Les Cahiers du GERAD ISSN: 0711–2440

Linear-quadratic mean field Stackelberg games: Master equations and decentralized feedback strategies

M. Huang, X. Yang

G–2023–21

June 2023

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2023
– Bibliothèque et Archives Canada, 2023

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2023
– Library and Archives Canada, 2023

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Linear-quadratic mean field Stackelberg games: Master equations and decentralized feedback strategies

Minyi Huang [a, b]

Xuwei Yang [c]

[a] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[b] *School of Mathematics and Statistics, Carleton University, Carleton (On), Canada, K1S 5B6*

[c] *Department of Mathematics and Statistics, McMaster University, Hamilton (On), Canada, L8S 4K1*

mhuang@math.carleton.ca
henryyangxuwei@gmail.com

June 2023
Les Cahiers du GERAD
G–2023–21

Abstract : This paper studies linear-quadratic Stackelberg games with a major player (leader) and N minor players (followers). To design decentralized strategies in the $N+1$ player model, we construct a mean field limit model consisting of the leader and a representative follower and use dynamic programming to derive two master equations. We analyze quadratic solutions to the master equations and characterize existence and uniqueness by a pair of Riccati ordinary differential equations. The master equation-based solution is time consistent and provides decentralized feedback strategies in finite populations. As in feedback solutions of standard two-player dynamic Stackelberg games, the leader's equilibrium strategy in the mean field model does not have global optimality in minimizing its cost, and this feature makes the equilibrium analysis much more intricate than in mean field games (Huang, 2010). To characterize the performance of the decentralized strategies, we extend a procedure of Ekeland and Lazrak (2006) introduced for time inconsistent optimal control, so that the game of $N+1$ players is interpreted as being played by a stream of short-lived agents. Subsequently, the set of decentralized strategies is shown to be an ε_N-Stackelberg equilibrium, where $\varepsilon_N = o(1)$.

Keywords : Mean field models, Stackelberg games, decentralized control, time consistency

1 Introduction

Mean field game (MFG) theory has been developed to deal with non-cooperative dynamic decision-making in a large population of comparably small players [36, 43]. By addressing the macroscopic behavior of the overall population, the theory relates finite large populations to a simpler continuum population model, leading to a tractable solution. This methodology overcomes the dimensionality difficulty and provides decentralized strategies for the original model with a finite number of players. Since the inception of this paradigm, there has accumulated an enormous literature; see [7, 11, 13, 17] and references therein. The basic framework of MFG modeling and analysis has been enriched by extensions along different directions such as common noise models [13, 15, 17], discrete states [30, 53], mean field of controls [16, 46], partial information [54], cooperative agents [27, 35], learning algorithms [31, 45], among others.

1.1 Major player models

Among the generalizations of MFG modeling, of particular interest is a population with a major player, which has strong influence on a large number of minor players. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space. Consider a major player $\mathcal{A}_0$ and N minor players $\mathcal{A}_i$, $1 \le i \le N$, which have state processes described by the stochastic differential equations (SDEs):

$$dX_t^0 = f_0(X_t^0, \mu_t^{(N)}, u_t^0)dt + \sigma_0 dW_t^0, \tag{1}$$

$$dX_t^i = f(X_t^i, X_t^0, \mu_t^{(N)}, u_t^i, u_t^0)dt + \sigma dW_t^i, \qquad 1 \le i \le N, \tag{2}$$

where $\{W_t^j, 0 \le j \le N\}$ are $N+1$ independent Brownian motions and $\mu_t^{(N)} := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ is the empirical distribution of the states of the minor players. For simplicity, we take σ_0 and σ as constant matrices of suitable dimensions.

The cost functionals with initial time 0 are given by

$$J_0^{N+1}(0, u^0, \mathbf{u}) = \mathbb{E}\int_0^T L_0(X_t^0, \mu_t^{(N)}, u_t^0)dt + \mathbb{E}g_0(X_T^0, \mu_T^{(N)}), \tag{3}$$

$$J_i^{N+1}(0, u^i, u^0, \mathbf{u}^{-i}) = \mathbb{E}\int_0^T L(X_t^i, X_t^0, \mu_t^{(N)}, u_t^i, u_t^0)dt + \mathbb{E}g(X_T^i, X_T^0, \mu_T^{(N)}), \qquad 1 \le i \le N, \tag{4}$$

where we denote by $\mathbf{u} = (u^1, \cdots, u^N)$ the strategies of all minor players and write $\mathbf{u}^{-i} = (u^1, \cdots, u^{i-1}, u^{i+1}, \cdots, u^N)$. Here $\mathcal{A}_0$ is called the major player due to its significant impact on all other players. In contrast, each minor player's impact on others is negligible when N is large. Note that the functions f_0, f, L_0 and L are defined as a function of the states (i.e., x^0, x^i), the controls (i.e., u^0, u^i) and the measure-valued variable μ from a certain space.

This class of major-minor player mean field games was initially introduced in a linear-quadratic (LQ) setting [34]. It developed the so-called Nash certainty equivalence approach treating a mean field limit model consisting of the major player and a representative minor player. This method obtains decentralized strategies for the original finite population, which have an ϵ-Nash equilibrium property. Within the LQ framework, the reader is referred to [49] for non-uniform minor players, [12, 28] for partially information, [38] for an application to optimal execution in finance, and [40] for random entrance of agents. Meanwhile, the study of major-minor players in nonlinear models can be found in [7, 8, 10, 18, 19, 50], where the stochastic maximum principle or dynamic programming plays a key role in analyzing an agent's best response control law. Sen and Caines [54] considered partial information and control with nonlinear filtering. Lasry and Lions [44] introduced master equations for a nonlinear major-minor player model. Cardaliaguet et al [14] analyzed a convergence problem for a major-minor player model, where a pair of master equations was obtained as the limit of the Hamilton–Jacobi–Bellman (HJB) equations for the $N+1$ players as $N \to \infty$. In the above contributions, all players choose their strategies simultaneously.

1.2 Stackelberg games and mean field Stackelberg games

Some noncooperative decision problems demonstrate a hierarchy among the players. Historically, von Stackelberg [56] introduced such a game model, later named after him, with a leader and a follower, where the leader moves first while incorporating the follower's response. Dynamic Stackelberg games started in the fundamental work [20, 55], and found important applications in economic theory since 1970s [41, 42].

The study of dynamic Stackelberg games has been mainly founded on three solution concepts and their associated information structures: (i) open-loop solutions, (ii) closed-loop solutions, and (iii) feedback solutions [3, 41, 55]. The open-loop solution can naturally be extended to stochastic dynamic Stackelberg games such that the players can use information of the underlying filtration on the probability space. Each player's control in the open-loop solution is just a function of time in a deterministic model, or a random process adapted to the underlying filtration in a stochastic model [60]. In a closed-loop solution, the leader's strategy is selected as a feedback rule on the entire time horizon all at once while taking into consideration the reaction of the follower(s). The feedback solution is determined by dynamic programming and specifies a feedback rule for each player, but is so-called for a distinction with the closed-loop solution which cannot be characterized by dynamic programming [55]. A further related solution notion is the global Stackelberg solution where the leader announces, under an accompanying information structure, its strategy on the whole time horizon and commits to that strategy, which the follower responds to [3]. The closed-loop solution may be viewed as a global Stackelberg solution under a specific information structure. For feedback solutions, see [3, sec. 7.6] for a continuous-time two-player Stackelberg game, and [41] for a discrete-time model with a leader and N followers playing a Stackelberg–Nash game. The feedback solution of Stackelberg games has time consistency. Informally stated, the solution, as a decision rule on a whole time interval $[0, T]$, still solves the decision problem restricted to any remaining period $[t, T]$.

Within the setting of major-minor players, a natural solution notion is to consider leadership of the major player while all minor players act as followers. In the recent literature on major player models, the analysis of leadership or Stackelberg equilibria can be found in [58] for a discrete time model, [5, 33, 48] for LQ Stackelberg games via stochastic calculus of variations, [6] for nonlinear dynamics with control delay, where each follower has delay in collecting the information of the leader, [26] for mean field principal-agent problems via the stochastic maximum principle, [2] for an application to epidemic control, and [39] for evolutionary inspection games under a major player's pressure. Also see [29, 47], which study leadership in the mean field setting via the stochastic maximum principle and so essentially adopt an open-loop solution. Moreover, different from the major-minor player modeling, both the leader and the follower in the mean field type models of [29, 47] can directly influence the mean field.

1.3 The LQ mean field Stackelberg model

In this paper we consider an LQ mean field Stackelberg game. For this purpose, in (1)–(4) we take the drift terms:

$$\begin{cases} f_0(x_0, \mu, u^0) & = A_0 x_0 + B_0 u^0 + F_0 \langle y \rangle_\mu, \\ f(x_i, x_0, \mu, u^i, u^0) & = A x_i + B u^i + B_1 u^0 + F \langle y \rangle_\mu + G x_0, \end{cases} \tag{5}$$

where $\langle y \rangle_\mu := \int_{\mathbb{R}^n} y\mu(dy)$, and running costs and terminal costs:

$$\begin{cases} L_0(x_0, \mu, u^0) & = |x_0 - \Gamma_0 \langle y \rangle_\mu|^2_{Q_0} + |u^0|^2_{R_0}, \\ L(x_i, x_0, \mu, u^i, u^0) & = |x_i - \Gamma_1 x_0 - \Gamma_2 \langle y \rangle_\mu|^2_Q \\ & \quad + |u^i|^2_R + |u^0|^2_{R_1} + 2u^{iT} R_2 u^0, \\ g_0(x_0, \mu) & = |x_0 - \Gamma_{0f} \langle y \rangle_\mu|^2_{Q_{0f}}, \\ g(x_i, x_0, \mu) & = |x_i - \Gamma_{1f} x_0 - \Gamma_{2f} \langle y \rangle_\mu|^2_{Q_f}, \end{cases} \tag{6}$$

where $|y|^2_M := y^T M y$ for a symmetric matrix $M \geq 0$. For this LQ model, $X^j_t \in \mathbb{R}^n$ and $u^j_t \in \mathbb{R}^{n_1}$ are, respectively, the state and control of $\mathcal{A}_j$, $0 \leq j \leq N$. The initial states $\{X^j_0, 0 \leq j \leq N\}$ are independent with finite second moment. 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 $A_0, B_0, F_0, \sigma_0, A, B, B_1, F, G, \sigma$, Γ_0, Q_0, R_0, Γ_1, Γ_2, Q, R, R_1, R_2, Γ_{0f}, Q_{0f}, Γ_{1f}, Γ_{2f}, and Q_f have compatible dimensions, where Q_0, R_0, Q, R, R_1, Q_{0f}, Q_f are symmetric and $Q_0 \geq 0$, $Q \geq 0$, $Q_{0f} \geq 0$, $Q_f \geq 0$, $R_0 > 0$, $R > 0$, $R_1 > 0$. It is possible to consider a more general form for J^{N+1}_0 and J^{N+1}_i. For instance, one may use $|x_0 - \Gamma_0 \langle y \rangle_\mu - \eta_0|^2_{Q_0}$ with a constant vector η_0 in J^{N+1}_0 and similarly generalize J^{N+1}_i. These more general cases can be easily handled by the method developed in this paper.

A desirable solution of this dynamic Stackelberg game of $N + 1$ players is to seek some form of low-complexity feedback strategies. With the $N + 1$ players, one might try a direct solution for a Stackelberg–Nash equilibrium by dynamic programming. This direct solution, however, becomes unfeasible for large N due to high complexity. Instead, we employ dynamic programming in the mean field limit model consisting of the leader and a representative follower. This approach may be viewed as the mean field counterpart of the feedback solution of Stackelberg games, and leads to the so-called master equations. Master equations have been an important tool to analyze mean field games; see e.g. [4, 7, 13, 17, 21, 44]. Preliminary results of our master equation-based approach have been presented at the conferences [37, 59], and in this paper we provide complete analysis.

By our approach, the solution has time-consistency. In contrast, several other contributions studying mean field Stackelberg games [6, 8, 33, 48] rely on calculus of variations or the stochastic maximum principle, and the resulting equilibria do not have time-consistency. When a decision rule is time inconsistent, a decision-maker lacking commitment will not stick to it under replanning in the future.

After the $N + 1$ players apply the master equation-based strategies φ^0 and φ for the leader and followers, respectively, a fundamental question is how to characterize the equilibrium properties of these strategies. The resulting performance issue, however, becomes intricate. The difficulty stems from the fact that the leader's strategy φ^0, while taking into account instantaneous reactions of all followers, is not guaranteed a minimizer of its cost on $[0, T]$, denoted as $\bar{J}_0(0, u^0, \psi^{u^0})$, where ψ^{u^0} solves the mean field game for all followers when u^0 is announced (see (57)) for details).

To overcome the above difficulty with asymptotic equilibrium characterization, we adopt a method of Ekeland and Lazrak [23] to view the game of $N + 1$ players as being played by a stream of short-lived agents. Accordingly, we only need to consider control perturbation on infinitesimally small time intervals rather than on the whole interval $[0, T]$. Remarkably, the method of [23] was originally introduced to obtain time consistent policies for time-inconsistent optimal control problems. The time inconsistency phenomena of decision problems were observed very early by economists; see e.g. [1, 57] . Later Pollak [51] suggested an equilibrium approach for replanning at a set of discrete times, where the decision maker at different stages is identified as a different entity acting non-cooperatively with others. The extension of the above equilibrium approach to continuous time was developed only much later by Ekeland and Lazrak [23] for optimal control problems with a non-exponential discount. They introduced the notion of t-selves (the decision-maker labelled by time t) making decisions sequentially, and characterized a sub-game perfect equilibrium [23] by use of spike variations of the equilibrium policy. This sub-game perfect equilibrium approach was adopted in [22] to mean field games without Markovian dynamics. For further references applying the technique in [23] to overcome time-inconsistency in optimal control problems, see [9, 24, 25, 32, 61].

The main contributions of this paper are outlined as follows.

- For a mean field limit model of a leader and a representative follower, we introduce a pair of master equations and give a necessary and sufficient condition, in terms of two Riccati ordinary differential equations (ODEs), for the existence of a unique quadratic solution. The connection with LQ mean field games is illuminated when the leader's strategy is based on the master equations and fixed and only the continuum of followers seek alternative strategies.
- To our best knowledge, this is the first contribution to achieve a time-consistent solution in continuous-time mean field Stackelberg games.
- We use the solution of the master equations to construct decentralized strategies for the $N+1$ players, and further evaluate their performance.
- Motivated by the Stackelberg–Nash equilibrium for $N+1$ players, we introduce the notion of ε_N-Stackelberg equilibrium, where each player's feedback strategy is modified following the procedure of Ekeland and Lazrak [23]. By considering agents alive on a short period and using a spike variation of their strategies, we show that the set of decentralized strategies is an ε_N-Stackelberg equilibrium.

1.4 Organization of the paper

The paper is organized as follows. Section 2 introduces the mean field limit model and master equations, which lead to a set of decentralized strategies. Section 3 analyzes quadratic solutions of the master equations and determines existence and uniqueness of such solutions. Section 4 applies these decentralized strategies to the finite-population model. Section 5 analyzes the performance of the decentralized strategies and establishes an ϵ-Stackelberg equilibrium result. Section 6 concludes the paper.

1.5 Notation

Let $C^2(\mathbb{R}^n;\mathbb{R})$ be the set of twice continuously differentiable functions. Let $C_b^2(\mathbb{R}^n;\mathbb{R}^k)$ be the set of $\mathbb{R}^k$-valued functions h with continuous and bounded second order partial derivatives (so that h has at most quadratic growth). For $h \in C^2(\mathbb{R}^n;\mathbb{R})$, its gradient is denoted by $\partial_y h(y)$ or h_y as a row vector, and its Hessian by $\partial_y^2 h$ or h_{yy}.

Denote $\langle \mu, h\rangle = \int h(y)\mu(dy)$, and $\langle y\rangle_\mu = \int y\mu(dy)$ for probability measure μ and function h if the integral is finite. We may indicate dy as in $\langle \mu(dy), h(x,y)\rangle$ when h involves more than one spatial variable. Let $\mathcal{P}_2(\mathbb{R}^n)$ be the set of Borel probability measures on $\mathbb{R}^n$ with finite second moment. 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, $\nu \in \mathcal{P}_2(\mathbb{R}^n)$ as the marginals. Then $(\mathcal{P}_2(\mathbb{R}^n), W_2)$ is a complete metric space [13]. For a function $h(t,x,\mu)$, we follow [15] to define $(\delta h/\delta\mu)(t,x,\mu,y)$ as the derivative with respect to the measure μ when it exists, and will use the short notation $\delta_\mu h(t,x,\mu;y)$. Here δ_μ indicates differentiation and shall not be confused with the dirac measure.

For a vector or matrix Y, its Frobenius norm is denoted by $|Y|$. For a symmetric matrix $M \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{n\times k}$, we denote $[\![Y]\!]_M^2 = Y^TMY$; if we further have $M \geq 0$ and $y \in \mathbb{R}^n$, we denote $|y|_M^2 = y^TMy$. For vectors and/or matrices, we sometimes write a product vw as $v\cdot w$ for ease of reading, especially when the product involves long expressions.

Letting $h(\epsilon)$ be nonnegative and defined on $[0,\epsilon_0]$ for some small $\epsilon_0 > 0$, we use $\mathcal{O}_{k\times l}(h(\epsilon))$ to denote a $k\times l$ matrix such that for a fixed constant $C>0$, $|\mathcal{O}_{k\times l}(h(\epsilon))| \leq Ch(\epsilon)$ holds for all sufficiently small

$\epsilon \geq 0$. Sometimes we drop $k \times l$ when the dimensions are clear from the context. If $k = l = 1$, we simply write $\mathcal{O}(h(\epsilon))$ as $O(h(\epsilon))$. For a function $h_1(\epsilon)$, we write $h_1(\epsilon) = o(h(\epsilon))$ if $\lim_{\epsilon\to 0^+} h_1(\epsilon)/h(\epsilon) = 0$. We similarly define $O(c_k)$ and $o(c_k)$ for a sequence $\{c_k, k \geq 1\}$.

Let I and 0 be, respectively, identity matrices and zero matrices of compatible dimensions. For N vectors Y_k, $1 \leq k \leq N$, of the same dimension, denote $Y^{(N)} = \frac{1}{N}\sum_{k=1}^N Y_k$ and $Y^{(-k)} = \frac{1}{N}\sum_{j=1,j\neq k}^N Y_j$. In some proofs, a vector or scalar function $h(t)$ of $t \in [0, T]$ (such as $\hat{\mathbb{P}}(t)$, $\hat{r}(t)$) is sometimes written as h_t. Denote $\Sigma_w = \sigma\sigma^T$ and $\Sigma_{w^0} = \sigma_0\sigma_0^T$.

Throughout the paper, we use the agent index from $\{0, \cdots, N\}$ to label a variable, a process, a function, and so on (for instance, X_s^i, u_t^0, Z_t^i, $\mathbb{P}^0$). They are always interpreted as a superscript, but not as an exponent. We use C, C_0, C_1, etc, to denote generic constants, which do not depend on (t, N) and may vary from place to place. When entries like u^i, $\mathbf{u}^F$, etc, appear as an argument of a cost functional (such as $J_0^{N+1}(0, u^0, \mathbf{u}^F)$), they mean a control law or a strategy, not a vector in a Euclidean space.

2 Mean field limit model and master equations

Based on the $(N+1)$-player model specified by (1)–(6), we consider a mean field limit Stackelberg game model which involves the leader $\mathcal{A}_0$ with state Z_s^0 and a representative follower $\mathcal{A}_i$ with state Z_s^i. The two players $\mathcal{A}_0$ and $\mathcal{A}_i$ have the mean field limit dynamics:

$$dZ_s^0 = f_0(Z_s^0, \mu_s, u_s^0)ds + \sigma_0 dW_s^0, \tag{7}$$
$$dZ_s^i = f(Z_s^i, Z_s^0, \mu_s, u_s^i, u_s^0)ds + \sigma dW_s^i, \tag{8}$$

where μ_s is the mean field generated at time s by a continuum of followers. For convenience of later analysis, here we use s to denote time. The initial condition is given by (X_0^0, X_0^i). Equations (7)–(8) are obtained from (1)–(2) after approximating $\mu_s^{(N)}$ by μ_s. The measure flow $\{\mu_s, s \geq 0\}$ drives the evolution of (Z_s^0, Z_s^i).

For this two-player mean field model, we take (Z_s^0, Z_s^i, μ_s) as the state variable. We consider closed-loop perfect state (CLPS) information and adopt state feedback strategies of the form

$$u_s^0 = \psi^0(s, Z_s^0, \mu_s), \qquad u_s^i = \psi(s, Z_s^i, Z_s^0, \mu_s), \tag{9}$$

where we have $\psi^0 : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}^{n_1}$ and $\psi : [0, T] \times \mathbb{R}^{2n} \times \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}^{n_1}$.

Following the consistent mean field approximation methodology in mean field games [36], below we specify the dynamics of μ_s. We apply feedback strategies of the form (9) to (7)–(8) for $1 \leq i \leq N$. So (8) is replicated to generate N processes driven by independent Brownian motions. Let $\mu_s^{(N)}$ be the empirical distribution of $(Z_s^1, \cdots, Z_s^N)$. For $h \in C_b^2(\mathbb{R}^n; \mathbb{R})$, by (8) and Itô's formula we have

$$\begin{aligned} d\langle \mu_s^{(N)}, h\rangle =& \langle \mu_s^{(N)}(dy),\ \partial_y h(y) \cdot f(y, Z_s^0, \mu_s, \psi(s, y, Z_s^0, \mu_s), \psi^0(s, Z_s^0, \mu_s)) \\ &+ (1/2)\mathrm{Tr}[\partial_y^2 h(y)\Sigma_w]\rangle ds + \frac{1}{N}\sum_{i=1}^N \partial_y h(Z_s^i) \cdot \sigma dW_s^i. \end{aligned} \tag{10}$$

The consistency condition stipulates that as $N \to \infty$, $\mu_s^{(N)}$ has μ_s as its limit, and we formally write (10) in a limit form, which is the following ODE in a weak form:

$$\begin{aligned} \frac{d}{ds}\langle \mu_s, h\rangle =& \langle \mu_s(dy),\ \partial_y h(y) \cdot f(y, Z_s^0, \mu_s, \psi(s, y, Z_s^0, \mu_s), \psi^0(s, Z_s^0, \mu_s)) \\ &+ (1/2)\mathrm{Tr}[\partial_y^2 h(y)\Sigma_w]\rangle, \end{aligned} \tag{11}$$

where the initial condition $\mu_0 = \mu_0^X$ is determined by the initial states of the continuum of followers. We call (11) the consistency condition. The pair of strategies (ψ^0, ψ) is called admissible if the resulting closed-loop system has a well-defined solution (Z_s^0, Z_s^i, μ_s) for (7)–(8) and (11).

Let (u^0, u^i) be a pair of admissible feedback strategies defined on $[0, T]$. Now consider (7)–(8) and (11) with a general initial time $t \in [0, T)$ and initial states $Z_t^0 = z_0$, $Z_t^i = z_i$, $\mu_t = \mu \in \mathcal{P}_2(\mathbb{R}^n)$. Define

$$\bar{J}_0(t, z_0, \mu, u^0, u^i) = \mathbb{E}\int_t^T L_0(Z_s^0, \mu_s, u_s^0)ds + \mathbb{E}g_0(Z_T^0, \mu_T), \tag{12}$$

$$\bar{J}_i(t, z_i, z_0, \mu, u^i, u^0) = \mathbb{E}\int_t^T L(Z_s^i, Z_s^0, \mu_s, u_s^i, u_s^0)ds + \mathbb{E}g(Z_T^i, Z_T^0, \mu_T). \tag{13}$$

Due to the arbitrary choice of z_i and μ at time t, μ_s in general is not equal to the distribution (or the conditional distribution given $\{Z_\tau^0, \tau \le s\}$) of Z_s^i.

Below we elaborate on the determination of the feedback strategies $(u^{0*}, u^{i*}) = (\varphi^0(t, z_0, \mu), \varphi(t, z_i, z_0, \mu))$, if they exist, by dynamic programming equations. Let the value functions be

$$V_0(t, z_0, \mu) = \bar{J}_0(t, z_0, \mu, u^{0*}, u^{i*}), \tag{14}$$

$$V(t, z_i, z_0, \mu) = \bar{J}_i(t, z_i, z_0, \mu, u^{i*}, u^{0*}), \tag{15}$$

where $t \in [0, T]$, $z_0, z_i \in \mathbb{R}^n$, and $\mu \in \mathcal{P}_2(\mathbb{R}^n)$.

2.1 Dynamic programming

For fixed vectors $u^0, u^i, v \in \mathbb{R}^{n_1}$, define the following differential operators associated with the processes (7)–(8):

$$(\mathcal{L}_0^{u^0} h)(z_0) = h_{z_0}(z_0) \cdot f_0(z_0, \mu, u^0) + \frac{1}{2}\mathrm{Tr}[h_{z_0 z_0}(z_0)\Sigma_{w^0}], \tag{16}$$

$$(\mathcal{L}^{u^i, u^0} h)(z_i) = h_{z_i}(z_i) \cdot f(z_i, z_0, \mu, u^i, u^0) + \frac{1}{2}\mathrm{Tr}[h_{z_i z_i}(z_i)\Sigma_w], \tag{17}$$

$$(\mathcal{L}_{\mathrm{mf}}^{v, u^0} h)(y) = h_y(y) \cdot f(y, z_0, \mu, v, u^0) + \frac{1}{2}\mathrm{Tr}[h_{yy}(y)\Sigma_w], \tag{18}$$

where $h \in C^2(\mathbb{R}^n; \mathbb{R})$. Here v is the control of a generic follower with state y from an infinite population. Throughout Sections 2.1 and 2.2 we follow the rule in (16)–(18) as to which variable is used in the differentiation.

We proceed to apply dynamic programming by assuming that the value functions (V_0, V) are well-defined on $[0, T]$ with sufficient regularity. We introduce the HJB equation system

$$\begin{cases} 0 = \partial_t V_0(t, z_0, \mu) + \mathcal{L}_0^{u^{0*}} V_0 + L_0(z_0, \mu, u^{0*}) \\ \qquad\qquad + \langle \mu(dy),\ \mathcal{L}_{\mathrm{mf}}^{\varphi(t,y,z_0,\mu), u^{0*}} \delta_\mu V_0(t, z_0, \mu; y)\rangle, \\ 0 = \partial_t V(t, z_i, z_0, \mu) + (\mathcal{L}_0^{u^{0*}} + \mathcal{L}^{u^{i*}, u^{0*}})V + L(z_i, z_0, \mu, u^{i*}, u^{0*}) \\ \qquad\qquad + \langle \mu(dy),\ \mathcal{L}_{\mathrm{mf}}^{\varphi(t,y,z_0,\mu), u^{0*}} \delta_\mu V(t, z_i, z_0, \mu; y)\rangle, \end{cases} \tag{19}$$

where $(t, z_i, z_0, \mu) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)$,

$$V_0(T, z_0, \mu) = g_0(z_0, \mu), \qquad V(T, z_i, z_0, \mu) = g(z_i, z_0, \mu).$$

We still need to specify $u^{0*} = \varphi^0(t, z_0, \mu)$ and $u^{i*} = \varphi(t, z_i, z_0, \mu)$ in terms of (V_0, V). The description of the procedure is postponed to Section 2.2.

We also call (19) the *master equations*. Here $\mathcal{L}_{\mathrm{mf}}^{v, u^0}$ acts on $\delta_\mu V_0$ and $\delta_\mu V$ via the y variable, with (t, z_i, z_0, μ) fixed. Note that $\delta_\mu V_0$ and $\delta_\mu V$ have the extra independent variable y.

Remark 2.1. For the master equations to be meaningful, we look for a solution pair (V_0, V) with those properties: (i) for fixed $\mu \in \mathcal{P}_2(\mathbb{R}^n)$, we have $V_0 \in C^{1,2}([0,T]\times\mathbb{R}^n)$ and $V \in C^{1,2}([0,T]\times\mathbb{R}^{2n})$; (ii) the functions $\delta_\mu V_0$, $(\delta_\mu V_0)_y$, $(\delta_\mu V_0)_{yy}$ (resp., $\delta_\mu V$, $(\delta_\mu V)_y$, $(\delta_\mu V)_{yy}$) are continuous in (t, z_0, μ, y) (resp., (t, z_i, z_0, μ, y)); (iii) the integrands for $\mu(dy)$ in (19), as a function of y, have quadratic growth.

If the strategies $(u^{0*}, u^{i*}) = (\varphi^0(t, z_0, \mu), \varphi(t, z_i, z_0, \mu))$ exist for the master equations in (19), they are called mean field feedback Stackelberg strategies.

2.2 Selection of reaction functions

Step 1. For each vector $u^0 \in \mathbb{R}^{n_1}$, let $\phi^{u^0}(t, y, z_0, \mu)$ be the instantaneous reaction function (with arguments (t, y, z_0, μ, u^0)) of a generic follower, where y is its state value. As a means to find $\phi^{u^0}(t, y, z_0, \mu)$, below we determine the reaction function $\hat{u}^i = \phi^{u^0}(t, z_i, z_0, \mu)$ of the representative player $\mathcal{A}_i$. Denote

$$
\begin{aligned}
\mathcal{H} =& L(z_i, z_0, \mu, u^i, u^0) + (\mathcal{L}_0^{u^0} + \mathcal{L}^{u^i,u^0})V(t, z_i, z_0, \mu) \\
&+ \langle \mu(dy),\ \mathcal{L}_{\mathrm{mf}}^{\phi^{u^0}(t,y,z_0,\mu),u_0} \delta_\mu V(t, y, z_0, \mu; y)\rangle,
\end{aligned}
$$

We interpret $-\mathcal{H}$ as the Hamiltonian of $\mathcal{A}_i$. Player $\mathcal{A}_i$ optimizes u^i, which is only contained in L and $\mathcal{L}^{u^i,u^0}V(t, z_i, z_0, \mu)$. Since $R > 0$, the minimizer of $\mathcal{H}$ is determined by the first order condition:

$$
\begin{aligned}
u^i &= \phi^{u^0}(t, z_i, z_0, \mu) \\
&= -\frac{1}{2}R^{-1}[B^T\partial_{z_i}^T V(t, z_i, z_0, \mu) + 2R_2u^0],
\end{aligned}
\tag{20}
$$

which is the instantaneous reaction of $\mathcal{A}_i$ to $u^0 \in \mathbb{R}^{n_1}$ given (t, z_i, z_0, μ).

Step 2. Next we consider the leader's reaction when all followers have adopted ϕ^{u^0} in (20) by matching with their own states. In view of (19), denote

$$
\begin{aligned}
\mathcal{H}_0 =& L_0(z_0, \mu, u^0) + \mathcal{L}_0^{u^0} V_0(t, z_0, \mu) \\
&+ \langle \mu(dy),\ \mathcal{L}_{\mathrm{mf}}^{\phi^{u^0}(t,y,z_0,\mu),u^0} \delta_\mu V_0(t, z_0, \mu; y)\rangle.
\end{aligned}
$$

Since $R_0 > 0$, the minimizer of $\mathcal{H}_0$ is determined by the first order condition:

$$
\begin{aligned}
u^{0*} &= -\frac{1}{2}R_0^{-1}[B_0^T\partial_{z_0}^T V_0(t, z_0, \mu) + \widetilde{B}_1^T\langle \mu(dy),\ \partial_y^T\delta_\mu V_0(t, z_0, \mu; y)\rangle] \\
&=: \varphi^0(t, z_0, \mu),
\end{aligned}
\tag{21}
$$

where

$$
\widetilde{B}_1 = B_1 - BR^{-1}R_2. \tag{22}
$$

Substituting (21) into (20) gives

$$
\begin{aligned}
u^{i*} &= -\frac{1}{2}R^{-1}B^T\partial_{z_i}^T V(t, z_i, z_0, \mu) + \frac{1}{2}R^{-1}R_2R_0^{-1}\cdot \\
&\quad \big[B_0^T\partial_{z_0}^T V_0(t, z_0, \mu) + \widetilde{B}_1^T\langle \mu(dy),\ \partial_y^T\delta_\mu V_0(t, z_0, \mu; y)\rangle\big] \\
&=: \varphi(t, z_i, z_0, \mu).
\end{aligned}
\tag{23}
$$

Remark 2.2. The master equations in (19) can be formally derived by combining the local expansion of the value function and the above reaction function selection in (21) and (23). We show how the integral terms in (19) arise. Let u^0 be fixed. On $[t, t+\epsilon]$ with the initial condition (z^i, z^0, μ) at t, we take a Taylor expansion of $V(t+\epsilon, Z_{t+\epsilon}^i, Z_{t+\epsilon}^0, \mu_{t+\epsilon})$. In particular, we have the first order approximation term

$$
\int_{\mathbb{R}^n} \delta_\mu V(t, z_i, z_0, \mu; y)(\mu_{t+\epsilon}(dy) - \mu(dy))
$$

$$= \epsilon\langle\mu(dy),\ \partial_y\delta_\mu V(t,z_i,z_0,\mu;y)\cdot f(y,z_0,\mu,v,u^0) \\ + (1/2)\mathrm{Tr}\big[\partial_y^2\delta_\mu V(t,z_i,z_0,\mu;y)\Sigma_w\big]\rangle + o(\epsilon).$$

Subsequently, we determine the reaction functions so that the pair (u^0, v) within f is taken as the equilibrium strategies $(\varphi^0(t,z_0,\mu), \varphi(t,y,z_0,\mu))$. The integral term in the equation of V_0 arises for similar reasons.

This section only constructs the equations in (19). The existence analysis for these equations will be investigated in the next section.

The pair in (21) and (23) is called the mean field feedback Stackelberg strategies for the mean field Stackelberg game specified by (7)–(8) and (12)–(13). Under the equilibrium strategies in (21)–(23), we may further write the closed-loop dynamics for (Z_s^0, Z_s^i, μ_s).

3 Quadratic solutions of the master equations

We exploit the linear-quadratic structure to write the master equations in a more explicit form. By (5) and (6), we reduce (19) to the following two equations:

$$\begin{aligned}
-\partial_t V_0(t,z_0,\mu) &= \partial_{z_0}V_0\cdot(A_0z_0 + F_0\langle y\rangle_\mu) \\
&\quad + \frac{1}{2}\mathrm{Tr}\big(\partial_{z_0}^2 V_0\Sigma_{w^0}\big) + |z_0 - \Gamma_0\langle y\rangle_\mu|_{Q_0}^2 \\
&\quad + \langle\mu(dy),\ \partial_y\delta_\mu V_0(t,z_0,\mu;y)\cdot(Ay + F\langle y\rangle_\mu + Gz_0)\rangle \\
&\quad - \frac{1}{2}\langle\mu(dy),\ \partial_y\delta_\mu V_0(t,z_0,\mu;y)\cdot BR^{-1}B^T\partial_y^T V(t,y,z_0,\mu)\rangle \\
&\quad - \frac{1}{4}\Big|B_0^T\partial_{z_0}^T V_0 + \widetilde{B}_1^T\langle\mu(dy),\ \partial_y^T\delta_\mu V_0(t,z_0,\mu;y)\rangle\Big|_{R_0^{-1}}^2 \\
&\quad + \frac{1}{2}\mathrm{Tr}\big[\langle\mu(dy),\ \partial_y^2\delta_\mu V_0(t,z_0,\mu;y)\Sigma_w\rangle\big],
\end{aligned} \tag{24}$$

$$\begin{aligned}
-\partial_t V(t,z_i,z_0,\mu) &= \partial_{z_0}V\cdot(A_0z_0 + F_0\langle y\rangle_\mu) + \partial_{z_i}V\cdot(Az_i + F\langle y\rangle_\mu + Gz_0) \\
&\quad - \frac{1}{4}\partial_{z_i}VBR^{-1}B^T\partial_{z_i}^T V \\
&\quad + \frac{1}{2}\mathrm{Tr}(\partial_{z_0}^2 V\Sigma_{w^0} + \partial_{z_i}^2 V\Sigma_w) + |z_i - \Gamma_1 z_0 - \Gamma_2\langle y\rangle_\mu|_Q^2 \\
&\quad + \langle\mu(dy),\ \partial_y\delta_\mu V(t,z_i,z_0,\mu;y)\cdot(Ay + F\langle y\rangle_\mu + Gz_0)\rangle \\
&\quad - \frac{1}{2}\langle\mu(dy),\ \partial_y\delta_\mu V(t,z_i,z_0,\mu;y)\cdot BR^{-1}B^T\partial_y^T V(t,y,z_0,\mu)\rangle \\
&\quad + \frac{1}{4}\Big[\!\Big[R_0^{-1}\big[B_0^T\partial_{z_0}^T V_0 + \widetilde{B}_1^T\langle\mu(dy),\ \partial_y^T\delta_\mu V_0(t,z_0,\mu;y)\rangle\big]\Big]\!\Big]_{R_{12}}^2 \\
&\quad - \frac{1}{2}\big[\partial_{z_0}V_0\cdot B_0 + \langle\mu(dy),\ \partial_y\delta_\mu V_0(t,z_0,\mu;y)\rangle\widetilde{B}_1\big]R_0^{-1}\cdot \\
&\qquad \big[B_0^T\partial_{z_0}^T V + \widetilde{B}_1^T\big\{\partial_{z_i}^T V + \langle\mu(dy),\ \partial_y^T\delta_\mu V(t,z_i,z_0,\mu;y)\rangle\big\}\big] \\
&\quad + \frac{1}{2}\mathrm{Tr}\big[\langle\mu(dy),\ \partial_y^2\delta_\mu V(t,z_i,z_0,\mu;y)\Sigma_w\rangle\big],
\end{aligned} \tag{25}$$

where

$$R_{12} = R_1 - R_2^T R^{-1} R_2, \tag{26}$$

and the terminal conditions are

$$\begin{aligned}
V_0(T,z_0,\mu) &= |z_0 - \Gamma_{0f}\langle y\rangle_\mu|_{Q_{0f}}^2, \\
V(T,z_i,z_0,\mu) &= |z_i - \Gamma_{1f}z_0 - \Gamma_{2f}\langle y\rangle_\mu|_{Q_f}^2.
\end{aligned}$$

Denote $\mathbf{z}_0 = (z_0^T, \bar{z}^T)^T \in \mathbb{R}^{2n}$ and $\mathbf{z} = (z_i^T, z_0^T, \bar{z}^T)^T \in \mathbb{R}^{3n}$, where $\bar{z} = \langle y \rangle_\mu$. We are interested in solutions of the form

$$V_0(t, z_0, \mu) = \mathbf{z}_0^T \mathbb{P}^0(t)\mathbf{z}_0 + r^0(t), \tag{27}$$

$$V(t, z_i, z_0, \mu) = \mathbf{z}^T \mathbb{P}(t)\mathbf{z} + r(t), \qquad 0 \le t \le T, \tag{28}$$

where $\mathbb{P}^0$ and $\mathbb{P}$ are symmetric matrix functions of $t \in [0, T]$ and have the partition

$$\mathbb{P}^0 = (P^0_{kl})_{1\le k,l,\le 2}, \qquad \mathbb{P} = (P_{kl})_{1\le k,l\le 3}.$$

Each submatrix P^0_{kl} or P_{kl} has dimensions $n \times n$. Moreover, $r^0(t)$ and $r(t)$ are functions from $[0, T]$ to $\mathbb{R}$.

Definition 3.1. We call the pair (V_0, V) in (27)–(28) a quadratic solution for master Equations (24)–(25) if the pair satisfies (24)–(25).

Denote

$$\mathbb{A}_0(t) = \begin{bmatrix} A_0 & F_0 \\ G - BR^{-1}B^T P_{12} & A + F - BR^{-1}B^T(P_{11} + P_{13}) \end{bmatrix},$$

$$\mathbb{A}(t) = \begin{bmatrix} A & (G, F) \\ 0 & \mathbb{A}_0 \end{bmatrix}, \quad \mathbb{B}_0 = \begin{bmatrix} B_0 \\ \widetilde{B}_1 \end{bmatrix}, \quad \mathbb{B} = \begin{bmatrix} B \\ 0_{2n\times n_1} \end{bmatrix} \quad \mathbb{B}_1 = \begin{bmatrix} \widetilde{B}_1 \\ \mathbb{B}_0 \end{bmatrix},$$

$$\mathbb{J}_1 = [0_{2n\times n}, I_{2n}],$$

where $\widetilde{B}_1$ is given by (22). Let R_{12} be defined by (26). We introduce the Riccati ODE system:

$$-\dot{\mathbb{P}}^0 = \mathbb{P}^0\mathbb{A}_0 + \mathbb{A}_0^T\mathbb{P}^0 - \mathbb{P}^0\mathbb{B}_0R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0 + [\![(I, -\Gamma_0)]\!]^2_{Q_0}, \tag{29}$$

$$\begin{aligned} -\dot{\mathbb{P}} =& \mathbb{P}\mathbb{A} + \mathbb{A}^T\mathbb{P} - \mathbb{P}\mathbb{B}R^{-1}\mathbb{B}^T\mathbb{P} + [\![(I, -\Gamma_1, -\Gamma_2)]\!]^2_Q \\ &- \mathbb{P}\mathbb{B}_1R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1 - \mathbb{J}_1^T\mathbb{P}^0\mathbb{B}_0R_0^{-1}\mathbb{B}_1^T\mathbb{P} \\ &+ \mathbb{J}_1^T\mathbb{P}^0\mathbb{B}_0R_0^{-1}R_{12}R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1, \end{aligned} \tag{30}$$

where

$$\mathbb{P}^0(T) = [\![(I, -\Gamma_{0f})]\!]^2_{Q_{0f}}, \quad \mathbb{P}(T) = [\![(I, -\Gamma_{1f}, -\Gamma_{2f})]\!]^2_{Q_f}.$$

Note that the coefficient matrices $\mathbb{A}_0$ and $\mathbb{A}$ depend on $\mathbb{P}$. We further introduce the following ODEs:

$$-\dot{r}^0 = \mathrm{Tr}(P^0_{11}\Sigma_{w^0}), \quad r^0(T) = 0, \tag{31}$$

$$-\dot{r} = \mathrm{Tr}(P_{22}\Sigma_{w^0} + P_{11}\Sigma_w), \quad r(T) = 0. \tag{32}$$

Remark 3.2. If the system (29)–(30) admits a solution $(\mathbb{P}^0, \mathbb{P})$ on $[0, T]$, the solution is unique since the vector field of the ODE system is locally Lipschitz along the solution trajectory. Subsequently, we further obtain a unique solution (r^0, r) on $[0, T]$.

Theorem 3.3. The master equation system (24)–(25) has a quadratic solution of the form (27)–(28) on $[0, T]$ if and only if $(\mathbb{P}^0, \mathbb{P})$ is a solution of (29)–(30) on $[0, T]$.

Proof. See appendix A. □

If $(\mathbb{P}^0, \mathbb{P})$ is a solution of (29)–(30), we construct the quadratic solution (27)–(28), which has the regularity properties in Remark 2.1.

3.1 Feedback strategies

We proceed to analyze the closed-loop systems under feedback strategies. Although the master equations give the control laws as a function of the agent states and the mean field μ, it turns out we only need to use the first moment of μ, which simplifies the computation and implementation of the control laws.

Proposition 3.4. Suppose that the system (29)–(30) has a solution $(\mathbb{P}^0, \mathbb{P})$ on $[0, T]$. Then for master equations in (19), the pair

$$\varphi^0(t, z_0, \mu) = K_1^0 z_0 + K_2^0 \langle y \rangle_\mu, \tag{33}$$

$$\varphi(t, z_i, z_0, \mu) = K_1 z_i + K_2 z_0 + K_3 \langle y \rangle_\mu, \tag{34}$$

$$t \in [0, T],\ z_0, z_i \in \mathbb{R}^n,\ \mu \in \mathcal{P}_2(\mathbb{R}^n),$$

gives mean field feedback Stackelberg strategies, where

$$[K_1^0, K_2^0] = -R_0^{-1} \mathbb{B}_0^T \mathbb{P}^0, \tag{35}$$

$$[K_1, K_2, K_3] = -R^{-1} \mathbb{B}^T \mathbb{P} + R^{-1} R_2 R_0^{-1} \mathbb{B}_0^T \mathbb{P}^0 \mathbb{J}_1. \tag{36}$$

Proof. If (29)–(30) admits a solution $(\mathbb{P}^0, \mathbb{P})$ on $[0, T]$, the master equation system (24)–(25) has a unique quadratic solution (27)–(28) by Theorem 3.3 and Remark 3.2. We substitute (27)–(28) into (21) and (23) to obtain (33)–(34). □

The feedback strategies are implemented using the actual states (Z_t^i, Z_t^0, μ_t) of the mean field limit model. The resulting control laws take the form

$$u_t^{0*} = K_1^0 Z_t^0 + K_2^0 \langle y \rangle_{\mu_t}, \tag{37}$$

$$u_t^{i*} = K_1 Z_t^i + K_2 Z_t^0 + K_3 \langle y \rangle_{\mu_t}, \tag{38}$$

and $\bar{Z}_t := \langle y \rangle_{\mu_t}$ satisfies the following equation

$$\begin{aligned} \dot{\bar{Z}}_t = {} & [A + F + B(K_1 + K_3) + B_1 K_2^0] \bar{Z}_t \\ & + (G + BK_2 + B_1 K_1^0) Z_t^0. \end{aligned} \tag{39}$$

The initial condition is $\bar{Z}_0 = \langle y \rangle_{\mu_0^X}$, where μ_0^X is the limit empirical state distribution of the followers. The ODE of $\bar{Z}_t$ is constructed from (11) by taking $h(y) = y$ and setting the control laws $\psi^0 = u^{0*}$ and $\psi = u^{i*}$. The process $\bar{Z}_t$ is interpreted as the average state of the continuum of followers.

Below we evaluate the costs under the equilibrium strategies (37)–(38). We take a general initial condition $(t, z_i, z_0, \bar{z})$ for the system

$$dZ_s^i = (AZ_s^i + Bu_s^{i*} + B_1 u_s^{0*} + F\bar{Z}_s + GZ_s^0) ds + \sigma dW_s^i, \tag{40}$$

$$dZ_s^0 = (A_0 Z_s^0 + B_0 u_s^{0*} + F_0 \bar{Z}_s) ds + \sigma_0 dW_s^0, \tag{41}$$

$$\dot{\bar{Z}}_s = [A + F + B(K_1 + K_3) + B_1 K_2^0] \bar{Z}_s + (G + BK_2 + B_1 K_1^0) Z_s^0, \tag{42}$$

$$s \in [t, T],$$

which has a unique strong solution. Under the equilibrium strategies, the costs in (12)–(13) are now written as

$$\bar{J}_0^*(t, z_0, \bar{z}) = \mathbb{E} \int_t^T (|Z_s^0 - \Gamma_0 \bar{Z}_s|_{Q_0}^2 + |u_s^{0*}|_{R_0}^2) ds + \mathbb{E}|Z_T^0 - \Gamma_{0f} \bar{Z}_T|_{Q_{0f}}^2 \tag{43}$$

$$\begin{aligned} \bar{J}_i^*(t, z_i, z_0, \bar{z}) = \mathbb{E} \int_t^T \Big(& |Z_s^i - \Gamma_1 Z_s^0 - \Gamma_2 \bar{Z}_s|_Q^2 + |u_s^{i*}|_R^2 + |u_s^{0*}|_{R_1}^2 \\ & + 2u_s^{i*T} R_2 u_s^{0*} \Big) ds + \mathbb{E}|Z_T^i - \Gamma_{1f} Z_T^0 - \Gamma_{2f} \bar{Z}_T|_{Q_f}^2. \end{aligned} \tag{44}$$

Proposition 3.5. Assume that the system (29)–(30) has a solution $(\mathbb{P}^0, \mathbb{P})$ on $[0, T]$. Then we have

$$\begin{cases} \bar{J}_0^*(t, z_0, \bar{z}) = \mathbf{z}_0^T \mathbb{P}^0(t)\mathbf{z}_0 + r^0(t), \\ \bar{J}_i^*(t, z_i, z_0, \bar{z}) = \mathbf{z}^T \mathbb{P}(t)\mathbf{z} + r(t), \end{cases} \tag{45}$$

where $\mathbf{z}_0 = (z_0^T, \bar{z}^T)^T \in \mathbb{R}^{2n}$ and $\mathbf{z} = (z_i^T, z_0^T, \bar{z}^T)^T \in \mathbb{R}^{3n}$.

Proof. The representation in (45) follows from setting $\mu = \delta_{\bar{z}}$ in $V_0(t, z_0, \mu)$ and $V(t, z_i, z_0, \mu)$. □

Remark 3.6. An alternative method can be used to obtain the cost representation in Proposition 3.5. Let $(\mathbb{P}^0, r^0)$ be given. By Lemma B.1, we write $\bar{J}_0^*(t, z_0, \bar{z}) = \mathbf{z}_0^T \bar{\mathbb{P}}^0(t)\mathbf{z}_0 + \bar{r}^0(t)$, where $\bar{\mathbb{P}}^0$ and $\bar{r}^0$ are determined by two linear ODEs with a unique solution. Specifically, we have

$$\begin{aligned} &-\dot{\bar{\mathbb{P}}}^0 = \bar{\mathbb{P}}^0 \widetilde{\mathbb{A}}_0 + \widetilde{\mathbb{A}}_0^T \bar{\mathbb{P}}^0 + [\![(I, -\Gamma_0)]\!]^2_{Q_0} + [\![(K_1^0, K_2^0)]\!]^2_{R_0}, \\ &\bar{\mathbb{P}}^0(T) = [\![(I, -\Gamma_{0f})]\!]^2_{Q_{0f}}, \end{aligned}$$

where $\widetilde{\mathbb{A}}_0$ is the coefficient matrix of $(Z_s^0, \bar{Z}_s)$ for the equation system (41)–(42). Then we can verify that $\mathbb{P}^0$ is a solution for the ODE of $\bar{\mathbb{P}}^0$. Furthermore, we can take the solution $\bar{r}^0 = r^0$. Hence this verifies the first equality in (45). We similarly re-derive the expression of $\bar{J}_i^*$.

3.2 A further characterization of strategies of followers

Suppose that the leader announces its strategy $u_t^{0*} = K_1^0 Z_t^0 + K_2^0 \bar{Z}_t$ for the time interval $[0, T]$. We let the continuum of followers re-choose their strategies to solve a mean field game. A natural question is what strategies the followers will take.

Consider the representative agent $\mathcal{A}_i$ with control u^i when the leader applies u_t^{0*} and the continuum of other followers have generated a mean field process $\bar{Z}_t$. In this case, player $\mathcal{A}_i$ solves an optimal control problem with dynamics

$$d \begin{bmatrix} Z_t^i \\ Z_t^0 \\ \bar{Z}_t \end{bmatrix} = \overline{\mathbb{A}} \begin{bmatrix} Z_t^i \\ Z_t^0 \\ \bar{Z}_t \end{bmatrix} dt + \mathbb{B} u_t^i dt + \begin{bmatrix} \sigma dW_t^i \\ \sigma_0 dW_t^0 \\ 0 \end{bmatrix}, \quad 0 \le t \le T, \tag{46}$$

where $Z_0^i = X_0^i$, $Z_0^0 = X_0^0$, $\bar{Z}_0 = \langle y \rangle_{\mu_0^X}$, and

$$\overline{\mathbb{A}} = \begin{bmatrix} A & G + B_1 K_1^0 & F + B_1 K_2^0 \\ 0 & A_0 + B_0 K_1^0 & F_0 + B_0 K_2^0 \\ 0 & \overline{G} & \overline{A} \end{bmatrix}, \quad \mathbb{B} = \begin{bmatrix} B \\ 0_{2n \times n_1} \end{bmatrix}. \tag{47}$$

In the above, K_1^0 and K_2^0 are known and defined by (35). The matrix functions $\overline{A}(t)$ and $\overline{G}(t)$ will be determined below as part of the solution of the mean field game. The equations of $(Z_t^i, \bar{Z}_t)$ within (46) have been modified from (40) and (42) by setting a general control u_t^i.

The cost functional of player $\mathcal{A}_i$ is given by

$$\begin{aligned} \bar{J}_i^{\mathrm{mfg}}(u^i) = \mathbb{E} \int_0^T &\Big[|Z_t^i - \Gamma_1 Z_t^0 - \Gamma_2 \bar{Z}_t|_Q^2 + |K_1^0 Z_t^0 + K_2^0 \bar{Z}_t|_{R_1}^2 \\ &+ |u_t^i|_R^2 + 2 u_t^{iT} R_2 (K_1^0 Z_t^0 + K_2^0 \bar{Z}_t) \Big] dt \\ &+ \mathbb{E} |Z_T^i - \Gamma_{1f} Z_T^0 - \Gamma_{2f} \bar{Z}_T|_{Q_f}^2, \end{aligned} \tag{48}$$

which is constructed from (44) after taking initial condition $(0, z_i, z_0, \bar{z})$ and replacing control u_t^{i*} by u_t^i. Due to the last term in the integrand, the control problem may be indefinite.

The solution procedure of the mean field game consists of two steps.

Step 1. Assume that $(\overline{A},\overline{G})$ has been known. We start by introducing the following Riccati equation

$$
\begin{aligned}
-\dot{\overline{\mathbb{P}}}(t) &= \overline{\mathbb{P}}\,\overline{\mathbb{A}} + \overline{\mathbb{A}}^T\overline{\mathbb{P}} - \overline{\mathbb{P}}\mathbb{B}R^{-1}\mathbb{B}^T\overline{\mathbb{P}} \\
&\quad - \overline{\mathbb{P}}\mathbb{B}R^{-1}R_2[0,K_1^0,K_2^0] - [0,K_1^0,K_2^0]^T R_2^T R^{-1}\mathbb{B}^T\overline{\mathbb{P}} \\
&\quad + [\![(I,-\Gamma_1,-\Gamma_2)]\!]_Q^2 + [\![(0,K_1^0,K_2^0)]\!]_{R_1}^2 - [\![(0,K_1^0,K_2^0)]\!]_{R_2^TR^{-1}R_2}^2, \\
\overline{\mathbb{P}}(T) &= [\![(I,-\Gamma_{1f},-\Gamma_{2f})]\!]_{Q_f}^2.
\end{aligned}
\tag{49}
$$

This Riccati equation is derived by a formal application of dynamic programming to the optimal control problem specified by (46) and (48). If (49) has a solution on $[0,T]$, the optimal control law is given by

$$
\bar{u}_t^i = -R^{-1}(\mathbb{B}^T\overline{\mathbb{P}} + R_2[0,K_1^0,K_2^0])\boldsymbol{Z}_t, \tag{50}
$$

where $\boldsymbol{Z}_t = [Z_t^{iT}, Z_t^{0T}, \bar{Z}_t]^T$.

Step 2. After determining $\overline{\mathbb{P}}$, denote

$$
[\overline{K}_1,\overline{K}_2,\overline{K}_3] = -R^{-1}(\mathbb{B}^T\overline{\mathbb{P}} + R_2[0,K_1^0,K_2^0]). \tag{51}
$$

Now following the method in [34, sec 4.3] we impose the consistency condition

$$
\begin{cases}
\overline{A} = A + F + B_1K_2^0 + B(\overline{K}_1 + \overline{K}_3), \\
\overline{G} = G + B_1K_1^0 + B\overline{K}_2.
\end{cases}
\tag{52}
$$

This condition is due to the requirement that $Z_t^{(N)}$, as the average of the closed-loop states $(Z_t^1,\cdots,Z_t^N)$ of N agents under the best response control laws $\bar{u}^i$, should regenerate $\bar{Z}_t$ in (46) as $N\to\infty$.

Finally the solution of the mean field game reduces to a solution of (49) subject to (47) and (52).

Proposition 3.7. If the ODE system (29)–(30) has a solution $(\mathbb{P}^0,\mathbb{P})$ on $[0,T]$, then the mean field game solution system consisting of (49), (47) and (52) has a unique solution as $\overline{\mathbb{P}}=\mathbb{P}$. Moreover, $\bar{u}_t^i$ in (50) is the equilibrium strategy and is equivalent to u_t^{i*} in (38).

Proof. We rewrite (49) in the equivalent form

$$
\begin{aligned}
-\dot{\overline{\mathbb{P}}}(t) &= \overline{\mathbb{P}}\{\overline{\mathbb{A}} - \mathbb{B}R^{-1}\mathbb{B}^T\overline{\mathbb{P}} - \mathbb{B}R^{-1}R_2[0,K_1^0,K_2^0]\} \\
&\quad + \{\overline{\mathbb{A}} - \mathbb{B}R^{-1}\mathbb{B}^T\overline{\mathbb{P}} - \mathbb{B}R^{-1}R_2[0,K_1^0,K_2^0]\}^T\overline{\mathbb{P}} \\
&\quad + \overline{\mathbb{P}}\mathbb{B}R^{-1}\mathbb{B}^T\overline{\mathbb{P}} + [\![(I,-\Gamma_1,-\Gamma_2)]\!]_Q^2 + [\![(0,K_1^0,K_2^0)]\!]_{R_1}^2 \\
&\quad - [\![(0,K_1^0,K_2^0)]\!]_{R_2^TR^{-1}R_2}^2.
\end{aligned}
\tag{53}
$$

Next we write equation (30) in the equivalent form

$$
\begin{aligned}
-\dot{\mathbb{P}} &= \mathbb{P}(\mathbb{A} - \mathbb{B}R^{-1}\mathbb{B}^T\mathbb{P} - \mathbb{B}_1R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1) \\
&\quad + (\mathbb{A} - \mathbb{B}R^{-1}\mathbb{B}^T\mathbb{P} - \mathbb{B}_1R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1)^T\mathbb{P} + \mathbb{P}\mathbb{B}R^{-1}\mathbb{B}^T\mathbb{P} \\
&\quad + \mathbb{J}_1^T\mathbb{P}^0\mathbb{B}_0R_0^{-1}R_{12}R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1 + [\![(I,-\Gamma_1,-\Gamma_2)]\!]_Q^2, \\
\mathbb{P}(T) &= [\![(I,-\Gamma_{1f},-\Gamma_{2f})]\!]_{Q_f}^2.
\end{aligned}
\tag{54}
$$

With $\mathbb{P}^0$ and $\mathbb{P}$ as given functions, we may use (35) to show

$$
\begin{aligned}
&[\![(0,K_1^0,K_2^0)]\!]_{R_1}^2 - [\![(0,K_1^0,K_2^0)]\!]_{R_2^TR^{-1}R_2}^2 \\
&\qquad = \mathbb{J}_1^T\mathbb{P}^0\mathbb{B}_0R_0^{-1}R_{12}R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1.
\end{aligned}
\tag{55}
$$

Now we take $\mathbb{P}$ as a candidate solution of $\overline{\mathbb{P}}$. After setting $\overline{\mathbb{P}} = \mathbb{P}$ for (51)–(52) to determine $\overline{\mathbb{A}}$, we can directly show

$$\overline{\mathbb{A}} - \mathbb{B}R^{-1}R_2[0, K_1^0, K_2^0] = \mathbb{A} - \mathbb{B}_1 R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1. \tag{56}$$

As a result, if we take $\overline{\mathbb{P}} = \mathbb{P}$, the equality in (53) holds. Therefore, (49) has a solution given as $\mathbb{P}$. Its solution is clearly unique by the local Lipschitz property of the vector field in (49), where $(\overline{A}, \overline{G})$ contained in $\overline{\mathbb{A}}$ is expressed in term of $\overline{\mathbb{P}}$ via (51)–(52).

Once we have determined $\overline{\mathbb{P}}$, for the best response optimal control problem of Step 1, we can show optimality of $\overline{u}_t^i$ in (50) using the method of completion-of-squares of the cost functionals; see [52, Corollary 3.2]. Since $\overline{\mathbb{P}} = \mathbb{P}$, we see that $\overline{u}_t^i$ and u_t^{i*} are the same feedback law. In other words, the strategy u_t^{i*} is the solution of the mean field game among the followers. □

Remark 3.8. When u_t^{0*} has been fixed and the game is played only among the followers, the strategy u_t^{i*} becomes a mean field Nash equilibrium in the sense (i) that u_t^{i*} is the optimal control with dynamics (46) and cost (48), and (ii) that $\bar{Z}_t$ has been generated by a continuum of followers applying u^{i*} with their own states.

Proposition 3.7 shows that $\varphi = u^{i*}$ minimizes $\bar{J}_i^{\text{mfg}}$ among all feedback control laws defined on $[0, T]$ as long as the closed-loop system has a well-defined solution. The leader's optimizing behavior, however, is different. Let the initial condition at time 0 be (Z_0^0, μ_0). Based on (12), denote the cost of the leader by $\bar{J}_0(0, \psi^0, \psi)$, which is integrated on $[0, T]$. In general, the leader's strategy $\varphi^0 = u^{0*}$ does not have a global minimizer property, resulting in

$$\bar{J}_0(0, \varphi^0, \varphi) > \inf_{u^0} \bar{J}_0(0, u^0, \psi^{u^0}), \tag{57}$$

for feedback strategies $u^0(t, z_0, \mu)$, where $\psi^{u^0}(t, z_i, z_0, \mu)$ is the strategy of the mean field game of the followers when $u^0(t, z_0, \mu)$ is announced for $[0, T]$. Thus it is possible for the leader to strictly improve for itself by taking a strategy u^0 different from φ^0. For standard dynamic Stackelberg games, it is well-known that the feedback solution is different from the global Stackelberg solution, where the leader's equilibrium strategy is a global minimizer while taking into account the follower's optimal response, and which cannot be solved by dynamic programming [3, p. 413]. Similarly, the determination of the global optimizer $u^{0\text{opt}}$ for the right hand side of (57) needs to anticipate the solution of the mean field game on $[0, T]$, and so cannot be achieved by dynamic programming as used in finding (φ^0, φ).

4 Decentralized strategies for the $N + 1$ players

We introduce the following assumptions.

Assumption 1. The ODE system (29)–(30) has a solution $(\mathbb{P}^0, \mathbb{P})$ on $[0, T]$.

Assumption 2. The initial states $\{X_0^j, j \le N\}$ are independent and satisfy

$$\sup_N \max_{j \le N} \mathbb{E}|X_0^j|^2 \le C_0^X, \qquad \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \mathbb{E}X_0^i = m_0^X.$$

Based on strategies (37)–(38) for the mean field limit model, we construct decentralized strategies $(\hat{u}^0, \hat{u}^1, \cdots, \hat{u}^N)$ for the $(N+1)$-player model (1)–(6) as follows. The idea is to replace $\bar{Z}_t$ in (39) by a new process $\overline{X}_t$:

$$\begin{aligned}\dot{\overline{X}}_t = {} & [A + F + B(K_1 + K_3) + B_1 K_2^0]\overline{X}_t \\ & + (G + BK_2 + B_1 K_1^0)X_t^0,\end{aligned} \tag{58}$$

where X_t^0 is now generated by the $(N+1)$-player model and we take the initial condition $\overline{X}_0 = m_0^X$. For (1)–(6), denote the strategies

$$\hat{u}_t^0 = K_1^0 X_t^0 + K_2^0 \overline{X}_t, \tag{59}$$

$$\hat{u}_t^i = K_1 X_t^i + K_2 X_t^0 + K_3 \overline{X}_t, \qquad 1 \le i \le N. \tag{60}$$

These strategies are called decentralized since each follower uses the state information $(X_t^i, X_t^0, \overline{X}_t)$ and the leader uses $(X_t^0, \overline{X}_t)$, instead of the overall state vector of a high dimensional system.

Denote

$$\begin{aligned}
&\widehat{\mathbb{A}}_{11} = A + BK_1, && \widehat{\mathbb{A}}_{12} = G + BK_2 + B_1 K_1^0, && \\
&\widehat{\mathbb{A}}_{13} = F, && \widehat{\mathbb{A}}_{14} = BK_3 + B_1 K_2^0, && \\
&\widehat{\mathbb{A}}_{22} = A_0 + B_0 K_1^0, && \widehat{\mathbb{A}}_{23} = F_0, && \widehat{\mathbb{A}}_{24} = B_0 K_2^0, \\
&\widehat{\mathbb{A}}_{32} = \widehat{A}_{12}, && \widehat{\mathbb{A}}_{33} = A + F + BK_1, && \widehat{\mathbb{A}}_{34} = \widehat{\mathbb{A}}_{14}, \\
&\widehat{\mathbb{A}}_{42} = \widehat{\mathbb{A}}_{12}, && \widehat{\mathbb{A}}_{43} = 0, && \widehat{\mathbb{A}}_{44} = A + F + B(K_1 + K_3) + B_1 K_2^0,
\end{aligned}$$

and

$$\widehat{\mathbb{A}}_0 = (\widehat{\mathbb{A}}_{kl})_{2 \le k,l \le 4}, \qquad \widehat{\mathbb{A}} = \begin{bmatrix} \widehat{\mathbb{A}}_{11} & [\widehat{\mathbb{A}}_{12}, \ \widehat{\mathbb{A}}_{13}, \ \widehat{\mathbb{A}}_{14}] \\ 0 & \widehat{\mathbb{A}}_0 \end{bmatrix}. \tag{61}$$

Under the set of strategies (59)–(60) for (1)–(6) on $[0,T]$, we use $\hat{X}_t^0$ and $\hat{X}_t^i$, $1 \le i \le N$, to denote the closed-loop state processes. Further denote $\hat{X}_t^{(N)} = \frac{1}{N}\sum_{k=1}^N \hat{X}_t^k$ and $W_t^{(N)} = \frac{1}{N}\sum_{k=1}^N W_t^k$. Fixing $i \in \{1, \cdots, N\}$, we obtain

$$d \begin{bmatrix} \hat{X}_t^i \\ \hat{X}_t^0 \\ \hat{X}_t^{(N)} \\ \hat{\overline{X}}_t \end{bmatrix} = \widehat{\mathbb{A}} \begin{bmatrix} \hat{X}_t^i \\ \hat{X}_t^0 \\ \hat{X}_t^{(N)} \\ \hat{\overline{X}}_t \end{bmatrix} dt + \begin{bmatrix} \sigma dW_t^i \\ \sigma_0 dW_t^0 \\ \sigma dW_t^{(N)} \\ 0 \end{bmatrix}, \tag{62}$$

where $\hat{\overline{X}}_t$ is given by (58) after replacing X_t^0 by $\hat{X}_t^0$. The initial states are $\hat{X}_0^j = X_0^j$ for $0 \le j \le N$, and $\overline{X}_0 = m_0^X$. The linear SDE system (62) admits a unique strong solution on $[0,T]$.

Lemma 4.1. Under Assumptions 1 and 2, for system (62), we have

$$\sup_{t \in [0,T]} \mathbb{E}|\hat{X}_t^{(N)} - \hat{\overline{X}}_t|^2 \le C \left(\mathbb{E}|X_0^{(N)} - m_0^X|^2 + 1/N \right), \tag{63}$$

where the constant C is independent of N.

Proof. From (62) we obtain

$$d(\hat{X}_t^{(N)} - \hat{\overline{X}}_t) = (A + F + BK_1)(\hat{X}_t^{(N)} - \hat{\overline{X}}_t)dt + \sigma dW_t^{(N)}.$$

We apply Itô's formula to get

$$\begin{aligned}
\mathbb{E}|\hat{X}_t^{(N)} - \hat{\overline{X}}_t|^2 =& \mathbb{E}|X_0^{(N)} - m_0^X|^2 + t|\sigma|^2/N \\
&+ 2\int_0^t \mathbb{E}[(\hat{X}_s^{(N)} - \hat{\overline{X}}_s)^T (A + F + BK_1)(\hat{X}_s^{(N)} - \hat{\overline{X}}_s)]ds.
\end{aligned}$$

Then (63) follows from elementary SDE estimates with Grönwall's lemma. □

4.1 Costs under decentralized strategies

Subsequently, we evaluate the performance of $\mathcal{A}_0$ and $\mathcal{A}_i$ among the $N+1$ players. Here the initial condition has been given as X_0^j, $0 \le j \le N$. Our method is to embed the cost evaluation problem into a family of problems with different initial conditions as in dynamic programming. The key observation is that if the term $\hat{X}^{(N)}$ in (62) is not defined as $\frac{1}{N}\sum_{k=1}^N \hat{X}_t^k$ but instead specified as an independent component of the state vector of the SDE (62), the resulting SDE is still well-defined on $[t_0, T]$ as long as an initial condition is selected at $t_0 \in [0, T]$.

For convenience of specifying the cost functionals below, we use a set of new variables $(Z_s^i, Z_s^0, \hat{Z}_s, \overline{Z}_s)$ to rewrite (62) as follows:

$$d\begin{bmatrix} Z_s^i \\ Z_s^0 \\ \hat{Z}_s \\ \overline{Z}_s \end{bmatrix} = \widehat{\mathbb{A}} \begin{bmatrix} Z_s^i \\ Z_s^0 \\ \hat{Z}_s \\ \overline{Z}_s \end{bmatrix} ds + \begin{bmatrix} \sigma dW_s^i \\ \sigma_0 dW_s^0 \\ \sigma dW_s^{(N)} \\ 0 \end{bmatrix}, \tag{64}$$

with initial condition $Z_t^i = z_i$, $Z_t^0 = z_0$, $\hat{Z}_t = \hat{z}$, and $\overline{Z}_t = \bar{z}$ at time $t \in [0, T]$. The above variables $(Z_s^i, Z_s^0, \overline{Z}_s)$ are used as generic notation and are different from those in (40)–(42). This reuse of notation shall cause no confusion. Here we take arbitrary $z_i, z_0, \hat{z}, \bar{z} \in \mathbb{R}^n$. Denote $\breve{u}_s^0 = K_1^0 Z_s^0 + K_2^0 \overline{Z}_s$ and $\breve{u}_s^i = K_1 Z_s^i + K_2 Z_s^0 + K_3 \overline{Z}_s$. Define

$$\hat{V}_0(t, z_0, \hat{z}, \bar{z}) = \mathbb{E}\int_t^T (|Z_s^0 - \Gamma_0 \hat{Z}_s|_{Q_0}^2 + |\breve{u}_s^0|_{R_0}^2) ds + \mathbb{E}|Z_T^0 - \Gamma_{0f}\hat{Z}_T|_{Q_{0f}}^2, \tag{65}$$

$$\begin{aligned} \hat{V}(t, z_i, z_0, \hat{z}, \bar{z}) = \mathbb{E}\int_t^T \Big(&|Z_s^i - \Gamma_1 Z_s^0 - \Gamma_2 \hat{Z}_s|_Q^2 + |\breve{u}_s^i|_R^2 + |\breve{u}_s^0|_{R_1}^2 \\ &+ 2\breve{u}_s^{iT} R_2 \breve{u}_s^0\Big) ds + \mathbb{E}|Z_T^i - \Gamma_{1f} Z_T^0 - \Gamma_{2f}\hat{Z}_T|_{Q_f}^2. \end{aligned} \tag{66}$$

Under Assumption 1, $\hat{V}_0$ and $\hat{V}$ are well-defined for $t \in [0, T]$.

Denote $\hat{\mathbf{z}}_0 = (z_0^T, \hat{z}^T, \bar{z}^T)^T$ and $\hat{\mathbf{z}} = (z_i^T, z_0^T, \hat{z}^T, \bar{z}^T)^T$. Below we will determine $\hat{V}_0$ and $\hat{V}$ in the form

$$\hat{V}_0(t, z_0, \hat{z}, \bar{z}) = \hat{\mathbf{z}}_0^T \hat{\mathbb{P}}^0(t)\hat{\mathbf{z}}_0 + \hat{r}^0(t), \tag{67}$$

$$\hat{V}(t, z_i, z_0, \hat{z}, \bar{z}) = \hat{\mathbf{z}}^T \hat{\mathbb{P}}(t)\hat{\mathbf{z}}^T + \hat{r}(t), \tag{68}$$

where $\hat{\mathbb{P}}^0(t)$ and $\hat{\mathbb{P}}(t)$ are symmetric matrix functions with the partition

$$\hat{\mathbb{P}}^0 = (\hat{P}_{kl}^0)_{1\le k,l\le 3}, \qquad \hat{\mathbb{P}} = (\hat{P}_{kl})_{1\le k,l\le 4}.$$

Each submatrix above is a function of $t \in [0, T]$ with value in $\mathbb{R}^{n\times n}$. To determine $\hat{V}_0$ and $\hat{V}$, we introduce the following linear ODE system:

$$-\frac{d}{dt}\hat{\mathbb{P}}^0 = \hat{\mathbb{P}}^0\widehat{\mathbb{A}}_0 + \widehat{\mathbb{A}}_0^T\hat{\mathbb{P}}^0 + [\![(K_1^0, 0, K_2^0)]\!]_{R_0}^2 + [\![(I, -\Gamma_0, 0)]\!]_{Q_0}^2, \tag{69}$$

$$\begin{aligned} -\frac{d}{dt}\hat{\mathbb{P}} = \hat{\mathbb{P}}\widehat{\mathbb{A}} + \widehat{\mathbb{A}}^T\hat{\mathbb{P}} &+ [\![(I, -\Gamma_1, -\Gamma_2, 0)]\!]_Q^2 \\ &+ [\![(K_1, K_2, 0, K_3)]\!]_R^2 + [\![(0, K_1^0, 0, K_2^0)]\!]_{R_1}^2 \\ &+ (K_1, K_2, 0, K_3)^T R_2 (0, K_1^0, 0, K_2^0) \\ &+ (0, K_1^0, 0, K_2^0)^T R_2^T (K_1, K_2, 0, K_3), \end{aligned} \tag{70}$$

$$-\frac{d}{dt}\hat{r}^0 = \mathrm{Tr}[\hat{P}_{11}^0 \Sigma_{w^0} + (1/N)\hat{P}_{22}^0 \Sigma_w], \tag{71}$$

$$-\frac{d}{dt}\hat{r} = \mathrm{Tr}\{[\hat{P}_{11} + (1/N)(\hat{P}_{33} + 2\hat{P}_{13})]\Sigma_w\} + \mathrm{Tr}(\hat{P}_{22}\Sigma_{w^0}), \tag{72}$$

where terminal conditions are

$$\hat{\mathbb{P}}^0(T) = [\![(I, -\Gamma_0, 0)]\!]^2_{Q_0}, \quad \hat{\mathbb{P}}(T) = [\![(I, -\Gamma_1, -\Gamma_2, 0)]\!]^2_{Q},$$
$$\hat{r}^0(T) = \hat{r}(T) = 0.$$

The matrix functions $(K_1^0, K_2^0, K_1, K_2, K_3)$ in (69)–(70) have been defined by (35)–(36).

Lemma 4.2. Under Assumption 1, the ODE system (69)–(72) has a unique solution $(\hat{\mathbb{P}}^0, \hat{\mathbb{P}}, \hat{r}^0, \hat{r})$ on $[0, T]$, and the functions $\hat{V}_0$ and $\hat{V}$ in (65)–(66) have the representation (67)–(68).

Proof. Once $(\mathbb{P}^0, \mathbb{P})$ is given, we find a unique solution $(\hat{\mathbb{P}}^0, \hat{\mathbb{P}})$ on $[0, T]$ from two linear ODEs with bounded coefficients, and further uniquely obtain $(\hat{r}^0, \hat{r})$ on $[0, T]$. The last part of the lemma follows from Lemma B.1. □

Lemma 4.3. Under Assumption 1, the solution $(\hat{\mathbb{P}}^0, \hat{\mathbb{P}})$ of (69)–(70) satisfies

$$\mathbb{P}^0(t) = \mathbb{J}_2^T \hat{\mathbb{P}}^0(t) \mathbb{J}_2, \quad \mathbb{P}(t) = \mathbb{J}_3^T \hat{\mathbb{P}}(t) \mathbb{J}_3, \quad \text{for all } t \in [0, T]. \tag{73}$$

where

$$\mathbb{J}_2 = \begin{bmatrix} I_{2n} \\ [0_{n\times n}, I_n] \end{bmatrix}, \quad \mathbb{J}_3 = \begin{bmatrix} I_{3n} \\ [0_{n\times n}, 0_{n\times n}, I_n] \end{bmatrix}.$$

Proof. For the purpose of relating $(\hat{\mathbb{P}}^0, \hat{\mathbb{P}})$ to $(\mathbb{P}^0, \mathbb{P})$, it is sufficient to consider the special case of deterministic dynamics. We evaluate the costs by taking $\sigma = 0$, $\sigma_0 = 0$, and initial conditions $(t, z_i, z_0, \bar{z})$ in system (40)–(42) and $(t, z_i, z_0, \bar{z}, \bar{z})$ in system (64). In this case, we evaluate $\bar{J}_0^*, \bar{J}_i^*$ in (44)–(43) and $\hat{V}_0$, $\hat{V}$ in (65)–(66) without using expectation. Then we easily see that the pair (Z_s^i, Z_s^0) is the same in both systems. Moreover, we have $\bar{Z}_s = \hat{Z}_s = \overline{Z}_s$ for all $s \in [t, T]$. Subsequently, we have

$$\bar{J}_0^*(t, z_0, \bar{z}) = \hat{V}_0(t, z_0, \bar{z}, \bar{z}) \quad \text{for all } t \in [0, T],\ z_0, \bar{z} \in \mathbb{R}^n, \tag{74}$$

since they use the same cost integrand and terminal cost.

On the other hand, by Proposition 3.5 and Lemma 4.2, when $\sigma = 0$ and $\sigma_0 = 0$, we have

$$\bar{J}_0^*(t, z_0, \bar{z}) = (z_0^T, \bar{z}^T) \mathbb{P}^0(t) \begin{bmatrix} z_0 \\ \bar{z} \end{bmatrix}, \tag{75}$$

$$\hat{V}_0(t, z_0, \bar{z}, \bar{z}) = (z_0^T, \bar{z}^T, \bar{z}^T) \hat{\mathbb{P}}^0(t) \begin{bmatrix} z_0 \\ \bar{z} \\ \bar{z} \end{bmatrix} = (z_0^T, \bar{z}^T) \mathbb{J}_2^T \hat{\mathbb{P}}^0(t) \mathbb{J}_2 \begin{bmatrix} z_0 \\ \bar{z} \end{bmatrix}. \tag{76}$$

Since z_0 and $\bar{z}$ are arbitrary, the first part of the lemma follows from (74)–(76). The second part is proved in a similar manner. □

The following theorem compares the costs of the finite population under the decentralized strategies (59)–(60) with the costs in (45) for the mean field limit model. The mean field limit model only involves the leader $\mathcal{A}_0$ and the representative follower $\mathcal{A}_i$ with initial condition (X_0^0, X_0^i, m_0^X) at $t = 0$. Recall that we take $\overline{X}_0 = m_0^X$. We evaluate J_k^{N+1}, $0 \le k \le N$, with the initial condition $(X_0^0, X_0^1, \cdots, X_0^N)$ and $\overline{X}_0$. For $\bar{J}_0^*$ and $\bar{J}_i^*$ we set the initial condition (X_0^0, X_0^i, m_0^X).

Theorem 4.4. Under Assumptions 1 and 2, if the decentralized strategies $\hat{u}^0$ and $\hat{\mathbf{u}}^F = (\hat{u}^1, \cdots, \hat{u}^N)$ in (59)–(60) are applied to the $(N+1)$-player model (1)–(6), then we have

$$\begin{aligned} &|J_0^{N+1}(0, \hat{u}^0, \hat{\mathbf{u}}^F) - \bar{J}_0^*(0, X_0^0, m_0^X)| \\ &\qquad = O((\mathbb{E}|X_0^{(N)} - m_0^X|^2)^{1/2} + 1/N), \end{aligned} \tag{77}$$

$$\begin{aligned} &|J_i^{N+1}(0, \hat{u}^i, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}) - \bar{J}_i^*(0, X_0^i, X_0^0, m_0^X)| \\ &\qquad = O((\mathbb{E}|X_0^{(N)} - m_0^X|^2)^{1/2} + 1/N). \end{aligned} \tag{78}$$

Proof. It follows from (67), (27) and (73) that

$$\begin{aligned}
&J_0^{N+1}(0,\hat u^0,\hat{\mathbf{u}}^F)-\bar J_0^*(0,X_0^0,m_0^X)\\
&\quad=\mathbb{E}\Big\{[X_0^{0T},X_0^{(N)T},\overline{X}_0^T]\hat{\mathbb{P}}^0(0)[X_0^{0T},X_0^{(N)T},\overline{X}_0^T]^T+\hat r^0(0)\Big\}\\
&\qquad-\mathbb{E}\Big\{[X_0^{0T},\overline{X}_0^T]\mathbb{P}^0(0)[X_0^{0T},\overline{X}_0^T]^T+r^0(0)\Big\}\\
&\quad=\mathbb{E}\Big\{[X_0^{0T},X_0^{(N)T},\overline{X}_0^T]\hat{\mathbb{P}}^0(0)[X_0^{0T},X_0^{(N)T},\overline{X}_0^T]^T+\hat r^0(0)\Big\}\\
&\qquad-\mathbb{E}\Big\{[X_0^{0T},\overline{X}_0^T,\overline{X}_0^T]\hat{\mathbb{P}}^0(0)[X_0^{0T},\overline{X}_0^T,\overline{X}_0^T]^T-r^0(0)\\
&\quad=\mathbb{E}\Big\{[0,X_0^{(N)T}-\overline{X}_0^T,0]\hat{\mathbb{P}}^0(0)[2X_0^{0T},\overline{X}_0^T+X_0^{(N)T},2\overline{X}_0^T]^T\Big\}\\
&\qquad+\hat r^0(0)-r^0(0).
\end{aligned}$$

Note that (73) implies $\hat P_{11}^0=P_{11}^0$ on $[0,T]$. So comparing (31) and (71), by Grönwall's lemma we have that $\sup_{t\in[0,T]}|\hat r^0(t)-r^0(t)|=O(1/N)$. The estimate (77) then follows. The estimate (78) can be shown in a similar manner. □

4.2 Improving performance via alternative strategies

For the $N+1$ player model, suppose u^0 is fixed as $\hat u^0$ for the leader. Player $\mathcal{A}^i$ attempts a different strategy $u^i\in\mathcal{U}^c$ consisting of feedback control laws of the form $\psi(t,X_t^0,\cdots,X_t^N,\overline{X}_t)$ which is continuous in its arguments $(t,x_0,\cdots,x_N,\bar x)\in[0,T]\times\mathbb{R}^{(N+2)n}$ and is Lipschitz continuous in $(x_0,\cdots,x_N,\bar x)$. We have the following ϵ-Nash equilibrium property for $(\hat u^1,\cdots,\hat u^N)$.

Proposition 4.5. Under Assumptions 1 and 2, we have

$$J_i^{N+1}(0,\hat u^i,\hat u^0,\hat{\mathbf{u}}^{F,-i})\le\inf_{u^i\in\mathcal{U}^c}J_i^{N+1}(0,u^i,\hat u^0,\hat{\mathbf{u}}^{F,-i})+\varepsilon_N,\qquad 1\le i\le N,\tag{79}$$

where $\varepsilon_N\le C[(\mathbb{E}|X_0^{(N)}-m_0^X|^2)^{1/2}+1/\sqrt{N}]$.

Proof. We have the key observation from Proposition 3.7 that u_t^{i*} in (38) is the solution of a mean field game (with a continuum of players) and that $\hat u_t^i$ is based on u_t^{i*} by using the states $(X_t^i,X_t^0,\overline{X}_t)$. Then we follow the standard method (see e.g. [34, sec 6]) to estimate the performance in the model of N followers. The detail is routine and omitted here. □

5 Asymptotic equilibrium in finite populations

We proceed to analyze the performance of the decentralized strategies (59)–(60) applied to the $(N+1)$-player game. In this case one expects that an asymptotic Stackelberg equilibrium property holds as $N\to\infty$. But the lack of global optimality of $u^{0*}=\varphi^0$ for minimizing $\bar J_0(0,u^0,\psi^{u^0})$ (see (57)) implies that we should not attempt to identify an equilibrium by testing strategies that allow modification on the whole interval $[0,T]$. To seek a feasible equilibrium concept, we adopt the idea of t-selves suggested in [23] so that each player $\mathcal{A}_i$ is distinguished as a different agent at differen time instants. At time t, the player is called the t-$\mathcal{A}_i$ agent. Therefore the game is played by a stream of $N+1$ t-indexed agents as incarnations of the original $N+1$ players.

To characterize the equilibrium of the t-indexed agents, one may try to relate the selection of (u^{0*},u^{i*}) to an optimization problem of $N+1$ agents alive only at time t. Such an optimization problem is, however, not meaningful if we only consider a single point of time. To overcome this difficulty, following [23], we consider a very small time interval as the following. Given the system initial condition at t, for each i, a coalition of followers denoted by $\mathcal{A}_i^{[t,t+\epsilon]}$, consisting of all s-$\mathcal{A}_i$ agents, $s\in[t,t+\epsilon]$, optimizes its cost defined on $[t,T]$ while it only acts on $[t,t+\epsilon]$. Then we check how the controls of the $N+1$ coalitions on that small interval behave as ϵ approaches 0.

5.1 Decentralized feedback strategies and spike variation

Consider the following system

$$dX_s^0 = (A_0X_s^0 + B_0u_s^0 + F_0X_s^{(N)})ds + \sigma_0 dW_s^0, \tag{80}$$

$$dX_s^i = (AX_s^i + Bu_s^i + B_1u_s^0 + FX_s^{(N)} + GX_s^0)ds + \sigma dW_s^i, \tag{81}$$

$$d\overline{X}_s = \bar{f}(s, X_s^0, \overline{X}_s)ds, \quad \overline{X}_0 = m_0^X, \quad 1 \le i \le N, \quad 0 \le s \le T, \tag{82}$$

where the initial states $\{X_0^j, 0 \le j \le N\}$ satisfy Assumption 2. Denote $\boldsymbol{X}_s^c = (X_s^0, X_s^1, \cdots, X_s^N, \overline{X}_s)$, where the superscript in $\boldsymbol{X}_s^c$ indicates centralized information. For specifying strategies of interest, we regard $\bar{f} : [0,T] \times \mathbb{R}^{2n} \to \mathbb{R}^n$ as part of the strategy design, and for convenience of presentation will call it a control.

Our plan here is to specify a class of feedback strategies that contains the set of decentralized strategies (59)–(60) and serves as a space to search for an asymptotic equilibrium to be introduced subsequently. We take u^0 and u^i, $1 \le i \le N$, as decentralized strategies of the form: $u^0 = \psi^0(s, X_s^0, \overline{X}_s)$ and $u^i = \psi(s, X_s^i, X_s^0, \overline{X}_s)$, where ψ is shared by all followers as N increases. We call $(u^0, \mathbf{u}^F, \bar{f})$ a set of admissible decentralized (feedback) strategies if the following conditions hold: (i) the functions $\bar{f}$, ψ^0, and ψ are continuous, and Lipschitz continuous in their spatial variables (i.e., those for $X_s^j, \overline{X}_s$); (ii) $\sup_{0\le s\le T} \mathbb{E}|X_s^{(N)} - \overline{X}_s|^2 = o(1)$ as $N \to \infty$. Condition (ii) is due to the consistency requirement that $\overline{X}_s$ be an approximation of $X_s^{(N)}$.

Let $\mathcal{U}_{0,T}^{fd}$ consist of all $(u^0, \mathbf{u}^F, \bar{f})$, each as a set of admissible decentralized strategies defined on $[0,T]$. When $\bar{f}$ is already given, we also say that $(u^0, \mathbf{u}^F)$ is in $\mathcal{U}_{0,T}^{fd}$ and is a set of admissible decentralized strategies.

Suppose that $(u^0, \mathbf{u}^F, \bar{f}) \in \mathcal{U}_{0,T}^{fd}$ is chosen for (80)–(82), which generates a unique solution $\boldsymbol{X}_s^c$, $s \in [0,T]$. At time t, taking the initial condition $\boldsymbol{X}_t^c$, we define the costs:

$$J_0^{N+1}(t, \boldsymbol{X}_t^c, u^0, \mathbf{u}^F, \bar{f}) = \mathbb{E}\int_t^T L_0(X_s^0, \mu_s^{(N)}, u_s^0)ds + \mathbb{E}g_0(X_T^0, \mu_T^{(N)}), \tag{83}$$

$$\begin{aligned} J_i^{N+1}(t, \boldsymbol{X}_t^c, u^i, u^0, \mathbf{u}^{F,-i}, \bar{f}) =& \mathbb{E}\int_t^T L(X_s^i, X_s^0, \mu_s^{(N)}, u_s^i, u_s^0)ds \\ &+ \mathbb{E}g(X_T^i, X_T^0, \mu_T^{(N)}), \qquad 1 \le i \le N. \end{aligned} \tag{84}$$

Now, under $(u^0, \mathbf{u}^F, \bar{f}) \in \mathcal{U}_{0,T}^{fd}$ and with the initial condition $(t, \boldsymbol{X}_t^c)$, we introduce two types of perturbed strategies. Denote $\bar{f}_s = \bar{f}(s, X_s^0, \overline{X}_s)$.

(a) On $[t, t+\epsilon]$, for a single player $\mathcal{A}_i$, we set $u_s^i = v^i$.

(b) On $[t, t+\epsilon]$, the set of controls $(u_s^0, \cdots, u_s^N, \bar{f}_s)$ is replaced by $(v^0, \cdots, v^N, \check{f}_s)$, where $\check{f}_s$ is used as a new drift term for $\overline{X}_s$ in (82).

Case (a). For the single player $\mathcal{A}_i$, we take

$$u_s^{i,\epsilon} := \begin{cases} v^i, & s \in [t, t+\epsilon], \\ u_s^i, & s \in [0,T] \setminus [t, t+\epsilon], \end{cases}$$

with v^i being represented as a Borel measurable function of $\boldsymbol{X}_t^c$ and $\mathbb{E}|v^i|^2 < \infty$. Such a random vector v^i is called admissible. The other components $\mathbf{u}^{F,-i}$ and $\bar{f}$ remain the same. We accordingly have the cost $J_i^{N+1}(t, \boldsymbol{X}_t^c, u^{i,\epsilon}, u^0, \mathbf{u}^{F,-i}, \bar{f})$. This method of modifying a feedback control implemented on $[t,T]$ by using a constant control on a small interval was initially introduced in [23]. We call $u^{i,\epsilon}$ a modification of u^i with the spike variation v^i on $[t, t+\epsilon]$. Thus for a given $\omega \in \Omega$, $u_s^{i,\epsilon}$ remains the same on $[t, t+\epsilon]$.

Case (b). We introduce a spike variation on $[t, t+\epsilon]$ for each strategy u^j defined on $[0, T]$, which leads to the new strategy

$$u_s^{j,\epsilon} := \begin{cases} v^j, & s \in [t, t+\epsilon], \\ u_s^j, & s \in [0,T] \setminus [t, t+\epsilon], \end{cases} \qquad 0 \le j \le N, \tag{85}$$

with v^j being a Borel measurable function of $\boldsymbol{X}_t^c$ and $\mathbb{E}|v^j|^2 < \infty$. Note that the choice of $\bar{f}$ in (80)–(82) has been part of the design of strategies. Now given $(v^0, \cdots, v^N)$, we further define

$$\bar{f}_s^\epsilon = \begin{cases} \check{f}_s. & s \in [t, t+\epsilon], \\ \bar{f}_s & s \in [0,T] \setminus [t, t+\epsilon], \end{cases} \tag{86}$$

where $\check{f}_s := \check{f}(s, X_s^0, \overline{X}_s, v^0, v^{(N)})$. The function $\check{f} : [0,T] \times \mathbb{R}^{2n+2n_1} \to \mathbb{R}^n$ is continuous, and Lipschitz continuous in $(x_0, \bar{x})$ (for $(X_s^0, \overline{X}_s)$), with linear growth in all four spatial variables. Since as in (82) our approach is still to use $\overline{X}_s$ to provide information on $X_s^{(N)}$, we consider the above form for $\check{f}$, instead of making $\check{f}$ as general as possible. Denote $\mathbf{u}^{F,\epsilon} = (u^{1,\epsilon}, \cdots, u^{N,\epsilon})$ and $\mathbf{u}^{F,-i,\epsilon}$ similarly.

For both case (a) and case (b), in our further analysis, the spike variation associated with a specific perturbed strategy $u_s^{j,\epsilon}$ will be clear from the context.

For the resulting solution $\overline{X}_s$ on $[t, t+\epsilon]$ in case (b) to be relevant, we need to further specify $\check{f}$ in an appropriate form. The idea is to use $\check{f}$ to adjust $\overline{X}_{t+\epsilon}$ to closely approximate $X_{t+\epsilon}^{(N)}$. Note that $\overline{X}_0$ has been selected under a similar requirement (see (58)). Adjusting $\overline{X}_{t+\epsilon}$ to the *right* position is crucial since otherwise it would be irrelevant to the game resumed at time $t+\epsilon$. We introduce the following definition.

Definition 5.1. We call $\check{f}$ in (86) compatible if it ensures that for some fixed constant C independent of $(t, N, v^0, \cdots, v^N)$,

$$\mathbb{E}|X_{t+\epsilon}^{(N)} - \overline{X}_{t+\epsilon}|^2 \le C\mathbb{E}|X_t^{(N)} - \overline{X}_t|^2 + C\epsilon/N,$$

where $(X_s^0, \cdots, X_s^N, \overline{X}_s)$, $0 \le s \le T$, is the solution of (80)–(82) with strategies $(u^{0,\epsilon}, \mathbf{u}^{F,\epsilon}, \bar{f}^\epsilon)$ specified by (85)–(86). Accordingly, we call $(u^{0,\epsilon}, \mathbf{u}^{F,\epsilon}, \bar{f}^\epsilon)$ a set of admissible perturbed strategies.

Remark 5.2. The compatibility condition in Definition 5.1 translates into a certain constraint on the selection of $\check{f}$. Since $(u^0, \mathbf{u}^F, \bar{f}) \in \mathcal{U}_{0,T}^{fd}$ ensures $\sup_{0 \le s \le t} \mathbb{E}|X_s^{(N)} - \overline{X}_s|^2 = o(1)$, now with $\check{f}$ being compatible, we have $\mathbb{E}|X_{t+\epsilon}^{(N)} - \overline{X}_{t+\epsilon}|^2 = o(1)$ as $N \to \infty$.

Note that $(u^{0,\epsilon}, \mathbf{u}^\epsilon, \bar{f}^\epsilon)$ in general is not in $\mathcal{U}_{0,T}^{fd}$ but still ensures a well-defined solution in $L^2_{\mathcal{F}^{WX_0}}(0, T; \mathbb{R}^{(N+2)n})$ for the SDE system (80)–(82), where $\mathcal{F}_t^{WX_0}$ is the σ-field generated by $(W_s^j, X_0^j, 0 \le j \le N, s \le t)$. So $J_0^{N+1}(t, \boldsymbol{X}_t^c, u^{0,\epsilon}, \mathbf{u}^{F,\epsilon}, \bar{f}^\epsilon)$ and $J_i^{N+1}(t, \boldsymbol{X}_t^c, u^{i,\epsilon}, u^{0,\epsilon}, \mathbf{u}^{F,-i,\epsilon}, \bar{f}^\epsilon)$ are well-defined.

Regardless of the specific form of $\bar{f}$, we adopt a natural construction of $\check{f}$:

$$\begin{aligned} \dot{\overline{X}}_s &= GX_s^0 + (A+F)\overline{X}_s + B_1 v^0 + Bv^{(N)} \\ &=: \check{f}_s, \qquad t \le s \le t+\epsilon, \end{aligned} \tag{87}$$

where the initial state $\overline{X}_t$ has been determined as the last component of $\boldsymbol{X}_t^c$. This ODE is constructed by approximating the SDE of $X_s^{(N)}$:

$$dX_s^{(N)} = [GX_s^0 + (A+F)X_s^{(N)} + B_1 v^0 + Bv^{(N)}]ds + \sigma dW_s^{(N)}, \quad t \le s \le t+\epsilon,$$

where $v^{(N)} = (1/N)\sum_{i=1}^N v^i$ and $W_s^{(N)} = (1/N)\sum_{i=1}^N W_s^i$.

Proposition 5.3. Under $(u^{0,\epsilon}, \mathbf{u}^{F,\epsilon}, \bar{f}^\epsilon)$ taking $\check{f}$ in (87), we have

$$\sup_{t \le s \le t+\epsilon} \mathbb{E}|X_s^{(N)} - \overline{X}_s|^2 \le C\mathbb{E}|X_t^{(N)} - \overline{X}_t|^2 + C\epsilon/N, \tag{88}$$

and $\check{f}$ is compatible.

Proof. The estimate (88) follows by checking the SDE of $X_s^{(N)} - \overline{X}_s$. Then $\check{f}$ is clearly compatible. $\square$

For determinacy, we will take $\check{f}$ of the form in (87) in all subsequent analysis. We introduce the notion of ε_N-Stackelberg equilibrium for the game of $N+1$ players in (80)–(82).

Definition 5.4. A set of strategies $(u^0, \cdots, u^N, \bar{f}) \in \mathcal{U}_{0,T}^{fd}$ is an ε_N*-Stackelberg equilibrium* if for each given $t \in [0, T)$ it fulfills the following two conditions:

(i) For each $1 \le i \le N$,

$$\limsup_{\epsilon \downarrow 0} \frac{1}{\epsilon}[J_i^{N+1}(t, \boldsymbol{X}_t^c, u^i, u^0, \mathbf{u}^{F,-i}, \bar{f}) - J_i^{N+1}(t, \boldsymbol{X}_t^c, u^{i,\epsilon}, u^0, \mathbf{u}^{F,-i}, \bar{f})] \le \varepsilon_N, \tag{89}$$

for any $u^{i,\epsilon}$ being a modification of u^i by taking an admissible v^i on $[t, t+\epsilon]$.

(ii) For admissible perturbed strategies $(u^{0,\epsilon}, \mathbf{u}^{F,\epsilon}, \bar{f}^\epsilon)$ with $\check{f}$ given by (87),

$$\limsup_{\epsilon \downarrow 0} \frac{1}{\epsilon}[J_0^{N+1}(t, \boldsymbol{X}_t^c, u^0, \mathbf{u}^F, \bar{f}) - J_0^{N+1}(t, \boldsymbol{X}_t^c, u^{0,\epsilon}, \mathbf{u}^{F,\epsilon}, \bar{f}^\epsilon)] \le \varepsilon_N, \tag{90}$$

where $u^{0,\epsilon}$ is any modification of u^0 by taking its admissible v^0 on $[t, t+\epsilon]$, and $\mathbf{u}^{F,\epsilon} = (u^{1,\epsilon}, \cdots, u^{N,\epsilon})$ is a modification of $\mathbf{u}^F = (u^1, \cdots, u^N)$ with $\mathbf{v}^F = (v^1, \cdots, v^N)$ on $[t, t+\epsilon]$ such that $\mathbf{u}^{F,\epsilon}$ is a Nash equilibrium for the N followers when v^0 has been selected and fixed for the leader (so $\mathbf{v}^F$ depends on v^0 and may be denoted $\mathbf{v}^F(v^0)$).

Remark 5.5. For very small ϵ and ε_N and the game played on $[t, t+\epsilon]$, condition (i) means that $(u^0, \cdots, u^N)$ restricted on that interval is almost a Nash equilibrium among the followers. Condition (ii) further indicates that u^0 nearly attains the lowest cost of all scenarios in which the followers react on the time interval $[t, t+\epsilon]$ by playing a Nash game among themselves.

For $N \ge 1$, denote

$$d_N^X = (\mathbb{E}|X_0^{(N)} - m_0^X|^2)^{1/2} + 1/\sqrt{N}. \tag{91}$$

Under Assumption 2, we have $d_N^X = o(1)$. A key result of the equilibrium analysis is the following theorem.

Theorem 5.6. Under Assumptions 1 and 2, the set of decentralized strategies $(\hat{u}^0, \hat{\mathbf{u}}^F) = (\hat{u}^0, \hat{u}^1, \cdots, \hat{u}^N)$ given by (59)–(60) constitutes an ε_N-Stackelberg equilibrium for the $N+1$ player model (1)–(6), where $\varepsilon_N = O(d_N^X)$.

In view of Lemma 4.1, it is easy to show that $(\hat{u}^0, \hat{\mathbf{u}}^F)$ belongs to $\mathcal{U}_{0,T}^{fd}$. The proof of Theorem 5.6 is postponed to the end of this section. For convenience of presentation, we will first check condition (ii) in Definition 5.4 for the cost of the leader.

5.2 The auxiliary static Stackelberg–Nash game

To prove Theorem 5.6, we introduce an auxiliary static game with $N+1$ players. Denote

$$Y_s^0 = [X_s^{0T},\ X_s^{(N)T},\ \overline{X}_s^T]^T, \qquad Y_s^i = [X_s^{iT},\ Y_s^{0T}]^T, \tag{92}$$

and

$$\boldsymbol{A}_0 = \begin{bmatrix} A_0 & F_0 & 0 \\ G & A+F & 0 \\ G & 0 & A+F \end{bmatrix}, \qquad \boldsymbol{A} = \begin{bmatrix} A & (G,\ F,\ 0) \\ 0 & \boldsymbol{A}_0 \end{bmatrix},$$

$$\boldsymbol{B}_0 = \begin{bmatrix} B_0 \\ B_1 \\ B_1 \end{bmatrix}, \qquad \boldsymbol{B}_0^a = \begin{bmatrix} 0 \\ B \\ B \end{bmatrix}, \qquad \boldsymbol{D}_0 = \begin{bmatrix} \sigma_0 \\ 0 \\ 0 \end{bmatrix}, \qquad \boldsymbol{D}_0^a = \begin{bmatrix} 0 \\ \sigma \\ 0 \end{bmatrix},$$

$$\boldsymbol{B}_1 = \begin{bmatrix} B_1 \\ \boldsymbol{B}_0 \end{bmatrix}, \qquad \boldsymbol{B}_N^a = \begin{bmatrix} B \\ \boldsymbol{B}_0^a/N \end{bmatrix}, \qquad \boldsymbol{B}^a = \begin{bmatrix} 0 \\ \boldsymbol{B}_0^a \end{bmatrix},$$
$$\boldsymbol{D} = \begin{bmatrix} 0 \\ \boldsymbol{D}_0 \end{bmatrix}, \qquad \boldsymbol{D}_N = \begin{bmatrix} \sigma \\ \boldsymbol{D}_0^a/N \end{bmatrix}, \qquad \boldsymbol{D}^a = \begin{bmatrix} 0 \\ \boldsymbol{D}_0^a \end{bmatrix},$$
$$\widetilde{\boldsymbol{B}}_0 = \begin{bmatrix} \mathbb{B}_0 \\ \widetilde{B}_1 \end{bmatrix}, \qquad \boldsymbol{B}_0^b = \begin{bmatrix} 0 \\ B \\ 0 \end{bmatrix}, \qquad \boldsymbol{B}_N^b = \begin{bmatrix} B \\ \boldsymbol{B}_0^b/N \end{bmatrix}.$$

Let $\hat{\boldsymbol{X}}_t^c := (\hat{X}_t^0, \hat{X}_t^1, \cdots, \hat{X}_t^N, \hat{\overline{X}}_t)$ be generated by (62) under the set of strategies $(\hat{u}^0, \cdots, \hat{u}^N)$ applied on $[0, t]$. Define $\hat{u}^{j,\epsilon}$ with spike variation v^j on $[t, t+\epsilon]$. When the system (1)–(6) takes the set of controls $(v^0, v^1, \cdots, v^N)$ on $[t, t+\epsilon]$ that depends only on $(Y_t^0, \cdots, Y_t^N)$, the processes $(Y_s^0, \cdots, Y_s^N)$ follow the dynamics on $[t, t+\epsilon]$:

$$dY_s^0 =(\boldsymbol{A}_0 Y_s^0 + \boldsymbol{B}_0 v^0 + \boldsymbol{B}_0^a v^{(N)})ds + \boldsymbol{D}_0 dW_s^0 + \boldsymbol{D}_0^a dW_s^{(N)}, \tag{93}$$
$$\begin{aligned} dY_s^i =&(\boldsymbol{A} Y_s^i + \boldsymbol{B}_1 v^0 + \boldsymbol{B}_N^a v^i + \boldsymbol{B}^a v^{(-i)})ds \\ &+ \boldsymbol{D} dW_s^0 + \boldsymbol{D}_N dW_s^i + \boldsymbol{D}^a dW_s^{(-i)}, \quad 1 \le i \le N, \end{aligned} \tag{94}$$

with the initial conditions

$$Y_t^0 =[\hat{X}_t^{0T}, \hat{X}_t^{(N)T}, \hat{\overline{X}}_t^T]^T, \quad Y_t^i = [\hat{X}_t^{iT}, Y_t^{0T}]^T, \tag{95}$$

where we denote $v^{(N)} = (1/N)\sum_{i=1}^N v^i$ and $v^{(-i)} = (1/N)\sum_{j=1, j\neq i}^N v^j$, and likewise for $W^{(N)}$ and $W^{(-i)}$. The initial conditions in (95) are determined using (62) on $[0, t]$. Both (93) and (94) share the last component $\overline{X}_s$ described by (87), which is re-displayed below:

$$\dot{\overline{X}}_s = G X_s^0 + (A + F)\overline{X}_s + B_1 v^0 + B v^{(N)}, \quad t \le s \le t+\epsilon.$$

See Remark 5.2 for the requirement underlying the construction of this equation.

For a real matrix M, define

$$\varPsi_M(s_2, s_1) = \int_{s_1}^{s_2} \exp\{(s_2 - s)M\} ds, \qquad \text{for all } 0 \le s_1 \le s_2 \le T.$$

By the variation of constants formula, for $s \in [t, t+\epsilon]$ we have

$$\begin{aligned} Y_s^0 =& \exp\{(s-t)\boldsymbol{A}_0\} Y_t^0 + \varPsi_{\boldsymbol{A}_0}(s,t)(\boldsymbol{B}_0 v^0 + \boldsymbol{B}_0^a v^{(N)}) \\ &+ \int_t^s \exp\{(s-\tau)\boldsymbol{A}_0\}(\boldsymbol{D}_0 dW_\tau^0 + \boldsymbol{D}_0^a dW_\tau^{(N)}), \end{aligned} \tag{96}$$
$$\begin{aligned} Y_s^i =& \exp\{(s-t)\boldsymbol{A}\} Y_t^i + \varPsi_{\boldsymbol{A}}(s,t)(\boldsymbol{B}_1 v^0 + \boldsymbol{B}_N^a v^i + \boldsymbol{B}^a v^{(-i)}) \\ &+ \int_t^s \exp\{(s-\tau)\boldsymbol{A}\}(\boldsymbol{D} dW_\tau^0 + \boldsymbol{D}_N dW_\tau^i + \boldsymbol{D}^a dW_\tau^{(-i)}). \end{aligned} \tag{97}$$

We proceed to specify the auxiliary game model. The $(N+1)$-player static Stackelberg–Nash game has actions $(v^0, v^1, \cdots, v^N)$ on $[t, t+\epsilon]$ and costs

$$J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,\epsilon}), \quad J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon}, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,-i,\epsilon}), \quad 1 \le i \le N, \tag{98}$$

where the leader chooses its action v^0 and the random vector $\hat{\boldsymbol{X}}_t^c$ is observed by all players; in analogue to (83)–(84), we define

$$J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,\epsilon}) =\mathbb{E}\int_t^T L_0(X_s^0, \mu_s^{(N)}, \hat{u}_s^{0,\epsilon}) ds + \mathbb{E} g_0(X_T^0, \mu_T^{(N)}), \tag{99}$$

$$J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon}, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,-i,\epsilon}) = \mathbb{E}\int_t^T L(X_s^i, X_s^0, \mu_s^{(N)}, \hat{u}_s^{i,\epsilon}, \hat{u}_s^{0,\epsilon})ds + \mathbb{E}g(X_T^i, X_T^0, \mu_T^{(N)}). \tag{100}$$

In this static game, we adapt Definition C.1 of the Stackelberg–Nash equilibrium by allowing general action spaces. Due to the particular form of the control laws $(\hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,\epsilon})$, and in view of (96)–(97), here J_0^{N+1} and J_i^{N+1} depend on $\hat{\boldsymbol{X}}_t^c$ through $(Y_t^0, v^0, \mathbf{v}^F)$ and $(Y_t^i, v^0, \mathbf{v}^F)$, where $\mathbf{v}^F = (v^1, \cdots, v^N)$, respectively, and so we write the costs in the form of (98). Similarly, for $(\hat{u}^0, \hat{\mathbf{u}}^F)$, we write the costs $J_0^{N+1}(t, Y_t^0, \hat{u}^0, \hat{\mathbf{u}}^F)$ and $J_i^{N+1}(t, Y_t^i, \hat{u}^i, \hat{u}^0, \hat{\mathbf{u}}^{F,-i})$ for $1 \le i \le N$.

Denote

$$\begin{aligned} &S_0^{N+1}(t, \epsilon, Y_t^0, v^0, \mathbf{v}^F) \\ &\quad = \mathbb{E}\left[\int_t^{t+\epsilon} L_0(X_s^0, \mu_s^{(N)}, v^0)ds + \hat{V}_0(t+\epsilon, Y_{t+\epsilon}^0)\middle|\hat{\boldsymbol{X}}_t^c\right], \end{aligned} \tag{101}$$

$$\begin{aligned} &S_i^{N+1}(t, \epsilon, Y_t^i, v^i, v^0, \mathbf{v}^{F,-i}) \\ &\quad = \mathbb{E}\left[\int_t^{t+\epsilon} L(X_s^i, X_s^0, \mu_s^{(N)}, v^i, v^0)ds + \hat{V}(t+\epsilon, Y_{t+\epsilon}^i)\middle|\hat{\boldsymbol{X}}_t^c\right], \end{aligned} \tag{102}$$

where $\mathbf{v}^{F,-i}$ is obtained from $\mathbf{v}^F$ by excluding v^i. The above processes $\{Y_s^j, t \le s \le t+\epsilon\}$, $0 \le j \le N$, are determined by (96)–(97). The functions $\hat{V}_0$ and $\hat{V}$ have been specified by Lemma 4.2. For each fixed t, S_0^{N+1} and S_i^{N+1} are Borel measurable function of $\hat{\boldsymbol{X}}_t^c$. A more concrete form of S_0^{N+1} and S_i^{N+1} will be derived in our further analysis; see (D.1) and (D.2). Then we have the relation

$$J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,\epsilon}) = \mathbb{E}[S_0^{N+1}(t, \epsilon, Y_t^0, v^0, \mathbf{v}^F)], \tag{103}$$

$$J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon}, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,-i,\epsilon}) = \mathbb{E}[S_i^{N+1}(t, \epsilon, Y_t^i, v^i, v^0, \mathbf{v}^{F,-i})]. \tag{104}$$

5.3 Performance of the leader

Before proving that (90) holds for the set of decentralized strategies $(\hat{u}^0, \cdots, \hat{u}^N)$, we make a cost comparison when $(\hat{u}^0, \cdots, \hat{u}^N)$ is perturbed in a particular way.

To avoid the notation from becoming too heavy, below we will often write $\hat{\mathbb{P}}^0(t)$ as $\hat{\mathbb{P}}_t^0$, and $\hat{\mathbb{P}}(t)$ as $\hat{\mathbb{P}}_t$. This is similarly done for $\hat{r}^0$ and $\hat{r}$. Throughout Sections 5.3 and 5.4, each term of the form $\mathcal{O}(h(\epsilon))$ satisfies $|\mathcal{O}(h(\epsilon))| \le C_0 h(\epsilon)$ for a constant C_0 independent of (t, N). For instance, we have $|\mathcal{O}(\epsilon)| \le C_0\epsilon$. Also, each term $\mathcal{O}(\epsilon)$ (or $\mathcal{O}(1)$, etc) is deterministic and can be determined in a concrete form, but for our purpose it is adequate to indicate the generic form $\mathcal{O}(\epsilon)$.

Lemma 5.7. Under Assumption 1, there exists $\epsilon^* > 0$ such that for all $0 < \epsilon \le \epsilon^*$, the static Stackelberg–Nash game with costs (103)–(104) has a unique equilibrium, which takes the form

$$v^{0*} = -R_0^{-1}\widetilde{\boldsymbol{B}}_0^T\hat{\mathbb{P}}^0(t)Y_t^0 + \mathcal{O}(\epsilon)Y_t^0, \tag{105}$$

$$v^{i*} = -[R^{-1}\boldsymbol{B}_N^{aT}\hat{\mathbb{P}}(t) + \mathcal{O}(\epsilon)]Y_t^i + R^{-1}R_2R_0^{-1}\widetilde{\boldsymbol{B}}_0^T\hat{\mathbb{P}}^0(t)Y_t^0, \quad 1 \le i \le N, \tag{106}$$

where the term $\mathcal{O}(\epsilon)$ in (106) is the same for all $i \in \{1, \cdots, N\}$.

Proof. See appendix D. □

Lemma 5.8. Suppose $\epsilon \in (0, \epsilon^*]$ for ϵ^* specified in Lemma 5.7. Under Assumptions 1 and 2, let $(v^{0*}, v^{1*}, \cdots, v^{N*})$ be the Stackelberg–Nash equilibrium of the static game with costs (103)–(104). For $0 \le j \le N$, define $\hat{u}^{j,\epsilon*}$ on $[0, T]$ by

$$\hat{u}_s^{j,\epsilon*} := \begin{cases} v^{j*}, & s \in [t, t+\epsilon], \\ \hat{u}_s^j, & s \in [0, T] \setminus [t, t+\epsilon], \end{cases}$$

where $\hat{u}^j$ is given by (59)–(60). Then it holds that

$$\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\Big[J_0^{N+1}(t,Y_t^0,\hat{u}^0,\hat{\mathbf{u}}^F)-J_0^{N+1}(t,Y_t^0,\hat{u}^{0,\epsilon*},\hat{\mathbf{u}}^{F,\epsilon*})\Big]=O(d_N^X), \tag{107}$$

where the limit on the left hand side exists and $\hat{\mathbf{u}}^{F,\epsilon*}:=(\hat{u}^{1,\epsilon*},\cdots,\hat{u}^{N,\epsilon*})$. The term d_N^X is given by (91).

Proof. See appendix D. □

5.4 Performance of the followers

Select a single player $\mathcal{A}_i$ and consider its auxiliary optimization problem as follows. It chooses its control ν^i on $[t,t+\epsilon]$ while the leader and the other $N-1$ followers take the strategies $(\hat{u}^0,\hat{\mathbf{u}}^{F,-i})$ on $[t,t+\epsilon]$. All $N+1$ players take the strategies $(\hat{u}^0,\hat{\mathbf{u}}^F)$ on $(t+\epsilon,T]$. Denote

$$\boldsymbol{A}_1=\begin{bmatrix} A & G+B_1K_1^0 & F & B_1K_2^0\\ 0 & A_0+B_0K_1^0 & F_0 & B_0K_2^0\\ -BK_1/N & G+B_1K_1^0+BK_2(1-1/N) & A+F+BK_1 & \boldsymbol{A}_{1,34}\\ 0 & G+B_1K_1^0+BK_2 & 0 & \boldsymbol{A}_{1,44}\end{bmatrix},$$

where

$$\begin{aligned}\boldsymbol{A}_{1,34}=&B_1K_2^0+BK_3(1-1/N),\\ \boldsymbol{A}_{1,44}=&A+F+B_1K_2^0+B(K_1+K_3).\end{aligned}$$

Player $\mathcal{A}_i$ chooses ν^i to minimize the cost $J_i^{N+1}(t,Y_t^i,\hat{u}^{i,\epsilon},\hat{u}^0,\hat{\mathbf{u}}^{F,-i})$, which is defined similarly to (100) and may be written as

$$J_i^{N+1}(t,Y_t^i,\hat{u}^{i,\epsilon},\hat{u}^0,\hat{\mathbf{u}}^{F,-i})=\mathbb{E}\left[\int_t^{t+\epsilon}L(X_s^i,X_s^0,\mu_s^{(N)},\nu^i,\hat{u}_s^0)ds+\hat{V}(t+\epsilon,Y_{t+\epsilon}^i)\right], \tag{108}$$

where, for Y_s^i defined in (92) and $t\le s\le t+\epsilon$,

$$dY_s^i=(\boldsymbol{A}_1Y_s^i+\boldsymbol{B}_N^b\nu^i)dt+\boldsymbol{D}dW_s^0+\boldsymbol{D}_NdW_s^i+\boldsymbol{D}^a dW_s^{(-i)}. \tag{109}$$

Lemma 5.9. Under Assumption 1, there exists $\bar{\epsilon}>0$ such that for all $0<\epsilon\le\bar{\epsilon}$, the optimization problem on $[t,t+\epsilon]$ to minimize (108) has a unique minimizer $\nu^{i\star}$, which takes the form

$$\nu^{i\star}=-R^{-1}\{R_2[0,K_1^0,0,K_2^0]+\boldsymbol{B}_N^{bT}\hat{\mathbb{P}}(t)+\mathcal{O}(\epsilon)\}Y_t^i. \tag{110}$$

Proof. See appendix D. □

Remark 5.10. For (106) and (110), we can show that for some constant C,

$$|v^{i*}-\nu^{i\star}|\le C|X_t^{(N)}-\overline{X}_t|+C|Y_t^i|\epsilon.$$

Denote

$$\hat{u}_s^{i,\epsilon\star}:=\begin{cases}\nu^{i\star}, & s\in[t,t+\epsilon],\\ \hat{u}_s^i, & s\in[0,T]\setminus[t,t+\epsilon].\end{cases} \tag{111}$$

Lemma 5.11. For the strategy given by (111),

$$\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}[J_i^{N+1}(t,Y_t^i,\hat{u}^i,\hat{u}^0,\hat{\mathbf{u}}^{F,-i})-J_i^{N+1}(t,Y_t^i,\hat{u}^{i,\epsilon\star},\hat{u}^0,\hat{\mathbf{u}}^{F,-i})]=O(d_N^X),$$

where the limit on the left hand side exists and d_N^X is given by (91).

Proof. See appendix D. □

5.5 Proof of Theorem 5.6

Since $(v^{0*}, \mathbf{v}^{F*}) = (v^{0*}, v^{1*}, \cdots, v^{N*})$ in (105)–(106) is a Stackelberg–Nash equilibrium for the static game of $N+1$ players with costs (103)–(104) and actions $(v^0, \cdots, v^N)$ on $[t, t+\epsilon]$, it follows that

$$J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon*}, \hat{\mathbf{u}}^{F,\epsilon*}) \le J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,\epsilon}),$$

for any $\hat{u}^{0,\epsilon}$ (with its associated variation v^0 on $[t, t+\epsilon]$) and $\hat{\mathbf{u}}^{F,\epsilon} = (\hat{u}^{1,\epsilon}, \cdots, \hat{u}^{N,\epsilon})$, where $\hat{\mathbf{u}}^{F,\epsilon}$ with its associated $\mathbf{v}^F = (v^1, \cdots, v^N)$ is a Nash equilibrium among the followers for the given $\hat{u}^{0,\epsilon}$. Therefore,

$$\begin{aligned}
&\limsup_{\epsilon\downarrow 0} \frac{1}{\epsilon}\Big[J_0^{N+1}(t, Y_t^0, \hat{u}^0, \hat{\mathbf{u}}^F) - J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon}, \hat{\mathbf{u}}^{F,\epsilon})\Big] \\
&\le \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}\Big[J_0^{N+1}(t, Y_t^0, \hat{u}^0, \hat{\mathbf{u}}^F) - J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon*}, \hat{\mathbf{u}}^{F,\epsilon*})\Big].
\end{aligned}$$

In view of the above inequality and Lemma 5.8, condition (90) holds for $(\hat{u}^0, \hat{\mathbf{u}}^F)$.

Let $\hat{u}^{i,\epsilon}$ be a modification of $\hat{u}^i$ with any admissible v^i. For $\hat{u}^{i,\epsilon\star}$ given by (111), Lemma 5.9 implies

$$J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon\star}, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}) \le J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon}, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}).$$

Hence we have

$$\begin{aligned}
&\limsup_{\epsilon\downarrow 0} \frac{1}{\epsilon}\Big[J_i^{N+1}(t, Y_t^i, \hat{u}^i, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}) - J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon}, \hat{u}^0, \hat{\mathbf{u}}^{F,-i})\Big] \\
&\le \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}\Big[J_i^{N+1}(t, Y_t^i, \hat{u}^i, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}) - J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon\star}, \hat{u}^0, \hat{\mathbf{u}}^{F-i})\Big].
\end{aligned}$$

which combined with Lemma 5.11 implies that condition (89) holds for $(\hat{u}^0, \hat{\mathbf{u}}^F)$.

6 Concluding remarks

We study linear-quadratic mean field Stackelberg games with a leader and a large number of followers. We use master equations to determine decentralized strategies for the finite-population model and establish an ε_N-Stackelberg equilibrium property for such strategies when the game is played by a stream of t-selves of the original $N+1$ players. For future work, it will be of interest to generalize the asymptotic equilibrium analysis to nonlinear models.

Appendix A: Proof of Theorem 3.3

Proof. Suppose (24)–(25) has a quadratic solution of the form (27)–(28). Recall $\mathbf{z}_0 = (x_0^T, \bar{z}^T)^T$ and $\mathbf{z} = (z_i^T, z_0^T, \bar{z}^T)^T$, where $\bar{z} = \langle y \rangle_\mu$. We have

$$\begin{cases}
\partial_{z_0} V_0 = 2z_0^T P_{11}^0 + 2\bar{z}^T P_{21}^0, \quad \partial_{z_0}^2 V_0 = 2P_{11}^0, \\
\delta_\mu V_0(t, z_0, \mu; y) = 2\bar{z}^T P_{22}^0 y + 2z_0^T P_{12}^0 y, \\
\partial_y \delta_\mu \partial V_0(t, z_0, \mu; y) = 2\bar{z}^T P_{22}^0 + 2z_0^T P_{12}^0, \\
\partial_y^2 \delta_\mu V_0(t, z_0, \mu; y) = 0, \\
\partial_{z_i} V = 2z_i^T P_{11} + 2z_0^T P_{21} + 2\bar{z}^T P_{31}, \quad \partial_{z_i}^2 V = 2P_{11}, \\
\partial_{z_0} V = 2z_0^T P_{22} + 2z_i^T P_{12} + 2\bar{z}^T P_{32}, \quad \partial_{z_0}^2 V = 2P_{22}, \\
\delta_\mu V(t, z_i, z_0, \mu; y) = 2\bar{z}^T P_{33} y + 2z_i^T P_{13} y + 2z_0^T P_{23} y, \\
\partial_y \delta_\mu V(t, z_i, z_0, \mu; y) = 2\bar{z}^T P_{33} + 2z_i^T P_{13} + 2z_0^T P_{23}, \\
\partial_y^2 \delta_\mu V(t, z_i, z_0, \mu; y) = 0.
\end{cases} \tag{A.1}$$

Denote the right hand side of (24) as $s_0 = \sum_{k=1}^{7} s_{0k}$, where the components s_{0k} correspond to the seven terms on the right hand side of (24) in the same order. Similarly, denote the right hand side of (25) as $s = \sum_{k=1}^{10} s_k$.

We substitute (A.1) into the right hand side of (24) to get

$$\begin{aligned}
s_{01} &= \mathbf{z}_0^T \left\{ \begin{bmatrix} P_{11}^0 \\ P_{21}^0 \end{bmatrix} [A_0, F_0] + \begin{bmatrix} A_0^T \\ F_0^T \end{bmatrix} [P_{11}^0,\ P_{12}^0] \right\} \mathbf{z}_0, \\
s_{04} + s_{05} &= \mathbf{z}_0^T \left\{ \begin{bmatrix} P_{12}^0 \\ P_{22}^0 \end{bmatrix} [\mathbb{A}_{0,21},\ \mathbb{A}_{0,22}] + \begin{bmatrix} \mathbb{A}_{0,21}^T \\ \mathbb{A}_{0,22}^T \end{bmatrix} [P_{21}^0,\ P_{22}^0] \right\} \mathbf{z}_0, \\
s_{02} &= \mathrm{Tr}(P_{11}^0 \Sigma_{w^0}), \qquad s_{03} = \mathbf{z}_0^T [\![(I,\ -\Gamma_0)]\!]_{Q_0}^2 \mathbf{z}_0, \\
s_{06} &= -\mathbf{z}_0^T \mathbb{P}^0 \mathbb{B}_0 R_0^{-1} \mathbb{B}_0^T \mathbb{P}^0 \mathbf{z}_0, \qquad s_{07} = 0,
\end{aligned} \tag{A.2}$$

where $\mathbb{A}_{0,21}$ and $\mathbb{A}_{0,22}$ are, respectively, the lower left and lower right submatrices within the partition of $\mathbb{A}_0$. Note that $s_{01} + s_{04} + s_{05} = \mathbb{P}^0\mathbb{A}_0 + \mathbb{A}_0^T\mathbb{P}^0$. By the above calculations, the right hand side of (24) is

$$s_0 = \mathbf{z}_0^T [\mathbb{P}^0\mathbb{A}_0 + \mathbb{A}_0^T\mathbb{P}^0 - \mathbb{P}^0\mathbb{B}_0 R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0 + [\![(I, -\Gamma_0)]\!]_{Q_0}^2]\mathbf{z}_0 + \mathrm{Tr}(P_{11}^0 \Sigma_{w^0}). \tag{A.3}$$

Substituting (A.1) into the left hand side of (24) gives

$$-\mathbf{z}_0^T \dot{\mathbb{P}}^0(t)\mathbf{z}_0 - \dot{r}^0(t). \tag{A.4}$$

Equating (A.4) to (A.3), we have that $\mathbb{P}^0$ satisfies (29) on $[0, T]$.

We substitute (28) into the right hand side of (28) to get

$$\begin{aligned}
s_1 &= \mathbf{z}^T \left\{ \begin{bmatrix} P_{12} \\ P_{22} \\ P_{32} \end{bmatrix} [0, A_0, F_0] + \begin{bmatrix} 0 \\ A_0^T \\ F_0^T \end{bmatrix} [P_{21}, P_{22}, P_{23}] \right\} \mathbf{z}, \\
s_2 &= \mathbf{z}^T \left\{ \begin{bmatrix} P_{11} \\ P_{21} \\ P_{31} \end{bmatrix} [A, G, F] + \begin{bmatrix} A^T \\ G^T \\ F^T \end{bmatrix} [P_{11}, P_{12}, P_{13}] \right\} \mathbf{z}, \\
s_3 &= -\mathbf{z}^T \begin{bmatrix} P_{11} \\ P_{21} \\ P_{31} \end{bmatrix} B R^{-1} B^T [P_{11}, P_{12}, P_{13}] \mathbf{z} \\
&= -\mathbf{z}^T \mathbb{P}\mathbb{B} R^{-1} \mathbb{B}^T \mathbb{P} \mathbf{z}, \\
s_6 + s_7 &= \mathbf{z}^T \left\{ \begin{bmatrix} P_{13} \\ P_{23} \\ P_{33} \end{bmatrix} [0, \mathbb{A}_{0,21}, \mathbb{A}_{0,22}] + \begin{bmatrix} 0 \\ \mathbb{A}_{0,21}^T \\ \mathbb{A}_{0,22}^T \end{bmatrix} [P_{31}, P_{32}, P_{33}] \right\} \mathbf{z}, \\
s_4 &= \mathrm{Tr}(P_{22}\Sigma_{w^0} + P_{11}\Sigma_w), \quad s_5 = \mathbf{z}^T [\![(I, -\Gamma_1, -\Gamma_2)]\!]_Q^2 \mathbf{z}, \\
s_8 &= \mathbf{z}_0^T [\![R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0]\!]_{R_{12}}^2 \mathbf{z}_0 = \mathbf{z}^T \mathbb{J}_1^T [\![R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0]\!]_{R_{12}}^2 \mathbb{J}_1 \mathbf{z}, \\
s_9 &= -2\mathbf{z}_0^T \mathbb{P}^0\mathbb{B}_0 R_0^{-1}\mathbb{B}_1^T\mathbb{P}\mathbf{z} = -2\mathbf{z}^T \mathbb{J}_1^T \mathbb{P}^0\mathbb{B}_0 R_0^{-1}\mathbb{B}_1^T\mathbb{P}\mathbf{z} \\
&= -\mathbf{z}^T (\mathbb{J}_1^T \mathbb{P}^0\mathbb{B}_0 R_0^{-1}\mathbb{B}_1^T\mathbb{P} + \mathbb{P}\mathbb{B}_1 R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1)\mathbf{z}, \\
s_{10} &= 0.
\end{aligned} \tag{A.5}$$

It is easy to show $s_1 + s_2 + s_6 + s_7 = \mathbb{P}\mathbb{A} + \mathbb{A}^T\mathbb{P}$. By the above calculations, the right hand side of (25) is

$$\begin{aligned}
s = &\mathbf{z}^T \{\mathbb{P}\mathbb{A} + \mathbb{A}^T\mathbb{P} - \mathbb{P}\mathbb{B}R^{-1}\mathbb{B}^T\mathbb{P} - \mathbb{P}\mathbb{B}_1 R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1 \\
&- \mathbb{J}_1^T\mathbb{P}^0\mathbb{B}_0 R_0^{-1}\mathbb{B}_1^T\mathbb{P} + \mathbb{J}_1^T\mathbb{P}^0\mathbb{B}_0 R_0^{-1} R_{12} R_0^{-1}\mathbb{B}_0^T\mathbb{P}^0\mathbb{J}_1
\end{aligned}$$

$$+ [\![(I, -\Gamma_1, -\Gamma_2)]\!]_Q^2\}\mathbf{z} + \mathrm{Tr}(P_{22}\Sigma_{w^0} + P_{11}\Sigma_w). \tag{A.6}$$

Substituting (28) into the left hand side of (25) gives

$$-\mathbf{z}^T\dot{\mathbb{P}}(t)\mathbf{z} - \dot{r}(t). \tag{A.7}$$

By equating (A.7) to (A.6), we have that $\mathbb{P}$ satisfies (30) on $[0, T]$.

Conversely, if (29)–(30) has a solution $(\mathbb{P}^0, \mathbb{P})$, we can substitute $(\mathbb{P}^0, \mathbb{P})$ into (31)–(32) to solve for a unique solution (r^0, r) on $[0, T]$. By (A.3)–(A.4) and (A.6)–(A.7), the pair in (27)–(28) is verified to be a solution of the master equations (24)–(25). □

Appendix B

Consider the linear SDE

$$dX_s = [A(s)X_s + k(s)]ds + \sigma(s)dW_s, \qquad s \in [t, T], \tag{B.1}$$

with initial condition $X_t = x$ for $t \in [0, T]$ and $x \in \mathbb{R}^d$. The standard Brownian motion W_s is in $\mathbb{R}^{d_1}$. The cost functional is

$$\begin{aligned} V(t, x) =&\mathbb{E}\int_t^T [X_s^T Q(s)X_s + 2\eta^T(s)X_s + h(s)]ds \\ &+ \mathbb{E}(X_T^T Q_f X_T + 2\eta_f^T X_T + h_f). \end{aligned} \tag{B.2}$$

The matrix or vector (or scalar) functions $A(t), k(t), \sigma(t), Q(t), \eta(t)$ and $h(t)$ are deterministic, bounded and Lebesgue measurable on $[0, T]$. The parameters Q_f, η_f and h_f are deterministic. We introduce the ODE system

$$\begin{cases} 0 = \dot{P}(t) + A^T P + PA + Q, \\ 0 = \dot{S}(t) + A^T S + Pk + \eta, \\ 0 = \dot{r}(t) + 2S^T k + h + \mathrm{Tr}(P\sigma\sigma^T), \end{cases} \qquad 0 \le t \le T, \tag{B.3}$$

where $P(T) = Q_f$, $S(T) = \eta_f$ and $r(T) = h_f$. The ODE system (B.3) has a unique solution on $[0, T]$.

Lemma B.1. For V defined in (B.2), we have the representation

$$V(t, x) = x^T P(t)x + 2x^T S(t) + r(t), \qquad t \in [0, T],\ x \in \mathbb{R}^d. \tag{B.4}$$

Proof. Let $\hat{V}(t, x)$ be the function defined by right hand side of (B.4). Applying Itô's formula and using (B.3), we get an SDE of $d\hat{V}(s, X_s)$ for $s \in [t, T]$. Integrating the resulting SDE on $[t, T]$ and taking expectation, we obtain

$$\mathbb{E}\hat{V}(T, X_T) - \hat{V}(t, x) = -\mathbb{E}\int_t^T [X_s^T Q(s)X_s + 2\eta^T(s)X_s + h(s)]ds. \tag{B.5}$$

Writing $\hat{V}(T, X_T)$ using the terminal conditions Q_f, η_f, h_f, we see that $\hat{V}(t, x)$ is equal to the right hand side of (B.2). □

Appendix C

This part formulates a static Stackelberg–Nash game with a leader $\mathcal{A}_0$ and N followers $\mathcal{A}_i$, $1 \le i \le N$. Player $\mathcal{A}_j$ has cost $\mathcal{J}_j$ and chooses action $u_j \in \mathbb{R}^{n_1}$. Their costs depend continuously on a small parameter $\varepsilon \ge 0$. After the leader's action u_0 is announced, the followers play a Nash game among themselves. Denote the two symmetric matrices $M^{0,\epsilon} = (M^{0,\epsilon}_{kl})_{1\le k,l\le 3}$ and $M^{F,\epsilon} = (M^{F,\epsilon}_{kl})_{1\le k,l\le 4}$, where the partitions satisfy $M^{0,\epsilon}_{kk}, M^{F,\epsilon}_{ll} \in \mathbb{R}^{n_1\times n_1}$ for all $1 \le k \le 2$, $1 \le l \le 3$, and $M^{0,\epsilon}_{33}, M^{F,\epsilon}_{44} \in \mathbb{R}^{n\times n}$. Denote $\mathbf{u}_F = (u_1, \cdots, u_N)$. Let $\mathbf{u}_{F,-i}$ be obtained from $\mathbf{u}_F$ by excluding u_i. Define $u^{(N)} = (1/N)\sum_{i=1}^N u_i$ and $u^{(-i)} = (1/N)\sum_{k=1,k\neq i}^N u_k$.

The costs are given by

$$\mathcal{J}_0(u_0, \mathbf{u}_F) = |(u_0^T, u^{(N)T}, y_0^T)^T|^2_{M^{0,\epsilon}}, \tag{C.1}$$

$$\mathcal{J}_i(u_i, u_0, \mathbf{u}_{F,-i}) = |(u_i^T, u_0^T, u^{(-i)T}, y_i^T)^T|^2_{M^{F,\epsilon}}, \qquad 1 \le i \le N, \tag{C.2}$$

where $y_j \in \mathbb{R}^n$, $0 \le j \le N$ are constants. The parameters $(y_0, \cdots, y_N)$ have been included in the quadratic forms for convenience of applying the obtained results to Section 5.

Below we consider $M^{0,\epsilon} = (M^{0,\epsilon}_{kl})$ and $M^{F,\epsilon} = (M^{F,\epsilon}_{kl})$ satisfying

$$M^{0,\epsilon} = \begin{bmatrix} \mathcal{O}(1) & \mathcal{O}(\epsilon) & \mathcal{O}(1) \\ \mathcal{O}(\epsilon) & \mathcal{O}(\epsilon) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \end{bmatrix}$$

and

$$M^{F,\epsilon} = \begin{bmatrix} \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(\epsilon) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(\epsilon) & \mathcal{O}(1) \\ \mathcal{O}(\epsilon) & \mathcal{O}(\epsilon) & \mathcal{O}(\epsilon) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \end{bmatrix}.$$

Denote $M^0_{kl} = M^{0,\epsilon}_{kl}|_{\epsilon=0}$ and $M^F_{kl} = M^{F,\epsilon}_{kl}|_{\epsilon=0}$ for all k and l.

Assumption C.1. Both $M^{0,\epsilon}$ and $M^{F,\epsilon}$ are Lipschitz continuous in $\epsilon \in [0, \bar\epsilon]$ for some $\bar\epsilon > 0$; M^0_{11} and M^F_{11} are positive definite.

Definition C.1. We call $(u_0^*, \cdots, u_N^*)$ a Stackelberg–Nash equilibrium if the two conditions hold: (i) for each i, $\mathcal{J}_i(u_i^*, u_0^*, \mathbf{u}^*_{F,-i}) \le \mathcal{J}_i(v_i, u_0^*, \mathbf{u}^*_{F,-i})$ for all $v_i \in \mathbb{R}^{n_1}$ (i.e., given u_0^*, $\mathbf{u}_F^*$ is a Nash equilibrium for the followers), (ii) $\mathcal{J}_0(u_0^*, \mathbf{u}_F^*) \le \mathcal{J}_0(v_0, \mathbf{u}_F^{ne}(v_0))$ for all $v_0 \in \mathbb{R}^{n_1}$, whenever $\mathbf{u}_F^{ne}(v_0)$ is a Nash equilibrium for the N followers with v_0 given.

In Lemma C.2 and its proof, all terms written as $\mathcal{O}(\epsilon)$ may be uniquely determined using $M^{0,\epsilon}, M^{F,\epsilon}, y_0, \cdots, y_N$. Their specific forms are immaterial for our purpose of application.

Lemma C.2. Under Assumption C.1, there exists $\epsilon^* \in (0, \bar\epsilon]$ such that for all $0 < \epsilon \le \epsilon^*$, the Stackelberg–Nash game (C.1)–(C.2) has a unique equilibrium with u_0^* and $\mathbf{u}_F^* = (u_1^*, \cdots, u_N^*)$ taking the following form

$$\begin{aligned} u_0^* =& -(M^0_{11})^{-1}[M^0_{13} - M^F_{21}(M^F_{11})^{-1}M^0_{23}]y_0 \\ &+ \mathcal{O}(\epsilon)y_0 + \mathcal{O}(\epsilon)y^{(N)}, \end{aligned} \tag{C.3}$$

$$\begin{aligned} u_i^* =& (M^F_{11})^{-1}M^F_{12}(M^0_{11})^{-1}[M^0_{13} - M^F_{21}(M^F_{11})^{-1}M^0_{23}]y_0 \\ &- (M^F_{11})^{-1}M^F_{14}y_i + \mathcal{O}(\epsilon)y_i + \mathcal{O}(\epsilon)y_0 + \mathcal{O}(\epsilon)y^{(N)}, \quad 1 \le i \le N. \end{aligned} \tag{C.4}$$

Proof. Since $M^F_{11} > 0$ and $M^{F,\epsilon}_{13} = \mathcal{O}(\epsilon)$, there exists $\epsilon_1 > 0$ such that $M^{F,\epsilon}_{11} > 0$, $\det(M^{F,\epsilon}_{11} + (1 - 1/N)M^{F,\epsilon}_{13}) > 0$, and $\det(M^{F,\epsilon}_{11} - M^{F,\epsilon}_{13}/N) > 0$ for all $0 < \epsilon \le \epsilon_1$ and $N \ge 2$.

Step 1. For u_0 fixed, we check the Nash equilibrium $\tilde{\mathbf{u}}_F = (\tilde u_1, \cdots, \tilde u_N)$ of the N followers by the first order condition

$$0 = M^{F,\epsilon}_{11}\tilde u_i + M^{F,\epsilon}_{12}u_0 + M^{F,\epsilon}_{13}\tilde u^{(-i)} + M^{F,\epsilon}_{14}y_i, \quad 1 \le i \le N, \tag{C.5}$$

for all $0 < \epsilon \le \epsilon_1$ and for all $N \ge 2$. From (C.5) we obtain

$$\tilde u^{(N)} = -[M^{F,\epsilon}_{11} + (1 - 1/N)M^{F,\epsilon}_{13}]^{-1}[M^{F,\epsilon}_{12}u_0 + M^{F,\epsilon}_{14}y^{(N)}]. \tag{C.6}$$

We further uniquely determine $\tilde{u}_i$. This gives the Nash equilibrium of the followers.

Step 2. When the followers take their Nash equilibrium strategies $\tilde{\mathbf{u}}_F$ in response to u_0, the leader's cost is $\mathcal{J}_0(u_0, \tilde{\mathbf{u}}_F) = |(u_0^T, \tilde{u}^{(N)T}, y_0^T)^T|^2_{M^{0,\epsilon}}$. Our next step is to rewrite $\mathcal{J}_0(u_0, \tilde{\mathbf{u}}_F)$ in terms of $(u_0, y_0, \cdots, y_N)$. Denote $\hat{M}_{11}^{F,\epsilon} = M_{11}^{F,\epsilon} + (1 - 1/N)M_{13}^{F,\epsilon}$ and

$$U^\epsilon = \begin{bmatrix} I & 0 & 0 \\ -(\hat{M}_{11}^{F,\epsilon})^{-1}M_{12}^{F,\epsilon} & -(\hat{M}_{11}^{F,\epsilon})^{-1}M_{14}^{F,\epsilon} & 0 \\ 0 & 0 & I \end{bmatrix}, \quad \widetilde{M}^{0,\epsilon} = U^{\epsilon T} M^{0,\epsilon} U^\epsilon.$$

We have the relation

$$[u_0^T, \tilde{u}^{(N)T}, y_0^T]^T = U^\epsilon [u_0^T, y^{(N)T}, y_0^T]^T.$$

The submatrices of the symmetric matrix $\widetilde{M}^{0,\epsilon} = (\widetilde{M}_{kl}^{0,\epsilon})_{1 \le k,l \le 3}$ satisfy

$$\begin{aligned} \widetilde{M}_{11}^{0,\epsilon} &= M_{11}^0 + \mathcal{O}(\epsilon), \\ \widetilde{M}_{31}^{0,\epsilon} &= M_{31}^0 - M_{32}^0 (M_{11}^F)^{-1} M_{12}^F + \mathcal{O}(\epsilon), \\ \widetilde{M}_{32}^{0,\epsilon} &= -M_{32}^0 (M_{11}^F)^{-1} M_{14}^F + \mathcal{O}(\epsilon), \\ \widetilde{M}_{33}^{0,\epsilon} &= M_{33}^0 + \mathcal{O}(\epsilon), \end{aligned}$$

and the three unspecified submatrices $\widetilde{M}_{kl}^{0,\epsilon}$ are $\mathcal{O}(\epsilon)$.

Then we have

$$\mathcal{J}_0(u_0, \tilde{\mathbf{u}}_F) = |(u_0^T, y^{(N)T}, y_0^T)^T|^2_{\widetilde{M}^{0,\epsilon}}.$$

Since $M_{11}^0 > 0$, there exists $\epsilon_2 > 0$ such that for all $0 \le \epsilon \le \epsilon_2$, we have $\widetilde{M}_{11}^{0,\epsilon} > 0$. For all $0 < \epsilon \le \min(\epsilon_1, \epsilon_2)$, the first order condition

$$0 = \widetilde{M}_{11}^{0,\epsilon} u_0^* + \widetilde{M}_{12}^{0,\epsilon} y^{(N)} + \widetilde{M}_{13}^{0,\epsilon} y_0$$

determines the leader's unique optimal action

$$u_0^* = -(\widetilde{M}_{11}^{0,\epsilon})^{-1} (\widetilde{M}_{13}^{0,\epsilon} y_0 + \widetilde{M}_{12}^{0,\epsilon} y^{(N)}), \tag{C.7}$$

which can be written in the form (C.3).

Step 3. After setting $u_0 = u_0^*$ in (C.5), we obtain $(u_1^*, \cdots, u_N^*)$. Substituting u_0^* in (C.7) for u_0 in (C.6), we have that

$$\begin{aligned} u^{*(N)} =& (M_{11}^F)^{-1} M_{12}^F (M_{11}^0)^{-1} [M_{13}^0 - M_{21}^F (M_{11}^F)^{-1} M_{23}^0] y_0 \\ &- (M_{11}^F)^{-1} M_{14}^F y^{(N)} + \mathcal{O}(\epsilon) y^{(N)} + \mathcal{O}(\epsilon) y_0. \end{aligned} \tag{C.8}$$

Substituting (C.7) and (C.8) into (C.5) after writing $\tilde{u}^{(-i)} = \tilde{u}^{(N)} - (1/N)\tilde{u}_i$, we have

$$0 = (M_{11}^{F,\epsilon} - M_{13}^{F,\epsilon}/N) u_i^* + M_{12}^{F,\epsilon} u_0^* + M_{13}^{F,\epsilon} u^{*(N)} + M_{14}^{F,\epsilon} y_i,$$

which gives (C.4). □

Appendix D

This appendix provides the proofs of some lemmas used in Section 5.

Proof of Lemma 5.7. By (67) and (96), we compute the conditional expectation in (101) to obtain

$$S_0^{N+1}(t,\epsilon,Y_t^0,v^0,\mathbf{v}^F)=\epsilon|(v^{0T},v^{(N)T},Y_t^{0T})^T|^2_{M^{0,\epsilon}}+\zeta_0^\epsilon(t), \tag{D.1}$$

where

$$\zeta_0^\epsilon(t)=\epsilon\{\mathrm{Tr}[\hat{\mathbb{P}}_t^0(\boldsymbol{D}_0\boldsymbol{D}_0^T+\boldsymbol{D}_0^a\boldsymbol{D}_0^{aT}/N)]+\mathcal{O}(\epsilon)\}+\hat{r}^0_{t+\epsilon}+Y_t^{0T}\hat{\mathbb{P}}^0_{t+\epsilon}Y_t^0.$$

The matrix $M^{0,\epsilon}=(M^{0,\epsilon}_{kl})_{1\le k,l\le 3}$ is symmetric and satisfies

$$M^{0,\epsilon}_{11}=R_0+\mathcal{O}(\epsilon), \qquad M^{0,\epsilon}_{13}=\boldsymbol{B}_0^T\hat{\mathbb{P}}_t^0+\mathcal{O}(\epsilon),$$
$$M^{0,\epsilon}_{23}=\boldsymbol{B}_0^{aT}\hat{\mathbb{P}}_t^0+\mathcal{O}(\epsilon),$$
$$M^{0,\epsilon}_{33}=\hat{\mathbb{P}}_t^0\boldsymbol{A}_0+\boldsymbol{A}_0^T\hat{\mathbb{P}}_t^0+[\![(I,-\Gamma_0,0)]\!]^2_{Q_0}+\mathcal{O}(\epsilon),$$

and the three unspecified submatrices $M^{0,\epsilon}_{kl}$ are of the form $\mathcal{O}(\epsilon)$.

By (68), (97) and (102), we represent S_i^{N+1} in the form

$$S_i^{N+1}(t,\epsilon,Y_t^i,v^i,v^0,\mathbf{v}^{F,-i})=\epsilon|(v^{iT},v^{0T},v^{(-i)T},Y_t^{iT})^T|^2_{M^\epsilon}+\zeta^\epsilon(t), \tag{D.2}$$

where

$$\begin{aligned}\zeta^\epsilon(t)=&\epsilon\{\mathrm{Tr}[\hat{\mathbb{P}}_t(\boldsymbol{D}\boldsymbol{D}^T+\boldsymbol{D}_N\boldsymbol{D}_N^T+((N-1)/N^2)\boldsymbol{D}^a\boldsymbol{D}^{aT})]+\mathcal{O}(\epsilon)\}\\&+\hat{r}_{t+\epsilon}+Y_t^{iT}\hat{\mathbb{P}}_{t+\epsilon}Y_t^i.\end{aligned} \tag{D.3}$$

The matrix $M^{F,\epsilon}=(M^{F,\epsilon}_{kl})_{1\le k,l\le 4}$ is symmetric and satisfies

$$M^{F,\epsilon}_{11}=R+\mathcal{O}(\epsilon), \qquad M^{F,\epsilon}_{12}=R_2+\mathcal{O}(\epsilon), \qquad M^{F,\epsilon}_{22}=R_1+\mathcal{O}(\epsilon),$$
$$M^{F,\epsilon}_{14}=\boldsymbol{B}_N^{aT}\hat{\mathbb{P}}_t+\mathcal{O}(\epsilon), \qquad M^{F,\epsilon}_{24}=\boldsymbol{B}_1^T\hat{\mathbb{P}}_t+\mathcal{O}(\epsilon),$$
$$M^{F,\epsilon}_{34}=\boldsymbol{B}^{aT}\hat{\mathbb{P}}_t+\mathcal{O}(\epsilon),$$
$$M^{F,\epsilon}_{44}=\hat{\mathbb{P}}_t\boldsymbol{A}+\boldsymbol{A}^T\hat{\mathbb{P}}_t+[\![(I,-\Gamma_1,-\Gamma_2,0)]\!]^2_Q+\mathcal{O}(\epsilon),$$

and the four unspecified submatrices $M^{F,\epsilon}_{kl}$ are of the form $\mathcal{O}(\epsilon)$.

By the static nature of the game, the solution with costs J_k^{N+1}, $0\le k\le N$, in (103)–(104) is equivalent to the solution with costs S_k^{N+1}, $0\le k\le N$, where the actions $(v^0,\cdots,v^N)$ are optimized with respect to each individual sample point $\omega\in\Omega$.

By Lemma C.2, there exists $\epsilon^*>0$ such that for all $0<\epsilon\le\epsilon^*$, the Stackelberg–Nash game with costs (D.1)–(D.2) admits a unique equilibrium $(v^{0*},\cdots,v^{N*})$. We further obtain (105)–(106) from (C.3)–(C.4). □

Proof of Lemma 5.8. First of all, by Lemma 4.2 and (103) we have

$$J_0^{N+1}(t,Y_t^0,\hat{u}^0,\hat{\mathbf{u}}^F)=\mathbb{E}(Y_t^{0T}\hat{\mathbb{P}}_t^0Y_t^0)+\hat{r}_t^0, \tag{D.4}$$
$$J_0^{N+1}(t,Y_t^0,\hat{u}^{0,\epsilon*},\hat{\mathbf{u}}^{F,\epsilon*})=\mathbb{E}S_0^{N+1}(t,\epsilon,Y_t^0,v^{0*},\mathbf{v}^{F*}). \tag{D.5}$$

Now by (D.1), we have

$$\begin{aligned}&S_0^{N+1}(t,\epsilon,Y_t^0,v^{0*},\mathbf{v}^{F*})\\=&\epsilon v^{0*T}[R_0+\mathcal{O}(\epsilon)]v^{0*}+2\epsilon Y_t^{0T}[M^{0,\epsilon}_{31}v^{0*}+M^{0,\epsilon}_{32}v^{*(N)}]\\&+\epsilon Y_t^{0T}[\hat{\mathbb{P}}_t^0\boldsymbol{A}_0+\boldsymbol{A}_0^T\hat{\mathbb{P}}_t^0+[\![(I,-\Gamma_0,0)]\!]^2_{Q_0}+\mathcal{O}(\epsilon)]Y_t^0\\&+2\epsilon v^{0*T}M^{0,\epsilon}_{12}v^{*(N)}+\epsilon v^{*(N)T}M^{0,\epsilon}_{22}v^{*(N)}\\&+\zeta_0^\epsilon(t).\end{aligned} \tag{D.6}$$

By (59)–(60), we have

$$\widehat{\mathbb{A}}_0 Y_t^0 = \boldsymbol{A}_0 Y_t^0 + \boldsymbol{B}_0 \hat{u}_t^0 + \boldsymbol{B}_0^a \hat{u}_t^{(N)} + \boldsymbol{B}_0^c K_1(\hat{\overline{X}}_t - \hat{X}_t^{(N)}), \tag{D.7}$$

$$|\hat{u}_t^0|_{R_0}^2 = Y_t^{0T} [\![(K_1^0, 0, K_2^0)]\!]_{R_0}^2 Y_t^0, \tag{D.8}$$

where $\boldsymbol{B}_0^c = [0, 0, B^T]^T$. It follows from (D.7) that

$$\begin{aligned} Y_t^{0T}(\hat{\mathbb{P}}_t^0 \boldsymbol{A}_0 + \boldsymbol{A}_0^T \hat{\mathbb{P}}_t^0) Y_t^0 =& Y_t^{0T}(\hat{\mathbb{P}}_t^0 \widehat{\mathbb{A}}_0 + \widehat{\mathbb{A}}_0^T \hat{\mathbb{P}}_t^0) Y_t^0 - 2 Y_t^{0T} \hat{\mathbb{P}}_t^0 \\ & \cdot [\boldsymbol{B}_0 \hat{u}_t^0 + \boldsymbol{B}_0^a \hat{u}_t^{(N)} + \boldsymbol{B}_0^c K_1(\hat{\overline{X}}_t - \hat{X}_t^{(N)})]. \end{aligned} \tag{D.9}$$

Combining (D.6) and (D.9), and recalling the linear ODE (69), now we have

$$\begin{aligned} & S_0^{N+1}(t, \epsilon, Y_t^0, v^{0*}, \mathbf{v}^{F*}) \\ =& \epsilon v^{0*T} R_0 v^{0*} + 2\epsilon Y_t^{0T}(M_{31}^{0,\epsilon} v^{0*} + M_{32}^{0,\epsilon} v^{*(N)}) \\ & + \epsilon Y_t^{0T} \Big\{ -\frac{d}{dt}\hat{\mathbb{P}}_t^0 - [\![(K_1^0, 0, K_2^0)]\!]_{R_0}^2 + \mathcal{O}(\epsilon) \Big\} Y_t^0 \\ & - 2\epsilon Y_t^{0T} \hat{\mathbb{P}}_t^0 (\boldsymbol{B}_0 \hat{u}_t^0 + \boldsymbol{B}_0^a \hat{u}_t^{(N)} + \boldsymbol{B}_0^c K_1(\hat{\overline{X}}_t - \hat{X}_t^{(N)})) \\ & + 2\epsilon v^{0*T} M_{12}^{0,\epsilon} v^{*(N)} + \epsilon v^{*(N)T} M_{22}^{0,\epsilon} v^{*(N)} \\ & + \zeta_0^\epsilon(t). \end{aligned} \tag{D.10}$$

By (D.10) we calculate

$$\begin{aligned} \varDelta_0^{*,\epsilon} :=& \frac{1}{\epsilon}[S_0^{N+1}(t, \epsilon, Y_t^0, v^{0*}, \mathbf{v}^{F*}) - (Y_t^{0T} \hat{\mathbb{P}}_t^0 Y_t^0 + \hat{r}_t^0)] \\ =& \{ v^{0*T} R_0 v^{0*} - Y_t^{0T} [\![(K_1^0, 0, K_2^0)]\!]_{R_0}^2 Y_t^0 \} \\ & + 2 Y_t^{0T} [M_{31}^{0,\epsilon} v^{0*} + M_{32}^{0,\epsilon} v^{*(N)} - \hat{\mathbb{P}}_t^0 (\boldsymbol{B}_0 \hat{u}_t^0 + \boldsymbol{B}_0^a \hat{u}_t^{(N)})] \\ & + Y_t^{0T} \Big\{ -\frac{d}{dt}\hat{\mathbb{P}}_t^0 + (\hat{\mathbb{P}}_{t+\epsilon}^0 - \hat{\mathbb{P}}_t^0)/\epsilon + \mathcal{O}(\epsilon) \Big\} Y_t^0 \\ & + [2 v^{0*T} M_{12}^{0,\epsilon} v^{*(N)} + v^{*(N)T} M_{22}^{0,\epsilon} v^{*(N)} - 2 Y_t^{0T} \hat{\mathbb{P}}_t^0 \boldsymbol{B}_0^c K_1(\hat{\overline{X}}_t - \hat{X}_t^{(N)})] \\ & + \mathrm{Tr}\{ [\hat{\mathbb{P}}_t^0 (\boldsymbol{D}_0 \boldsymbol{D}_0^T + \boldsymbol{D}_0^a \boldsymbol{D}_0^{aT}/N)] + \mathcal{O}(\epsilon) \} + \frac{1}{\epsilon}(\hat{r}_{t+\epsilon}^0 - \hat{r}_t^0) \\ =:& \sum_{k=1}^5 \varDelta_{0k}^{*,\epsilon}, \end{aligned} \tag{D.11}$$

where each term $\varDelta_{0k}^{*,\epsilon}$ stands for one line. Now we have

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon}[J_0^{N+1}(t, Y_t^0, \hat{u}^0, \hat{\mathbf{u}}^F) - J_0^{N+1}(t, Y_t^0, \hat{u}^{0,\epsilon*}, \mathbf{u}^{F,\epsilon*})] \\ & \quad = \lim_{\epsilon \downarrow 0} \mathbb{E} \varDelta_0^{*,\epsilon}. \end{aligned} \tag{D.12}$$

Comparing (59)–(60) and (105)–(106), we have

$$v^{0*} - \hat{u}_t^0 = R_0^{-1} \widetilde{\boldsymbol{B}}_0^T \hat{\mathbb{P}}_t^0 [0, (\hat{\overline{X}}_t - \hat{X}_t^{(N)})^T, 0]^T + \mathcal{O}(\epsilon) Y_t^0, \tag{D.13}$$

and

$$\begin{aligned} v^{i*} - \hat{u}_t^i =& R^{-1} \boldsymbol{B}^T \hat{\mathbb{P}}_t [0, 0, (\hat{\overline{X}}_t - \hat{X}_t^{(N)})^T, 0]^T \\ & - R^{-1} R_2 R_0^{-1} \widetilde{\boldsymbol{B}}_0^T \hat{\mathbb{P}}_t^0 [0, (\hat{\overline{X}}_t - \hat{X}_t^{(N)})^T, 0]^T \\ & + R^{-1} (\boldsymbol{B} - \boldsymbol{B}_N^a)^T \hat{\mathbb{P}}_t Y_t^i + \mathcal{O}(\epsilon) Y_t^i, \end{aligned} \tag{D.14}$$

where $\boldsymbol{B} = [B^T, 0_{n_1\times 3n}]^T$. By elementary SDE estimates for (62), we find a constant C_0 independent of (N, t) such that

$$\mathbb{E}[|Y_t^0|^2 + |v^{0*}|^2 + |v^{*(N)}|^2] \le C_0. \tag{D.15}$$

There exists $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0]$, we have

$$|\varDelta_0^{*,\epsilon}| \le C_a + C_b|Y_t^0|^2, \tag{D.16}$$

for two constants C_a and C_b. Moreover, $\varDelta_0^{*,\epsilon}$ converges almost surely as $\epsilon \downarrow 0$. By (D.12), (D.15), (D.16), and dominated convergence, the limit in (107) exists.

By Lemma 4.1 and (D.13)–(D.15), we further obtain

$$\mathbb{E}|v^{0*} - \hat{u}_t^0|^2 \le C\mathbb{E}|X_0^{(N)} - m_0^X|^2 + C(\epsilon^2 + 1/N), \tag{D.17}$$

$$\mathbb{E}|v^{*(N)} - \hat{u}_t^{(N)}|^2 \le C\mathbb{E}|X_0^{(N)} - m_0^X|^2 + C(\epsilon^2 + 1/N), \tag{D.18}$$

where C does not depend on (t, N). Subsequently, in view of (D.8), we obtain

$$\mathbb{E}|\varDelta_{01}^{*,\epsilon}| \le C[\mathbb{E}|X_0^{(N)} - m_0^X|^2]^{1/2} + C/\sqrt{N} + C\epsilon. \tag{D.19}$$

By the representation of $M^{0,\epsilon}$ and (D.17), (D.18), and Lemma 4.1, we have

$$\mathbb{E}(|\varDelta_{02}^{*,\epsilon}| + |\varDelta_{04}^{*,\epsilon}|) \le C[\mathbb{E}|X_0^{(N)} - m_0^X|^2]^{1/2} + C/\sqrt{N} + C\epsilon. \tag{D.20}$$

Next, by use of the ODEs of $\hat{\mathbb{P}}^0$ and $\hat{r}^0$, we have

$$\mathbb{E}|\varDelta_{03}^{*,\epsilon}| + |\varDelta_{05}^{*,\epsilon}| \le C\epsilon. \tag{D.21}$$

Finally, (107) follows from (D.12), (D.19), (D.20) and (D.21). □

Proof of Lemma 5.9. Denote

$$\begin{aligned}&\breve{S}_i^{N+1}(t, \epsilon, Y_t^i, \nu^i)\\ &= \mathbb{E}\left[\int_t^{t+\epsilon} L(X_s^i, X_s^0, \mu_s^{(N)}, \nu^i, \hat{u}_s^0)ds + \hat{V}(t+\epsilon, Y_{t+\epsilon}^i)\middle|\hat{\boldsymbol{X}}_t^c\right].\end{aligned} \tag{D.22}$$

Then we have

$$J_i^{N+1}(t, Y_t^i, \hat{u}^{i,\epsilon}, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}) = \mathbb{E}[\breve{S}_i^{N+1}(t, \epsilon, Y_t^i, \nu^i)], \tag{D.23}$$

and by Lemma 4.2,

$$J_i^{N+1}(t, Y_t^i, \hat{u}^i, \hat{u}^0, \hat{\mathbf{u}}^{F,-i}) = \mathbb{E}(Y_t^{iT}\hat{\mathbb{P}}_t Y_t^i) + \hat{r}_t. \tag{D.24}$$

Let $\varPhi(\cdot,\cdot)$ be the fundamental solution matrix of the differential equation $\dot{z} = \boldsymbol{A}_1(t)z$. Then for $s \in [t, t+\epsilon]$, we have

$$\begin{aligned}Y_s^i =& \varPhi(s,t)Y_t^i + \int_t^s \varPhi(s,\tau)\boldsymbol{B}_N^b\nu^i d\tau\\ &+ \int_t^s \varPhi(s,\tau)[\boldsymbol{D}dW_\tau^0 + \boldsymbol{D}_N dW_\tau^i + \boldsymbol{D}^a dW_\tau^{(-i)}].\end{aligned} \tag{D.25}$$

By (D.25), we calculate

$$\mathbb{E}\left[\int_t^{t+\epsilon} L(X_s^i, X_s^0, \mu_s^{(N)}, \nu^i, \hat{u}_s^0)ds\middle|\hat{\boldsymbol{X}}_t^c\right] \tag{D.26}$$

$$
\begin{aligned}
&= \epsilon Y_t^{iT}\{[\![(I,-\varGamma_1,-\varGamma_2,0)]\!]_Q^2 + [\![(0,K_1^0,0,K_2^0)]\!]_{R_1}^2 + \mathcal{O}(\epsilon)\}Y_t^i \\
&\quad + 2\epsilon\nu^{iT}\{R_2[0,K_1^0,0,K_2^0] + \mathcal{O}(\epsilon)\}Y_t^i \\
&\quad + \epsilon\nu^{iT}[R+\mathcal{O}(\epsilon)]\nu^i + O(\epsilon^2).
\end{aligned}
$$

By Lemma 4.2,

$$
\hat{V}(t+\epsilon, Y_{t+\epsilon}^i) = Y_{t+\epsilon}^{iT}\hat{\mathbb{P}}_{t+\epsilon}Y_{t+\epsilon}^i + \hat{r}_{t+\epsilon}.
$$

We use (109) to calculate

$$
\begin{aligned}
&\mathbb{E}\left[Y_{t+\epsilon}^{iT}\hat{\mathbb{P}}_{t+\epsilon}Y_{t+\epsilon}^i \middle| \hat{\boldsymbol{X}}_t^c\right] \\
&= Y_t^{iT}\hat{\mathbb{P}}_tY_t^i + \epsilon Y_t^{iT}[\hat{\mathbb{P}}_t\boldsymbol{A}_1(t) + \boldsymbol{A}_1^T(t)\hat{\mathbb{P}}_t + \mathcal{O}(\epsilon)]Y_t^i \\
&\quad + 2\epsilon Y_t^{iT}[\hat{\mathbb{P}}_t\boldsymbol{B}_N^b + \mathcal{O}(\epsilon)]\nu^i + \nu^{iT}\mathcal{O}(\epsilon^2)\nu^i \\
&\quad + \epsilon\mathrm{Tr}\big\{\hat{\mathbb{P}}_t[\boldsymbol{D}\boldsymbol{D}^T + \boldsymbol{D}_N\boldsymbol{D}_N^T + ((N-1)/N^2)\boldsymbol{D}^a\boldsymbol{D}^{aT} + \mathcal{O}(\epsilon)]\big\}.
\end{aligned} \tag{D.27}
$$

Now by (D.26)–(D.27), we obtain

$$
\breve{S}_i^{N+1}(t,\epsilon,Y_t^i,\nu^i) = \epsilon|(\nu^{iT},Y_t^{iT})^T|_{\widehat{M}^\epsilon}^2 + \zeta_i^\epsilon(t),
$$

where the symmetric matrix $\widehat{M}^\epsilon = (\widehat{M}_{kl}^\epsilon)_{1\le k,l\le 2}$ satisfies

$$
\begin{aligned}
\widehat{M}_{11}^\epsilon =& R + \mathcal{O}(\epsilon), \\
\widehat{M}_{12}^\epsilon =& \boldsymbol{B}_N^{bT}\hat{\mathbb{P}}_t + R_2[0,K_1^0,0,K_2^0] + \mathcal{O}(\epsilon), \\
\widehat{M}_{21}^\epsilon =& (\widehat{M}_{12}^\epsilon)^T, \\
\widehat{M}_{22}^\epsilon =& \hat{\mathbb{P}}_t\boldsymbol{A}_1 + \boldsymbol{A}_1^T\hat{\mathbb{P}}_t + [\![(0,K_1^0,0,K_2^0)]\!]_{R_1}^2 \\
& + [\![(I,-\varGamma_1,-\varGamma_2,0)]\!]_Q^2 + \mathcal{O}(\epsilon).
\end{aligned}
$$

and

$$
\begin{aligned}
\zeta_i^\epsilon(t) =& \epsilon\mathrm{Tr}\{\hat{\mathbb{P}}_t[\boldsymbol{D}\boldsymbol{D}^T + \boldsymbol{D}_N\boldsymbol{D}_N^T + ((N-1)/N^2)\boldsymbol{D}^a\boldsymbol{D}^{aT} + \mathcal{O}(\epsilon)]\} \\
& + \hat{r}_{t+\epsilon} + Y_t^{iT}\hat{\mathbb{P}}_{t+\epsilon}Y_t^i.
\end{aligned}
$$

Since $R>0$, there exists $\bar{\epsilon}>0$ such that for all $0<\epsilon\le\bar{\epsilon}$, $R+\mathcal{O}(\epsilon)>0$ holds and the optimal $\nu^{i\star}$ is uniquely determined by the first order condition

$$
0 = \widehat{M}_{11}^\epsilon\nu^{i\star} + \widehat{M}_{12}^\epsilon Y_t^i,
$$

which implies (110). □

Proof of Lemma 5.11. By (D.23)–(D.24), we have

$$
\begin{aligned}
&J_i^{N+1}(t,Y_t^i,\hat{u}^{i,\epsilon\star},\hat{u}^0,\hat{\mathbf{u}}^{F,-i}) - J_i^{N+1}(t,Y_t^i,\hat{u}^i,\hat{u}^0,\hat{\mathbf{u}}^{F,-i}) \\
&\quad = \mathbb{E}\breve{S}_i^{N+1}(t,\epsilon,Y_t^i,\nu^{i\star}) - \mathbb{E}(Y_t^{iT}\hat{\mathbb{P}}_tY_t^i + \hat{r}_t).
\end{aligned} \tag{D.28}
$$

For the $\nu^{i\star}$ dependent terms in $\breve{S}_i^{N+1}$, by Lemma 5.9 we obtain

$$
|\nu^{i\star}|_R^2 + 2\nu^{i\star T}(R_2[0,K_1^0,0,K_2^0] + \boldsymbol{B}_N^{bT}\hat{\mathbb{P}}_t)Y_t^i = -\xi_R + Y_t^{iT}\mathcal{O}(\epsilon)Y_t^i, \tag{D.29}
$$

where

$$
\xi_R = |(R_2[0,K_1^0,0,K_2^0] + \boldsymbol{B}_N^{bT}\hat{\mathbb{P}}_t)Y_t^i|_{R^{-1}}^2. \tag{D.30}
$$

We use (D.29) to calculate

$$\begin{aligned}
\Delta^\epsilon :=& \frac{1}{\epsilon}[\breve{S}_i^{N+1}(t,\epsilon,Y_t^i,\nu^{i\star}) - (Y_t^{iT}\hat{\mathbb{P}}_t Y_t^i + \hat{r}_t)]\\
=& \frac{1}{\epsilon} Y_t^{iT}(\hat{\mathbb{P}}_{t+\epsilon} - \hat{\mathbb{P}}_t)Y_t^i + \frac{1}{\epsilon}(\hat{r}_{t+\epsilon} - \hat{r}_t)\\
&+ \mathrm{Tr}\{\hat{\mathbb{P}}_t[\boldsymbol{D}\boldsymbol{D}^T + \boldsymbol{D}_N\boldsymbol{D}_N^T + ((N-1)/N^2)\boldsymbol{D}^a\boldsymbol{D}^{aT} + \mathcal{O}(\epsilon)]\}\\
&+ Y_t^{iT}\{[\![(I,-\varGamma_1,-\varGamma_2,0)]\!]_Q^2 + [\![(I,K_1^0,K_2^0,0)]\!]_{R_1}^2\}Y_t^i\\
&+ Y_t^{iT}[\hat{\mathbb{P}}_t\boldsymbol{A}_1(t) + \boldsymbol{A}_1^T(t)\hat{\mathbb{P}}_t + \mathcal{O}(\epsilon)]Y_t^i\\
&- \xi_R.
\end{aligned}$$

Our method below is to rewrite $\boldsymbol{B}_N^{bT}\hat{\mathbb{P}}_t$ within ξ_R so that it can be related to the ODE of $\hat{\mathbb{P}}$. Let $\boldsymbol{B}$ be specified as in (D.14). We have

$$\boldsymbol{B}_N^{bT}\hat{\mathbb{P}}_t Y_t^i = \boldsymbol{B}^T\hat{\mathbb{P}}_t Y_t^i + (1/N)[0,0,B^T,0]\hat{\mathbb{P}}_t Y_t^i. \tag{D.31}$$

Denote $Y_t^o = [\hat{X}_t^{iT}, \hat{X}_t^{0T}, \hat{\overline{X}}_t^T]^T$. By (36) and Lemma 4.3, we have

$$\begin{aligned}
\boldsymbol{B}^T\hat{\mathbb{P}}_t Y_t^i =& \mathbb{B}^T\mathbb{J}_3^T\hat{\mathbb{P}}_t(\mathbb{J}_3 Y_t^o + Y_t^i - \mathbb{J}_3 Y_t^o)\\
=& R(-[K_1,K_2,K_3] - R^{-1}R_2[0,K_1^0,K_2^0])Y_t^o\\
&+ \mathbb{B}^T\mathbb{J}_3^T\hat{\mathbb{P}}_t[0,0,\hat{X}_t^{(N)T} - \hat{\overline{X}}_t^T,0]^T\\
=& -(R[K_1,K_2,0,K_3] + R_2[0,K_1^0,0,K_2^0])Y_t^i\\
&+ \mathbb{B}^T\mathbb{J}_3^T\hat{\mathbb{P}}_t[0,0,\hat{X}_t^{(N)T} - \hat{\overline{X}}_t^T,0]^T.
\end{aligned} \tag{D.32}$$

Therefore, it follows that

$$\begin{aligned}
\boldsymbol{B}_N^{bT}\hat{\mathbb{P}}_t Y_t^i =& (-R[K_1,K_2,0,K_3] - R_2[0,K_1^0,0,K_2^0])Y_t^i\\
&+ \mathbb{B}^T\mathbb{J}_3^T\hat{\mathbb{P}}_t[0,0,\hat{X}_t^{(N)T} - \hat{\overline{X}}_t^T,0]^T\\
&+ (1/N)[0,0,B^T,0]\hat{\mathbb{P}}_t Y_t^i.
\end{aligned}$$

We further use (D.31) to obtain

$$\begin{aligned}
\xi_R =& Y_t^{iT}[\![(K_1,K_2,0,K_3)]\!]_R^2 Y_t^i + [0,0,\hat{X}_t^{(N)T} - \hat{\overline{X}}_t^T,0]\mathcal{O}(1)Y_t^i\\
&+ (1/N)Y_t^{iT}\mathcal{O}(1)Y_t^i.
\end{aligned} \tag{D.33}$$

For $\widehat{\mathbb{A}}$ in (61), we rewrite

$$\widehat{\mathbb{A}} = \begin{bmatrix} A+BK_1 & G+B_1K_1^0+BK_2 & F & B_1K_2^0+BK_3\\ 0 & A_0+B_0K_1^0 & F_0 & B_0K_2^0\\ 0 & G+B_1K_1^0+BK_2 & A+F+BK_1 & B_1K_2^0+BK_3\\ 0 & G+B_1K_1^0+BK_2 & 0 & A+F+B_1K_2^0+B(K_1+K_3)\end{bmatrix}.$$

By the form of $\boldsymbol{A}_1$, we have

$$\begin{aligned}
Y_t^{iT}\hat{\mathbb{P}}_t\boldsymbol{A}_1(t)Y_t^i =& Y_t^{iT}\hat{\mathbb{P}}_t\widehat{\mathbb{A}}(t)Y_t^i - Y_t^{iT}\hat{\mathbb{P}}_t\boldsymbol{B}[K_1,K_2,0,K_3]Y_t^i\\
&+ Y_t^{iT}\mathcal{O}(1/N)Y_t^i.
\end{aligned} \tag{D.34}$$

By (D.32), we have

$$-Y_t^{iT}\hat{\mathbb{P}}_t\boldsymbol{B}[K_1,K_2,0,K_3]Y_t^i \tag{D.35}$$

$$= Y_t^i(R[K_1,K_2,0,K_3]+R_2[0,K_1^0,0,K_2^0])^T[K_1,K_2,0,K_3]Y_t^i$$
$$+ [0,0,\hat{X}_t^{(N)T}-\hat{\overline{X}}_t^T,0]\mathcal{O}(1)Y_t^i.$$

Now by (D.33)–(D.34) and (D.35), we have

$$\begin{aligned}
\Delta^\epsilon &= \frac{1}{\epsilon}(\hat{r}_{t+\epsilon}-\hat{r}_t)\\
&+ \mathrm{Tr}\{\hat{\mathbb{P}}_t[\boldsymbol{D}\boldsymbol{D}^T+\boldsymbol{D}_N\boldsymbol{D}_N^T+((N-1)/N^2)\boldsymbol{D}^a\boldsymbol{D}^{aT}+\mathcal{O}(\epsilon)]\}\\
&+ Y_t^{iT}\Big\{\frac{1}{\epsilon}(\hat{\mathbb{P}}_{t+\epsilon}-\hat{\mathbb{P}}_t)+\hat{\mathbb{P}}_t\widehat{\mathbb{A}}(t)+\widehat{\mathbb{A}}^T(t)\hat{\mathbb{P}}_t\\
&+ [K_1,K_2,0,K_3]^T(R[K_1,K_2,0,K_3]+R_2[0,K_1^0,0,K_2^0])\\
&+ (R[K_1,K_2,0,K_3]+R_2[0,K_1^0,0,K_2^0])^T[K_1,K_2,0,K_3]\\
&+ [\![(I,-\Gamma_1,-\Gamma_2,0)]\!]_Q^2+[\![(I,K_1^0,K_2^0,0)]\!]_{R_1}^2\\
&- [\![(K_1,K_2,0,K_3)]\!]_R^2\Big\}Y_t^i\\
&+ [0,0,\hat{X}_t^{(N)T}-\hat{\overline{X}}_t^T,0]\mathcal{O}(1)Y_t^i\\
&+ (1/N)Y_t^{iT}\mathcal{O}(1)Y_t^i.
\end{aligned}$$

We check Δ^ϵ and use bounded convergence as in the proof of Lemma 5.8 to show the existence of the limit of the lemma. By the ODEs of $\hat{\mathbb{P}}$ and $\hat{r}$, we obtain for a fixed constant C such that

$$\begin{aligned}
\mathbb{E}|\Delta^\epsilon| &\le C(\epsilon+1/N)(1+\mathbb{E}|Y_t^i|^2)\\
&\quad + C(\mathbb{E}|\hat{X}_t^{(N)}-\hat{\overline{X}}_t|^2)^{1/2}(\mathbb{E}|Y_t^i|^2)^{1/2}.
\end{aligned} \tag{D.36}$$

From (D.28), (D.36) and Lemma 4.1, we conclude

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}[J_i^{N+1}(t,Y_t^i,\hat{u}^i,\hat{u}^0,\hat{\mathbf{u}}^{F,-i})-J_i^{N+1}(t,Y_t^i,\hat{u}^{i,\epsilon\star},\hat{u}^0,\hat{\mathbf{u}}^{F,-i})]\\
&\quad = O((\mathbb{E}|X_0^{(N)}-m_0^X|^2)^{1/2}+1/\sqrt{N}).
\end{aligned}$$

This completes the proof. □

References

[1] M. Allais (1947), Economie et Interet, Paris: Imprimerie National.

[2] A. Aurell, R. Carmona, G. Dayanikli, and M. Lauriere (2022), Optimal incentives to mitigate epidemics: a Stackelberg mean field game approach, SIAM J. Control Optim., 60(2):S294–S322.

[3] T. Başar and G. J. Olsder (1999), Dynamic Noncooperative Game Theory, 2nd ed. Philadelphia : SIAM

[4] E. Bayraktar and A. Cohen (2018), Analysis of a finite state many player game using its master equation, SIAM J. Control Optim., 56:3538–3568.

[5] A. Bensoussan, M. Chau, Y. Lai and P. Yam (2017), Linear-quadratic mean field Stackelberg games with state and control delays. SIAM J. Control Optim., 55(4):2748–2781.

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

[7] A. Bensoussan, J. Frehse and P. Yam (2013). Mean Field Games and Mean Field Type Control Theory, New York: Springer.

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

[9] T. Björk and A. Murgoci (2008), A general theory of Markovian time inconsistent stochastic control problems, preprint.

[10] R. Buckdahn, J. Li, and S. Peng (2014), Nonlinear stochastic differential games involving a major player and a large number of collectively acting minor agents. SIAM Journal on Control and Optimization, 52(1):451–492.

[11] P. E. Caines, M. Huang, and R. P. Malhamé (2017), Mean field games, in Handbook of Dynamic Game Theory, T. Başar and G. Zaccour, Eds., pp. 345–372, Berlin: Springer.

[12] P. E. Caines and A. C. Kizilkale (2017), ϵ-Nash equilibria for partially observed LQG mean field games with a major player. IEEE Trans. Autom. Control, 62(7):3225–3234.

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

[14] P. Cardaliaguet, M. Cirant and A. Porretta (2020), Remarks on Nash equilibria in mean field game models with a major player, Proc. Amer. Math. Soc.,148(10):4241–4255.

[15] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions (2019), The master equation and the convergence problem in mean field games, Princeton Univ. Press.

[16] P. Cardaliaguet and C.-A. Lehalle (2018), Mean field game of controls and an application to trade crowding, Math. Financ. Econ., 12(3):335–363.

[17] R. Carmona and F. Delarue (2018), Probabilistic Theory of Mean Field Games with Applications I–II. Cham: Springer.

[18] R. Carmona and P. Wang (2017), An alternative approach to mean field game with major and minor players, and applications to herders impacts. Applied Math. Optim., 76(1):5–27.

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

[20] C.I. Chen and J. B. Cruz, Jr. (1972), Stackelberg solution for two-person games with biased information patterns, IEEE Trans. Autom. Control, AC–17(6):791–798.

[21] F. Delarue, D. Lacker, and K. Ramanan (2019), From the master equation to mean field game limit theory: a central limit theorem, Electron. J. Probab., 24:1–54, 2019, https://doi.org/10.1214/19-EJP298

[22] B. Djehiche and M. Huang (2016), A characterization of sub-game perfect equilibria for SDEs of mean-field type, Dyn. Games Appl., 6:55–81.

[23] I. Ekeland and A. Lazrak (2006), Being serious about non-commitment: subgame perfect equilibrium in continuous time, arXiv:math/0604264.

[24] I. Ekeland, O. Mbodji, and T. A. Pirvu (2012), Time-consistent portfolio management, SIAM J. Financial Math., 3(1):1–32.

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

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

[27] X. Feng, Y. Hu, and J. Huang (2023), A unified approach to linear-quadratic-Gaussian mean-field team: homogeneity, heterogeneity and quasi-exchangeability, Ann. Appl. Probab., published online.

[28] D. Firoozi and P. E. Caines (2021), ϵ-Nash equilibria for major–minor LQG mean field games with partial observations of all agents, IEEE Trans. Autom. Control, 66(6):2778–2786.

[29] G. Fu and U. Horst (2020), Mean-field leader-follower games with terminal state constraint, SIAM J. Control Optim., 58(4):2078–2113.

[30] D. A. Gomes, J. Mohr, and R. R. Souza (2013), Continuous time finite state mean field games, Appl. Math. Optim. 68(1):99–143.

[31] X. Guo, A. R. Hu, R. Y. Xu, and J. Z. Zhang (2019). Learning mean-field games, Proc. NeurIPS.

[32] Y. Hu, H. Jin, and X. Y. Zhou (2012), Time-inconsistent stochastic linear-quadratic control, SIAM J. Control Optim., 50(3):1548–1572.

[33] J. Huang, K. Si, and Z. Wu (2021), Linear-quadratic mixed Stackelberg–Nash stochastic differential game with major-minor agents, Applied Math. Optim., 84:2445–2494.

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

[35] M. Huang, P.E. Caines, and R.P. Malhamé (2012), Social optima in mean field LQG control: centralized and decentralized strategies, IEEE Trans. Autom. Control, 57(7):1736–1751.

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

[37] M. Huang and X. Yang (2020), Mean field Stackelberg games: state feedback equilibrium. Proc. 21st IFAC World Congress (virtual), Berlin, pp. 2237–2242.

[38] X. Huang, S. Jaimungal and M. Nourian (2019), Mean-field game strategies for optimal execution. Applied Mathematicl Finance, 26(2):153–185.

[39] V. Kolokoltsov (2017), The evolutionary game of pressure (or interference), resistance and collaboration. Math. Oper. Res., 42(4):915–944.

[40] I. Kordonis and G. P. Papavassilopoulos (2015), LQ Nash games with random entrance: an infinite horizon major player and minor players of finite horizons. IEEE Trans. Automat. Control, 60(6):1486–1500.

[41] F. Kydland (1977), Equilibrium solutions in dynamic dominant-player models, J. Economic Theory, 15:307–324.

[42] F.E. Kydland and E. C. Prescott (1977), Rules rather than discretion: the inconsistency of optimal plans. Journal of Political Economy, 85(3):473–492.

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

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

[45] M. Lauriere, S. Perrin, M. Geist, and O. Pietquin (2022), Learning mean field games: A survey, arXiv:2205.12944v2.

[46] M. Li, C. Mou, Z. Wu, and C. Zhou (2023), Linear-quadratic mean field games of controls with non-monotone data, Trans. Amer. Math. Soc., avail. online.

[47] Y. Lin, X. Jiang, and W. Zhang (2019), Open-loop Stackelberg strategy for the linear quadratic mean-field stochastic differential game. IEEE Trans. Autom. Control, 64(1):97–110.

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

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

[50] M. Nourian and P. E. Caines (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.

[51] R. A. Pollak (1968), Consistent planning, Rev. Economic Studies, 35(2):201–208.

[52] M. A. Rami, J.B. Moore, and X.Y. Zhou (2001), Indefinite stochastic linear quadratic control and generalized differential Riccati equation, SIAM J. Control Optim., 40(4):1296–1311.

[53] N. Saldi, T. Basar, and M. Raginsky (2018), Markov–Nash equilibria in mean-field games with discounted cost, SIAM Journal on Control and Optimization 56(6):4256–4287.

[54] N. Sen and P.E. Caines (2016), Mean field game theory with a partially observed major agent. SIAM J. Control Optim., 54(6):3174–3224.

[55] M. Simaan and J.B. Cruz, Jr. (1973), Additional aspects of the Stackelberg strategy in nonzero-sum games, Journal of Optimization Theory and Applications, 11:613–626.

[56] H. von Stackelberg (1934), Marktform und Gleichgewicht. Wien New York: Springer-Verlag.

[57] R. H. Strotz (1955), Myopia and inconsistency in dynamic utility maximization, The Review of Economic Studies, 23(3):165–180.

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

[59] X. Yang and M. Huang (2021), Linear quadratic mean field Stackelberg games: Master equations and time consistent feedback strategies. Proc. the 60th IEEE CDC, Austin, TX, pp. 171–176.

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

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