005.
[11] B. Bouchard, A stochastic target formulation for optimal switching problems in finite horizon, Stochastics 81(2) (2009), pp. 171–197.
[12] A. Braverman and E. Stiglitz, Sharecropping and the interlinking of agrarian markets, Amer. Econ. Rev. 72 (1982), pp. 695–715.
[13] R. Buckdahn, B. Djehiche, J. Li, and S. Peng, Mean-Field backward stochastic differential equations: A limit approach, Ann. Probab. 37(4) (2009), pp. 1524–1565.
[14] R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications I, Springer International Publishing, 2018.
[15] R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications II, Springer International Publishing, 2018.
[16] R. Carmona and D. Lacker, A probabilistic weak formulation of mean field games and applications, Ann. Appl. Probab. 25(3) (2015), pp. 1189–1231.
[17] R. Carmona and P. Wang, Finite-State contract theory with a principal and a field of agents, Manage. Sci. 67(8) (2021), pp. 4725–4741.
[18] J.F. Chassagneux, R. Elie, and I. Kharroubi, A note on existence and uniqueness for solutions of multidimensional reflected BSDEs, Electron. Commun. Probab. 16 (2011), pp. 120–128.
[19] J. Cvitanić, D. Possamai, and N. Touzi, Moral hazard in dynamic risk management, Manage. Sci. 63(107) (2017), pp. 3328–3346.
[20] J. Cvitanić, D. Possamai, and N. Touzi, Dynamic programming approach to principal-agent problems, Finance Stoch. 22(1) (2018), pp. 1–37.
[21] J. Cvitanić and H. Xing, Asset pricing under optimal contracts, J. Econom. Theory 173 (2018), pp. 142–180.
[22] J. Cvitanić and J. Zhang, Contract theory in continuous-time models. Springer Finance, Springer, Heidelberg, 2013.
[23] A. Dixit, G.M. Grossman, and E. Helpman, Common agency and coordination: General theory and application to government policy making, J. Political Econ. 105 (1997), pp. 752–769.
[24] N. El Karoui, S. Peng, and M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7(1) (1997), pp. 1–71.
[25] R. Elie, E. Hubert, T. Mastrolia, and D. Possamaï, Mean-field moral hazard for optimal energy demand response management, Math. Finance 31(1) (2021), pp. 399–473.
[26] R. Elie and I. Kharroubi, Probabilistic representation and approximation for coupled systems of variational inequalities, Statist. Probab. Lett. 80(17-18) (2010), pp. 1388–1396.
[27] R. Elie, T. Mastrolia, and D. Possamaï, A tale of a principal and many, many agents, Math. Oper. Res. 44(2) (2019), pp. 440–467.
[28] R. Elie and D. Possamai, Contracting theory with competitive interacting agents, SIAM J. Control Optim. 57(2) (2019), pp. 1157–1188.
[29] O.E. Euch, T. Mastrolia, M. Rosenbaum, and N. Touzi, Optimal make-take fees for market making regulation, Math. Finance 31(1) (2021), pp. 109–148.
[30] M. Fuhrman and M.A. Morlais, Optimal switching problems with an infinite set of modes: An approach by randomization and constrained backward SDEs, Stoch. Process. Appl. 130(5) (2020), pp. 3120–3153.
[31] S. Hamadène and J.P. Lepeltier, Backward equations, stochastic control and zero-sum stochastic differential games, Stoch. Stoch. Rep. 54(3-4) (1995), pp. 221–231.
[32] S. Hamadène and J. Zhang, Switching problem and related system of reflected backward SDEs, Stoch. Process. Appl. 120(4) (2010), pp. 403–426.
[33] U.G. Haussmann and J.P. Lepeltier, On the existence of optimal controls, SIAM J. Control Optim.28(4)(1990),pp.851–902.
[34] C. Hernández, Me, myself and I: Time-inconsistent stochastic control, contract theory and backward stochastic Volterra integral equations, Ph.D. thesis., Columbia University, New York, USA, 2021.
[35] C. Hernández and D. Possamaï, Me, myself and I: A general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. To appear in Ann. Appl. Probab. (2022).
[36] B. Holmström and P. Milgrom, Aggregation and linearity in the provision of intertemporal incentives, Econometrica 55(2) (1987), pp. 303–328.
[37] Y. Hu and S. Tang, Multi-dimensional BSDE with oblique reflection and optimal switching, Probab. Theory. Relat. Fields. 147(1-2) (2010), pp. 89–121.
[38] E. Hubert, T. Mastrolia, D. Possamai, and X. Warin, Incentives, lockout, and testing: From Thucydides’s analysis to the COVID-19 pandemic. To appear in J. Math. Biol. (2022).
[39] J. Jacod, Multivariate point processes: Predictable projection, radon-Nikodym derivatives, representation of martingales, Z. Wahrscheinlichkeitstheorie Verw 31(3) (1975), pp. 235–253.
[40] H.K. Koo, G. Shim, and J. Sung, Optimal multi-agent performance measures for team contracts, Math. Finance 18(4) (2008), pp. 649–667.
[41] D. Lacker, Mean field games via controlled martingale problems: Existence of markovian equilibria, Stoch. Process. Their. Appl. 125(7) (2015), pp. 2856–2894.
[42] J.M. Lasry and P.L. Lions, Mean field games, Jpn. J. Math. 2(1) (2007), pp. 229–260.
[43] M. Laurière and L. Tangpi, Backward propagation of chaos. Electronic Journal of Probability 27 (2022), pp. 1–30.
[44] Y. Lin, Z. Ren, N. Touzi, and J. Yang, Random horizon principal-Agent problems, SIAM J. Control Optim. 60(1) (2022), pp. 355–384.
[45] T. Mastrolia and Z. Ren, Principal-agent problem with common agency without communication, SIAM J. Financial Math. 9(2) (2018), pp. 775–799.
[46] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, Syst. Control. Lett. 14(1) (1990), pp. 55–61.
[47] H. Pham, L.V. Vathana, and X.Y. Zhou, Optimal switching over multiple regimes, SIAM J. Control Optim. 48(4) (2009), pp. 2217–2253.
[48] D. Possamaï and N. Touzi, Is there a Golden Parachute in Sannikov’s principal-agent problem?, Preprint (2020).
[49] Y. Sannikov, A continuous-time version of the principal-agent problem, Rev. Econom. Stud. 75(3) (2008), pp. 957–984.
[50] J. Zhang, Backward Stochastic Differential Equations, From Linear to Fully Nonlinear Theory, Springer, 2017.