2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Can Mean Field Game Equilibria Amongst Exchangeable Agents Survive Under Partial Observability of Their Competitors' States? *

Farid Rajabali, Roland Malhamé

Abstract— **Classical mean field games (MFG) have been concerned with large games amongst symmetrically influential agents with asymptotically negligible weight. In the absence of a common driving noise, propagation of chaos occurs. The analysis assumes that the initial agent's state probability distribution is known, making its future deterministic and computable via a fixed-point calculation under a limiting equilibrium policy, if it exists. However, oftentimes, despite equal mutual influence, a given agent can only observe a limited number of neighboring agents due to the agent observability structure characterized by an information access graph. This graph may have a low degree even with a large number of agents. The main question addressed is whether an MFG equilibrium can still potentially emerge asymptotically over time. The answer is affirmative, contingent on specific conditions that rely on the stability properties of agents' dynamics and the relative speed of communication to reactions, as derived in this study. The focus is on independent linear scalar agents correlated through a quadratic cost related to the mean state of the agents, which remains unobservable. To tackle convergence to a mean field equilibrium, the proposed model involves a fast communication time scale using a consensus algorithm, alongside a slower agent dynamic time scale. The research explores agents' ability to accurately estimate the system mean as both time and agent numbers increase.**

I. Introduction

Mean Field Game (MFG) problems gained a lot of interest over the past decade, leading to the development of various methods for solving different setups of MFG problems. A substantial body of literature on MFGs has emerged, building upon foundational works by Lasry, Lions [1, 2], Caines, Huang, and Malhamé [3, 4] who approached the analysis from a PDE perspective, and Carmona, and Delarue [5, 6] who adopted a probabilistic viewpoint. In this context, we will focus on papers directly relevant to the specific research issues we aim to address. MFG problems involve non-cooperative agents trying to minimize their cost functions, leading to a system of coupled Hamilton-Jacobi-Bellman equations. In large stochastic games with diminishing individual influence, agents become stochastically independent. In such cases, their joint probability distribution follows a Fokker-Planck-Kolmogorov (FPK) equation. In MFG problems, limiting equilibria are described by a system of coupled forward-backward partial differential equations [5].

A. Literature Review

We review MFG literature based on their application to our problem. At the outset, studies delve into the core MFG problem, introducing key concepts such as ε-Nash equilibrium and Nash certainty equivalence (NCE). Two main approaches are used to find Nash Equilibrium: the bottom-up method, which directly solves the finite game and derives limiting equations as the population size approaches infinity, and the top-down method, which involves solving an optimal control problem using a representative agent while assuming mass behavior for other agents. An equilibrium is reached when all agents follow their best response policy, and the FPK equation replicates the assumed mass behavior, forming the basis for NCE [7]. Other pertinent studies closely aligned with this research are stochastic games on graphs where agents' mutual influences are mediated by a weighted graph [8, 9]. Paper [9] is particularly instructive in that it provides a closed-form solution for the Nash equilibrium in a class of linear quadratic games where agents attempt to follow, at least cost, a weighted combination of the states of their direct neighbors in a so-called transitive graph (essentially a graph which looks "similar" as seen by any agent). Note that this is a radical departure from the classical MFG formulation where agents within possibly distinct classes are exchangeable. An important feature of this paper is that the Nash equilibrium solution is computed under the assumption that agents can observe all other agent states at all times. *It is precisely this somewhat unrealistic assumption that we wish to do without in our proposed research.* Finally, papers that consider non-cooperative aggregative games on networks. Aggregative games are static games where agent costs depend on both their actions and an aggregate measure of all the other agent actions, typically their mean [10]. In a series of papers [10-16], graph-based information exchanges by agents in the form of consensus algorithms were assumed with the objective of helping achieve distributed computation of their Nash equilibria. While, unlike MFGs, the games in [17-23] are static, they relate from the modeling point of view to the question that we are attempting to explore here.

B. Contribution of this Research

In this preliminary work, we wish to explore scalar linear quadratic (LQ) MFGs where agents try to track system mean, while able to exchange information only with a limited number of agents over a transitive so-called information access graph (transitivity will preserve equivalent views of the graph as seen by arbitrary agents). The main question we are trying to address is the following: Can a mean field effect (agents guided by the population mean) still take hold in large populations of non-cooperative agents even though agents are

*Research supported by NSERC-RGPIN-2022-0540.

F. Rajabali and R. Malhamé are with the Department of Electrical Engineering at Polytechnique Montréal and GERAD, Canada. Emails: farid.rajabali@polymtl.ca and roland.malhame@polymtl.ca

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

not able to always observe the statistics of the complete population?

Motivated by what we believe to be a mean field effect taking hold in fish schools which display a high degree of nimbleness in reforming purposeful groups following disturbances, we contribute here: (i) a modeling framework where we consider a separation of time scales between "communication dynamics" over an information access graph, (it characterizes the agents whose states are observable at all times by a given agent, and with whom that agent can exchange information) and "motion dynamics" of the agents; (ii) an algorithm for estimating the current global state which relies on a consensus-based approach to improve the initial mean system state estimate, followed by a forecast. The forecast assumes that agents will use a "certainty equivalent" control law structure, i.e., based on their best response under the assumption of full state observation (iii) an analysis of the dependence of the bias and variance of the mean estimate based on the graph structure, the assumptions on the random nature of the initial state distribution, and the time at which the mean is estimated; (iv) how these quantities evolve as the number of agents increases to infinity. Note that under these assumptions, we assimilate a mean field effect taking place with the ability of agents to estimate at some point in time the system mean with reasonable accuracy.

Remark: This research focuses on a non-cooperative game where individual agents aim to minimize their own cost functions. Despite being non-cooperative, the agents are assumed to share specific information. This sharing of information can be understood either as a scenario where agents seek decentralized control strategies while desiring cooperation or as agents recognizing the mutual benefit of exchanging states, as seen in apps like Google Maps or Waze.

Notation: In this paper, $E[x]$ stands for the expected value of a random variable x. The $N \times 1$ column vector of all ones is denoted $\mathbf{1}$.Vector e_i is $N \times 1$ vector whose i^{th} element is 1 ($e_i = [0 \ \cdots \ 1 \ \cdots \ 0]^T$). Also, small letters are used for scalar variables and capital letters for vectors or matrices. Set $\mathbb{R}$ denotes real numbers.

II. BACKGROUND

In our analysis of MFG with information access graphs, we shall follow the bottom-up approach, i.e., starting from finite agent population games and moving towards infinite population games. Thus, in this section, under full state observation assumptions (or equivalently a full information access graph) and based on [7, 17], we first summarize useful results on the existence and Nash equilibrium (NE) control policy in linear quadratic games.

A. NE Policies in Finite Population, Finite Horizon, Scalar LQ Games with Full State Observations [7]

Consider a non-cooperative game in a population of N agents that are uniform and have scalar dynamics. The dynamics equation for agent i is written in (1) which is a linear and stochastic differential equation.

$$dx_i(t) = \big(ax_i(t) + bu_i(t)\big)dt + \sigma dw_i(t), \quad t \geq 0 \tag{1}$$

In (1), $x_i(t)$ is the state of agent i and $u_i(t)$ is the control input or action of agent i. Coefficients a, b are in $\mathbb{R}$ and σ is nonnegative finite value. Noises $w_i(t)$, $i = 1,2,\dots,N$ are scalar mutually independent zero mean Wiener processes and independent from initial states. The agents' initial conditions are assumed to be random with finite variance.

Agents wish to track $\phi\big(\bar{x}^N(t)\big)$ which is taken to be an affine function of the empirical mean with the cost given by:

$$J_i(u_i, x_i, \bar{x}) = E\left[\int_0^T \left[q\left(x_i(t) - \phi\big(\bar{x}^N(t)\big)\right)^2 + ru_i^2(t)\right]dt + h\left(x_i(T) - \phi\big(\bar{x}^N(T)\big)\right)^2 |x_i(0)\right] \tag{2}$$

$$\phi(\bar{x}^N) = \Gamma\bar{x}^N + \eta, \;\; \bar{x}^N = \frac{1}{N}\sum_{j=1}^{N} x_j \tag{3}$$

Coefficients q and h are non-negative real numbers, and r is a positive real number.
By solving a system of N coupled Hamilton-Jacobi-Bellman equations, it is possible to show (see [7, 17]) that a NE policy for an arbitrary agent i can be written as:

$$u_i(t) = -\frac{b}{r}\big(p(t)x_i(t) + \alpha(t)\bar{x}^N(t) + \beta(t)\big) \tag{4}$$

$$\frac{dp(t)}{dt} = \frac{b^2}{r}p^2(t) - 2ap(t) - q \tag{5}$$
$$p(T) = h$$

$$\frac{d\alpha(t)}{dt} = -2\left(a - \frac{b^2}{r}p(t)\right)\alpha(t) + \frac{b^2}{r}\alpha^2(t) + q\Gamma \tag{6}$$
$$\alpha(T) = -h\Gamma$$

In (5), $p(t)$ is the solution of a standard Riccati differential equation guaranteed to exist under our assumptions, while, given $p(t), \alpha(t)$ in (6) corresponds to the solution of another Riccati differential equation which in general is not guaranteed to exist. However, we will show later that if $\Gamma < 1$, the solution to $\alpha(t)$ exists, and discuss motivations for $\Gamma < 1$.
Remark: For simplicity and without loss of generality, we assume $\eta = 0$, so $\beta(t)$ will be zero over time [7].
Remark: In this paper, we consider $\Gamma \leq 1$, since if $\Gamma > 1$ finite escape time happens and one should be cautious in the time interval of the game to avoid finite escape times [7].

Remark: An application for $\Gamma < 1$ is decentralized power control of cellular phones within the same cell. The latter compete to enhance their signal to noise ratio. Increasing power levels generates more noise on other phones within the cell, compelling those phones to boost their power as well. This scenario can be viewed as a static game whose dynamic version has been the inspiration behind [18]. A motivation for $\Gamma = 1$ is an attempt to reproduce collective dynamics within fish schools, with individual fishes aiming to follow their school with least effort.

For a known (considered deterministic) initial empirical mean $\bar{x}^N(0)$, and if the solution of (6) exists, $\bar{x}^N(t)$ evolves according to:

$$d\bar{x}^N = \left(a - \frac{b^2}{r}\big(\alpha(t) + p(t)\big)\right)\bar{x}^N(t)dt + \sigma d\bar{w}^N(t) \tag{7}$$

where $\bar{w}^N(t) = \frac{1}{N}\sum_{i=1}^{N} w_i(t)$.
Denoting $\bar{x}(t) = E[\bar{x}^N(t)]$, and using the state transition matrix $\varphi_{\bar{x}}(t, t_0)$ (it is scalar) for the interval $[t_0, t]$ we have:

$$\bar{x}(t) = \varphi_{\bar{x}}(t, t_0)\bar{x}(t_0) \quad (8)$$

Remark: Note that the implementation of the NE control law in (4) requires that besides an agent's own state, the agent be able to observe the empirical global mean state at all times. Furthermore, as N grows without bounds, the law of large numbers dictates that it becomes sufficient to only know the initial empirical global mean. This indeed becomes the MFG solution concept. However, these global mean quantities are no longer available in the presence of a sparse information access graph.

Remark: The stability characteristics of the transition function $\varphi_{\bar{x}}(t_n, t_0)$ are determined by the properties of the Riccati equation derived from the addition of equations (5) and (6):

$$\frac{d\big(p(t) + \alpha(t)\big)}{dt} = \frac{b^2}{r}(p+\alpha)^2 - 2a(p+\alpha) - q(-\Gamma+1)$$
$$p(T) + \alpha(T) = h(1-\Gamma) \quad (9)$$

We analyze the stability behavior of $p(t) + \alpha(t)$ in two scenarios for the parameter Γ. Case 1: for $\Gamma < 1$, the boundary condition $h(1-\Gamma)$ is positive and $q(-\Gamma+1)$ is also positive. These two conditions are adequate to ensure the existence of a positively bounded solution for the Riccati equation, leading to a stabilizing gain [19], and thus $\varphi_{\bar{x}}(t_n, t_0)$ will be stable. Case 2: for $\Gamma = 1$, it can be readily demonstrated that $\alpha(t) + p(t)$ equals zero. The stability of either $\bar{x}^N$ or $\varphi_{\bar{x}}(t_n, t_0)$ hinges on the sign of parameter a.

B. *Consensus Algorithm*

In this section, we briefly review some notions of consensus algorithms (see for example [20, 21]). Consensus algorithms are quite useful when an average is to be estimated from a collection of observation points and agents can only exchange information over a graph.

Assume agents observe each other through an undirected information access graph (IAG) G. The continuous time consensus algorithm dynamics is written as follows:

$$\frac{dX^*}{dt} = -LX^* \quad (10)$$

where X^* is an $N \times 1$ vector comprising the scalar states of all agents, and L is the Laplacian matrix of the graph G. We assume λ_i for $i = 1, \dots, N$ are eigenvalues of L in ascending order and V_i for $i = 1, \dots, N$ are corresponding eigenvectors.

Two important lemmas regarding the transition matrix of the Laplacian matrix, $\varphi_L(t, t_0)$, are stated in the following. We do not provide their proofs here since these are well established materials in the consensus algorithm references [20, 21].

Lemma 1: Assume L is the Laplacian matrix of an undirected and connected graph, then the following formula holds.

$$\mathbf{1}^T\varphi_L(t, t_0) = \mathbf{1}^T$$

Lemma 2: In the transition matrix $\varphi_L(t, t_0)$, in the asymptotic cases when time goes to infinity, the following equation holds:

$$\lim_{t\to\infty} \varphi_L(t, t_0) = \frac{1}{N}\mathbf{1}\mathbf{1}^T$$

III. A Mathematical Model for Mean Field Games on a Partial IAG

In subsection A above, we approached the original MFG problem on a finite control horizon through a bottom-up approach. The latter is more restrictive than the top-down approach, starts from finite games, and views the MFG equilibrium as the limit if it exists of the Nash equilibria, if they exist, of the finite games. [7] developed for the LQ case, sufficient conditions for the bottom-up and top-down approaches to produce equivalent results. Our interest in the bottom-up approach is that it produces a state feedback policy involving the current agent state and *current* global empirical mean. Thus, it clearly displays the role of global empirical state observations needed at all times to compute the NE control action. In the limit MFG situation where the number of agents goes to infinity, the empirical global agents mean needs to be observed at least at the initial time. However, this is no longer possible when the agents interact on a partial IAG. In what follows, we propose a mathematical model where inspired again by the fish school situation, we consider separate dynamics for information exchange and physical motion. More precisely, one could hypothesize that in large groups of interacting agents, information could travel at a speed much higher than the rates of change of agent states. Thus, for example, in a fish school, fish exchange visual and pressure signals which travel much faster than they can move. The same goes for individuals moving as part of a crowd, cars within traffic, or banks reacting to market signals. This indicates that while the assumption of a complete IAG is inadequate in general, an intermediary more realistic model is one where agents on a partial IAG can exchange information about their current state (voluntarily or otherwise) at speeds which are significant relative to that of their state dynamics.

In this context of partial information access, we simplify analysis by assuming that agent state actions evolve discretely over time intervals of length $\tilde{t}$. During these intervals, agents exchange information about a system quantity and use it to synthesize a better-informed control action. The dynamics of information exchange follow a continuous time consensus process (as in (10)), with a relative information exchange rate characterized by the coefficient $r_{c/d}$, acting as an accelerator of consensus dynamics. Agents employ a "certainty equivalent" best response policy analogous to (4), replacing the empirical mean state with its best non-anticipative estimate based on all past and current information at time $t_n = t_0 + n\tilde{t}$, $n = 0, 1, 2, \dots$. Consequently, consensus algorithm dynamics including the ratio between communication time scale and dynamics time scale, $r_{c/d}$, is rewritten in (11) as:

$$\frac{dX^*}{dt} = -r_{c/d}LX^* \quad (11)$$

Remark: We should notice that the presence of $r_{c/d}$ in the differential equation of the consensus algorithm results in multiplying eigenvalues of L by $r_{c/d}$ which leads to faster convergence if $r_{c/d} > 1$.

IV. The Special Case of Initial Global Empirical Mean Information Exchange

In this study, we use an interlaced information exchange/motion dynamics model to improve estimates of the initial global empirical mean state (GEMS) over consensus cycles $\tilde{t}$. Agents use these estimates to forecast the most current GEMS for their individualized certainty equivalent control law (4). We analyze the quality of the forecast-based estimator and how it evolves over time. As time progresses, two competing effects influence the estimator quality: (i) More agents reached through consensus propagation lead to improved initial GEMS estimation, and (ii) the current forecast becomes more compromised by process noise due to increasing forecast intervals.

A. Empirical Mean Estimation Procedure

At time $t_n = t_0 + n\tilde{t}$, generic agent i calculates its current estimate $\hat{\bar{x}}_{0i}^N(t_n)$ of the initial global empirical mean state $\bar{x}^N(t_0)$. Using this estimate, agent i produces a forecast $\hat{\bar{x}}_i^N(t_n)$ of the GEMS $\bar{x}^N(t)$, assuming all agents follow trajectories dictated by (1) and (4) with $\bar{x}^N(t_n)$ replaced by $\hat{\bar{x}}_j^N(t_n), j = 1, \dots, N$. The dynamics of GEMS as observed by agent i remain governed by (8) under the proposed certainty equivalent control policy. Thus, agent i uses (8) initialized with $\hat{\bar{x}}_{0i}^N(t_n)$ to estimate GEMS at time t_n.

The summary of how the calculations proceed over time is written in table Algorithm 1.

Algorithm 1: Empirical Mean Estimation Using Consensus Algorithm

Initialization: $t = t_0$, $\hat{\bar{x}}_{0i}^N(t_0) = x_i(t_0), i = 1,2, \dots, N$

Iteration: For $n = 0,1, \dots, round(\frac{T}{\tilde{t}})$

$t \in [t_n, t_{n+1}]$

Action: Use $\hat{\bar{x}}_i^N(t_n)$ for the calculation of $u_i(t)$ based on (4) and (8) initialized with $\hat{\bar{x}}_{0i}^N(t_n)$

Communication and Estimation: Use $\hat{\bar{x}}_{0i}^N(t_n)$ to communicate with neighbors until t_{n+1} to get an estimation of $\hat{\bar{x}}_{0i}^N(t_{n+1})$ at t_{n+1}

Forecast: compute forecast, $\hat{\bar{x}}_i^N(t_{n+1})$, using $\hat{\bar{x}}_0^N(t_{n+1})$ based on (8)

Before stating Theorem 1 in the next section, we delve into the assumptions needed for it.

Assumption 1: The agent population is made up of N homogeneous agents in a connected partial IAG which is time-invariant, undirected, and transitive. Also, agents do not have prior information on the initial state distribution of the population. Initial agent states are arbitrary random variables with finite first and second moments.

Assumption 2: The transition function $\varphi_{\bar{x}}(t_n, t_0)$ exhibits stable dynamics associated with condition $\Gamma < 1$.

B. Evaluation of Quality of GEMS Estimation by Individual Agents

In what follows, we develop an expression for the mean of the GEMS estimation error by a generic agent i. In order to follow the solution easily, we denote $\hat{\bar{X}}_{0n}^N$ and $\hat{\bar{X}}_n^N$ the $N \times 1$ concatenated vectors of respectively initial GEMS estimates after running the consensus algorithm for time duration $n\tilde{t}$, and current GEMS estimates of agents at the time t_n. The error between the estimated GEMS and its true value for agent i at time t is denoted $err_i(t)$.

$$err_i(t) = \bar{x}^N(t) - \hat{\bar{x}}_i^N(t) \tag{12}$$

Theorem 1*:* Under Assumptions 1 and 2 and provided that agents employ Algorithm 1 for mean estimation and control action: (i) The GEMS achieves asymptotic unbiasedness as time increases indefinitely; (ii) Convergence to zero bias is geometric with a rate governed by the second smallest eigenvalue of L (Fiedler eigenvalue).

Proof: The proof follows from calculating the error err_i and showing that its expectation converges to zero as time grows indefinitely. We assume the global vector $X(t)$ contains all agent states. The empirical mean $\bar{x}^N$ is written based on X:

$$X(t) = [x_1(t) \quad \cdots \quad x_N(t)]^T,\ \bar{x}^N(t) = \frac{1}{N}\mathbf{1}^T X(t) \tag{13}$$

The initial condition for the consensus algorithm differential equation in (11) is $X^*(t_0) = X(t_0)$. The solution for the consensus state at the time t_n is expressed as

$$X^*(t_n) = \varphi_L(t_n, t_0)X^*(t_0) = \varphi_L(t_n, t_0)X(t_0) \tag{14}$$

At time t_n, agents find their n^{th} estimation of the GEMS for time t_0 by using the state of consensus at t_n in (14).

$$\hat{\bar{X}}_{0n}^N = X^*(t_n) \tag{15}$$

Employing equations (8), (14), and (15), agents forecast the mean for time t_n: $\hat{\bar{X}}_n^N = \varphi_{\bar{x}}(t_n, t_0)\hat{\bar{X}}_{0n}^N$. The predicted mean value for agent i, $\hat{\bar{x}}_i^N(t_n)$, corresponds to time t_n.

$$\hat{\bar{X}}_n^N = [\hat{\bar{x}}_1^N \quad \cdots \quad \hat{\bar{x}}_N^N]^T, \hat{\bar{x}}_i^N(t_n) = e_i^T \varphi_{\bar{x}}(t_n, t_0)\varphi_L(t_n, t_0)X(t_0) \tag{16}$$

In order to calculate $\bar{x}^N$, we need to average the states of agents considering the estimated value of the mean in their dynamics. The details of the calculation for $\bar{x}^N$ is given in Appendix A.

$$\bar{x}^N(t) = \varphi_{\bar{x}}(t, t_0)\frac{1}{N}\mathbf{1}^T X(t_0) + \sigma\int_{t_0}^{t}\varphi(t, s)d\bar{w}^N(s) \tag{17}$$

$$\varphi(t, t_0) = exp\left(\int_{t_0}^{t}\left(a - \frac{b^2}{r}p(s)\right)ds\right)$$

We have calculated the exact value for the mean and that estimated by agent i. Now, we calculate the error at t_n.

$$err_i(t_n) = \bar{x}^N(t_n) - \hat{\bar{x}}_i^N(t_n) = \varphi_{\bar{x}}(t_n, t_0)\frac{1}{N}\mathbf{1}^T X(t_0) +$$

$$\sigma\int_{t_0}^{t_n}\varphi(t_n, s)d\bar{w}^N(s) - e_i^T\varphi_{\bar{x}}(t_n, t_0)\varphi_L(t_n, t_0)X(t_0) \tag{18}$$

Since $E[d\bar{w}^N] = 0$, the expected value of error is:

$$E[err_i(t_n)] = \varphi_{\bar{x}}(t_n, t_0)\left[\frac{1}{N}\mathbf{1}^T - e_i^T\varphi_L(t_n, t_0)\right] E[X(t_0)] \tag{19}$$

Based on Lemma 2, when n goes to ∞,

$$\lim_{n\to\infty} e_i^T\varphi_L(t_n, t_0) = \frac{1}{N}\mathbf{1}^T \Rightarrow \lim_{n\to\infty} E[err_i(t_n)] = 0 \qquad \blacksquare$$

Remark: In Theorem 1, as indicated by equation (19), the convergence rate of $E[err_i(t_n)]$ is determined by $\exp\left(\int_{t_0}^{t_n}\left(a - \frac{b^2}{r}(\alpha(\tau) + p(\tau))\right)d\tau - r_{c/d}\lambda_2(t_n - t_0)\right)$. With the condition $\Gamma < 1$, the stability of $\varphi_{\bar{x}}(t_n, t_0)$ is assured, resulting in faster convergence of consensus facilitated by the stable dynamics of the agents.

Theorem 2: Given Assumption 1 and $\Gamma = 1$, for $E[err_i(t_n)]$ to converge to zero as n goes to infinity, it is necessary and sufficient that $r_{c/d} > \frac{a}{\lambda_2}$.

Proof: When $\Gamma = 1$, as previously discussed, $\varphi_{\bar{x}}(t, t_0)$ simplifies to $\exp(a(t - t_0))$. Furthermore, according to equation (19), the convergence rate of $E[err_i(t_n)]$ is represented by $\exp\left((a - r_{c/d}\lambda_2)(t_n - t_0)\right)$. Therefore, for the objective of driving $E[err_i(t_n)]$ towards zero, it is essential that $a - r_{c/d}\lambda_2 < 0$. This condition ensures the desired outcome. $\blacksquare$

Remark: Among partial IAGs, cycle graphs offer a particularly illustrative worst-case scenario. Indeed, in these graphs, each agent is linked to only two neighbors. The eigenvalues of cycle graphs are expressed by equation (20):

$$\lambda_j = 4\,sin^2\left(\frac{\pi(j-1)}{N}\right),\ \ j = 1, \dots, N \tag{20}$$

For cycle graphs, as the number of agents approaches infinity, it follows that λ_2 tends towards $4\left(\frac{\pi}{N}\right)^2$. This observation highlights that for cycle graphs as N increases, λ_2 diminishes in a manner proportional to N^{-2}, consequently leading to a reduction in the convergence rate of the proposed algorithm.

V. Properties of GEMS Estimator

In this section, assuming an i.i.d. initial distribution for agents, we shall provide an analysis regarding the expected value and variance of the error in (12).

Assumption 3: Agents have i.i.d. initial distribution with finite mean $E[x_i(t_0)] = x_0$ and finite variance $E[(x_i(t_0) - x_0)^2] = \sigma_0^2\ (or\ \Sigma_{X(t_0)} = \sigma_0^2 I)$.

Proposition 1: Under Assumptions 1 and 3 the GEMS estimated by agent i, $\hat{\bar{x}}_i^N(t_n)$, is an unbiased estimator of $\bar{x}^N(t_n)$ at all times.

Proof: We have:

$$E[err_i(t_n)] = \varphi_{\bar{x}}(t_n, t_0)\left[\frac{1}{N}\mathbf{1}^{\boldsymbol{T}} - e_i^T\varphi_L(t_n, t_0)\right] E[X(t_0)]$$

In view of the i.i.d. initial state distribution assumption, Lemma 2, and the symmetry of matrix L, one can write:

$$E[X(t_0)] = x_0\mathbf{1},\ \ e_i^T\varphi_L(t_n, t_0)\mathbf{1}x_0 = x_0$$

As a result, $E[err_i(t_n)] = 0$, and $\hat{\bar{x}}_i^N(t_n)$ is an unbiased estimator of $\bar{x}^N(t_n)$. ■

A. Error Variance Analysis

In this subsection, we derive the variance of error estimation. Our goal is to analyze the evolution of the GEMS estimator variance over time.

$$\Sigma_{err_i}(t_n) = E[(err_i(t_n) - E[err_i(t_n)])^2] = \frac{\sigma^2}{N}\int_{t_0}^{t_n}\varphi^2(t_n, s)ds + \varphi_{\bar{x}}^2(t_n, t_0)\left[\frac{1}{N}\mathbf{1}^{\boldsymbol{T}} - e_i^T\varphi_L(t_n, t_0)\right]\Sigma_{X(t_0)}\left[\frac{1}{N}\mathbf{1}^{\boldsymbol{T}} - e_i^T\varphi_L(t_n, t_0)\right]^T$$

In the following, the error variance is calculated under Assumption 3 using eigenvalues and eigenvectors of L.

$$\Sigma_{err_i}(t_n) = \frac{\sigma^2}{N}\int_{t_0}^{t_n}\varphi^2(t_n, s)ds + \varphi_{\bar{x}}^2(t_n, t_0)\sigma_0^2\sum_{j=2}^{N} exp^2\left(-r_{c/d}\lambda_j(t_n - t_0)\right) e_i^T V_j V_j^T e_i \tag{21}$$

We now discuss the behavior of GEMS estimation variance based on the stability properties of $\varphi_{\bar{x}}(t_n, t_0)$.

Remark: The first term in (21) is the variance contributed by the Wiener noise term in the GEMS forecast based on (1). It is affected by the state transition function $\varphi(t_n, t_0)$ which is guaranteed to fall to zero exponentially due to the stabilization properties of the associated Riccati equation-based gain. The second term is the variance associated with the forecast of the consensus-based initial GEMS estimate.

Proposition 2: Given Assumptions 1, 3, and the condition $\Gamma < 1$, it can be established that the variance of the error as defined in equation (21) remains bounded.

Proof: As previously mentioned, in the scenario where $\Gamma < 1$, the function $\varphi_{\bar{x}}(t, t_0)$ exhibits exponential decay as time progresses. This behavior leads to the convergence of the second term in equation (21) to zero, leaving only the first term in play. Consequently, variance remains bounded.

Proposition 3: Under Assumptions 1 and 3, when $\Gamma = 1$, the variance experiences a downward trend as time progresses if the communication dynamics, represented by $r_{c/d}\lambda_2$, outpace the system dynamics, represented by a, signifying that $r_{c/d}$ should be greater than $\frac{a}{\lambda_2}$.

Proof: The long term behavior of the consensus-induced variance, as represented in (21), is determined by the term $\varphi_{\bar{x}}^2(t_n, t_0)\,exp^2\left(-r_{c/d}\lambda_2(t_n - t_0)\right)$, which equates to $exp^2\left((a - r_{c/d}\lambda_2)(t_n - t_0)\right)$. Analyzing this growth rate, it becomes evident that in order to ensure the boundedness of the variance, the condition $r_{c/d}\lambda_2 > a$ must be satisfied.

Remark: In Theorem 2 and Proposition 3, as discussed earlier, in a worst case cycle graph, when the number of agents N approaches infinity, λ_2 diminishes at a rate proportional to N^{-2}. For these assertions to hold, it becomes necessary for the communication parameter $r_{c/d}$ to escalate in proportion to N^2. This observation underscores the need for communication speed to increase quadratically as the agents' network expands while the IAG remains sparse, ensuring that agents remain well-informed about the population mean for effective decision-making.

B. Simulation of Error Variance

In Figure 1, error variance (21) is plotted for $\Gamma = 0.6$ across different population sizes. As N increases, noise-induced variance weakens, causing an overall reduction in variance. Additionally, higher N leads to a longer time to reach minimum variance. These trends suggest agents might eventually trust their GEMS estimates, implying a Nash equilibrium.

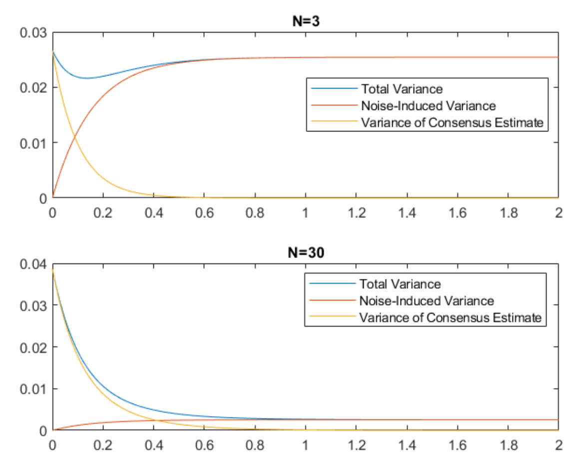

Figure 1. The error variance for different N. Parameters of the simulation: a=0, b=1, q=1, r=0.1, h=1, σ=0.7, Γ=0.6, $r_{c/d} = 1$,$\sigma_0 = 0.2$

VI. Conclusion

In this paper, we discussed the question of possible convergence to a Nash equilibrium amongst large systems of exchangeable agents interacting through an incomplete but connected information access graph. An observations-dynamics model was proposed to capture the possible separation of time scales between communications and dynamics. The communications part is assumed to be effectively equivalent to a consensus algorithm the information of which reaches the controller with some time delay. The agents use a particular estimation scheme to evaluate their cost and produce a Nash certainty equivalent policy. For the worst-case Partial IAG -cycle graphs- analysis suggests that for stable mean field dynamics, convergence to an ideal MFG equilibrium over time is helped by the consensus process and will always occur. For unstable mean field dynamics over a finite horizon, convergence for a given large network size can be helped by increasing the relative communication to dynamics speed. In future work, we shall explore more thoroughly the role of IAG structure, communication speed, and possibly more sophisticated estimation schemes on the cost regret relative to a complete IAG when the number of agents grows to infinity.

APPENDIX

In the following, we will calculate $\bar{x}^N(t)$ by averaging $x_i(t)$.

$$dx_i = \left(\left(a - \frac{b^2}{r}p(t)\right)x_i - \frac{b^2}{r}\alpha(t)\hat{\hat{x}}_i^N(t)\right)dt + \sigma dw_i$$

$$x_i(t) = \varphi(t,t_0)x_i(t_0) - \frac{b^2}{r}\int_{t_0}^{t}\varphi(t,s)\alpha(s)\hat{\hat{x}}_i^N(s)ds + \sigma\int_{t_0}^{t}\varphi(t,s)dw_i(s)$$

Lemma 3: By using Lemma 1, we can easily show that

$$\frac{1}{N}\sum_{i=1}^{N}\hat{\hat{x}}_{\text{in}}^{\text{N}}(t) = \frac{1}{N}\mathbf{1}^{T}X(t_0)\varphi_{\bar{x}}(t,t_0) = E[\bar{x}^N(t)]$$

Using Lemma 3, $\bar{x}^N$ is calculated.

$$\bar{x}^N(t) = \varphi(t,t_0)\frac{1}{N}\mathbf{1}^{T}X(t_0) - \frac{b^2}{r}\frac{1}{N}\mathbf{1}^{T}X(t_0)\int_{t_0}^{t}\varphi(t,s)\varphi_{\bar{x}}(s,t_0)\alpha(s)ds$$
$$+\frac{\sigma}{N}\int_{t_0}^{t}\varphi(t,s)\sum_{i=1}^{N}dw_i(s)$$

Using the following equation to simplify the integral.

$$\varphi(t,s)\varphi_{\bar{x}}(s,t_0) = \varphi(t,t_0)\,exp\left(\int_{t_0}^{s} -\frac{b^2}{r}\alpha(l)dl\right)$$

So, $\bar{x}^N(t)$ is rewritten.

$$\bar{x}^N(t) = \varphi(t,t_0)\frac{1}{N}\mathbf{1}^{T}X(t_0)\left(1 - \frac{b^2}{r}\int_{t_0}^{t}\alpha(s)\,exp\left(\int_{t_0}^{s} -\frac{b^2}{r}\alpha(l)dl\right)ds\right)$$
$$+\frac{\sigma}{N}\int_{t_0}^{t}\varphi(t,s)d\bar{w}^N(s)$$

Using the fact that: $1 - \frac{b^2}{r}\int_{t_0}^{t}\alpha(s)\exp\left(\int_{t_0}^{s} -\frac{b^2}{r}\alpha(l)dl\right)ds = \exp\left(\int_{t_0}^{t} -\frac{b^2}{r}\alpha(s)ds\right)$.

$$\bar{x}^N(t) = \varphi(t,t_0)\frac{1}{N}\mathbf{1}^{T}X(t_0)\,exp\left(\int_{t_0}^{t} -\frac{b^2}{r}\alpha(s)ds\right) + \sigma\int_{t_0}^{t}\varphi(t,s)d\bar{w}^N(s)$$

Finally, using $\varphi(t,t_0)\exp\left(\int_{t_0}^{t} -\frac{b^2}{r}\alpha(s)ds\right) = \varphi_{\bar{x}}(t,t_0)$, one can find $\bar{x}^N(t)$.

$$\bar{x}^N(t) = \varphi_{\bar{x}}(t,t_0)\frac{1}{N}\mathbf{1}^{T}X(t_0) + \sigma\int_{t_0}^{t}\varphi(t,s)d\bar{w}^N(s)$$

REFERENCES

[1] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics,* vol. 2, no. 1, pp. 229-260, 2007/03/01 2007, doi: 10.1007/s11537-007-0657-8.

[2] O. Guéant, J.-M. Lasry, and P.-L. Lions, "Mean Field Games and Applications," in *Paris-Princeton Lectures on Mathematical Finance 2010*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 205-266.

[3] M. Huang, R. P. Malhamé, and P. E. Caines, "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.

[4] M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized $\boldsymbol{\varepsilon}$-Nash equilibria," *IEEE transactions on automatic control,* vol. 52, no. 9, pp. 1560-1571, 2007.

[5] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications I-II*. Springer.

[6] R. Carmona, F. Delarue, and D. Lacker, "Mean field games with common noise," *The Annals of Probability,* vol. 44, no. 6, pp. 3740-3803, 2016.

[7] M. Huang and M. Zhou, "Linear quadratic mean field games: Asymptotic solvability and relation to the fixed point approach," *IEEE Transactions on Automatic Control,* vol. 65, no. 4, pp. 1397-1412, 2019.

[8] F. Delarue, "Mean field games: A toy model on an Erdös-Renyi graph," *ESAIM: Proceedings and Surveys,* vol. 60, pp. 1-26, 2017.

[9] D. Lacker and A. Soret, "A case study on stochastic games on large graphs in mean field and sparse regimes," *Mathematics of Operations Research,* 2021.

[10] R. Zhu, J. Zhang, and K. You, "Networked aggregative games with linear convergence," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021: IEEE, pp. 3381-3386.

[11] J. Koshal, A. Nedić, and U. V. Shanbhag, "Distributed algorithms for aggregative games on graphs," *Operations Research,* vol. 64, no. 3, pp. 680-704, 2016.

[12] S. Liang, P. Yi, and Y. Hong, "Distributed Nash equilibrium seeking for aggregative games with coupled constraints," *Automatica,* vol. 85, pp. 179-185, 2017/11/01/ 2017.

[13] F. Parise, S. Grammatico, B. Gentile, and J. Lygeros, "Distributed convergence to Nash equilibria in network and average aggregative games," *Automatica,* vol. 117, p. 108959, 2020/07/01/ 2020.

[14] G. Belgioioso, A. Nedić, and S. Grammatico, "Distributed Generalized Nash Equilibrium Seeking in Aggregative Games on Time-Varying Networks," *IEEE Transactions on Automatic Control,* vol. 66, no. 5, pp. 2061-2075, 2021, doi: 10.1109/TAC.2020.3005922.

[15] H. Kebriaei, S. J. Sadati-Savadkoohi, M. Shokri, and S. Grammatico, "Multipopulation Aggregative Games:Equilibrium Seeking via Mean-Field Control and Consensus," *IEEE Transactions on Automatic Control,* vol. 66, no. 12, pp. 6011-6016, 2021, doi: 10.1109/TAC.2021.3057063.

[16] R. Zhu, J. Zhang, K. You, and T. Başar, "Asynchronous Networked Aggregative Games," *Automatica,* vol. 136, p. 110054, 2022/02/01/ 2022.

[17] T. Başar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1998.

[18] H. Minyi, P. E. Caines, and R. P. Malhame, "Uplink power adjustment in wireless communication systems: a stochastic control analysis," *IEEE Transactions on Automatic Control,* vol. 49, no. 10, pp. 1693-1708, 2004, doi: 10.1109/TAC.2004.835388.

[19] D. H. Jacobson, "New conditions for boundedness of the solution of a matrix Riccati differential equation," *Journal of Differential Equations,* vol. 8, no. 2, pp. 258-263, 1970/09/01/ 1970.

[20] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control,* vol. 49, no. 9, pp. 1520-1533, 2004, doi: 10.1109/TAC.2004.834113.

[21] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE,* vol. 95, no. 1, pp. 215-233, 2007.