Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Mean Field Stackelberg Games: State Feedback Equilibrium

Minyi Huang * **Xuwei Yang** **

* *School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: mhuang@math.carleton.ca)*
** *School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: xuweiyang@cunet.carleton.ca)*

Abstract: We study mean field Stackelberg games between a major player (the leader) and a large population of minor players (the followers). By treating the mean field as part of the dynamics of the major player and a representative minor player, we Markovianize the decision problems and employ dynamic programming to determine the equilibrium strategy in a state feedback form. We show that for linear quadratic (LQ) models, the feedback equilibrium strategy is time consistent. We further give the explicit solution in a discrete-time LQ model.

Keywords: Mean field Stackelberg game, feedback strategy, linear-quadratic model, time consistency

1. INTRODUCTION

To tackle complexity in large-population decision problems, mean field games exploit ideas in statistical physics to approximate the collective behavior of a large number of rational agents; see Huang et al. (2006); Lasry and Lions (2006); Cardaliaguet (2013); Caines et al. (2017); Carmona and Delarue (2018). An important generalization of mean field game modeling is to include a major player to interact with a large number of minor players. This is initially introduced in Huang (2010) for an LQ model. Nguyen and Huang (2012) study nonuniform minor players parametrized by a continuum set. Partial information is addressed in Firoozi and Caines (2015). Ma and Huang (2020) apply multi-scale analysis. Important nonlinear extensions are developed by Nourian and Caines (2013); Carmona and Zhu (2016); Lasry and Lions (2018).

For the major-minor player model, a problem of interest is to consider leadership of the major player. In this paper, we will analyze a mean field Stackelberg model with a major player and a large number of minor players. There has been a long history to address leadership in games. Stackelberg competition of a leader and a follower is attributed to von Stackelberg (1934). Başar and Olsder (1999) give a general introduction to dynamic Stackelberg games. Yong (2002) solves an LQ stochastic differential game between a leader and a follower by the stochastic maximum principle. Bensoussan et al. (2015b) derive a maximum principle for stochastic Stackelberg differential games between a leader and a follower under the adapted closed-loop memoryless information structure.

Moon and Başar (2016, 2018) study LQ mean field Stackelberg games in discrete-time and continuous-time settings, respectively. They start with an $(N+1)$-player model and take mean field approximations to derive decentralized strategies. Bensoussan et al. (2015a) study a class of mean field Stackelberg games, in which each minor player has delay in collecting the information of the major player. Bensoussan et al. (2016) study a mean field game model between a (dominating) major player and a continuum of minor players.

The solutions in Bensoussan et al. (2015a, 2016); Moon and Başar (2016, 2018) to various extent rely on calculus of variations or the stochastic maximum principle. The resulting equilibria in general do not have time-consistency. In this mean field game context, a set of strategies is called time consistent if it still has the equilibrium property when implemented in a remaining time horizon. Also see Elie et al. (2019); Fu and Horst (2018) for studying leadership via the stochastic maximum principle. In our work, we introduce a different approach by dynamic programming in an augmented state space so that time consistency can be achieved. Wang and Zhang (2014) use dynamic programming for a discrete time system. But they consider a special class of simple dynamics and costs without control penalty. Their method does not augment the dynamics as in our work and can not handle our model.

Time consistency is an important issue in decision problems including optimal control and dynamic games. Ekeland and Lazrak (2006) analyze a deterministic optimal control problem in continuous time, which has time-inconsistency due to non-exponential discounting. They take a game theoretic point of view by considering t-selves as different decision makers and characterize a subgame perfect equilibrium. For further references overcoming time-inconsistency, see Ekeland and Pirvu (2008); Björk and Murgoci (2008); Djehiche and Huang (2016); Yong (2017).

1.1 Contribution and organization of this paper

Starting from an $N+1$ player Stackelberg game, we construct a limiting model with a major player and a representative minor player and look for their equilibrium strategies. Within an augmented state space, we introduce two dynamic programming equations, which can be called the master equations. This paper is organized as follows. Section 2 introduces the finite population Stackelberg game. In Section 3 we study a mean field limit model and obtain feedback strategies. Section 4 solves an LQ model to illustrate the approach in Section 3. Section 5 analyzes a discrete-time LQ mean field limit model.

1.2 Notation

Let $\mathcal{P}_2(\mathbb{R}^n)$ be the set of Borel probability measures on $\mathbb{R}^n$ with finite second moment; $C^2_{b2d}(\mathbb{R}^n;\mathbb{R}^k)$ be the set of $\mathbb{R}^k$-valued functions with continuous and bounded second order partial derivatives. Denote $\langle \mu, g\rangle := \int g(y)\mu(dy)$, and $\langle y\rangle_\mu := \int y\mu(dy)$ for probability measure μ and function g if the integral is finite. Denote $|y|_M = (y^T My)^{1/2}$, $[\![Y]\!]^2_M = Y^T MY$ for $y \in \mathbb{R}^n$, $Y \in \mathbb{R}^{n\times k}$ and symmetric $n \times n$ matrix $M \geq 0$.

2. FINITE-POPULATION STACKELBERG GAME

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete filtered probability space. Consider a major player $\mathcal{A}_0$ and N minor players $\mathcal{A}_i$, $1 \leq i \leq N$, described by the stochastic differential equations (SDEs):

$$dX^0_t = f_0(X^0_t, \mu^{(N)}_t, u^0_t)dt + \sigma_0 dW^0_t, \tag{1}$$

$$dX^i_t = f(X^0_t, X^i_t, \mu^{(N)}_t, u^0_t, u^i_t)dt + \sigma dW^i_t,\ 1 \leq i \leq N, \tag{2}$$

where $X^j_t \in \mathbb{R}^n$ and $u^j_t \in \mathbb{R}^{n_1}$ are the state and control of $\mathcal{A}_j$, $0 \leq j \leq N$, and $\mu^{(N)}_t := \frac{1}{N}\sum_{i=1}^N \delta_{X^i_t}$ is the empirical distribution of all minor players' states. All initial states are independent with $E|X^j_0|^2 \leq C$ for some fixed C. The $\mathbb{R}^{n_2}$-valued Brownian motions $\{W^j : 0 \leq j \leq N\}$ are mutually independent and also independent of the initial states. The constant matrices σ_0 and σ are $n \times n_2$. For two closed subsets U^0 and U of $\mathbb{R}^{n_1}$, $u^0_t \in U^0$, $u^i_t \in U$, for $1 \leq i \leq N$. Let J^{N+1}_j be the cost functional of player $\mathcal{A}_j$ and

$$J^{N+1}_0(u^0, u^1, \cdots, u^N) = E\int_0^T \mathrm{e}^{-\rho t}\, L_0(X^0_t, \mu^{(N)}_t, u^0_t)dt, \tag{3}$$

$$J^{N+1}_i(u^0, u^i, u^{-i}) = E\int_0^T \mathrm{e}^{-\rho t}\, L(X^0_t, X^i_t, \mu^{(N)}_t, u^0_t, u^i_t)dt, \tag{4}$$

where $\rho > 0, 1 \leq i \leq N$ and $u^{-i} = (u^1, ..., u^{i-1}, u^{i+1}, ..., u^N)$. For simplicity, the terminal costs are taken as zero.

On $\mathcal{P}_2(\mathbb{R}^n)$ we define the Wasserstein metric $W_2(\mu, \nu) = \inf_{\gamma\in\Gamma(\mu,\nu)}(\int_{\mathbb{R}^{2n}} |x - y|^2\gamma(dx, dy))^{1/2}$, where $\Gamma(\mu, \nu)$ is the set of probability distributions on $\mathbb{R}^{2n}$ that have μ and ν as the first and second marginals, respectively. Then $(\mathcal{P}_2(\mathbb{R}^n), W_2)$ is a complete metric space.

We introduce the following assumption (A1).

(A1) The following functions

$f_0 : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times U^0 \to \mathbb{R}^n,\ L_0 : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times U^0 \to \mathbb{R},$
$f : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times U^0 \times U \to \mathbb{R}^n,$
$L : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \times U^0 \times U \to \mathbb{R},$

are continuous, and there exists a constant C_0 such that for $\phi_f = f_0, f$, and $\psi_L = L_0, L$,

$$|\phi_f(x_0, x_1, \mu, u_0, u_1) - \phi_f(\hat{x}_0, \hat{x}_1, \hat{\mu}, \hat{u}_0, \hat{u}_1)| \leq C_0(|x_0 - \hat{x}_0| + |x_1 - \hat{x}_1| + |u_0 - \hat{u}_0| + |u_1 - \hat{u}_1| + W_2(\mu, \hat{\mu})),$$

$$|\psi_L(x_0, x_1, \mu, u_0, u_1)| \leq C_0(1 + |x_0|^2 + |x_1|^2 + |u_0|^2 + |u_1|^2 + \langle |y|^2\rangle_\mu),$$

$\forall x_0, \hat{x}_0, x_1, \hat{x}_1 \in \mathbb{R}^n,\ \forall \mu, \hat{\mu} \in \mathcal{P}_2(\mathbb{R}^n),\ (u_0, u_1) \in U^0 \times U.$

Denote $\mathbf{X}_t = (X^0_t, \cdots, X^N_t)$. If we choose u^0_t and u^i_t as continuous functions of $(t, \mathbf{X}_t)$, with Lipschitz continuity in $\mathbf{X}_t$, then the SDE system (1)-(2) has a well defined solution.

A basic solution of the Stackelberg game is to consider state feedback strategies. For the major player, as the leader, and N minor players, one may try to adapt the dynamic programming approach in Başar and Olsder (1999) for a two-player Stackelberg game. This method, however, becomes unfeasible for large N due to high complexity.

3. MEAN FIELD LIMIT MODEL

Based on the $(N + 1)$-player game, we consider a mean field limit Stackelberg model which involves the major player $\mathcal{A}_0$, a representative minor player $\mathcal{A}_1$, and the distribution μ_t determined by a continuum of minor players. Fix $u^0_t \equiv u^0$ and $u^i_t \equiv u$ in (1)-(2). For $g \in C^2_{b2d}(\mathbb{R}^n)$, by Itô's formula we have

$$d\langle\mu^{(N)}_t, g\rangle = \langle\mu^{(N)}, g'(\cdot)f(X^0_t, \cdot, \mu^{(N)}, u^0, u) + \frac{1}{2}\mathrm{Tr}[g''(\cdot)\sigma\sigma^T]\rangle dt + \frac{1}{N}\sum_{i=1}^N g'(X^i_t)\sigma dW^i_t. \tag{5}$$

We consider the following system:

$$dX^0_s = f_0(X^0_s, \mu_s, u^0_s)dt + \sigma_0 dW^0_s, \tag{6}$$

$$dX^1_s = f(X^0_s, X^1_s, \mu_s, u^0_s, u^1_s)ds + \sigma dW^1_s, \tag{7}$$

$$\frac{d}{ds}\int_{\mathbb{R}^n} g(y)\mu_s(dy) = \int_{\mathbb{R}^n}[f^T(X^0_s, y, \mu_s, u^0_s, u^1_s)g'(y) + \tfrac{1}{2}\,\mathrm{Tr}(g''(y)\sigma\sigma^T)]\mu_s(dy), \tag{8}$$

where $s \geq t$, $X^0_t = x_0$, $X^1_t = x_1$, $\mu_t = \mu \in \mathcal{P}_2(\mathbb{R}^n)$, and $g \in C^2_{b2d}(\mathbb{R}^n)$. Equations (6)-(7) are obtained from (1)-(2) after approximating $\mu^{(N)}_t$ by μ_t. The measure flow $\{\mu_t, t \geq 0\}$ on one hand drives the evolution of (X^0_s, X^1_s), and on the other is regenerated by the empirical distribution of a large population of similar minor players with appropriate initial states. The differential equation (8) for μ_s in a weak form, which may be viewed as a limiting form of (5), essentially results from (7). It is informative to list it separately. Due to the arbitrary choice of x_1 and μ, μ_s in general is not equal to the distribution (or the conditional distribution given $\{X^0_h, h \leq s\}$) of X^1_s.

Similarly we define

$$J_0(t, x_0, \mu, u^0, u^1) = E\int_t^T \mathrm{e}^{-\rho(s-t)}\, L_0(X^0_s, \mu_s, u^0_s)ds, \tag{9}$$

$$J_1(t, x_0, x_1, \mu, u^0, u^1) = E\int_t^T \mathrm{e}^{-\rho(s-t)}\, L(X^0_s, X^1_s, \mu_s, u^0_s, u^1_s)ds. \tag{10}$$

We view (X^0_t, X^1_t, μ_t) as a state variable and look for feedback strategies of the following form, $u^0 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \to U^0$, and $u^1 : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \to U$. The pair of strategies is admissible if the resulting closed loop system has a well defined solution. Let the value functions be

$$V_0(t, x_0, \mu) = J_0(t, x_0, \mu, u^{0*}, u^{1*}), \tag{11}$$

$$V_1(t, x_0, x_1, \mu) = J_1(t, x_0, x_1, \mu, u^{0*}, u^{1*}). \tag{12}$$

Below we elaborate on the determination of the Stackelberg strategies (u^{0*}, u^{1*}) by dynamic programming equations of the value functions. The reader may consult (Başar and Olsder, 1999, sec. 7.6) for this approach applied to two-player Stackelberg differential games. Define the following differential operators associated with the processes (6)-(7):

$$\mathcal{L}_0^{u^0}\cdot = f_0^T(x_0,\mu,u^0)\tfrac{\partial}{\partial x_0}\cdot + \tfrac{1}{2}\operatorname{Tr}[(\tfrac{\partial^2}{\partial x_0^2}\cdot)\sigma_0\sigma_0^T],$$
$$\mathcal{L}_1^{u^0,u^1}\cdot = f^T(x_0,x_1,\mu,u^0,u^1)\tfrac{\partial}{\partial x_1}\cdot + \tfrac{1}{2}\operatorname{Tr}[(\tfrac{\partial^2}{\partial x_1^2}\cdot)\sigma\sigma^T],$$
$$\mathcal{L}_{\mathrm{mf}}^{u^0,u^1}\cdot = f^T(x_0,y,\mu,u^0,u^1)\tfrac{\partial}{\partial y}\cdot + \tfrac{1}{2}\operatorname{Tr}[(\tfrac{\partial^2}{\partial y^2}\cdot)\sigma\sigma^T].$$

The mean field Stackelberg equilibrium (u^{0*},u^{1*}), if it exists, is characterized by the Hamilton-Jacobi-Bellman (HJB) equation system

$$\begin{cases}\rho V_1 = \frac{\partial V_1}{\partial t} + L(x_0,x_1,\mu,u^{0*},u^{1*}) + (\mathcal{L}_0^{u^{0*}} + \mathcal{L}_1^{u^{0*},u^{1*}})V_1 \\ \qquad + \int_{\mathbb{R}^n}\mathcal{L}_{\mathrm{mf}}^{u^{0*},u^{1*}}\partial_\mu V_1(t,x_0,x_1,\mu;y)\mu(dy), \\ \rho V_0 = \frac{\partial V_0}{\partial t} + L_0(x_0,\mu,u^{0*}) + \mathcal{L}_0^{u^{0*}}V_0 \\ \qquad + \int_{\mathbb{R}^n}\mathcal{L}_{\mathrm{mf}}^{u^{0*},u^{1*}}\partial_\mu V_0(t,x_0,\mu;y)\mu(dy), \end{cases} \tag{13}$$

where $(t,x_0,x_1,\mu)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{P}_2(\mathbb{R}^n)$ and $V_1 = V_0 = 0$ at $t=T$. We may also called them the master equations. Here $\mathcal{L}_{\mathrm{mf}}^{u^0,u^1}$ acts on $\partial_\mu V_i$, $i=0,1$ via the y variable, with (t,x_0,x_1,μ) fixed. Note that $\partial_\mu V_i$ has the extra independent variable y. For instance, for a fixed continuous function g of quadratic growth, $G(\mu)=\int g(x)\mu(dx)$ is a function of $\mu\in\mathcal{P}_2(\mathbb{R}^n)$. Then $\partial_\mu G(\mu)=g(\cdot)$. For related references, see Cardaliaguet et al. (2015); Huang (2019). For this section, we make the hypotheses i) the value functions V_0 and V_1 have sufficient smoothness; ii) the derivatives $\partial_\mu V_0,\ \partial_\mu V_1$ exist and have sufficient smoothness with respect to y; iii) integral terms involve functions of mild growth rate to ensure a well defined term; iv) the minimizers in (14)-(15) below are well defined.

These hypotheses are satisfied by the LQ model in Section 4.

For each $u^0\in U^0$, let $\hat u^1$ be the best response of the continuum of minor players including $\mathcal{A}_1$. We determine $\hat u^1$ by studying the optimizing behavior of $\mathcal{A}_1$. Denote

$$H_1 = L(x_0,x_1,\mu,u^0,u^1) + (\mathcal{L}_0^{u^0} + \mathcal{L}_1^{u^0,u^1})V_1(t,x_0,x_1,\mu) + \int_{\mathbb{R}^n}\mathcal{L}_{\mathrm{mf}}^{u^0,\hat u^1}\partial_\mu V_1(t,x_0,x_1,\mu;y)\mu(dy),$$

where $\hat u^1$ has been applied by all other minor players to determine the mean field. Player $\mathcal{A}_1$ optimizes u^1 only through L and $\mathcal{L}_1^{u^0,u^1}V_1(t,x_0,x_1,\mu)$. Let the minimizer of H_1 be

$$\hat u^1 = \varphi_1(x_0,x_1,\mu,u^0,\tfrac{\partial V_1}{\partial x_1}), \tag{14}$$

which is the best response of $\mathcal{A}_1$ at (t,x_0,x_1,μ) to $u^0\in U^0$.

Next we consider the major player's optimizer when all minor players have adopted (14) by matching their own states. Denote

$$H_0 = L_0(x_0,\mu,u^0) + \mathcal{L}_0^{u^0}V_0(t,x_0,\mu) + \int_{\mathbb{R}^n}\mathcal{L}_{\mathrm{mf}}^{u^0,\hat u^1}\partial_\mu V_0(t,x_0,\mu;y)\mu(dy).$$

Let the minimizer of H_0 be

$$u^{0*} = \varphi_0(x_0,\mu,\tfrac{\partial V_0}{\partial x_0},\partial_\mu V_0(t,x_0,\mu;\cdot),\tfrac{\partial V_1}{\partial x_1}(t,x_0,\cdot,\mu)). \tag{15}$$

Substituting (15) into (14) gives

$$u^{1*} = \varphi_1(x_0,x_1,\mu,u^{0*},\tfrac{\partial V_1}{\partial x_1}). \tag{16}$$

The selection of (u^{0*},u^{1*}) may be viewed as optimization problems of t-selves. For instance, given (t,x_0,x_1,μ), a coalition of minor players, i.e., s-$\mathcal{A}^1$ agents with $s\in[t,t+\epsilon]$, optimizes its cost defined on $[t,T]$ while it only acts on $[t,t+\epsilon]$. Then we let $\epsilon\to 0$. The pair (15)-(16) is called a feedback Stackelberg equilibrium strategy for the mean field Stackelberg game specified by (6)-(8) and (9)-(10). Under the equilibrium strategy (15)-(16), we may further write the closed-loop dynamics for (X_s^0,X_s^1,μ_s). This section only constructs the HJB equations (13). The existence analysis for these equations together with the closed loop system is an interesting subject. We will not give in-depth analysis here, but will use the LQ case to illustrate computations.

Remark 1. We give some detail about the integral terms in (13). Let u^0 be fixed. On $[t,t+\epsilon]$, we take a Taylor expansion of $V_1(t+\epsilon,X_{t+\epsilon}^0,X_{t+\epsilon}^1,\mu_{t+\epsilon})$. In particular, we have the first order approximation term

$$\int_{\mathbb{R}^n}(\partial_\mu V_1)(t,x_0,x_1,\mu;y)(\mu_{t+\epsilon}(dy)-\mu(dy))$$
$$= \int_{\mathbb{R}^n} f^T(x_0,y,\mu,u^0,\hat u^1)\tfrac{\partial}{\partial y}(\partial_\mu V_1)(t,x_0,x_1,\mu;y)\mu(dy)\epsilon$$
$$+ \int_{\mathbb{R}^n}\tfrac{1}{2}\operatorname{Tr}(\tfrac{\partial^2}{\partial y^2}(\partial_\mu V_1)(t,x_0,x_1,\mu;y)\sigma\sigma^T)\mu(dy)\epsilon + o(\epsilon).$$

In the end, $(u^0,\hat u^1)$ will be taken as (u^{0*},u^{1*}). The integral term in the HJB equation of V_0 arises for similar reasons.

4. LINEAR QUADRATIC MEAN FIELD LIMIT MODEL

We consider an LQ mean field limit model, and follow the steps in Section 3 to search for an explicit solution of the mean field Stackelberg equilibrium. Now equations (6)-(7) take drift terms

$$f_0(X_t^0,\mu_t,u_t^0) = A_0X_t^0 + B_0u_t^0 + F_0\langle y\rangle_{\mu_t},$$
$$f(X_t^0,X_t^1,\mu_t,u_t^0,u_t^1) = AX_t^1 + Bu_t^1 + Du_t^0 + F\langle y\rangle_{\mu_t} + GX_t^0.$$

The instantaneous cost functions take the quadratic forms

$$L_0(X_t^0,\mu_t,u_t^0) = |X_t^0 - \Gamma_0\langle y\rangle_{\mu_t}|_{Q_0}^2 + |u_t^0|_{R_0}^2, \tag{17}$$
$$L(X_t^0,X_t^1,\mu_t,u_t^0,u_t^1) = |X_t^1 - \Gamma_1X_t^0 - \Gamma_2\langle y\rangle_{\mu_t}|_Q^2 + |u_t^1|_R^2 + |u_t^0|_{R_1}^2 + 2u_t^{0T}R_2u_t^1. \tag{18}$$

The matrices A_0, A, B_0, B, F_0, F, D, G, Γ_0, Γ_1, Γ_2, Q_0, Q, $R>0$, $R_0>0$, R_1 and R_2 have compatible dimensions.

4.1 Stackelberg equilibrium strategy

Denote $\widetilde D = D - BR^{-1}R_2^T$. By (15)-(16), the Stackelberg equilibrium strategy is

$$u^{0*} = -\tfrac{1}{2}R_0^{-1}[B_0^T\tfrac{\partial V_0}{\partial x_0} + \widetilde D^T\int_{\mathbb{R}^n}\tfrac{\partial}{\partial y}(\partial_\mu V_0)\mu(dy)], \tag{19}$$
$$u^{1*} = -\tfrac{1}{2}R^{-1}[B^T\tfrac{\partial V_1}{\partial x_1} + 2R_2^Tu^{0*}]. \tag{20}$$

As a special case of (13), the HJB equations are

$$\rho V_0 = \tfrac{\partial V_0}{\partial t} + (\tfrac{\partial V_0}{\partial x_0})^T(A_0x_0 + B_0u^{0*} + F_0\langle y\rangle_\mu)$$
$$+ \int_{\mathbb{R}^n}(Gx_0 + Ax_1 + F\langle y\rangle_\mu + Bu^{1*} + Du^{0*})|_{x_1=y}^T\cdot$$
$$\tfrac{\partial}{\partial y}(\partial_\mu V_0)\mu(dy) + L_0(x_0,\mu,u^{0*})$$
$$+ \tfrac{1}{2}\operatorname{Tr}[(\tfrac{\partial^2 V_0}{\partial x_0^2})\sigma_0\sigma_0^T + \int_{\mathbb{R}^n}\tfrac{\partial^2}{\partial y^2}(\partial_\mu V_0)\sigma\sigma^T\mu(dy)], \tag{21}$$

$$\begin{aligned}\rho V_1 =& \tfrac{\partial V_1}{\partial t} + (\tfrac{\partial V_1}{\partial x_0})^T(A_0x_0 + B_0u^{0*} + F_0\langle y\rangle_\mu)\\ &+ (\tfrac{\partial V_1}{\partial x_1})^T(Gx_0 + Ax_1 + F\langle y\rangle_\mu + Bu^{1*} + Du^{0*})\\ &+ \int_{\mathbb{R}^n}(Gx_0 + Ax_1 + F\langle y\rangle_\mu + Bu^{1*} + Du^{0*})|^T_{x_1=y}\cdot\\ &\quad \frac{\partial}{\partial y}(\partial_\mu V_1)\mu(dy) + L(x_0, x_1, \mu, u^{0*}, u^{1*})\\ &+ \tfrac{1}{2}\,\mathrm{Tr}[\tfrac{\partial^2 V_1}{\partial x_0^2}\sigma_0\sigma_0^T + \tfrac{\partial^2 V_1}{\partial x_1^2}\sigma\sigma^T + \int_{\mathbb{R}^n}\frac{\partial^2}{\partial y^2}(\partial_\mu V_1)\sigma\sigma^T\mu(dy)].\end{aligned} \tag{22}$$

Assume V_0 and V_1 take the following forms

$$\begin{aligned}V_0(t, x_0, \mu) =& x_0^TP_0^0(t)x_0 + \langle y\rangle_\mu^T P_1^0(t)\langle y\rangle_\mu\\ &+ 2x_0^TP_{01}^0(t)\langle y\rangle_\mu + r_0(t),\end{aligned} \tag{23}$$

$$\begin{aligned}V_1(t, x_0, x_1, \mu) =& x_0^TP_0(t)x_0 + x_1^TP_1(t)x_1 + \langle y\rangle_\mu^TP_2(t)\langle y\rangle_\mu\\ &+ 2(x_0^TP_{01}(t)x_1 + x_0^TP_{02}(t)\langle y\rangle_\mu + x_1^TP_{12}(t)\langle y\rangle_\mu) + r_1(t).\end{aligned} \tag{24}$$

We substitute (23)-(24) into (19)-(20) to obtain

$$u^{0*} = K_0^0x_0 + K_1^0\langle y\rangle_\mu, \tag{25}$$

$$u^{1*} = K_0x_0 + K_1x_1 + K_2\langle y\rangle_\mu, \tag{26}$$

where

$$\begin{aligned}K_0^0 &= -R_0^{-1}(B_0^TP_0^0 + \widetilde{D}^TP_{01}^{0T}),\\ K_1^0 &= -R_0^{-1}(B_0^TP_{01}^0 + \widetilde{D}^TP_1^0),\\ K_0 &= -R^{-1}(R_2^TK_0^0 + B^TP_{01}^T),\\ K_1 &= -R^{-1}B^TP_1, \quad K_2 = -R^{-1}(R_2^TK_1^0 + B^TP_{12}).\end{aligned}$$

The pair (u^{0*}, u^{1*}) obtained by (25)-(26) is a feedback Stackelberg equilibrium. Substituting (23)-(26) into (21)-(22) gives the following Riccati ODE system on $[0, T]$:

$$\begin{aligned}\dot{P}_0 =& \rho P_0 - P_0(A_0 + B_0K_0^0) - (A_0 + B_0K_0^0)^TP_0\\ &- (P_{01} + P_{02})(G + BK_0 + DK_0^0)\\ &- (G + BK_0 + DK_0^0)^T(P_{01} + P_{02})^T\\ &- K_0^{0T}R_2K_0 - K_0^TR_2^TK_0^0 - [\![K_0]\!]^2_R - [\![K_0^0]\!]^2_{R_1} - [\![\Gamma_1]\!]^2_Q,\end{aligned}$$

$$\dot{P}_1 = \rho P_1 - P_1(A + BK_1) - (A + BK_1)^TP_1 - [\![K_1]\!]^2_R - Q,$$

$$\begin{aligned}\dot{P}_2 =& \rho P_2 - P_{02}^T(B_0K_1^0 + F_0) - (B_0K_1^0 + F_0)^TP_{02}\\ &- P_{12}^T(BK_2 + DK_1^0 + F) - (BK_2 + DK_1^0 + F)^TP_{12}\\ &- (A + BK_1 + BK_2 + DK_1^0 + F)^TP_2\\ &- P_2(A + BK_1 + BK_2 + DK_1^0 + F)\\ &- [\![K_2]\!]^2_R - [\![K_1^0]\!]^2_{R_1} - K_1^{0T}R_2K_2 - K_2^TR_2^TK_1^0 - [\![\Gamma_2]\!]^2_Q,\end{aligned}$$

$$\begin{aligned}\dot{P}_{01} =& \rho P_{01} - (A_0 + B_0K_0^0)^TP_{01} - P_{01}(A + BK_1)\\ &- (G + BK_0 + DK_0^0)^T(P_1 + P_{12}^T) - K_0^TRK_1\\ &- K_0^{0T}R_2K_1 + \Gamma_1^TQ,\end{aligned}$$

$$\begin{aligned}\dot{P}_{02} =& \rho P_{02} - 2P_0(B_0K_1^0 + F_0) - (A_0 + B_0K_0^0)^TP_{02}\\ &- P_{01}(BK_2 + DK_1^0 + F)\\ &- (G + BK_0 + DK_0^0)^T(P_{12} + P_2)\\ &- P_{02}(A + BK_1 + BK_2 + DK_1^0 + F) - K_0^TRK_2\\ &- K_0^{0T}R_1K_1^0 - K_0^{0T}R_2K_2 - K_0^TR_2^TK_1^0 - \Gamma_1^TQ\Gamma_2,\end{aligned}$$

$$\begin{aligned}\dot{P}_{12} =& \rho P_{12} - P_{01}^T(B_0K_1^0 + F_0) - P_1(BK_2 + DK_1^0 + F)\\ &- P_{12}(A + BK_1 + BK_2 + DK_1^0 + F)\\ &- (A + BK_1)^TP_{12} - K_1^TRK_2 - K_1^TR_2^TK_1^0 + Q\Gamma_2,\end{aligned}$$

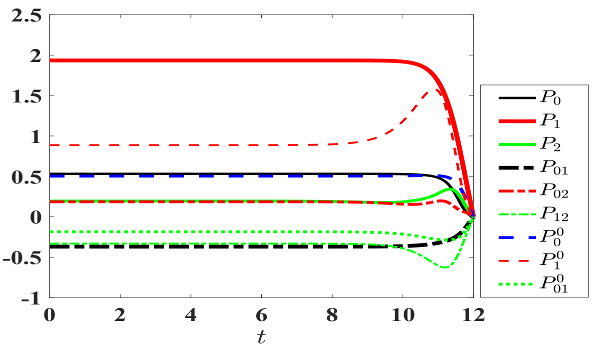

Fig. 1. The solution of $P_0, P_1, \cdots, P_{01}^0$ in Section 4.3.

$$\begin{aligned}\dot{P}_0^0 =& \rho P_0^0 - P_0^0(A_0 + B_0K_0^0) - (A_0 + B_0K_0^0)^TP_0^0\\ &- P_{01}^0(G + BK_0 + DK_0^0) - (G + BK_0 + DK_0^0)^TP_{01}^{0T}\\ &- [\![K_0^0]\!]^2_{R_0} - Q_0,\end{aligned}$$

$$\begin{aligned}\dot{P}_1^0 =& \rho P_1^0 - P_{01}^{0T}(B_0K_1^0 + F_0) - (B_0K_1^0 + F_0)^TP_{01}^0\\ &- (A + BK_1 + BK_2 + DK_1^0 + F)^TP_1^0\\ &- P_1^0(A + BK_1 + BK_2 + DK_1^0 + F) - [\![K_1^0]\!]^2_{R_0} - [\![\Gamma_0]\!]^2_{Q_0},\end{aligned}$$

$$\begin{aligned}\dot{P}_{01}^0 =& \rho P_{01}^0 - P_0^0(B_0K_1^0 + F_0) - (A_0 + B_0K_0^0)^TP_{01}^0\\ &- P_{01}^0(A + BK_1 + BK_2 + DK_1^0 + F)\\ &- (G + BK_0 + DK_0^0)^TP_1^0 - K_0^{0T}R_0K_1^0 + Q_0\Gamma_0,\end{aligned} \tag{27}$$

where all of $P_0, P_1, \cdots, P_{01}^0$ are equal to 0 at T.

Theorem 1. If the Riccati ODE system (27) has a solution on $[0, T]$, then (25)-(26) is a feedback Stackelberg equilibrium on $[0, T]$. □

If (27) has a solution on $[0, T]$, the closed-loop system of (X_s^0, X_s^1) under (25)-(26) admits a unique strong solution.

4.2 Time consistency

A strategy is called time consistent on $[0, T]$, if for any subgame on $[t_1, T]$, $\forall t_1 \in (0, T)$, it is still an equilibrium.

Theorem 2. If the ODE system (27) has a solution on $[0, T]$, the Stackelberg equilibrium strategy (25)-(26) is time consistent.

Proof. When one restricts to a remaining period $[t_0, T]$ for any $t_0 \in (0, T)$ and re-solves a mean field Stackelberg game, the same Riccati ODE system is still valid. □

4.3 Numerical illustration

MatLab ODE solver `ode45` is used to solve the ODE system (27) of $P_0, P_1, \cdots, P_{01}^0$ on $[0, T]$, with terminal condition 0 and parameter values $A_0 = 1$, $B_0 = 2$, $F_0 = 0.5$, $A = 0.5$, $B = 1$, $D = 1$, $F = 0.2$, $G = 0.4$, $\Gamma_0 = 0.8$, $\Gamma_1 = 0.3$, $\Gamma_2 = 0.5$, $Q = 2$, $Q_0 = 1$, $R = 1$, $R_0 = 0.5$, $R_1 = 1$, $R_2 = 0.5$, $T = 12$, and $\rho = 0.1$. The solution is shown in Fig. 1

5. DISCRETE-TIME MEAN FIELD LIMIT MODEL

Consider a mean field Stackelberg game with discrete time horizon $\{0, 1, \cdots, T\}$. The major player $\mathcal{A}_0$, a representative minor player $\mathcal{A}_1$, and the mean field state $\overline{X}$ have the dynamics:

$$X^0_{t+1} = A_0X^0_t + B_0u^0_t + F_0\overline{X}_t + W^0_t, \quad (28)$$
$$X^1_{t+1} = AX^1_t + Bu^1_t + Du^0_t + F\overline{X}_t + GX^0_t + W^1_t, \quad (29)$$
$$\overline{X}_{t+1} = (A + F)\overline{X}_t + B\overline{u}_t + Du^0_t + GX^0_t, \quad (30)$$

where W^i_t, $0 \le t \le T-1$ are i.i.d. random variables with zero mean and finite variance. The control mean field $\overline{u}_t = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N u^i_t$, where each u^i_t is a minor player's control. For discount factor $\alpha \in (0,1)$, the cost functionals are

$$J_0(u^0, u^1) = E[\sum_{t=0}^{T-1} \alpha^t(|X^0_t - \Gamma_0\overline{X}_t|^2_{Q_0} + |u^0_t|^2_{R_0})], \quad (31)$$
$$J_1(u^0, u^1) = E[\sum_{t=0}^{T-1} \alpha^t(|X^1_t - \Gamma_1X^0_t - \Gamma_2\overline{X}_t|^2_Q + |u^1_t|^2_R + |u^0_t|^2_{R_1} + 2u^{0T}_t R_2u^1_t)], \quad (32)$$

Remark 2. The underlying $N+1$ player model is specified by replacing $\overline{X}_t$ by $X^{(N)}_t = \frac{1}{N}\sum_{i=1}^N X^i_t$ in (28)-(29) and J_0, J_1.

5.1 Minimizer of the minor player

By dynamic programming,

$$V_1(t, x_0, x_1, \overline{x}) = \min_{u^1}\{|X^1_t - \Gamma_1X^0_t - \Gamma_2\overline{X}_t|^2_Q + |u^1_t|^2_R + |u^0_t|^2_{R_1} + 2u^{0T}_t R_2u^1_t + \alpha E[V_1(t+1, X^0_{t+1}, X^1_{t+1}, \overline{X}_{t+1})]_{(X^0_t, X^1_t, \overline{X}_t)=(x_0,x_1,\overline{x})}\}. \quad (33)$$

Given $u^0_t = u^0$ for the major player, we assume for $k = t+1$,

$$V_1(k, x_0, x_1, \overline{x}) = x^T_0\Phi_{0,k}x_0 + x^T_1\Phi_{1,k}x_1 + \overline{x}^T\Phi_{2,k}\overline{x} + r^1_k + 2(x^T_0\Phi_{01,k}x_1 + x^T_0\Phi_{02,k}\overline{x} + x^T_1\Phi_{12,k}\overline{x}).$$

By the first order condition, the minimizer in (33) satisfies

$$\begin{aligned} 0 =& (R + \alpha B^T\Phi_{1,t+1}B)u^1_t + \alpha B^T\Phi_{12,t+1}B\bar{u}_t \\ &+ \{R^T_2 + \alpha B^T[(\Phi_{1,t+1} + \Phi_{12,t+1})D + \Phi^T_{01,t+1}B_0]\}u^0_t \\ &+ \alpha B^T[(\Phi_{1,t+1} + \Phi_{12,t+1})G + \Phi^T_{01,t+1}A_0]x_0 \\ &+ \alpha B^T\Phi_{1,t+1}Ax_1 \\ &+ \alpha B^T[\Phi_{1,t+1}F + \Phi^T_{01,t+1}F_0 + \Phi_{12,t+1}(A+F)]\overline{x}. \end{aligned}$$

If $R + \alpha B^T\Phi_{1,t+1}B$ is invertible, then

$$\begin{aligned} u^1_t =& -(R + \alpha B^T\Phi_{1,t+1}B)^{-1}\alpha B^T\cdot \\ &\{[(\Phi_{1,t+1} + \Phi_{12,t+1})G + \Phi^T_{01,t+1}A_0]x_0 + \Phi_{1,t+1}Ax_1 \\ &+ [\Phi_{1,t+1}F + \Phi^T_{01,t+1}F_0 + \Phi_{12,t+1}(A+F)]\overline{x} \\ &+ \Phi_{12,t+1}B\bar{u}_t\} - (R + \alpha B^T\Phi_{1,t+1}B)^{-1}\cdot \\ &\{R^T_2 + \alpha B^T[(\Phi_{1,t+1} + \Phi_{12,t+1})D + \Phi^T_{01,t+1}B_0]\}u^0_t. \end{aligned} \quad (34)$$

By the *consistency condition* $\overline{u}_t = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N u^i_t$, with each u^i_t being a copy of u^1_t given by (34), we determine

$$\begin{aligned} \overline{u}_t =& -[R + \alpha B^T(\Phi_{1,t+1} + \Phi_{12,t+1})B]^{-1}\alpha B^T\cdot \\ &\{[(\Phi_{1,t+1} + \Phi_{12,t+1})G + \Phi^T_{01,t+1}A_0]x_0 \\ &+ [(\Phi_{1,t+1} + \Phi_{12,t+1})(A+F) + \Phi^T_{01,t+1}F_0]\overline{x}\} \\ &- [R + \alpha B^T(\Phi_{1,t+1} + \Phi_{12,t+1})B]^{-1}\cdot \\ &\{R^T_2 + \alpha B^T[(\Phi_{1,t+1} + \Phi_{12,t+1})D + \Phi^T_{01,t+1}B_0]u^0_t\}. \end{aligned}$$

We write the above $\overline{u}_t$ in the form

$$\overline{u}_t = \xi^0_{t+1}X^0_t + \xi^2_{t+1}\overline{X}_t + \xi^3_{t+1}u^0_t, \quad (35)$$

and further substitute (35) back into (34) to get

$$u^1_t = \theta^0_{t+1}X^0_t + \theta^1_{t+1}X^1_t + \theta^2_{t+1}\overline{X}_t + \theta^3_{t+1}u^0_t, \quad (36)$$

where ξ^i_{t+1} and θ^i_{t+1} are functions of $\Phi_{i,t+1}$. The control (36) is the best response of the minor player to u^0 of the major player. (35) is the control mean field for given u^0.

5.2 Minimizer of the major player

Taking into account the best response (36), we solve the major player's minimization problem. By dynamic programming,

$$V_0(t, x_0, \overline{x}) = \min_{u^0}\{|x_0 - \Gamma_0\overline{x}|^2_{Q_0} + |u^0_t|^2_{R_0} + \alpha E[V_0(t+1, X^0_{t+1}, \overline{X}_{t+1})]_{(X^0_t, \overline{X}_t)=(x_0,\overline{x})}\}. \quad (37)$$

Assume that V_0 takes the following form for $k = t+1$,

$$V_0(k, x_0, \overline{x}) = x^T_0\Phi^0_{0,k}x_0 + \overline{x}^T\Phi^0_{2,k}\overline{x} + 2x^T_0\Phi^0_{02,k}\overline{x} + r^0_k.$$

By the first order condition, the minimizer in (37) satisfies that

$$\begin{aligned} 0 =& R_0u^0_t + \alpha[B^T_0\Phi^0_{0,t+1} + (B\xi^3_{t+1} + D)^T\Phi^{0T}_{02,t+1}]\cdot \\ &(A_0x_0 + B_0u^0_t + F_0\overline{x}) \\ &+ \alpha[B^T_0\Phi^0_{02,t+1} + (B\xi^3_{t+1} + D)^T\Phi^{0T}_{2,t+1}]\cdot \\ &[(A+F)\overline{x} + B(\xi^0_{t+1}x_0 + \xi^2_{t+1}\overline{x} + \xi^3_{t+1}u^0_t) + Du^0_t + Gx_0]. \end{aligned}$$

If $R_0 + \alpha[B^T_0\Phi^0_{0,t+1} + (B\xi^3_{t+1} + D)^T\Phi^{0T}_{02,t+1}](B_0 + D_0)$ is invertible, we obtain from the first order condition that

$$u^0_t = \beta^0_{t+1}x_0 + \beta^2_{t+1}\overline{x}, \quad (38)$$

where the β^i_{t+1} are functions of $\Phi_{i,t+1}$ and $\Phi^0_{i,t+1}$.

5.3 Stackelberg equilibrium

We substitute (38) into (36) and (35) to obtain

$$u^1_t = \widehat{\theta}^0_{t+1}x_0 + \widehat{\theta}^1_{t+1}x_1 + \widehat{\theta}^2_{t+1}\overline{x}, \quad (39)$$
$$\overline{u}_t = \widehat{\xi}^0_{t+1}x_0 + \widehat{\xi}^2_{t+1}\overline{x}, \quad (40)$$

where $\widehat{\theta}^i_{t+1}$ and $\widehat{\xi}^i_{t+1}$ are functions of $\Phi_{i,t+1}$ and $\Phi^0_{i,t+1}$. The pair (u^0_t, u^1_t) in (38)-(39) is a Stackelberg equilibrium, and (40) is the control mean field.

We substitute the equilibrium $(u^0_t, u^1_t, \overline{u}_t)$ given by (38)-(40) into (33) to obtain equations for $\Phi_{i,t}$ and $\Phi^0_{i,t}$, which depend on $\Phi_{i,t+1}$ and $\Phi^0_{i,t+1}$. We can solve for $\Phi_{i,t}$ and $\Phi^0_{i,t}$ backwards in time with zero terminal condition at T.

$$\begin{aligned} \Phi_{0,t} =& [\![\Gamma_1]\!]^2_Q + [\![\theta^0_{t+1}]\!]^2_R + [\![\beta^0_{t+1}]\!]^2_{R_1} + \alpha[\![A^0_{t+1}]\!]^2_{\Phi_{0,t+1}} \\ &+ \alpha[\![G_{t+1}]\!]^2_{\Phi_{1,t+1}} + \alpha[\![\overline{G}_{t+1}]\!]^2_{\Phi_{2,t+1}} + 2\alpha A^{0T}_{t+1}\Phi_{01,t+1}G_{t+1} \\ &+ 2\alpha A^{0T}_{t+1}\Phi_{02,t+1}\overline{G}_{t+1} + 2\alpha G^T_{t+1}\Phi_{12,t+1}\overline{G}_{t+1}, \end{aligned}$$

$$\Phi_{1,t} = Q + [\![\widehat{\theta}^1_{t+1}]\!]^2_R + \alpha[\![A_{t+1}]\!]^2_{\Phi_{1,t+1}},$$

$$\begin{aligned} \Phi_{2,t} =& [\![\Gamma_2]\!]^2_Q + [\![\widehat{\theta}^2_{t+1}]\!]^2_R + [\![\beta^2_{t+1}]\!]^2_{R_1} + \alpha[\![F^0_{t+1}]\!]^2_{\Phi_{0,t+1}} \\ &+ \alpha[\![F_{t+1}]\!]^2_{\Phi_{1,t+1}} + \alpha[\![\overline{A}_{t+1}]\!]^2_{\Phi_{2,t+1}} + 2\alpha F^{0T}_{t+1}\Phi_{01,t+1}F_{t+1} \\ &+ 2\alpha F^{0T}_{t+1}\Phi_{02,t+1}\overline{A}_{t+1} + 2\alpha F^T_{t+1}\Phi_{12,t+1}\overline{A}_{t+1}, \end{aligned}$$

$$\begin{aligned} \Phi_{01,t} =& -\Gamma^T_1Q + \widehat{\theta}^{0T}_{t+1}R\widehat{\theta}^1_{t+1} + \alpha G^T_{t+1}\Phi_{1,t+1}A_{t+1} \\ &+ \alpha A^{0T}_{t+1}\Phi_{01,t+1}A_{t+1} + \alpha\overline{G}^T_{t+1}\Phi^T_{12,t+1}A_{t+1}, \end{aligned}$$

$$\begin{aligned} \Phi_{02,t} =& \Gamma_1Q\Gamma_2 + \widehat{\theta}^{0T}_{t+1}R\widehat{\theta}^2_{t+1} + \beta^{0T}_{t+1}R_1\beta^2_{t+1} + \alpha A^{0T}_{t+1}\Phi_{0,t+1}F^0_{t+1} \\ &+ \alpha[G^T_{t+1}\Phi_{1,t+1}F_{t+1} + \overline{G}^T_{t+1}\Phi_{2,t+1}\overline{A}_{t+1}] \\ &+ \alpha[A^{0T}_{t+1}\Phi_{01,t+1}F_{t+1} + G^T_{t+1}\Phi^T_{01,t+1}F^0_{t+1}] \\ &+ \alpha[A^{0T}_{t+1}\Phi_{02,t+1}\overline{A}_{t+1} + \overline{G}^T_{t+1}\Phi^T_{02,t+1}F_{t+1}] \\ &+ \alpha[G^T_{t+1}\Phi_{12,t+1}\overline{A}_{t+1} + \overline{G}^T_{t+1}\Phi^T_{12,t+1}F_{t+1}], \end{aligned}$$

$$\begin{aligned}\Phi_{12,t} =& - Q\Gamma_2 + \widehat{\theta}^{1T}_{t+1} R\widehat{\theta}^2_{t+1} + \alpha A^T_{t+1}\Phi_{1,t+1}F_{t+1}\\ &+ \alpha[A^T_{t+1}\Phi^T_{01,t+1}F^0_{t+1} + A^{0T}_{t+1}\Phi_{02,t+1}\overline{A}_{t+1}]\\ &+ \alpha[\overline{G}^T_{t+1}\Phi^T_{02,t+1}F^0_{t+1} + A^T_{t+1}\Phi_{12,t+1}\overline{A}_{t+1}],\\ \Phi^0_{0,t} =& [\![\Gamma_0]\!]^2_{Q_0} + [\![\beta^0_{t+1}]\!]^2_{R_0} + \alpha[\![A^0_{t+1}]\!]^2_{\Phi^0_{0,t+1}} + [\![\overline{G}_{t+1}]\!]^2_{\Phi^0_{2,t+1}}\\ &+ 2\alpha A^{0T}_{t+1}\Phi^0_{02,t+1}\overline{G}_{t+1},\\ \Phi^0_{2,t} =& [\![\Gamma_0]\!]^2_{Q_0} + [\![\beta^2_{t+1}]\!]^2_{R_0} + \alpha[[\![F^0_{t+1}]\!]^2_{\Phi^0_{0,t+1}} + [\![\overline{A}_{t+1}]\!]^2_{\Phi^0_{2,t+1}}]\\ &+ 2\alpha A^{0T}_{t+1}\Phi^0_{02,t+1}\overline{A}_{t+1},\\ \Phi^0_{02,t} =& - Q_0\Gamma_0 + \beta^{0T}_{t+1}R_0\beta^2_{t+1} + \alpha A^{0T}_{t+1}\Phi^0_{0,t+1}F^0_{t+1}\\ &+ \alpha[\overline{G}^T_{t+1}\Phi^0_{2,t+1}\overline{A}_{t+1} + A^{0T}_{t+1}\Phi^0_{02,t+1}\overline{A}_{t+1}]\\ &+ \alpha\overline{G}^T_{t+1}\Phi^{0T}_{02,t+1}F^0_{t+1}.\end{aligned}$$

In the above equations, the matrices A^0_{t+1}, F^0_{t+1}, A_{t+1}, F_{t+1}, G_{t+1}, $\overline{A}_{t+1}$ and $\overline{G}_{t+1}$ are defined as

$$\begin{aligned}&A^0_{t+1} = A_0 + B_0\beta^0_{t+1}, \quad F^0_{t+1} = F_0 + B_0\beta^2_{t+1},\\ &A_{t+1} = A + B\widehat{\theta}^1_{t+1}, \quad F_{t+1} = F + B\widehat{\theta}^2_{t+1} + D\beta^2_{t+1},\\ &G_{t+1} = G + B\widehat{\theta}^0_{t+1} + D\beta^0_{t+1},\\ &\overline{A}_{t+1} = A + F + B\widehat{\xi}^2_{t+1} + D\beta^2_{t+1},\\ &\overline{G}_{t+1} = G + B\widehat{\xi}^0_{t+1} + D\beta^0_{t+1}.\end{aligned}$$

Remark 3. The continuous-time LQ model needs the control-coupling term $u^{0T}R_2u^1$ in (18) so that the best response of the minor player explicitly involves u^0, which enables the major player's leadership. In the discrete-time case, leadership can be generated by dynamic coupling alone without control-coupling.

6. CONCLUDING REMARKS

For mean field Stackelberg games with a major player, we apply dynamic programming to find feedback equilibrium strategies. For future work, it is of interest to analyze the performance of the decentralized strategies applied by a finite population.

REFERENCES

Başar, T. and Olsder, G.J. (1999). *Dynamic Noncooperative Game Theory*. Academic Press, New York, 2 edition.

Bensoussan, A., Chau, M.H.M., and Yam, S.C.P. (2015a). Mean field Stackelberg games: Aggregation of delayed instructions. *SIAM J. Control Optim.*, 53(4), 2237–2266.

Bensoussan, A., Chau, M.H.M., and Yam, S.C.P. (2016). Mean filed games with a dominating player. *Appl. Math. Optim.*, 74(1), 91–128.

Bensoussan, A., Chen, S., and Sethi, S.P. (2015b). The maximum principle for global solutions of stochastic Stackelberg differential games. *SIAM J. Contr. Optim.*, 53(4), 1956–1981.

Björk, T. and Murgoci, A. (2008). A general theory of Markovian time inconsistent stochastic control problems. Preprint.

Caines, P.E., Huang, M., and Malhamé, R.P. (2017). Mean field games. In T. Başar and G. Zaccour (eds.), *Handbook of Dynamic Game Theory*. Springer, Berlin.

Cardaliaguet, P., Delarue, F., Lasry, J.M., and Lions, P.L. (2015). The master equation and the convergence problem in mean field games. Preprint.

Cardaliaguet, P. (2013). *Notes on mean field games*. University of Paris, Dauphine.

Carmona, R. and Zhu, X. (2016). A probabilistic approach to mean field games with major and minor players. *Ann. Appl. Probab.*, 26(3), 1535–1580.

Carmona, R. and Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer.

Djehiche, B. and Huang, M. (2016). A characterization of subgame perfect equilibria for SDEs of mean-field type. *Dyn. Games Appl.*, 6, 55–81.

Ekeland, I. and Lazrak, A. (2006). Being serious about non-commitment: subgame perfect equilibrium in continuous time. Preprint.

Ekeland, I. and Pirvu, T.A. (2008). Investment and consumption without commitment. *Math. Finan. Econ.*, 2, 57–86.

Elie, R., Mastrolia, T., and Possamaï, D. (2019). A tale of a principle and many, many agents. *Mathematics of Operations Research*, 44(2), 440–467.

Firoozi, D. and Caines, P.E. (2015). ϵ-Nash equilibria for partially observed LQG mean field games with major agent: Partial observations by all agents. In *Proc. 54th IEEE CDC*, 4430–4437. Osaka, Japan.

Fu, G. and Horst, U. (2018). Mean-field leader-follower games with terminal state constraint. arXiv:1809.04401.

Huang, M., Malhamé, R.P., and Caines, P.E. (2006). Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inform. Systems*, 6(3), 221–252.

Huang, M. (2010). Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.*, 48(5), 3318–3353.

Huang, M. (2019). Stochastic optimal control with Markovian lossy state observations. In *Proc. 58th IEEE CDC*, 678–683. Nice, France.

Lasry, J.M. and Lions, P.L. (2018). Mean-field games with a major player. *C. R. Math. Acad. Sci. Paris*, 356(8), 886–890.

Lasry, J.M. and Lions, P.L. (2006). Jeux à champ moyen. I - Le cas stationaire. *C. R. Math.*, 343(9), 619–625.

Ma, Y. and Huang, M. (2020). Linear quadratic mean field games with a major player: The multi-scale approach. *Automatica*, 113(3).

Moon, J. and Başar, T. (2016). Discrete-time stochastic Stackelberg dynamic games with a large number of followers. In *Proc. 55th IEEE CDC*, 3578–3583. Las Vegas, NV.

Moon, J. and Başar, T. (2018). Linear quadratic mean field Stackelberg differential games. *Automatica*, 97, 200–213.

Nguyen, S.L. and Huang, M. (2012). Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players. *SIAM J. Control Optim.*, 50(5), 2907–2937.

Nourian, M. and Caines, P.E. (2013). ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM J. Control Optim.*, 51(4), 3302–3331.

von Stackelberg, H. (1934). *Marktform und Gleichgewicht*. Springer-Verlag, Wien New York.

Wang, B.C. and Zhang, J.F. (2014). Hierarchical mean field games for multiagent systems with tracking-type costs: Distributed ϵ-Stackelberg equilibria. *IEEE Trans. Autom. Control*, 59(8), 2241–2247.

Yong, J. (2002). A leader-follower stochastic linear quadratic differential game. *SIAM J. Contr. Optim.*, 41(4), 1015–1041.

Yong, J. (2017). Linear-quadratic optimal control problems for mean-field stochastic differential equations – time-consistent solutions. *Trans. Amer. Math. Society*, 369(8), 5467–5523.