J Syst Sci Complex (2021) 34: 2003–2035

Linear Quadratic Mean Field Games: Decentralized $O(1/N)$-Nash Equilibria*

HUANG Minyi · YANG Xuwei

DOI: 10.1007/s11424-021-1266-y
Received: 20 July 2021

Abstract This paper studies an asymptotic solvability problem for linear quadratic (LQ) mean field games with controlled diffusions and indefinite weights for the state and control in the costs. The authors employ a rescaling approach to derive a low dimensional Riccati ordinary differential equation (ODE) system, which characterizes a necessary and sufficient condition for asymptotic solvability. The rescaling technique is further used for performance estimates, establishing an $O(1/N)$-Nash equilibrium for the obtained decentralized strategies.

Keywords Asymptotic solvability, decentralized strategies, ε-Nash equilibria, linear quadratic, mean field games, Riccati equations.

1 Introduction

Since its inception[1, 2], mean field game theory has undergone a phenomenal growth and found applications in diverse areas[3–16]. The theory is inspired by ideas in statistical physics and overcomes the dimensionality difficulty in competitive decision problems involving a large population of agents. The reader is referred to [17–20] for an overview of basic theory and applications.

While mean field games have been developed with very different modelling frameworks, linear quadratic (LQ) mean field games are of particular importance and have been extensively studied due to their elegant closed-form solutions[13, 21–23]. Huang, et al.[23] adopt infinite horizon discounted costs and use the infinite population limit model to design decentralized strategies for the actual model with a large but finite population. Li and Zhang[24] study decentralized strategies with ergodic costs. Wang and Zhang[25] introduce Markov jumps in the system dynamics and costs. Bardi and Priuli[26] study LQ N-person games and their mean field limit with ergodic costs. Huang, et al.[22] adopt backward stochastic differential

HUANG Minyi · YANG Xuwei
School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada.
Email: mhuang@math.carleton.ca; xuweiyang@cunet.carleton.ca.
*This research was supported by Natural Sciences and Engineering Research Council (NSERC) of Canada.
◇*This paper was recommended for publication by Guest Editor ZHANG Ji-Feng.*

equations for modelling state processes. Moon and Başar[27] consider risk sensitive costs and address robustness. Huang and Huang[28] consider linear diffusion dynamics including model uncertainty treated as an adversarial player. Tchuendom[29] shows nonuniqueness can arise, but interestingly, uniqueness can be restored by the presence of common noise. LQ mean field games have an extension by including a major player[30, 31]. This modelling framework is introduced by Huang[30]. Bensoussan, et al.[32] consider Stackelberg equilibria under state and control delays. Caines and Kizikale[33] consider partial information and filtering based strategies for an LQ model with a major player.

In this paper we study a class of LQ mean field games with common noise and indefinite weight matrices (simply called weights below) in the cost functional. We adopt the so-called asymptotic solvability framework in [34]. Starting with feedback perfect state information, this approach aims to determine feedback Nash strategies under such centralized information and next study how the solutions behave when the number of players increases. It uses a rescaling method to derive a set of Riccati ordinary differential equations (ODEs), which characterizes a necessary and sufficient condition for asymptotic solvability[34]. This method can be extended to LQ mean field games with a major player[35]. Recently, Huang and Yang[36] extend this asymptotic solvability notion to mean field social optimization, where the agents cooperatively optimize a social cost. That work further develops a method of asymptotic analysis to obtain tight estimates of optimality loss when decentralized strategies are implemented. For our current model, the test of asymptotic solvability reduces to checking two Riccati ODEs in a low dimensional space, which, as a result of the controlled individual and common noises, have higher nonlinearity than those Riccati equations in [34].

In the analysis of mean field games, a crucial step is to examine how the strategies obtained in the mean field limit model perform when implemented in the actual model with a large but finite population. This can be addressed by establishing the so-called ε-Nash equilibrium property, where $\varepsilon \to 0$ as $N \to \infty$. For LQ models[21–23, 25] as well as some nonlinear cases[37], one can obtain an $O(1/\sqrt{N})$-Nash equilibrium when all players are symmetric. This typically results from cost estimates by the Cauchy-Schwarz inequality. To our best knowledge on the existing literature, probably only Basna, et al.[38] have obtained an $O(1/N)$-Nash equilibrium result in a finite state mean field game. We will establish an $O(1/N)$-Nash equilibrium for the decentralized strategies obtained from the LQ mean field limit model; our approach is different from that in [38] which relies on perturbation estimates of generators of continuous-time controlled Markov chains. We will directly treat the best response control problem of the unilateral agent in a high dimensional space and then employ the rescaling method to obtain accurate information about its performance improvement. We will develop extensive asymptotic error estimates by building upon techniques in the companion paper[36] on social optimization. In a convergence problem of mean field games with common noise, Cardaliaguet, et al.[39] prove that the value functions of N players converge in an average sense to the solution of the master equation, and the averaged error disappears by rate $1/N$ as $N \to \infty$. But their error bound is different from the $O(1/N)$-Nash equilibrium notion.

It will be helpful to briefly explain the route that we will follow in the analysis. For the LQ

Nash game with indefinite weights, we apply dynamic programming to derive a set of large-scale Riccati equations, which is used to formulate the asymptotic solvability problem of the N-player game. In order to get useful information from the large Riccati equations, we exploit their symmetries to achieve dimension reduction and next use a rescaling technique to derive two key Riccati equations, which completely characterize asymptotic solvability. By taking the mean field limit of the solution of the N-player game, we construct a set of decentralized strategies, which are then applied to the N-player model. We further obtain explicit formulas for the per agent cost for three scenarios: i) The N players apply the Nash equilibrium strategies $(\widehat{u}_1, \widehat{u}_2, \cdots, \widehat{u}_N)$; ii) The N players apply decentralized strategies $(\check{u}_1, \check{u}_2, \cdots, \check{u}_N)$ obtained from the mean field limit model; iii) The player in question takes its best response while the other $N-1$ players apply these decentralized strategies. When $N \to \infty$, the three cases have the same limit for the per agent cost. The comparison of the costs in scenarios ii) and iii) establishes the $O(1/N)$-Nash equilibrium property. A comparison of the per agent costs for the mean field game and the mean field social optimum enables us to quantify the efficiency loss of the mean field game with respect to the social optimum; see the comparison in the companion paper [36].

1.1 Organization of the Paper

Section 2 introduces the N-player LQ Nash game with indefinite weight matrices in the cost functionals. The set of feedback Nash equilibrium strategies is characterized using a system of Riccati ODEs in Section 3. The asymptotic solvability problem is studied in Section 4 and a necessary and sufficient condition is derived. Section 5 constructs a set of decentralized strategies for the N-player game, and Section 6 proves an $O(1/N)$-Nash equilibrium theorem. A numerical example is presented in Section 7. Section 8 concludes the paper.

1.2 Notation

Let $\mathcal{S}^n$ be the set of $n \times n$ real symmetric matrices. We denote the quadratic form $[x]_M^2 = x^{\mathrm{T}} M x$ for $M \in \mathcal{S}^n$ and $x \in \mathbb{R}^n$. We use I to denote an identity matrix of compatible dimensions, and sometimes write I_k to indicate the $k \times k$ identity matrix. We use 0 to denote either the scalar zero or a zero vector/matrix of compatible dimensions.

We denote by $|F|$ the Euclidean norm of a vector or matrix F, by $\mathbf{1}_{k\times l}$ a $k \times l$ matrix with all entries equal to 1, by $\otimes$ the Kronecker product, and by the column vectors $\{e_1^k, e_2^k, \cdots, e_k^k\}$ the canonical basis of $\mathbb{R}^k$. For a function $f(t,x)$, we may write partial derivatives $\partial f/\partial t$ as $\partial_t f$; $\partial f/\partial x$ as $\partial_x f$; and $\partial^2 f/\partial x^2$ as $\partial_x^2 f$.

2 The LQ Nash Game

Consider a system of N players (or called agents) denoted by $\mathcal{A}_i$, $1 \le i \le N$. The state process $X_i(t)$ satisfies the following stochastic differential equation (SDE)

$$\begin{aligned} dX_i(t) = {} & (AX_i(t) + Bu_i(t) + GX^{(N)}(t))dt + (B_1 u_i(t) + D)dW_i(t) \\ & + (B_0 u^{(N)}(t) + D_0)dW_0(t), \end{aligned} \tag{1}$$

where we have the state $X_i(t) \in \mathbb{R}^n$, the control $u_i(t) \in \mathbb{R}^{n_1}$, the mean field state $X^{(N)} := (1/N)\sum_{i=1}^N X_i$ and the control mean field $u^{(N)} := (1/N)\sum_{i=1}^N u_i$. The initial states $\{X_i(0) : 1 \le i \le N\}$ are independent with $\mathbb{E}|X_i(0)|^2 < \infty$. The individual noise processes $\{W_i : 1 \le i \le N\}$ are 1-dimensional independent standard Brownian motions, which are also independent of $\{X_i(0) : 1 \le i \le N\}$. The common noise W_0 is a 1-dimensional standard Brownian motion independent of $\{W_i : 1 \le i \le N\}$ and $\{X_i(0) : 1 \le i \le N\}$. In contrast to [34, 40], each individual noise is affected by that player's control, and the model contains a common noise affected by the control mean field.

The individual cost functional (simply called cost) of $\mathcal{A}_i$, $1 \le i \le N$, is given by

$$\begin{aligned} &J_i(u_1, u_2, \cdots, u_N) \\ =&\mathbb{E}\left[\int_0^T \Big([X_i(t) - \Gamma X^{(N)}(t)]_Q^2 + [u_i(t)]_R^2\Big)dt + [X_i(T) - \Gamma_f X^{(N)}(T)]_{Q_f}^2\right], \end{aligned} \tag{2}$$

where we denote $[x]_M^2 = x^{\mathrm{T}}Mx$ for $M \in \mathcal{S}^n$ and $x \in \mathbb{R}^n$. The constant matrices A, B, B_0 B_1, D, D_0, G, Γ, Q, R, Γ_f, Q_f above have compatible dimensions, and Q, Q_f, R are symmetric, possibly indefinite, matrices.

Define

$$\begin{aligned} &X(t) = \begin{bmatrix} X_1(t) \\ \vdots \\ X_N(t) \end{bmatrix} \in \mathbb{R}^{Nn}, \quad u_{-i} = (u_1, u_2, \cdots, u_{i-1}, u_{i+1}, \cdots, u_N), \\ &\boldsymbol{A} = \mathrm{diag}[A, A, \cdots, A] + \mathbf{1}_{N\times N} \otimes \frac{G}{N} \in \mathbb{R}^{Nn\times Nn}, \\ &\boldsymbol{B}_0 = \mathbf{1}_{N\times 1} \otimes \frac{B_0}{N} \in \mathbb{R}^{Nn\times n_1}, \quad \boldsymbol{D}_0 = \mathbf{1}_{N\times 1} \otimes D_0 \in \mathbb{R}^{Nn\times 1}, \\ &\widehat{\boldsymbol{B}}_k = e_k^N \otimes B \in \mathbb{R}^{Nn\times n_1}, \quad \boldsymbol{B}_k = e_k^N \otimes B_1 \in \mathbb{R}^{Nn\times n_1}, \\ &\boldsymbol{D}_k = e_k^N \otimes D \in \mathbb{R}^{Nn\times 1}, \quad 1 \le k \le N. \end{aligned}$$

Then $X = (X_1^{\mathrm{T}}, X_2^{\mathrm{T}}, \cdots, X_N^{\mathrm{T}})^{\mathrm{T}}$ has the following dynamics

$$\begin{aligned} dX(t) =& \Big(\boldsymbol{A}X(t) + \sum_{i=1}^N \widehat{\boldsymbol{B}}_i u_i(t)\Big)dt + \sum_{i=1}^N (\boldsymbol{B}_i u_i(t) + \boldsymbol{D}_i)dW_i \\ &+ \Big(\boldsymbol{B}_0 \sum_{i=1}^N u_i(t) + \boldsymbol{D}_0\Big)dW_0. \end{aligned} \tag{3}$$

We denote

$$\begin{aligned} &\boldsymbol{K}_i = [0, \cdots, 0, I_n, 0, \cdots, 0] - (1/N)[\Gamma, \Gamma, \cdots, \Gamma] \in \mathbb{R}^{n\times Nn}, \\ &\boldsymbol{K}_{if} = [0, \cdots, 0, I_n, 0, \cdots, 0] - (1/N)[\Gamma_f, \Gamma_f, \cdots, \Gamma_f], \\ &\boldsymbol{Q}_i = \boldsymbol{K}_i^{\mathrm{T}} Q \boldsymbol{K}_i, \quad \boldsymbol{Q}_{if} = \boldsymbol{K}_{if}^{\mathrm{T}} Q_f \boldsymbol{K}_{if}. \end{aligned}$$

The individual cost (2) can be written as

$$J_i(u_i, u_{-i}) = \mathbb{E}\left[\int_0^T \left([X(t)]^2_{\boldsymbol{Q}_i} + [u_i(t)]^2_R\right)\, dt + [X(T)]^2_{\boldsymbol{Q}_{if}}\right]. \tag{4}$$

We begin by solving the LQ Nash game under closed-loop perfect state (CLPS) information, where the full state vector $X(t)$ is observed by each player. The players seek a set of Nash equilibrium strategies $(\widehat{u}_1, \widehat{u}_2, \cdots, \widehat{u}_N)$.

For notational simplicity, Sections 3–6 will treat a simplified model (1)–(2) with $D = D_0 = 0$. The extension to the general case will be discussed in Section 6.

3 Riccati Equations and Feedback Nash Strategies

Based on (4), we may naturally define the cost $J(t, \boldsymbol{x}, u_1, \cdots, u_N)$ where the running cost is integrated on $[t, T]$ instead of $[0, T]$ with initial state $X(t) = \boldsymbol{x} = (x_1^{\mathrm{T}}, x_2^{\mathrm{T}}, \cdots, x_N^{\mathrm{T}})^{\mathrm{T}}$. Let $V_i(t, \boldsymbol{x})$ denote the value function of player $\mathcal{A}_i$. The Hamilton-Jacobi-Bellman (HJB) equations of the N players associated with (3)–(4) (taking $D_0 = D = 0$) are

$$\begin{aligned} -\frac{\partial V_i}{\partial t} &= \frac{\partial^{\mathrm{T}} V_i}{\partial \boldsymbol{x}}\left(\boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^N \widehat{\boldsymbol{B}}_k \widehat{u}_k\right) + \frac{1}{2}\left(\sum_{k=1}^N \widehat{u}_k\right)^{\mathrm{T}} \boldsymbol{B}_0^{\mathrm{T}} \frac{\partial^2 V_i}{\partial \boldsymbol{x}^2} \boldsymbol{B}_0 \left(\sum_{k=1}^N \widehat{u}_k\right) \\ &\quad + \frac{1}{2}\sum_{k=1}^N (\boldsymbol{B}_k \widehat{u}_k)^{\mathrm{T}} \frac{\partial^2 V_i}{\partial \boldsymbol{x}^2} (\boldsymbol{B}_k \widehat{u}_k) + \boldsymbol{x}^{\mathrm{T}} \boldsymbol{Q}_i \boldsymbol{x} + \widehat{u}_i^{\mathrm{T}} R \widehat{u}_i, \\ V_i(T, \boldsymbol{x}) &= \boldsymbol{x}^{\mathrm{T}} \boldsymbol{Q}_{if} \boldsymbol{x}, \quad 1 \le i \le N. \end{aligned} \tag{5}$$

Each $\widehat{u}_i$ is the minimizer in the HJB equation of $V_i(t, \boldsymbol{x})$ as specified below. Taking $(u_1, u_2, \cdots, u_N)$ in place of $(\widehat{u}_1, \widehat{u}_2, \cdots, \widehat{u}_N)$, we write the right hand side of (5) in the form:

$$\mathcal{H}(\boldsymbol{x}, \partial_{\boldsymbol{x}} V_i, \partial^2_{\boldsymbol{x}} V_i, u_i, u_{-i}).$$

Then we require

$$\widehat{u}_i = \arg\min_{u_i} \mathcal{H}(\boldsymbol{x}, \partial_{\boldsymbol{x}} V_i, \partial^2_{\boldsymbol{x}} V_i, u_i, \widehat{u}_{-i}), \quad \forall i. \tag{6}$$

We will calculate $\widehat{u}_i$ under the following conditions: for all $(t, \boldsymbol{x}) \in [0, T] \times \mathbb{R}^{Nn}$,

$$R + \frac{1}{2}\boldsymbol{B}_i^{\mathrm{T}} \frac{\partial^2 V_i(t, \boldsymbol{x})}{\partial \boldsymbol{x}^2} \boldsymbol{B}_i > 0, \tag{7}$$

$$I + \frac{1}{2}\sum_{k=1}^N \left(R + \frac{1}{2}\boldsymbol{B}_k^{\mathrm{T}} \frac{\partial^2 V_k(t, \boldsymbol{x})}{\partial \boldsymbol{x}^2} \boldsymbol{B}_k\right)^{-1} \boldsymbol{B}_0^{\mathrm{T}} \frac{\partial^2 V_k(t, \boldsymbol{x})}{\partial \boldsymbol{x}^2} \boldsymbol{B}_0 \text{ is invertible}, \tag{8}$$

$$R + \frac{1}{2}\boldsymbol{B}_i^{\mathrm{T}} \frac{\partial^2 V_i(t, \boldsymbol{x})}{\partial \boldsymbol{x}^2} \boldsymbol{B}_i + \frac{1}{2}\boldsymbol{B}_0^{\mathrm{T}} \frac{\partial^2 V_i(t, \boldsymbol{x})}{\partial \boldsymbol{x}^2} \boldsymbol{B}_0 > 0. \tag{9}$$

By (6), we derive

$$0 = \widehat{\boldsymbol{B}}_i^{\mathrm{T}} \frac{\partial V_i}{\partial \boldsymbol{x}} + \boldsymbol{B}_0^{\mathrm{T}} \frac{\partial^2 V_i}{\partial \boldsymbol{x}^2} \boldsymbol{B}_0 \sum_{i \ne k=1}^N \widehat{u}_k + \boldsymbol{B}_0^{\mathrm{T}} \frac{\partial^2 V_i}{\partial \boldsymbol{x}^2} \boldsymbol{B}_0 \widehat{u}_i + \boldsymbol{B}_i^{\mathrm{T}} \frac{\partial^2 V_i}{\partial \boldsymbol{x}^2} \boldsymbol{B}_i \widehat{u}_i + 2R\widehat{u}_i, \tag{10}$$

which implies that

$$\widehat{u}_i = -\frac{1}{2}\left(R + \frac{1}{2}\boldsymbol{B}_i^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_i\right)^{-1}\left(\widehat{\boldsymbol{B}}_i^{\mathrm{T}}\frac{\partial V_i}{\partial \boldsymbol{x}} + \boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\sum_{k=1}^{N}\widehat{u}_k\right). \tag{11}$$

Adding up the N equations in (11) leads to

$$\sum_{i=1}^{N}\widehat{u}_i = -\frac{1}{2}\sum_{i=1}^{N}\left(R + \frac{1}{2}\boldsymbol{B}_i^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_i\right)^{-1}\left(\widehat{\boldsymbol{B}}_i^{\mathrm{T}}\frac{\partial V_i}{\partial \boldsymbol{x}} + \boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\sum_{k=1}^{N}\widehat{u}_k\right),$$

which under the condition (8) yields

$$\begin{aligned}\sum_{k=1}^{N}\widehat{u}_k =& -\frac{1}{2}\left[I + \frac{1}{2}\sum_{k=1}^{N}\left(R + \frac{1}{2}\boldsymbol{B}_k^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k\right)^{-1}\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\right]^{-1} \\ &\cdot\sum_{k=1}^{N}\left(R + \frac{1}{2}\boldsymbol{B}_k^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k\right)^{-1}\widehat{\boldsymbol{B}}_k^{\mathrm{T}}\frac{\partial V_k}{\partial \boldsymbol{x}} \\ =:& \boldsymbol{M}.\end{aligned} \tag{12}$$

Combining (11) and (12) gives that

$$\widehat{u}_i = -\frac{1}{2}\left(R + \frac{1}{2}\boldsymbol{B}_i^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_i\right)^{-1}\left(\widehat{\boldsymbol{B}}_i^{\mathrm{T}}\frac{\partial V_i}{\partial \boldsymbol{x}} + \boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M}\right). \tag{13}$$

We substitute (13) into the right hand side of (5) to obtain

$$\begin{aligned}-\frac{\partial V_i}{\partial t} =& \frac{1}{4}\left[\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M} - \widehat{\boldsymbol{B}}_i^{\mathrm{T}}\frac{\partial V_i}{\partial \boldsymbol{x}}\right]^{\mathrm{T}}\left(R + \frac{1}{2}\boldsymbol{B}_i^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_i\right)^{-1}\left[\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M} + \widehat{\boldsymbol{B}}_i^{\mathrm{T}}\frac{\partial V_i}{\partial \boldsymbol{x}}\right] \\ &- \frac{1}{2}\frac{\partial^{\mathrm{T}} V_i}{\partial \boldsymbol{x}}\sum_{i\neq k=1}^{N}\widehat{\boldsymbol{B}}_k\left(R + \frac{1}{2}\boldsymbol{B}_k^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k\right)^{-1}\left[\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M} + \widehat{\boldsymbol{B}}_k^{\mathrm{T}}\frac{\partial V_k}{\partial \boldsymbol{x}}\right] \\ &+ \frac{1}{8}\sum_{i\neq k=1}^{N}\left[\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M} + \widehat{\boldsymbol{B}}_k^{\mathrm{T}}\frac{\partial V_k}{\partial \boldsymbol{x}}\right]^{\mathrm{T}}\left(R + \frac{1}{2}\boldsymbol{B}_k^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k\right)^{-1} \\ &\cdot\boldsymbol{B}_k^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k\left(R + \frac{1}{2}\boldsymbol{B}_k^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k\right)^{-1}\left[\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_k}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M} + \widehat{\boldsymbol{B}}_k^{\mathrm{T}}\frac{\partial V_k}{\partial \boldsymbol{x}}\right] \\ &+ \frac{1}{2}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\boldsymbol{M} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_i\boldsymbol{x} + \frac{\partial^{\mathrm{T}} V_i}{\partial \boldsymbol{x}}\boldsymbol{A}\boldsymbol{x}, \qquad (14) \\ V_i(T,\boldsymbol{x}) =& \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_{if}\boldsymbol{x}, \quad 1 \le i \le N,\end{aligned}$$

subject to the conditions (7), (8), and (9).

We are interested in a solution of the form

$$V_i(t,\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{x}, \quad 1 \le i \le N, \tag{15}$$

where $\boldsymbol{P}_i(t)$ is a symmetric matrix function of $t \in [0,T]$ and is differentiable in t. Substituting (15) into (14), we obtain the ODE system for $\boldsymbol{P}_i$, $1 \le i \le N$:

$$\begin{cases} -\dot{\boldsymbol{P}}_i = \boldsymbol{P}_i\boldsymbol{A} + \boldsymbol{A}^{\rm T}\boldsymbol{P}_i + \boldsymbol{Q}_i - \boldsymbol{P}_i\widehat{\boldsymbol{B}}_i(R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i)^{-1}\widehat{\boldsymbol{B}}_i^{\rm T}\boldsymbol{P}_i \\ \qquad +\boldsymbol{M}_0^{\rm T}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_0(R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i)^{-1}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_0\boldsymbol{M}_0 \\ \qquad -\boldsymbol{P}_i\displaystyle\sum_{i\neq k=1}^{N}\widehat{\boldsymbol{B}}_k(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_k)^{-1}(\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_0\boldsymbol{M}_0+\widehat{\boldsymbol{B}}_k^{\rm T}\boldsymbol{P}_k) \\ \qquad -\displaystyle\sum_{i\neq k=1}^{N}(\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_0\boldsymbol{M}_0+\widehat{\boldsymbol{B}}_k^{\rm T}\boldsymbol{P}_k)^{\rm T}(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_k)^{-1}\widehat{\boldsymbol{B}}_k^{\rm T}\boldsymbol{P}_i \\ \qquad +\displaystyle\sum_{i\neq k=1}^{N}(\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_0\boldsymbol{M}_0+\widehat{\boldsymbol{B}}_k^{\rm T}\boldsymbol{P}_k)^{\rm T}(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_k)^{-1}\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_k \\ \qquad\quad \cdot(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_k)^{-1}(\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_0\boldsymbol{M}_0+\widehat{\boldsymbol{B}}_k^{\rm T}\boldsymbol{P}_k)+\boldsymbol{M}_0^{\rm T}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_0\boldsymbol{M}_0, \\ \boldsymbol{P}_i(T) = \boldsymbol{Q}_{if}, \end{cases} \tag{16}$$

subject to

$$\begin{cases} \text{(i)} \quad R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i(t)\boldsymbol{B}_i > 0, \quad \forall t\in[0,T], \\ \text{(ii)} \quad R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i(t)\boldsymbol{B}_i+\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i(t)\boldsymbol{B}_0 > 0, \quad \forall t\in[0,T], \\ \text{(iii)} \quad I+\sum_{k=1}^{N}\left(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k(t)\boldsymbol{B}_k\right)^{-1}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k(t)\boldsymbol{B}_0 \text{ is invertible}, \quad \forall t\in[0,T], \end{cases} \tag{17}$$

where

$$\boldsymbol{M}_0 := -\left[I+\sum_{k=1}^{N}\left(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_k\right)^{-1}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_0\right]^{-1}\sum_{k=1}^{N}\left(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k\boldsymbol{B}_k\right)^{-1}\widehat{\boldsymbol{B}}_k^{\rm T}\boldsymbol{P}_k.$$

In further analysis, if we just say $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ is a solution of (16), that means (17) is in effect unless otherwise indicated. Condition (17) (ii) is not used in the vector field of the Riccati equation, but will play a role in the best response control problem later.

Remark 3.1 If the ODE system (16) admits a solution $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ on $[0,T]$, then it is the unique solution since the vector field of the ODE system has a local Lipschitz property along the solution trajectory satisfying (17) (i) and (iii).

The following theorem gives a sufficient condition for the existence of feedback Nash strategies in terms of the Riccati equations (16). These strategies are called centralized due to the use of full state information by each player.

Theorem 3.2 *If* (16) *has a solution* $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ *on* $[0,T]$*, then the Nash game* (3)–(4) *has a set of feedback Nash strategies* $(\widehat{u}_1, \widehat{u}_2, \cdots, \widehat{u}_N)$ *given by*

$$\widehat{u}_i(t) = -[R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i(t)\boldsymbol{B}_i]^{-1}[\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i(t)\boldsymbol{B}_0\boldsymbol{M}_0(t)+\widehat{\boldsymbol{B}}_i^{\rm T}\boldsymbol{P}_i(t)]X(t), \quad 1\le i\le N. \tag{18}$$

Proof See Appendix 1. ∎

The best response control problem in the proof of Theorem 3.2 amounts to LQ optimal control with indefinite weights in the cost. The HJB equation (5) is only used for constructing

(16). The rigorous proof of $\widehat{u}_1$ as a best response strategy on $[t,T]$ given $(t,\boldsymbol{x},\widehat{u}_{-1})$ has been solely based on the Riccati equation system (16) itself.

4 Asymptotic Solvability

We start with a representation of the matrix $\boldsymbol{P}_i$ if the ODE system (16) has a solution. Write the $Nn\times Nn$ identity matrix I_{Nn} as $I_{Nn}=\mathrm{diag}[I_n,I_n,\cdots,I_n]$. Let J_{ij} denote the matrix obtained by exchanging the ith and jth rows of submatrices in I_{Nn}.

Lemma 4.1 *Suppose* (16) *has a solution* $(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N)$ *on* $[0,T]$. *Then* $\boldsymbol{P}_i$, $1\le i\le N$, *have the representation*

$$\boldsymbol{P}_1=\begin{bmatrix}\Pi_1^N & \Pi_2^N & \Pi_2^N & \cdots & \Pi_2^N\\ \Pi_2^{N\mathrm{T}} & \Pi_3^N & \Pi_4^N & \cdots & \Pi_4^N\\ \Pi_2^{N\mathrm{T}} & \Pi_4^N & \Pi_3^N & \cdots & \Pi_4^N\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \Pi_2^{N\mathrm{T}} & \Pi_4^N & \Pi_4^N & \cdots & \Pi_3^N\end{bmatrix},\quad \boldsymbol{P}_i=J_{1i}^{\mathrm{T}}\boldsymbol{P}_1J_{1i},\quad \forall 2\le i\le N, \tag{19}$$

where $\Pi_1^N(t)$, $\Pi_3^N(t)$, $\Pi_4^N(t)\in\mathcal{S}^n$, *and* $\Pi_2^N(t)\in\mathbb{R}^{n\times n}$.

Proof See Appendix 1. ∎

Following the route in [34], we introduce the notion of asymptotic solvability of the Nash game (1)–(2) (with $D=D_0=0$).

Definition 4.2 The Nash game (1)–(2) is asymptotically solvable if there exist $N_0>0$ and $c_0>0$ such that the ODE system (16)–(17) has a solution $(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N)$ on $[0,T]$ for all $N\ge N_0$, and that

$$\sup_{N\ge N_0}\sup_{0\le t\le T}(|\Pi_1^N|+N|\Pi_2^N|+N^2|\Pi_3^N|+N^2|\Pi_4^N|)<\infty, \tag{20}$$

$$R+\boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{P}_i\boldsymbol{B}_i\ge c_0I,\quad \forall t\in[0,T],\quad \forall N\ge N_0, \tag{21}$$

$$I+\sum_{k=1}^N\left(R+\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{P}_k\boldsymbol{B}_k\right)^{-1}\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_k\boldsymbol{B}_0 \text{ is invertible},\quad \forall t\in[0,T],\ \forall N\ge N_0. \tag{22}$$

Remark 4.3 The conditions (20)–(21) imply that

$$R+\boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_i+\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_0\ge(c_0/2)I,\quad \forall t,$$

as long as a sufficiently large N_0 is chosen.

Define the mapping $\mathcal{R}_1:\mathcal{S}^n\to\mathcal{S}^n$ by

$$\mathcal{R}_1(Z)=R+B_1^{\mathrm{T}}ZB_1,\quad \text{for } Z\in\mathcal{S}^n.$$

For $\Lambda_k \in \mathcal{S}^n$, $k = 1, 3, 4$, and $\Lambda_2 \in \mathbb{R}^{n\times n}$, we define the mappings:

$$\begin{aligned}
\Psi_1(\Lambda_1) &= \Lambda_1 BHB^{\mathrm{T}}\Lambda_1 - \Lambda_1 A - A^{\mathrm{T}}\Lambda_1 - Q,\\
\Psi_2(\Lambda_1,\Lambda_2) &= -\Lambda_1 G - \Lambda_2(A+G) - A^{\mathrm{T}}\Lambda_2 + \Lambda_1 BHB^{\mathrm{T}}\Lambda_2 + \Lambda_2 BHB^{\mathrm{T}}\Lambda_1\\
&\quad + \Lambda_2 BHB^{\mathrm{T}}\Lambda_2 + Q\Gamma,\\
\Psi_3(\Lambda_1,\Lambda_2,\Lambda_3,\Lambda_4) &= -\Lambda_3 A - A^{\mathrm{T}}\Lambda_3 - (\Lambda_2^{\mathrm{T}}+\Lambda_4)G - G^{\mathrm{T}}(\Lambda_2+\Lambda_4)\\
&\quad - (\Lambda_1+\Lambda_2^{\mathrm{T}})BHB_0^{\mathrm{T}}(\Lambda_1+\Lambda_2+\Lambda_2^{\mathrm{T}}+\Lambda_4)B_0HB^{\mathrm{T}}(\Lambda_1+\Lambda_2)\\
&\quad + \Lambda_3 BHB^{\mathrm{T}}\Lambda_1 + \Lambda_1 BHB^{\mathrm{T}}\Lambda_3 + \Lambda_4 BHB^{\mathrm{T}}\Lambda_2\\
&\quad + \Lambda_2^{\mathrm{T}}BHB^{\mathrm{T}}(\Lambda_2+\Lambda_4) - \Lambda_1 BHB_1^{\mathrm{T}}\Lambda_3 B_1 HB^{\mathrm{T}}\Lambda_1 - \Gamma^{\mathrm{T}}Q\Gamma,\\
\Psi_4(\Lambda_1,\Lambda_2,\Lambda_4) &= -\Lambda_4 A - A^{\mathrm{T}}\Lambda_4 - (\Lambda_2^{\mathrm{T}}+\Lambda_4)G - G^{\mathrm{T}}(\Lambda_2+\Lambda_4) - \Gamma^{\mathrm{T}}Q\Gamma\\
&\quad - (\Lambda_1+\Lambda_2^{\mathrm{T}})BHB_0^{\mathrm{T}}(\Lambda_1+\Lambda_2+\Lambda_2^{\mathrm{T}}+\Lambda_4)B_0HB^{\mathrm{T}}(\Lambda_1+\Lambda_2)\\
&\quad + \Lambda_4 BHB^{\mathrm{T}}(\Lambda_1+\Lambda_2) + (\Lambda_1+\Lambda_2^{\mathrm{T}})BHB^{\mathrm{T}}\Lambda_4 + \Lambda_2^{\mathrm{T}}BHB^{\mathrm{T}}\Lambda_2,
\end{aligned}$$

where we denote $H = (\mathcal{R}_1(\Lambda_1))^{-1}$ provided that the inverse matrix exists. It is clear that Ψ_k, $k = 1, 3, 4$, are $\mathcal{S}^n$-valued.

We introduce the following ODE system

$$\begin{cases}\dot\Lambda_1 = \Psi_1(\Lambda_1),\\ \Lambda_1(T) = Q_f, \quad \mathcal{R}_1(\Lambda_1(t)) > 0, \quad \forall t\in[0,T],\end{cases} \tag{23}$$

$$\dot\Lambda_2 = \Psi_2(\Lambda_1,\Lambda_2), \qquad \Lambda_2(T) = -Q_f\Gamma_f, \tag{24}$$

$$\dot\Lambda_3 = \Psi_3(\Lambda_1,\Lambda_2,\Lambda_3,\Lambda_4), \qquad \Lambda_3(T) = \Gamma_f^{\mathrm{T}}Q_f\Gamma_f, \tag{25}$$

$$\dot\Lambda_4 = \Psi_4(\Lambda_1,\Lambda_2,\Lambda_4), \qquad \Lambda_4(T) = \Gamma_f^{\mathrm{T}}Q_f\Gamma_f. \tag{26}$$

Remark 4.4 Note that (23) is the Riccati equation associated with an optimal control problem with controlled diffusion. If (23)–(24) admits a solution (Λ_1,Λ_2), substituting (Λ_1,Λ_2) into (25)–(26) gives a first order linear ODE system of (Λ_3,Λ_4), which then admits a unique solution on $[0,T]$.

Remark 4.5 If $B_1 = 0$ and (23)–(24) has a solution on $[0,T]$, from (25)–(26) we obtain a first order linear homogeneous ODE of $\Lambda_3-\Lambda_4$ with zero terminal condition $\Lambda_3(T)-\Lambda_4(T) = 0$, which implies that $\Lambda_3-\Lambda_4 = 0$ on $[0,T]$. Such a representation by three submatrices is similar to [34, Theorem 3].

The following theorem characterizes asymptotic solvability of the Nash game (1)–(2) in terms of the low-dimensional ODE system (23)–(24). The proof is postponed near the end of this section.

Theorem 4.6 *The Nash game* (1)–(2) *has asymptotic solvability if and only if* (23)–(24) *has a solution* (Λ_1,Λ_2) *on* $[0,T]$.

Following the rescaling method in [34–36], we define

$$\Lambda_1^N(t) = \Pi_1^N(t), \quad \Lambda_2^N(t) = N\Pi_2^N(t), \quad \Lambda_3^N(t) = N^2\Pi_3^N(t), \quad \Lambda_4^N(t) = N^2\Pi_4^N(t). \tag{27}$$

We introduce the following ODE system for $(\Lambda_1^N, \Lambda_2^N, \cdots, \Lambda_4^N)$:

$$\begin{cases} \dot{\Lambda}_1^N = \Psi_1(\Lambda_1^N) + g_1^N, \\ \Lambda_1^N(T) = (I - \Gamma_f^{\mathrm{T}}/N)Q_f(I - \Gamma_f/N), \quad \mathcal{R}_1(\Lambda_1^N(t)) > 0, \quad \forall t \in [0, T], \end{cases} \tag{28}$$

$$\begin{cases} \dot{\Lambda}_2^N = \Psi_2(\Lambda_1^N, \Lambda_2^N) + g_2^N, \\ \Lambda_2^N(T) = -(I - \Gamma_f^{\mathrm{T}}/N)Q_f\Gamma_f, \end{cases} \tag{29}$$

$$\begin{cases} \dot{\Lambda}_3^N = \Psi_3(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N) + g_3^N, \\ \Lambda_3^N(T) = \Gamma_f^{\mathrm{T}}Q_f\Gamma_f, \end{cases} \tag{30}$$

$$\begin{cases} \dot{\Lambda}_4^N = \Psi_4(\Lambda_1^N, \Lambda_2^N, \Lambda_4^N) + g_4^N, \\ \Lambda_4^N(T) = \Gamma_f^{\mathrm{T}}Q_f\Gamma_f, \end{cases} \tag{31}$$

where g_k^N, $1 \le k \le 4$, are perturbation terms. We have

$$\begin{aligned} g_1^N = &- [\Lambda_2^N BK^N B^{\mathrm{T}} S_{12}^N + S_{12}^{N\mathrm{T}} BK^{N\mathrm{T}} B^{\mathrm{T}} \Lambda_2^{N\mathrm{T}}](N-1)/N^3 \\ &+ [\Lambda_2^N BH^N B^{\mathrm{T}} \Lambda_2^N + \Lambda_2^{N\mathrm{T}} BH^N B^{\mathrm{T}} \Lambda_2^{N\mathrm{T}}](N-1)/N^2 \\ &- [S_{12}^{N\mathrm{T}} BK^{N\mathrm{T}}/N - \Lambda_2^{N\mathrm{T}} BH^N] B_1^{\mathrm{T}} \Lambda_3^N B_1 [K^N B^{\mathrm{T}} S_{12}^N/N - H^N B^{\mathrm{T}} \Lambda_2^N](N-1)/N^4 \\ &- [\Lambda_1^N G + G^{\mathrm{T}} \Lambda_1^N]/N - [\Lambda_2^N G + G^{\mathrm{T}} \Lambda_2^{N\mathrm{T}}](N-1)/N^2 \\ &- S_{12}^{N\mathrm{T}} BF^N B^{\mathrm{T}} S_{12}^N/N^2 - (\Gamma^{\mathrm{T}} Q\Gamma/N - \Gamma^{\mathrm{T}} Q - Q\Gamma)/N \end{aligned}$$

and

$$\begin{aligned} H^N &= (R + B_1^{\mathrm{T}} \Lambda_1^N B_1)^{-1}, \\ S^N &= \Lambda_1^N + (\Lambda_2^N + \Lambda_2^{N\mathrm{T}})(N-1)/N + [\Lambda_3^N + \Lambda_4^N(N-2)](N-1)/N^2, \\ S_{12}^N &= \Lambda_1^N + \Lambda_2^N(N-1)/N, \\ S_{34}^N &= \Lambda_3^N/N^2 + \Lambda_4^N(N-2)/N^2, \\ K^N &= H^N B_0^{\mathrm{T}} S^N B_0 (I + H^N B_0^{\mathrm{T}} S^N B_0/N)^{-1} H^N, \\ F^N &= H^N (I + B_0^{\mathrm{T}} S^N B_0 H^N/N)^{-1} (B_0^{\mathrm{T}} S^N B_0 + B_0^{\mathrm{T}} S^N B_0 H^N B_0^{\mathrm{T}} S^N B_0/N^2) \\ &\quad \cdot (I + H^N B_0^{\mathrm{T}} S^N B_0/N)^{-1} H^N. \end{aligned}$$

The other terms g_k^N, $k = 2, 3, 4$, are not displayed due to limited space and can be found in [41]. They depend on S_{34}^N above. The mappings g_k^N, $1 \le k \le 4$, are defined for $\Lambda_k^N \in \mathcal{S}^n$, $k = 1, 3, 4$, and $\Lambda_2^N \in \mathbb{R}^{n\times n}$. If (28)–(31) has a solution on $[0, T]$, then $\Lambda_k^N(t)$ is $\mathcal{S}^n$-valued for $k = 1, 3, 4$. The ODE system (28)–(31) is essentially derived from (16) by use of the new variables (27). However, (28)–(31) can stand alone without being immediately related to (16). If $(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N)$ is a solution, the inverse $(I + H^N B_0^{\mathrm{T}} S^N B_0/N)^{-1}$ necessarily exists for all $t \in [0, T]$; such a solution is unique.

For $\Lambda_k^N \in \mathcal{S}^n$, $k=1,3,4$, and $\Lambda_2^N \in \mathbb{R}^{n\times n}$, define the mappings

$$\begin{aligned}\xi(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N) =&\mathcal{R}_1(\Lambda_1^N)+B_0^{\rm T}[\Lambda_1^N+(\Lambda_2^N+\Lambda_2^{N{\rm T}})(N-1)/N\\ &+\Lambda_3^N(N-1)/N^2+\Lambda_4^N(N-1)(N-2)/N^2]B_0/N^2, \qquad (32)\\ \xi_0(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N) =&I+\big(\mathcal{R}_1(\Lambda_1^N)\big)^{-1}B_0^{\rm T}[\Lambda_1^N+(\Lambda_2^N+\Lambda_2^{N{\rm T}})(N-1)/N\\ &+\Lambda_3^N(N-1)/N^2+\Lambda_4^N(N-1)(N-2)/N^2]B_0/N. \qquad (33)\end{aligned}$$

It is easy to show that

$$I+H^NB_0^{\rm T}S^NB_0/N=\xi_0(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N). \qquad (34)$$

Lemma 4.7 (i) *Suppose* (16)–(17) *has a solution* $(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N)$ *on* $[0,T]$*, and let* $(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N)$ *be defined using* (19) *and* (27). *Then* $(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N)$ *satisfies* (28)–(31).

(ii) *Conversely, if* (28)–(31) *admits a solution* $(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N)$ *on* $[0,T]$*, and such a solution further satisfies*

$$\xi(\Lambda_1^N(t),\Lambda_2^N(t),\Lambda_3^N(t),\Lambda_4^N(t))>0$$

for all $t\in[0,T]$*, then* (16)–(17) *has a solution* $(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N)$ *on* $[0,T]$*. Moreover,* $\boldsymbol{P}_i$ *may be determined in terms of the above* $(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N)$ *using* (19).

Proof (i) By (19) and (27), we have

$$R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i=\mathcal{R}_1(\Lambda_1^N).$$

By the condition (17) (i), $R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i>0$. Therefore, $\mathcal{R}_1(\Lambda_1^N(t))>0$ on $[0,T]$. It can be shown that

$$I+\sum_{k=1}^N\big(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k(t)\boldsymbol{B}_k\big)^{-1}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k(t)\boldsymbol{B}_0=\xi_0(\Lambda_1^N(t),\Lambda_2^N(t),\Lambda_3^N(t),\Lambda_4^N(t)). \qquad (35)$$

We substitute (19) into (16) and change to the variables Λ_k^N, $1\le k\le 4$, to verify the equalities (28)–(31), for which the inverse $(I+H^NB_0^{\rm T}S^NB_0/N)^{-1}$ exists by condition (17) (iii), (34) and (35).

(ii) If (28)–(31) admits a solution $(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N)$ on $[0,T]$, let $\boldsymbol{P}_i$ be defined by (19) and (27). By $\mathcal{R}_1(\Lambda_1^N)>0$ in (28), we have $R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i>0$. We can verify

$$R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i+\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_0=\xi(\Lambda_1^N,\Lambda_2^N,\Lambda_3^N,\Lambda_4^N) \qquad (36)$$

so that $R+\boldsymbol{B}_i^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_i+\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_i\boldsymbol{B}_0>0$ for all $t\in[0,T]$. Note that $(I+H^NB_0^{\rm T}S^NB_0/N)^{-1}$ in (28)–(31) exists for all $t\in[0,T]$. Recalling (33)–(34) and (35), we see that

$$I+\sum_{k=1}^N\big(R+\boldsymbol{B}_k^{\rm T}\boldsymbol{P}_k(t)\boldsymbol{B}_k\big)^{-1}\boldsymbol{B}_0^{\rm T}\boldsymbol{P}_k(t)\boldsymbol{B}_0$$

is invertible for all $t \in [0,T]$. Now it is straightforward to verify that $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ defined above solves (16) subject to (17). ∎

Proof of Theorem 4.6 (i) Necessity. Suppose the game (1)–(2) has asymptotic solvability, where N_0 and $c_0 > 0$ have been selected in (20)–(22). By Lemma 4.7 (i), for all $N \geq N_0$, (28)–(31) has a solution $(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N)$ on $[0,T]$, and by (20)–(21) and (27), we have

$$\sup_{N \geq N_0} \sup_{0 \leq t \leq T} (|\Lambda_1^N| + |\Lambda_2^N| + |\Lambda_3^N| + |\Lambda_4^N|) < \infty, \tag{37}$$

$$\mathcal{R}_1(\Lambda_1^N(t)) \geq c_0 I, \quad \forall t \in [0,T], \quad \forall N \geq N_0. \tag{38}$$

We write (28) in the integral form

$$\Lambda_1^N(t) = \Lambda_1^N(T) - \int_t^T [\Psi_1(\Lambda_1^N) + g_1^N] d\tau,$$

and do the same for Λ_2^N, Λ_3^N and Λ_4^N. By (37)–(38) we obtain $\sup_{0 \leq t \leq T, k \leq 4} |g_k^N| = O(1/N)$. Then the functions $\{(\Lambda_1^N(\cdot), \Lambda_2^N(\cdot)), \Lambda_3^N(\cdot)), \Lambda_4^N(\cdot))\}_{N \geq N_0}$ are uniformly bounded and equicontinuous on $[0,T]$. By Arzelà-Ascoli theorem[42], there exists a subsequence $\{(\Lambda_1^{N_j}(\cdot), \Lambda_2^{N_j}(\cdot), \Lambda_3^{N_j}(\cdot), \Lambda_4^{N_j}(\cdot))\}_{j \geq 1}$ that converges to $(\Lambda_1^*, \Lambda_2^*, \Lambda_3^*, \Lambda_4^*)$ uniformly on $[0,T]$ as $j \to \infty$. It is easy to see that $(\Lambda_1^*, \Lambda_2^*, \Lambda_3^*, \Lambda_4^*)$ solves the system (23)–(26) and $\mathcal{R}_1(\Lambda_1^*(t)) \geq c_0 I$ for all $t \in [0,T]$.

(ii) Sufficiency. Suppose (23)–(24) has a solution so that we can obtain $(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4)$ from (23)–(26). We proceed to check the solution of (28)–(31), which now stands alone without using (16). Following the method in the sufficiency proof of Theorem 3.1 in [36], we specify a thin "tube", surrounding the solution trajectory $(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4)$, $t \in [0,T]$, of this form:

$$\mathcal{C} = \left\{ (t, Z_1, Z_2, Z_3, Z_4) \in [0,T] \times \mathcal{S}^n \times \mathbb{R}^{n \times n} \times \mathcal{S}^n \times \mathcal{S}^n : \sum_{k \leq 4} |Z_k - \Lambda_k(t)| < \delta_0 \right\}, \tag{39}$$

where $\delta_0 > 0$ is a sufficiently small but fixed constant, and next show that for all sufficiently large N, the solution of (28)–(31) starting from the terminal condition will always remain in this tube. This establishes the global existence of solutions on $[0,T]$, and the detailed steps are exactly the same as in [36]. Specifically, it can be shown that there exist $\widehat{N}_0$ and $c_0 > 0$ such that we have the following: (a) (28)–(31) has a solution $(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N)$ remaining in the tube (39) on $[0,T]$ for all $N \geq \widehat{N}_0$; (b)

$$\sup_{N \geq \widehat{N}_0} \sup_{0 \leq t \leq T} (|\Lambda_1^N| + |\Lambda_2^N| + |\Lambda_3^N| + |\Lambda_4^N|) < \infty, \tag{40}$$

$$\mathcal{R}_1(\Lambda_1^N(t)) \geq c_0 I, \quad \forall t \in [0,T], \quad \forall N \geq \widehat{N}_0; \tag{41}$$

(c) for $\xi_0(\cdot)$ defined in (33), $\xi_0(\Lambda_1^N(t), \Lambda_2^N(t), \Lambda_3^N(t), \Lambda_4^N(t))$ is invertible for all $N \geq \widehat{N}_0$, so that the term $(I + H^N B_0^{\mathrm{T}} S^N B_0/N)^{-1}$ in (28)–(31) is well defined.

If $\widehat{N}_1 > \widehat{N}_0$ is sufficiently large, by (41) we can ensure that

$$\xi(\Lambda_1^N(t), \Lambda_2^N(t), \Lambda_3^N(t), \Lambda_4^N(t)) > (c_0/2)I, \quad \forall t \in [0,T], \ \forall N \geq \widehat{N}_1, \tag{42}$$

where $\xi(\cdot)$ is defined in (32). By (42), we apply Lemma 4.7 (ii) to obtain $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ for (16) whenever $N \geq \widehat{N}_1$. By (42) and (35), we see that (22) holds for all $N \geq \widehat{N}_1$. Subsequently, asymptotic solvability holds. ∎

Corollary 4.8 *If (23)–(24) has a solution (Λ_1, Λ_2) on $[0, T]$, then there exists $\widehat{N}_0 > 0$ such that for each $N \geq \widehat{N}_0$, (28)–(31) has a solution $(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N)$ on $[0, T]$ and moreover, $\sup_{t\in[0,T],k\leq 4} |\Lambda_k^N(t) - \Lambda_k(t)| = O(1/N)$.*

Proof Since (23)–(26) has a solution on $[0, T]$, we take a sufficiently thin tube as in (39). Then by the sufficiency proof of Theorem 4.6, there exists $\widehat{N}_0 > 0$ such that for each $N \geq \widehat{N}_0$, (28)–(31) has a solution $(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N)$ on $[0, T]$, which is always within the tube. The desired result then follows from Grönwall's lemma. See similar estimates in [36, Corollary 3.1]. ∎

5 Decentralized Strategies

By Theorem 4.6, the Nash game (1)–(2) has asymptotic solvability if and only if (23)–(24) admits a solution (Λ_1, Λ_2) on $[0, T]$. We introduce the following assumptions:

Assumption 5.1 The ODE system (23)–(24) has a solution (Λ_1, Λ_2) on $[0, T]$.

For $X_i(0)$, denote the covariance matrix $\Sigma_0^i = \mathbb{E}\{[X_i(0) - \mathbb{E}X_i(0)][X_i(0) - \mathbb{E}X_i(0)]^{\mathrm{T}}\}$.

Assumption 5.2 The initial states $\{X_i(0), i \geq 0\}$ are independent. There exist $\mu_0 \in \mathbb{R}^n$ and a constant C_Σ, both independent of N, such that $\mathbb{E}X_i(0) = \mu_0$ and $|\Sigma_0^i| \leq C_\Sigma$ for all i.

Under Assumption 5.1, the sufficiency proof of Theorem 4.6 shows that there exists $\widehat{N}_1$ such that (28)–(31) has a solution $(\Lambda_1^N, \Lambda_2^N, \Lambda_3^N, \Lambda_4^N)$ for all $N \geq \widehat{N}_1$. By Lemma 4.7 (ii), we determine $\boldsymbol{P}$ in (16) by using (19) and (27), and obtain the Nash equilibrium strategies (18), which are displayed below:

$$\widehat{u}_i(t) = -[R + B_1^{\mathrm{T}} \Lambda_1^N(t) B_1]^{-1} [\boldsymbol{B}_0^{\mathrm{T}} \boldsymbol{P}_i(t) \boldsymbol{B}_0 \boldsymbol{M}_0(t) + \widehat{\boldsymbol{B}}_i^{\mathrm{T}} \boldsymbol{P}_i(t)] X(t), \quad 1 \leq i \leq N.$$

Throughout this section we assume $N \geq \widehat{N}_1$. Before further analysis we introduce some notation:

$$\begin{aligned}
&\Theta(t) = (\mathcal{R}_1(\Lambda_1(t)))^{-1} B^{\mathrm{T}} \Lambda_1(t), \quad \Theta_1(t) = (\mathcal{R}_1(\Lambda_1(t)))^{-1} B^{\mathrm{T}} \Lambda_2(t), \\
&\widehat{\Theta}(t) = I_N \otimes \Theta(t), \quad \widehat{\Theta}_1 = \mathbf{1}_{N\times 1} \otimes \Theta_1, \\
&\boldsymbol{e}_i = (e_i^N \otimes I_n)^{\mathrm{T}} = (0, \cdots, 0, I_n, 0, \cdots, 0) \in \mathbb{R}^{n\times Nn}, \\
&\widehat{\boldsymbol{B}} = (\widehat{\boldsymbol{B}}_1, \widehat{\boldsymbol{B}}_2, \cdots, \widehat{\boldsymbol{B}}_N) \in \mathbb{R}^{Nn\times Nn_1}, \quad \boldsymbol{I} = (I_n, I_n, \cdots, I_n) \in \mathbb{R}^{n\times Nn}.
\end{aligned}$$

By using the closed-loop dynamics under $(\widehat{u}_1, \widehat{u}_2, \cdots, \widehat{u}_N)$, we consider the SDE of $X^{(N)}$ and let $N \to \infty$. This gives the mean field limit state $\overline{X}$ as follows:

$$d\overline{X} = (A + G - B(\Theta + \Theta_1))\overline{X} dt - B_0(\Theta + \Theta_1)\overline{X} dW_0, \quad t \geq 0, \tag{43}$$

where $\overline{X}(0) = \mu_0$. We denote the set of decentralized feedback strategies

$$\check{u}_i(t) = -\Theta(t)\boldsymbol{e}_i X(t) - \Theta_1(t)\overline{X}(t), \quad 1 \leq i \leq N. \tag{44}$$

The state dynamics under the decentralized strategies (44) follows

$$dX(t) =(\boldsymbol{A}X - \widehat{\boldsymbol{B}}(\widehat{\Theta}X + \widehat{\Theta}_1\overline{X}))dt - \sum_{i=1}^{N}\boldsymbol{B}_i(\Theta X_i + \Theta_1\overline{X})dW_i$$
$$- \boldsymbol{B}_0\sum_{i=1}^{N}(\Theta X_i + \Theta_1\overline{X})dW_0, \quad t \geq 0,$$

where the initial state $X(0) = (X_1^{\mathrm{T}}(0), X_2^{\mathrm{T}}(0), \cdots, X_N^{\mathrm{T}}(0))^{\mathrm{T}}$ is the same as in (3).

Below we evaluate the cost with more general initial conditions. When all the N players take the decentralized strategies (44), the cost of player $\mathcal{A}_i$ with initial condition $(X(t), \overline{X}(t)) = (\boldsymbol{x}, \overline{x})$ is denoted by $\check{V}_i(t, \boldsymbol{x}, \overline{x})$, $t \in [0, T]$, $\boldsymbol{x} \in \mathbb{R}^{Nn}$, $\overline{x} \in \mathbb{R}^n$. The Feynman-Kac formula [43, Sec. 1.3, 3.5] gives the following equation that $\check{V}_i$ satisfies:

$$\begin{cases} -\dfrac{\partial \check{V}_i}{\partial t} = \dfrac{\partial^{\mathrm{T}} \check{V}_i}{\partial \boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x} - \widehat{\boldsymbol{B}}(\widehat{\Theta}\boldsymbol{x} + \widehat{\Theta}_1\overline{x})) + \dfrac{\partial^{\mathrm{T}} \check{V}_i}{\partial \overline{x}}(A + G - B(\Theta + \Theta_1))\overline{x} \\ \qquad +(1/2)(\Theta\boldsymbol{I}\boldsymbol{x} + \boldsymbol{I}\widehat{\Theta}_1\overline{x})^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\dfrac{\partial^2 \check{V}_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0(\Theta\boldsymbol{I}\boldsymbol{x} + \boldsymbol{I}\widehat{\Theta}_1\overline{x}) \\ \qquad +(1/2)((\Theta + \Theta_1)\overline{x})^{\mathrm{T}}B_0^{\mathrm{T}}\dfrac{\partial^2 \check{V}_i}{\partial \overline{x}^2}B_0(\Theta + \Theta_1)\overline{x} \\ \qquad +(1/2)\displaystyle\sum_{k=1}^{N}(\Theta\boldsymbol{e}_k\boldsymbol{x} + \Theta_1\overline{x})^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}\dfrac{\partial^2 \check{V}_i}{\partial \boldsymbol{x}^2}\boldsymbol{B}_k(\Theta\boldsymbol{e}_k\boldsymbol{x} + \Theta_1\overline{x}) \\ \qquad +(\Theta\boldsymbol{I}\boldsymbol{x} + \boldsymbol{I}\widehat{\Theta}_1\overline{x})^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\dfrac{\partial^2 \check{V}_i}{\partial \boldsymbol{x}\partial \overline{x}}B_0(\Theta + \Theta_1)\overline{x} \\ \qquad +(\Theta\boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x})^{\mathrm{T}}R(\Theta\boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x}) + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_i\boldsymbol{x}, \\ \check{V}_i(T, \boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_{if}\boldsymbol{x}. \end{cases} \tag{45}$$

Assume $\check{V}_i(t, \boldsymbol{x}, \overline{x})$ takes the following form

$$\check{V}_i(t, \boldsymbol{x}, \overline{x}) = \boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{P}}_1^i(t)\boldsymbol{x} + 2\boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^i(t)\overline{x} + \overline{x}^{\mathrm{T}}\check{\boldsymbol{P}}_2^i(t)\overline{x}, \quad 1 \leq i \leq N. \tag{46}$$

Substituting (46) into (45) gives the following system of ODEs for $\check{\boldsymbol{P}}_1^i$, $\check{\boldsymbol{P}}_{12}^i$ and $\check{\boldsymbol{P}}_2^i$:

$$\begin{cases} -\dfrac{d}{dt}\check{\boldsymbol{P}}_1^i = \check{\boldsymbol{P}}_1^i(\boldsymbol{A} - \widehat{\boldsymbol{B}}\widehat{\Theta}) + (\boldsymbol{A} - \widehat{\boldsymbol{B}}\widehat{\Theta})^{\mathrm{T}}\check{\boldsymbol{P}}_1^i + \boldsymbol{I}^{\mathrm{T}}\Theta^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\check{\boldsymbol{P}}_1^i\boldsymbol{B}_0\Theta\boldsymbol{I} \\ \qquad + \displaystyle\sum_{k=1}^{N}\boldsymbol{e}_k^{\mathrm{T}}\Theta^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}\check{\boldsymbol{P}}_1^i\boldsymbol{B}_k\Theta\boldsymbol{e}_k + \boldsymbol{e}_i^{\mathrm{T}}\Theta^{\mathrm{T}}R\Theta\boldsymbol{e}_i + \boldsymbol{Q}_i, \\ \check{\boldsymbol{P}}_1^i(T) = \boldsymbol{Q}_{if}, \end{cases} \tag{47}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\boldsymbol{P}}_{12}^i = -\check{\boldsymbol{P}}_1^i\widehat{\boldsymbol{B}}\widehat{\Theta}_1 + (\boldsymbol{A} - \widehat{\boldsymbol{B}}\widehat{\Theta})^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^i + \check{\boldsymbol{P}}_{12}^i(A + G - B(\Theta + \Theta_1)) \\ \qquad +\boldsymbol{I}^{\mathrm{T}}\Theta^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\check{\boldsymbol{P}}_1^i\boldsymbol{B}_0\boldsymbol{I}\widehat{\Theta}_1 + \displaystyle\sum_{k=1}^{N}\boldsymbol{e}_k^{\mathrm{T}}\Theta^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}\check{\boldsymbol{P}}_1^i\boldsymbol{B}_k\Theta_1 \\ \qquad +\boldsymbol{I}^{\mathrm{T}}\Theta^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^iB_0(\Theta + \Theta_1) + \boldsymbol{e}_i^{\mathrm{T}}\Theta^{\mathrm{T}}R\Theta_1, \\ \check{\boldsymbol{P}}_{12}^i(T) = 0, \end{cases} \tag{48}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\boldsymbol{P}}_2^i = -\check{\boldsymbol{P}}_{12}^{i\mathrm{T}}\widehat{\boldsymbol{B}}\widehat{\Theta}_1 - \widehat{\Theta}_1^{\mathrm{T}}\widehat{\boldsymbol{B}}^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^i + \check{\boldsymbol{P}}_2^i(A+G-B(\Theta+\Theta_1)) \\ \qquad +(A+G-B(\Theta+\Theta_1))^{\mathrm{T}}\check{\boldsymbol{P}}_2^i + (\boldsymbol{B}_0\boldsymbol{I}\widehat{\Theta}_1)^{\mathrm{T}}\check{\boldsymbol{P}}_1^i\boldsymbol{B}_0\boldsymbol{I}\widehat{\Theta}_1 \\ \qquad +\sum_{k=1}^N(\boldsymbol{B}_k\Theta_1)^{\mathrm{T}}\check{\boldsymbol{P}}_1^i\boldsymbol{B}_k\Theta_1 + (\Theta+\Theta_1)^{\mathrm{T}}B_0^{\mathrm{T}}\check{\boldsymbol{P}}_2^iB_0(\Theta+\Theta_1) \\ \qquad +(\boldsymbol{B}_0\boldsymbol{I}\widehat{\Theta}_1)^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^iB_0(\Theta+\Theta_1) + (B_0(\Theta+\Theta_1))^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^{i\mathrm{T}}\boldsymbol{B}_0\boldsymbol{I}\widehat{\Theta}_1 \\ \qquad +\Theta_1^{\mathrm{T}}R\Theta_1, \\ \check{\boldsymbol{P}}_2^i(T) = 0. \end{cases} \tag{49}$$

Remark 5.1 (47)–(49) is a first order linear ODE system and admits a unique solution. We have the following submatrix partition of the matrices $\check{\boldsymbol{P}}_1^i$, $\check{\boldsymbol{P}}_{12}^i$, $1 \le i \le N$.

Lemma 5.2 *For* (47) *and* (48), *the solution* $(\check{\boldsymbol{P}}_1^i, \check{\boldsymbol{P}}_{12}^i)$, $1 \le i \le N$, *has the representation*

$$\check{\boldsymbol{P}}_1^1 = \begin{bmatrix} \check{\Pi}_1^N & \check{\Pi}_2^N & \check{\Pi}_2^N & \cdots & \check{\Pi}_2^N \\ \check{\Pi}_2^{N\mathrm{T}} & \check{\Pi}_3^N & \check{\Pi}_4^N & \cdots & \check{\Pi}_4^N \\ \check{\Pi}_2^{N\mathrm{T}} & \check{\Pi}_4^N & \check{\Pi}_3^N & \cdots & \check{\Pi}_4^N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \check{\Pi}_2^{N\mathrm{T}} & \check{\Pi}_4^N & \check{\Pi}_4^N & \cdots & \check{\Pi}_3^N \end{bmatrix}, \qquad \check{\boldsymbol{P}}_1^i = J_{1i}^{\mathrm{T}}\check{\boldsymbol{P}}_1^1J_{1i}, \quad \forall 2 \le i \le N, \tag{50}$$

$$\check{\boldsymbol{P}}_{12}^1 = \left[\check{\Pi}_{11}^{N\mathrm{T}},\ \check{\Pi}_{12}^{N\mathrm{T}},\ \cdots,\ \check{\Pi}_{12}^{N\mathrm{T}}\right]^{\mathrm{T}}, \quad \check{\boldsymbol{P}}_{12}^i = J_{1i}^{\mathrm{T}}\check{\boldsymbol{P}}_{12}^1, \quad \forall 2 \le i \le N, \tag{51}$$

where $\check{\Pi}_1^N(t)$, $\check{\Pi}_3^N(t)$, $\check{\Pi}_4^N(t) \in \mathcal{S}^n$, *and* $\check{\Pi}_2^N(t)$, $\check{\Pi}_{11}^N(t)$, $\check{\Pi}_{12}^N(t) \in \mathbb{R}^{n\times n}$.

Proof The proof is similar to that of Lemma 4.1 or [34, Theorem 3] and is omitted. ∎

We define the new variables:

$$\begin{cases} \check{\Lambda}_1^N = \check{\Pi}_1^N, \quad \check{\Lambda}_2^N = N\check{\Pi}_2^N, \quad \check{\Lambda}_3^N = N^2\check{\Pi}_3^N, \quad \check{\Lambda}_4^N = N^2\check{\Pi}_4^N, \\ \check{\Lambda}_{11}^N = \check{\Pi}_{11}^N, \quad \check{\Lambda}_{12}^N = N\check{\Pi}_{12}^N, \quad \check{\Lambda}_{22}^N = \check{\boldsymbol{P}}_2^i. \end{cases} \tag{52}$$

Substituting (50)–(51) into (45) and next converting into the new variables in (52), we derive

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_1^N = \check{\Lambda}_1^N(A-B\Theta) + (A-B\Theta)^{\mathrm{T}}\check{\Lambda}_1^N + \Theta^{\mathrm{T}}(R + B_1^{\mathrm{T}}\check{\Lambda}_1^NB_1)\Theta \\ \qquad +Q + \check{g}_1^N, \\ \check{\Lambda}_1^N(T) = (I - \Gamma_f^{\mathrm{T}}/N)Q_f(I - \Gamma_f/N), \end{cases} \tag{53}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_2^N = \check{\Lambda}_1^NG + \check{\Lambda}_2^N(A+G-B\Theta) + (A-B\Theta)^{\mathrm{T}}\check{\Lambda}_2^N - Q\Gamma + \check{g}_2^N, \\ \check{\Lambda}_2^N(T) = -(I - \Gamma_f^{\mathrm{T}}/N)Q_f\Gamma_f, \end{cases} \tag{54}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_3^N = \Theta^{\mathrm{T}}[B_1^{\mathrm{T}}\check{\Lambda}_3^NB_1 + B_0^{\mathrm{T}}(\check{\Lambda}_1^N + \check{\Lambda}_2^N + \check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N)B_0]\Theta \\ \qquad +(\check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N)G + G^{\mathrm{T}}(\check{\Lambda}_2^N + \check{\Lambda}_4^N) + \check{\Lambda}_3^N(A-B\Theta) \\ \qquad +(A-B\Theta)^{\mathrm{T}}\check{\Lambda}_3^N + \Gamma^{\mathrm{T}}Q\Gamma + \check{g}_3^N, \\ \check{\Lambda}_3^N(T) = \Gamma_f^{\mathrm{T}}Q_f\Gamma_f, \end{cases} \tag{55}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_4^N = \check{\Lambda}_2^{N\mathrm{T}} G + G^{\mathrm{T}} \check{\Lambda}_2^N + \check{\Lambda}_4^N (A + G - B\Theta) + (A + G - B\Theta)^{\mathrm{T}} \check{\Lambda}_4^N \\ \qquad\qquad + \Theta^{\mathrm{T}} B_0^{\mathrm{T}} (\check{\Lambda}_1^N + \check{\Lambda}_2^N + \check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N) B_0 \Theta + \Gamma^{\mathrm{T}} Q \Gamma + \check{g}_4^N, \\ \check{\Lambda}_4^N(T) = \Gamma_f^{\mathrm{T}} Q_f \Gamma_f, \end{cases} \tag{56}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_{11}^N = \check{\Lambda}_{11}^N (A + G - B(\Theta + \Theta_1)) + (A - B\Theta)^{\mathrm{T}} \check{\Lambda}_{11}^N \\ \qquad\qquad - (\check{\Lambda}_1^N + \check{\Lambda}_2^N) B \Theta_1 + \Theta^{\mathrm{T}} (R + B_1^{\mathrm{T}} \check{\Lambda}_1^N B_1) \Theta_1 + \check{g}_{11}^N, \\ \check{\Lambda}_{11}^N(T) = 0, \end{cases} \tag{57}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_{12}^N = -(\check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N) B \Theta_1 + (A + G - B\Theta)^{\mathrm{T}} \check{\Lambda}_{12}^N + G^{\mathrm{T}} \check{\Lambda}_{11}^N \\ \qquad\qquad + \check{\Lambda}_{12}^N (A + G - B(\Theta + \Theta_1)) \\ \qquad\qquad + \Theta^{\mathrm{T}} B_0^{\mathrm{T}} (\check{\Lambda}_1^N + \check{\Lambda}_2^N + \check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N) B_0 \Theta_1 \\ \qquad\qquad + \Theta^{\mathrm{T}} B_0^{\mathrm{T}} (\check{\Lambda}_{11}^N + \check{\Lambda}_{12}^N) B_0 (\Theta + \Theta_1) + \check{g}_{12}^N, \\ \check{\Lambda}_{12}^N(T) = 0, \end{cases} \tag{58}$$

$$\begin{cases} -\dfrac{d}{dt}\check{\Lambda}_{22}^N = -(\check{\Lambda}_{11}^N + \check{\Lambda}_{12}^N + \check{\Lambda}_{22}^N)^{\mathrm{T}} B \Theta_1 - \Theta_1^{\mathrm{T}} B^{\mathrm{T}} (\check{\Lambda}_{11}^N + \check{\Lambda}_{12}^N + \check{\Lambda}_{22}^N) \\ \qquad\qquad + \check{\Lambda}_{22}^N (A + G - B\Theta) + (A + G - B\Theta)^{\mathrm{T}} \check{\Lambda}_{22}^N \\ \qquad\qquad + \Theta^{\mathrm{T}} B_0^{\mathrm{T}} \check{\Lambda}_{22}^N B_0 \Theta + \Theta^{\mathrm{T}} B_0^{\mathrm{T}} (\check{\Lambda}_{11}^N + \check{\Lambda}_{12}^N + \check{\Lambda}_{22}^N)^{\mathrm{T}} B_0 \Theta_1 \\ \qquad\qquad + \Theta_1^{\mathrm{T}} B_0^{\mathrm{T}} (\check{\Lambda}_{11}^N + \check{\Lambda}_{12}^N + \check{\Lambda}_{22}^N) B_0 \Theta + \Theta_1^{\mathrm{T}} (R + B_1^{\mathrm{T}} \check{\Lambda}_1^N B_1) \Theta_1 \\ \qquad\qquad + \Theta_1^{\mathrm{T}} B_0^{\mathrm{T}} (\check{\Lambda}_1^N + \check{\Lambda}_2^N + \check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N \\ \qquad\qquad + \check{\Lambda}_{11}^N + \check{\Lambda}_{12}^N + \check{\Lambda}_{11}^{N\mathrm{T}} + \check{\Lambda}_{12}^{N\mathrm{T}} + \check{\Lambda}_{22}^N) B_0 \Theta_1 + \check{g}_{22}^N, \\ \check{\Lambda}_{22}^N(T) = 0. \end{cases} \tag{59}$$

We have the perturbation terms $\check{g}_1^N, \check{g}_2^N, \cdots, \check{g}_{22}^N$, and

$$\begin{aligned} \check{g}_1^N :=& (\check{\Lambda}_1^N G + G^{\mathrm{T}} \check{\Lambda}_1^N)/N + (\check{\Lambda}_2^N G + G^{\mathrm{T}} \check{\Lambda}_2^N)(N-1)/N^2 \\ &+ (\Gamma^{\mathrm{T}} Q \Gamma / N - \Gamma^{\mathrm{T}} Q - Q\Gamma)/N + \Theta^{\mathrm{T}} B_0^{\mathrm{T}} \check{S}^N B_0 \Theta / N^2, \\ \check{S}^N =& \check{\Lambda}_1^N + (\check{\Lambda}_2^N + \check{\Lambda}_2^{N\mathrm{T}})(N-1)/N + \check{\Lambda}_3^N (N-1)/N^2 + \check{\Lambda}_4^N (N-1)(N-2)/N^2. \end{aligned}$$

The remaining perturbation terms can be found in [41].

Remark 5.3 Under Assumption 5.1, the system (53)–(59) is a first order linear ODE system and admits a unique solution $(\check{\Lambda}_1^N, \check{\Lambda}_2^N, \cdots, \check{\Lambda}_{22}^N)$ on $[0, T]$.

Remark 5.4 Let ψ^N stand for any of the functions $\check{\Lambda}_1^N$, $\check{\Lambda}_2^N$, $\check{\Lambda}_3^N$, $\check{\Lambda}_4^N$, $\check{\Lambda}_{11}^N$, $\check{\Lambda}_{12}^N$ and $\check{\Lambda}_{22}^N$. Due to the bounded coefficients in the ODE system (53)–(59), $\sup_{N \geq \widehat{N}_1, 0 \leq t \leq T} |\psi^N| \leq C$ for some fixed constant C.

Remark 5.5 Let h^N stand for any of the functions $\check{g}_1^N$, $\check{g}_2^N$, $\check{g}_3^N$, $\check{g}_4^N$, $\check{g}_{11}^N$, $\check{g}_{12}^N$ and $\check{g}_{22}^N$. Then $\sup_{t \in [0,T]} |h^N(t)| = O(1/N)$.

Let $(\check{\Lambda}_1^N, \check{\Lambda}_2^N, \cdots, \check{\Lambda}_{22}^N)$ be obtained from (53)–(59). By substituting (50) into (46), which is further expressed in terms of $(\check{\Lambda}_1^N, \check{\Lambda}_2^N, \cdots, \check{\Lambda}_{22}^N)$ via (52), we obtain an explicit representation of a player's cost when all the players take the set of decentralized strategies $(\check{u}_1, \check{u}_2, \cdots, \check{u}_N)$ in (44). The cost of player $\mathcal{A}_i$ is

$$\begin{aligned} J_i(\check{u}_i, \check{u}_{-i}) =& \mathbb{E}\big[\check{V}_i(0, X(0), \overline{X}(0))\big] \\ =& \mathbb{E}\big[X^{\mathrm{T}}(0)\check{\boldsymbol{P}}_1^i(0)X(0) + 2X^{\mathrm{T}}(0)\check{\boldsymbol{P}}_{12}^i(0)\overline{X}(0) + \overline{X}^{\mathrm{T}}(0)\check{\boldsymbol{P}}_2^i(0)\overline{X}(0)\big]. \end{aligned} \tag{60}$$

Denote $\mathcal{N}_{-i} = \{1, 2, \cdots, N\} \setminus \{i\}$. Under Assumption 5.2, the first term on the right hand side of (60) is

$$\begin{aligned} &\mathbb{E}\big[X^{\mathrm{T}}(0)\check{\boldsymbol{P}}_1^i(0)X(0)\big] \\ =&\mathbb{E}\Bigg[X_i^{\mathrm{T}}(0)\check{\Lambda}_1^N(0)X_i(0) + (2/N)\sum_{j\in\mathcal{N}_{-i}} X_i^{\mathrm{T}}(0)\check{\Lambda}_2^N(0)X_j(0) \\ &+ \frac{1}{N^2}\sum_{j\in\mathcal{N}_{-i}} X_j^{\mathrm{T}}(0)\check{\Lambda}_3^N(0)X_j(0) + \frac{1}{N^2}\sum_{j,k\in\mathcal{N}_{-i}, j\neq k} X_j^{\mathrm{T}}(0)\check{\Lambda}_4^N(0)X_k(0)\Bigg] \\ =& \operatorname{Tr}[\check{\Lambda}_1^N(0)\Sigma_0^i] + (1/N^2)\sum_{j\in\mathcal{N}_{-i}} \operatorname{Tr}[\check{\Lambda}_3^N(0)\Sigma_0^j] \\ &+ \mu_0^{\mathrm{T}}[\check{\Lambda}_1^N(0) + \check{\Lambda}_2^N(0) + \check{\Lambda}_2^{N\mathrm{T}}(0) + \check{\Lambda}_4^N(0)]\mu_0 \\ &+ \mu_0^{\mathrm{T}}[-(\check{\Lambda}_2^N(0) + \check{\Lambda}_2^{N\mathrm{T}}(0))/N + \check{\Lambda}_3^N(0)(N-1)/N^2 \\ &\quad + \check{\Lambda}_4^N(0)(2-3N)/N]\mu_0. \end{aligned} \tag{61}$$

The second term on the right hand side of (60) can be written as

$$\begin{aligned} &2\mathbb{E}\big[X^{\mathrm{T}}(0)\check{\boldsymbol{P}}_{12}^i(0)\overline{X}(0)\big] \\ =&\mu_0^{\mathrm{T}}[\check{\Lambda}_{11}^N(0) + \check{\Lambda}_{11}^{N\mathrm{T}}(0) + \check{\Lambda}_{12}^N(0) + \check{\Lambda}_{12}^{N\mathrm{T}}(0)]\mu_0 - \mu_0^{\mathrm{T}}[\check{\Lambda}_{12}^N(0) + \check{\Lambda}_{12}^{N\mathrm{T}}(0)]\mu_0/N. \end{aligned} \tag{62}$$

The third term on the right hand side of (60) can be written as

$$\mathbb{E}[\overline{X}^{\mathrm{T}}(0)\check{\boldsymbol{P}}_2^i(0)\overline{X}(0)] = \mu_0^{\mathrm{T}}\check{\Lambda}_{22}^N(0)\mu_0. \tag{63}$$

Denote

$$\check{Y}^N := \check{\Lambda}_1^N + \check{\Lambda}_2^N + \check{\Lambda}_2^{N\mathrm{T}} + \check{\Lambda}_4^N + \check{\Lambda}_{11}^N + \check{\Lambda}_{11}^{N\mathrm{T}} + \check{\Lambda}_{12}^N + \check{\Lambda}_{12}^{N\mathrm{T}} + \check{\Lambda}_{22}^N. \tag{64}$$

Substituting (61), (62), and (63) into (60) gives

$$\begin{aligned} J_i(\check{u}_i, \check{u}_{-i}) =& \mu_0^{\mathrm{T}}\check{Y}^N(0)\mu_0 + \operatorname{Tr}[\check{\Lambda}_1^N(0)\Sigma_0^i] + (1/N^2)\sum_{j\in\mathcal{N}_{-i}} \operatorname{Tr}[\check{\Lambda}_3^N(0)\Sigma_0^j] \\ &+ \mu_0^{\mathrm{T}}\{-(\check{\Lambda}_2^N(0) + \check{\Lambda}_2^{N\mathrm{T}}(0))/N + \check{\Lambda}_3^N(0)(N-1)/N^2 \\ &+ \check{\Lambda}_4^N(0)(2-3N)/N^2 - (\check{\Lambda}_{12}^N(0) + \check{\Lambda}_{12}^{N\mathrm{T}}(0))/N\}\mu_0. \end{aligned}$$

6 Decentralized $O(1/N)$-Nash Equilibrium Strategies

In this section we show that the set of decentralized strategies in (44) has an $O(1/N)$-Nash equilibrium property. More precisely, when the game (1)–(2) is asymptotically solvable and all other players take the decentralized strategies (44), the extra benefit that a player obtains by unilaterally deviating from the strategy (44) is at most $O(1/N)$.

Theorem 6.1 *Under Assumptions* 5.1 *and* 5.2*, the set of decentralized strategies* $(\check{u}_1, \check{u}_2, \cdots, \check{u}_N)$ *given by* (44) *is an* $O(1/N)$*-Nash equilibrium of the Nash game* (1)–(2)*, i.e.,*

$$J_i(\check{u}_i, \check{u}_{-i}) \le J_i(u_i, \check{u}_{-i}) + O(1/N), \quad \forall 1 \le i \le N, \tag{65}$$

where u_i *is any admissible control under CLPS information such that the closed-loop system under* $(u_i, \widehat{u}_{-i})$ *has a well defined solution.*

We will prove Theorem 6.1 after some technical preparations. Without loss of generality, we prove (65) for player $\mathcal{A}_1$. Suppose that players $\mathcal{A}_i$, $2 \le i \le N$, use decentralized strategies given by (44). Player $\mathcal{A}_1$ seeks its best response strategy u_1^b with respect to $\check{u}_{-1}$ so that $J_1(u_1^b, \check{u}_{-1}) = \inf_{u_1} J_1(u_1, \widehat{u}_{-1})$, where J_1 is defined by (2). This leads to the optimal control problem with dynamics

$$\begin{aligned} dX(t) =& \big[\boldsymbol{A}X + \widehat{\boldsymbol{B}}_1 u_1 - \widehat{\boldsymbol{B}}_{-1}(\widehat{\Theta}X + \widehat{\Theta}_1\overline{X})\big]dt + \boldsymbol{B}_1 u_1 dW_1 \\ &- \sum_{i=2}^N \boldsymbol{B}_i(\Theta \boldsymbol{e}_i X + \Theta_1\overline{X})dW_i + \boldsymbol{B}_0\Big(u_1 - \sum_{i=2}^N(\Theta \boldsymbol{e}_i X + \Theta_1\overline{X})\Big)dW_0, \end{aligned}$$

where we denote $\widehat{\boldsymbol{B}}_{-1} = (0, \widehat{\boldsymbol{B}}_2, \widehat{\boldsymbol{B}}_3, \cdots, \widehat{\boldsymbol{B}}_N)$ and the mean field limit state $\overline{X}$ follows the dynamics (43). The best response u_1^b is to be determined.

We employ a dynamic programming approach to solve player $\mathcal{A}_1$'s optimal control problem. Let $V_1^b(t, \boldsymbol{x}, \overline{x})$ be the value function of $\mathcal{A}_1$ with initial state $(X(t), \overline{X}(t)) = (\boldsymbol{x}, \overline{x})$, associated with the cost $J_1(u_1, \check{u}_{-1})$. Now $V_1^b(t, \boldsymbol{x}, \overline{x})$ is formally solved from the following dynamic programming equation:

$$\begin{aligned} -\frac{\partial V_1^b}{\partial t} = \min_{u_1\in\mathbb{R}^{n_1}} \Big[& \frac{\partial^{\mathrm{T}} V_1^b}{\partial \boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x} + \widehat{\boldsymbol{B}}_1 u_1 - \widehat{\boldsymbol{B}}_{-1}(\widehat{\Theta}\boldsymbol{x} + \widehat{\Theta}_1\overline{x})) \\ &+ \frac{\partial^{\mathrm{T}} V_1^b}{\partial \overline{x}}(A + G - B(\Theta + \Theta_1))\overline{x} \\ &+ \frac{1}{2}\Big(u_1 - \sum_{i=2}^N(\Theta \boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x})\Big)^{\mathrm{T}} \boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\Big(u_1 - \sum_{i=2}^N(\Theta \boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x})\Big) \\ &+ \frac{1}{2}(\boldsymbol{B}_1 u_1)^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_1 u_1 + \frac{1}{2}\sum_{i=2}^N(\boldsymbol{B}_i(\Theta \boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x}))^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_i(\Theta \boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x}) \\ &+ \frac{1}{2}(B_0(\Theta + \Theta_1)\overline{x})^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \overline{x}^2}B_0(\Theta + \Theta_1)\overline{x} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_1\boldsymbol{x} + u_1^{\mathrm{T}}Ru_1 \\ &- \Big(u_1 - \sum_{i=2}^N(\Theta \boldsymbol{e}_i\boldsymbol{x} + \Theta_1\overline{x})\Big)^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}\partial \overline{x}}B_0(\Theta + \Theta_1)\overline{x}\Big], \end{aligned} \tag{66}$$

$$V_1^b(T, \boldsymbol{x}, \overline{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_{1f}\boldsymbol{x}, \quad t \in [0, T],\ \boldsymbol{x} \in \mathbb{R}^{Nn},\ \overline{x} \in \mathbb{R}^n.$$

The first order condition with respect to u_1 gives

$$
\begin{aligned}
u_1^b = &-\frac{1}{2}\left(R+\frac{1}{2}\boldsymbol{B}_1^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_1+\frac{1}{2}\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\right)^{-1} \\
&\cdot\left[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\frac{\partial V_1^b}{\partial \boldsymbol{x}}-\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})-\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}\partial \overline{x}}B_0(\Theta+\Theta_1)\overline{x}\right].
\end{aligned} \tag{67}
$$

We substitute (67) into (66) to obtain

$$
\begin{aligned}
-\frac{\partial V_1^b}{\partial t} = &-\left[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\frac{\partial V_1^b}{\partial \boldsymbol{x}}-\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})-\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}\partial \overline{x}}B_0(\Theta+\Theta_1)\overline{x}\right]^{\mathrm{T}} \\
&\cdot\frac{1}{4}\left(R+\frac{1}{2}\boldsymbol{B}_1^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_1+\frac{1}{2}\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\right)^{-1} \\
&\cdot\left[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\frac{\partial V_1^b}{\partial \boldsymbol{x}}-\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})-\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}\partial \overline{x}}B_0(\Theta+\Theta_1)\overline{x}\right] \\
&+\frac{\partial^{\mathrm{T}} V_1^b}{\partial \boldsymbol{x}}\left(\boldsymbol{A}\boldsymbol{x}-\sum_{i=2}^{N}\widehat{\boldsymbol{B}}_i(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})\right)+\frac{\partial^{\mathrm{T}} V_1^b}{\partial \overline{x}}(A+G-B(\Theta+\Theta_1))\overline{x} \\
&+\frac{1}{2}\left(\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})\right)^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_0\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x}) \\
&+\frac{1}{2}\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})^{\mathrm{T}}\boldsymbol{B}_i^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}^2}\boldsymbol{B}_i(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})+\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_1\boldsymbol{x} \\
&+\frac{1}{2}(B_0(\Theta+\Theta_1)\overline{x})^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \overline{x}^2}B_0(\Theta+\Theta_1)\overline{x} \\
&+\left(\boldsymbol{B}_0\sum_{i=2}^{N}(\Theta \boldsymbol{e}_i\boldsymbol{x}+\Theta_1\overline{x})\right)^{\mathrm{T}}\frac{\partial^2 V_1^b}{\partial \boldsymbol{x}\partial \overline{x}}B_0(\Theta+\Theta_1)\overline{x},
\end{aligned} \tag{68}
$$

$$
V_1^b(T,\boldsymbol{x},\overline{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}_{1f}\boldsymbol{x}.
$$

Assume V_1^b takes the form

$$
V_1^b(t,\boldsymbol{x},\overline{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}_1^b(t)\boldsymbol{x}+2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}_{12}^b(t)\overline{x}+\overline{x}^{\mathrm{T}}\boldsymbol{P}_2^b(t)\overline{x}. \tag{69}
$$

We denote $\boldsymbol{I}_{-1}=(0,I_n,I_n,\cdots,I_n)\in\mathbb{R}^{n\times Nn}$, and substitute (69) into (68) to obtain ODEs for $\boldsymbol{P}_1^b$, $\boldsymbol{P}_{12}^b$ and $\boldsymbol{P}_2^b$:

$$
\begin{cases}
-\dot{\boldsymbol{P}}_1^b = -(\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_1^b-\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1})^{\mathrm{T}}(R+\boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_1+\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0)^{-1} \\
\qquad\cdot(\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_1^b-\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1})+\boldsymbol{P}_1^b(\boldsymbol{A}-\widehat{\boldsymbol{B}}_{-1}\widehat{\Theta}) \\
\qquad+(\boldsymbol{A}-\widehat{\boldsymbol{B}}_{-1}\widehat{\Theta})^{\mathrm{T}}\boldsymbol{P}_1^b+(\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1})^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1} \\
\qquad+\sum_{k=2}^{N}(\boldsymbol{B}_k\Theta\boldsymbol{e}_k)^{\mathrm{T}}\boldsymbol{P}_1^b(\boldsymbol{B}_k\Theta\boldsymbol{e}_k)+\boldsymbol{Q}_1, \\
\boldsymbol{P}_1^b(T)=\boldsymbol{Q}_1, \quad R+\boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{P}_1^b(t)\boldsymbol{B}_1+\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b(t)\boldsymbol{B}_0>0, \quad \forall t\in[0,T],
\end{cases} \tag{70}
$$

$$\begin{cases} -\dot{\boldsymbol{P}}_{12}^b = (\boldsymbol{A} - \widehat{\boldsymbol{B}}_{-1}\widehat{\Theta})^{\mathrm{T}}\boldsymbol{P}_{12}^b + \boldsymbol{P}_{12}^b(A + G - B(\Theta + \Theta_1)) \\ \qquad\quad +(\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1})^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1 - \boldsymbol{P}_1^b\widehat{\boldsymbol{B}}_{-1}\widehat{\Theta}_1 \\ \qquad\quad +\sum_{k=2}^{N}(\boldsymbol{B}_k\Theta\boldsymbol{e}_k)^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_k\Theta_1 + (\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1})^{\mathrm{T}}\boldsymbol{P}_{12}^bB_0(\Theta + \Theta_1) \\ \qquad\quad -(\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_1^b - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\Theta\boldsymbol{I}_{-1})^{\mathrm{T}}(R + \boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_1 + \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0)^{-1} \\ \qquad\qquad \cdot[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_{12}^b - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1 - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_{12}^bB_0(\Theta + \Theta_1)], \\ \boldsymbol{P}_{12}^b(T) = 0, \end{cases} \tag{71}$$

$$\begin{cases} -\dot{\boldsymbol{P}}_2^b = -[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_{12}^b - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1 - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_{12}^bB_0(\Theta + \Theta_1)]^{\mathrm{T}} \\ \qquad\quad \cdot(R + \boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_1 + \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0)^{-1} \\ \qquad\quad \cdot[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_{12}^b - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1 - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_{12}^bB_0(\Theta + \Theta_1)] \\ \qquad\quad -\boldsymbol{P}_{12}^{b\mathrm{T}}\widehat{\boldsymbol{B}}_{-1}\widehat{\Theta}_1 - \widehat{\Theta}_1^{\mathrm{T}}\widehat{\boldsymbol{B}}_{-1}^{\mathrm{T}}\boldsymbol{P}_{12}^b + (\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1)^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1 \\ \qquad\quad +\boldsymbol{P}_2^b(A + G - B(\Theta + \Theta_1)) + (A + G - B(\Theta + \Theta_1))^{\mathrm{T}}\boldsymbol{P}_2^b \\ \qquad\quad +\sum_{k=2}^{N}(\boldsymbol{B}_k\Theta_1)^{\mathrm{T}}\boldsymbol{P}_1^b(\boldsymbol{B}_k\Theta_1) + (\Theta + \Theta_1)^{\mathrm{T}}B_0^{\mathrm{T}}\boldsymbol{P}_2^bB_0(\Theta + \Theta_1) \\ \qquad\quad +(\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1)^{\mathrm{T}}\boldsymbol{P}_{12}^bB_0(\Theta + \Theta_1) + (\Theta + \Theta_1)^{\mathrm{T}}B_0^{\mathrm{T}}\boldsymbol{P}_{12}^{b\mathrm{T}}\boldsymbol{B}_0\boldsymbol{I}_{-1}\widehat{\Theta}_1, \\ \boldsymbol{P}_2^b(T) = 0. \end{cases} \tag{72}$$

Proposition 6.2 *Suppose that Assumption* 5.1 *holds and that* (70) *has a solution* $\boldsymbol{P}_1^b$ *on* $[0, T]$. *Then we may uniquely solve* (71)–(72), *and the best response strategy for* $\mathcal{A}_1$ *is*

$$\begin{aligned} u_1^b(t) = &-\left(R + \boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_1 + \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\right)^{-1}\Bigg[\widehat{\boldsymbol{B}}_1^{\mathrm{T}}(\boldsymbol{P}_1^bX(t) + \boldsymbol{P}_{12}^b\overline{X}(t)) \\ &- \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1^b\boldsymbol{B}_0\sum_{i=2}^{N}(\Theta\boldsymbol{e}_iX(t) + \Theta_1\overline{X}(t)) - \boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_{12}^bB_0(\Theta + \Theta_1)\overline{X}(t)\Bigg]. \end{aligned} \tag{73}$$

Proof If (70) admits a (unique) solution $\boldsymbol{P}_1^b$ on $[0, T]$, then we can substitute $\boldsymbol{P}_1^b$ into (71) and solve a first order linear ODE for a unique $\boldsymbol{P}_{12}^b$. Given $(\boldsymbol{P}_1^b, \boldsymbol{P}_{12}^b)$, $\boldsymbol{P}_2^b$ is again solved from a linear ODE. Note that the LQ optimal control problem of player $\mathcal{A}_1$ has its Riccati equation given by (70)–(72). It then follows from [44, Theorem 6.6.1] that player $\mathcal{A}_1$'s optimal control problem is solvable with the optimal control given by (73). ∎

We will later show that for all sufficiently large N, (70) indeed has a solution on $[0, T]$ (see Lemma 6.8). The next lemma is parallel to Lemma 4.1.

Lemma 6.3 *Suppose* (70) *has a solution* $\boldsymbol{P}_1^b$ *on* $[0, T]$. *Then for* (70) *and* (71), $\boldsymbol{P}_1^b$ *and*

$\boldsymbol{P}_{12}^b$ *have the representations*

$$\boldsymbol{P}_1^b = \begin{bmatrix} \Pi_1^{bN} & \Pi_2^{bN} & \Pi_2^{bN} & \cdots & \Pi_2^{bN} \\ (\Pi_2^{bN})^{\mathrm{T}} & \Pi_3^{bN} & \Pi_4^{bN} & \cdots & \Pi_4^{bN} \\ (\Pi_2^{bN})^{\mathrm{T}} & \Pi_4^{bN} & \Pi_3^{bN} & \cdots & \Pi_4^{bN} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (\Pi_2^{bN})^{\mathrm{T}} & \Pi_4^{bN} & \Pi_4^{bN} & \cdots & \Pi_3^{bN} \end{bmatrix}, \tag{74}$$

$$\boldsymbol{P}_{12}^b = \Big[(\Pi_{11}^{bN})^{\mathrm{T}},\ (\Pi_{12}^{bN})^{\mathrm{T}},\ \cdots,\ (\Pi_{12}^{bN})^{\mathrm{T}}\Big]^{\mathrm{T}}, \tag{75}$$

where $\Pi_1^{bN}(t)$, $\Pi_3^{bN}(t)$, $\Pi_4^{bN}(t) \in \mathcal{S}^n$, *and* $\Pi_2^{bN}(t)$, $\Pi_{11}^{bN}(t)$, $\Pi_{12}^{bN}(t) \in \mathbb{R}^{n\times n}$.

Proof The proof is similar to that of Lemma 4.1, and is thus omitted here. ∎

We define new variables:

$$\begin{cases} \Lambda_1^{bN} = \Pi_1^{bN}, \quad \Lambda_2^{bN} = N\Pi_2^{bN}, \quad \Lambda_3^{bN} = N^2\Pi_3^{bN}, \quad \Lambda_4^{bN} = N^2\Pi_4^{bN}, \\ \Lambda_{11}^{bN} = \Pi_{11}^{bN}, \quad \Lambda_{12}^{bN} = N\Pi_{12}^{bN}, \quad \Lambda_{22}^{bN} = \boldsymbol{P}_2^b, \end{cases} \tag{76}$$

and suppose (70) has a solution $\boldsymbol{P}_1^b$ on $[0,T]$. We substitute (74) and (75) into (70)–(72) and take a change of variables by (76) to obtain (under the additional condition that $R+B_1^{\mathrm{T}}\Lambda_1^{bN}(t)B_1 > 0$) the following ODEs:

$$\begin{cases} \dot{\Lambda}_1^{bN} = \Lambda_1^{bN}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_1^{bN} - \Lambda_1^{bN}A - A^{\mathrm{T}}\Lambda_1^{bN} - Q + g_1^{bN}, \\ \Lambda_1^{bN}(T) = (I - \Gamma_f^{\mathrm{T}}/N)Q_f(I-\Gamma_f/N), \\ R + B_1^{\mathrm{T}}\Lambda_1^{bN}(t)B_1 > 0, \quad \forall t\in[0,T], \end{cases} \tag{77}$$

$$\begin{cases} \dot{\Lambda}_2^{bN} = \Lambda_1^{bN}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_2^{bN} - (\Lambda_1^{bN}+\Lambda_2^{bN})G \\ \qquad\quad -A^{\mathrm{T}}\Lambda_2^{bN} - \Lambda_2^{bN}(A-B\Theta) + Q\Gamma + g_2^{bN}, \\ \Lambda_2^{bN}(T) = -(I-\Gamma_f^{\mathrm{T}}/N)Q_f\Gamma_f, \end{cases} \tag{78}$$

$$\begin{cases} \dot{\Lambda}_3^{bN} = (\Lambda_2^{bN})^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_2^{bN} - (\Lambda_2^{bN}+\Lambda_4^{bN})^{\mathrm{T}}G \\ \qquad\quad -G^{\mathrm{T}}(\Lambda_2^{bN}+\Lambda_4^{bN}) - \Lambda_3^{bN}(A-B\Theta) - (A-B\Theta)^{\mathrm{T}}\Lambda_3^{bN} \\ \qquad\quad -\Theta^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_1^{bN}+\Lambda_2^{bN}+(\Lambda_2^{bN})^{\mathrm{T}}+\Lambda_4^{bN})B_0\Theta \\ \qquad\quad -\Theta^{\mathrm{T}}B_1^{\mathrm{T}}\Lambda_3^{bN}B_1\Theta - \Gamma^{\mathrm{T}}Q\Gamma + g_3^{bN}, \\ \Lambda_3^{bN}(T) = \Gamma_f^{\mathrm{T}}Q_f\Gamma_f, \end{cases} \tag{79}$$

$$\begin{cases}\dot{\Lambda}_4^{bN}=(\Lambda_2^{bN})^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_2^{bN}-(\Lambda_2^{bN})^{\mathrm{T}}G-G^{\mathrm{T}}\Lambda_2^{bN}\\ \qquad -\Lambda_4^{bN}(A+G-B\Theta)-(A+G-B\Theta)^{\mathrm{T}}\Lambda_4^{bN}\\ \qquad -\Theta^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_1^{bN}+\Lambda_2^{bN}+(\Lambda_2^{bN})^{\mathrm{T}}+\Lambda_4^{bN})B_0\Theta\\ \qquad -\Gamma^{\mathrm{T}}Q\Gamma+g_4^{bN},\\ \Lambda_4^{bN}(T)=\Gamma_f^{\mathrm{T}}Q_f\Gamma_f,\end{cases} \tag{80}$$

$$\begin{cases}\dot{\Lambda}_{11}^{bN}=\Lambda_1^{bN}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_{11}^{bN}+\Lambda_2^{bN}B\Theta_1-A^{\mathrm{T}}\Lambda_{11}^{bN}\\ \qquad -\Lambda_{11}^{bN}(A+G-B(\Theta+\Theta_1))+g_{11}^{bN},\\ \Lambda_{11}^{bN}(T)=0,\end{cases} \tag{81}$$

$$\begin{cases}\dot{\Lambda}_{12}^{bN}=(\Lambda_2^{bN})^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_{11}^{bN}-G^{\mathrm{T}}(\Lambda_{11}^{bN}+\Lambda_{12}^{bN})\\ \qquad -(A^{\mathrm{T}}-\Theta^{\mathrm{T}}B^{\mathrm{T}})\Lambda_{12}^{bN}-\Lambda_{12}^{bN}(A+G-B(\Theta+\Theta_1))\\ \qquad -\Theta^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_1^{bN}+\Lambda_2^{bN}+(\Lambda_2^{bN})^{\mathrm{T}}+\Lambda_4^{bN})B_0\Theta_1\\ \qquad -\Theta^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_{11}^{bN}+\Lambda_{12}^{bN})B_0(\Theta+\Theta_1)+\Lambda_4^{bN}B\Theta_1+g_{12}^{bN},\\ \Lambda_{12}^{bN}(T)=0,\end{cases} \tag{82}$$

$$\begin{cases}\dot{\Lambda}_{22}^{bN}=(\Lambda_{11}^{bN})^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}\Lambda_{11}^{bN}\\ \qquad -\Lambda_{22}^{bN}(A+G-B(\Theta+\Theta_1))+(\Lambda_{12}^{bN})^{\mathrm{T}}B\Theta_1\\ \qquad -(A+G-B(\Theta+\Theta_1))^{\mathrm{T}}\Lambda_{22}^{bN}+\Theta_1^{\mathrm{T}}B^{\mathrm{T}}\Lambda_{12}^{bN}\\ \qquad -\Theta_1^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_1^{bN}+\Lambda_2^{bN}+(\Lambda_2^{bN})^{\mathrm{T}}+\Lambda_4^{bN})B_0\Theta_1\\ \qquad -(\Theta+\Theta_1)^{\mathrm{T}}B_0^{\mathrm{T}}\Lambda_{22}^{bN}B_0(\Theta+\Theta_1)\\ \qquad -\Theta_1^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_{11}^{bN}+\Lambda_{12}^{bN})B_0(\Theta+\Theta_1)\\ \qquad -(\Theta+\Theta_1)^{\mathrm{T}}B_0^{\mathrm{T}}(\Lambda_{11}^{bN}+\Lambda_{12}^{bN})^{\mathrm{T}}B_0\Theta_1+g_{22}^{bN},\\ \Lambda_{22}^{bN}(T)=0.\end{cases} \tag{83}$$

The perturbation terms g_k^{bN}, $1\le k\le 4$, g_{11}^{bN}, g_{12}^{bN} and g_{22}^{bN} are functions of $(N,\Lambda_1^{bN},\Lambda_2^{bN},\cdots,\Lambda_{22}^{bN})$, and we have

$$\begin{aligned}g_1^{bN}=&\Lambda_1^{bN}B[(\mathcal{R}_1(\Lambda_1^{bN})+B_0^{\mathrm{T}}S^{bN}B_0/N^2)^{-1}-(\mathcal{R}_1(\Lambda_1^{N}))^{-1}]B^{\mathrm{T}}\Lambda_1^{bN}\\ &-(\Lambda_1^{bN}G+G^{\mathrm{T}}\Lambda_1^{bN})/N-(\Lambda_2^{bN}G+G^{\mathrm{T}}\Lambda_2^{bN})(N-1)/N^2\\ &-(\Gamma^{\mathrm{T}}Q\Gamma/N-\Gamma^{\mathrm{T}}Q-Q\Gamma)/N,\\ S^{bN}=&\Lambda_1^{bN}+(\Lambda_2^{bN}+(\Lambda_2^{bN})^{\mathrm{T}})(N-1)/N+\Lambda_3^{bN}(N-1)/N^2\\ &+\Lambda_4^{bN}(N-1)(N-2)/N^2.\end{aligned}$$

The remaining perturbation terms can be found in [41].

Let $(\Lambda_1^b,\Lambda_2^b,\cdots,\Lambda_{22}^b)$ be determined by the ODE system (104)–(110) in Appendix 2.

Lemma 6.4 *Under Assumption* 5.1 *we have*

$$\Lambda_1^b(t) = \Lambda_1(t), \tag{84}$$
$$\Lambda_2^b(t) + \Lambda_{11}^b(t) = \Lambda_2(t), \tag{85}$$
$$\zeta^b(t) = \Lambda_2(t) + \Lambda_2^{\mathrm{T}}(t) + \Lambda_4(t) \tag{86}$$

for all $t \in [0, T]$, *where*

$$\zeta^b(t) := (\Lambda_2^b + \Lambda_2^{b\mathrm{T}} + \Lambda_4^b + \Lambda_{11}^b + \Lambda_{11}^{b\mathrm{T}} + \Lambda_{12}^b + \Lambda_{12}^{b\mathrm{T}} + \Lambda_{22}^b)(t).$$

Proof (84) is already stated in the proof of Lemma 8.1. By considering the ODE of $\Lambda_2^b + \Lambda_{11}^b - \Lambda_2$ and next applying Gröwnwall's lemma, we establish $\sup_{0\le t\le T} |\Lambda_2^b(t) + \Lambda_{11}^b(t) - \Lambda_2(t)| = 0$, which implies (85).

Define $\zeta(t) = \Lambda_2(t) + \Lambda_2^{\mathrm{T}}(t) + \Lambda_4(t)$. By use of (24), (26), and (105)–(110) we write the ODEs:

$$\dot\zeta(t) = \Phi(\Lambda_2, \Lambda_4),$$
$$\dot\zeta^b(t) = \Phi^b(\Lambda_2^b, \Lambda_4^b, \Lambda_{11}^b, \Lambda_{12}^b, \Lambda_{22}^b),$$

where the two vector fields are not fully displayed but can be easily determined. Note that $\Lambda_1(t)$ and $(\Lambda_1(t), \Lambda_2(t))$ appear in Φ and Φ^b, respectively, and are treated as known functions of time. Letting $H = (\mathcal{R}_1(\Lambda_1))^{-1}$, we have

$$\begin{aligned}\Phi - \Phi^b =&(\zeta^b - \zeta)(A + G) + (A + G)^{\mathrm{T}}(\zeta^b - \zeta)\\ &+ (\Theta + \Theta_1)^{\mathrm{T}} B_0^{\mathrm{T}}(\zeta^b - \zeta)B_0(\Theta + \Theta_1)\\ &- \Lambda_1 BHB^{\mathrm{T}}(\zeta^b - \zeta) - (\zeta^b - \zeta)BHB^{\mathrm{T}}\Lambda_1\\ &- (\zeta^b - \zeta)BHB^{\mathrm{T}}\Lambda_2 + \Delta_\Phi,\end{aligned} \tag{87}$$

where we have used (84)–(85) to derive the last line to get

$$\begin{aligned}\Delta_\Phi =&\Lambda_2^{\mathrm{T}} BHB^{\mathrm{T}}\Lambda_2^{\mathrm{T}} + \Lambda_2^{\mathrm{T}} BHB^{\mathrm{T}}\Lambda_4\\ &- \Theta_1^{\mathrm{T}} B^{\mathrm{T}}(\Lambda_2^{b\mathrm{T}} + \Lambda_{11}^{b\mathrm{T}} + \Lambda_{12}^{b\mathrm{T}} + \Lambda_4^b + \Lambda_{22}^{b\mathrm{T}} + \Lambda_{12}^b).\end{aligned}$$

By use of the definition of ζ and (85), we obtain $\Delta_\Phi = \Lambda_2^{\mathrm{T}} BHB^{\mathrm{T}}(\zeta - \zeta^b)$. By the ODE of $\zeta - \zeta^b$ and Grönwall's lemma, we obtain $\sup_{t\in[0,T]} |\zeta(t) - \zeta^b(t)| = 0$. ∎

Although the system (77)–(83) has been constructed based on (70)–(72), it can stand alone for its existence analysis without using the latter.

Lemma 6.5 *Under Assumption* 5.1, *there exists* $N_1 > 0$ *such that for all* $N \ge N_1$, (77)–(83) *admits a solution* $(\Lambda_1^{bN}, \Lambda_2^{bN}, \cdots, \Lambda_{22}^{bN})$ *on* $[0, T]$ *satisfying*

$$(\mathcal{R}_1(\Lambda_1^{bN}) + B_0^{\mathrm{T}} S^{bN} B_0/N^2)(t) > \varepsilon_0 I, \quad \forall t \in [0, T], \tag{88}$$

for some small constant $\varepsilon_0 > 0$. *In addition,* $\sup_{t\in[0,T]} |\Lambda_\iota^{bN} - \Lambda_\iota^b| = O(1/N)$ *for* $\iota = 1, 2, \cdots, 22$, *where* Λ_1^b, Λ_2^b, $\cdots$, Λ_{22}^b *are given in Appendix* 2.

Proof We view (77)–(83) as a slightly perturbed version of (104)–(110). By the same thin tube method as in the sufficiency proof of Theorem 4.6, we establish the existence and uniqueness of a solution of (77)–(83) for all sufficiently large N. We may ensure (88) due to $\mathcal{R}_1(\Lambda_1^b) > 0$ for all $t \in [0,T]$ and a continuity argument. The error bound of $O(1/N)$ is obtained by applying Grönwall's lemma as in Corollary 4.8. ∎

Remark 6.6 Let ψ^N stand for any of the functions Λ_1^{bN}, Λ_2^{bN}, Λ_3^{bN}, Λ_4^{bN}, Λ_{11}^{bN}, Λ_{12}^{bN} and Λ_{22}^{bN}. Then $\sup_{N\geq N_1, 0\leq t\leq T} |\psi^N| \leq C$ for some fixed constant C.

Remark 6.7 Let h^N stand for any of the functions g_1^{bN}, g_2^{bN}, g_3^{bN}, g_4^{bN}, g_{11}^{bN}, g_{12}^{bN} and g_{22}^{bN}. Then $\sup_{t\in[0,T]} |h^N(t)| = O(1/N)$.

Lemma 6.8 *Under Assumption* 5.1*, the ODE system* (70)*–*(72) *has a solution on* $[0,T]$ *for all* $N \geq N_1$*, where* N_1 *is specified in Lemma* 6.5*.*

Proof After obtaining $(\Lambda_1^{bN}, \Lambda_2^{bN}, \cdots, \Lambda_{22}^{bN})$ by Lemma 6.5, we define $\boldsymbol{P}_1^b$ using (74) and (76). Then we can directly verify that $\boldsymbol{P}_1^b$ satisfies (70), where $R + \boldsymbol{B}_1^{\rm T}\boldsymbol{P}_1^b(t)\boldsymbol{B}_1 + \boldsymbol{B}_0^{\rm T}\boldsymbol{P}_1^b(t)\boldsymbol{B}_0 > 0$ holds for all $t \in [0,T]$ since this matrix is equal to the term $\mathcal{R}_1(\Lambda_1^{bN}) + B_0^{\rm T} S^{bN} B_0/N^2$ appearing in (77). Note that (88) holds. Then we further uniquely solve (71)–(72). ∎

Combining Lemma 6.8 with Proposition 6.2 and Lemma 6.3, we have the following facts. Under Assumption 5.1, for all sufficiently large N, the best response control problem for player $\mathcal{A}_1$ has a solution. Next, the value function of the best response control problem can be specified using (77)–(83), which has a well defined solution.

Lemma 6.9 $\sup_{t\in[0,T]} |\Lambda_1^{bN}(t) - \Lambda_1(t)| = O(1/N)$.

Proof The lemma follows from Lemma 6.5 and (84). ∎

Lemma 6.10 $\sup_{t\in[0,T]} |\check{\Lambda}_1^{N}(t) - \Lambda_1(t)| = O(1/N)$.

Proof Taking the difference of (53) and (23) gives

$$\begin{cases} \dfrac{d}{dt}(\check{\Lambda}_1^N - \Lambda_1) = -\Theta^{\rm T} B_1^{\rm T}(\check{\Lambda}_1^N - \Lambda_1)B_1\Theta - (\check{\Lambda}_1^N - \Lambda_1)(A - B\Theta) \\ \qquad\qquad -(A - B\Theta)^{\rm T}(\check{\Lambda}_1^N - \Lambda_1) - \check{g}_1^N, \\ \check{\Lambda}_1^N(T) - \Lambda_1(T) = (I - \Gamma_f^{\rm T}/N)Q_f(I - \Gamma_f/N) - Q_f. \end{cases}$$

By Remark 5.5, $\sup_{t\in[0,T]} |\check{g}_1^N(t)| = O(1/N)$. The desired result follows from Grönwall's lemma. ∎

Lemma 6.11 $\sup_{t\in[0,T]} |\Lambda_2^{bN}(t) + \Lambda_{11}^{bN}(t) - \Lambda_2(t)| = O(1/N)$.

Proof The lemma follows from Lemma 6.5 and (85). ∎

Lemma 6.12 *Let* $\check{Y}^N$ *be defined by* (64)*, and denote*

$$Y^{bN} := \Lambda_1^{bN} + \Lambda_2^{bN} + (\Lambda_2^{bN})^{\rm T} + \Lambda_4^{bN} + \Lambda_{11}^{bN} + (\Lambda_{11}^{bN})^{\rm T} + \Lambda_{12}^{bN} + (\Lambda_{12}^{bN})^{\rm T} + \Lambda_{22}^{bN}.$$

Then $\sup_{t\in[0,T]} |\check{Y}^N(t) - Y^{bN}(t)| = O(1/N)$.

Proof Combining the ODEs (53)–(59) and (77)–(83), we obtain the following ODE of $\check{Y}^N - Y^{bN}$:

$$
\begin{aligned}
&\frac{d}{dt}(\check{Y}^N - Y^{bN})\\
=&(Y^{bN} - \check{Y}^N)(A + G - B(\Theta + \Theta_1)) + (A + G - B(\Theta + \Theta_1))^{\mathrm{T}}(Y^{bN} - \check{Y}^N)\\
&+ (\Theta + \Theta_1)^{\mathrm{T}}[B_1^{\mathrm{T}}(\Lambda_1^{bN} - \Lambda_1)B_1 + B_0^{\mathrm{T}}(Y^{bN} - \check{Y}^N)B_0](\Theta + \Theta_1)\\
&- (\Lambda_1^{bN} + \Lambda_2^{bN} + \Lambda_{11}^{bN} - \Lambda_1 - \Lambda_2)^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1))^{-1}B^{\mathrm{T}}(\Lambda_1^{bN} + \Lambda_2^{bN} + \Lambda_{11}^{bN} - \Lambda_1 - \Lambda_2)\\
&- (\Lambda_1 + \Lambda_2)^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1))^{-1}B_1^{\mathrm{T}}(\Lambda_1^{bN} - \Lambda_1)B_1(\mathcal{R}_1(\Lambda_1))^{-1}B^{\mathrm{T}}(\Lambda_1 + \Lambda_2)\\
&- (\Lambda_1^{bN} + \Lambda_2^{bN} + \Lambda_{11}^{bN})^{\mathrm{T}}B(\mathcal{R}_1(\Lambda_1))^{-1}B_1^{\mathrm{T}}(\Lambda_1 - \Lambda_1^{bN})B_1\\
&\cdot (\mathcal{R}_1(\Lambda_1^{bN}))^{-1}B^{\mathrm{T}}(\Lambda_1^{bN} + \Lambda_2^{bN} + \Lambda_{11}^{bN}) + \rho^N,\\
&\check{Y}^N(T) - Y^{bN}(T) = 0,
\end{aligned}
$$

where

$$
\begin{aligned}
\rho^N :=& - (\check{g}_1^N + \check{g}_2^N + \check{g}_2^{N\mathrm{T}} + \check{g}_4^N + \check{g}_{11}^N + \check{g}_{11}^{N\mathrm{T}} + \check{g}_{12}^N + \check{g}_{12}^{N\mathrm{T}} + \check{g}_{22}^N)\\
&- [g_1^{bN} + g_2^{bN} + (g_2^{bN})^{\mathrm{T}} + g_4^{bN} + g_{11}^{bN} + (g_{11}^{bN})^{\mathrm{T}} + g_{12}^{bN} + (g_{12}^{bN})^{\mathrm{T}} + g_{22}^{bN}].
\end{aligned}
$$

The coefficients of the term $\check{Y}^N - Y^{bN}$ are bounded. By Lemmas 6.9 and 6.11, we have $\sup_{t\in[0,T]}|\Lambda_1^{bN}(t) - \Lambda_1(t)| = O(1/N)$ and $\sup_{t\in[0,T]}|\Lambda_1^{bN} + \Lambda_2^{bN} + \Lambda_{11}^{bN} - \Lambda_1 - \Lambda_2| = O(1/N)$. By Remarks 5.5 and 6.7, we have that $\sup_{t\in[0,T]}|\rho^N| = O(1/N)$. The lemma is then proven by applying Grönwall's lemma to the integral form of the ODE of $\check{Y}^N - Y^{bN}$. ∎

Proof of Theorem 6.1 When all other players $\mathcal{A}_i$, $2 \le i \le N$, take the decentralized strategies $\check{u}_{-1} = (\check{u}_2, \check{u}_3, \cdots, \check{u}_N)$, we compare the cost of player $\mathcal{A}_1$ under u_1^b with the cost under $\check{u}_1$. The cost $J_1(u_1^b, \check{u}_{-1})$ of $\mathcal{A}_1$ is

$$
\begin{aligned}
J_1(u_1^b, \check{u}_{-1}) =&\mathbb{E}\big[V_1^b(0, X(0), \overline{X}(0))\big]\\
=&\mathbb{E}\big[X^{\mathrm{T}}(0)\boldsymbol{P}_1^b(0)X(0) + 2X^{\mathrm{T}}(0)\boldsymbol{P}_{12}^b(0)\overline{X}(0) + \overline{X}^{\mathrm{T}}(0)\boldsymbol{P}_2^b(0)\overline{X}(0)\big]\\
=&\mu_0^{\mathrm{T}}Y^{bN}(0)\mu_0 + \mathrm{Tr}[\Lambda_1^{bN}(0)\Sigma_0^1] + (1/N^2)\sum_{i=2}^N \mathrm{Tr}[\Lambda_3^{bN}(0)\Sigma_0^i]\\
&+ \mu_0^{\mathrm{T}}\{-(\Lambda_2^{bN}(0) + (\Lambda_2^{bN}(0))^{\mathrm{T}})/N + \Lambda_3^{bN}(0)(N-1)/N^2\\
&+ \Lambda_4^{bN}(0)(2 - 3N)/N^2 - (\Lambda_{12}^{bN}(0) + (\Lambda_{12}^{bN}(0))^{\mathrm{T}})/N\}\mu_0.
\end{aligned} \tag{89}
$$

The cost $J_1(\check{u}_1, \check{u}_{-1})$ can be obtained from (60). Then we have

$$
\begin{aligned}
J_1(u_1^b, \check{u}_{-1}) - J_1(\check{u}_1, \check{u}_{-1}) =&\mu_0^{\mathrm{T}}[Y^{bN}(0) - \check{Y}^N(0)]\mu_0 + \mathrm{Tr}[(\Lambda_1^{bN}(0) - \check{\Lambda}_1^N(0))\Sigma_0^1]\\
&+ (1/N^2)\sum_{i=2}^N \mathrm{Tr}[(\Lambda_3^{bN}(0) - \check{\Lambda}_3^N(0))\Sigma_0^i] + O(1/N),
\end{aligned} \tag{90}
$$

where we obtain the estimate $O(1/N)$ using Remarks 5.4 and 6.6. By Lemma 6.12, we have

$$
|\mu_0^{\mathrm{T}}[Y^{bN}(0) - \check{Y}^N(0)]\mu_0| = O(1/N). \tag{91}
$$

From Lemmas 6.9 and 6.10, we have $\sup_{t\in[0,T]}|\Lambda_1^{bN}(0)-\check{\Lambda}_1^N(0)|=O(1/N)$ and thus

$$|\operatorname{Tr}[(\Lambda_1^{bN}(0)-\check{\Lambda}_1^N(0))\Sigma_0^1]|=O(1/N). \tag{92}$$

By Assumption 5.2 and Remarks 5.4 and 6.6, we have

$$(1/N^2)\Big|\sum_{i=2}^N \operatorname{Tr}[(\Lambda_3^{bN}(0)-\check{\Lambda}_3^N(0))\Sigma_0^i]\Big|=O(1/N). \tag{93}$$

It follows from (90) and (91), (92) and (93) that

$$0\le J_1(\check{u}_1,\check{u}_{-1})-J_1(u_1^b,\check{u}_{-1})=O(1/N). \tag{94}$$

Note that the term $O(1/N)$ in (94) does not depend on which player is selected to apply its best response. This completes the proof. ∎

Let u_i^b denote the best response strategy of $\mathcal{A}_i$ when all other players apply their strategies $\check{u}_{-i}$.

Theorem 6.13 *Under Assumptions* 5.1 *and* 5.2*, we have*

$$\max_{1\le i\le N}|J_i(u_i^b,\check{u}_{-i})-J_i(\widehat{u}_i,\widehat{u}_{-i})|=O(1/N), \tag{95}$$

$$\max_{1\le i\le N}|J_i(\check{u}_i,\check{u}_{-i})-J_i(\widehat{u}_i,\widehat{u}_{-i})|=O(1/N). \tag{96}$$

Proof By using the value function of the N-player Nash game, we have

$$\begin{aligned}J_i(\widehat{u}_i,\widehat{u}_{-i}) =&\mu_0^{\mathrm T}[\Lambda_1^N(0)+\Lambda_2^N(0)+\Lambda_2^{N\mathrm T}(0)+\Lambda_4^N(0)]\mu_0+\operatorname{Tr}[\Lambda_1^N(0)\Sigma_0^i] \quad (97)\\ &+(1/N^2)\sum_{i\ne j=1}^N \operatorname{Tr}[\Lambda_3^N(0)\Sigma_0^j]+\mu_0^{\mathrm T}\{-(\Lambda_2^N(0)+\Lambda_2^{N\mathrm T}(0))/N\\ &+\Lambda_3^N(0)(N-1)/N^2+\Lambda_4^N(0)(2-3N)/N^2\}\mu_0\\ =&\mu_0^{\mathrm T}[\Lambda_1^N(0)+\Lambda_2^N(0)+\Lambda_2^{N\mathrm T}(0)+\Lambda_4^N(0)]\mu_0\\ &+\operatorname{Tr}[\Lambda_1^N(0)\Sigma_0^i]+O(1/N)\\ =&\mu_0^{\mathrm T}[\Lambda_1(0)+\Lambda_2(0)+\Lambda_2^{\mathrm T}(0)+\Lambda_4(0)]\mu_0+\operatorname{Tr}[\Lambda_1(0)\Sigma_0^i]+O(1/N), \quad (98)\end{aligned}$$

where the last equality follows from Corollary 4.8. Similarly, we use (89) and Lemma 6.5 to obtain

$$\begin{aligned}J_i(u_i^b,\check{u}_{-i}) =&\mu_0^{\mathrm T}Y^{bN}(0)\mu_0+\operatorname{Tr}[\Lambda_1^{bN}(0)\Sigma_0^i]+O(1/N)\\ =&\mu_0^{\mathrm T}[\Lambda_1(0)+\zeta^b(0)]\mu_0+\operatorname{Tr}[\Lambda_1(0)\Sigma_0^i]+O(1/N).\end{aligned} \tag{99}$$

The term $O(1/N)$ in all estimates obtained above does not depend on i. By (98)–(99) and (86) in Lemma 6.4, we obtain (95), which combined with Theorem 6.1 yields (96). ∎

6.1 The General Model

Now we consider a general LQ model where D and D_0 in (1) may be nonzero and where the cost (2) is modified by using the running cost $[X_i(t) - \Gamma X^{(N)}(t) - \eta]_Q^2 + [u_i(t)]_R^2$ and the terminal cost $[X_i(T) - \Gamma_f X^{(N)}(T) - \eta_f]_{Q_f}^2$ for $\eta, \eta_f \in \mathbb{R}^n$. Then all the previous analysis in Sections 3–6 may be easily adapted to this general model.

The value function in (15) is now replaced by the form

$$V^G(t, \boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{P}_i(t) \boldsymbol{x} + 2\boldsymbol{x}^{\mathrm{T}} \boldsymbol{S}_i^G(t) + \boldsymbol{r}_i^G(t), \quad 1 \le i \le N.$$

The same ODE system (16)–(17) is used for $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$. If $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ is given on $[0, T]$, then $(\boldsymbol{S}_1^G, \boldsymbol{S}_2^G, \cdots, \boldsymbol{S}_N^G)$ is uniquely solved from a linear ODE system. Finally, given $(\boldsymbol{P}_i, \boldsymbol{S}_i^G)$, $1 \le i \le N$, on $[0, T]$, each $\boldsymbol{r}_i^G$ is again solved from a linear ODE. For this general model, Definition 4.2 about asymptotic solvability remains valid, and Theorem 4.6 still holds. The asymptotic analysis can be extended to treat $\{\boldsymbol{S}_i^G, \boldsymbol{r}_i^G, 1 \le i \le N\}$. We can accordingly determine the Nash equilibrium strategies $\widehat{u}_i^G$, $1 \le i \le N$, the decentralized strategies $\check{u}_i^G$, $1 \le i \le N$, and the best response strategy u_i^{Gb} given $\check{u}_{-i}^G$, which are further used to establish Theorems 6.1 and 6.13. We summarize the following result:

Corollary 6.14 *Under Assumptions* 5.1 *and* 5.2*, Theorems* 6.1 *and* 6.13 *still hold for the general model with parameters* (D, D_0, η, η_f).

7 Numerical Example

We present a numerical example to illustrate asymptotic solvability and individual costs. The parameter values are $A = -1$, $B = 1$, $B_0 = -2$, $B_1 = 4$, $G = 1$, $R = -1$, $Q = 8$, $\Gamma = 0.8$, $Q_f = 8$, $\Gamma_f = 0.8$, and $T = 2$. We take the initial conditions $X_i(0) = 1$ for all $i \ge 1$, and so $\overline{X}(0) = 1$.

When (23)–(24) admits a solution (Λ_1, Λ_2) on $[0, T]$, we use MATLAB ODE solver `ode45` to solve (23)–(26) to obtain the solution $(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4)$. At $t = 0$, we obtain $\Lambda_1(0) = 3.9435$, $\Lambda_2(0) = -2.3751$, $\Lambda_3(0) = 1.8351$ and $\Lambda_4(0) = 1.7786$. Figure 1 (left panel) shows that (23)–(24) admits a solution (Λ_1, Λ_2) on $[0, T]$ so that the Nash game (1)–(2) has asymptotic solvability. By the initial conditions and (98), under Nash strategies the asymptotic per agent cost is $\lim_{N\to\infty} J_i(\widehat{u}_i, \widehat{u}_{-i}) = \Lambda_1(0) + 2\Lambda_2(0) + \Lambda_4(0) = 0.9719$, which is indicated by the dashed horizonal line in Figure 1 (right panel). Figure 1 (right panel) shows that as N increases, the cost $J_i(\check{u}_i, \check{u}_{-i})$ of player $\mathcal{A}_i$ under the set of decentralized strategies approaches $\lim_{N\to\infty} J_i(\widehat{u}_i, \widehat{u}_{-i})$, as asserted by Theorem 6.13.

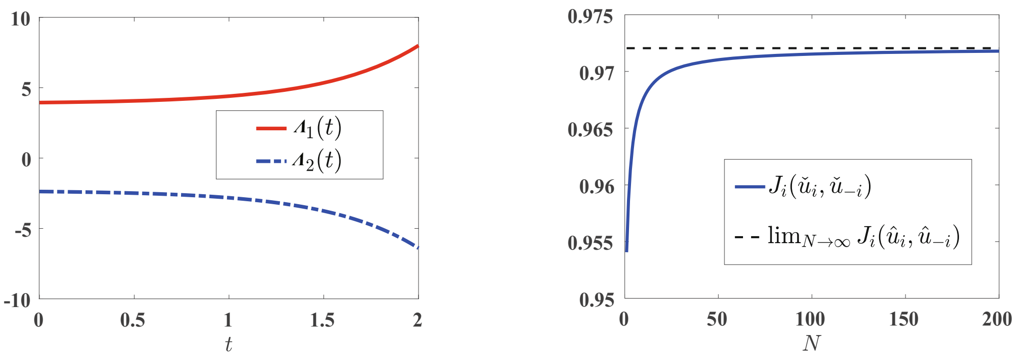

Figure 1 Left panel: (Λ_1, Λ_2) admits a solution on $[0, T]$ with $T = 2$. Right panel: The cost of player $\mathcal{A}_i$ under the set of decentralized strategies $(\check{u}_i, \check{u}_{-i})$ converges to a limit as $N \to \infty$

8 Conclusion

This paper studies an asymptotic solvability problem for LQ mean field games with controlled diffusions and indefinite cost weights. By a rescaling approach we derive a necessary and sufficient condition for asymptotic solvability. We further establish an $O(1/N)$-Nash equilibrium property for the obtained decentralized strategies.

References

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

[2] Lasry J M and Lions P L, Mean field games, *Japan. J. Math.*, 2007, **2**(1): 229–260.

[3] Bauso D, Tembine H, and Basar T, Opinion dynamics in social networks through mean-field games, *SIAM J. Control Optim.*, 2016, **54**(6): 3225–3257.

[4] Carmona R, Fouque J P, and Sun L H, Mean field games and systemic risk, *Communications in Mathematical Sciences*, 2015, **13**: 911–933.

[5] Chan P and Sircar R, Fracking, renewables, and mean field games, *SIAM Review*, 2017, **59**(3): 588–615.

[6] De Paola A, Angeli D, and Strbac G, Distributed control of micro-storage devices with mean field games, *IEEE Transactions on Smart Grid*, 2016, **7**(2): 1119–1127.

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

[8] Lachapelle A and Wolfram M T, On a mean field game approach modeling congestion and aversion in pedestrian crowds, *Transportation Research Part B: Methodological*, 2011, **45**(10): 1572–1589.

[9] Lacker D and Zariphopoulou T, Mean field and n-agent games for optimal investment under relative performance criteria, *Math. Finance*, 2019, **29**(4): 1003–1038.

[10] Laguzet L and Turinici G, Individual vaccination as Nash equilibrium in an SIR model with application to the 2009–2010 influenza A (H1N1) epidemic in France, *Bulletin of Mathematical Biology*, 2015, **77**(10): 1955–1984.

[11] Li Z, Reppen A M, and Sircar R, A mean field games model for cryptocurrency mining, arXiv: 1912.01952, 2019.

[12] Ma Z, Callaway D S, and Hiskens I A, Decentralized charging control of large populations of plug-in electric vehicles, *IEEE Transactions on Control Systems Technology*, 2013, **21**(1): 67–78.

[13] Salhab R, Malhamé R P, and Ny J L, A dynamic game model of collective choice in multiagent systems, *IEEE Trans. Autom. Control*, 2018, **63**(3): 768–782.

[14] Swiecicki I, Gobron T, and Ullmo D, Schrödinger approach to mean field games, *Physical Review Letters*, 2016, **116**(12): 128701.

[15] Wang B and Huang M, Mean field production output control with sticky prices: Nash and social solutions, *Automatica*, 2019, **100**: 90–98.

[16] Yin H, Mehta P G, Meyn S P, et al., Synchronization of coupled oscillators is a game, *IEEE Trans. Autom. Control*, 2012, **57**(4): 920–935.

[17] Bensoussan A, Frehse J, and Yam S C P, *Mean Field Games and Mean Field Type Control Theory*, Springer, New York, 2013.

[18] Caines P E, Huang M, and Malhamé R P, Mean field games, Eds. by Başar T and Zaccour G, *Handbook of Dynamic Game Theory*, Springer, Berlin, 2017, 345–372.

[19] Cardaliaguet P, *Notes on Mean Field Games*, University of Paris, Dauphine, 2013.

[20] Carmona R and Delarue F, *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer, Cham, Switzerland, 2018.

[21] Bensoussan A, Sung K C J, Yam S C P, et al., Linear-quadratic mean-field games, *J. Optim. Theory Appl.*, 2016, **169**(2): 496–529.

[22] Huang J, Wang S, and Wu Z, Backward mean-field linear-quadratic-gaussian (LQG) games: Full and partial information, *IEEE Trans. Autom. Control*, 2016, **61**(12): 3784–3796.

[23] Huang M, Caines P E, and Malhamé R P, Large-population cost-coupled LQG problems with non-uniform agents: Individual-mass behavior and decentralized ε-Nash equilibria, *IEEE Trans. Autom. Control*, 2007, **52**(9): 1560–1571.

[24] Li T and Zhang J F, Asymptotically optimal decentralized control for large population stochastic multiagent systems, *IEEE Trans. Autom. Control*, 2008, **53**(7): 1643–1660.

[25] Wang B C and Zhang J F, Mean field games for large-population multiagent systems with Markov jump parameters, *SIAM J. Control Optim.*, 2012, **50**(4): 2308–2334.

[26] Bardi M and Priuli F S, Linear-quadratic N-person and mean-field games with ergodic cost, *SIAM J. Control Optim.*, 2014, **52**(5): 3022–3052.

[27] Moon J and Başar T, Linear quadratic risk-sensitive and robust mean field games, *IEEE Trans. Autom. Control*, 2017, **62**(3): 1062–1077.

[28] Huang J and Huang M, Robust mean field linear-quadratic-Gaussian games with unknown L^2-disturbance, *SIAM J. Control Optim.*, 2017, **55**(5): 2811–2840.

[29] Tchuendom R F, Uniqueness for linear-quadratic mean field games with common noise, *Dyn. Games Appl.*, 2018, **8**(1): 199–210.

[30] Huang M, Large-population LQG games involving a major player: The Nash certainty equivalence

principle, *SIAM J. Control Optim.*, 2010, **48**(5): 3318–3353.

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

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

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

[34] Huang M and Zhou M, Linear quadratic mean field games: Asymptotic solvability and relation to the fixed point approach, *IEEE Trans. Autom. Control*, 2020, **65**(4): 1397–1412.

[35] Ma Y and Huang M, Linear quadratic mean field games with a major player: The multi-scale approach, *Automatica*, 2020, **113**(3): 108774.

[36] Huang M and Yang X, Linear quadratic mean field social optimization: Asymptotic solvability and decentralized control, *Appl. Math. Optim.*, 2021, (accepted).

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

[38] Basna R, Hilbert A, and Kolokoltsov V, An epsilon-Nash equilibrium for non-linear Markov games of mean-field-type on finite spaces, *Communications on Stochastic Analysis*, 2014, **8**(4): 449–468.

[39] Cardaliaguet P, Delarue F, Lasry J M, et al., The master equation and the convergence problem in mean field games, arXiv: 1509.02505, 2015.

[40] Huang M and Zhou M, Linear quadratic mean field games — Part I: The asymptotic solvability problem, *Proc. 23rd Internat. Symp. Math. Theory Networks and Systems*, Hong Kong, China, July, 2018, 489–495.

[41] Huang M and Yang X, Linear quadratic mean field games: Decentralized $O(1/N)$-Nash equilibria, arXiv: 2107.09168, 2021.

[42] Yosida K, *Functional Analysis*, Springer-Verlag, Berlin, 6th edition, 1980.

[43] Pham H, *Continuous-Time Stochastic Control and Optimization with Financial Applications*, Springer Science & Business Media, Berlin, 2009.

[44] Yong J and Zhou X Y, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.

Appendix 1

Proof of Theorem 3.2 To show the feedback Nash equilibrium property, we let $\mathcal{A}_k$, $k = 2, 3, \cdots, N$, take the strategies in (18) and $\mathcal{A}_1$ unilaterally improves for itself. We need to show the optimality of $\widehat{u}_1$ for minimizing $J_1(t, \boldsymbol{x}, u_1, \widehat{u}_{-1})$ for any given $(t, \boldsymbol{x})$.

Step 1 Denote the Riccati ODEs (16) in the form

$$-\dot{\boldsymbol{P}}_i = \Phi_i(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N), \quad \boldsymbol{P}_i(T) = \boldsymbol{Q}_{if}, \quad 1 \leq i \leq N. \tag{100}$$

Let $(\boldsymbol{P}_2, \boldsymbol{P}_3, \cdots, \boldsymbol{P}_N)$ still be specified by (100). We will derive a new but equivalent ODE for $\boldsymbol{P}_1$. It is necessary to do so since the best response control problem will give rise to a Riccati equation not exactly in the form of (16) with $i = 1$. For parameter $\boldsymbol{x} \in \mathbb{R}^{Nn}$, based on (10) we consider the following equation system

$$0=\widehat{\boldsymbol{B}}_i^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{x}+\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_0\sum_{k\neq i}^{N}u_k^{\boldsymbol{x}}+[\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_0+\boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_i+R]u_i^{\boldsymbol{x}},\quad 1\le i\le N,\tag{101}$$

which under (17) has a unique solution

$$\begin{aligned}u_i^{\boldsymbol{x}}&=-[R+\boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_i]^{-1}[\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_i(t)\boldsymbol{B}_0\boldsymbol{M}_0(t)+\widehat{\boldsymbol{B}}_i^{\mathrm{T}}\boldsymbol{P}_i(t)]\boldsymbol{x}\\&=:\boldsymbol{K}_i(t)\boldsymbol{x},\quad 1\le i\le N.\end{aligned}$$

Now for $i=1$, we further use (101) to obtain

$$\begin{aligned}u_1^{\boldsymbol{x}}=&-[R+\boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{P}_1(t)\boldsymbol{B}_1+\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1(t)\boldsymbol{B}_0]^{-1}\Big\{\widehat{\boldsymbol{B}}_1^{\mathrm{T}}\boldsymbol{P}_1(t)\boldsymbol{x}+\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1(t)\boldsymbol{B}_0\sum_{k=2}^{N}\boldsymbol{K}_k(t)\boldsymbol{x}\Big\}\\=:&\ \widetilde{\boldsymbol{K}}_1(t)\boldsymbol{x}.\end{aligned}$$

Since $\boldsymbol{x}$ is arbitrary, we obtain the identity

$$\boldsymbol{K}_1(t)=\widetilde{\boldsymbol{K}}_1(t),\quad \forall t\in[0,T].$$

Although the Riccati equation system (16) may be written down without using the HJB equation (5), the following observation is useful. For each $\boldsymbol{P}_i$, the vector field in (16) may be constructed from the quadratic form determined by the right hand side of (5). For illustration, take $k\neq i$. Then $\partial_{\boldsymbol{x}}^{\mathrm{T}}V_i\widehat{\boldsymbol{B}}_k\widehat{u}_k$ in (5) contributes $\boldsymbol{P}_i\widehat{\boldsymbol{B}}_k\boldsymbol{K}_k(t)+\boldsymbol{K}_k^{\mathrm{T}}(t)\widehat{\boldsymbol{B}}_k^{\mathrm{T}}\boldsymbol{P}_i$ contained in the right hand side of (16). Now for the ODE (16) of $\boldsymbol{P}_1$, whenever a term originates from $\widehat{u}_1$ so that $\boldsymbol{K}_1(t)$ is used in the vector field, we replace $\boldsymbol{K}_1(t)$ by $\widetilde{\boldsymbol{K}}_1(t)$. For example, now $\boldsymbol{M}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1\boldsymbol{B}_0\boldsymbol{M}_0$ is replaced by

$$\Big(\widetilde{\boldsymbol{K}}_1(t)+\sum_{j=2}^{N}\boldsymbol{K}_j(t)\Big)^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}\boldsymbol{P}_1\boldsymbol{B}_0\Big(\widetilde{\boldsymbol{K}}_1(t)+\sum_{j=2}^{N}\boldsymbol{K}_j(t)\Big),$$

which is ultimately expressed in terms of $(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N)$. By the above substitution of $\boldsymbol{K}_1(t)$ by $\widetilde{\boldsymbol{K}}_1(t)$, we see that $\boldsymbol{P}_1$ satisfies the new equation

$$-\dot{\boldsymbol{P}}_1=\varPhi_1^{\mathrm{new}}(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N),\quad \boldsymbol{P}_1(T)=\boldsymbol{Q}_{1f}.\tag{102}$$

The vector field $\varPhi_1^{\mathrm{new}}$ is not fully displayed here to save space, but can be easily determined. We note that the term $\boldsymbol{P}_i\widehat{\boldsymbol{B}}_k\boldsymbol{K}_k(t)+\boldsymbol{K}_k^{\mathrm{T}}(t)\widehat{\boldsymbol{B}}_k^{\mathrm{T}}\boldsymbol{P}_i$, $k\neq i$, mentioned above remains in $\varPhi_1^{\mathrm{new}}$. Then $(\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N)$ is uniquely solved from the ODE system specified by $\varPhi_1^{\mathrm{new}}$, $\varPhi_2,\cdots,\varPhi_N$.

Step 2 Consider the initial time and state pair $(t,\boldsymbol{x})$, $t\in[0,T)$. Suppose players $\mathcal{A}_k$, $k\ge 2$, apply the strategies in (18) while $\mathcal{A}_1$ minimizes the cost $J_1(t,\boldsymbol{x},u_1,\widehat{u}_{-1})$ with the initial condition $(t,\boldsymbol{x})$. The resulting optimal control u_1^{br} on $[t,T]$ is its best response. By considering the state process $X(s)$, $s\in[t,T]$ under $(u_1,\widehat{u}_{-1})$ and the cost $J_1(t,\boldsymbol{x},u_1,\widehat{u}_{-1})$, it is straightforward to determine the Riccati equation of this optimal control problem in the form

$$\dot{\boldsymbol{P}}_1^{\mathrm{br}}=\varPhi^{\mathrm{br}}(\boldsymbol{P}_1^{\mathrm{br}};\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N),\quad \boldsymbol{P}_1^{\mathrm{br}}(T)=\boldsymbol{Q}_{1f},\quad s\in[t,T],\tag{103}$$

where $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ specifies coefficients of (103) and has been solved from (102) and (100) with $i \geq 2$. By comparing the structure of (102) and (103), we see that (103) is verified by taking $\boldsymbol{P}_1^{\rm br} = \boldsymbol{P}_1$. In particular, for the inverse term $[R + \boldsymbol{B}_1^{\rm T}\boldsymbol{P}_1^{\rm br}(t)\boldsymbol{B}_1 + \boldsymbol{B}_0^{\rm T}\boldsymbol{P}_1^{\rm br}(t)\boldsymbol{B}_0]^{-1}$ appearing in (103), we have

$$R + \boldsymbol{B}_1^{\rm T}\boldsymbol{P}_1^{\rm br}(t)\boldsymbol{B}_1 + \boldsymbol{B}_0^{\rm T}\boldsymbol{P}_1^{\rm br}(t)\boldsymbol{B}_0 > 0,$$

since $R + \boldsymbol{B}_1^{\rm T}\boldsymbol{P}_1(t)\boldsymbol{B}_1 + \boldsymbol{B}_0^{\rm T}\boldsymbol{P}_1(t)\boldsymbol{B}_0 > 0$ holds and we have taken $\boldsymbol{P}_1^{\rm br} = \boldsymbol{P}_1$. By uniqueness, we see that $\boldsymbol{P}_1^{\rm br}$ must be equal to $\boldsymbol{P}_1$ on $[t, T]$. The best response is well defined on $[t, T]$, and we use (103) and $\boldsymbol{P}_1^{\rm br}$ to determine

$$u_1^{\rm br}(s) = \widetilde{\boldsymbol{K}}_1(s)X(s) = \boldsymbol{K}_1(s)X(s), \quad s \in [t, T].$$

The optimality of $u_1^{\rm br}$ may be shown by using (103) and applying completion of squares to the cost (see [44, Theorem 6.6.1]). Hence $\widehat{u}_1$ gives the best response for $\mathcal{A}_i$ on $[t, T]$.

Step 3 The same best response property holds for $\widehat{u}_i$ when any other single player $\mathcal{A}_i$ is chosen for unilateral performance improvement. We conclude that the feedback Nash equilibrium property holds. ∎

Proof of Lemma 4.1 The proof is carried out in the same manner as that of [34, Theorem 3].

Step 1 For each $2 \leq j < l \leq N$, denote $\boldsymbol{P}_i^{\dagger} = J_{jl}^{\rm T}\boldsymbol{P}_i J_{jl}$, $1 \leq i \leq N$. We have that

$$(\boldsymbol{P}_1^{\dagger}, \boldsymbol{P}_2^{\dagger}, \cdots, \boldsymbol{P}_{j-1}^{\dagger},\ \boldsymbol{P}_l^{\dagger},\ \boldsymbol{P}_{j+1}^{\dagger},\ \cdots,\ \boldsymbol{P}_{l-1}^{\dagger},\ \boldsymbol{P}_j^{\dagger},\ \boldsymbol{P}_{l+1}^{\dagger},\ \cdots,\ \boldsymbol{P}_N^{\dagger})$$

satisfies the same ODE system (16)–(17) as $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ does. Thus $J_{jl}^{\rm T}\boldsymbol{P}_1 J_{jl} = \boldsymbol{P}_1$ for all $2 \leq j < l \leq N$. Denote $\boldsymbol{P}_i = (P_{jl}^i)_{1 \leq j,l \leq N}$, where each P_{jl}^i is an $n \times n$ matrix. Then we have that

$$\begin{aligned} &P_{12}^1 = P_{13}^1 = \cdots = P_{1N}^1, \quad P_{21}^1 = P_{31}^1 = \cdots = P_{N1}^1, \\ &P_{22}^1 = P_{33}^1 = \cdots = P_{NN}^1, \quad P_{j_1 l_1}^1 = P_{j_2 l_2}^1, \quad \forall 2 \leq j_1 \neq l_1 \leq N,\ \forall 2 \leq j_2 \neq l_2 \leq N. \end{aligned}$$

This proves the representation of $\boldsymbol{P}_1$ in (19).

Step 2 For each $2 \leq j \leq N$, denote $\boldsymbol{P}_i^{\ddagger} = J_{1j}^{\rm T}\boldsymbol{P}_i J_{1j}$, $1 \leq i \leq N$. We have that

$$(\boldsymbol{P}_j^{\ddagger},\ \boldsymbol{P}_2^{\ddagger}, \cdots, \boldsymbol{P}_{j-1}^{\ddagger},\ \boldsymbol{P}_1^{\ddagger},\ \boldsymbol{P}_{j+1}^{\ddagger},\ \cdots,\ \boldsymbol{P}_N^{\ddagger})$$

and $(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ both satisfy (16)–(17). This implies $\boldsymbol{P}_j = J_{1j}^{\rm T}\boldsymbol{P}_1 J_{1j}$. ∎

Appendix 2 A Limit ODE System

We introduce the following ODE system:

$$\begin{cases} \dot{\Lambda}_1^b = \Lambda_1^b B(\mathcal{R}_1(\Lambda_1^b))^{-1}B^{\rm T}\Lambda_1^b - \Lambda_1^b A - A^{\rm T}\Lambda_1^b - Q, \\ \Lambda_1^b(T) = Q_f, \quad \mathcal{R}_1(\Lambda_1^b(t)) > 0, \quad \forall t \in [0, T], \end{cases} \tag{104}$$

$$\begin{cases} \dot{\Lambda}_2^b = \Lambda_1^b B(\mathcal{R}_1(\Lambda_1^b))^{-1}B^{\rm T}\Lambda_2^b - (\Lambda_1^b + \Lambda_2^b)G \\ \qquad\quad -A^{\rm T}\Lambda_2^b - \Lambda_2^b(A - B\Theta) + Q\Gamma, \\ \Lambda_2^b(T) = -Q_f\Gamma_f, \end{cases} \tag{105}$$

$$\begin{cases}\dot{\Lambda}_3^b = \Lambda_2^{b\mathrm{T}} B(\mathcal{R}_1(\Lambda_1^b))^{-1} B^{\mathrm{T}} \Lambda_2^b - (\Lambda_2^b + \Lambda_4^b)^{\mathrm{T}} G \\ \qquad -G^{\mathrm{T}}(\Lambda_2^b + \Lambda_4^b) - \Lambda_3^b(A - B\Theta) - (A - B\Theta)^{\mathrm{T}} \Lambda_3^b \\ \qquad -\Theta^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_1^b + \Lambda_2^b + \Lambda_2^{b\mathrm{T}} + \Lambda_4^b) B_0 \Theta \\ \qquad -\Theta^{\mathrm{T}} B_1^{\mathrm{T}} \Lambda_3^b B_1 \Theta - \Gamma^{\mathrm{T}} Q \Gamma, \\ \Lambda_3^b(T) = \Gamma_f^{\mathrm{T}} Q_f \Gamma_f, \end{cases} \tag{106}$$

$$\begin{cases}\dot{\Lambda}_4^b = \Lambda_2^{b\mathrm{T}} B(\mathcal{R}_1(\Lambda_1^b))^{-1} B^{\mathrm{T}} \Lambda_2^b - \Lambda_2^{b\mathrm{T}} G - G^{\mathrm{T}} \Lambda_2^b \\ \qquad -\Lambda_4^b(A + G - B\Theta) - (A + G - B\Theta)^{\mathrm{T}} \Lambda_4^b \\ \qquad -\Theta^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_1^b + \Lambda_2^b + \Lambda_2^{b\mathrm{T}} + \Lambda_4^b) B_0 \Theta - \Gamma^{\mathrm{T}} Q \Gamma, \\ \Lambda_4^b(T) = \Gamma_f^{\mathrm{T}} Q_f \Gamma_f, \end{cases} \tag{107}$$

$$\begin{cases}\dot{\Lambda}_{11}^b = \Lambda_1^b B(\mathcal{R}_1(\Lambda_1^b))^{-1} B^{\mathrm{T}} \Lambda_{11}^b + \Lambda_2^b B \Theta_1 - A^{\mathrm{T}} \Lambda_{11}^b \\ \qquad -\Lambda_{11}^b(A + G - B(\Theta + \Theta_1)), \\ \Lambda_{11}^b(T) = 0, \end{cases} \tag{108}$$

$$\begin{cases}\dot{\Lambda}_{12}^b = \Lambda_2^{b\mathrm{T}} B(\mathcal{R}_1(\Lambda_1^b))^{-1} B^{\mathrm{T}} \Lambda_{11}^b - G^{\mathrm{T}}(\Lambda_{11}^b + \Lambda_{12}^b) \\ \qquad -(A - B\Theta)^{\mathrm{T}} \Lambda_{12}^b - \Lambda_{12}^b(A + G - B(\Theta + \Theta_1)) \\ \qquad -\Theta^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_1^b + \Lambda_2^b + \Lambda_2^{b\mathrm{T}} + \Lambda_4^b) B_0 \Theta_1 \\ \qquad -\Theta^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_{11}^b + \Lambda_{12}^b) B_0(\Theta + \Theta_1) + \Lambda_4^b B \Theta_1, \\ \Lambda_{12}^b(T) = 0, \end{cases} \tag{109}$$

$$\begin{cases}\dot{\Lambda}_{22}^b = \Lambda_{11}^{b\mathrm{T}} B(\mathcal{R}_1(\Lambda_1^b))^{-1} B^{\mathrm{T}} \Lambda_{11}^b \\ \qquad -\Lambda_{22}^b(A + G - B(\Theta + \Theta_1)) + \Lambda_{12}^{b\mathrm{T}} B \Theta_1 \\ \qquad -(A + G - B(\Theta + \Theta_1))^{\mathrm{T}} \Lambda_{22}^b + \Theta_1^{\mathrm{T}} B^{\mathrm{T}} \Lambda_{12}^b \\ \qquad -\Theta_1^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_1^b + \Lambda_2^b + \Lambda_2^{b\mathrm{T}} + \Lambda_4^b) B_0 \Theta_1 \\ \qquad -(\Theta + \Theta_1)^{\mathrm{T}} B_0^{\mathrm{T}} \Lambda_{22}^b B_0(\Theta + \Theta_1) \\ \qquad -\Theta_1^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_{11}^b + \Lambda_{12}^b) B_0(\Theta + \Theta_1) \\ \qquad -(\Theta + \Theta_1)^{\mathrm{T}} B_0^{\mathrm{T}}(\Lambda_{11}^b + \Lambda_{12}^b)^{\mathrm{T}} B_0 \Theta_1, \\ \Lambda_{22}^b(T) = 0. \end{cases} \tag{110}$$

Under Assumption 5.1, the coefficients in (104)–(110) are defined on $[0, T]$. We may regard (104)–(110) as the limit ODE system for (77)–(83).

Lemma 8.1 *Under Assumption* 5.1*, the ODE system* (104)–(110) *admits a unique solution on* $[0, T]$.

Proof We have that (104) admits a unique solution $\Lambda_1^b = \Lambda_1$ on $[0, T]$. With Λ_1^b obtained from solving (104), (105) is a first order linear ODE and admits a unique solution Λ_2^b on $[0, T]$. Given $(\Lambda_1^b, \Lambda_2^b)$ on $[0, T]$, the ODE system (106)–(109) is a first order linear ODE system and admits a unique solution $(\Lambda_3^b, \Lambda_4^b, \cdots, \Lambda_{12}^b)$ on $[0, T]$. Finally, we further uniquely solve (110) on $[0, T]$. ∎