IEEE Control Systems Letters paper presented at
2023 American Control Conference (ACC)
San Diego, CA, USA. May 31 - June 2, 2023

Large Population Games with Timely Scheduling over Constrained Networks

Shubham Aggarwal, Muhammad Aneeq uz Zaman, Melih Bastopcu, Tamer Başar

Abstract— **In this paper, we consider a discrete-time multi-agent system involving N cost-coupled networked rational agents solving a consensus problem, and a central Base Station (BS), scheduling agent communications over a network. Due to an average bandwidth constraint on the number of transmissions, the BS can let at most $R_d < N$ agents to access their state information through the network on average. For the scheduling problem, we propose a novel *weighted age of information (WAoI)* metric. Then, under standard information structures, we are able to separate the estimation and control problems for each agent. We first solve an unconstrained MDP problem and then compute an optimal policy for the original problem using the solution to the former problem. Next, we solve the consensus problem using the mean-field game framework wherein we first design decentralized control policies for a limiting case of the N–agent system as $N \to \infty$, and prove the existence of a unique mean-field equilibrium. Consequently, we show that the obtained equilibrium policies constitute an ϵ–Nash equilibrium for the finite agent system. Finally, we validate the performance of both the scheduling and the control policies through numerical simulations.**

I. Introduction

With the emergence of time-critical applications such as real-time monitoring in surveillance systems, autonomous vehicular systems, internet-of-things and cyberphysical security [1], *networked control systems* (NCS) promise interesting research directions within both the communication and the controls community. Such systems involve large populations of spatially distributed agents, and allow for remote information sharing and decentralized execution of the designated tasks. However, while decentralization reduces the storage complexity of the servers, limited information availability directly affects the system performance. Thus, appropriate information structures need to be assigned to each system component alongside *timely* and *accurate* transmission of time-sensitive sensor measurements to the corresponding control units. All these factors are critical in determining the optimal scheduling and control policies, the design of which forms the major objective of this paper.

In this paper, we consider a discrete-time problem among $N+1$ players (N agents and a Base Station (BS)). The N cost-coupled agents solve a consensus problem [2], where each agent constitutes two active decision makers–a controller and an estimator, which have access to their state information through a wireless communication network, controlled by the BS as shown in Fig. 1. As in the case of the real world wireless communication systems, the medium connecting the BS with the agents has a limited average bandwidth of $R_d < N$ units. Due to this bottleneck, the agents may have intermittent access to their state information. To optimally schedule the wireless communication between the agents, we propose a novel Age of Information (AoI) based performance metric for the BS called the *Weighted*-AoI (WAoI). This poses a novel $N+1$ agent game where the N agents are individually trying to solve a consensus problem among themselves while the BS is trying to minimize a WAoI based metric using a scheduling policy. Due to the presence of a large population of agents communicating over the network, we solve the consensus problem using a mean-field game (MFG) setting, and finally provide approximate Nash policies for the N–agent game problem.

Related Literature: Since the seminal works [3], [4] established separation properties in systems with limited information, there has been quite a lot of work on information structures in stochastic decision making problems [5]. Further, works dealing with networked control problems include (but are not limited to) [6] with uninterrupted, and [7]–[9] with intermittent communications under contention and/or constraints on the resource acquisitions or the number of transmissions. These works, however, contain no strategic interaction between the agents as in a game setting, where the scalability is the primary concern as the complexity increases exponentially with the number of agents.

The novel idea of MFGs [10], [11] solves the scalability problem by considering the limiting case as the number of agents grows, i.e., $N \to \infty$. As N gets large, individual deviations have negligible effect on the mass of agents (mean-field), causing strategic interaction between agents to disappear. The game can now be characterized by the interaction between a generic agent and the mean-field. While the earliest works in MFGs dealt with continuous-time systems, research on discrete-time setting has recently gained momentum, especially the benchmark Linear-Quadratic MFGs (LQ-MFGs) [12], [13]. Except for intermittent communication in [12], most works in LQ-MFGs consider continuous and reliable communication. A recent work [2] deals with the case where agents have access to their states through a noisy channel controlled by a fixed scheduler. In this paper, however, we utilize the emerging notion of AoI to devise optimal scheduling policies over a bandwidth constrained network.

Age of information was introduced to measure the timeliness of information in communication networks. Scheduling problems in AoI have received significant attention recently.

The authors are affiliated with the Coordinated Science Lab, University of Illinois at Urbana-Champaign, Urbana, IL, USA 61801. Emails: {sa57,mazaman2,bastopcu,basar1}@illinois.edu.
Research of the authors was supported in part by ARO MURI Grant AG285 and in part by AFOSR Grant FA9550-19-1-0353.

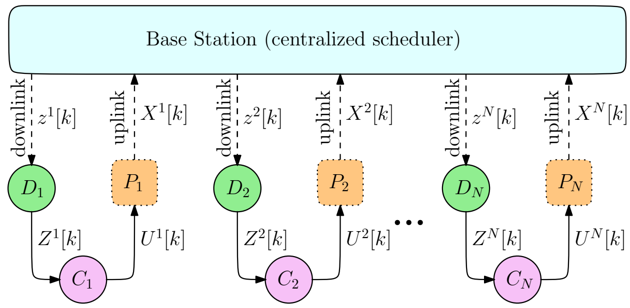

Fig. 1: A prototypical NCS showing hierarchy of decision making at the Base Station and the agents. Solid blocks (Base Station, Decoders, Controllers) denote active decision makers while dotted blocks (plants) denote passive components. Dashed lines denote a wireless transfer medium while bold lines denote wired transfer.

Among the policies employed for solving scheduling problems include the maximum-age-first policy [14], and the maximum-weighted-age-reduction policy [15], among others. Age-optimal scheduling policies have been obtained by using MDP formulation in [16], [17]. Recently, AoI has also been studied in the context of networked control problems. In [18], the authors compared the performances of AoI and value of information (VoI) based scheduling algorithms. While [19] proposes a discounted error scheduler by using MDP approach for a truncated state-space of the AoI, [20] studies a resource allocation problem to minimize MSE through AoI. A detailed literature review on AoI can be found in [21].

Contribution: In this work, we solve a large population game problem involving N networked dynamical agents interacting strategically with each other to minimize individual coupled cost functions and a BS aiming to minimize a performance measure by scheduling communication over a bandwidth-constrained wireless medium. For the scheduling problem, inspired by the Age of Incorrect Information (AoII) metric [22], we first introduce a weighted AoI metric as the performance measure for the BS, which is a function of the agent's estimation error (due to intermittent transmissions by the BS) and the AoI at the controller. Then, we construct a randomized scheduling policy to solve the scheduling problem at the BS. Finally, due to the scalability issue in solving the consensus problem for large N, we use the MFG framework to construct ϵ-Nash strategies with $\epsilon = \mathcal{O}(1/\sqrt{\min_{\theta\in\Theta}|N_\theta|})$, where $|N_\theta|$ is the number of agents of type $\theta \in \Theta$, obtained by the existence, uniqueness and linearity of the Mean-Field Equilibrium.

Organization: We formulate the $N+1$ player game problem in Sec. II. In Sec. III, we analyze the centralized scheduling problem of the BS and construct an optimal policy for the same in Sec. IV. In Sec. V, we solve the consensus problem and show its ϵ–Nash property for the finite agent game problem. In Sec. VI, we provide numerical analysis and conclude the paper in Sec. VII. The proofs of the supporting lemmas, propositions and theorems can be found in the full version of the paper [23] along with detailed additional simulations.

Notations: $\mathbb{Z}^+$ denotes the set of non-negative integers. For a matrix S and a vector x, $\|x\|_S^2 := x^\top Sx$. Further, $[N] := \{1, 2, \cdots, N\}$ and $tr(\cdot)$ denotes the trace of its argument. The Euclidean norm for vectors, or the induced 2-norm for matrices is denoted by $\|\cdot\|$, and $\|\cdot\|_F$ denotes the Frobenius norm of its argument. Further, $Y_{a:a+k} := \{Y[a], \cdots, Y[a+k]\},\ k \geq 0$. All empty summations are set to 0.

II. System Model and Problem Formulation

We start by formulating the $N+1$ player game, which can be expressed as 1) a consensus problem between N agents and 2) a centralized scheduling problem of the BS.

A. Consensus Problem

We consider a discrete-time N-agent game on an infinite horizon, communicating over a wireless network. Each agent's dynamics evolves as:

$$X^i[k+1] = A(\theta_i)X^i[k] + B(\theta_i)U^i[k] + W^i[k], \qquad (1)$$

for a time-step $k \in \mathbb{Z}^+$ and agent $i \in [N]$. Here $X^i[k] \in \mathbb{R}^n$ and $U^i[k] \in \mathbb{R}^m$ are the state and control action, respectively of agent i. The noise process for agent i, $W^i[k] \in \mathbb{R}^n$ has zero mean and bounded positive-definite covariance $K_W(\theta_i)$. Agent i's initial state $X^i[0]$ is independent of the noise process and is assumed to have a symmetric density with mean $\nu_{\theta_i,0}$ and bounded positive-definite covariance Σ_x. The time-invariant system matrices $A(\theta_i) \in \mathbb{R}^{n\times n}$ and $B(\theta_i) \in \mathbb{R}^{n\times m}$ depend on the agent type $\theta_i \in \Theta := \{\theta_1, \cdots, \theta_p\}$ which is chosen according to the empirical probability mass function $\mathbb{P}_N(\theta = \theta_i),\quad \theta \in \Theta$. It is further assumed that $|\mathbb{P}_N(\theta) - \mathbb{P}(\theta)| = \mathcal{O}(1/N)$ for all $\theta \in \Theta$, where $\mathbb{P}(\theta)$ is the limiting distribution. We now state the following assumption on information transmission over the network.

Assumption 1. *The wireless links connecting system components are error-free and the BS can send measurements to the corresponding controllers instantaneously.*

Due to Assumption 1, the information can be transmitted from the plant to the controller for its next action without any delay, if the BS decides to send an update for that agent.

The $N+1$-player system is shown in Fig. 1. For each agent, the full-state information of the plant P_i is relayed to the decoder D_i through a (noiseless) two-hop network (called uplink and downlink) via a centralized BS, which is then communicated to the controller C_i for generating an actuation signal. The uplink in the wireless network is ideal while the downlink can transmit only $R_d < N$ users on average, which acts as a bottleneck for transmission of information from the plants to their respective controllers. Under Assumption 1, the state of the i^{th} plant as observed by the i^{th} decoder is given as $z^i[k] = X^i[k]$, if $\gamma^{d,i}[k] = 1$, or $z^i[k] = \varphi$, if $\gamma^{d,i}[k] = 0$, where $\gamma^{d,i}[k] = 1$ denotes that *current state* information is transmitted to the i^{th} decoder (over the downlink) while $\gamma^{d,i}[k] = 0$ stands for no transmission (or φ). Additionally, between transmission times, the decoder calculates the minimum mean square estimate $Z^i[k] = \mathbb{E}\left[X^i[k] \mid I^{d,i}[k]\right]$ based on its information history $I^{d,i}[k] := \left\{z^i_{0:k}, \gamma^{d,i}_{0:k}, U^i_{0:k-1}\right\}$, which is then sent to the controller. Typically, in the game problems as formulated above, the control action of agent i can depend on other agents' state and control actions, and hence the information history of the i^{th} controller would be given by $I^{c,con,i}[k] := \{U^i_{0:k-1}, Z^i_{0:k}\}_{i\in[N]}$. Here, $I^{c,con,i}[k]$ denotes a centralized information structure [2] where the controller has knowledge of not only its own but also of other agents' controller states and actions. This entails that $\mathcal{M}^{c,con}_i = \{\pi^i \mid \pi^i \text{ is adapted to } \sigma(I^{c,con,i}[s],\ s = 0, \cdots, k)\}$, where $\sigma(\cdot)$ is the σ–algebra adapted to its argument and $\mathcal{M}^{c,con}_i$ is the space of admissible centralized control policies for agent i.

Now, each agent i aims to minimize its infinite-horizon average cost function

$$J^N_i(\pi^i, \pi^{-i}) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1} \|X^i[k] - \frac{1}{N}\sum_{j=1}^{N} X^j[k]\|^2_{Q(\theta_i)} + \|U^i[k]\|^2_{R(\theta_i)}\right], \tag{2}$$

where $Q(\theta_i) \geq 0$ and $R(\theta_i) > 0$. The consensus-like term $\frac{1}{N}\sum_{j=1}^{N} X^j[k]$ couples the agents' cost. The cost function penalizes deviations from the consensus term and large control effort. We define $\pi^{-i} := (\pi^1, \cdots, \pi^{i-1}, \pi^{i+1}, \cdots, \pi^N)$, where $\pi^i := (\pi^i[1],\ \pi^i[2], \cdots,) \in \mathcal{M}^{c,con}_i$ denotes a control policy for the i^{th} agent. Finally, the expectation in (2) is taken with respect to the noise statistics and the initial state distribution. We note now that having access to (and keeping track of) the information of other agents in a large population setting is quite difficult, and hence we will resort to the MFG framework (in Section V) to characterize decentralized control policies whereby decisions are made based on an agent's local information. We also remark here that the estimator and the controller *work together* in a team setting to minimize (2), and as we will see, can be designed independently of each other. Finally, the BS centrally schedules transmissions over the downlink as discussed next.

B. Centralized Scheduling Problem

Consider the most recently received observations by controller i as $z^i[s^i[k]]$, where $s^i[k] = \sup\{s \in \mathbb{Z}^+ : s \leq k, z^i[s] \neq \varphi\}$ denotes the latest transmission time. By definition, the AoI is the time elapsed since the generation time-stamp of the most recent packet at the plant. Thus, the AoI $\Delta^i[k]$ at the controller C_i is given as $\Delta^i[k] = k - s^i[k]$. More precisely, we have $\Delta^i[k+1] = 0$ if $\gamma^{d,i}[k] = 1$, and $\Delta^i[k+1] = \Delta^i[k] + 1$, if $\gamma^{d,i}[k] = 0$. Thus, the constrained scheduling problem is defined as:

Problem 1.

$$\inf_{\zeta^d\in\mathcal{Z}^d} \ J^S(\zeta^d) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{T-1}\sum_{i=1}^{N} \eta^i[k]\Delta^i[k]\right]$$

$$\text{s.t.} \quad \limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1}\sum_{i=1}^{N} \gamma^{d,i}[k] \leq R_d,$$

where $\zeta^d = [\zeta^{d,1}, \cdots, \zeta^{d,N}]^\top$, $\mathcal{Z}^d$ is the space of scheduling policies across all agents, and $\gamma^{d,i}[k]$ is chosen from the policy $\zeta^{d,i}$ at instant k. Further, we define $\eta^i[k]$ as the importance weight associated with agent i given by $\eta^i[k] := \mathbb{E}[\|e^i[k]\|^2]$, where $e^i[k] := X^i[k] - Z^i[k]$ is the estimation error between the plant state and the controller state at instant k. Finally, the expectation is taken over the stochasticity induced by (possibly) randomized policies.

III. Centralized Scheduling Problem Analysis

We start this section by computing the best estimate at the decoder. Also, to avoid cluttering notations, we use the shorthands $A_i := A(\theta_i)$, $B_i := B(\theta_i)$, and $K_{W^i} := K_W(\theta_i)$, unless specified otherwise.

Based on its input $z^i[k]$, the decoder computes the best estimate of the state as $Z^i[k] = X^i[k]$, if $\gamma^{d,i}[k] = 1$, and $Z^i[k] = \mathbb{E}\left[X^i[k] \mid I^{d,i}[k], \gamma^{d,i}[k] = 0\right]$ if $\gamma^{d,i}[k] = 0$, where $\mathbb{E}\left[X^i[k] \mid I^{d,i}[k], \gamma^{d,i}[k] = 0\right] = A_i\mathbb{E}[X^i[k-1] \mid I^{d,i}[k-1]] + B_iU^i[k-1] + \hat{W}^i[k-1]$ and $\hat{W}^i[k-1] := \mathbb{E}\left[W^i[k-1] \mid \gamma^{d,i}[k] = 0\right]$. Note that the presence of transmission instants in the conditioning leads to the extra term $\hat{W}^i[k-1]$. While this conveys additional information on the state of the plant in the absence of any communication between the BS and the decoder, it is typically hard to compute optimally. Here, however, it is easy to show using similar arguments as in [7], and assumptions of symmetric densities of initial state and the noise process, that $\hat{W}^i[k-1] = 0$. Hence, the optimal decoder is given by

$$Z^i[k] = \begin{cases} X^i[k], & \text{if } \gamma^{d,i}[k] = 1, \\ A_iZ^i[k-1] + B_iU^i[k-1], & \text{if } \gamma^{d,i}[k] = 0, \end{cases} \tag{3}$$

where $A_iZ^i[k-1] + B_iU^i[k-1]$ can be thought of as the recursive estimate calculated between transmission instants based on the information history. The signal $Z^i[k]$ can now be computed easily using (3). We summarize the above results in the following lemma.

Lemma 1. *The estimation error $e^i[k]$ for all agents is independent of the control inputs, and hence there is no dual effect of control [2]. Moreover,*

$$e^i[k] = \begin{cases} 0, & \textit{if } \ \gamma^{d,i}[k] = 1, \\ \sum_{l=1}^{\Delta^i[k]} A_i^{l-1}W^i[k-l], & \textit{if } \ \gamma^{d,i}[k] = 0, \end{cases} \tag{4}$$

and the covariance of the estimation error can be formulated in terms of the AoI, i.e., $\mathbb{E}[\|e^i[k]\|^2] := h(\Delta^i[k], A_i, K_{W^i}) = \sum_{l=1}^{\Delta^i[k]} tr\left({A_i^{l-1}}^\top A_i^{l-1}K_{W^i}\right)$.

Thus, Problem 1 can be equivalently written as:

Problem 2.

$$\inf_{\zeta^d\in\mathcal{Z}^d} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{T-1}\sum_{i=1}^{N} \underbrace{h(\Delta^i[k], A_i, K_{W^i})\Delta^i[k]}_{=:g(\Delta^i[k], A_i, K_{W^i})}\right]$$

$$\text{s.t.} \quad \limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1}\sum_{i=1}^{N} \gamma^{d,i}[k] \leq R_d. \tag{5}$$

We note that Problem 2 does not only depend on the AoI but also on the system parameters of each agent such as A_i, and K_{W^i}. This entails that $g(\Delta^i[k], A_i, K_{W^i})$ takes into account the system dynamics as well. Hence, we refer to it as *Control-Aware AoI* [18]. Additionally, as a consequence of this reformulation, the scheduler does not have to store the plant states and the controller actions for any agent. Thus, no additional storage space is required for the scheduler.

Remark 1. *Note that the AoI, $\Delta^i[k]$, may not be omitted at the cost function in Problem 2. This is because the error over an infinite horizon may approach a finite limit (for instance, consider stable agents), which can cause the AoI of those agents to approach infinity since a trigger for information transmission may never be generated for these agents. This can lead to poor regulation. Thus, the use of a weighted metric as in Problem 2 penalizes both the error as well as the AoI and is appropriately justified.*

Next, we observe that the constraint (5) entails that more than R_d users are able to transmit over the channel at some times as long as the average transmissions satisfy the constraint over the infinite horizon. Thus, to solve Problem 2, we reformulate the N–agent scheduling problem as a single agent discrete-time MDP to find the optimal scheduling strategies. To this end, we introduce the Lagrangian function as $\mathcal{L}(\zeta^d, \lambda) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}[\frac{1}{N}\sum_{k=0}^{T-1}\sum_{i=1}^{N}[g(\Delta^i[k], A_i, K_{W^i}) + \lambda\gamma^{d,i}[k] - \lambda\frac{R_d}{N}]]$, where $\lambda \geq 0$ is the Lagrange multiplier. Such a multiplier can be thought of as the cost of scheduling for each agent over the channel. Thus, for a fixed λ, the decoupled single user optimization problem is defined as:

Problem 3.

$$\inf_{\zeta^{d,i}\in\mathcal{Z}^{d,i}} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1} g(\Delta^i[k], A_i, K_{W^i}) + \lambda\gamma^{d,i}[k]\right].$$

Since Problem 3 is solved for a single user, we henceforth drop the superscript i until mentioned otherwise.

IV. Decentralized Scheduling Problem

We now formulate Problem 3 into a discrete-time controlled MDP $\mathbb{M}$, defined by the quadruplet $\mathbb{M} := (\mathcal{S}, \mathcal{A}, \mathcal{P}, C)$. The *state space* $\mathcal{S} = \mathbb{Z}^+$ is the set of all possible AoI of the agent and is countably infinite. The *action space* is $\mathcal{A} = \{0, 1\}$, where $a = 1$ denotes that the agent is connected over the channel while $a = 0$ stands for no transmission. Note here that a is different from $\gamma^d[k]$, which is the action under a constrained problem. The *probability transition function* $\mathcal{P}$ denotes the evolution of the states of the controlled system. When $a = 0$, we have $\mathcal{P}(\Delta \to \Delta + 1) = 1$. When $a = 1$, we have $\mathcal{P}(\Delta \to 0) = 1$. We further note that although the states evolve deterministically, writing them in the form of an MDP will simplify the notation. The *one-stage cost* $C(\Delta, a) = g(\Delta, A, K_W) + \lambda a$ denotes the cost incurred when an action a is taken at the state Δ. Next, we formally define the decentralized scheduling problem.

Problem 4.

$$\inf_{\zeta^d\in\mathcal{Z}^d} V(\Delta, \gamma^d) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1} C(\Delta[k], a[k])\right]. \quad (6)$$

In the following, we provide the solution to Problem 4.

A. Single-agent Deterministic Scheduling Policy

We first provide the following definition to be used in Theorem 1 for characterizing optimality of a threshold policy.

Definition 1. *A scheduling policy ζ^d for the MDP $\mathbb{M}$ is $g(\Delta)$–optimal if it infimizes the time-average cost $V(\Delta, \zeta^d)$.*

We now state the following theorem regarding existence of optimal policy and its threshold structure.

Theorem 1. *Given $\lambda \in \mathbb{R}$, there exists a $g(\Delta)$–optimal stationary deterministic policy solving Problem 4, and has a threshold structure, i.e., $\exists\tau := \tau(\lambda, A, K_W)$ such that*

$$a = \begin{cases} 1, & \Delta \geq \tau, \\ 0, & \Delta < \tau. \end{cases} \quad (7)$$

Having established the threshold structure of the $g(\Delta)$–optimal policy for the single-agent scheduling Problem 4, we can restrict our attention to the finite-state MDP with the state space $\mathcal{S}' = \{0, 1, \cdots, \tau\}$. Then, the one-stage cost for the time-average cost function satisfies the Bellman's equation given by $V(\Delta) + \sigma^* = \min\{C(\Delta, 0) + V(\Delta+1), C(\Delta, 1) + V(0)\}$, where σ^* is the average cost under the $g(\Delta)$–optimal policy, and $V(\Delta)$ is the optimal cost.

Next, we obtain an analytical expression for finding τ as a function of λ and the system parameters. We know from Theorem 1 that $a = 1$ is optimal at $\Delta = \tau$. Then, from the Bellman equation, we have $C(\tau, 1) + V(0) - \sigma^* < C(\tau, 0) + V(\Delta+1) - \sigma^*$, which yields

$$V(0) + \lambda < V(\tau + 1). \quad (8)$$

Further, at $\Delta = \tau - 1$, $a = 0$ is optimal. Thus, by the same argument, we have that $V(\tau) \leq \lambda + V(0)$, which on combining with (8) leads to $V(\tau) \leq \lambda + V(0) < V(\tau+1)$. Furthermore, by using $a = 1$ at $\Delta = \tau$, we get

$$V(\Delta) = g(\Delta, A, K_W) + \lambda + V(0) - \sigma^*. \quad (9)$$

Since $V(\Delta)$ is monotonically non-decreasing in Δ, $\exists\eta \in [0, 1)$ such that $V(\tau + \eta) = \lambda + V(0)$, and by using (9), $\sigma^* = g(\tau + \eta, A, K_W)$. Further, for $\Delta < \tau$, we have $V(\Delta + 1) - V(\Delta) = \sigma^* - g(\Delta, A, K_W)$, which on summing from $\Delta = 0$ to $\Delta = \tau - 1$ gives

$$V(\tau) = V(0) + \sigma^*\tau - \sum_{\Delta=0}^{\tau-1} g(\Delta, A, K_W). \quad (10)$$

Next, we have that $g(\Delta, A, K_W) = \sum_{r=1}^{\Delta} tr((A^{r-1})^\top A^{r-1}K_W)\Delta = \sum_{r=1}^{\Delta}\|A^{r-1}K_W^{0.5}\|_F^2\Delta$. Substituting this in (10), we can calculate the value of τ by using (9) and (10).

Now, we provide a simplified expression to compute τ for scalar systems ($n = 1$ in (1)). Equating the values of $V(\tau)$ from (9) and (10), we arrive at the equation $(\tau+1)g(\tau + \eta, A, K_W) - \sum_{l=0}^{\tau} g(l, A, K_W) = \lambda$. By substituting the expression for $g(\cdot, \cdot, \cdot)$ in this equation yields

$$\begin{cases} f_1(\tau, A, K_W, \lambda) = 0, & A \neq 1, \\ f_2(\tau, K_W, \lambda) = 0, & A = 1, \end{cases} \quad (11)$$

where the closed-form expressions for (11) are provided in [23, (23)] We note that (11) is an implicit equation in τ and

η for given values of γ, A and K_W. Thus, the value of τ can be calculated by plotting η vs τ, and choosing the integer value of τ for an admissible η.

B. Multi-agent Randomized Scheduling Policy

In the previous subsection, we showed the existence of a single-agent stationary deterministic policy for a fixed λ. In this subsection, we return to the multi-agent case and obtain the optimal value of λ. Consequently, we use the threshold characterization of the deterministic policy to obtain the optimal solution to Problem 2. Henceforth, we also resume the use of superscript i to denote the agent index.

Consider the threshold for the i^{th} agent given by $\tau^i(\lambda) := \tau^i(\lambda, A_i, K_{W^i})$. Then the agent is connected to its respective controller every $(\tau^i(\lambda)+1)$–th instant. Thus, under the constraint (5), we have that $W(\lambda) := \sum_{i=1}^{N} \frac{1}{\tau^i(\lambda)+1} \leq R_d$.

In order to find the optimal value of the Lagrange multiplier solving Problem 2, we use the Bisection search procedure [24], which we summarize next. Since $\lambda \geq 0$, we initialize two parameters $\lambda_l^{(0)} = 0$ and $\lambda_u^{(0)} = 1$. We then calculate the threshold parameters $\tau^i(\lambda_u^{(0)})$ for all i, by using the piece-wise definition in (11). Consequently, we iterate by setting $\lambda_l^{(j+1)} = \lambda_u^{(j)}$ and $\lambda_u^{(j+1)} = 2\lambda_u^{(j)}$ until the constraint $W(\lambda) \leq R_d$ is satisfied for $\lambda_u^{(r)}$, for some integer r. Then, we define the interval $[\lambda_l^{(r)}, \lambda_u^{(r)}]$. This interval contains the optimal value of the multiplier λ^*, which can be calculated using the *Bisection method*. The iteration stops when $|\lambda_u^{(m)} - \lambda_l^{(m)}| \leq \epsilon$, for the iterating index m and for a suitably chosen $\epsilon > 0$.

We next construct a stationary randomized policy solving Problem 2. Define $\lambda_l^* = \lambda_l^{(m)}$ and $\lambda_u^* = \lambda_u^{(m)}$ as obtained above. Further, let the stationary deterministic policies $\gamma_{D1}^{d,i}$ and $\gamma_{D2}^{d,i}$ as obtained from Theorem 1 be those corresponding to λ_l^* and λ_u^*, respectively, where $\lambda_l^* \mapsto \tau_l(\lambda_l^*) := \{\tau_l^1(\lambda_l^*), \cdots, \tau_l^N(\lambda_l^*)\}^\top$ and $\lambda_u^* \mapsto \tau_u(\lambda_u^*) := \{\tau_u^1(\lambda_u^*), \cdots, \tau_u^N(\lambda_u^*)\}^\top$. Also, we define R_d^u and R_d^l as the total bandwidth used corresponding to the multipliers λ_u^* and λ_l^*, respectively. We note that τ_l^i differs from τ_u^i by at most one state. Then, we define the probability p and the deterministic policies for all i as:

$$p := (R_d - R_d^u)/(R_d^l - R_d^u), \tag{12}$$

$$\zeta_{D1}^{d,i}(\Delta^i) := \begin{cases} 1, & \Delta^i \geq \tau_l^i(\lambda_l^*, A_i, K_{W^i}) \\ 0, & \Delta^i < \tau_l^i(\lambda_l^*, A_i, K_{W^i}) \end{cases}, \tag{13}$$

$$\zeta_{D2}^{d,i}(\Delta^i) := \begin{cases} 1, & \Delta^i \geq \tau_u^i(\lambda_u^*, A_i, K_{W^i}) \\ 0, & \Delta^i < \tau_u^i(\lambda_u^*, A_i, K_{W^i}) \end{cases}. \tag{14}$$

The randomized policy γ^d for the relaxed Problem 2 can then be obtained as:

$$\zeta^{d,i} = p\zeta_{D1}^{d,i} + (1-p)\zeta_{D2}^{d,i}, \ \forall i \in [N]. \tag{15}$$

Next, we have the following proposition which establishes optimality of the above randomized policy for Problem 2.

Proposition 1 (Optimality of Randomized Policy)**.** *Under Assumption 1, the policy* (12)-(15) *is optimal for Problem 2.*

Now, since the solution to the scheduling problem is complete, in the next section, we proceed to establish the ϵ–Nash property of the mean-field game solution.

V. N Agent Consensus Problem

In this section, we solve the consensus problem using the BS's policy. To this end, we first consider the limiting game called the mean-field game (MFG) as $N \to \infty$. Under this setting, the empirical coupling term in (2) is approximated by a known deterministic sequence (or the MF trajectory), whose closeness is justified by analysis. This principle is the well known Nash certainty equivalence principle [10] and reduces the game problem to a stochastic optimal control problem of a generic agent with a coupled consistency condition. The equilibrium solution obtained (called the mean-field equilibrium (MFE)) will be shown to constitute an ϵ–Nash approximation to the finite agent game.

A. Stochastic Optimal Tracking Control

Consider a generic agent of type θ from an infinite population with dynamics

$$X[k+1] = A(\theta)X[k] + B(\theta)U[k] + W[k], \ k \in \mathbb{Z}^+, \tag{16}$$

where $X[k] \in \mathbb{R}^n$ and $U[k] \in \mathbb{R}^m$ are the state vector and the control input, respectively. Further, the initial state $X[0]$ is assumed to have a symmetric density function with $\mathbb{E}[X[0]] = \nu_{\theta,0}$ and $cov(X[0]) = \Sigma_x$ is bounded. Next, $W_k \in \mathbb{R}^n$ is a zero-mean i.i.d. Gaussian noise with finite covariance $K_W(\theta)$. All covariance matrices are assumed to be positive-definite. The objective function of the generic agent is given by

$$J(\mu,\bar{X}) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1} \|X[k]-\bar{X}[k]\|^2_{Q(\theta)} + \|U[k]\|^2_{R(\theta)}\right] \tag{17}$$

where $\mu := (\mu[1], \mu[2], \cdots,) \in \mathcal{M}^{d,con}$ and is an admissible control policy of the generic agent. Further, the admissible set $\mathcal{M}^{d,con} := \{\mu \mid \mu \text{ is adapted to } \sigma(I^{d,con}[s], s = 0, 1, \cdots, k)\}$ is the space of *decentralized control* policies for the generic agent and $I^{d,con}[0] := \{Z[0]\}$, $I^{d,con}[k] := \{U_{0:k-1}, Z_{0:k}\}, k \geq 1$, is the local information history of the generic agent. Recall that this is in contrast to $I^{c,con}[k]$, which includes information of other agents as well. The information structure for the generic agent's decoder is defined similar to that in Subsection II-A, except with the superscript i removed. Further, $\bar{X} = (\bar{X}[1], \bar{X}[2], \cdots,)$ is the MF trajectory and denotes the infinite player approximation to the consensus term in (2) and serves to decouple the otherwise coupled game problem into a tracking (LQT) problem. Finally, the expectation above is taken with respect to the noise statistics and the initial state distribution.

To solve the LQT problem with dynamics in (16) and the cost in (17), we first state the following assumption.

Assumption 2. *(i) The pair $(A(\theta), B(\theta))$ is controllable and $(A(\theta), \sqrt{Q(\theta)})$ is observable.*

(ii) The MF trajectory belongs to the space of bounded functions, i.e., $\bar{X} \in \mathcal{X} := \{\bar{X}[k] \in \mathbb{R}^n \mid \|\bar{X}\|_\infty := \sup_{k\geq 0} \|\bar{X}[k]\| < \infty\}$.

We remark here that Assumption 2 is standard in the MF-LQG literature [12], [25]. Next, we define a MFE by introducing the following operators [13]: (1) $\Xi : \mathcal{X} \to \mathcal{M}^{d,con}$, defined as $\Xi(\bar{X}) = \text{argmin}_{\mu \in \mathcal{M}^{d,con}} J(\mu, \bar{X})$, gives the optimal control policy for a given MF trajectory, and (2) the consistency operator $\Lambda : \mathcal{M}^{d,con} \to \mathcal{X}$ that regenerates a MF trajectory for an admissible control policy $\mu \in \mathcal{M}^{d,con}$.

Definition 2 (Mean-Field Equilibrium (MFE) [2])**.** *The pair $(\mu^*, \bar{X}^*) \in \mathcal{M}^{d,con} \times \mathcal{X}$ is a MFE if, $\mu^* = \Xi(\bar{X}^*)$ and $\bar{X}^* = \Lambda(\mu^*)$. In other words, $\bar{X}^* = \Lambda \circ \Xi(\bar{X}^*)$.*

Now, the central scheduling policy is fixed from the previous section, similar to [2], we have the following result for the optimal tracking control of a generic agent.

Proposition 2. *Consider the generic agent* (16) *with controller state as in* (3) *and cost function in* (17)*. Then, under Assumptions 1-2, the following hold true:*

1) The optimal decentralized control action of a generic agent is given as

$$U^*[k] = -\Pi(\theta)Z[k] - L(\theta)r[k+1], \tag{18}$$

where $L(\theta) = (R(\theta) + B(\theta)^\top K(\theta)B(\theta))^{-1}B(\theta)^\top$, $\Pi(\theta) = L(\theta)K(\theta)A(\theta)$, and $K(\theta)$ is the unique positive definite solution to the algebraic Riccati equation $K(\theta) = A(\theta)^\top[K(\theta)A(\theta) - K(\theta)^\top B(\theta)\Pi(\theta)] + Q(\theta)$. Further, the trajectory $r[k]$ satisfies the backward dynamics $r[k] = H(\theta)^\top r[k+1] - Q(\theta)\bar{X}[k]$, with the initial condition $r[0] = -\sum_{j=0}^{\infty} (H(\theta)^j)^\top Q(\theta)\bar{X}[j]$ and $H(\theta) = A(\theta) - B(\theta)\Pi(\theta)$ is Hurwitz.

2) The difference equation for $r[k]$ has a unique solution in $\mathcal{X}$, given as $r[k] = -\sum_{j=k}^{\infty} (H(\theta)^{j-k})^\top Q(\theta)\bar{X}[j]$.

3) The optimal cost $J(\mu^*, \bar{X})$ is bounded from above by a function of system parameters.

With the above result, we proceed with the analysis of the existence and uniqueness of the MFE in the next subsection.

B. Mean-Field Analysis

We now establish existence and uniqueness of the MFE by constructing the MF operator as follows. Consider the closed-loop system in (3) with the control policy in (18):

$$Z[k+1] = \begin{cases} H(\theta)Z[k] - B(\theta)L(\theta)r[k+1] + W[k], & \gamma^d[k+1] = 1, \\ H(\theta)Z[k] - B(\theta)L(\theta)r[k+1], & \gamma^d[k+1] = 0, \end{cases}$$

where $\gamma^d[k]$ is chosen from the policy ζ^d of Section IV-B. The above can be rewritten as

$$X[k+1] = H(\theta)X[k] - B(\theta)L(\theta)r[k+1] + B(\theta)\Pi(\theta)e[k] + W[k],$$

where $e[k]$ is defined in (4). By taking expectation on both sides and substituting $r[k]$ from Proposition 2, we get $\hat{X}_\theta[k] = H(\theta)^k \nu_{\theta,0} + \sum_{j=0}^{k-1} H(\theta)^{k-j-1} B(\theta)L(\theta) \sum_{s=j+1}^{\infty} (H(\theta)^{s-j-1})^\top Q(\theta)\bar{X}[s]$, where $\hat{X}_\theta[k] = \mathbb{E}[X[k]]$ is the aggregate dynamics across agents of type θ and we use the fact that $\mathbb{E}[e[k]] = 0$. Next, using the empirical distribution from Section II, we define the MF operator as $\mathcal{T}(\bar{X})[k] := \sum_{\theta \in \Theta} \hat{X}_\theta[k]\mathbb{P}(\theta)$, and state the following assumption before we present the main result.

Assumption 3. *We assume $\check{H}(\theta) := \|H(\theta)\| + \upsilon < 1, \forall \theta \in \Theta$, where $\upsilon = \sum_{\theta \in \Theta} \frac{\|Q(\theta)\| \|B(\theta)L(\theta)\|}{(1 - \|H(\theta)\|)^2} \mathbb{P}(\theta)$.*

It is common in the literature [10], [13] to invoke the above assumption; however, it is stronger than the corresponding assumption in [12] and leads to the linearity property of the MF trajectory, which can then be easily computed offline.

Theorem 2. *Under Assumptions 1-3, the following statements hold true: (1) The operator $\mathcal{T}(\bar{X}) \in \mathcal{X}, \quad \forall \bar{X} \in \mathcal{X}$. Furthermore, there exists a unique $\bar{X}^* \in \mathcal{X}$ such that $\mathcal{T}(\bar{X}^*) = \bar{X}^*$, and (2) $\bar{X}^*[k]$ follows linear dynamics, i.e., $\exists\, \mathscr{E}^* \in \mathcal{E} := \{\mathscr{E} \in \mathbb{R}^{n \times n} : \|\mathscr{E}\| < 1, \bar{X}^*[k+1] = \mathscr{E}\bar{X}^*[k]\}$, where $\bar{X}^*[k]$ is the equilibrium MF trajectory of the agents, and $\bar{X}^*[0] = \sum_{\theta \in \Theta} \nu_{\theta,0}\mathbb{P}(\theta)$.*

We note that while Theorem 2 (1) gives us a unique MFE, the linearity of the MF trajectory in (2) makes the computation of this equilibrium trajectory tractable, which would otherwise have involved a non-causal infinite sum. Further, as a result of the linearity, we arrive at an explicit convergence rate between the equilibrium trajectory and the coupling term in (2) in Lemma 3.

C. ϵ–Nash Analysis

In this subsection, we show that the decentralized equilibrium control policy obtained from the MF analysis is approximately Nash for the finite-agent system. We start by stating the following lemmas, which show that the closed-loop system is stable under the MFE solution, and that the equilibrium MF trajectory approximates the finite-agent state average in the mean-square sense.

Lemma 2. *Suppose that Assumptions 1-3 hold. Then, the closed-loop system* (1) *is uniformly mean-square stable under the decentralized equilibrium control policy* (18)*, i.e., $\sup_{N \geq 1} \max_{i \in [N]} \limsup_{T \to \infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1} \left\|X^{*,i}[k]\right\|^2\right] < \infty$.*

Lemma 3. *Under Assumptions 1-3, the equilibrium MF trajectory converges (in the mean-square sense) to the coupling term in* (2) *with a rate of $\mathcal{O}\big(\frac{1}{\min_{\theta \in \Theta} |N_\theta|}\big)$, where $N_\theta \subset [N]$ is the number of agents of type θ.*

We next define ϵ-Nash equilibrium as follows.

Definition 3. *[2] The set of control policies $\{\mu^{*,i}, 1 \leq i \leq N\}$ constitute an ϵ–Nash equilibrium with respect to the cost functions $\{J_i^N, 1 \leq i \leq N\}$, if there exists $\epsilon > 0$, such that*

$$J_i^N(\mu^{*,i}, \mu^{*,-i}) \leq \inf_{\pi^i \in \mathcal{M}_i^{c,con}} J_i^N(\pi^i, \mu^{*,-i}) + \epsilon, \quad \forall i \in [N].$$

Then, we have the following theorem stating that the control laws prescribed by the MFE constitute an ϵ-Nash equilibrium in the finite population case.

Theorem 3. *Suppose Assumptions 1-3 hold. Then the set of decentralized control policies $\{\mu^{*,i}, 1 \leq i \leq N\}$, constitute an ϵ–Nash equilibrium for the N–agent LQ-mean field game with bandwidth limits. In particular, we have that*

$$J_i^N(\mu^{*,i}, \mu^{*,-i}) \leq \inf_{\pi^i \in \mathcal{M}_i^{c,con}} J_i^N(\pi^i, \mu^{*,-i}) + \mathcal{O}\left(1/\sqrt{\min_{\theta \in \Theta} |N_\theta|}\right).$$

Before concluding, we finally note that since Theorem 3 establishes that the decentralized equilibrium policy provides an ϵ–Nash equilibrium for the centralized policy structure in the N-agent game, it does so for the decentralized policy structure also, as formulated in Section II-A.

VI. Simulations

In this section, we provide an empirical analysis of the theoretical results. We first demonstrate the performance of the scheduling policy $\gamma_R^{d,*}$. We simulate the behavior of WAoI under $\gamma_R^{d,*}$ with 4 different values of $R_d = 0.6N$, $0.7N, 0.8N, 0.9N$. We plot the average WAoI in Fig. 2(a), which shows that the WAoI increases as the number of agents being transmitted over the downlink decreases. Next, we empirically evaluate the behavior of $N = 800$ agents under the MFE policy (18). To this end, we plot the average cost of an agent as a function of R_d in Fig. 2(b). The figure shows a box plot depicting the median (red line) and spread (box) of the average cost per agent over 100 runs for each value of α. From the figure, we see that the average cost decreases as the available bandwidth increases, aligned with intuition.

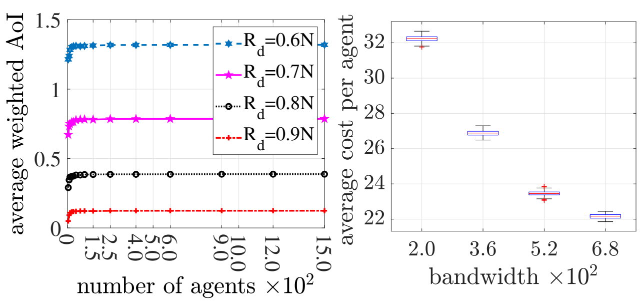

Fig. 2: (a) shows that the average WAoI for decreases as R_d increases, and (b) shows that average cost per agent decreases as R_d increases.

VII. Conclusion

In this paper, we have studied an $N + 1$ player game problem in which N agents aim to achieve consensus while the BS schedules information over a rate-constrained network to these agents using a WAoI metric. To solve the scheduling problem, we have constructed a stationary randomized optimal scheduling policy for it by decoupling the problem and following an MDP approach. Next, we have solved the game problem between the N–agents using the mean-field game approach and employing the obtained scheduling policy. By considering a limiting system as $N \to \infty$, we have first proved the existence of a unique mean-field equilibrium and then have shown the ϵ–Nash property of the equilibrium solution for the finite-agent system. Finally, we have validated the theoretical results with simulations.

References

[1] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Contrl Sys. Mag.*, vol. 35, no. 1, pp. 20–23, 2015.

[2] S. Aggarwal, M. A. uz Zaman, and T. Başar, "Linear quadratic mean-field games with communication constraints," in *IEEE ACC*, June 2022, pp. 1323–1329.

[3] A. A. Feldbaum, "Dual control theory," in *Control Theory: Twenty-Five Seminal Papers (1932-1981)*, T. Başar, Ed. Wiley-IEEE Press, 2001, ch. 10, pp. 874–880.

[4] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, 1971.

[5] A. A. Malikopoulos, "On team decision problems with nonclassical information structures," *IEEE Transactions on Automatic Control*, 2022.

[6] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.

[7] C. Ramesh, H. Sandberg, and K. H. Johansson, "Design of state-based schedulers for a network of control loops," *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 1962–1975, 2013.

[8] A. Molin and S. Hirche, "On the optimality of certainty equivalence for event-triggered control systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 470–474, 2012.

[9] O. C. Imer and T. Başar, "Optimal estimation with limited measurements," *Internationa Journal of Systems, Control and Communications*, vol. 2, no. 1-3, pp. 5–29, 2010.

[10] 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.

[11] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[12] J. Moon and T. Başar, "Discrete-time LQG mean field games with unreliable communication," in *IEEE CDC*, December 2014, pp. 2697–2702.

[13] M. A. uz Zaman, K. Zhang, E. Miehling, and T. Başar, "Reinforcement learning in non-stationary discrete-time linear-quadratic mean-field games," in *IEEE CDC*, December 2020, pp. 2278–2284.

[14] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Age-optimal sampling and transmission scheduling in multi-source systems," in *ACM Mobihoc*, July 2019, pp. 121–130.

[15] J. Zhong, W. Zhang, R. D. Yates, A. Garnaev, and Y. Zhang, "Age-aware scheduling for asynchronous arriving jobs in edge applications," in *IEEE INFOCOM*, May 2019, pp. 674–679.

[16] H. Tang, J. Wang, Z. Tang, and J. Song, "Scheduling to minimize age of synchronization in wireless broadcast networks with random updates," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4023–4037, 2020.

[17] M. Hatami, M. Leinonen, Z. Chen, N. Pappas, and M. Codreanu, "On-demand AoI minimization in resource-constrained cache-enabled IoT networks with energy harvesting sensors," *Available on arXiv:2201.12277*, 2022.

[18] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-information vs. value-of-information scheduling for cellular networked control systems," in *ACM/IEEE ICCPS*, April 2019, pp. 109–117.

[19] O. Ayan, M. Vilgelm, and W. Kellerer, "Optimal scheduling for discounted age penalty minimization in multi-loop networked control," in *IEEE CCNC*, January 2020, pp. 1–7.

[20] O. Ayan, A. Ephremides, and W. Kellerer, "Age of information: An indirect way to improve control system performance," in *IEEE INFOCOM*, May 2021, pp. 1–7.

[21] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[22] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2215–2228, 2020.

[23] S. Aggarwal, M. A. uz Zaman, M. Bastopcu, and T. Başar, "Weighted age of information based scheduling for large population games on networks," *Available on arXiv:2209.12888*, September 2022.

[24] Y. Chen and A. Ephremides, "Minimizing age of incorrect information for unreliable channel with power constraint," in *IEEE GLOBECOM*, December 2021, pp. 1–6.

[25] M. A. u. Zaman, E. Miehling, and T. Başar, "Reinforcement learning for non-stationary discrete-time linear–quadratic mean-field games in multiple populations," *Dynamic Games and Applications*, vol. 13, pp. 118–164, 2023.