Secure Discrete-Time Linear-Quadratic Mean-Field Games*

Muhammad Aneeq uz Zaman[1], Sujay Bhatt[1], and Tamer Başar[1]

Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana IL 61801-2307, USA

Abstract. In this paper, we propose a framework for strategic interaction among a large population of agents. The agents are linear stochastic control systems having a communication channel between the sensor and the controller for each agent. The strategic interaction is modeled as a Secure Linear-Quadratic Mean-Field Game (SLQ-MFG), within a consensus framework, where the communication channel is noiseless, but, is susceptible to eavesdropping by adversaries. For the purposes of security, the sensor shares only a sketch of the states using a private key. The controller for each agent has the knowledge of the private key, and has fast access to the sketches of states from the sensor. We propose a secure communication mechanism between the sensor and controller, and a state reconstruction procedure using multi-rate sensor output sampling at the controller. We establish that the state reconstruction is noisy, and hence the Mean-Field Equilibrium (MFE) of the SLQ-MFG does not exist in the class of linear controllers. We introduce the notion of an approximate MFE (ϵ-MFE) and prove that the MFE of the standard (non-secure) LQ-MFG is an ϵ-MFE of the SLQ-MFG. Also, we show that $\epsilon \to 0$ as the estimation error in state reconstruction approaches 0. Furthermore, we show that the MFE of LQ-MFG is also an $(\epsilon + \varepsilon)$-Nash equilibrium for the finite population version of the SLQ-MFG; and $(\epsilon + \varepsilon) \to 0$ as the estimation error approaches 0 and the number of agents $n \to \infty$. We empirically investigate the performance sensitivity of the $(\epsilon + \varepsilon)$-Nash equilibrium to perturbations in sampling rate, model parameters, and private keys.

1 Introduction

In this paper, we study large scale multi-agent interaction, where a large number of *sensing systems* (a.k.a agents) interact with each other, to solve a *consensus problem* in a decentralized manner. The individual agents in such a multi-agent system comprise of a sensor and a controller with a communication channel between them. The communication channel is noiseless, but, is susceptible to eavesdropping by adversaries; so a secure communication mechanism is desired. The eavesdropping adversary is not a strategic player, and hence there is no

* Research support in part by Grant FA9550-19-1-0353 from AFOSR, and in part by US Army Research Laboratory (ARL) Cooperative Agreement W911NF-17-2-0196.

game being played against the adversary, as in robust MFGs [1]. The sensors collect and/or generate various sensory data over time and the controllers analyze the data streams to discover new information, derive future insights, and take sequential control decisions.

Mean-Field Games (MFGs) are a framework for analyzing large scale strategic interaction between rational agents optimizing their accumulated returns over time. Estimating the solution of finite population games with a large number of agents is prohibitive in most cases, being exponential in the number of agents [2, 3]. Mean-field games was proposed to address this scalability issue in the seminal works of [4] and [5]. In the mean-field setting, a generic agent interacts with a mass of infinitely many agents, modeled as an exogenous signal, also called the mean-field trajectory. The solution concept analogous to the Nash equilibrium in MFGs is that of the mean-field equilibrium, where the generic agent reacts optimally to the mean-field trajectory and the mean-field trajectory in turn models the *average* behavior of the agent.

Linear Quadratic MFGs (LQ-MFGs) [6–8] are a significant benchmark in the area of MFGs which inherit much of the advantages of the MFG formalism, while allowing analytical expressions for MFE. While there have been several works on the continuous time formulation of LQ-MFGs [6–8], the discrete-time setting has recently gained momentum due to its application to digital systems [9] and reinforcement learning [10–12]. Following this line of work, we adopt a discrete-time LQ framework to study the secure multi-agent interaction.

1.1 Agent Model & Objective

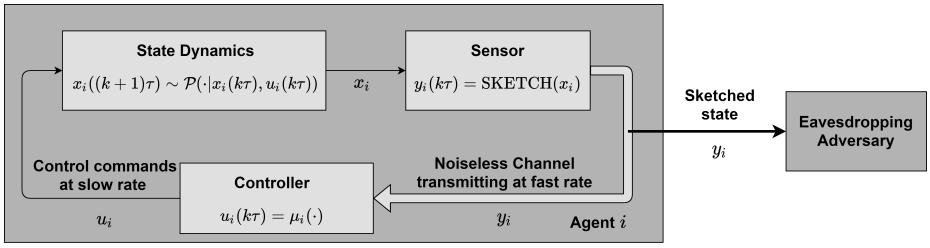


Fig. 1: Schematic representation of an agent in the multi-agent system. The sensor measures the state and provides a multi-rate access to the controller over the noiseless communication channel. The channel is susceptible to eavesdropping by adversaries. Hence, the sensor and controller employ a secure communication mechanism (sketching and multi-rate sampling) for decision making.

Figure 1 shows the schematic representation of an agent in the multi-agent system under consideration. The agents aim to solve a consensus problem in a decentralized manner. Namely, the cost function of each agent penalizes (i)

deviation of the agent's state from the aggregate behavior, and (ii) high control effort. For each agent, the communication link between the sensor and the controller is assumed to be noiseless, but susceptible to eavesdropping by adversaries. The sensor collects information on the underlying high-dimensional stochastic process, modelled as a state, and the controller takes actions to affect the evolution of the stochastic process. However, for the purposes of security and fast communication, the sensor shares only a *sketch* of the high-dimensional states. Sketches obfuscate the original data using a transformation (called the *private key*) and thus have the flexibility to reduce the dimensionality of the shared data while providing privacy features [13, 14]. The controller reconstructs the states, using the private key and multi-rate sensor output sampling, for decision making. We assume that there is a finite set of private keys that the agents in the multi-agent system can choose from. Each agent chooses a key from the set randomly for the purposes of secure sketching and state reconstruction.

In this paper, we model the agents as *linear* control systems perturbed by a noise process, which is not necessarily Gaussian. For ease of implementation, we restrict the controllers of the secure LQ game to the class of linear controllers. There is a communication channel between the sensor and the controller; a model similar to [15]. However, unlike [15], the channel is noiseless but susceptible to eavesdropping by adversaries, and the sensor provides the controller with multi-rate (fast with respect to the rate of control input) access to the linear sketch of the underlying states. The controller reconstructs the state, which is noisy owing to finite-sampling rates, and decides the course of action guided by a decentralized feedback control law. In a multi-agent system made up of agents as in Figure 1, we consider the following problem: *How do rational agents solve a consensus problem in a decentralized manner, when they have security concerns?*

1.2 A Motivating Application

Internet of Battlefield Things (IoBT) is a paradigm that is becoming increasingly important, partly due to the advantages it can offer in battlefield scenarios and largely due to the success of Internet of Things. Entities ("things") are more useful and effective when they are smarter, and even more so when they can interact with each other [16].

Multi-agent systems can model the interaction of entities and their capabilities including controlled sensing and processing of information, undertaking coordinated defensive actions against adversaries, and effecting offensive measures to achieve desired objectives. This is achieved by coordinating, jointly planning and executing the decisions of the agents that constitute the Internet of Battlefield Things. Successful leverage of large scale multi-agent systems to build battlefield solutions requires learning optimal policies in a *secure* and *decentralized* manner; which increases reliability, decreases computational overhead, and facilitates interaction of heterogeneous agents.

1.3 Main Results & Organization

The paper is organized as follows:

1. In Section 2, we propose a Secure Linear Quadratic Mean-Field Game (SLQ-MFG), for studying the multi-agent interaction with possibly infinitely many agents solving a consensus problem. We also derive insights in the secure (finite) $n-$agent dynamic game using this analysis.
2. In Section 3, we discuss a secure communication mechanism for information transfer between the sensor and the controller for each agent; see Figure 1. This (noisy) reconstructed state, obtained using multi-rate sensor output sampling, is used by the controller for decentralized decision making.
3. In Section 4, we establish that Mean-Field Equilibrium (MFE), a notion that formalizes the notion of consensus in multi-agent systems, does not exist in SLQ-MFGs in the class of linear controllers. Hence, we introduce the notion of ϵ-MFE and $(\epsilon+\varepsilon)$-Nash equilibrium to characterize consensus in secure multi-agent interactions. We prove that MFE of (standard) LQ-MFG, in which the controller has perfect state information, corresponds to ϵ-MFE of the SLQ-MFG, and an $(\epsilon+\varepsilon)$-Nash equilibrium for the secure $n-$agent dynamic game.
4. In Section 5 we empirically investigate the performance of $(\epsilon+\varepsilon)$-Nash equilibrium (deduced in Section 4) and its sensitivity to perturbations in sampling rate, model parameters and private keys.

2 Secure LQ-MFG: Model & Objective

In this section, we discuss the model and objective for dynamic games considered in this paper for analyzing the multi-agent interaction. Section 2.1 discusses the model and objective of Secure n-agent LQ game and Section 2.2 discusses the model and objective of Secure LQ-MFG.

2.1 Secure n-agent Linear Quadratic (LQ) game

We formulate the secure n-agent LQ game by defining the state dynamics and objectives of the agents, where the agents are coupled by a consensus-like term in their objectives.

Consider the following secure $n-$agent discrete-time Linear Quadratic (LQ) game [17]. Each agent i, $i \in \{1, 2, \cdots, n\}$ in the game has the following dynamics:

$$\begin{aligned} x_i(k\tau + (j+1)\Delta) &= Ax_i(k\tau + j\Delta) + Bu_i(k\tau) + \tilde{w}_i(k\tau + j\Delta), \\ y_i(k\tau + j\Delta) &= \text{SKETCH}(x_i(k\tau + j\Delta)) = C_i x_i(k\tau + j\Delta), \end{aligned} \tag{1}$$

where $j \in \{0, 1, \ldots, N-1\}$ is the index of the fast time scale and $k \in \{0, 1, 2, \ldots\}$ is the index of the slow time scale. We denote by $x_i \in \mathbb{R}^m$ the state of agent i, by $u_i \in \mathbb{R}^p$ the control action and by $y_i \in \mathbb{R}^q$ the garbled state (a sketch a.k.a. observation) which is revealed to the agent. The initial state $x_i(0)$ has

mean ν_0 and covariance matrix Ξ_0. A and B are state transition and control matrices of appropriate dimensions, and $\tilde{w}_i$ are i.i.d random vectors generated by distribution with mean zero and covariance $\tilde{\Sigma}$, denoted by $\mathcal{D}(0, \tilde{\Sigma})$. Note that since we are restricting our attention to linear controllers we don't require the noise to be necessarily Gaussian. At time $k\tau$ each agent has access to knowledge $I_i(k\tau)$, which corresponds to observations and actions of the agent i and the observations of all other agents upto time $k\tau$. Here SKETCH$(\cdot)$ is any function that obfuscates the data. In this paper, we choose a random linear sketch [13] as a private key to transmit the state from the sensor to the controller. Sketching has been used for dimensionality reduction by using random projections [13] and providing privacy features [14]. In this paper, sketching is employed to obfuscate the data by making it hard for any adversary to guess/reconstruct the states.

Let the set of private keys available for each agent be denoted by:

$$\mathcal{C} = \{C^{(i)} | i \in \{1, 2, \dots, M\}, M < \infty, C^{(i)} \in \mathbb{R}^{q \times m}\}, \tag{2}$$

We assume that the each agent chooses a private key $C^{(i)}$ from $\mathcal{C}$, uniformly at random. In Section 4.4 we discuss how careful construction of $\mathcal{C}$ positively impacts the performance of the proposed control strategies.

Multi-rate setup: It is assumed that the sensor and hence the controller has multi-rate (fast) $1/\Delta$ access to the states, while the system is controlled at a slower rate $1/\tau$, such that $\tau = N\Delta$ for an integer $N > 0$. The motivation for this two time-scaled approach is that the obfuscated state received by the controller is not readily amenable for decision making as the private key C_i is not invertible. So using ideas as in [17], we use multi-rate sensor output sampling approach to reconstruct the state. Note that this approach is different from using filtering [13] which is not applicable in this context as there is no observation noise due to the channel being noiseless. In [17], it is shown that for the controller to be able to reconstruct the state using the observations, N should be greater than the observability index of the system (A, C_i).

The cost of agent i is a consensus like cost which couples the agent with the other agents by penalizing deviation from the aggregate behavior of the other agents while minimizing control effort. This is the standard objective function for the original LQ-MFGs problem [6] and has been used as a framework in many follow-up studies, such as [1]. The cost of agent i in the n-agent game under policy $\pi^i \in \Pi^i$, where Π^i is the set of all policies $\pi^i : I_i \to \mathbb{R}^p$, while other agents are following the set of policies $\pi^{-i} \in \Pi^{-i}, \Pi^{-i} := \{\Pi^j\}_{j \neq i}$ is given by,

$$J_i^n(\pi^i, \pi^{-i}) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \Big\{ \sum_{k=0}^{T-1} ||x_i(k\tau) - \bar{x}_i^n(k\tau)||_Q^2 + ||u_i(k\tau)||_R^2 \Big\}, \tag{3}$$

where

$$\bar{x}_i^n(k\tau) = \frac{1}{n-1} \sum_{j \neq i} x_j(k\tau), \tag{4}$$

and $Q \geq 0$ and $R > 0$ are symmetric matrices of appropriate dimensions and I_i is the information set of agent i. The expectation $\mathbb{E}_\pi$ is with respect to the joint control law $\pi = (\pi^i, \pi^{-i})$. The quantity $\bar{x}_i^n$ captures the aggregate behavior of agents other than i and is also called the *empirical mean-field trajectory*.

2.2 Secure Linear Quadratic Mean-Field Game (SLQ-MFG)

Owing to the difficulty of finding ϵ-Nash equilibria in $n-$agent games, we now introduce a limiting case of infinite population game, called the mean-field game [4], [5]. Specifically, we formulate SLQ-MFG by describing the state dynamics and objective of the generic agent interacting with the mass of infinitely many agents, referred to as the *mean-field trajectory* [6].

The dynamics of a generic agent are

$$\begin{aligned} x(k\tau + (j+1)\Delta) &= Ax(k\tau + j\Delta) + Bu(k\tau) + \tilde{w}(k\tau + j\Delta), \\ y(k\tau + j\Delta) &= Cx(k\tau + j\Delta) \end{aligned} \tag{5}$$

where $x \in \mathbb{R}^m$ denotes the state, $u \in \mathbb{R}^p$ the control action and $y \in \mathbb{R}^q$ the observation of the generic agent. As in the finite population game (Section 2.1), j is the index of the fast time scale and k is the index of the slow time scale. Matrices A and B denote the state transition and control matrices of appropriate dimensions and C belongs to the finite set $\mathcal{C}$ as defined in (2). As the generic agent in the SLQ-MFG has the same structure as the agent shown in Figure 1, it is susceptible to surveillance and hence it is necessary to sketch the state of the agent. As a result, the multi-rate setup is required to reconstruct the sketched state. The initial condition $x(0)$ has mean ν_0 and covariance matrix Ξ_0. The noise process $\tilde{w}$ is generated i.i.d. with distribution $\mathcal{D}(0, \tilde{\Sigma})$. The generic agent's controller μ at time t depends on the observations and actions of the agent upto time $k\tau$ and the mean-field trajectory. The set of all such controllers is denoted by $\mathcal{M}$. The generic agent's cost under controller μ and mean-field trajectory $\bar{x} = (\bar{x}(0), \bar{x}(\tau), \bar{x}(2\tau) \ldots)$ is

$$J(\mu, \bar{x}) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\mu \Big\{ \sum_{k=0}^{T-1} ||x(k\tau) - \bar{x}(k\tau)||_Q^2 + ||u(k\tau)||_R^2 \Big\}, \tag{6}$$

The mean-field trajectory is assumed to belong to the set of bounded sequences ℓ^∞. In the context of MFGs, the appropriate solution concept is that of Mean-Field Equilibrium (MFE), which is the infinite population analog of Nash equilibrium. To define the mean-field equilibirium (MFE) we use an operator $\Lambda : \mathcal{M} \to \ell^\infty$ (as defined in [18]) which maps a controller μ to a mean-field trajectory $\bar{x}$. If $\bar{x} = \Lambda(\mu)$, $\bar{x}$ is also referred to as being *generated* by controller μ. Now we state the definition of MFE.

Definition 1. *The tuple* $(\mu^*, \bar{x}^*) \in \mathcal{M} \times \ell^\infty$ *is an MFE if* $\bar{x}^* = \Lambda(\mu^*)$ *and*

$$J(\mu^*, \bar{x}^*) \leq J(\mu, \bar{x}^*), \quad \forall \mu \in \mathcal{M}$$

MFE is the infinite population analog to the Nash equilibrium where any agent (represented by the generic agent) has no incentive to deviate from the MFE given all other agents are following the same policy.

3 State Reconstruction using Multi-Rate Sensor Output Sampling

Since the multi-rate setup plays a crucial role in state reconstruction, we describe the reconstruction mechanism (inspired by [17]) employed by the controller to reconstruct the state of the agent. The method in [17] is for deterministic dynamic systems, and we extend it to the case where the agent dynamics are stochastic. This reconstruction is shown to reproduce the state of the agent with some estimation error.

In the multi-rate setup, the observation rate is faster than the control rate. We suppress subscripts for clarity in this subsection. As a consequence of (5),

$$x(k\tau + j\Delta) = A^j x(k\tau) + \sum_{i=0}^{j-1} A^i B u(k\tau) + \sum_{i=0}^{j-1} A^i \tilde{w}(k\tau + (j-1-i)\Delta)$$

and since $x((k+1)\tau) = x(k\tau + N\Delta)$ we can deduce the dynamics of the state at the slower input rate $1/\tau$,

$$\begin{aligned} x((k+1)\tau) &= A^N x(k\tau) + \sum_{i=0}^{N-1} A^i B u(k\tau) + \sum_{i=0}^{N-1} A^i \tilde{w}(k\tau + (N-1-i)\Delta) \\ &= A_0 x(k\tau) + B_0 u(k\tau) + w^0(k\tau) \end{aligned} \tag{7}$$

where $A_0 = A^N$, $B_0 = \sum_{i=0}^{N-1} A^i B$ and $w^0(k\tau)$ is an i.i.d. random vectors such that $w^0(k\tau) = \sigma w_{[k]}$ where $\sigma = [A^{N-1}, A^{N-2}, \ldots, I]$ and $w_{[k]} = [\tilde{w}^T(k\tau), \tilde{w}^T(k\tau + \Delta), \ldots, \tilde{w}^T(k\tau + (N-1)\Delta)]^T$. The vectors $w^0(k\tau)$ and $w_{[k]}$ have distributions $w_{[k]} \sim \mathcal{D}(0, I \otimes \tilde{\Sigma})$ and $w^0 \sim \mathcal{D}(0, \Sigma^0)$ where $\Sigma^0 = \sigma(I \otimes \tilde{\Sigma})\sigma^T$. Let us define $y_{[k]}$ such that $y_{[k]} := [y^T((k-1)\tau), y^T((k-1)\tau + \Delta), \ldots, y^T((k-1)\tau + (N-1)\Delta)]^T$. The vector $y_{[k+1]}$ can be written down as,

$$y_{[k+1]} = C_0 x(k\tau) + D_0 u(k\tau) + C_d w_{[k]} \tag{8}$$

where

$$C_0 = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{bmatrix}, D_0 = \begin{bmatrix} 0 \\ CB \\ C(AB+B) \\ \vdots \\ C\sum_{i=0}^{N-2} A^i B \end{bmatrix}, C_d = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ C & 0 & 0 & \ldots & 0 \\ CA & C & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{N-2} & CA^{N-3} & CA^{N-4} & \ldots & 0 \end{bmatrix}$$

Multiplying C_0^T on both sides of (8) we get,

$$C_0^T y_{[k+1]} = C_0^T (C_0 x(k\tau) + D_0 u(k\tau) + C_d w_{[k]})$$

As N is greater than the observability index of the system, $C_0^T C_0$ is invertible. Hence,

$$x(k\tau) = (C_0^T C_0)^{-1} C_0^T \left[y_{[k+1]} - D_0 u(k\tau) - C_d w_{[k]} \right] \tag{9}$$

Using (7) and (9) we get,

$$x((k+1)\tau) = A_0 (C_0^T C_0)^{-1} C_0^T \Big[y_{[k+1]} - D_0 u(k\tau) \Big] + B_0 u(k\tau) + E_C w_{[k]}$$

where $E_C = \sigma - A_0 (C_0^T C_0)^{-1} C_0^T C_d$. Thus the state $x(k\tau)$ can be expressed as,

$$x(k\tau) = L_y y_{[k]} + L_u u((k-1)\tau) + w(k\tau)$$

where $L_y = A_0 (C_0^T C_0)^{-1} C_0^T$, $L_u = B_0 - L_y D_0$ and $w(k\tau) = E_C w_{[k-1]}$ is a zero mean random vector with covariance matrix

$$\Sigma_C = E_C (I \otimes \tilde{\Sigma}) E_C^T. \tag{10}$$

As the controller has access to $y_{[k]}$ and $u((k-1)\tau)$ at time $(k\tau)^-$ it can reconstruct the state as,

$$\hat{x}(k\tau) = L_y y_{[k]} + L_u u((k-1)\tau)$$

Hence the estimation error $\hat{x}(k\tau) - x(k\tau) = -w(k\tau)$ is a zero mean random vector with covariance matrix Σ_C. Note that since Σ_C depends on the key C, it belongs to a finite set. We denote this set by $\mathcal{E}_C$, which has one-to-one correspondence with $\mathcal{C}$ and hence has cardinality M.

4 Equilibria of Secure LQ Games

In this section, we establish the equilibrium notions and derive the (approximately) optimal policies for the agents in the multi-agent system. Section 4.1 shows that the optimal control problem (which is a part of the MFE) is a non-standard optimal control problem. Due to this non-standard problem, the SLQ-MFG does not permit an MFE in the class of linear controllers. Section 4.2 introduces the concept of ϵ-MFE and establishes that the MFE of the LQ-MFG is an ϵ-MFE for the SLQ-MFG. The variable ϵ depends on the estimation error in the state reconstruction. Section 4.3 shows that MFE for LQ-MFG is also an $(\epsilon + \varepsilon)$-Nash equilibrium for the secure n-agent LQ game, where ϵ and ε depend on the estimation error and the number of agents n.

4.1 MFE of the SLQ-MFG

We will show that the MFE of the SLQ-MFG does not exist in the class of linear controllers. As is common in the stochastic LQ setting [18], we restrict the

controllers to the class of linear feedback controllers μ_K. The controller has the form

$$u(k\tau) = K_1\hat{x}(k\tau) + K_2\bar{x}(k\tau)$$

where $\hat{x}(k\tau)$ is the reconstructed state of the agent and $\bar{x}(k\tau)$ is the mean-field trajectory.

The choice of the controller is similar to [18], where the MFE of the LQ-MFG was derived. See in Appendix (Section 7.1) for a brief overview of the results. These controllers generate mean-field trajectories which follow linear dynamics [18], formally if $\bar{x} = \Lambda(\mu_K)$ where $K = [K_1^T, K_2^T]^T \in \mathbb{R}^{p\times 2m}$, then $\bar{x}((k+1)\tau) = F\bar{x}(k\tau)$ where $F = A_0 - B_0(K_1 + K_2)$. So for this paper we focus our attention on linear feedback controllers and mean-field trajectories with linear dynamics. Moreover, with some slight abuse of notation $J(\mu, \bar{x})$ will be referred to by $J(K, F)$ and $\bar{x} = \Lambda(\mu_K) \iff F = \Lambda(K)$, since a linear feedback controller μ_K is completely characterized by K and a linear mean-field trajectory $\bar{x}$ by F.

Let us define an augmented state by $z(k\tau) = [x^T(k\tau), \bar{x}^T(k\tau)]^T$ where $\bar{x}^T(k\tau)$ is the mean-field trajectory. Assuming that the mean-field trajectory has linear dynamics given by state matrix F, using (7) the augmented system can be written down as,

$$z((k+1)\tau) = \bar{A}z(k\tau) + \bar{B}u(k\tau) + \bar{w}(k\tau)$$

where,

$$\bar{A} = \begin{bmatrix} A_0 & 0 \\ 0 & F \end{bmatrix}, \bar{B} = \begin{bmatrix} B_0 \\ 0 \end{bmatrix}, \bar{w}(k\tau) = \begin{bmatrix} w^0(k\tau) \\ 0 \end{bmatrix} \tag{11}$$

The random variable $\bar{w}(k\tau)$ has distribution $\mathcal{D}(0, \bar{\Sigma})$ where $\bar{\Sigma} = \begin{bmatrix} \Sigma^0 & 0 \\ 0 & 0 \end{bmatrix}$. Similarly using (6) the cost function of the generic agent can be expressed as,

$$J(K, F) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_K\Big\{ \sum_{k=0}^{T-1} \|z(k\tau)\|_{\bar{Q}}^2 + \|u(k\tau)\|_R^2 \Big\}, \text{where } \bar{Q} = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} \geq 0 \tag{12}$$

Now we investigate the MFE of the SLQ-MFG (Definition 1). As per the definition, one part of finding the MFE corresponds to finding the K which minimizes cost (12) for a given F.

In what follows we show that such a K does not exist. For stabilizing linear feedback controllers $u(k\tau) = -K\hat{z}(k\tau) = -K(z(k\tau) - \hat{w}(k\tau))$ where $\hat{w}(k\tau) = [w^T(k\tau), 0]^T$, the closed-loop dynamics of a system $z(k\tau)$ are

$$z((k+1)\tau) = \bar{A}z(k\tau) + \bar{B}u(k\tau) + \bar{w}(k\tau) = (\bar{A} - \bar{B}K)z(k\tau) + \bar{B}K\hat{w}(k\tau) + \bar{w}(k\tau) \tag{13}$$

where $\hat{w}(k\tau) \sim \mathcal{D}(0, \hat{\Sigma}_C)$, with

$$\hat{\Sigma}_C = \begin{bmatrix} \Sigma_C & 0 \\ 0 & 0 \end{bmatrix} \tag{14}$$

As Σ_C (as given in equation (10)) belongs to the finite set $\mathcal{E}_C$, $\hat{\Sigma}_C$ also belongs to the finite set $\hat{\mathcal{E}}_C$ which has the same cardinality M and has one-to-one correspondence with $\mathcal{E}_C$. Using the definitions of $\hat{w}(k\tau)$ and $\bar{w}(k\tau)$, the random vector $(BK\hat{w}(k\tau) + \bar{w}(k\tau)) \sim \mathcal{D}(0, \Psi_K)$, where $\Psi_K = \bar{\Sigma} + BK\hat{\Sigma}_C(BK)^T$. The stationary distribution of the closed-loop system (13) is $\mathcal{D}(0, \Sigma_K)$, where Σ_K satisfies

$$\Sigma_K = \Psi_K + (\bar{A} - \bar{B}K)\Sigma_K(\bar{A} - \bar{B}K)^T$$

and is guaranteed to be positive semi-definite and unique for any stabilizing K. Using techniques similar to [19] the cost of controller K is

$$\begin{aligned} J(K,F) &= \mathbb{E}_{z(k\tau)\sim\mathcal{D}(0,\Sigma_K)}[z^T(k\tau)(\bar{Q} + K^T R K)z(k\tau)] + \mathrm{Tr}(K\hat{\Sigma}_C K^T R), \\ &= \mathrm{Tr}((\bar{Q} + K^T R K)\Sigma_K) + \mathrm{Tr}(K\hat{\Sigma}_C K^T R) \\ &= \mathrm{Tr}((\bar{Q} + K^T R K)\mathcal{T}_K^T(\Psi_K)) + \mathrm{Tr}(K\hat{\Sigma}_C K^T R), \\ &= \mathrm{Tr}(\mathcal{T}_K(\bar{Q} + K^T R K)\Psi_K) + \mathrm{Tr}(K\hat{\Sigma}_C K^T R), \\ &= \mathrm{Tr}(P_K\bar{\Sigma}) + \mathrm{Tr}((B^T P_K B + R)K\hat{\Sigma}_C K^T) \end{aligned} \tag{15}$$

where the operators $\mathcal{T}_K(M)$ and $\mathcal{T}_K^T(M)$ are defined as

$$\begin{aligned} \mathcal{T}_K(M) &= \sum_{i=0}^{\infty}((\bar{A} - \bar{B}K)^T)^i M(\bar{A} - \bar{B}K)^i, \\ \mathcal{T}_K^T(M) &= \sum_{i=0}^{\infty}(\bar{A} - \bar{B}K)^i M((\bar{A} - \bar{B}K)^T)^i \end{aligned}$$

and P_K is the solution to the Lyapunov equation,

$$\begin{aligned} P_K &= (\bar{A} - \bar{B}K)^T P_K(\bar{A} - \bar{B}K) + (\bar{Q} + K^T R K) \\ &= \bar{A}^T P_K \bar{A} + \bar{Q} + K^T(\bar{B}^T P_K \bar{B} + R)K - \bar{A}^T P_K \bar{B}K - (\bar{B}K)^T P_K \bar{A} \end{aligned} \tag{16}$$

To find the K which minimizes cost (15), we define the Hamiltonian,

$$H = \mathrm{Tr}(P_K\bar{\Sigma}) + \mathrm{Tr}((B^T P_K B + R)K\hat{\Sigma}_C K^T) + \mathrm{Tr}(GS)$$

where

$$G = -P_K + (\bar{A} - \bar{B}K)^T P_K(\bar{A} - \bar{B}K) + (\bar{Q} + K^T R K) = 0$$

Using the minimum principle,

$$\frac{\partial H}{\partial S} = -P_K + (\bar{A} - \bar{B}K)^T P_K (\bar{A} - \bar{B}K) + (\bar{Q} + K^T R K) = 0 \tag{17}$$

$$\frac{\partial H}{\partial P_K} = \Psi_K - S + (\bar{A} - \bar{B}K) S (\bar{A} - \bar{B}K)^T = 0 \tag{18}$$

$$\begin{aligned}\frac{\partial H}{\partial K} &= \frac{\partial}{\partial K}[\mathrm{Tr}(K^T(\bar{B}^T P_K \bar{B} + R) K \hat{\Sigma}_C) + \mathrm{Tr}(GS)] \\ &= \frac{\partial}{\partial K}[\mathrm{Tr}((-\bar{A}^T P_K \bar{A} + P_K - \bar{Q} + \bar{A}^T P_K \bar{B} K + (\bar{B}K)^T P_K \bar{A})\hat{\Sigma}_C) \\ &\qquad + \mathrm{Tr}(GS)] \text{ (using equation (16))} \\ &= \bar{B}^T P_K \bar{A}(\hat{\Sigma}_C - \Sigma_K) + (\bar{B}^T P_K \bar{B} + R) K \Sigma_K = 0 \end{aligned} \tag{19}$$

Equation (17) recovers the Lyapunov equation (16), and from equation (18) we can deduce that $S \equiv \Sigma_K$ since K is a stabilizing controller. Equation (19) gives the form of the optimal controller (if it exists) for a given F

$$\hat{K}_F = (\bar{B}^T \hat{P} \bar{B} + R)^{-1} \bar{B}^T \hat{P} \bar{A}(I - \hat{\Sigma}_C \Sigma_{\hat{K}_F}^{-1})$$

where $\hat{P}$ is the solution to the Lyapunov equation,

$$\hat{P} = (\bar{A} - \bar{B}\hat{K}_F)^T \hat{P} (\bar{A} - \bar{B}\hat{K}_F) + (\bar{Q} + \hat{K}_F^T R \hat{K}_F) \tag{20}$$

Since

$$\begin{aligned}\Sigma_K &= \mathcal{T}_K^T(\Psi_K) \\ &= \sum_{i=0}^{\infty} \begin{bmatrix} A_0 - B_0 K_1 & -B_0 K_2 \\ 0 & F \end{bmatrix}^i \begin{bmatrix} B_0 K_1 \Sigma_C (B_0 K_1)^T + \Sigma^0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_0 - B_0 K_1 & -B_0 K_2 \\ 0 & F \end{bmatrix}^{T^i}\end{aligned}$$

it is clear that Σ_K will be a block diagonal matrix with the second block as all zeros. Hence Σ_K is guaranteed to have at least a zero eignevalue which results in Σ_K being non-invertible. This means that $\hat{K}_F$ does not exist and as a result the MFE does not exist within the class of linear controllers.

To formulate useful strategies for the SLQ-MFG its useful to recall the MFE for the LQ-MFG. The MFE of the LQ-MFG is defined by the tuple (K^*, F^*) where K^* is the controller and F^* is the matrix which defines the mean-field trajectoy $\bar{x}^*$. These results are obtained from a previous work [18] and the reader can refer to the Appendix (Section 7.1) for details of the MFE of LQ-MFG.

4.2 ϵ-MFE of the SLQ-MFG

We introduce the concept of ϵ-MFE for the SLQ-MFG and the MFE of LQ-MFG is shown to satisfy this definition. We start by formally proposing the ϵ-MFE of the SLQ-MFG. This is followed by the result that the MFE of the LQ-MFG is also the ϵ-MFE of the SLQ-MFG where $\epsilon \to 0$ as estimation error due to state reconstruction approaches 0. Consequently, if the state sketching is performed in a manner that estimation error is minimized (while obfuscating the state) the MFE of LQ-MFG is a close-to-optimal strategy for the SLQ-MFG.

Definition 2. *The tuple* $(K', F') \in \mathbb{R}^{p\times 2m} \times \mathbb{R}^{m\times m}$ *is an* ϵ*-MFE if* $F' = \Lambda(K')$ *and*

$$J(K', F') \leq J(K, F') + \epsilon, \;\; \forall K \in \mathbb{R}^{p\times 2m}, \epsilon > 0$$

This is the analog of the ϵ-Nash equilibrium for the infinite population case. An ϵ-MFE proposes strategies where an agent has at most ϵ incentive to deviate from the ϵ-MFE. Now we note the conditions required for the existence and uniqueness of the MFE, which are a carry-over from previous work [18].

Assumption 1. *With* P *given as the unique positive definite solution to the Discrete-Time Riccati Equation (DARE),*

$$P = A_0^T P A_0 + Q - A_0^T P B_0 (R + B_0^T P B_0)^{-1} B_0^T P A_0$$

and furthermore that $G_P := -(R + B_0^T P B_0)^{-1} B_0^T$ *and* $H_P := A_0^T (I + P B_0 G_P)$*, we have*

$$\|H_P\|_2 + \frac{\|B_0 G_P\|_2 \|Q\|_2}{(1 - \|H_P\|_2)^2} < 1$$

Now we present the result that the MFE of LQ-MFG is also the ϵ-MFE of the SLQ-MFG. Moreover $\epsilon \to 0$ if estimation error approaches 0.

Theorem 1. *Under Assumption 1, the MFE of LQ-MFG* (K^*, F^*) *is also the* ϵ*-MFE of the SLQ-MFG where* $\epsilon = \mathcal{O}(\mathrm{Tr}(\hat{\Sigma}_C))$ *and* $\hat{\Sigma}_C$ *(as given in equation* (14)*) belongs to the finite set* $\hat{\mathcal{E}}_C$*.*

Proof. We start by defining the cost of the tuple (K, F) in the LQ-MFG. From equation (28), we have

$$\tilde{J}(K, F) = \mathrm{Tr}(P_K \bar{\Sigma}),$$

where P_K satisfies the Lyapunov equation (16). Hence, $J(K, F) \geq \tilde{J}(K, F)$ for a given K and F and as a consequence $\inf_K J(K, F) \geq \inf_K \tilde{J}(K, F)$. Trying to characterize ϵ-MFE,

$$\begin{aligned}
&J(K^*, F^*) - \inf_K J(K, F^*) \\
&\quad \leq J(K^*, F^*) - \inf_K \tilde{J}(K, F^*) = J(K^*, F^*) - \tilde{J}(K^*, F^*) \\
&\quad = \mathrm{Tr}(P^* \bar{\Sigma}) + \mathrm{Tr}((B^T P^* B + R) K^* \hat{\Sigma}_C K^{*T}) - \mathrm{Tr}(P^* \bar{\Sigma}) \\
&\quad = \mathrm{Tr}\left(\left(\bar{A}^T P^* \bar{B} (R + \bar{B}^T P^* \bar{B})^{-1} \bar{B}^T P^* \bar{A} \right) \hat{\Sigma}_C \right)
\end{aligned}$$

The first equality in the above equation is due to equation (29), the second one is obtained by using (15), (28), (16) and the third equality is a result of using the definition of K^* (26). Hence,

$$\begin{aligned}
J(K^*, F^*) - \inf_K J(K, F^*) \leq &\mathrm{Tr}\left(\left(\bar{A}^T P^* \bar{B} (R + \bar{B}^T P^* \bar{B})^{-1} \bar{B}^T P^* \bar{A} \right) \hat{\Sigma}_C \right) \\
= &\epsilon = \mathcal{O}(\mathrm{Tr}(\hat{\Sigma}_C))
\end{aligned}$$

Now we prove that F^* is generated by controller K^*. A generic agent using controller K^* will have closed-loop dynamics,

$$x((k+1)\tau) = (A_0 - B_0K_1^*)x(k\tau) - B_0K_2^*\bar{x}^*(k\tau) + B_0K_1^*w(k\tau) + w^0(k\tau),$$

where $\bar{x}^*(k\tau)$ is the mean-field at time t generated by the controller K^*. Aggregating the closed-loop dynamics we get the dynamics of $\bar{x}^*$,

$$\bar{x}^*((k+1)\tau) = (A_0 - B_0(K_1^* + K_2^*))\bar{x}^*(k\tau) = F^*\bar{x}^*(k\tau)$$

Hence we have completed the proof of the second part of the ϵ-MFE definition, that is F^* is generated by controller K^*. □

4.3 $(\epsilon + \varepsilon)$-Nash Equilibrium of the secure n-agent LQ game

We prove that the MFE of the LQ-MFG is also an $(\epsilon + \varepsilon)$-Nash Equilibrium of the secure n-agent LQ game. We refer to the control law generated by MFE of LQ-MFG (F^*, K^*) by μ^*. We now show that for the n-agent game μ^* is an $(\epsilon + \varepsilon)$-Nash equilibrium where $\epsilon \to 0$ if the estimation error goes to zero and $\varepsilon \to 0$ if number of agents $n \to \infty$. As a result, if the state sketching is performed so that the estimation error is minimized (while obfuscating the state) and the number of agents is large enough, then the MFE of the LQ-MFG is close-to-optimal strategy for the secure n-agent LQ game. Before stating the theorem, since we are considering linear controllers, we define the set $\Pi_K^i \subset \Pi^i$, which is the set of controllers π^i that are linear in their arguments.

Theorem 2. *Under Assumption 1, let the cost of secure n-agent LQ game under controller μ^* be $J_i^n(\mu^{*i}, \mu^{*-i})$, then*

$$J_i^n(\mu^{*i}, \mu^{*-i}) - \inf_{\pi^i \in \Pi_K^i} J_i^n(\pi^i, \mu^{*-i}) < \epsilon + \varepsilon$$

where $\epsilon = \mathcal{O}(\hat{\sigma}_{max})$ and $\hat{\sigma}_{max} := \max_{\hat{\Sigma}_C \in \hat{\mathcal{E}}_C} \mathrm{Tr}(\hat{\Sigma}_C)$ and $\varepsilon = \mathcal{O}(\hat{\sigma}_{max}/\sqrt{n-1})$.

Proof. For an agent i,

$$J_i^n(\mu^{*i}, \mu^{*-i}) - \inf_{\pi^i \in \Pi_K^i} J_i^n(\pi^i, \mu^{*-i}) =$$
$$J_i^n(\mu^{*i}, \mu^{*-i}) - J(K^*, F^*) + J(K^*, F^*) - \inf_{\pi^i \in \Pi_K^i} J_i^n(\pi^i, \mu^{*-i}) \quad (21)$$

We start by bounding the first expression on the RHS of (21). Let us define $\bar{x}^*(k\tau)$ as the linear mean-field trajectory defined by the state matrix F^* (also by definition generated by controller K^*) and $\bar{x}_i^{n*}(k\tau)$ as the empirical mean-field trajectory if all agents are following controller μ^*. From [9] we obtain the expression,

$$J_i^n(\mu^{*i}, \mu^{*-i}) - J(K^*, F^*) = \mathcal{O}\left(\sqrt{\limsup_{T\to\infty} \sum_{k=0}^{T-1} \mathbb{E}_\mu\big(\|\bar{x}_i^{n*}(k\tau) - \bar{x}^*(k\tau)\|_2^2\big)/T}\right) \quad (22)$$

To upper bound the expression, we write the dynamics of $\bar{x}_i^{n*}$,

$$\bar{x}_i^{n*}((k+1)\tau) = (A_0 - B_0 K_1^*)\bar{x}_i^{n*}(k\tau) - B_0 K_2 \bar{x}(k\tau) + \bar{w}_i^n(k\tau)$$

where $\bar{w}_i^n(k\tau) = \sum_{j\neq i}(\bar{w}_j(k\tau)+B_0K^*\hat{w}_j(k\tau))/(n-1)$ with distribution $\bar{w}_i^n(k\tau) \sim \mathcal{D}(0, \bar{\Sigma}_i^n)$ where,

$$\bar{\Sigma}_i^n = \sum_{j\neq i} \frac{\bar{\Sigma}}{(n-1)^2} + \sum_{j\neq i} \frac{BK^*\hat{\Sigma}_{C_j}(BK^*)^T}{(n-1)^2} = \frac{\bar{\Sigma}}{n-1} + \sum_{j\neq i} \frac{BK^*\hat{\Sigma}_{C_j}(BK^*)^T}{(n-1)^2}$$

Since $\hat{\Sigma}_{C_j}$ belongs to the finite set $\hat{\mathcal{E}}_C$, we define a constant $\hat{\sigma}_{max} := \max_{\hat{\Sigma}_C \in \hat{\mathcal{E}}_C} \mathrm{Tr}(\hat{\Sigma}_C)$. Using this constant we can bound $\mathrm{Tr}(\bar{\Sigma}_i^n)$

$$\mathrm{Tr}(\bar{\Sigma}_i^n) \leq \frac{\mathrm{Tr}(\bar{\Sigma})}{n-1} + \frac{m_1\hat{\sigma}_{max}}{n-1}$$

where m_1 is a constant. Using the same techniques as [18] the expression on RHS of (22) is $\mathcal{O}(\hat{\sigma}_{max}/\sqrt{n-1})$. Now we bound the second expression on the RHS in (21). Using techniques used in [18] for any $\pi^i \in \Pi_K^i$,

$$J_i^n(\pi^i, \mu^{*-i}) \geq J(\pi^i, \bar{x}^*) + \lim_{T\to\infty} \frac{2}{T} \sum_{k=0}^{T-1} \mathbb{E}[(x_i(k\tau) - \bar{x}^*(k\tau))^T C_Z(\bar{x}^*(k\tau) - \bar{x}_i^{n*}(k\tau))]$$

$$\geq J(K^*, F^*) - \epsilon_i + \lim_{T\to\infty} \frac{2}{T} \sum_{k=0}^{T-1} \mathbb{E}[(x_i(k\tau) - \bar{x}^*(k\tau))^T C_Z(\bar{x}^*(k\tau) - \bar{x}_i^{n*}(k\tau))]$$

$$\geq J(K^*, F^*) - \epsilon + \lim_{T\to\infty} \frac{2}{T} \sum_{k=0}^{T-1} \mathbb{E}[(x_i(k\tau) - \bar{x}^*(k\tau))^T C_Z(\bar{x}^*(k\tau) - \bar{x}_i^{n*}(k\tau))]$$

where ϵ_i is obtained from Theorem 1 and $\epsilon_i = \mathcal{O}(\mathrm{Tr}(\hat{\Sigma}_{C_i}))$, and hence $\max_i \epsilon_i =: \epsilon = \mathcal{O}(\hat{\sigma}_{max})$. Moreover using techniques similar to [18]

$$\lim_{T\to\infty} \frac{2}{T} \sum_{k=0}^{T-1} \mathbb{E}[(x_i(k\tau) - \bar{x}^*(k\tau))^T C_Z(\bar{x}^*(k\tau) - \bar{x}_i^{n*}(k\tau))] = \mathcal{O}(\hat{\sigma}_{max}/\sqrt{n-1}) \tag{23}$$

Using (22)-(23) we get

$$J_i^n(\mu^{*i}, \mu^{*-i}) - \inf_{\pi^i \in \Pi_K^i} J_i^n(\pi^i, \mu^{*-i}) < \epsilon + \varepsilon$$

where $\epsilon = \mathcal{O}(\hat{\sigma}_{max})$ and $\varepsilon = \mathcal{O}(\hat{\sigma}_{max}/\sqrt{n-1})$. □

4.4 Summary & Discussion

The results in Section 4 provide decentralized feedback control laws for the agents in the multi-agent system, and provide conditions under which they are

approximately optimal for the Secure LQ games. For ease of implementation, we restrict the controllers of the secure LQ game to the class of linear controllers. The approximation is characterized in terms of the estimation error and the number of agents.

In Section 4.1, we have shown that for the class of linear controllers, the MFE does not exist for the SLQ-MFG, as the estimation error in state reconstruction leads to a non-standard optimal control problem. This problem is overcome by proposing the idea of an ϵ-MFE (approximate MFE) in Section 4.2. Then the MFE of the (standard) LQ-MFG is shown to be an ϵ-MFE of the SLQ-MFG.

It is shown that, $\epsilon = \mathcal{O}(\mathrm{Tr}(\hat{\Sigma}_C))$. As $\hat{\Sigma}_C$ is dependent on the private key C chosen uniformly from set $\mathcal{C}$, by careful choice of $\mathcal{C}$ in (2) the estimation error can be minimized (while obfuscating the state) resulting in close-to-optimal strategies for the agent.

In Section 4.3, the MFE of LQG has been shown to be $(\epsilon + \varepsilon)$-Nash Equilibrium of the secure n-agent LQ game. Moreover, $\epsilon = \mathcal{O}(\hat{\sigma}_{max})$ where $\hat{\sigma}_{max} := \max_{\hat{\Sigma}_j \in \hat{\mathcal{E}}_C} \mathrm{Tr}(\hat{\Sigma}_j)$ and $\varepsilon = \mathcal{O}(\hat{\sigma}_{max}/\sqrt{n-1})$. Similar to Section 4.2, through a careful choice of set $\mathcal{C}$ in (2), ϵ can be minimized. It can also be seen that if the number of agents n is large enough, ε is also small. Hence, MFE of LQ-MFG will be close-to-optimal for the secure n-agent LQ game.

In the IoBT setting (Section 1.2) for example, owing to the scale of the multi-agent systems, the computation of optimal (Nash) decentralized feedback laws for the agents is prohibitive. Results in Section 4 provide a way to design decentralized feedback control laws by first deriving the results for the infinite agent case, which is computationally tractable, and then establishing that the same control laws perform well (are approximately optimal) for the considered large scale finite agent setting.

5 Empirical Studies

In this section, we empirically investigate the performance and sensitivity of the $(\epsilon+\varepsilon)$-Nash policies (Section 4) to perturbations in the parameters: (i) Sampling rate (N), (ii) Model parameters (A, B), and (iii) Private keys ($\mathcal{C}$).

We use the *average accumulated cost* [20] as a metric to measure the performance of the $(\epsilon + \varepsilon)$-Nash policies. The average accumulated cost, $J^{n,T}$ is obtained by first simulating the secure n-agent LQ game, under the $(\epsilon+\varepsilon)$-Nash policies. The average accumulated cost is then defined as [20]:

$$J^{n,T} = \frac{1}{T}\sum_{i=1}^{n}\frac{1}{n}\Big\{\sum_{k=0}^{T-1}||x_i(k\tau) - \bar{x}_i^n(k\tau)||_Q^2 + ||u_i(k\tau)||_R^2\Big\},$$

where x_i and u_i are the state and control trajectories of agent i and $\bar{x}_i^n$ is the empirical mean-field trajectory defined in equation (4).

The cost $J^{n,T}$ is an empirical approximation of the cost per agent J_i^n (Section 2), for T sufficiently large. Hence (for high enough T) a low value of $J^{n,T}$ implies a low value of J_i^n and hence indicates good performance by the $(\epsilon + \varepsilon)$-Nash policies.

5.1 Performance sensitivity w.r.t. sampling rate

First we explore the effect of increasing the sampling rate N on the cost $J^{n,T}$. As shown in (15), the effect of N on the cost per agent (through the covariance matrix $\bar{\Sigma}_C$) is quite involved and hence hard to analyze in closed form. Due to this reason, we examine this effect using empirical studies.

We simulate the behavior of $n = 500$ agents, where each agent follows linear dynamics with states $x_i \in \mathbb{R}^{10}$, control actions $u_i \in \mathbb{R}^4$ and sketched state $y_i \in \mathbb{R}^2$, using a fixed set of private keys $\mathcal{C}$ (generated randomly with $m = 4$) and the $(\epsilon+\varepsilon)$-Nash policies stated in Section 4.3. Figure 2 presents the effect of sampling rate N on the average accumulated cost $J^{n,T}$ for $T = \{400, 425, 450, 475, 500\}$. The values of $J^{n,T}$ reach steady-state for $T > 500$.

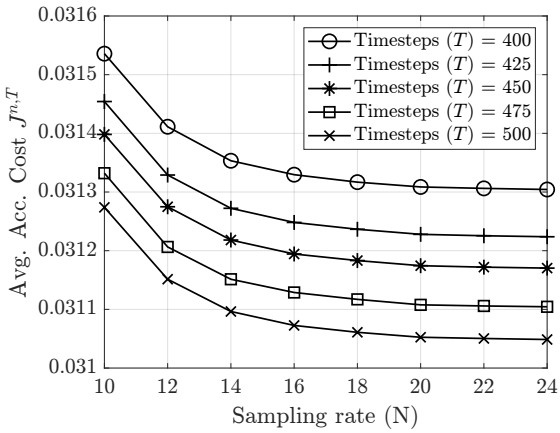


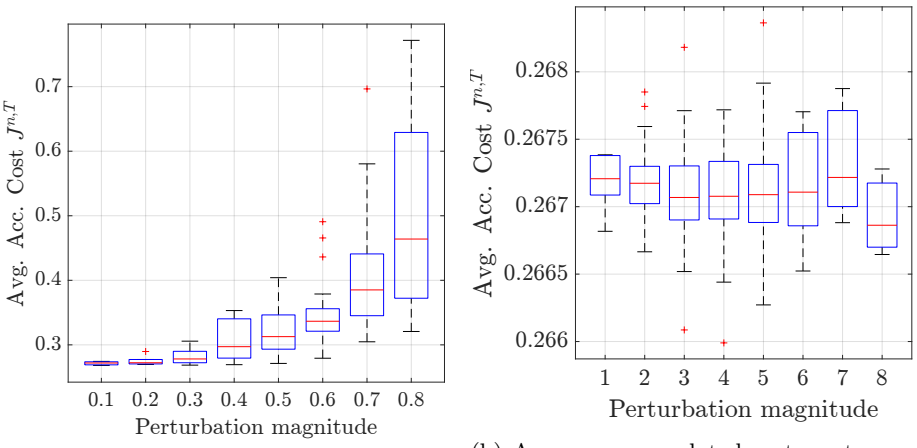
Fig. 2: Average accumulated cost w.r.t. change in sampling rate N.

Observations: Figure 2 shows that for a fixed sampling rate N, the cost $J^{n,T}$ decreases to a steady state value, as T increases. This decrease is due to the stabilizing nature of the $(\epsilon + \varepsilon)$-Nash policies. Furthermore, from Figure 2 we also observe that for a fixed T and increasing N, the cost $J^{n,T}$ decreases to a steady state value. This suggests that higher sampling rates, N, lead to better performance by the $(\epsilon + \varepsilon)$-Nash policies, but high sampling rates may not be achievable due to limited bandwidth available at the channel (see Figure 1). This indicates a trade-off between reduction in cost and limitations of the channel, which calls for a judicious choice.

5.2 Performance sensitivity w.r.t. model parameters and private keys

Using the same setup as before, we next investigate the average accumulated cost, under perturbations in model parameters (A, B) (Figure 3a) and set of

private keys $\mathcal{C}$ (Figure 3b). In the simulation, the perturbed parameters are obtained by adding randomly generated matrices to A, B and $C^{(i)}$ for $i \in [M]$. Figure 3a shows the boxplot[1] for perturbation of model parameters (A, B) and Figure 3b for perturbation of set of private keys $\mathcal{C}$. The boxes in these figures are ordered by increasing perturbation magnitude, where perturbation magnitude is defined as the Frobenius norm of the perturbation. The x-axes of Figures 3a and 3b show the perturbation magnitudes, for their respective boxplots.



(a) Average accumulated cost w.r.t. perturbation of the A and B matrices.

(b) Average accumulated cost w.r.t. perturbation of the set of keys $\mathcal{C}$.

Fig. 3: Distribution of average accumulated cost $J^{n,T}$ w.r.t. magnitude of perturbation of model parameters (A, B) and set of keys $\mathcal{C}$, with $T = 500$.

Observations: Figure 3a shows that as the perturbation in the model parameters (A, B) increases, the median and variance of the cost $J^{n,T}$ increases as well, signifying a decrease in performance. The increase in median cost $J^{n,T}$ is quite intuitive, as the $(\epsilon + \varepsilon)$-Nash policies have been generated for the model parameters (A, B) and any perturbation of these parameters should cause a loss in performance. The increase in variance of cost $J^{n,T}$ is due to random nature of the perturbation, higher perturbation magnitudes result in a bigger spread of the perturbations, resulting in a bigger spread of the average accumulated cost $J^{n,T}$. On the other hand, the median and variance of cost $J^{n,T}$ is shown to be quite insensitive to small perturbations of the set of keys $\mathcal{C}$ (Figure 3b). This result is quite interesting and opens a path for future studies to investigate online methods to maintain security in the face of adapting adversaries.

[1] The red line in the boxplot represents the median, the box represents the 1st and 3rd quartile and the whiskers represent the max and min values, with outliers shown as red crosses.

In this section, we have empirically investigated the performance of the $(\epsilon+\varepsilon)$-Nash policies and inferred the following. There is a clear trade-off when choosing the sampling rate N, as the performance improves for higher N but it might be bounded from above due to channel limitations. Furthermore, we have discovered that the performance of $(\epsilon + \varepsilon)$-Nash policies deteriorates with perturbations in model parameters (A, B) but is insensitive to perturbations in the set of private keys $\mathcal{C}$.

6 Conclusion

In this paper, we have proposed the framework of Secure Linear Quadratic Mean-Field Games (SLQ-MFGs) to analyze multi-agent interactions between agents solving a consensus problem. Each agent has a sensor and a controller, with communication carried out over a noiseless channel, which however is susceptible to eavesdropping. The agents are coupled through their objective functions as in consensus problems. We have proposed a multi-rate sensor output sampling mechanism for the controller to reconstruct the state, albeit with some estimation error. We showed that this estimation error results in non-existence of the Mean-Field Equilibrium (MFE) of the SLQ-MFG for the class of linear controllers, and hence introduced the notions of ϵ-MFE and $(\epsilon + \varepsilon)$-Nash equilibria to characterize consensus in secure multi-agent interactions. Moreover, we have established that MFE of (standard) LQ-MFG, in which the controller has perfect state information, corresponds to ϵ-MFE of the SLQ-MFG, and an $(\epsilon+\varepsilon)$-Nash equilibrium for the secure $n-$agent dynamic game. Finally, we have empirically demonstrated that the performance of the $(\epsilon + \varepsilon)$-Nash equilibrium improves with increasing sampling rate N, deteriorates with variations in model parameters (A, B), and is insensitive to small perturbations in the set of private keys $\mathcal{C}$.

A number of extensions are being considered for future work: (i) A secure and robust n-agent LQ game where the adversary is strategic and can inject malicious signals into the communication channels of the agents to manipulate them into desired behavior, (ii) Design of the set of private keys $\mathcal{C}$ such that the estimation error is minimized while ensuring the obfuscation of the state from the adversary,(iii) Optimal state reconstruction (e.g. MMSE) strategy for the multi-rate setup with noise, (iv) Learning in secure n-agent LQ games where the agents have incomplete knowledge of its dynamic system and/or cost function.

References

1. J. Moon and T. Başar, "Linear quadratic risk-sensitive and robust mean field games," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1062–1077, 2016.
2. R. Breban, R. Vardavas, and S. Blower, "Mean-field analysis of an inductive reasoning game: Application to influenza vaccination," *Physical Review E*, vol. 76, no. 3, p. 031127, 2007.

3. R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, "Electrical vehicles in the smart grid: A mean field game analysis," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1086–1096, 2012.
4. M. Huang, R. P. Malhamé, P. E. Caines *et al.*, "Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.
5. J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
6. M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.
7. A. Bensoussan, K. Sung, S. C. P. Yam, and S.-P. Yung, "Linear-quadratic mean field games," *Journal of Optimization Theory and Applications*, vol. 169, no. 2, pp. 496–529, 2016.
8. M. Huang and M. Zhou, "Linear quadratic mean field games–part I: The asymptotic solvability problem," *arXiv preprint arXiv:1811.00522*, 2018.
9. J. Moon and T. Başar, "Discrete-time LQG mean field games with unreliable communication," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 2697–2702.
10. X. Guo, A. Hu, R. Xu, and J. Zhang, "Learning mean-field games," in *Advances in Neural Information Processing Systems*, 2019.
11. Z. Fu, Z. Yang, Y. Chen, and Z. Wang, "Actor-critic provably finds Nash equilibria of linear-quadratic mean-field games," in *International Conference on Learning Representation*, 2020.
12. R. Elie, J. Pérolat, M. Laurière, M. Geist, and O. Pietquin, "Approximate fictitious play for mean field games," *arXiv preprint arXiv:1907.02633*, 2019.
13. D. Berberidis and G. B. Giannakis, "Data sketching for large-scale Kalman filtering," *IEEE Transactions on Signal Processing*, vol. 65, no. 14, pp. 3688–3701, 2017.
14. J. Blocki, A. Blum, A. Datta, and O. Sheffet, "The Johnson-Lindenstrauss transform itself preserves differential privacy," in *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*. IEEE, 2012, pp. 410–419.
15. S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1549–1561, 2004.
16. A. Kott, A. Swami, and B. J. West, "The internet of battle things," *Computer*, vol. 49, no. 12, pp. 70–75, 2016.
17. S. Janardhanan and B. Bandyopadhyay, "Output feedback sliding-mode control for uncertain systems using fast output sampling technique," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 5, pp. 1677–1682, 2006.
18. M. Zaman, K. Zhang, E. Miehling, and T. Başar, "Reinforcement learning in non-stationary discrete-time linear-quadratic mean-field games," *in 59th IEEE Conference on Decision and Control,*, 2020 (to appear).
19. Z. Yang, Y. Chen, M. Hong, and Z. Wang, "Provably global convergence of actor-critic: A case for linear quadratic regulator with ergodic cost," in *Advances in Neural Information Processing Systems*, 2019, pp. 8351–8363.
20. M. Zaman, K. Zhang, E. Miehling, and T. Başar, "Approximate equilibrium computation for discrete-time linear-quadratic mean-field games," *arXiv preprint arXiv:2003.13195*, 2020.

7 Appendix

In this section, we provide some necessary background material for completeness.

7.1 MFE of the LQ-MFG

Here we briefly discuss the MFE of the LQ-MFG which has been developed in a previous work [18]. We note from [18] that the dynamics of the generic agent in the LQ-MFG are given by,

$$x((k+1)\tau) = A_0 x(k\tau) + B_0 u(k\tau) + w^0(k\tau) \tag{24}$$

Although w^0 is assumed to be non-Gaussian (Section 2.2) the results of [18] (which assume Gaussian distribution) still hold, since we restrict our attention to the class of linear controllers. In the standard LQ-MFG, the multi-rate setup is not required since the controller has access to the true state of the agent. The generic agent aims to minimize the cost function,

$$\tilde{J}(\mu, \bar{x}) = \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}_\mu \Big\{ \sum_{k=0}^{T-1} ||x(k\tau) - \bar{x}(k\tau)||_Q^2 + ||u(k\tau)||_R^2 \Big\}, \tag{25}$$

where $\bar{x}$ is the mean-field trajectory. Next we restate the existence and uniqueness guarantees of MFE for the LQ-MFG.

Proposition 1 ([18]). *Under Assumption 1 the LQ-MFG ((24)-(25)) admits the unique MFE given by the tuple* $(K^*, F^*) \in \mathbb{R}^{p\times 2m} \times \mathbb{R}^{m\times m}$. *The matrix* $F^* = \Lambda(K^*) = A_0 - B_0(K_1^* + K_2^*)$, *and controller* K^* *is defined as,*

$$K^* = (\bar{B}^T P^* \bar{B} + R)^{-1} \bar{B}^T P^* \bar{A}^*, \text{ where } \bar{A}^* = \begin{bmatrix} A_0 & 0 \\ 0 & F^* \end{bmatrix} \tag{26}$$

and P^* *is the solution to the DARE,*

$$P^* = {\bar{A}^*}^T P^* \bar{A}^* + \bar{Q} - {\bar{A}^*}^T P^* \bar{B}(R + \bar{B}^T P^* \bar{B})^{-1} \bar{B}^T P^* \bar{A}^* \tag{27}$$

and $\bar{B}$ *and* $\bar{Q}$ *as defined in* (11) *and* (12), *respectively.*

The DARE is obtained by substituting K^* in the Lyapunov equation (16) hence $P^* = P_{K^*}$. An important point to note is that in the LQ-MFG the estimation error is 0, as the controller has perfect access to the state of the agent. This translates to the covariance matrix of estimation error $\Sigma_C = 0$ and hence $\hat{\Sigma}_C = 0$ for LQ-MFG. Using (15) the cost of linear controller K and linear trajectory defined by matrix F for the LQ-MFG will be

$$\tilde{J}(K, F) = \mathrm{Tr}(P_K \bar{\Sigma}) \tag{28}$$

where P_K is the solution to the Lyapunov equation (15). Furthermore it can also be verified that

$$K^* = \operatorname*{argmin}_K \tilde{J}(K, F^*) \tag{29}$$

This MFE (K^*, F^*) can be obtained by using the mean-field update operator as discussed in [18].