Convex Analysis for LQG Systems with Applications to Major Minor LQG Mean Field Game Systems

Dena Firoozi, Sebastian Jaimungal, Peter E. Caines

1 Abstract

A convex analysis method is used to rederive the solutions to LQG optimal control problems. Then the methodology is applied to major minor LQG mean field game (MM LQG MFG) systems to retrieve the best response strategies for the major agent and each individual minor agent which collectively yield an $\epsilon$-Nash equilibrium for the entire system.

2 Introduction

In the literature, various approaches such as calculus of variations, (stochastic) maximum principle, dynamic programming, and change of functional have been used to address deterministic linear quadratic (LQ) and stochastic linear quadratic (LQG) optimal control problems [1-3].

In a convex analysis approach to optimization for static systems, the Gâteaux derivative of the functional to be optimized is used to solve the problem (see e.g., [4], [5]). In [6], the relationship between the Gâteaux derivative of the cost functional of a dynamic system and its Hamiltonian is established. A stochastic tracking problem in finance is studied in [7] using the convex analysis approach, while an algorithmic trading problem is investigated in [8] and the best response trading strategies are obtained for a large number of heterogeneous traders using the convex analysis approach.

In this work, a convex analysis method is used to rederive the solutions to LQG optimal control problems. Then the methodology is applied to major minor LQG
mean field game (MM LQG MFG) systems to retrieve the best response strategies for the major agent and each individual minor agent addressed in [9].

3 Convex Analysis

Let $V$ be a reflexive Banach space with the dual space $V^*$ and $\mathcal{V}$ be a non-empty closed convex subset of $V$.

**Definition 1** (Gâteaux Derivative). The function $J$ defined on a neighbourhood of $u \in V$ with values in $\mathbb{R}$ is differentiable in the sense of Gâteaux at $u$ in the direction of $\omega$, if there exists $J'(u) \in V^*$ such that

$$
\langle J'(u), \omega \rangle = \lim_{\epsilon \to 0} \frac{J(u + \epsilon \omega) - J(u)}{\epsilon}.
$$

(1)

The function $J'(u)$ is called the Gâteaux derivative of $J$ at $u$.

**Theorem 1** (Euler Inequality). Assume that the function $J$ is convex, continuous, proper, and Gâteaux differentiable with continuous derivative $J'(u)$. Then

$$
J(u) = \inf_{v \in \mathcal{U}} J(v),
$$

(2)

if and only if

$$
\langle J'(u), v - u \rangle \geq 0, \quad \forall v \in \mathcal{V}.
$$

(3)

□

Proof of Theorem 1 may be found in [4] and [5].

**Remark 1** (Euler Equality). In the case where $\mathcal{V} = V$, $\omega = v - u$ produces the whole space of $V$, and therefore (3) reduces to Euler equality

$$
\langle J'(u), \omega \rangle = 0, \quad \forall \omega \in V,
$$

(4)

which implies that

$$
J'(u) = 0.
$$

(5)

We note that in this paper $V$ is the Hilbert space $\mathbb{R}^m$ whose norm is determined by an inner product.
4 Single-Agent LQG Problems

In this section, the solutions to single-agent LQG problems are rederived using a convex analysis method.

4.1 Dynamics

Consider single-agent LQG systems with governing dynamics

\[
dx_t = (Ax_t + Bu_t + b(t))dt + \sigma(t)dw_t,
\]

where \( t \geq 0, x_t \in \mathbb{R}^n, u_t \in \mathbb{R}^m, \) and \( w_t \in \mathbb{R}^r \) denote, respectively, the state, the control action, and a standard Wiener process. Moreover, \( A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, \) and \( b(t) \in \mathbb{R}^n, \sigma(t) \in \mathbb{R}^{n \times r}, \) are deterministic functions of time.

4.1.1 Control \( \sigma \)-Fields

We denote by \( \mathcal{F} := (\mathcal{F}_t)_{t \in [0,T]} \) the natural filtration generated by the agent’s state \( (x_t)_{t \in [0,T]} \). Then, we introduce the admissible control set \( \mathcal{U} \) to be the set of feedback control laws with second moment lying in \( L^1[0,T] \), for any finite \( T \), which are adapted to \( \mathcal{F} \).

4.2 Cost Functional

The cost functional to be minimized is given by

\[
J(u) = \frac{1}{2} \mathbb{E} \left[ e^{-\rho T} x_T^T G x_T + \int_0^T e^{-\rho t} \left\{ x_t^T Q x_t + 2 x_t^T N u_t + u_t^T R u_t - 2 x_t^T \eta - 2 u_t^T \eta \right\} dt \right],
\]

where \( \rho \) denotes the discount rate.

**Assumption 1.** For cost the functional (7) to be convex, it is assumed that \( G \geq 0, R > 0, \) and \( Q - NR^{-1}N^T > 0 \).

4.3 Optimal Control Action

The system dynamics (6) together with the cost functional (7) constitute an LQ stochastic optimal control problem, which is solved for using the following theorem.
Theorem 2. (Gâteaux Derivative of Cost for LQG Systems) For system (6)-(7), the Gâteaux derivative of the cost functional is given by

\[
\langle J'(u), \omega \rangle = \mathbb{E} \left[ \int_0^T \omega^T \left\{ e^{-\rho_t} N^T x_t^u + e^{-\rho_t} R u_t - e^{-\rho_t} n \right. \right.
\]
\[
+ B^T \left( e^{-A^T t} M_t - \int_0^t e^{-\rho_s} e^{A^T (s-t)} (Q x_s^u + N u_s - \eta) ds \right) \left. \right\} dt \bigg| \mathcal{F}_t \bigg].
\]

where \( M_t \) is a martingale process given by

\[
M_t = \mathbb{E} \left[ e^{-\rho T} e^{A^T T} G x_T^u + \int_0^T e^{-\rho_s} e^{A^T s} (Q x_s^u + N u_s - \eta) ds \bigg| \mathcal{F}_t \right].
\]

Proof. The Gâteaux derivative \( J'(u) \) of (7) is computed as follows.

The solution \( x_t^u \) to the state representation of the system (6) subject to the control action \( u_t \) is given by

\[
x_t^u = e^{A t} x_0 + \int_0^t e^{A(t-s)} \left( B u_s + b(s) \right) ds + \int_0^t e^{A(t-s)} \sigma(s) dw_s,
\]

where \( x_0 \in \mathbb{R}^n \) and \( \phi(t, s) = e^{A(t-s)}, s \leq t \leq T \), denote, respectively, the initial state and the state transition matrix for the system (6).

Let \( x_t^{u+\epsilon \omega} \) denote the solution to (6) subject to a perturbed control action \( u_t + \epsilon \omega_t \) in the direction of \( \omega_t \) given by

\[
x_t^{u+\epsilon \omega} = e^{A t} x_0 + \int_0^t e^{A(t-s)} \left( B u_s + b(s) \right) ds + \int_0^t e^{A(t-s)} \sigma(s) dw_s + \epsilon \int_0^t e^{A(t-s)} B \omega_s ds.
\]

To find the relation between \( x_t^u \) and \( x_t^{u+\epsilon \omega} \), (10) is substituted in (11) which yields

\[
x_t^{u+\epsilon \omega} = x_t^u + \epsilon \int_0^t e^{A(t-s)} B \omega_s ds.
\]

Then by differentiating both sides of (12), the evolution of \( x^{u+\epsilon \omega}(t) \) in terms of \( x^u(t) \) is given by

\[
dx_t^{u+\epsilon \omega} = dx_t^u + \epsilon B \omega_t dt + \epsilon A \int_0^t e^{A(t-s)} B \omega_s ds.
\]
The cost induced by the perturbed control action $u_t + \epsilon \omega_t$ and, subsequently, the perturbed state $x_t^{u+\omega}$ is given by

$$J(u + \epsilon \omega) = \frac{1}{2} \mathbb{E} \left[ e^{-\rho T} (x_T^{u+\omega})^T G x_T^{u+\omega} + \int_0^T e^{-\rho s} \left\{ (x_s^{u+\omega})^T Q x_s^{u+\omega} + 2(x_s^{u+\omega})^T N (u_s + \epsilon \omega_s) + (u_s + \epsilon \omega_s)^T R (u_s + \epsilon \omega_s) - 2(x_s^{u+\omega})^T \eta - 2(u_s + \epsilon \omega_s)^T n \right\} ds \right],$$

(14)

where the terminal cost, by utilizing the integration by parts technique, can be presented in integral form as

$$e^{-\rho T} (x_T^{u+\omega})^T G x_T^{u+\omega} = (x_0)^T G x_0 + \int_0^T \rho \left[ e^{-\rho s} (x_s^{u+\omega})^T G x_s^{u+\omega} \right] ds.$$

(15)

To write $J(u + \epsilon \omega)$ in terms of $J(u_0)$, $u_t$ and $x_t^u$, first (15), and then (12)-(13) are substituted in (14) which gives rise to

$$J(u + \epsilon \omega) = J(u) + \epsilon \int_0^T e^{-\rho s} \left\{ \left( \int_0^s e^{A(s-t)} B \omega_t dt \right)^T (Gdx^u_s + (Qx^u_s + Nu_s + ATGx^u_s - \rho Gx^u_s - \eta) ds) + (x_s^u)^T N \omega_s + (x_s^u)^T GB \omega_s + (u_s)^T R \omega_s - n^T \omega_s \right\} ds.$$

(16)

Then the Gâteaux derivative of $J'(u)$ in the direction of $\omega$ is obtained by first taking $J(u)$ to the left hand side of (16), dividing both sides of the equation by $\epsilon$, and finally taking the limit as $\epsilon \rightarrow 0$, which yields

$$\langle J'(u), \omega \rangle = \int_0^T e^{-\rho s} \left\{ \left( \int_0^s e^{A(s-t)} B \omega_t dt \right)^T (Gdx^u_s + (Qx^u_s + Nu_s + ATGx^u_s - \rho Gx^u_s - \eta) ds) + ((x_s^u)^T N \omega_s + (x_s^u)^T GB \omega_s + (u_s)^T R \omega_s - n^T \omega_s) ds \right\}.$$

(17)
An application of Fubini’s theorem to change the order of integration in (17) results in

\[
\langle J'(u), \omega \rangle = E \left[ \int_0^T \omega_t^T \left\{ e^{-\rho T} B^T Gx_t^u + e^{-\rho T} N^T x_t^u + e^{-\rho T} Ru_t - e^{-\rho T} n \\
+ B^T \int_t^T e^{-\rho s} e^{A^T(s-t)} \left( Gdx_s^u + (A^T Gx_s^u - \rho Gx_s^u + Q x_s^u + Nu_s - \eta)ds \right) \right\} dt \right].
\] (18)

By using integration by parts again, we have

\[
\int_t^T e^{-\rho s} e^{A^T(s-t)} (A^T Gx_s ds + Gdx_s - \rho Gx_s^u ds) = \int_t^T d(e^{-\rho s} e^{A^T(s-t)} Gx_s) = e^{-\rho T} e^{A^T(T-t)} Gx_T - e^{-\rho T} Gx_t,
\] (19)

whose substitution in (18) yields

\[
\langle J'(u), \omega \rangle = E \left[ \int_0^T \omega_t^T \left\{ e^{-\rho T} B^T e^{A^T(T-t)} Gx_T^u + e^{-\rho T} N^T x_t^u + e^{-\rho T} Ru_t - e^{-\rho T} n \\
+ B^T \int_t^T e^{-\rho s} e^{A^T(s-t)} (Q x_s^u + Nu_s - \eta)ds \right\} dt \right].
\] (20)

Using the smoothing property of conditional expectation [10], the Gateaux derivative (20) may be rewritten as

\[
\langle J'(u), \omega \rangle = E \left[ \int_0^T \omega_t^T \left\{ e^{-\rho T} B^T e^{A^T(T-t)} Gx_T^u + e^{-\rho T} N^T x_t^u + e^{-\rho T} Ru_t - e^{-\rho T} n \\
+ B^T \mathbb{E} \left[ e^{-\rho T} e^{A^T(T-t)} Gx_T^u + \int_t^T e^{-\rho s} e^{A^T(s-t)} (Q x_s^u + Nu_s - \eta)ds \bigg| \mathcal{F}_t \right] \right\} dt \right].
\] (21)

Then the following martingale is defined

\[
M_t = E \left[ e^{-\rho T} e^{A^T s} Gx_T^u + \int_0^T e^{-\rho s} e^{A^T s} (Q x_s^u + Nu_s - \eta)ds \bigg| \mathcal{F}_t \right],
\] (22)
and is substituted in (21) to give

\[
\langle J'(u), \omega \rangle = \mathbb{E} \left[ \int_0^t \omega_t^T \left\{ e^{-\rho t} N^T x_t^u + e^{-\rho t} R u_t - e^{-\rho t} n \\
+ B^T \left( e^{-A^T t} M_t - \int_0^t e^{-\rho s} e^{A^T (s-t)} (Q x_s^u + N u_s - \eta) ds \right) \right\} dt \right].
\]

(23)

\[\Box\]

**Theorem 3** (LQG Optimal Control Action). Given Assumption [7] the optimal control action for LQG systems given by (6)-(7) is specified by

\[
u^*_t = -R^{-1} \left[ N^T x_t^* - n + B^T e^{\rho t} \left( e^{-A^T t} M_t - \int_0^t e^{-\rho s} e^{A^T (s-t)} (Q x_s^* + N u_s - \eta) ds \right) \right].
\]

(24)

**Proof.** As per Theorem [7] and Remark [7] the necessary condition for \( u^*(t) \) to be the optimal control is given by

\[
\langle J'(u^*), \omega \rangle = 0, \text{ a.s. for all possible paths of } \omega(t) \in U,
\]

which, according to (8), implies that

\[
u^*_t = -R^{-1} \left[ N^T x_t^* - n + B^T e^{\rho t} \left( e^{-A^T t} M_t - \int_0^t e^{-\rho s} e^{A^T (s-t)} (Q x_s^* + N u_s - \eta) ds \right) \right].
\]

(26)

Moreover, since Assumption [7] holds, (25) is the sufficient condition of optimality as well. \[\Box\]

**Theorem 4** (LQG State Feedback Optimal Control). For LQG systems governed by (6)-(7), the optimal control action is given by the linear state feedback control

\[
u_t = -R^{-1} B^T [\Pi(t) x_t + s(t)],
\]

(27)

where \( \Pi(t) \) and \( s(t) \) are given by

\[
\dot{\Pi}(t) + \Pi(t) A + A^T \Pi(t) - (B^T \Pi(t) + N)^T R^{-1} (B^T \Pi(t) + N) + Q = 0, \quad (28)
\]

\[
s(t) + [ (A - BR^{-1} N)^T - \Pi B R^{-1} B^T ] s(t) + \Pi(t) b(t) = 0. \quad (29)
\]

with terminal conditions \( \Pi(T) = G \) and \( s(T) = 0. \)
Proof. Let us define $p(t)$ as

$$p_t = e^{\rho t} (e^{-A^T t} M_t - \int_0^t e^{-\rho s} e^{A^T (s-t)} (Qx_s^* + Nu_s^* - \eta) ds),$$  \hspace{1cm} (30)

which is the adjoint process for the system (6)-(7) in the framework of stochastic maximum principle. Then the ansatz for $p_t^0$ is adopted to be

$$p_t = \Pi(t)x_t^* + s(t),$$  \hspace{1cm} (31)

and is substituted in (24) to give

$$u_t^* = -R^{-1} [Nx_t^* + B^T (\Pi(t)x_t^* + s(t))].$$  \hspace{1cm} (32)

To find $\Pi(t)$ and $s(t)$, both sides of (31) are first differentiated, and then (6) and (32) are substituted to yield

$$dp_t = \left[ (\dot{\Pi} + \Pi A - \Pi B R^{-1} N - \Pi B R^{-1} B^T \Pi) x_t^* - \Pi B R^{-1} B^T s(t) \right. \left. + \Pi b + \Pi B R^{-1} n + \dot{s}(t) \right] dt + \Pi \sigma(t) dw_t.$$  \hspace{1cm} (33)

Next, both sides of (30) are differentiated to give

$$dp_t = (\rho p_t - A^T p_t - Qx_t^* - Nu_t^* + \eta) dt + e^{\rho t} e^{-A^T t} dM_t,$$  \hspace{1cm} (34)

where according to the martingale representation theorem, the martingale $M_t$ may be written as

$$M_t = M_0 + \int_0^t Z_s dw_s,$$  \hspace{1cm} (35)

and hence

$$dM_t = z_t dw_t.$$  \hspace{1cm} (36)

Then, equations (31), (32) and (36) are substituted in (34) to get

$$dp_t = \left[ (\rho \Pi_0 - Q + NR^{-1} N^T + NR^{-1} B^T \Pi - A^T \Pi) x_t^* + \rho s(t) \right. \left. + (N^T R^{-1} B^T - A^T) s(t) + \eta - NR^{-1} n \right] dt + q_t dw_t.$$  \hspace{1cm} (37)
where \( q_t = e^{-A^T_t Z_t} \).

Finally, for (33) and (37) to be equal, the corresponding drifts and diffusions must be equal. Hence the following equations must hold

\[
q_t = \Pi \sigma(t), \quad (38)
\]

\[
\begin{align*}
\rho \Pi &= \dot{\Pi} + \Pi A + A^T \Pi - (B^T \Pi + N^T)^T R^{-1} (B^T \Pi + N) + Q, \\
\Pi(T) &= G, \quad (39)
\end{align*}
\]

\[
\begin{align*}
\rho s(t) &= \dot{s}(t) + [(A - BR^{-1} N^T)^T - \Pi BR^{-1} B^T]s(t) \\
&\quad + \Pi (b(t) + BR^{-1} n) + NR^{-1} n - \eta = 0, \quad (40)
\end{align*}
\]

\[
s(T) = 0,
\]

Remark 2 (Finite Horizon LQG Systems). Typically, the cost functional for finite horizon LQG systems is not discounted, i.e. \( \rho = 0 \), and hence the Riccati and offset equations (28)-(29) reduce to

\[
\begin{align*}
\dot{\Pi} &= \Pi A + A^T \Pi - (B^T \Pi + N^T)^T R^{-1} (B^T \Pi + N) + Q, \\
\dot{s}(t) &= [(A - BR^{-1} N^T)^T - \Pi BR^{-1} B^T]s(t) + \Pi (b(t) + BR^{-1} n) + NR^{-1} n - \eta,
\end{align*}
\]

subject to the terminal conditions \( \Pi(T) = G, s(T) = 0 \).

Remark 3 (Infinite Horizon LQG Systems). For infinite horizon LQG systems where the terminal time \( T \) in (7) is set to infinity, the terminal cost becomes zero. Hence, the infinite horizon cost functional is given by

\[
J(u) = \frac{1}{2} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left\{ x_t^T Q x_t + 2 x_t^T N u_t + u_t^T R u_t - 2 x_t^T \eta - 2 u_t^T n \right\} dt \right], \quad (42)
\]

The dynamics (6) remains the same in the infinite horizon LQG systems.

Assumption 2. The pair \((L, A - (\rho/2) I)\) is detectable where \( L = Q^{1/2} \).

Assumption 3. The pair \((A - (\rho/2) I, B)\) is stabilizable.

Given that Assumptions 2-3 hold, for infinite horizon LQG systems governed by (6) and (42), the optimal control action is given by (27), where the steady state Riccati matrix \( \Pi \) satisfies an algebraic Riccati equation given by

\[
\rho \Pi = \Pi A + A^T \Pi - (B^T \Pi + N^T)^T R^{-1} (B^T \Pi + N) + Q, \quad (43)
\]

and the steady state offset vector \( s_0 \) satisfies the differential equation

\[
\rho s(t) = \dot{s}(t) + [(A - BR^{-1} N^T)^T - \Pi BR^{-1} B^T]s(t) + \Pi (M(t) + BR^{-1} n) + NR^{-1} n - \eta. \quad (44)
\]
Major Minor LQG Mean Field Game Systems

In this section, the convex analysis method introduced in Section 4 is utilized to rederive the best response strategies for major minor LQG MFG problems addressed in [9]. A large population $N$ of minor agents with a major agent, where agents are subject to stochastic linear dynamics and quadratic cost functionals are considered. Each agent is coupled with other agents through their dynamics and cost functional with the average state of minor agents, i.e. the empirical mean field.

5.1 Dynamics

The dynamics of the major and minor agents are assumed to be given, respectively, by

$$dx_0^t = [A_0x_0^t + F_0x_i^{(N)} + B_0u_0^t + b_0(t)]dt + \sigma_0dw_0^t, \quad (45)$$

$$dx_i^t = [A_ix_i^t + F_ix_i^{(N)} + B_iu_i^t + b_i(t)]dt + \sigma_idw_i^t, \quad (46)$$

where $t \geq 0$, $i \in \mathcal{N}$, $\mathcal{N} = \{1, \ldots, N\}$, $N < \infty$, and the subscript $k$, $k \in \mathcal{K}$, $\mathcal{K} = \{1, \ldots, K\}$, $K \leq N$, denotes the type of a minor agent. Here $x_i^t \in \mathbb{R}^n$, $i \in \mathcal{N}_0$, $\mathcal{N}_0 = \{0, \ldots, N\}$, are the states, $u_i^t \in \mathbb{R}^m$, $i \in \mathcal{N}_0$ are the control inputs, $\{w_i^t, i \in \mathcal{N}_0\}$ denotes $(N + 1)$ independent standard Wiener processes in $\mathbb{R}^r$, where $w_i$ is progressively measurable with respect to the filtration $\mathcal{F}_t^w := (\mathcal{F}_t^w)_{t \in [0,T]}$. All matrices in (45) and (46) are constant and of appropriate dimension; vectors $b_0(t)$, and $b_i(t)$ are deterministic functions of time.

5.1.1 Agents types

Minor agents are given in $K$ distinct types with $1 \leq K < \infty$. The notation

$$\Psi_k \triangleq \Psi(\theta_i), \quad \theta_i = k$$

is introduced where $\theta_i \in \Theta$, with $\Theta$ being the parameter set, and $\Psi$ may be any dynamical parameter in (46) or weight matrix in the cost functional (49). The symbol $\mathcal{I}_k$ denotes

$$\mathcal{I}_k = \{i : \theta_i = k, i \in \mathcal{N}\}, \quad k \in \mathcal{K}$$

where the cardinality of $\mathcal{I}_k$ is denoted by $N_k = |\mathcal{I}_k|$. Then, $\pi^N = (\pi_1^N, \ldots, \pi_K^N)$, $\pi_k^N = \frac{N_k}{N}$, $k \in \mathcal{K}$, denotes the empirical distribution of the
parameters $(\theta_1, ..., \theta_N)$ sampled independently of the initial conditions and Wiener processes of the agents $A_i, i \in \mathcal{N}$. The first assumption is as follows.

**Assumption 4.** There exists $\pi$ such that $\lim_{N \to \infty} \pi^N = \pi$ a.s.

### 5.1.2 Control $\sigma$-Fields

We denote by $\mathcal{F}_i := (\mathcal{F}_i^t)_{t \in [0,T]}, i \in \mathcal{N}$, the natural filtration generated by the $i$-th minor agent’s state $(x_i^t)_{t \in [0,T]}$, by $\mathcal{F}^0 := (\mathcal{F}_i^0)_{t \in [0,T]}$ the natural filtration generated by the major agent’s state $(x_0^t)_{t \in [0,T]}$, and $\mathcal{F}^g := (\mathcal{F}_i^g)_{t \in [0,T]}$ the natural filtration generated by the states of all agents $((x_i^t)_{t \in \mathcal{N}}, x_0^t)_{t \in [0,T]}$.

Next, we introduce two admissible control sets. Let $\mathcal{U}_0^0$ denote the set of feedback control laws with second moment lying in $L^1[0,T]$, for any finite $T$, which are adapted to the local information set of the major agent $A_0$, i.e. $\mathcal{F}^0$. The set of control inputs $\mathcal{U}_i, i \in \mathcal{N}$, based upon the local information set of the minor agent $A_i, i \in \mathcal{N}$, consists of the feedback control laws adapted to the filtration $\mathcal{F}_i^r := (\mathcal{F}_i^r^t)_{t \in [0,T]}$, where $\mathcal{F}_i^r := \mathcal{F}_i^0 \vee \mathcal{F}^0, i \in \mathcal{N}$, while $\mathcal{U}_0^N$ is adapted to the general filtration $\mathcal{F}^g := (\mathcal{F}_i^g)_{t \in [0,T]}$, and the $L^1[0,T]$ constraint on second moments applies in each case.

### 5.2 Cost functionals

The individual (finite) large population finite horizon cost functional for the major agent is specified by

$$J_{0}^N(u^0, u^{-0}) = \frac{1}{2} \mathbb{E}\left[\|x_0^T - \Phi(x^{(N)}_T)\|^2_{G_0} + \int_0^T \left\{\|x_0^t - \Phi(x^{(N)}_t)\|^2_{Q_0} + 2(x_0^t - \Phi(x^{(N)}_t))^T N_0 u_0^t + \|u_0^t\|^2_{R_0}\right\} dt\right], \quad (47)$$

where

$$\Phi(.) := H_0 x^{(N)}_t + \eta_0. \quad (48)$$

**Assumption 5.** For the cost functional (47) to be convex, we assume that $G_0 \geq 0, R_0 > 0$, and $Q_0 - N_0 R_0^{-1} N_0^T > 0$.

The individual (finite) large population finite horizon cost functional for a minor agent $A_i, i \in \mathcal{N}$, is specified as
\[ J_i^N(u^i, u^{-i}) = \frac{1}{2} \mathbb{E} \left[ \| x_i^T - \Psi(x_i^{(N)}) \|^2_{G_k} + \int_0^T \left\{ \| x_i^t - \Psi(x_i^{(N)}) \|^2_{Q_k} \\
+ 2(x_i^t - \Psi(x_i^{(N)}))^T N_k u_i^t + \| u_i^t \|^2_{R_k} \right\} dt \right], \tag{49} \]

where
\[ \Psi(.) := H_k x_i^0 + \hat{H}_k x_i^{(N)} + \eta_k. \tag{50} \]

**Assumption 6.** For the cost functional (49) to be convex, we assume that \( G_k \geq 0, \) \( R_k > 0, \) and \( Q_k - N_k R_k^{-1} N_k^T > 0 \) for \( k \in \mathbb{K}. \)

We note that the major agent \( A_0 \) and minor agents \( A_i, \) \( i \in \mathfrak{N} \) are coupled with each other through the average term \( x_i^{(N)} = \frac{1}{N} \sum_{i=1}^N x_i^i \) in their dynamics and cost functionals given by, respectively, (45)-(46) and (47)-(49).

### 5.3 Solutions to Major Minor LQG MFG Problems

Following the mean field game methodology with a major agent [11], [9], the problem is first solved in the infinite population case where the average terms in the finite population dynamics and cost functional of each agent are replaced with their infinite population limit, i.e. the mean field. Then specializing to LQG MFG systems, the major agent’s state is extended with the mean field, while the minor agent’s state is extended with the major agent’s state, and mean field; this yields stochastic optimal control problems for each agent linked only through the major agent’s state and mean field. Finally the infinite population best response strategies are applied to the finite population system which yields an \( \epsilon \)-Nash equilibrium [9]. The following theorem specifies the control laws which yield the infinite population Nash equilibrium and their relation with the finite population \( \epsilon \)-Nash equilibrium.

**Theorem 5** (\( \epsilon \)-Nash Equilibrium for LQG MFG Systems). [9] Assume that the conditions of [9] for the existence and uniqueness of Nash equilibrium hold, then the system equations (45)-(49) together with the mean field equations (76)-(77) generate a set of control laws \( U_i^{\infty}_{MF} \triangleq \{ u_i^i; i \geq 0 \} \) where \( u_i^i \) is given by
\[ u_0^i,* = -R_0^{-1} \left[ (N_0^T + B_0^T \Pi_0(t)) \left[ (x_0^0)^T, (\bar{x}_0^0)^T \right]^T + B_0^T s_0(t) \right], \tag{51} \]
\[ u_i^i,* = -R_k^{-1} \left[ (N_k^T + B_k^T \Pi_k(t)) \left[ (x_i^0)^T, (\bar{x}_i^0)^T, (\bar{x}_i)^T \right]^T + B_k^T s_k(t) \right], \tag{52} \]
such that

\[ \text{12} \]
(i) the set of infinite population control laws $\mathcal{U}_\infty^{MF} \triangleq \{u^{i,*} ; i \geq 0\}$ yields the infinite population Nash equilibrium.

$$J_i^{\infty}(u^{i,*}, u^{-i,*}) = \inf_{u^i \in \mathcal{U}_i^{\infty}, L} J_i^{\infty}(u^i, u^{-i,*});$$

(ii) All agent systems $\mathcal{A}_i$, $i \in \mathcal{N}_0$, are second order stable.

(iii) the set of control laws $\mathcal{U}_N^{MF} \triangleq \{u^{i,*} ; i \in \mathcal{N}_0\}$, $1 \leq N < \infty$, yields an $\epsilon$-Nash equilibrium for all $\epsilon$, i.e. for all $\epsilon > 0$, there exists $N(\epsilon)$ such that for all $N \geq N(\epsilon)$,

$$J_i^N(u^{i,*}, u^{-i,*}) - \epsilon \leq \inf_{u^i \in \mathcal{U}_i^{N,L}} J_i^N(u^i, u^{-i,*}) \leq J_i^N(u^{i,*}, u^{-i,*}).$$

The proof of Theorem 5 consists of two parts: (i) the set of control laws $\mathcal{U}_\infty^{MF}$ yields the Nash equilibrium for the infinite population system, (ii) when a finite subset of the control laws $\mathcal{U}_N^{MF}$ is applied to the finite population system, all agent systems are second order stable and it yields an $\epsilon$-Nash equilibrium. In this section, a novel convex analysis approach is presented to retrieve the set of best response strategies $\mathcal{U}_\infty^{MF}$ which yields the Nash equilibrium.

### 5.3.1 Mean Field Evolution

We introduce the empirical state average as

$$x^{(N)}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} x^j_k, \quad k \in \mathcal{K},$$

and write $x^{(N)} = [x^{(N_1)}, x^{(N_2)}, ..., x^{(N_K)}]$, where the pointwise in time $L^2$ limit of $x^{(N)}$, if it exists, is called the mean field of the system and is denoted by $\bar{x} = [\bar{x}^1, ..., \bar{x}^K]$. We consider for each minor agent $\mathcal{A}_i$ of type $k$, $k \in \mathcal{K}$, a uniform (with respect to $i$) feedback control $u^i_k \in \mathcal{U}_{i,L} \subset \mathcal{U}_i$, where $\mathcal{U}_{i,L}$ consists of linear time invariant controls, as

$$u^i_k = L^i_1 x^i_k + \sum_{l=1}^{K} \sum_{j=1}^{N_l} L^i_l x^j_l + L^i_3 x^i_0 + m^i(t),$$

where $0 \leq t \leq \infty$, $L^i_1$, $L^i_2$, and $L^i_3$ are constant matrices, and $m^i(t)$ is a continuous bounded function of time. If we substitute $u^i_k$ in (46) for $i \in \mathcal{N}$, and take the
average of the states of closed loop systems of type \( k, \ k \in \mathbb{R}, \) and hence calculate \( x_t^{(N)} \), it can be shown that the \( L^2 \) limit \( \bar{x}_t \) of \( x_t^{(N)} \), i.e. the mean field satisfies

\[
d\bar{x}_t = \bar{A}\bar{x}_t dt + \bar{G}x_t^0 dt + \bar{m}(t) dt,
\]

where \( \bar{A}, \bar{G}, \) and \( \bar{m} \) are to be solved for in the tracking solution. By abuse of language, the mean value of the system’s Gaussian mean field given by the state process \( \bar{x}_t = [\bar{x}_t^1, \ldots, \bar{x}_t^K] \) shall also be termed the system’s mean field.

### 5.3.2 Major Agent: Infinite Population

To solve the infinite population tracking problem for the major agent \( A_0 \), first, its state is extended with the mean field process \( \bar{x}_t \), where this is assumed to exist. Then the dynamics of major agent’s extended state \( X_t^0 \triangleq [x_t^0, \bar{x}_t^{(N)}] \) is given as (see [9])

\[
dX_t^0 = A_0 X_t^0 dt + B_0 u_t^0 dt + M_0(t) dt + \Sigma_0 dW_t^0,
\]

where

\[
A_0 = \begin{bmatrix} A_0 & F_0 \\ G & \bar{A} \end{bmatrix}, \quad B_0 = \begin{bmatrix} B_0 \\ 0 \end{bmatrix}, \quad M_0(t) = \begin{bmatrix} b_0(t) \\ \bar{m}(t) \end{bmatrix}, \quad \Sigma_0 = \begin{bmatrix} \sigma_0 & 0 \\ 0 & 0 \end{bmatrix}, \quad W_t^0 = \begin{bmatrix} w_t^0 \\ 0 \end{bmatrix}.
\]

The infinite population individual cost functional for the major agent is given by

\[
J_\infty^0(u^0) = \frac{1}{2} \mathbb{E} \left( X_T^0 G_0 X_T^0 + \int_0^T \left( (X_s^0)^T Q_0 X_s^0 + 2(X_s^0)^T N_0 u_s^0 + (u_s^0)^T R_0 u_s^0 \\ - 2(X_s^0)^T \bar{n}_0 - 2(u_s^0)^T \bar{n}_0 \right) ds \right),
\]

where the corresponding weight matrices are specified by

\[
G_0 = [I_n, -H^\pi_0]^T G_0 [I_n, -H^\pi_0], \quad Q_0 = [I_n, -H^\pi_0]^T Q_0 [I_n, -H^\pi_0], \\
N_0 = [I_n, -H^\pi_0]^T N_0, \quad \bar{n}_0 = [I_n, -H^\pi_0]^T Q_0 \eta_0, \quad \bar{n}_0 = N_0^T \eta_0.
\]

The dynamics (54) together with the cost functional (56) constitute a stochastic LQ optimal control problem for the major agent \( A_0 \)’s extended system in the infinite population limit. To determine the optimal control \( u_t^{0,*} \), first \textit{Theorem}
\[ (J_0^\infty'(u^0), \omega^0) = \mathbb{E} \left[ \int_0^T (\omega_t^0) \left\{ N_0^T X_t^{0,u} + R_0 u_t^0 - \bar{n}_0 
 + \mathbb{B}_0^T \left( e^{-\bar{\kappa}_0^0 t} M_t^0 - \int_0^t e^{\bar{\kappa}_0^0 (s-t)} (Q_0 X_s^{0,u} + N_0 u_s^0 - \bar{n}_0) ds \right) \right\} dt \right], \tag{58} \]

where
\[ M_t^0 = \mathbb{E} \left[ e^{\bar{\Lambda}_0^0 T} \mathcal{G}_0 X_T^{0,u} + \int_0^T e^{\bar{\kappa}_0^0 s} (Q_0 X_s^{0,u} + N_0 u_s^0 - \bar{n}_0) ds \right] \mathcal{F}_t^0. \tag{59} \]

Then, as per Theorem 3, the optimal control action for the major agent’s extended system (54)-(57) in the infinite population limit is given by
\[ u_t^{0,*} = -R_0^{-1} \left[ N_0^T X_t^{0,*} - \bar{n}_0 + \mathbb{B}_0^T \left( e^{-\bar{\kappa}_0^0 t} M_t^0 - \int_0^t e^{\bar{\kappa}_0^0 (s-t)} (Q_0 X_s^{0,*} + N_0 u_s^{0,*} - \bar{n}_0) ds \right) \right], \tag{60} \]

Finally, using Theorem 2 (60) can be written in the state feedback form as
\[ u_t^{0,*} = -R_0^{-1} \left[ N_0^T X_t^{0} - \bar{n}_0 + \mathbb{B}_0^T (\Pi_0(t) X_t^{0} + s_0(t)) \right], \tag{61} \]

where
\[ \begin{align*}
-\Pi_0 &= \Pi_0 \Lambda_0 + \Lambda_0^T \Pi_0 - (\mathbb{B}_0^T \Pi_0 + N_0^T)^T R_0^{-1} (\mathbb{B}_0^T \Pi_0 + N_0) + Q_0, \\
\Pi_0(T) &= \mathcal{G}_0, \tag{62} \\
-\dot{s}_0(t) &= [(\Lambda_0 - \mathbb{B}_0 R_0^{-1} N_0^T)^T - \Pi_0 \mathbb{B}_0 R_0^{-1} \mathbb{B}_0^T] s_0(t) \\
&\quad + \Pi_0 (\mathcal{M}_0(t) + \mathbb{B}_0 R_0^{-1} \bar{n}_0) + N_0 R_0^{-1} \bar{n}_0 - \bar{n}_0, \tag{63} \\
s_0(T) &= 0.
\end{align*} \]

### 5.3.3 Minor Agent: Infinite Population

To solve the infinite population tracking problem for a minor agent \( A_i, i \in \mathcal{N} \), first, its state is extended with the major agent’s state and the mean field process \( \bar{x}_t \), where this is assumed to exist. Then the dynamics of minor agent \( A_i \)’s extended state \( X_t^i \equiv [ (x_t^i)^T, (x_t^0)^T, (\bar{x}_t)^T ]^T \) is given as (see [9])
\[ dX_t^i = \Lambda_k X_t^i dt + \mathbb{B}_k u_t^i dt + \mathcal{M}_k(t) dt + \Sigma_k dW_t^i, \tag{64} \]
Theorem 3

Then according to computed as

\[
A_k = \begin{bmatrix}
A_k & [H_k, F_k^\pi] \\
0 & A_0 - B_0 R_0^{-1} N_0 - B_0 R_0^{-1} B_0^T \Pi_0
\end{bmatrix}, \quad B_k = \begin{bmatrix}
B_k \\
0
\end{bmatrix},
\]

\[
M_k(t) = \begin{bmatrix}
b_k(t) \\
M_0(t) - B_0 R_0^{-1} B_0^T s_0(t)
\end{bmatrix}, \quad \Sigma_k = \begin{bmatrix}
\sigma_k & 0 \\
0 & \Sigma_0
\end{bmatrix}, \quad W_t = \begin{bmatrix}
w_t^i \\
W_t^0
\end{bmatrix}.
\]

The infinite population individual cost functional for minor agent \(A_i\), \(1 \leq i \leq N\), is given by

\[
J_i^\infty(u^i) = \frac{1}{2} \mathbb{E} \left[ (X_T^i)^T G_k X_T^i + \int_0^T \left\{ (X_s^i)^T Q_k X_s^i + 2(X_s^i)^T N_k u^i_s + (u_s^i)^T R_k u^i_s \\
- 2(X_s^i)^T \eta_k - 2(u_s^i)^T \bar{n}_k \right\} ds \right],
\]

where the corresponding weight matrices are specified by

\[
G_k = [I_n, -H_k, -\dot{H}_k^\pi]^T G_k [I_n, -H_k, -\dot{H}_k^\pi], \quad Q_k = [I_n, -H_k, -\dot{H}_k^\pi]^T Q_k [I_n, -H_k, -\dot{H}_k^\pi],
\]

\[
N_k = [I_n, -H_k, -\dot{H}_k^\pi]^T N_0, \quad \eta_k = [I_n, -H_k, -\dot{H}_k^\pi]^T \eta_k, \quad \bar{n}_k = N_k^T \eta_k.
\]

The dynamics (64) together with the cost functional (66) constitute a stochastic LQ optimal control problem for the infinite population of \(A_i\)’s extended system in the infinite population limit. To determine the optimal control \(u_i^i^*\) for minor agent \(A_i\), \(1 \leq i \leq N\), first, using Theorem 2, the Gâteaux derivative \(J_i^\infty(u^i)\) of (66) is computed as

\[
\langle J_i^\infty(u^i), \omega^i \rangle = \mathbb{E} \left[ \int_0^T (\omega_t^i)^T \left\{ N_k^T X_t^{i,u} + R_k u^i_t - \bar{n}_k \\
+ B_k e^{\dot{A}_k t} M_t^i - \int_0^t e^{\dot{A}_k (s-t)} (Q_k X_s^{i,u} + N_k u_s^i - \bar{n}_k) ds \right\} dt \right],
\]

where

\[
M_t^i = \mathbb{E} \left[ e^{\dot{A}_k T} G_k X_T^{i,u} + \int_0^T e^{\dot{A}_k (T-s)} (Q_k X_s^{i,u} + N_k u_s^i - \bar{n}_k) ds \left| \mathcal{F}_t \right. \right].
\]

Then according to Theorem 3, the optimal control action for minor agent \(A_i\), \(i \in \mathcal{N}\), is given by

\[
u_i^{i^*} = -R_k^{-1} \left[ N_k^T X_t^{i^*} - \bar{n}_k + B_k \left( e^{\dot{A}_k t} M_t^i - \int_0^t e^{\dot{A}_k (s-t)} (Q_k X_s^{i^*} + N_k u_s^{i^*} - \bar{n}_k) ds \right) \right].
\]
Finally, using Theorem 2, the control action (70) can be presented in linear state feedback form as

\[ u_t^i = -R_k^{-1}[N_k^T X_t^i - \bar{n}_k + B_k^T (\Pi_k(t) X_t^i + s_k(t))] \]

where

\[
\begin{aligned}
-\Pi_k &= \Pi_k A_k + A_k^T \Pi_k = - (B_k^T \Pi_k + N_k^T) R_k^{-1} (B_k^T \Pi_k + N_k) + Q_k, \\
\Pi_k(T) &= \mathcal{G}_k, \\
\end{aligned}
\]

(72)

\[
\begin{aligned}
-\dot{s}_k(t) &= [(A_k - B_k R_k^{-1} N_k^T)^T - \Pi_k B_k^T R_k^{-1} B_k^T] s_k(t) \\
&\quad + \Pi_k [M_k(t) + B_k R_k^{-1} \bar{n}_k] + N_k R_k^{-1} \bar{n}_k - \bar{n}_k, \\
s_k(T) &= 0.
\end{aligned}
\]

(73)

5.3.4 **Mean Field Consistency Conditions**

To obtain the consistency conditions, we substitute (71) into (46) which results in

\[
dx_t^i = \left( A_k x_t^i - B_k R_k^{-1} N_k^T [x_t^i, (x_0^i)^T, \bar{x}_t^i]^T - \bar{n}_k + B_k^T (\Pi_k [(x_t^i)^T, (x_0^i)^T, \bar{x}_t^i]^T + s_k) \right) \\
&\quad + H_k x_0^i + F_k^\pi \bar{x}_t + b_k \right) dt + \sigma_k dw_t^i.
\]

(74)

Let define

\[
\Pi_k = \begin{bmatrix}
\Pi_{k,11} & \Pi_{k,12} & \Pi_{k,13} \\
\Pi_{k,21} & \Pi_{k,22} & \Pi_{k,23} \\
\Pi_{k,31} & \Pi_{k,32} & \Pi_{k,33}
\end{bmatrix}, \quad k \in \mathcal{K},
\]

and \( e_k = [0_{n \times n}, \ldots, 0_{n \times n}, I_n, 0_{n \times n}, \ldots, 0_{n \times n}] \), where the \( n \times n \) identity matrix \( I_n \) is at the \( k \)th block.

If we take the average of (74) over subpopulation \( A_k, k \in \mathcal{K} \), and then take the \( L^2 \) limit as the number \( N_k \) of agents within the subpopulation goes to infinity (i.e. \( N_k \to \infty \)), we get

\[
d\bar{x}_t^k = \left( F_k^\pi + [A_k - B_k R_k^{-1} (N_{k,1}^T + B_k^T \Pi_{k,11})] e_k - B_k R_k^{-1} B_k^T \Pi_{k,13} \right) \bar{x}_t dt \\
&\quad + (H_k - B_k R_k^{-1} B_k^T \Pi_{k,12}) x_0^i dt + (b_k + B_k R_k^{-1} \bar{n}_k - B_k R_k^{-1} B_k^T s_k) dt.
\]

(75)
Hence, the major agent’s infinite horizon cost functional is given by
\[ J_\infty^M = \pi_0 \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( \| x_t \|_{Q_0}^2 + 2 (x_t - \Phi(x_t))^T N_0 x_t + \| u_t \|_{R_0}^2 \right) dt \right]. \] (78)

Similarly, the discounted infinite horizon cost functional for minor agent \( A_k \), \( 1 \leq i \leq N \) is given by
\[ J_i^N(u^i, u^{-i}) = \frac{1}{2} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( \| x_t^i - \Psi(x_t^n) \|_{Q_k}^2 + 2 (x_t^i - \Psi(x_t^n))^T N_k x_t^i + \| u_t^i \|_{R_k}^2 \right) dt \right]. \] (79)

The dynamics (45)-(46) for the major agent and minor agents remain the same in the infinite horizon LQG MFG systems.

Assumption 7. The pair \( (L_a, A_0 - (\rho/2)I) \) is detectable, and for each \( k \in \mathbb{R} \), the pair \( (L_b, A_k - (\rho/2)I) \) is detectable, where \( L_a = Q_0^{1/2}[I, -H_0^T] \) and \( L_b = Q_k^{1/2}[I, -H_k, -\hat{H}_k^T] \).
Assumption 8. The pair \((A_0 - (\rho/2) I, B_0)\) is stabilizable and \((A_k - (\rho/2) I, B_k)\) is stabilizable for each \(k \in K\).

Given that Assumptions 7-8 hold, for the major agent’s system (45), (78), the best response strategy is given by (61), where the steady state Riccati matrix \(\Pi_0\) satisfies an algebraic Riccati equation given by

\[
\rho \Pi_0 = \Pi_0 A_0 + A_0^T \Pi_0 - (B_0^T \Pi_0 + N_0^T R_0^{-1} B_0^T \Pi_0 + N_0) + Q_0, \quad (80)
\]

and the steady state offset vector \(s_0\) satisfies the differential equation

\[
\rho s_0(t) = s_0(t) + [(A_0 - B_0 R_0^{-1} N_0^T)^T - \Pi_0 B_0 R_0^{-1} B_0^T] s_0(t) + \Pi_0 (M_0(t) + B_0 R_0^{-1} \bar{n}_0) + N_0 R_0^{-1} \bar{n}_0 - \bar{\eta}_0. \quad (81)
\]

Similarly, for minor agent \(A_i\)’s system (46), (79), \(i \in N\), the best response strategy is given by (71), where the steady state Riccati matrix \(\Pi_k\) and offset matrix \(s_k\) satisfy the following algebraic Riccati equation and differential offset equation.

\[
\begin{align*}
\rho \Pi_k &= \Pi_k A_k + A_k^T \Pi_k - (B_k^T \Pi_k + N_k^T)^T R_k^{-1} (B_k^T \Pi_k + N_k) + Q_k, \quad \forall k; \\
\rho s_k(t) &= s_k(t) + [(A_k - B_k R_k^{-1} N_k^T)^T - \Pi_k B_k R_k^{-1} B_k^T] s_k(t) + \Pi_k (M_k(t) + B_k R_k^{-1} \bar{n}_k) + N_k R_k^{-1} \bar{n}_k - \bar{\eta}_k, \\
&\quad + N_k R_k^{-1} \bar{n}_k - \bar{\eta}_k, \quad s_k(T) = 0, \quad \forall k. \quad (82)
\end{align*}
\]

References

[1] D. E. Kirk, *Optimal control theory: An introduction*. New York: Prentice Hall, Inc., 1971.

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

[3] A. Bensoussan, *Stochastic control of partially observable systems*. Cambridge: Cambridge University Press, 1992.

[4] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*. Society for Industrial and Applied Mathematics, 1999.

[5] G. Allaire, *Numerical analysis and optimization: An introduction to mathematical modeling and numerical simulation*. Oxford University Press, 2007.
[6] R. Carmona, *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*. Society for Industrial and Applied Mathematics, 2016.

[7] P. Bank, H. M. Soner, and M. Voß, “Hedging with temporary price impact,” *Mathematics and Financial Economics*, vol. 11, no. 2, pp. 215–239, 2017.

[8] P. Casgrain and S. Jaimungal, “Algorithmic trading with partial information: A mean field game approach,” *arXiv*, 2018.

[9] M. Huang, “Large-population LQG games involving a major player: The Nash certainty equivalence principle,” *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, 2010.

[10] P. Caines, *Linear stochastic systems*. New York: John Wiley, 1988.

[11] M. Nourian and P. E. Caines, “$\epsilon$-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents,” *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 3302–3331, 2013.