games

MDPI

Article

Linear–Quadratic Mean-Field-Type Games: A Direct Method

Tyrone E. Duncan [1] and Hamidou Tembine [2,*]

[1] Department of Mathematics, University of Kansas, Lawrence, KS 66044, USA; duncan@math.ku.edu
[2] Learning and Game Theory Laboratory, New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, UAE
* Correspondence: tembine@nyu.edu

Received: 4 January 2018; Accepted: 31 January 2018; Published: 12 February 2018

Abstract: In this work, a multi-person mean-field-type game is formulated and solved that is described by a linear jump-diffusion system of mean-field type and a quadratic cost functional involving the second moments, the square of the expected value of the state, and the control actions of all decision-makers. We propose a direct method to solve the game, team, and bargaining problems. This solution approach does not require solving the Bellman–Kolmogorov equations or backward–forward stochastic differential equations of Pontryagin's type. The proposed method can be easily implemented by beginners and engineers who are new to the emerging field of mean-field-type game theory. The optimal strategies for decision-makers are shown to be in a state-and-mean-field feedback form. The optimal strategies are given explicitly as a sum of the well-known linear state-feedback strategy for the associated deterministic linear–quadratic game problem and a mean-field feedback term. The equilibrium cost of the decision-makers are explicitly derived using a simple direct method. Moreover, the equilibrium cost is a weighted sum of the initial variance and an integral of a weighted variance of the diffusion and the jump process. Finally, the method is used to compute global optimum strategies as well as saddle point strategies and Nash bargaining solution in state-and-mean-field feedback form.

Keywords: Nash bargaining solution; mean-field equilibrium; variance; direct method

1. Introduction

In 1952, Markowitz proposed a paradigm for dealing with risk issues concerning choices which involve many possible financial instruments [1]. Formally, it deals with only two discrete time periods (e.g., "now" and "3 months from now"), or equivalently, one accounting period (e.g., "3 months"). In this scheme, the goal of an Investor is to select the portfolio of securities that will provide the best distribution of future consumption, given their investment budget. Two measures of the prospects provided by such a portfolio are assumed to be sufficient for evaluating its desirability: the expected value at the end of the accounting period and the standard deviation or its square, the variance, of that value. If the initial investment budget is positive, there will be a one-to-one relationship between these end-of-period measures and comparable measures relating to the percentage change in value, or return over the period. Thus, Markowitz' approach is often framed in terms of the expected return of a portfolio and its standard deviation of return, with the latter serving as a measure of risk. A typical example of risk in the current market is the evolution of the prices [2,3] of the cryptocurrencies (bitcoin, litecoin, ethereum, dash, etc). The Markowitz paradigm (also termed as mean-variance paradigm) is often characterized as dealing with portfolio risk and (expected) return [4,5]. We address this problem when several entities are involved. Game problems in which the state dynamics is given by a linear stochastic system with a Brownian motion and a cost functional that is quadratic in the state and the control are often called linear–quadratic–Gaussian (LQG) games. For the continuous time LQG game

problem with positive coefficients, the optimal strategy is a linear state-feedback strategy which is identical to an optimal control for the corresponding deterministic linear–quadratic game problem, where the Brownian motion is replaced by the zero process. Moreover, the equilibrium cost only differs from the deterministic game problem's equilibrium cost by the integral of a function of time. For LQG control and LQG zero-sum games, it can be shown that a simple square completion method provides an explicit solution to the problem. It was successfully developed and applied by Duncan et al. [6–11] in the mean-field-free case. Interestingly, the method can be used beyond the class of LQG framework. Moreover, Duncan et al. extended the direct method to more general noises, including fractional Brownian noises and some non-quadratic cost functionals on spheres, torus, and more general spaces.

The main goal of this work is to investigate whether these techniques can be used to solve mean-field-type game problems which are non-standard problems [12]. To do so, we modify the state dynamics to include mean-field terms which are (i) the expected value of the state, (ii) the expected value of the control-actions, in the drift function. We also modify the instant cost and terminal cost function to include (iii) the square of the expected values of the state and (iv) the square of the expected values of the control action. When the state dynamics and/or the cost functional involve a mean-field term (such as the expected value of the state and/or expected values of the control actions), the game is said to be an LQG game of mean-field type, or MFT-LQG. We aim to study the behavior of such MFT-LQG game problems when mean-field terms are involved. If in addition the state dynamics is driven by a jump-diffusion process, then the problem is termed as an MFT-LQJD game problem.

For such game problems, various solution methods such as the stochastic maximum principle (SMP) ([12]) and the dynamic programming principle (DPP) with Hamilton–Jacobi–Bellman–Isaacs equation and Fokker–Planck–Kolmogorov equation have been proposed [12–14]. Most studies illustrated these solution methods in the linear–quadratic game with an infinite number of decision-makers [15–21]. These works assume indistinguishability within classes, and the cost functions were assumed to be identical or invariant per permutation of decision-makers indexes. Note that the indistinguishability assumption is not fulfilled for many interesting problems, such as variance reduction or and risk quantification problems, in which decision-makers have different sensitivity towards the risk. One typical and practical example is to consider an energy-efficient multi-level building in which every resident has its own comfort zone temperature and aims to use the Heating, ventilation, and air conditioning (HVAC) system to be closer to its comfort temperature and to maintain it within its own comfort zone. This problem clearly does not satisfy the indistinguishability assumption used in the previous works on mean-field games. Therefore, it is reasonable to look at the problem beyond the indistinguishability assumption. Here we drop these assumptions and solve the problem directly with an arbitrary finite number of decision-makers. In the LQ-mean-field-type game problems, the state process can be modeled by a set of linear stochastic differential equations of McKean–Vlasov, and the preferences are formalized by quadratic cost functions with mean-field terms. These game problems are of practical interest, and a detailed exposition of this theory can be found in [7,12,22–25]. The popularity of these game problems is due to practical considerations in signal processing, pattern recognition, filtering, prediction, economics, and management science [26–29].

To some extent, most of the risk-neutral versions of these optimal controls are analytically and numerically solvable [6,7,9,11,24]. On the other hand, the linear quadratic robust setting naturally appears if the decision makers' objective is to minimize the effect of a small perturbation and related variance of the optimally controlled nonlinear process. By solving a linear–quadratic game problem of mean-field type, and using the implied optimal control actions, decision-makers can significantly reduce the variance (and the cost) incurred by this perturbation. The variance reduction and minimax problems have very interesting applications in risk quantification problems under adversarial attacks and in security issues in interdependent infrastructures and networks [27,30–33]. Table 1 summarizes some recent developments in MF-LQ-related games.

Table 1. Some recent developments on mean-field-type linear–quadratic–Gaussian (MF-LQG)-related games.

Feature	State of-the-Art	This Work
Jump		yes
Diffusion	[15,16]	yes
Mean-Field Type	[12,31,32]	yes
One decision-maker	[12]	yes
Two or more decision-makers	[31,32]	yes
State-MF	[12]	yes
Control-Action-MF		yes
Bargaining		yes
Anonymity	[15,16]	relaxed
Indistinguishability	[15,16]	relaxed

In this work, we propose a simple argument that gives the best-response strategy and the Nash equilibrium cost for a class of MFT-LQJD games without the use of the well-known solution methods (SMP and DPP). We apply the square completion method in the risk-neutral mean-field-type game problems. It is shown that this method is well-suited to MF-LQJD games, as well as to variance reduction performance functionals. Applying the solution methodology related to the DPP or the SMP requires an involved (stochastic) analysis and convexity arguments to generate necessary and sufficient optimality criteria. We avoid all of this with this method.

1.1. Contribution of This Article

Our contribution can be summarized as follows. We formulate and solve a mean-field-type game described by a linear jump-diffusion dynamics and a mean-field-dependent quadratic or robust-quadratic cost functional for each generic decision-maker. The optimal strategies for the decision-makers are given semi-explicitly using a simple and direct method based on square completion, suggested by Duncan et al. (e.g., [7–9]) for the mean-field-free case. This approach does not use the well-known solution methods such as the stochastic maximum principle and the dynamic programming principle with Hamilton–Jacobi–Bellman–Isaacs equation and Fokker–Planck–Kolmogorov equation. It does not require extended backward–forward integro-partial differential equations (IPDEs) to solve the problem. In the risk-neutral linear–quadratic mean-field-type game, we show that there is generally a best response strategy to the mean of the state, and provide a sufficient condition of existence of mean-field Nash equilibrium. We also provide a global optimum solution to the problem in the case of full cooperation between the decision-makers. This approach gives a basic insight into the solution by providing a simple explanation for the additional term in the robust Riccati equation, compared to the risk-neutral Riccati equation. Sufficient conditions for the existence and uniqueness of mean-field equilibria are obtained when the horizon lengths are small enough and the Riccati coefficient parameters are positive. The method (see Figure 1) is then extended to the linear–quadratic robust mean-field-type games under disturbance, formulated as a minimax mean-field-type game.

Only a very limited amount of prior work seems to have been done on the MF-LQJD mean-field-type game problems. As indicated in Table 1, the jump term brings a new feature to the existing literature, and to the best of our knowledge, it is the first work that introduces and provides a bargaining solution [34] in mean-field-type games using a direct method.

The last section of this article is devoted to the validation of the novel equations derived in this article using other approaches. We confirm the validity of the optimal feedback strategies. In the Appendix we provide a basic example illustrating the sub-optimality of the mean-field game approach (which consists of freezing the mean-field term) compared with the mean-field-type game approach (in which an individual decision-maker can significantly influence the mean-field term).

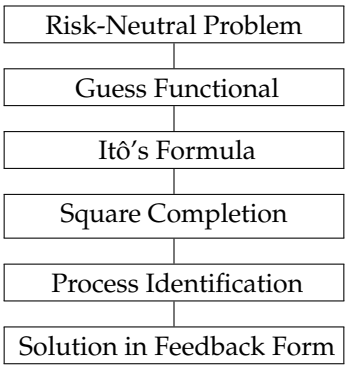

Figure 1. Methods developed in this work. Cooperative vs. noncooperative. Adversarial/robust vs. nonzero-sum mean-field-type games.

1.2. Structure

A brief outline of the article follows. The next section introduces the non-cooperative mean-field-type game problem and provides its solution. Then, the fully-cooperative game and the bargaining problems and their solutions are presented. The last part of the article is devoted to adversarial problems of mean-field type.

Notation and Preliminaries

Let $T > 0$ be a fixed time horizon and $(\Omega, \mathcal{F}, \mathbb{F}^{B,N}, \mathbb{P})$ be a given filtered probability space on which a one-dimensional standard Brownian motion $B = \{B(t)\}_{t\geq 0}$ is given, $\tilde{N}(dt, d\theta) = N(dt, d\theta) - \nu(d\theta)dt$ is a centered jump process with Lévy measure ν defined over Θ. The filtration $\mathbb{F} = \{\mathcal{F}_t^{B,N},\ 0 \leq t \leq T\}$ is the natural filtration generated by the union $\{B, N\}$ augmented by $\mathbb{P}-$null sets of $\mathcal{F}$. The processes B and N are mutually independent. In practice, B is used to capture smaller disturbance and N is used for larger jumps of the system.

We introduce the following notation:

- Let $k \geq 1$. $L^k(0, T; \mathbb{R})$ be the set of functions $f:\ [0, T] \to \mathbb{R}$ such that $\int_0^T |f(t)|^k dt < \infty$.
- $\mathcal{L}^k_{\mathbb{F}}(0, T; \mathbb{R})$ is the set of $\mathbb{F}$-adapted $\mathbb{R}$-valued processes $X(\cdot)$ such that $E\left[\int_0^T |X(t)|^k dt\right] < \infty$.
- $\bar{X}(t) = E[X(t)]$ denotes the expected value of the random variable $X(t)$.

An admissible control strategy u_i of decision-maker i is an $\mathbb{F}$-adapted and square-integrable process with values in a non-empty subset U_i of $\mathbb{R}$. We denote the set of all admissible controls by $\mathcal{U}_i$:

$$\mathcal{U}_i = \{u_i(\cdot) \in \mathcal{L}^2_{\mathbb{F}}(0, T; \mathbb{R});\ u_i(.) \in U_i\ a.e.\, t \in [0, T],\ \mathbb{P} - a.s.\}.$$

2. Non-Cooperative Problem

Consider n risk-neutral decision-makers ($n \geq 2$) and let $L_i(u_1, \ldots, u_n)$ be the objective functional of decision-maker i, given by

$$\begin{aligned} L_i(u_1, \ldots, u_n) &= \tfrac{1}{2} q_i(T) x^2(T) + \tfrac{1}{2} \bar{q}_i(T)[Ex(T)]^2 \\ &+ \tfrac{1}{2} \int_0^T \left\{ q_i(t) x^2(t) + \bar{q}_i(t)(E[x(t)])^2 + r_i(t) u_i^2(t) + \bar{r}_i(t)[Eu_i(t)]^2 \right\} dt. \end{aligned} \tag{1}$$

Then, the best-response of decision-maker i to the process $(u_{-i}, E[x]) := (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n, E[x])$ solves the following risk-neutral linear–quadratic mean-field-type control problem

$$\begin{cases} \inf_{u_i(\cdot)\in\mathcal{U}_i} E\left[L_i(u_1,\ldots,u_n)\right], \\ \text{subject to} \\ dx(t) = \left\{a(t)x(t) + \bar{a}(t)E[x(t)] + \sum_{i=1}^n b_i(t)u_i(t) + \sum_{i=1}^n \bar{b}_i\bar{u}_i(t)\right\} dt \\ \qquad +\sigma(t)dB(t) + \int_\Theta \mu(t,\theta)\tilde{N}(dt,d\theta), \\ x(0) := x_0, \end{cases} \tag{2}$$

where $Ex^2(0) < +\infty$, $q_i(t) \geq 0, q_i(t) + \bar{q}_i(t) \geq 0, r_i(t) > 0, r_i(t) + \bar{r}_i(t) \geq 0$, and $a(t), \bar{a}(t), b_i(t), \sigma(t)$ are real-valued functions, and where $E[x(t)]$ is the expected value of the state created by all decision-makers under the control action profile $(u_1,\ldots,u_n) \in \prod_{j=1}^n \mathcal{U}_j$. The method below can handle time-varying coefficients. For simplicity, we impose an integrability condition on these coefficient functions over $[0,T]$:

$$\int_0^T \left[|a(t)| + |\bar{a}(t)| + |b(t)| + |\bar{b}(t)| + \sigma^2(t) + \int_\Theta (|\mu(t,\theta)| + \mu^2(t,\theta))\nu(d\theta)\right] dt < +\infty. \tag{3}$$

Under condition (3), the state dynamics of (2) has a solution for each $u = (u_1,\ldots,u_n) \in \prod_{j=1}^n \mathcal{U}_j$. Note that we do not impose boundedness or Lipschitz conditions (because quadratic functionals are not necessarily Lipschitz).

Definition 1 (BR$_i$: Best Response of decision-maker i)**.** *Any strategy $u_i^*(\cdot) \in \mathcal{U}_i$ satisfying the infimum in (2) is called a risk-neutral best-response strategy of decision-maker i to the other decision-makers strategy $u_{-i} \in \prod_{j\neq i}\mathcal{U}_j$. The set of best-response strategies of i is denoted by $BR_i : \prod_{j\neq i}\mathcal{U}_j \to 2^{\mathcal{U}_i}$, where $2^{\mathcal{U}_i}$ denotes the set of subsets of $\mathcal{U}_i$.*

Note that if $b_i = 0 = r_i$, there are multiple optimizers of the best-response problem.

Definition 2 (Mean-Field Nash Equilibrium)**.** *Any strategy profile $(u_i^*,\ldots,u_n^*) \in \prod_i \mathcal{U}_i$ such that $u_i^* \in BR_i(u_{-i}^*)$ for every i and $\bar{x}^* = E[x^*]$ is called a Nash equilibrium of the LQ-MFJD game above.*

The risk-neutral mean-field-type Nash equilibrium problem we are concerned with is to find and characterize the processes $(x^*, u^*, E[x^*], E[u^*])$ such that for every decision-maker i, u_i^* is an optimizer of the best response problem (2) and the expected value of the resulting common state $E[x^*]$ created by all the decision-makers coincides with $\bar{x}$. This means that an equilibrium is a fixed-point of the best response correspondence $BR = (BR_1,\ldots,BR_n)$, where $BR_i : \prod_{j\neq i}\mathcal{U}_j \to 2^{\mathcal{U}_i}$ is the best-response correspondence of decision-maker i.

We rewrite the expected objective functional and the state coefficients in terms of $x - \bar{x}$ and $\bar{x}$:

$$\begin{aligned} EL_i &= \tfrac{1}{2}[q_i(T)E(x(T)-\bar{x}(T))^2 + [q_i(T)+\bar{q}_i(T)]\bar{x}^2(T) \\ &\qquad + E\int_0^T q_i(x-\bar{x})^2 + [q_i+\bar{q}_i]\bar{x}^2 + r_i(u_i-\bar{u}_i)^2 + (r_i+\bar{r}_i)[\bar{u}_i]^2 dt], \\ dx &= \left\{a(x-\bar{x}) + (a+\bar{a})\bar{x} + \sum_{i=1}^n b_i(u_i-\bar{u}_i) + \sum_{i=1}^n (b_i+\bar{b}_i)\bar{u}_i\right\} dt \\ &\quad + \sigma dB(t) + \int_\Theta \mu(t,\theta)\tilde{N}(dt,d\theta). \end{aligned} \tag{4}$$

Note that the expected value of the first term in the integral in L_i can be seen as a weighted variance *var* of the state, since $\bar{q}_i(t)E[(x(t)-E[x(t)])^2] = \bar{q}_i(t)var(x(t))$. Taking the expectation of the state dynamics, one arrives at the deterministic linear dynamics

$$\begin{aligned} &\tfrac{d}{dt}Ex = \dot{\bar{x}} = (a+\bar{a})\bar{x} + \sum_{i=1}^n (b_i+\bar{b}_i)\bar{u}_i, \\ &\bar{x}(0) = Ex(0). \end{aligned} \tag{5}$$

The direct method consists of writing a generic structure of the cost functional, with unknown deterministic functions to be identified. Inspired from the structure of the terminal cost function,

we try a generic solution in a quadratic form. Let $f_i(t,x) = \frac{1}{2}\alpha_i(t)(x-\bar{x})^2 + \frac{1}{2}\beta_i(t)\bar{x}^2 + \gamma_i(t)\bar{x} + \delta_i(t)$, where $\alpha, \beta, \gamma, \delta$ are deterministic functions of time, such that

$$f_i(T, x(T)) = \frac{1}{2}\{q_i(T)E(x(T)-\bar{x}(T))^2 + [q_i(T)+\bar{q}_i(T)]\bar{x}^2(T)\}.$$

At the final time T, one can identify $\alpha_i(T) = q_i(T),\ \beta_i(T) = q_i(T) + \bar{q}_i(T),\ \gamma_i(T) = \delta_i(T) = 0$. Recall that Itô's formula for the jump-diffusion process is

$$\begin{aligned}
&f_i(T, x(T)) = f_i(0, x(0)) \\
&+ \int_0^T [f_{i,t} + f_{i,x}D + f_{i,xx}\tfrac{\sigma^2}{2}]dt \\
&+ \int_0^T \sigma f_{i,x} dB \\
&+ \int_0^T \int_\Theta [f_i(t, x+\mu(t,\theta)) - f_i(t,x) - f_{i,x}\mu(t,\theta)]\nu(d\theta)dt \\
&+ \int_0^T \int_\Theta [f_i(t_-, x+\mu(t_-,\theta)) - f_i(t_-,x)]\tilde{N}(dt, d\theta),
\end{aligned} \tag{6}$$

where D is the drift term $D := a(x-\bar{x}) + (a+\bar{a})\bar{x} + \sum_{i=1}^n b_i(u_i - \bar{u}_i) + \sum_{i=1}^n (b_i + \bar{b}_i)\bar{u}_i$. We compute the derivative terms:

$$\begin{cases}
\dot{\bar{x}} = (a+\bar{a})\bar{x} + \sum_{i=1}^n (b_i + \bar{b}_i)\bar{u}_i, \\
f_{i,t} = \frac{1}{2}\dot{\alpha}_i(x-\bar{x})^2 + \frac{1}{2}\dot{\beta}_i\bar{x}^2 + \dot{\gamma}_i\bar{x} + \dot{\delta}_i - \alpha_i(x-\bar{x})\dot{\bar{x}} + \beta_i\bar{x}\dot{\bar{x}} + \gamma_i\dot{\bar{x}}, \\
f_{i,x} = \dot{\alpha}_i(x-\bar{x}), \\
f_{i,xx} = \alpha_i, \\
f_i(t, x+\mu) - f_i(t,x) - f_{i,x}\mu \\
= \frac{1}{2}\alpha_i(x+\mu-\bar{x})^2 - \frac{1}{2}\alpha_i(x-\bar{x})^2 - \alpha_i(x-\bar{x})\mu = \frac{1}{2}\alpha_i\mu^2.
\end{cases} \tag{7}$$

Using (7) in (6) and taking the expectation yields

$$\begin{aligned}
&E[f_i(T, x(T)) - f_i(0, x(0))] \\
&= \tfrac{1}{2}E\int_0^T \dot{\alpha}_i(x-\bar{x})^2 + \dot{\beta}_i\bar{x}^2 dt \\
&+\tfrac{1}{2}E\int_0^T 2\beta_i[(a+\bar{a})\bar{x}^2 + \textstyle\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j\bar{x}]dt \\
&+\tfrac{1}{2}E\int_0^T \left\{2a\alpha_i(x-\bar{x})^2 + 2\alpha_i\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j)(x-\bar{x})\right\}dt \\
&+\tfrac{1}{2}\int_0^T [\sigma^2 + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\alpha_i dt \\
&+E\int_0^T \dot{\gamma}_i\bar{x} + \gamma_i[(a+\bar{a})\bar{x} + \textstyle\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j]dt + \int_0^T \dot{\delta}_i dt,
\end{aligned} \tag{8}$$

where we have used the following equalities:

$$\begin{cases}
E[\alpha_i(x-\bar{x})\dot{\bar{x}}] = 0, \\
E\int_0^T \sigma f_{i,x} dB = 0, \\
E\int_0^T \int_\Theta [f_i(t, x+\mu(t,\theta)) - f_i(t,x)]\tilde{N}(dt, d\theta) = 0.
\end{cases} \tag{9}$$

We compute the gap between $E[L_i]$ and $E[f_i(0, x(0))]$ as

$$\begin{aligned}
&E[L_i - f_i(0, x(0))] \\
&= \tfrac{1}{2}(q_i(T) - \alpha_i(T))E(x(T)-\bar{x}(T))^2 \\
&+\tfrac{1}{2}[q_i(T) + \bar{q}_i(T) - \beta_i(T)]\bar{x}^2(T) \\
&+\tfrac{1}{2}E\int_0^T q_i(x-\bar{x})^2 + [q_i + \bar{q}_i]\bar{x}^2(t)dt \\
&+\tfrac{1}{2}E\int_0^T r_i(u_i - \bar{u}_i(t))^2 + (r_i + \bar{r}_i)[\bar{u}_i]^2 dt] \\
&+\tfrac{1}{2}E\int_0^T \dot{\alpha}_i(x-\bar{x})^2 + \dot{\beta}_i\bar{x}^2 + 2\beta_i(a+\bar{a})\bar{x}^2 dt \\
&+\tfrac{1}{2}E\int_0^T 2\beta_i(b_i+\bar{b}_i)\bar{u}_i\bar{x}dt \\
&+\tfrac{1}{2}E\int_0^T 2\beta_i\textstyle\sum_{j\neq i}^n (b_j+\bar{b}_j)\bar{u}_j\bar{x}dt \\
&+\tfrac{1}{2}E\int_0^T \left\{2a\alpha_i(x-\bar{x})^2 + 2\alpha_i b_i(u_i-\bar{u}_i)(x-\bar{x}) + 2\alpha_i\textstyle\sum_{j\neq i} b_j(u_j-\bar{u}_j)(x-\bar{x})\right\}dt \\
&+\tfrac{1}{2}\int_0^T [\sigma^2 + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\alpha_i dt \\
&+\tfrac{1}{2}E\int_0^T 2\dot{\gamma}_i\bar{x} + 2\gamma_i[(a+\bar{a})\bar{x} + \textstyle\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j]dt \\
&+\tfrac{1}{2}\int_0^T 2\dot{\delta}_i dt.
\end{aligned} \tag{10}$$

2.1. Best Response to Open-Loop Strategies

In this subsection, we compute the best-response of decision-maker i to open-loop strategies $(u_j)_{j\neq i}$. The information structure for the others players is limited to time and initial point; i.e., the mappings $(u_j)_{j\neq i}$ are measurable functions of time (and do not depend on x) and initial point x_0.

$$\begin{aligned}
&E[L_i - f_i(0, x(0))] \\
&= \tfrac{1}{2}[(q_i(T) - \alpha_i(T))E(x(T) - \bar{x}(T))^2 \\
&+\tfrac{1}{2}[q_i(T) + \bar{q}_i(T) - \beta_i(T)]\bar{x}^2(T) \\
&+\tfrac{1}{2}E\int_0^T \{\dot{\alpha}_i + 2a\alpha_i - \tfrac{b_i^2}{r_i}\alpha_i^2 + q_i\}(x(t) - \bar{x}(t))^2 \\
&+\tfrac{1}{2}E\int_0^T [\dot{\beta}_i + 2\beta_i(a + \bar{a}) - \tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\beta_i^2 + q_i + \bar{q}_i]\bar{x}^2 dt \\
&+\tfrac{1}{2}E\int_0^T r_i[u_i - \bar{u}_i + \tfrac{b_i}{r_i}\alpha_i(x - \bar{x})]^2 dt \\
&+\tfrac{1}{2}E\int_0^T (r_i + \bar{r}_i)[\bar{u}_i + \tfrac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}(\beta_i\bar{x} + \gamma_i)]^2 dt \\
&+\tfrac{1}{2}E\int_0^T \left\{2\dot{\gamma}_i + 2\gamma_i(a + \bar{a}) - \tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}2\beta_i\gamma_i + 2\beta_i\textstyle\sum_{j\neq i}(b_j + \bar{b}_j)\bar{u}_j\right\}\bar{x}dt \\
&+\tfrac{1}{2}E\int_0^T -\tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\gamma_i^2 + 2\gamma_i\textstyle\sum_{j\neq i}(b_j + \bar{b}_j)\bar{u}_j dt \\
&+\tfrac{1}{2}\int_0^T [\sigma^2 + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\alpha_i dt \\
&+\tfrac{1}{2}\int_0^T 2\mathring{\delta}_i\, dt,
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
&(r_i + \bar{r}_i)[\bar{u}_i]^2 + 2\beta_i\textstyle\sum_{j=1}^n (b_j + \bar{b}_j)\bar{u}_j\bar{x} + 2\gamma_i\sum_{j=1}^n (b_j + \bar{b}_j)\bar{u}_j \\
&= (r_i + \bar{r}_i)[\bar{u}_i]^2 + 2\bar{u}_i(b_i + \bar{b}_i)(\beta_i x + \bar{\gamma}_i) + 2\beta_i\textstyle\sum_{j\neq i}(b_j + \bar{b}_j)\bar{u}_j\bar{x} + 2\gamma_i\sum_{j\neq i}(b_j + \bar{b}_j)\bar{u}_j \\
&= (r_i + \bar{r}_i)[\bar{u}_i + \tfrac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}(\beta_i\bar{x} + \gamma_i)]^2 - \tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\{\beta_i^2\bar{x}^2 + 2\beta_i\gamma_i\bar{x} + \bar{\gamma}_i^2\} \\
&+2\beta_i\textstyle\sum_{j\neq i}(b_j + \bar{b}_j)\bar{u}_j\bar{x} + 2\gamma_i\sum_{j\neq i}(b_j + \bar{b}_j)\bar{u}_j.
\end{aligned} \tag{12}$$

The best response of decision-maker i to the open-loop strategies $(u_j)_{j\neq i}$ is $u_i = \bar{u}_i - \frac{b_i}{r_i}\alpha_i(x - \bar{x})$, and its expected value is $\bar{u}_i = -\frac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}(\beta_i\bar{x} + \gamma_i)$, where $\alpha_i, \beta_i, \gamma_i$ are deterministic functions of time t. Clearly, the best response to open-loop strategies is in state-and-mean-field feedback form. Here the mean-field feedback terms are the expected value of the state $E[x(t)]$ and the expected value of the control action $E[u_i(t)]$.

Therefore, we examine optimal strategies in state-and-mean-field feedback form in the next section.

2.2. Feedback Strategies

The information structure for feedback solution is as follows. The model and the objective functions are assumed to be common knowledge. We assume that the state is of perfect observation. We will show below that the mean-field term is computable (via the initial mean state and the model). If the other decision-makers play their optimal state-and-mean-field feedback strategies, then the functions $\gamma_1, \ldots, \gamma_n$ are identically zero at any given time. We compute again $E[L_i - f_i(0, x(0))]$ and complete the squares using the elements of $\{x - \bar{x},\ \bar{x}\}$.

$$\begin{aligned}
&E[L_i - f_i(0, x(0))] \\
&= \tfrac{1}{2}(q_i(T) - \alpha_i(T))E(x(T) - \bar{x}(T))^2 \\
&\quad +\tfrac{1}{2}[q_i(T) + \bar{q}_i(T) - \beta_i(T)]\bar{x}^2(T) \\
&+\tfrac{1}{2}E\int_0^T [\dot{\alpha}_i + 2a\alpha_i - \tfrac{b_i^2}{r_i}\alpha_i^2 - 2\alpha_i\textstyle\sum_{j\neq i}^n \tfrac{b_j^2}{r_j}\alpha_j + q_i](x - \bar{x})^2 \\
&+\tfrac{1}{2}E\int_0^T \left\{\dot{\beta}_i + 2\beta_i(a + \bar{a}) - \beta_i^2\tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i} - 2\beta_i\textstyle\sum_{j\neq i}^n \tfrac{(b_j+\bar{b}_j)^2}{r_j+\bar{r}_j}\beta_j + q_i + \bar{q}_i\right\}\bar{x}^2 dt \\
&+\tfrac{1}{2}E\int_0^T r_i[u_i - \bar{u}_i + \tfrac{b_i}{r_i}\alpha_i(x - \bar{x})]^2 dt \\
&+\tfrac{1}{2}E\int_0^T (r_i + \bar{r}_i)[\bar{u}_i + \beta_i\tfrac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}\bar{x}]^2 dt \\
&+\tfrac{1}{2}E\int_0^T [\sigma^2 + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\ \alpha_i dt,
\end{aligned} \tag{13}$$

where we have used the following square completions:

$$\begin{aligned}
&r_i(u_i-\bar{u}_i)^2+2\alpha_i\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j)(x-\bar{x})\\
&=r_i[u_i-\bar{u}_i+\tfrac{b_i}{r_i}\alpha_i(x-\bar{x})]^2-\tfrac{b_i^2}{r_i}\alpha_i^2(x-\bar{x})^2+2\alpha_i\textstyle\sum_{j\neq i}b_j(u_j-\bar{u}_j)(x-\bar{x}),\ \text{and}\\
&(r_i+\bar{r}_i)[\bar{u}_i]^2+2\beta_i\textstyle\sum_{j=1}^n(b_j+\bar{b}_j)\bar{u}_j\bar{x}\\
&=(r_i+\bar{r}_i)[\bar{u}_i+\beta_i\tfrac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}\bar{x}]^2-\beta_i^2\tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\bar{x}^2+2\beta_i\textstyle\sum_{j\neq i}(b_j+\bar{b}_j)\bar{u}_j\bar{x}.
\end{aligned}\tag{14}$$

It follows that

$$\begin{aligned}
&\inf_{u_i\in\mathcal{U}_i}E[L_i]=\tfrac{1}{2}\alpha_i(0)var(x(0))+\tfrac{1}{2}\beta_i(0)[Ex(0)]^2+\tfrac{1}{2}\textstyle\int_0^T[\sigma^2+\int_\Theta\mu^2(t,\theta)\nu(d\theta)]\,\alpha_i dt,\\
&u_i^*=-\tfrac{b_i}{r_i}\alpha_i(x-\bar{x})-\beta_i\tfrac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}\bar{x},\\
&\dot{\alpha}_i+2a\alpha_i-\tfrac{b_i^2}{r_i}\alpha_i^2-2\alpha_i\textstyle\sum_{j\neq i}\tfrac{b_j^2}{r_j}\alpha_j+q_i=0,\\
&\alpha_i(T)=q_i(T),\\
&\dot{\beta}_i+2\beta_i(a+\bar{a})-\beta_i^2\tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}-2\beta_i\textstyle\sum_{j\neq i}\tfrac{(b_j+\bar{b}_j)^2}{r_j+\bar{r}_j}\beta_j+q_i+\bar{q}_i=0,\\
&\beta_i(T)=q_i(T)+\bar{q}_i(T).\\
&\bar{x}(t)=\bar{x}(0)e^{\int_0^t\{(a+\bar{a})-\sum_{i=1}^n\beta_i\frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\}dt'}
\end{aligned}\tag{15}$$

provides a mean-field Nash equilibrium in feedback strategies.

These Riccati equations are different from those of open-loop control strategies. The coefficient of the coupling terms $\beta_i\beta_j$, $\alpha_i\alpha_j$ are different, reflecting the coupling through the state and the mean state. Notice that the optimal strategy is in state-and-mean-field feedback form, which is different from the standard LQG game solution. As $\bar{a},\bar{b},\bar{r},\bar{q}$ vanish in (15), one gets the Nash equilibrium of the corresponding stochastic differential game in closed-loop strategies with $\alpha_i=\beta_i$, and u_i becomes mean-field-free. When the diffusion coefficient σ and the jump rate μ vanish, one obtains the noiseless deterministic game problem, and the optimal strategy solution will be given by the equation in β_i because $x-\bar{x}=0$ in the deterministic case.

How to feedback the mean-field term $E[x(t)]$? Here the mean-field term can be explicitly computed if the initial mean state $\bar{x}(0)$ is given and the model known:

$$E[x(t)]=\bar{x}(t)=\bar{x}(0)e^{\int_0^t\{(a+\bar{a})-\sum_{i=1}^n\beta_i\frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\}dt'}.$$

3. Fully-Cooperative Solutions

In this section, we examine the global optimum and Nash bargaining solution [34] of the game.

3.1. Global Optimum

We now consider the fully cooperative scenario where all the decision-makers decide jointly to optimize a single global objective $L_0:=\sum_i L_i$ given by

$$\begin{aligned}
&\inf_{(u_1,\ldots,u_n)}E\textstyle\sum_i L_i,\\
&dx=\left\{a(x-\bar{x})+(a+\bar{a})\bar{x}+\sum_{i=1}^n b_i(u_i-\bar{u}_i)+\sum_{i=1}^n(b_i+\bar{b}_i)\bar{u}_i\right\}dt\\
&+\sigma dB(t)+\textstyle\int_\Theta\mu(t,\theta)\tilde{N}(dt,d\theta),\\
&x(0)=x_0.
\end{aligned}\tag{16}$$

Following the same methodology as above with $q_0=\sum_i q_i$, $\bar{q}_0=\sum_i\bar{q}_i$ and $f_0(t,x)=\frac{1}{2}\alpha_0(t)(x-\bar{x})^2+\frac{1}{2}\beta_0(t)\bar{x}^2$, we obtain:

$$
\begin{aligned}
&\inf_{(u_1,\ldots,u_n)} E\sum_i [L_i] \\
&= \tfrac{1}{2}\alpha_0(0)var(x(0)) + \tfrac{1}{2}\beta_0(0)[Ex(0)]^2 + \tfrac{1}{2}\int_0^T [\sigma^2(t) + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\ \alpha_0(t)dt, \\
&u_i^* = -\tfrac{b_i}{r_i}\alpha_0(x-\bar{x}) - \beta_0\tfrac{(b_i+\bar{b}_i)}{r_i+\bar{r}_i}\bar{x}, \\
&\dot{\alpha}_0 + 2a\alpha_0 - \alpha_0^2\textstyle\sum_i \frac{b_i^2}{r_i} + q_0 = 0, \\
&\alpha_0(T) = q_0(T) = \textstyle\sum_{i=1}^n q_i(T), \\
&\dot{\beta}_0 + 2\beta_0(a+\bar{a}) - \beta_0^2\textstyle\sum_{i=1}^n \frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i} + q_0 + \bar{q}_0 = 0, \\
&\beta_0(T) = \textstyle\sum_{i=1}^n q_i(T) + \bar{q}_i(T). \\
&\bar{x}(t) = \bar{x}(0)e^{\int_0^t \{(a+\bar{a}) - \beta_0 \sum_{i=1}^n \frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\}dt'}.
\end{aligned}
\tag{17}
$$

When the coefficients are constant (in time), α_0, β_0 are explicitly given by

$$
\begin{aligned}
&S := \textstyle\sum_i \frac{b_i^2}{r_i}, \\
&\alpha_0(t) = \tfrac{a}{S} + \sqrt{\tfrac{q_0}{S} + \tfrac{a^2}{S^2}}\left(-1 + \frac{q_0(T) - \frac{a}{S} + \sqrt{\frac{q_0}{S} + \frac{a^2}{S^2}}}{\Gamma}\right), \\
&\Gamma := \tfrac{1}{2}(q_0(T) - \tfrac{a}{S} + \sqrt{\tfrac{q_0}{S} + \tfrac{a^2}{S^2}}) - \tfrac{1}{2}(q_0(T) - \tfrac{a}{S} - \sqrt{\tfrac{q_0}{S} + \tfrac{a^2}{S^2}})e^{-2(T-t)\sqrt{\frac{q_0}{S} + \frac{a^2}{S^2}}}, \\
&\tilde{S} := \textstyle\sum_{i=1}^n \frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}, \\
&\beta_0(t) = \tfrac{a+\bar{a}}{\tilde{S}} + \sqrt{\tfrac{q_0+\bar{q}_0}{\tilde{S}} + \tfrac{a^2}{\tilde{S}^2}}\left(-1 + \frac{q_0(T) + \bar{q}_0(T) - \frac{a+\bar{a}}{\tilde{S}} + \sqrt{\frac{q_0+\bar{q}_0}{\tilde{S}} + \frac{(a+\bar{a})^2}{\tilde{S}^2}}}{\tilde{\Gamma}}\right), \\
&\tilde{\Gamma} := \tfrac{1}{2}(q_0(T) + \bar{q}_0(T) - \tfrac{a+\bar{a}}{\tilde{S}} + \sqrt{\tfrac{q_0+\bar{q}_0}{\tilde{S}} + \tfrac{(a+\bar{a})^2}{\tilde{S}^2}}) \\
&-\tfrac{1}{2}(q_0(T) + \bar{q}_0(T) - \tfrac{a+\bar{a}}{\tilde{S}} - \sqrt{\tfrac{q_0+\bar{q}_0}{\tilde{S}} + \tfrac{(a+\bar{a})^2}{\tilde{S}^2}}) \times e^{-2(T-t)\sqrt{\frac{q_0+\bar{q}_0}{\tilde{S}} + \frac{(a+\bar{a})^2}{\tilde{S}^2}}}.
\end{aligned}
\tag{18}
$$

The global optimum cost in the fully-cooperative case is

$$
L_0 = \tfrac{1}{2}\alpha_0(0)var(x(0)) + \tfrac{1}{2}\beta_0(0)[Ex(0)]^2 + \tfrac{1}{2}\int_0^T [\sigma^2(t) + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\ \alpha_0(t)dt, \tag{19}
$$

and is less than the total cost at the Nash equilibrium, which is

$$
\tfrac{1}{2}\left(\textstyle\sum_{i=1}^n \alpha_i(0)\right) var(x(0)) + \tfrac{1}{2}\left(\textstyle\sum_{i=1}^n \beta_i(0)\right)(\bar{x}(0))^2 + \tfrac{1}{2}\int_0^T [\sigma^2(t) + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\ \textstyle\sum_{i=1}^n \alpha_i(t)dt. \tag{20}
$$

This loss of efficiency of Nash equilibria was analyzed in [35], and is often termed as the price of anarchy [36,37].

3.2. Nash Bargaining Solution

Mean-field-type bargaining theory deals with the situation in which decision-makers can realize—through cooperation—other better outcomes than the one which becomes effective when they do not cooperate. This non-cooperative outcome is called the threatpoint $L^{NE} = (L_1^{NE}, \ldots, L_n^{NE})$. The question is which outcome might the decision-makers possibly agree to. Let $\mathcal{V}$ be the set of feasible outcomes of the benefit of bargaining [34]. We assume that if the agents unanimously agree on a point $v = (v_1, \ldots, v_n) \in \mathcal{V}$, they obtain v. Otherwise, they obtain $L^{NE} = (L_1^{NE}, \ldots, L_n^{NE})$. This presupposes that each decision-maker can enforce the threatpoint when he does not agree with a proposal. Generically, what decision-maker i can guarantee is $\inf_{u_i} \sup_{u_{-i}} L_i$, which is non-admissible in the quadratic setting. The outcome v the decision-makers will finally agree on is called the solution of the bargaining problem. Therefore, we have chosen the non-cooperation solution when there is a disagreement.

The Nash bargaining solution selects for a given set $\mathcal{V}$ the point at which the product of gains from L^{NE} is maximal.

$$NBS(\mathcal{V}, L^{NE}) = \arg\max_{v \in \mathcal{V}} \{ \prod_{i \in \mathcal{N}} [L_i^{NE} - v_i] \} \tag{21}$$

Since the function $v \mapsto \prod_{k \in \mathcal{N}} v_k$ is non-convex, Problem (21) is non-convex. Here we exploit the convexity of the functional $u \mapsto \langle w, L(u) \rangle$ for any given $w = (w_1, \ldots, w_n) \in \mathbb{R}^n_{++}$ such that $\sum_{i=1}^n w_i = 1$ to reach any point in the Pareto frontier of the game. The maximization (in w) of the product $P := \prod_{i=1}^n [L_i^{NE} - L_i(\hat{u}(w))]$ yields

$$\sum_{i=1}^n \frac{\partial}{\partial w_j} L_i(\hat{u}(w)).\{\prod_{j \neq i} [L_j^{NE} - L_j(\hat{u}(w))]\} = c.$$

This is equivalent to

$$\sum_{i=1}^n \frac{P}{L_i^{NE} - L_i(\hat{u}(w))} \frac{\partial}{\partial w_j} L_i(\hat{u}(w)) = c.$$

We set $y_i := \frac{\frac{P}{L_i^{NE} - L_i(\hat{u}(w))}}{\sum_{k=1}^n \left(\frac{P}{L_k^{NE} - L_k(\hat{u}(w))} \right)}$. Then, it follows that

$$\sum_{i=1}^n y_i \frac{\partial}{\partial w_j} L_i(\hat{u}(w)) = \frac{c}{\sum_{i=1}^n \left(\frac{P}{L_i^{NE} - L_i(\hat{u}(w))} \right)}.$$

Moreover, $y_i \geq 0$, $\sum_{i=1}^n y_i = 1$.

Assume the matrix $(\frac{\partial}{\partial w_j} L_i(\hat{u}(w)))_{(i,j) \in \mathcal{N}^2}$ has at least rank $n-1$. Then, the Nash bargaining solution is explicitly given by $v = (L_1(\hat{u}(w^*)), \ldots, L_n(\hat{u}(w^*)))$, with the weight

$$w_i^* = \frac{\prod_{j \neq i} [L_j^{NE} - L_j(\hat{u}(w^*))]}{\sum_{k=1}^n \prod_{j \neq k} [L_j^{NE} - L_j(\hat{u}(w^*))]} = y_i(w^*),$$

where the optimal bargaining strategy profile is $\hat{u}(w) \in \arg\min_u \{\langle w, L(u) \rangle\}$. It remains to compute the functional $\hat{u}(w) = (\hat{u}_1(w), \ldots, \hat{u}_n(w))$.

$$\begin{aligned}
&\inf_{(u_1, \ldots, u_n)} E \textstyle\sum_i w_i L_i, \\
&dx = \left\{ a(x - \bar{x}) + (a + \bar{a})\bar{x} + \sum_{i=1}^n b_i(u_i - \bar{u}_i) + \sum_{i=1}^n (b_i + \bar{b}_i)\bar{u}_i \right\} dt \\
&+ \sigma dB + \textstyle\int_\Theta \mu(t, \theta) \tilde{N}(dt, d\theta). \\
&x(0) = x_0,
\end{aligned} \tag{22}$$

Following the same methodology as above, we obtain:

$$\begin{aligned}
&\inf_{(u_1, \ldots, u_n)} E \textstyle\sum_i w_i L_i = \frac{1}{2}\alpha_0(0) var(x(0)) + \frac{1}{2}\beta_0(0)[\bar{x}(0)]^2 + \int_0^T [\sigma^2 + \int_\Theta \mu^2(t, \theta)\nu(d\theta)] \frac{\alpha_0}{2} dt, \\
&\hat{u}_i(w) = -\frac{b_i}{w_i r_i} \alpha_0 (x - \bar{x}) - \frac{(b_i + \bar{b}_i)}{w_i(r_i + \bar{r}_i)} (\beta_0 \bar{x}), \\
&\dot{\alpha}_0 + 2a\alpha_0 - \alpha_0^2 \textstyle\sum_i \frac{b_i^2}{w_i r_i} + \sum_i w_i q_i = 0, \\
&\alpha_0(T) = \textstyle\sum_i w_i q_i(T), \\
&\dot{\beta}_0 + 2\beta_0(a + \bar{a}) - \beta_0^2 \textstyle\sum_i \frac{(b_i + \bar{b}_i)^2}{w_i(r_i + \bar{r}_i)} + \sum_i w_i(q_i + \bar{q}_i) = 0, \\
&\beta_0(T) = \textstyle\sum_i w_i(q_i + \bar{q}_i(T)), \\
&\bar{x}(t) = \bar{x}(0) e^{\int_0^t [a + \bar{a} - \beta_0 \sum_{i=1}^n \frac{(b_i + \bar{b}_i)^2}{w_i(r_i + \bar{r}_i)}] dt'}, \\
&\bar{x}(0) \in \mathbb{R}.
\end{aligned} \tag{23}$$

4. LQ Robust Mean-Field-Type Games

We now consider a robust mean-field-type game with two decision-makers. Decision-maker 1 minimizes with respect to u_1 and Decision-maker 2 maximizes with respect to u_2. The minimax problem of mean-field type is given by

$$\begin{aligned}
&\inf_{u_1} \sup_{u_2} E[L(u_1, u_2)], \\
&dx = \{a(t)(x(t) - \bar{x}(t)) + (a(t) + \bar{a}(t))\bar{x}(t) \\
&+ \textstyle\sum_{i=1}^{2} b_i(t)(u_i(t) - \bar{u}_i(t)) + \sum_{i=1}^{2}(b_i(t) + \bar{b}_i(t))\bar{u}_i(t)\Big\} dt \\
&+\sigma(t)dB(t) + \textstyle\int_\Theta \mu(t,\theta)\tilde{N}(t, d\theta), \\
&x(0) = x_0 \in L^2(\mathbb{R}),
\end{aligned} \tag{24}$$

where the objective functional is

$$\begin{aligned}
L(u_1, u_2) = &\tfrac{1}{2}[q(T)(x(T) - \bar{x}(T))^2 + [q(T) + \bar{q}(T)]\bar{x}^2(T) \\
&+ \textstyle\int_0^T q(t)(x(t) - \bar{x}(t))^2 + [q(t) + \bar{q}(t)]\bar{x}^2(t)dt \\
&+ \textstyle\int_0^T r_1(t)(u_1(t) - \bar{u}_1(t))^2 + (r_1(t) + \bar{r}_1(t))[\bar{u}_1(t)]^2 dt \\
&+ \textstyle\int_0^T r_2(t)(u_2(t) - \bar{u}_2(t))^2 + (r_2(t) + \bar{r}_2(t))[\bar{u}_2(t)]^2 dt].
\end{aligned} \tag{25}$$

The risk-neutral robust mean-field-type equilibrium problem we are concerned with is to characterize the processes $(x^*, u_1^*, u_2^*, E[x^*])$ such that for every decision-maker, $\bar{u}_1^*$ is the minimizer and u_2^* is the maximum of the best response problem (24), and the expected value of the resulting common state created by all the decision-makers is $\bar{x}^*(t) = E[x^*(t)]$.

Below, we solve Problem (24) for $r_1(t) > 0, \bar{r}_1(t) > 0, r_2(t) < 0,\ \bar{r}_2(t) < 0$.

$$\begin{aligned}
&E[f(T, x(T)) - f(0, x(0))] \\
&= \tfrac{1}{2}E\textstyle\int_0^T \{\dot{\alpha} + 2a\alpha\}(x - \bar{x})^2 dt + \{\dot{\beta} + 2(a + \bar{a})\beta\}\bar{x}^2 dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T 2\alpha \sum_{i=1}^{2} b_i(u_i - \bar{u}_i)(x - \bar{x})dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T 2\beta \sum_{i=1}^{2}(b_i + \bar{b}_i)\bar{u}_i\bar{x}dt \\
&+\tfrac{1}{2}\textstyle\int_0^T [\sigma^2 + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\ \alpha dt
\end{aligned} \tag{26}$$

$$\begin{aligned}
&E[L - f(0, x(0))] \\
&= \tfrac{1}{2}[(q(T) - \alpha(T))]E(x(T) - \bar{x}(T))^2 \\
&+\tfrac{1}{2}[q(T) + \bar{q}(T) - \beta(T)]\bar{x}^2(T) \\
&+\tfrac{1}{2}E\textstyle\int_0^T \{\dot{\alpha} + 2a\alpha - (\frac{b_1^2}{r_1} + \frac{b_2^2}{r_2})\alpha^2 + q\}(x - \bar{x})^2 dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T \Big\{\dot{\beta} + 2(a + \bar{a})\beta \\
&\quad -(\tfrac{(b_1+\bar{b}_1)^2}{r_1+\bar{r}_1} + \tfrac{(b_2+\bar{b}_2)^2}{r_2+\bar{r}_2})\beta^2 + q + \bar{q}\Big\}\ \bar{x}^2 dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T r_1[u_1 - \bar{u}_1 + \frac{b_1}{r_1}\alpha(x - \bar{x})]^2\ dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T (r_1 + \bar{r}_1)[\bar{u}_1 + \frac{(b_1+\bar{b}_1)}{r_1+\bar{r}_1}\beta\bar{x}]^2\ dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T r_2[u_2 - \bar{u}_2 + \frac{b_2}{r_2}\alpha(x - \bar{x})]^2 dt \\
&+\tfrac{1}{2}E\textstyle\int_0^T (r_2 + \bar{r}_2)[\bar{u}_2 + \frac{(b_2+\bar{b}_2)}{r_2+\bar{r}_2}\beta\bar{x}]^2 dt \\
&+\tfrac{1}{2}\textstyle\int_0^T [\sigma^2 + \int_\Theta \mu^2(t,\theta)\nu(d\theta)]\ \alpha dt,
\end{aligned} \tag{27}$$

where we have used the following square completions:

$$\begin{aligned}
&r_1(u_1-\bar{u}_1)^2+(r_1+\bar{r}_1)[\bar{u}_1]^2\\
&+2\alpha b_1(u_1-\bar{u}_1)(x-\bar{x})+2\beta(b_1+\bar{b}_1)\bar{u}_1\bar{x}\\
&=r_1[u_1-\bar{u}_1+\tfrac{b_1}{r_1}\alpha(x-\bar{x})]^2-\tfrac{b_1^2}{r_1}\alpha^2(x-\bar{x})^2\\
&+(r_1+\bar{r}_1)[\bar{u}_1+\tfrac{(b_1+\bar{b}_1)}{r_1+\bar{r}_1}\beta\bar{x}]^2-\tfrac{(b_1+\bar{b}_1)^2}{r_1+\bar{r}_1}\beta^2\bar{x}^2,\\
&r_2(u_2-\bar{u}_2)^2+(r_2+\bar{r}_2)[\bar{u}_2]^2\\
&+2\alpha b_2(u_2-\bar{u}_2)(x-\bar{x})+2\beta(b_2+\bar{b}_2)\bar{u}_2\bar{x}\\
&=r_2[u_2-\bar{u}_2+\tfrac{b_2}{r_2}\alpha(x-\bar{x})]^2-\tfrac{b_2^2}{r_2}\alpha^2(x-\bar{x})^2\\
&+(r_2+\bar{r}_2)[\bar{u}_2+\tfrac{(b_2+\bar{b}_2)}{r_2+\bar{r}_2}\beta\bar{x}]^2-\tfrac{(b_2+\bar{b}_2)^2}{r_2+\bar{r}_2}\beta^2\bar{x}^2.
\end{aligned} \tag{28}$$

It follows that the equilibrium solution is

$$\begin{aligned}
&\inf_{u_1}\sup_{u_2} E[L(u_1,u_2)]=\tfrac{1}{2}\alpha(0)var(x(0))+\tfrac{1}{2}\beta(0)[E(x(0))]^2+\tfrac{1}{2}\int_0^T[\sigma^2(t)+\int_\Theta\mu^2(t,\theta)\nu(d\theta)]\,\alpha(t)dt,\\
&u_1^*=-\tfrac{b_1}{r_1}\alpha(x-\bar{x})-\tfrac{(b_1+\bar{b}_1)}{r_1+\bar{r}_1}\beta\bar{x},\\
&u_2^*=-\tfrac{b_2}{r_2}\alpha(x-\bar{x})-\tfrac{(b_2+\bar{b}_2)}{r_2+\bar{r}_2}\beta\bar{x},\\
&\dot{\alpha}+2a\alpha-(\tfrac{b_1^2}{r_1}+\tfrac{b_2^2}{r_2})\alpha^2+q=0,\\
&\alpha(T)=q(T),\\
&\dot{\beta}+2(a+\bar{a})\beta-(\tfrac{(b_1+\bar{b}_1)^2}{r_1+\bar{r}_1}+\tfrac{(b_2+\bar{b}_2)^2}{r_2+\bar{r}_2})\beta^2+q+\bar{q}=0,\\
&\beta(T)=q(T)+\bar{q}(T),\\
&\bar{x}(t)=\bar{x}(0)e^{\int_0^t\{(a+\bar{a})-\beta\sum_{i=1}^2\frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\}dt'}.
\end{aligned} \tag{29}$$

When $r_1>0,\ \bar{r}_1\geq 0,\ r_2<0,\bar{r}_2\geq 0,\ S_2=\sum_{i=1}^2\frac{b_i^2}{r_i}>0$ and $\tilde{S}_2=\sum_{i=1}^2\frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}>0$, the functions α,β are explicitly given by

$$\begin{aligned}
&\alpha(t)=\tfrac{a}{S_2}+\sqrt{\tfrac{q_0}{S_2}+\tfrac{a^2}{S_2^2}}\left(-1+\frac{q_0(T)-\frac{a}{S_2}+\sqrt{\frac{q_0}{S_2}+\frac{a^2}{S_2^2}}}{\Gamma}\right),\\
&\Gamma:=\tfrac{1}{2}(q_0(T)-\tfrac{a}{S_2}+\sqrt{\tfrac{q_0}{S_2}+\tfrac{a^2}{S_2^2}})-\tfrac{1}{2}(q_0(T)-\tfrac{a}{S_2}-\sqrt{\tfrac{q_0}{S_2}+\tfrac{a^2}{S_2^2}})e^{-2(T-t)\sqrt{\frac{q_0}{S_2}+\frac{a^2}{S_2^2}}},\\
&\beta(t)=\tfrac{a+\bar{a}}{\tilde{S}_2}+\sqrt{\tfrac{q_0+\bar{q}_0}{\tilde{S}_2}+\tfrac{a^2}{\tilde{S}_2^2}}\left(-1+\frac{q_0(T)+\bar{q}_0(T)-\frac{a+\bar{a}}{\tilde{S}_2}+\sqrt{\frac{q_0+\bar{q}_0}{\tilde{S}_2}+\frac{(a+\bar{a})^2}{\tilde{S}_2^2}}}{\tilde{\Gamma}}\right),\\
&\tilde{\Gamma}:=\tfrac{1}{2}(q_0(T)+\bar{q}_0(T)-\tfrac{a+\bar{a}}{\tilde{S}_2}+\sqrt{\tfrac{q_0+\bar{q}_0}{\tilde{S}_2}+\tfrac{(a+\bar{a})^2}{\tilde{S}_2^2}})\\
&-\tfrac{1}{2}(q_0(T)+\bar{q}_0(T)-\tfrac{a+\bar{a}}{\tilde{S}_2}-\sqrt{\tfrac{q_0+\bar{q}_0}{\tilde{S}_2}+\tfrac{(a+\bar{a})^2}{\tilde{S}_2^2}})\times e^{-2(T-t)\sqrt{\frac{q_0+\bar{q}_0}{\tilde{S}_2}+\frac{(a+\bar{a})^2}{\tilde{S}_2^2}}}.
\end{aligned} \tag{30}$$

Notice that under the conditions $r_1(t)>0,\ \bar{r}_1(t)>0,\ r_2(t)<0,\bar{r}_2(t)<0,\ S_2(t)>0,\tilde{S}_2(t)>0$, the minimax solution is also a maximin solution: there is a saddle point, and the saddle point is (u_1^*,u_2^*). It solves

$$\begin{aligned}
&E[L(u_1^*,u_2)]\leq E[L(u_1^*,u_2^*)]\leq E[L(u_1,u_2^*)],\\
&\forall(u_1,u_2)\in\mathcal{U}_1\times\mathcal{U}_2.
\end{aligned} \tag{31}$$

The value of the game is

$$E[L(u_1^*,u_2^*)]=\tfrac{1}{2}\alpha(0)var(x(0))+\tfrac{1}{2}\beta(0)[Ex(0)]^2+\tfrac{1}{2}\int_0^T[\sigma^2+\int_\Theta\mu^2(t,\theta)\nu(d\theta)]\,\alpha(t)dt. \tag{32}$$

5. Checking Our Results

In this section, we verify the validity of our results above using a Bellman system. Due to the non-Markovian nature of x, one needs to build an augmented state. A candidate augmented state is the measure m, since one can write the objective functionals in terms of the measure $m(dx)$. This leads to a dynamic programming principle in infinite dimensions. Below, we use functional derivatives with respect to $m(dx)$. The Bellman equilibrium system (in infinite dimension) is

$$\hat{V}_{i,t}(t,m) + \int_x H_i(x,m,\hat{V}_{i,m},\hat{V}_{i,xm},\hat{V}_{i,xxm})m(dx) = 0, \tag{33}$$

where the terminal equilibrium payoff functional at time T is

$$\hat{V}_i(T,m) = \tfrac{1}{2}q(T)\int_y (y-\bar{m})^2\, m(dy) + \tfrac{1}{2}[q(T)+\bar{q}(T)]\left[\int_y y\, m(dy)\right]^2, \tag{34}$$

and the integrand Hamiltonian is

$$\begin{aligned}
&H_i(x,m,\hat{V}_{i,m},\hat{V}_{i,xm},\hat{V}_{i,xmm})\\
&= \inf_{u_i}\Big\{\tfrac{1}{2}q_i(x-\bar{x})^2 + \tfrac{1}{2}[q_i+\bar{q}_i]\bar{x}^2 + \tfrac{1}{2}r_i(u_i-\bar{u}_i)^2 + \tfrac{1}{2}(r_i+\bar{r}_i)[\bar{u}_i]^2\\
&+[a(x-\bar{x}) + (a+\bar{a})\bar{x} + \textstyle\sum_{i=1}^n b_i(u_i-\bar{u}_i) + \sum_{i=1}^n (b_i+\bar{b}_i)\bar{u}_i]\hat{V}_{i,xm}\Big\}\\
&+\tfrac{\sigma^2}{2}\hat{V}_{i,xxm} + \int_\Theta[\hat{V}_{i,m}(t_-,x+\mu) - \hat{V}_{i,m} - \mu\hat{V}_{i,xm}]\nu(d\theta).
\end{aligned} \tag{35}$$

It is important to notice that the last term in the integrand Hamiltonian $\int_\Theta[\hat{V}_{i,m}(t,x+\mu) - \hat{V}_{i,m} - \mu\hat{V}_{i,xm}]\nu(d\theta)$ is coming from the jump process involved in the state dynamics. From this Hamiltonian. we deduce that, generically, the optimal strategy is state-and-mean-field feedback form, as the RHS of (35) is. We now solve explicitly McKean–Vlasov integro-partial differential equation above. Inspired by the structure of the final payoff $\hat{V}_i(T,m)$, we choose a guess functional in the following form:

$$\hat{V}_i(t,m) = \tfrac{\alpha_i}{2}\int_y (y-\bar{m})^2\, m(dy) + \tfrac{\beta_i}{2}\left[\int_y y\, m(dy)\right]^2 + \delta_i. \tag{36}$$

The reader may ask why the term $\tilde{\gamma}_i[(x-\bar{m})\,]\left[\int_y y\, m(dy)\right]$ is missing in the guess functional. This is because we are looking for the expected value optimization (risk-neutral case), and its expected value is zero. The term $\gamma_i\left[\int_y y\, m(dy)\right]$ does not appear because there is no constant shift in the drift and no cross-terms in the loss function.

We now utilize the functional directional derivative. Consider another measure $\tilde{m} \in L^2$ and compute $\hat{V}_i(t,m+\epsilon\tilde{m})$.

$$\begin{aligned}
&\hat{V}_i(t,m+\epsilon\tilde{m}) = \tfrac{\alpha_i}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}})^2\, m(dy) + \tfrac{\epsilon\alpha_i}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}})^2\, \tilde{m}(dy)\\
&+\tfrac{\beta_i}{2}\left[\int_y y\, m(dy) + \epsilon\int_y y\, \tilde{m}(dy)\right]^2 + \delta_i.
\end{aligned} \tag{37}$$

Differentiating the latter term with respect to ϵ yields

$$\begin{aligned}
&\tfrac{d}{d\epsilon}\hat{V}_i(t,m+\epsilon\tilde{m})\\
&= \tfrac{-2\alpha_i\bar{\tilde{m}}(t)}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}})\, m(t,dy)\\
&-\tfrac{2\epsilon^2\alpha_i\bar{\tilde{m}}}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}})\, \tilde{m}(t,dy)\\
&+\tfrac{\alpha_i}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}})^2\, \tilde{m}(dy)\\
&+\tfrac{2\beta_i\bar{\tilde{m}}}{2}\left[\int_y y\, m(dy) + \epsilon\int_y y\, \tilde{m}(dy)\right]\\
&= \tfrac{-2\alpha_i\bar{\tilde{m}}}{2}\int_y (y-\bar{m})\, m(dy)\\
&+\tfrac{\alpha_i}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}}(t))^2\, \tilde{m}(dy)\\
&+\tfrac{2\beta_i\bar{\tilde{m}}}{2}\left[\int_y y\, m(dy)\right]\\
&= \tfrac{\alpha_i}{2}\int_y (y-\bar{m}-\epsilon\bar{\tilde{m}})^2\, \tilde{m}(dy) + \tfrac{2\beta_i\bar{\tilde{m}}}{2}\left[\int_y y\, m(dy)\right].
\end{aligned} \tag{38}$$

We deduce the following equalities:

$$\begin{aligned}
&\hat{V}_{i,m}[t,\tilde{m}](x) = \tfrac{\alpha_i}{2}\left[x - \int_y y\, m(dy)\right]^2 + \beta_i x\left[\int_y y\, m(dy)\right],\\
&\hat{V}_{i,xm}[t,m](x) = \alpha_i\left[x - \int_y y\, m(dy)\right] + \beta_i\left[\int_y y\, m(dy)\right],\\
&\hat{V}_{i,xxm}[t,m](x) = \alpha_i,\\
&\hat{V}_{i,m}(t,x+\mu) - \hat{V}_{i,m} - \mu\hat{V}_{i,xm} = \tfrac{1}{2}\alpha_i\mu^2;
\end{aligned} \tag{39}$$

$$\begin{aligned}
&\tfrac{1}{2}r_i(u_i-\bar{u}_i)^2 + \tfrac{1}{2}(r_i+\bar{r}_i)[\bar{u}_i]^2 + [\textstyle\sum_{i=1}^n b_i(u_i-\bar{u}_i) + \sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_i]\hat{V}_{i,xm}\\
&= \tfrac{1}{2}r_i(u_i-\bar{u}_i)^2 + \tfrac{1}{2}(r_i+\bar{r}_i)[\bar{u}_i]^2 + [\textstyle\sum_{i=1}^n b_i(u_i-\bar{u}_i)]\left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm} + E\hat{V}_{i,xm}\right\}\\
&+[\textstyle\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j]\left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm} + E\hat{V}_{i,xm}\right\}\\
&= \tfrac{1}{2}r_i(u_i-\bar{u}_i)^2 + \tfrac{1}{2}(r_i+\bar{r}_i)[\bar{u}_i]^2 + \left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm}\right\}\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j)\\
&+[E\hat{V}_{i,xm}]\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j) + \left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm}\right\}\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j\\
&+[E\hat{V}_{i,xm}]\textstyle\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j,
\end{aligned} \tag{40}$$

where we have used the following orthogonal decomposition:

$$\hat{V}_{i,xm} = \hat{V}_{i,xm} - E\hat{V}_{i,xm} + E\hat{V}_{i,xm}.$$

Noting that the expected value of the following term

$$[E\hat{V}_{i,xm}]\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j) + \left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm}\right\}\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j \tag{41}$$

is zero, the optimization yields the optimization of

$$\begin{aligned}
&\tfrac{1}{2}r_i(u_i-\bar{u}_i)^2 + \tfrac{1}{2}(r_i+\bar{r}_i)[\bar{u}_i]^2 + \left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm}\right\}\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j)\\
&+[E\hat{V}_{i,xm}]\textstyle\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j.
\end{aligned} \tag{42}$$

Thus, the equilibrium strategy of decision-maker i is

$$\begin{aligned}
&u_i^* = \bar{u}_i - \tfrac{b_i}{r_i}\left\{\hat{V}_{i,xm} - E\hat{V}_{i,xm}\right\} = \bar{u}_i - \tfrac{b_i}{r_i}\alpha_i\left[x - \int_y y\, m(t,dy)\right]\\
&= -\tfrac{b_i+\bar{b}_i}{r_i+\bar{r}_i}\beta_i\left[\int_y y\, m(t,dy)\right] - \tfrac{b_i}{r_i}\alpha_i\left[x - \int_y y\, m(t,dy)\right],\\
&\bar{u}_i^* = -\tfrac{b_i+\bar{b}_i}{r_i+\bar{r}_i}E[\hat{V}_{i,xm}] = -\tfrac{b_i+\bar{b}_i}{r_i+\bar{r}_i}\beta_i\left[\int_y y\, m(t,dy)\right],
\end{aligned} \tag{43}$$

which are exactly the expressions of the optimal strategies obtained in (15). Based on the latter expressions, we refine our statement. The optimal strategy is state-and-(mean of) mean-field feedback form.

We now solve explicitly the McKean–Vlasov integro-partial differential equation above.

$$\begin{aligned}
&\tilde{H}_i(x,m,\hat{V}_{i,m},\hat{V}_{i,xm},\hat{V}_{i,xmm})\\
&= \tfrac{1}{2}q_i(x-\bar{x})^2 + \tfrac{1}{2}[q_i+\bar{q}_i]\bar{x}^2\\
&+a(x-\bar{x})(\hat{V}_{i,xm} - E\hat{V}_{i,xm}) + (a+\bar{a})\bar{x}E\hat{V}_{i,xm}\\
&+\tfrac{\sigma^2}{2}\hat{V}_{i,xxm} + \textstyle\int_\Theta[\hat{V}_{i,m}(t,x+\mu) - \hat{V}_{i,m} - \mu\hat{V}_{i,xm}]\nu(d\theta)\\
&+\inf_{u_i}\left\{\tfrac{1}{2}r_i(u_i-\bar{u}_i)^2 + \tfrac{1}{2}(r_i+\bar{r}_i)[\bar{u}_i]^2 + [\hat{V}_{i,xm} - E\hat{V}_{i,xm}]\textstyle\sum_{j=1}^n b_j(u_j-\bar{u}_j) + [E\hat{V}_{i,xm}]\sum_{j=1}^n (b_j+\bar{b}_j)\bar{u}_j\right\}\\
&= \tfrac{1}{2}q_i(x-\bar{x})^2 + \tfrac{1}{2}[q_i+\bar{q}_i]\bar{x}^2 + \tfrac{1}{2}2a\alpha_i(x-\bar{x})^2 + \tfrac{1}{2}2(a+\bar{a})\beta_i\bar{x}^2 + \tfrac{\sigma^2}{2}\alpha_i + \textstyle\int_\Theta \tfrac{1}{2}\mu^2(t,\theta)\alpha_i\nu(d\theta)\\
&+\tfrac{1}{2}\tfrac{b_i^2}{r_i}\alpha_i^2(x-\bar{x})^2 + \tfrac{1}{2}\tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\beta_i^2\bar{x}^2 - \alpha_i(x-\bar{x})^2\textstyle\sum_{j=1}^n \tfrac{b_j^2}{r_j}\alpha_j - \beta_i\bar{x}^2\sum_{j=1}^n \tfrac{(b_j+\bar{b}_j)^2}{(r_j+\bar{r}_j)}\beta_j\\
&= \tfrac{1}{2}\{2a\alpha_i - 2\alpha_i\textstyle\sum_{j=1}^n \tfrac{b_j^2}{r_j}\alpha_j + \tfrac{b_i^2}{r_i}\alpha_i^2 + q_i\}(x-\bar{x})^2\\
&+\tfrac{1}{2}\left\{2(a+\bar{a})\beta_i - 2\beta_i\textstyle\sum_{j=1}^n \tfrac{(b_j+\bar{b}_j)^2}{(r_j+\bar{r}_j)}\beta_j + \tfrac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\beta_i^2 + q_i + \bar{q}_i\right\}\bar{x}^2\\
&+\tfrac{1}{2}\alpha_i[\sigma^2 + \textstyle\int_\Theta \mu(t,\theta)\nu(d\theta)].
\end{aligned} \tag{44}$$

Using the time derivative of $\hat{V}_i(t,m)$,

$$\hat{V}_{i,t} = \frac{\dot{\alpha}_i}{2}\int_y (y-\bar{m})^2\, m(dy) + \frac{\dot{\beta}_i}{2}\left[\int_y y\, m(dy)\right]^2 + \dot{\delta}_i, \tag{45}$$

and identifying the coefficients, we arrive at

$$\begin{aligned}
&\dot{\alpha}_i + 2a\alpha_i - 2\alpha_i \textstyle\sum_{j=1}^n \frac{b_j^2}{r_j}\alpha_j + \frac{b_i^2}{r_i}\alpha_i^2 + q_i = 0,\\
&\alpha_i(T) = q_i(T),\\
&\dot{\beta}_i + 2(a+\bar{a})\beta_i - 2\beta_i \textstyle\sum_{j=1}^n \frac{(b_j+\bar{b}_j)^2}{(r_j+\bar{r}_j)}\beta_j + \frac{(b_i+\bar{b}_i)^2}{r_i+\bar{r}_i}\beta_i^2 + q_i + \bar{q}_i = 0,\\
&\beta_i(T) = q_i(T) + \bar{q}_i(T),\\
&\dot{\delta}_i + \tfrac{1}{2}\alpha_i[\sigma^2 + \textstyle\int_\Theta \mu(t,\theta)\nu(d\theta)] = 0,\\
&\delta_i(T) = 0.
\end{aligned} \tag{46}$$

We retrieve the expressions in (15), confirming the validity of our approach.

6. Conclusions

In this article, we have shown that a mean-field equilibrium can be determined in a semi-explicit way for the linear–quadratic game problem where the Brownian motion is replaced by a jump-diffusion process in which the drift is of mean-field type. The method does not require the sophisticated non-elementary extension (33) to backward–forward systems. It does not need IPDEs (33). It does not need SMPs. It is basic and applies the expectation of Itô's formula. The use of this simple method may open the accessibility of the tool to a broader audience including beginners and engineers to this emerging field of mean-field-type game theory. In our future work, we would like to investigate the extension of the method to include common noise, action-and-state-dependent jump-diffusion coefficients, matrix forms, operator forms, jump-fractional noise, and risk-sensitivity, among other interesting aspects [38,39]. Additionally, it would be interesting to investigate the non-quadratic mean-field-dependent setting, as Duncan et al. [24] have extended the direct method to some non-quadratic cost functionals on spheres, torus, and more general spaces. We also would like to investigate how the explicit solution provided in this article can be used to improve numerical methods in mean-field-type game theory.

Acknowledgments: This research work is supported by U.S. Air Force Office of Scientific Research under grant number FA9550-17-1-0259 and this research supported by NSF grant DMS 1411412 and AFOSR grant FA9550-17-1-0073.

Author Contributions: T.E.D. and H.T. conceived and designed the model; performed the analysis and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Difference with the Mean-Field Game Approach

We would like to highlight that the variance reduction problem is fundamentally different from the classical risk-neutral mean-field game approach. In the classical mean-field game literature, the mean-field term is frozen, as it is assumed to be resulting from an infinite number of decision-makers. However, when one wants to reduce the variance on own-state, one cannot freeze the mean state as it is resulting from own-control. A change in the individual action of a deviant decision-maker changes its own-state and therefore it changes its own mean state. To illustrate the difference between mean-field games and mean-field-type games approaches, we consider a basic example below. Through a simple example, we show that for a wide range of parameters, the mean-field game approach is sub-optimal and leads a much higher risk than the mean-field-type game approach.

Let $q > 0, \bar{q} > 0$ and consider the following mean-field game problem:

$$(mfg) \begin{cases} \inf_{u_i(\cdot)\in\mathcal{U}_i} E\left[qx_i^2(T) + \bar{q}m^2(T) + \int_0^T u_i^2 dt\right], \\ \text{subject to} \\ dx_i(t) = u_i(t)dt + x_i(t)dB_i(t), \\ x_i(0) \in \mathbb{R}, \\ m(t) = \liminf_{n\to+\infty} \frac{1}{n}\sum_{k=1}^n x_k(t). \end{cases} \tag{A1}$$

In the problem (mfg), the pair (u_i, x_i) of an individual decision-maker alone does not affect the limiting mean state m in (A1). The solution of the mean-field game problem is

$$(mfg) \begin{cases} u_i^{mfg}(t) = -\alpha(t)x_i(t), \\ m(t) = m(0)e^{-\int_0^t \alpha(s)ds} \\ L_{mfg} = E\alpha(0)x_i(0)^2 + \gamma(0), \\ \dot{\alpha} + \alpha - \alpha^2 = 0,\ \alpha(T) = q > 0. \end{cases} \tag{A2}$$

As we can see by freezing the mean-field term, the achieved cost is $L_{mfg} = \alpha(0)var(x(0)) + \alpha(0)[Ex(0)]^2 + \gamma(0)$.

Now consider the following one-decision-maker (say decision-maker i) problem:

$$(mftg) \begin{cases} \inf_{u_i\in\mathcal{U}_i} E\left[qx_i^2(T) + \bar{q}(\bar{x}_i)^2(T) + \int_0^T u_i^2 dt\right], \\ \text{subject to} \\ dx_i(t) = u_i(t)dt + x_i(t)dB_i(t), \\ x_i(0) \in \mathbb{R}, \\ \bar{x}_i(t) = E[x_i(t)]. \end{cases} \tag{A3}$$

In the problem (mftg), the pair (u_i, x_i) of an individual decision-maker alone does significantly affect the mean state $\bar{x}_i$ in (A3):

$$(mftg) \begin{cases} u_i^{mftg}(t) = -\alpha(t)(x_i(t) - \bar{x}_i(t)) - \beta(t)\bar{x}_i(t), \\ Eu_i^{mftg}(t) = -\beta(t)\bar{x}_i(t), \\ \bar{x}_i(t) = \bar{x}_i(0)e^{-\int_0^t \beta(s)ds}, \\ L_{mftg} = \alpha(0)var(x_i(0)) + \beta(0)[\bar{x}_i(0)]^2, \\ \dot{\alpha} + \alpha - \alpha^2 = 0,\ \alpha(T) = q > 0, \\ \dot{\beta} + \alpha - \beta^2 = 0,\ \beta(T) = q + \bar{q} > 0. \end{cases} \tag{A4}$$

When we do not freeze the mean-field term, the achieved cost is L_{mftg}. Taking the difference between the ordinary differential equations, it is not difficult to see that $\alpha - \beta$ satisfies

$$\begin{cases} \frac{d}{dt}[\alpha - \beta] \le (\alpha - \beta)(\alpha + \beta),\ \ t < T, \\ (\alpha - \beta)(T) = -\bar{q} < 0. \end{cases} \tag{A5}$$

Since $q < q + \bar{q}$ for $\bar{q} > 0$, it implies $\alpha(t) < \beta(t)$. We would like to compare $\alpha(0)[Ex(0)]^2 + \gamma(0)$ and $\beta(0)[\bar{x}(0)]^2$.

$$\begin{cases} L_{mfg} - L_{mftg} \\ = \alpha(0)[Ex(0)]^2 + \gamma(0) - \beta(0)[\bar{x}(0)]^2 \\ = [\alpha(0) - \beta(0)][Ex(0)]^2 + \gamma(0) \\ = [\alpha(0) - \beta(0)][Ex(0)]^2 + \bar{q}[Ex(0)]^2 e^{-2\int_0^T \alpha(s)ds} \\ = [\alpha(0) - \beta(0) + \bar{q}e^{-2\int_0^T \alpha(s)ds}][Ex(0)]^2. \end{cases} \tag{A6}$$

The latter expression being positive, we deduce that for any $T > 0 : L_{mfg} > L_{mftg}$. Thus, in this example, the mean-field game approach—which consists of freezing the mean-field term—is sub-optimal. On the other hand, the mean-field-type game approach coincides with the global optimization problem in the one-decision-maker case. Hence, (A4) is the global optimum.

Difference with Multi-Population Mean-Field Games

The model studied here differs from (non-cooperative) multi-population (multi-class or multi-type) mean-field games. In multi-population mean-field games, it is usually assumed that there is an infinite number of decision-makers, each of them having their own control action. In those models, a single decision-maker does not influence the population mean state within its class, since the class size is assumed infinite. On the other hand, in the mean-field-type game model presented here, there is a finite number of "true" decision-makers, and each decision-maker does have a non-negligible effect on the mean-field terms.

References

1. Markowitz, H.M. Portfolio Selection. *J. Financ.* **1952**, *7*, 77–91.
2. Roos, C.F. A Mathematical Theory of Competition. *Am. J. Math.* **1925**, *47*, 163–175.
3. Roos, C.F. A Dynamic Theory of Economics. *J. Polit. Econ.* **1927**, *35*, 632–656.
4. Markowitz, H.M. The Utility of Wealth. *J. Politi. Econ.* **1952**, *60*, 151–158.
5. Markowitz, H.M. *Portfolio Selection: Efficient Diversification of Investments*; John Wiley & Sons: New York, NY, USA, 1959.
6. Duncan, T.E.; Pasik-Duncan, B. Solvable stochastic differential games in rank one compact symmetric spaces. *Int. J. Control* **2017**, doi:10.1080/00207179.2016.1269947.
7. Duncan, T.E. Linear Exponential Quadratic Stochastic Differential Games. *IEEE Trans. Autom. Control* **2016**, *61*, 2550–2552.
8. Duncan, T.E.; Pasik-Duncan, B. Linear-quadratic fractional Gaussian control. *SIAM J. Control Optim.* **2013**, *51*, 4604–4619.
9. Duncan, T.E.; Pasik-Duncan, B. Linear-exponential-quadratic control for stochastic equations in a Hilbert space. *Dyn. Syst. Appl.* **2012**, *21*, 407–416.
10. Duncan, T.E. Linear-exponential-quadratic Gaussian control. *IEEE Trans. Autom. Control* **2013**, *58*, 2910–2911.
11. Duncan, T.E. Linear-quadratic stochastic differential games with general noise processes. In *Models and Methods in Economics and Management Science*; El Ouardighi, F., Kogan, K., Eds.; Operations Research and Management Series; Springer International Publishing: Cham, Switzerland, 2014; Volume 198, pp. 17–26.
12. Bensoussan, A.; Frehse, J.; Yam, S.C.P. *Mean Field Games and Mean Field Type Control Theory*; Springer: Berlin, Germany, 2013.
13. Lasry, J.M.; Lions, P.L. Mean field games. *Jpn. J. Math.* **2007**, *2*, 229–260.
14. Bensoussan, A.; Djehiche, B.; Tembine, H.; Yam, P. Risk-sensitive mean-field-type control. *arXiv* **2017**, arXiv:1702.01369.
15. Bardi, M. Explicit solutions of some linear-quadratic mean field games. *Netw. Heterog. Media* **2012**, *7*, 243–261.
16. Bardi, M.; Priuli, F.S. Linear-Quadratic *N*-person and Mean-Field Games with Ergodic Cost. *SIAM J. Control Optim.* **2014**, *52*, 3022–3052.
17. Tembine, H. Mean-field-type games. *AIMS Math.* **2017**, *2*, 706–735.
18. Djehiche, B.; Tembine, H. On the Solvability of Risk-Sensitive Linear-Quadratic Mean-Field Games. *arXiv* **2014**, arXiv:1412.0037v1.
19. Tembine, H.; Zhu, Q.; Basar, T. Risk-sensitive mean-field games. *IEEE Trans. Autom. Control* **2014**, *59*, 835–850.
20. Tembine, H.; Bauso, D.; Basar, T. Robust linear quadratic mean-field games in crowd-seeking social networks. In Proceedings of the 52nd Annual Conference on Decision and Control, Florence, Italy, 10–13 December 2013; pp. 3134–3139.
21. Kolokoltsov, V.N. Nonlinear Markov Games on a Finite State Space (Mean-field and Binary Interactions). *Int. J. Stat. Probab.* **2012**, *1*, 77–91.
22. Engwerda, J.C. On the open-loop Nash equilibrium in LQ-games. *J. Econ. Dyn. Control* **1998**, *22*, 729–762.

23. Bensoussan, A. Explicit solutions of linear quadratic differential games. In *Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems*; International Series in Operations Research and Management Science; Springer: Berlin, Germany, 2006; Volume 94, pp. 19–34.
24. Duncan, T.E.; Pasik-Duncan, B. A direct method for solving stochastic control problems. *Commun. Inf. Syst.* **2012**, *12*, 1–14.
25. Lukes, D.L.; Russell, D.L. A Global Theory of Linear-Quadratic Differential Games. *J. Math. Anal. Appl.* **1971**, *33*, 96–123.
26. Tembine, H. *Distributed Strategic Learning for Wireless Engineers*; CRC Press: Boca Raton, FL, USA, 2012; 496 p.
27. Tembine, H. Risk-sensitive mean-field-type games with p-norm drifts. *Automatica* **2015**, *59*, 224–237.
28. Djehiche, B.; Tembine, H.; Tempone, R. A Stochastic Maximum Principle for Risk-Sensitive Mean-Field-Type Control. *IEEE Trans. Autom. Control* **2014**, *60*, 2640–2649.
29. Dockner, E.J.; Jorgensen, S.; Van Long, N.; Sorger, G. *Differential Games in Economics and Management Science*; Cambridge University Press: Cambridge, UK, 2000.
30. Tembine, H. Nonasymptotic mean-field games. *IEEE Trans. Syst. Man Cybern. B Cybern.* **2014**, *44*, 2744–2756.
31. Djehiche, B.; Tcheukam, A.; Tembine, H. Mean-field-type games in engineering. *AIMS Electron. Electr. Eng.* **2017**, *1*, 18–73.
32. Başar, T.; Djehiche, B.; Tembine, H. *Mean-Field-Type Game Theory*; Springer: New York, NY, USA, 2019; under preparation.
33. Tembine, H. Energy-constrained mean-field games in wireless networks. *Strateg. Behav. Environ.* **2014**, *4*, 187–211.
34. Nash, J. The Bargaining Problem. *Econometrica* **1950**, *18*, 155–162.
35. Dubey, P. Inefficiency of Nash equilibria. *Math. Operat. Res.* **1986**, *11*, 1–8.
36. Koutsoupias, E.; Papadimitriou, C. Worst-case Equilibria. In Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science, Trier, Germany, 4–6 March 1999; pp. 404–413.
37. Koutsoupias, E.; Papadimitriou, C. Worst-case Equilibria. *Comput. Sci. Rev.* **2009**, *3*, 65–69.
38. Tsiropoulou, E.E.; Vamvakas, P.; Papavassiliou, S. Energy Efficient Uplink Joint Resource Allocation Non-cooperative Game with Pricing. In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC 2012), Shanghai, China, 1–4 April 2012.
39. Musku, M.R.; Chronopoulos, A.T.; Popescu, D.C. Joint rate and power control with pricing. In Proceedings of the IEEE Global Telecommunications Conference, St. Louis, MO, USA, 28 November–2 December 2005.