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

Large Population Games on Constrained Unreliable Networks

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

Abstract— **This paper studies an N–agent cost-coupled game where the agents are connected via an *unreliable capacity constrained* network. Each agent receives state information over that network which loses packets with probability p. A Base station (BS) actively schedules agent communications over the network by minimizing a weighted Age of Information (WAoI) based cost function under a capacity limit $\mathcal{C} < N$ on the number of transmission attempts at each instant. Under a standard information structure, we show that the problem can be decoupled into a *scheduling problem* for the BS and a *game problem* for the N agents. Since the scheduling problem is an NP hard combinatorics problem, we propose an approximately optimal solution which approaches the optimal solution as $N \to \infty$. In the process, we also provide some insights on the case without channel erasure. Next, to solve the large population game problem, we use the mean-field game framework to compute an approximate decentralized Nash equilibrium. Finally, we validate the theoretical results using a numerical example.**

I. Introduction

With the phenomenal expansion in data-traffic galvanized by the growing number of connected devices, Internet-of-Things (IoT) finds applications in diverse areas such as smart grids, autonomous vehicles, and monitoring systems [1], to name a few. A commonality among all of the above is the presence of distributed sensing and actuating devices communicating via a wireless network. While distributed systems can efficiently handle the growing network size compared to their centralized counterparts, they come with added challenges, such as limited channel capacities, network unreliability, and scalability concerns. These constraints might cause end-to-end latency or in a worse case, missing information at the end-user, which can lead to compromised reliability in safety-critical applications. Thus, there is an urgent need for the development of dependable and timeliness-aware communication technologies with the potential to mitigate the above posed concerns. In this work, we aim to propose strategies to mitigate the deleterious effects of unreliable capacity-constrained communication in networks involving a large number of decision-making agents.

Specifically, we consider a large population setting where N rational agents aim to form consensus while communicating intermittently over a network. This intermittency is caused by i) a capacity-constrained downlink connecting BS to the decoders, and ii) the possibility of *erasure* amidst transmission, after information is relayed by the BS. This results in an unreliable capacity-constrained network. As a result, the agents must maintain an estimate of their state to consequently compute control actions that can achieve consensus. Meanwhile, the BS, which is tasked with the scheduling of information, must carefully design policies to account for the heterogeneity in agent dynamics whilst also dealing with the possibility of erasure of the scheduled information. We formulate the BS's problem by proposing a Weighted Age of Information (WAoI) based cost function which is monotonically increasing in the average estimation error of the agents, thereby extending the setting of our earlier work [2] to erasure channels. Further, we improve upon the convergence guarantees in [2] for the case where the network is erasure free by proposing a novel scheduling policy. Finally, we employ this policy to construct an approximate Nash solution for the finite-agent consensus problem.

In literature, the early works [3], [4] have dealt with an optimal control problem with unreliable communication, albeit, for a single agent system and an unconstrained network. The work [5] extends the setting to multi-agent games; however, the considered network is unconstrained. In order to measure timeliness in communication networks, age of information (AoI) has been introduced as a potential metric. In the context of networked feedback systems, the AoI-based policies have been proposed for solving resource allocation and end-user uncertainty reduction problems as in [6]. Recently, age of incorrect information (AoII) has been proposed for solving multi-agent remote state estimation problems [7]. Age-optimal scheduling policies have been considered with Markovian error-prone channel state in [8], [9], with unknown erasure probabilities in [10], and over erroneous broadcast channels in [11]. A more detailed literature review on age-optimal scheduling policies can be found in [12].

To appropriately handle the concerns of increasing network interactions, one of the most relevant framework is that of mean-field games (MFGs) [13], [14]. It leads one to circumvent the issues posed by scalability, by allowing for a representative agent to play against the population, although, at the cost of entailing an *approximate* equilibrium solution for the finite-agent consensus problem. It has been well-studied in the regime of linear-quadratic systems [15], [16] and holds great potential to solve problems involving ultradense networks or massive machine-type communication [17]. For additional literature on large multi-agent systems with networked communication, we refer the reader to [2].

We list below the main contributions of this paper. We extend the setting of our previous works [2], [18] to the case of unreliable downlink communication. Since the scheduling problem belongs to the class of restless multi-armed bandits,

Research of the authors was supported in part by the ARO MURI Grant AG285 and in part by the AFOSR Grant FA9550-19-1-0353.

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.

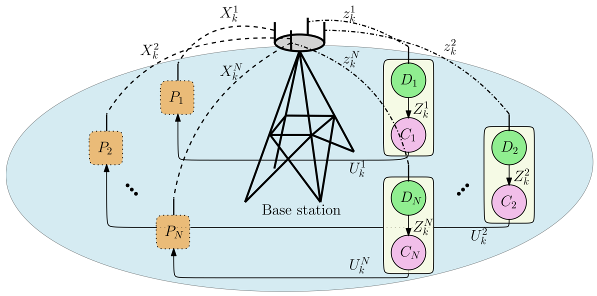

Fig. 1: A prototypical networked control system constituting a BS and N game playing agents. The BS, decoders, and controllers are active decision makers. Dashed lines denote an erasure-free wireless transfer, dotted-dashed lines denote erasure-prone one, and bold lines denote wired information transfer.

for which an optimal policy is hard to compute, we propose a novel suboptimal maximum age-based tie-breaking protocol (MATB-P) to solve the capacity-constrained scheduling problem of the BS (which is also different from the uniform sampling-based policy considered in [2]). We prove that this policy approaches optimality (exponentially fast) as N grows large, in contrast to the $\mathcal{O}(N^{-0.5})$ rate proposed in [2]. Further, we also provide high-probability guarantees on the tail of the AoI, which, in turn, provides guarantees on the freshness of information under high traffic. Additionally, in the special case with no channel erasure, we relax the assumption on the A matrix in the work [2] by proving a uniform upper bound on the AoI of all agents under MATB-P. Finally, using the policy constructed above, we solve the N–agent consensus problem by leveraging the MFG paradigm and getting ϵ–Nash policies for the agents where $\epsilon \xrightarrow{N\to\infty} 0$.

The rest of the paper is organized as follows. We formulate the $(N+1)$–player game problem in Sec. II. In Sec.III, we solve the BS-level scheduling problem, and provide its analysis in Sec. IV. Then, we solve the agent-level game problem in Sec. V, and provide a numerical example in Sec. VI. The paper is concluded in Sec. VII with major highlights. The proofs of Proposition 2 and Theorems 2 and 3 are omitted, and can be found in [19].

Notations: We let $[N] := \{1, 2, \cdots, N\}$ and $tr(\cdot)$ denote the trace of its argument matrix. The Euclidean 2-norm and the Frobenius norm are denoted by $\|\cdot\|$ and $\|\cdot\|_F$, respectively. All the empty summations are set to 0. For a vector x and a positive semi-definite matrix Q, $\|x\|_Q^2 := x^\top Qx$. We define the limit superior of a real sequence as $\overline{\lim} := \limsup$. Finally, $\mathbf{1}_A$ denotes the indicator function of the argument.

II. Problem Formulation

In this section, we set up the two sub-problems in the $(N+1)$–player game, namely, a) the agent-level game problem, and b) the BS-level scheduling problem.

Consider a multi-agent system consisting of N cost-coupled agents receiving information over an unreliable network. Each agent i constitutes a plant, a decoder and a controller, labeled as a tuple (P_i, D_i, C_i) as shown in Fig. 1. Dynamics of P_i evolve in discrete-time as

$$X_{k+1}^i = A(\phi_i)X_k^i + B(\phi_i)U_k^i + W_k^i,\ k \geq 0, \tag{1}$$

where $X_k^i \in \mathbb{R}^n$ is the state and $U_k^i \in \mathbb{R}^m$ is the control input, both for agent i. The exogeneous noise $W_k^i \in \mathbb{R}^n$ is zero mean with covariance $C_W(\theta_i) > 0$. The initial state X_0^i of agent i is assumed to have symmetric density with mean $x_{\phi_i,0}$ and covariance $\Sigma_x > 0$. Further, it is assumed to be independent of the noise process for all timesteps k. The system matrices $A(\phi_i)$, $B(\phi_i)$ are time-invariant with suitable dimensions. Further, they are chosen according to an empirical function $\mathbb{P}^N(\phi = \phi_i)$, where $\phi \in \Phi$ denotes the type of an agent chosen from a finite set $\Phi := \{\phi_1, \cdots, \phi_p\}$. We assume that $|\mathbb{P}^N(\phi) - \mathbb{P}(\phi)| = \mathcal{O}(1/N)$, $\forall \phi$, where $\mathbb{P}(\phi)$ denotes the limiting distribution.

The state of plant P_i is relayed to the decoder D_i via an ideal uplink to the BS, which then regulates agent communications over the downlink. The downlink is constrained by a capacity limit of $\mathcal{C} < N$ units on the number of transmissions and serves as a bottleneck from the plant to the decoder. Further, it is unreliable in the sense that a packet communicated over it may be lost according to a Bernoulli distributed signal $\beta_k \sim \text{Ber}(p)$, with p being the erasure probability. The decoder receives the information signal:

$$z_k^i := X_k^i \mathbf{1}_{E_k^i} + \emptyset \mathbf{1}_{(E_k^i)^c}, \tag{2}$$

where the event E_k^i denotes that state information is *successfully* transmitted. Further, the event $\left(E_k^i\right)^c$ denotes no transmission (or $\emptyset$), which can be either due to no transmission by the BS or a packet drop over the channel. Let us denote the instants of information reception by the decoder as $\wp_k^i := \zeta_k^i \beta_k$. Then, the information history of the decoder is defined as $I_k^{D_i} := \{z_{0:k}^i, \wp_{0:k}^i, U_{0:k-1}^i\}$, based on which it computes the minimum mean-squared (MMS) estimate $(\mathbb{E}[X_k^i \mid I_k^{D_i}])$ of the state X_k^i. We adopt the convention that $z_{-1}^i = Z_{-1}^i = U_{-1}^i = 0$, and $W_{-1}^i = X_0^i - Z_0^i$, for all i.

Next, each controller C_i receives the estimate from D_i and aims to minimize the average cost function

$$J_i(\pi_c) := \overline{\lim_{T\to\infty}} \frac{1}{T} \mathbb{E}\left[\sum_{k=0}^{T-1} \|X_k^i - \mu_k^N\|_{Q(\phi_i)} + \|U_k^i\|_{R(\phi_i)}\right], \tag{3}$$

where $Q(\phi_i) \geq 0$, $R(\phi_i) > 0$, and $\mu_k^N := \frac{1}{N}\sum_{j=1}^N X_k^j$ represents the coupling between agents. Due to this coupling, the cost J_i depends on the strategy $\pi_c := \{\pi_c^1, \cdots, \pi_c^N\}$ of the entire population. Further, $\pi_c^i \in \Pi_i := \{\pi_c^i \mid \pi_c^i \text{ is adapted to } \sigma(I_s^{C_i}, s = 0, \cdots, k)\}$, $\forall i$, where $I_k^{C_i} := \{U_{0:k-1}^j, Z_{0:k}^j\}_{j\in[N]}$ denotes the information history of C_i, and $\sigma(\cdot)$ denotes the sigma-algebra generated by its argument. We assume that the pair $(A(\phi_i), B(\phi_i))$ is controllable and the pair $(A(\phi_i), \sqrt{Q(\phi_i)})$ is observable [20]. Due to the difficulty in computing Nash equilibrium for the game (1)-(3), we will resort to the MFG framework (later in Section V) to compute decentralized ϵ-Nash policies where only local information will be required for decision-making and $\epsilon \to 0$ as $N \to \infty$. Next, we describe the BS-level problem, where the goal is to find an optimal scheduling policy of the BS.

The aim of the BS is to *efficiently* transmit information over the downlink. To this end, consider the most recent timestep when information was received by the i^{th} controller, which is defined as $\ell_k^i := \sup_{\ell \leq k}\{\ell \geq 0 \mid z_\ell^i \neq \emptyset\}$. Then, the AoI at the controller, which is the time elapsed since the generation of the most recent packet at the plant, is defined as $\tau_k^i := k - \ell_k^i$. Further, its evolution is given as $\tau_{k+1}^i = (\tau_k^i + 1)\mathbf{1}_{\{\wp_k^i = 0\}}$, i.e., the AoI drops to zero only when a transmission is attempted by the BS and the packet is not dropped by the network. With the above AoI evolution, we formally define the capacity-constrained scheduling problem at the BS as follows:

Problem 1.

$$\inf_{\gamma\in\Gamma} J^{BS}(\gamma) := \overline{\lim}_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{T-1}\sum_{i=1}^{N} w_k^i\tau_k^i\right], \text{s.t.} \sum_{i=1}^{N}\zeta_k^i \le \mathcal{C}$$

$\forall k$ where $\gamma := \{\gamma^1, \cdots, \gamma^N\}$ and $\Gamma := \{\gamma \mid \gamma$ is adapted to $\sigma(I_s^{BS}), s = 0, \cdots, k\}$ is the space of admissible scheduling policies with $I_k^{BS} := \{\tau_{0:k}^i, \zeta_{0:k-1}^i, \wp_{0:k-1}^i, Z_{0:k-1}^i\}_{i\in[N]}$ being the information history of the BS. Moreover, $w_k^i := \mathbb{E}[\|e_k^i\|^2]$ denote the importance weights associated to each agent and are functions of the estimation error $e_k^i := X_k^i - Z_k^i$ at the controller. Finally, the expectation is taken over the probabilistic scheduling due to the erasure-prone downlink and (possible) randomization in the scheduling policy.

We note here that the information history of the BS includes the information reception instants of the agent decoders. This can be easily facilitated by a TCP-like protocol [3], where the decoder sends a one-bit ACK/NACK information to acknowledge whether or not the transmitted information was received by it. Further, Problem 1 involves a hard-limit on the number of transmissions, which makes it a combinatorics problem. It belongs to the class of restless multi-armed bandit problems, computing an optimal policy for which is quite difficult. Thus, in the sequel, we first reformulate the problem using the AoIs of each agent and then solve a relaxed problem involving a time-averaged constraint. The solution to the latter problem will then lead to a sub-optimal policy for Problem 1, which we will finally show to approach the optimal policy as N increases.

To this end, we start by defining the shorthands $A_i := A(\phi_i)$, $B_i := B(\phi_i)$ and $C_{W^i} := C_W(\phi_i)$. Then, we construct the decoder's MMS estimate as

$$Z_k^i = X_k^i\mathbf{1}_{[\wp_k^i=1]} + \mathbb{E}[X_k^i \mid I_k^{D_i}]\mathbf{1}_{[\wp_k^i=0]}, \tag{4}$$

which upon using (1), yields

$$Z_k^i = X_k^i\mathbf{1}_{[\wp_k^i=1]} + (A_iZ_{k-1}^i + B_iU_{k-1}^i + \mathbb{E}_c[W_{k-1}^i])\mathbf{1}_{[\wp_k^i=0]},$$

where $\mathbb{E}_c[\cdot] := \mathbb{E}[\cdot \mid \zeta_k^i = 0]$. Then, using similar arguments as in [2], we can show that the term $\mathbb{E}_c[W_{k-1}^i] = 0$ under the assumption of symmetric densities of X_0^i and W_0^i. Hence, the estimate at the decoder can be easily computed as:

$$Z_k^i = X_k^i\mathbf{1}_{[\wp_k^i=1]} + (A_iZ_{k-1}^i + B_iU_{k-1}^i)\mathbf{1}_{[\wp_k^i=0]}. \tag{5}$$

With the above estimate, we can re-express the term w_k^i in Problem 1 using Lemma 1 from [2] as:

$$w_k^i := w_k^i(\tau_k^i, A_i, C_{W^i}) = \sum_{\ell=1}^{\tau_k^i} tr\left(A_i^{\ell-1\top}A_i^{\ell-1}C_{W^i}\right). \tag{6}$$

We note that the capacity constraint in Problem 1 makes the optimal policy difficult to compute. Thus, by defining the running cost $c(\tau_k^i, A_i, C_{W^i}) := w_k^i\tau_k^i$, we relax the problem to one with an average constraint:

Problem 2.

$$\inf_{\gamma\in\Gamma} J^{BS}(\gamma) := \overline{\lim}_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{T-1}\sum_{i=1}^{N} c(\tau_k^i, A_i, C_{W^i})\right]$$

$$\text{s.t. } \overline{\lim}_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1}\sum_{i=1}^{N}\zeta_k^i\right] \le \mathcal{C}.$$

We note that the constraint in the the above problem entails that the capacity constraint may be violated at any given timestep as long as it is satisfied in the long run. This is clearly a weaker constraint than the one in Problem 1 since the latter requires the capacity constraint to be satisfied at all timesteps k. Hence, it is indeed a relaxation of Problem 1. The objective now is to compute an optimal solution to Problem 2 and then utilize the solution to come up with an asymptotically optimal solution to Problem 1. To this end, we start by constructing the Lagrangian of Problem 2 (with $\lambda \ge 0$ the Lagrange multiplier):

$$\mathscr{L}(\gamma,\lambda) = \overline{\lim_{T\to\infty}}\frac{1}{T}\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{T-1}\sum_{i=1}^{N} c(\tau_k^i, A_i, C_{W^i}) + \lambda\left(\zeta_k^i - \frac{\mathcal{C}}{N}\right)\right],$$

where λ can be thought of as a price on the downlink utilization. Thus, given a fixed λ, we decouple the N–agent scheduling problem into N decoupled single-agent problems:

Problem 3. *For all $i \in [N]$,*

$$\inf_{\gamma^i\in\Gamma^i} V^i(\gamma) := \overline{\lim}_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{k=0}^{T-1} c(\tau_k^i, A_i, C_{W^i}) + \lambda\zeta_k^i\right].$$

In the next section, we will first solve Problem 3, for which we will cast the evolution of the AoI in an MDP framework and then construct a suboptimal policy for Problem 1. Additionally, we will also suppress the superscript i.

III. Solution to the BS-level Problem

We compute an optimal policy for Problem 3 by first defining it as a discrete-time MDP $\mathrm{M} := (\mathrm{S}, \mathrm{A}, \mathrm{P}, \mathrm{C})$. The state space S is the space of non-negative integers. The action set $\mathrm{A} = \{0, 1\}$. An action $a = 0$ denotes that a transmission is not attempted while $a = 1$ denotes that it is. The probability transition function P describes the evolution of the AoI, i.e., $\mathrm{P}(\tau_{k+1} = 0 \mid \tau_k) = a_k(1-p)$ and $\mathrm{P}(\tau_{k+1} = \tau_k + 1 \mid \tau_k) = 1 - a_k + a_kp$. Finally, with the per stage cost defined to be $C(\tau, a) := c(\cdot) + \lambda a$, the MDP objective is to infimize the function $V(\gamma) := \overline{\lim}_{T\to\infty}\frac{1}{T}\mathbb{E}\big[\sum_{k=0}^{T-1}\mathrm{C}(\tau_k, a_k)\big]$ for which we compute an optimal policy next.

A. Solution to Problem 2

We start by stating the following theorem, which characterizes an optimal policy solving Problem 3.

Theorem 1. *Given $\lambda \ge 0$, there exists a stationary policy γ_s solving the above MDP with an optimal cost of σ^*, which is independent of τ. Moreover, the optimal policy is given as $a := \mathbf{1}_{[\tau\ge\kappa]}$, for integer $\kappa := \kappa(A, C_W, \lambda)$.*

The proof follows in a similar manner as the proof of [2, Theorem 1]. The theorem says that the optimal policy for Problem 3 is a threshold policy. Next, we compute the threshold parameter κ by invoking the following condition, which links the erasure probability p and the instability in the agent's dynamics.

Assumption 1. *We have that $\|A\|_F^2 p < 1$.*

Notice that the above assumption is standard in the literature on unreliable communication [3]–[5] and formalizes the fact that a higher erasure probability restricts our ability to stabilize highly unstable agents. In the extreme case when no communication is possible (i.e., $p = 1$) it requires that all agents must be stable. Because of space limitations, we refer the reader to the corresponding arXiv paper [19, Appendix I] for detailed computation of κ. Now, with the deterministic single-agent policy as provided above, we proceed toward constructing an optimal policy for Problem 2 which, as we will see, will be a randomized policy since the optimal policy for such a constrained optimization problem may not, in general, lie in the class of stationary deterministic policies [21]. Henceforth, we resume the use of superscript i to denote the i^{th} agent.

We start by computing an optimal value of λ. To this end, consider the threshold parameter $\kappa^i(\lambda) := \kappa^i(A_i, C_{W^i}, \lambda)$ as in Theorem 1. Then, the expected return time of agent i starting from $\tau_k^i = 0$ can be found by $\mathrm{R}_0^i = [\sum_{r=0}^{\infty}(\kappa^i + r + 1)(1-p)^{r+1}p^r]^{-1}$ and is equal to:

$$\mathrm{R}_0^i = \frac{((1-p)p-1)^2}{(1-p)((p-1)p(\kappa^i+1)+(1-p)p+\kappa^i+1)}. \quad (7)$$

Then, under the average constraint in Problem 2, we have $R(\lambda) := \sum_{i=1}^{N} \mathrm{R}_0^i \leq \mathcal{C}$. Consequently, we can use the iterative Bisection search algorithm, as given in [2], [7], starting with the initial parameters $\underline{\lambda}^{(0)} = 0$, and $\overline{\lambda}^{(0)} = 1$. The algorithm terminates when $|\overline{\lambda}^{(m)} - \underline{\lambda}^{(m)}| \leq \epsilon$, for an iterating index m and a suitably chosen $\epsilon > 0$. Next, let us define $\underline{\lambda}^* = \underline{\lambda}^{(m)}$ and $\overline{\lambda}^* = \overline{\lambda}^{(m)}$ as obtained above, and the corresponding deterministic policies as $\gamma_{s_1}^i$ and $\gamma_{s_2}^i$, which are obtained from Theorem 1. More precisely, we have that $\underline{\lambda}^* \mapsto \underline{\kappa}(\underline{\lambda}^*) := \{\underline{\kappa}^1(\underline{\lambda}^*), \cdots, \underline{\kappa}^N(\underline{\lambda}^*)\}^\top$ and $\overline{\lambda}^* \mapsto \overline{\kappa}(\overline{\lambda}^*) := \{\overline{\kappa}^1(\overline{\lambda}^*), \cdots, \overline{\kappa}^N(\overline{\lambda}^*)\}^\top$. Also, let $\overline{\mathcal{C}}$ and $\underline{\mathcal{C}}$ be the total capacities used corresponding to the multipliers $\overline{\lambda}^*$ and $\underline{\lambda}^*$, respectively. Then, we define the deterministic policies:

$$\gamma_{s_1}^i(\tau^i) := \mathbf{1}_{[\tau^i \geq \underline{\kappa}^i(\cdot,\cdot,\underline{\lambda}^*)]}, \;\; \gamma_{s_2}^i(\tau^i) := \mathbf{1}_{[\tau^i \geq \overline{\kappa}^i(\cdot,\cdot,\overline{\lambda}^*)]}, \quad (8)$$

for all i, using which we can construct a randomized policy $\gamma_R := [\gamma_R^1, \cdots, \gamma_R^N]^\top$ for the relaxed Problem 2 as:

$$\gamma_R^i = q\gamma_{s_1}^i + (1-q)\gamma_{s_2}^i, \; \forall i, \quad (9)$$

where $q = (\mathcal{C} - \overline{\mathcal{C}})/(\underline{\mathcal{C}} - \overline{\mathcal{C}})$ is the probability of randomization. Next, in the following proposition, we state that the randomized policy obtained is indeed optimal for Problem 2.

Proposition 1. *[2] Under Assumption 1, the policy* (8)-(9) *is optimal for the relaxed minimization Problem 2.*

With the solution to Problem 2, in the next subsection, we propose a novel asymptotically optimal policy for Problem 1.

B. Solution to Problem 1

In this subsection, we provide a sub-optimal solution to Problem 1, using the solution to Problem 2, which is shown to be asymptotically optimal as $N \to \infty$. We refer to this policy as the maximum-age-first tie-breaking protocol (or MATB-P for short). Consider the solution γ_R^i to Problem 2 as computed in the previous subsection and let $a_k^i = \gamma_R^i(I_k^{BS})$ be the scheduling action at timestep k. Define $\Lambda_k := \{j \in [N] \mid a_k^j = 1\}$ as the set of agents scheduled to be transmitted at instant k and its cardinality to be n_k^λ. Then, the scheduling decision ζ_k^i under MATB-P (γ^i) is given as:

- If $n_k^\lambda \leq \mathcal{C}$, then $\zeta_k^i = a_k^i$
- If $n_k^\lambda > \mathcal{C}$, then $\zeta_k^i = 1$ for a subset $\Lambda_k^{max} \subset \Lambda_k$ of the agents, where the cardinality of Λ_k^{max} is $\mathcal{C}$ for all k, and it constitutes the agents with the maximum values of τ_k. The agents in the set $\Lambda_k \setminus \Lambda_k^{max}$ remain unselected.

In the next section, we provide a tail-bound analysis of the constructed MATB policy, first, for the special case with no channel erasure, and then, for the general case.

IV. Tail-bound Analysis & ε-Optimality

In this section, we show that the costs under γ_R and γ approach each other as $N \to \infty$. To this end, we first prove Proposition 2 for the case of an ideal downlink with $p = 0$, where we show that the maximum AoI is uniformly bounded independent of N, and then Theorem 3 for the general non-ideal downlink case, where we provide a high confidence bound on the maximum AoI, again, independent of N. Then, we finally show (using Theorems 2 and 4) that γ approaches the optimal policy as $N \to \infty$ in both cases.

To this end, consider the Markov chain induced by the relaxed policy γ_R^i for the i^{th} agent as

$$\tau_{k+1}^i = \begin{cases} \tau_k^i + 1, & \text{w.p. } 1, \quad \tau_k^i < \underline{\kappa}^i(\underline{\lambda}^*), \\ \begin{cases} \tau_k^i + 1, & \text{w.p. } (1-q)p, \\ 0, & \text{w.p. } 1-(1-q)p, \end{cases} & \tau_k^i = \underline{\kappa}^i(\underline{\lambda}^*), \\ \begin{cases} \tau_k^i + 1, & \text{w.p. } p, \\ 0, & \text{w.p. } 1-p, \end{cases} & \tau_k^i \geq \overline{\kappa}^i(\overline{\lambda}^*). \end{cases}$$

Then, since each state in the set S is reachable from every other state, the above Markov chain is irreducible, and hence admits a unique stationary distribution π^i. Now, we provide the following proposition which shows that the AoI under MATB-P for a deterministic channel (with $p = 0$) is uniformly bounded, independent of N.

Proposition 2. *Under a fixed $\alpha = \mathcal{C}/N$ and $p = 0$, the AoI τ_k^i of any agent $i \in [N]$ under MATB-P is bounded by $\mathcal{O}(\alpha^{-1})$.*

As a result of the above proposition, we next prove that MATB-P and the relaxed policy approach each other as $N \to \infty$, which would then (as a result of (11)) imply that MATB-P is asymptotically optimal for Problem 1. To this end, we define an auxiliary policy $\hat{\gamma}$, under which the AoI sample paths are the same as those under γ_R, but for each additional agent that is not supposed to be transmitted by MATB-P, it adds a penalty to the cost as:

$$\omega(y, A, C_W) = c(\bar{\Delta}, A, C_W) \times \mathbf{1}_{\{(1-\frac{\mathcal{C}}{n_k^\lambda})>0\}} \mathbf{1}_{\{\tau \geq y\}}. \quad (10)$$

Further, we let $\{\tilde{\tau}_k^i\}_{k=1}^\infty$ and $\{\tau_k^i\}_{k=1}^\infty$ to be the sequences of AoIs of the i^{th} agent under MATB-P and γ_R^i (or equivalently $\hat{\gamma}^i$), respectively. Then, it is easy to see that $\omega(\tilde{\tau}^i(t), A_i, C_{W^i})$ dominates $c(\tilde{\tau}^i(t), A_i, C_{W^i})$, $\forall i, k$. As a consequence, it follows that

$$J^{BS}(\gamma_R) \leq J^{BS}(\gamma^*) \leq J^{BS}(\gamma) \leq J^{BS}(\hat{\gamma}), \quad (11)$$

where γ^* is any optimal policy that solves Problem 1. Then, we have the following result.

Theorem 2. *Let α be fixed and suppose that Assumption 1 holds. Then, the difference in the scheduling cost under MATB-P and γ_R converges to 0 exponentially fast as a function of N. Consequently, as $N \to \infty$, MATB-P becomes asymptotically optimal for Problem 1.*

Remark 1. *As a consequence of Proposition 2, in the case of adeterministic channel, no assumptions are needed on the system parameters to prove the asymptotic optimality of the MATB protocol. This is thus a significant relaxation of the result given in [2], where an upper bound on $\|A(\theta)\|_F$ was required. Second, we note that Theorem 2 proposes an exponential order of convergence of MATB-P toward optimality as $N \to \infty$, which is sharper than the $\mathcal{O}(1/\sqrt{N})$ convergence bound obtained in [2].*

Next, we proceed to the general case with $p > 0$. The following theorem shows that under MATB-P, the AoI takes large values with arbitrarily small probability.

Theorem 3. *Let α be fixed and $p > 0$. Then, given $\delta \in (0, 1)$, the upper confidence bound on AoI $\tau_k^i = \mathcal{O}(\log(1/\delta))$, $\forall i$, $\forall k$ with probability at least $1 - \delta$.*

Theorem 3 shows that the AoI has a vanishing tail under MATB-P, which can be used to give high-probability guarantees on the freshness of information under high traffic. Next, to prove asymptotic optimality of MATB-P under the case with erasure-prone channel, we again consider an auxiliary policy $\tilde{\gamma}$, which transmits agents according to γ_R, except that, for each agent which is not supposed to be transmitted by MATB-P, it adds an additional penalty to the cost, which (by a slight abuse of notation) is defined as:

$$\omega(y, A, C_W) = \sum_{\ell=1}^{\infty} p^{\ell} c(\tau + \ell, A, C_W) \times \mathbf{1}_{\{(1-\frac{c}{n_k^{\lambda}})>0\}} \mathbf{1}_{\{\tau \geq y\}},$$

such that $\omega(\tilde{\tau}_k^i, A_i, C_{W^i})$ dominates the expected WAoI $c(\tilde{\tau}_k^i, A_i, C_{W^i})$, for all i, k. Further, $\omega(y, A, C_W) < \infty$ as a consequence of Assumption 1. Thus, using similar arguments as for Theorem 2, we can prove the following main result.

Theorem 4. *Let α be fixed and $0 < p < 1$. Then, MATB-P approaches the optimal policy γ^* for Problem 1 exponentially fast as the number of agents grows.*

Remark 2. *We note here that in literature such as [6], [22], the authors have relied on truncating the AoI state space to a sufficiently large value and consequently worked with a finite space to derive the corresponding scheduling policies. Here, however, we do not require such a truncation on the state space. This is more natural as in communication systems with erasure channels, the AoI can always exceed the truncation value, even if the probability of the same tends to 0.*

The solution to the original capacity-constrained problem is thus completely characterized, and we next proceed to solving the finite-agent game problem.

V. Solution to Agent-Level Game Problem

In this section, we only provide an outline of the proof for the solution to the agent-level game problem (1)-(3) under the MATB-P scheduling policy as constructed in the previous section and the details can be found in [19]. Due to the difficulty in computing a centralized Nash equilibrium (NE) for a large population game [23], we instead compute a decentralized ϵ-NE by considering the MFG with infinite players. Consider a generic agent of type ϕ in the infinite population game, whose plant dynamics evolve as

$$X_{k+1} = A(\phi)X_k + B(\phi)U_k + W_k(\phi), \ k \geq 0,$$

where $X_k \in \mathbb{R}^n$ and $U_k \in \mathbb{R}^m$ denote, respectively, the state and control input of the generic agent. The initial state X_0 and the noise process $W_k(\phi) \in \mathbb{R}^n$ have, respectively, the same distributions as X_0^i and W_k^i in Section II for the corresponding type ϕ. The decoder and the controller information structures are also the same as in Section II. The goal of the generic agent is to minimize the cost function

$$J(\xi, \mu) := \overline{\lim_{T\to\infty}} \frac{1}{T} \mathbb{E}\bigg[\sum_{k=0}^{T-1} \|X_k - \mu_k\|^2_{Q(\phi)} + \|U_k\|^2_{R(\phi)}\bigg],$$

where the policy $\xi \in \Xi$ is adapted to the decentralized information structure $I_0^{d,con} := Z_0$, $I_k^{d,con} := \{U_{0:k-1}, Z_{0:k}\}$, $\forall k \geq 1$. Note that this is different from the centralized information structure of Section II. The sequence $\mu = (\mu_k)_{k\geq 0} \in \mathcal{M}$, also called the MF trajectory, is a bounded sequence which is the infinite agent approximation to the aggregate term $(\mu_k^N)_{k\geq 0}$ in (3). The analog to NE in the MFG is called the Mean-Field Equilibrium (MFE) and consists of the MFE controllers ξ_ϕ^*, $\forall \phi \in \Phi$ and MFE trajectory μ^* (full definition can be found in [19]). Next we characterize the MFE controller ξ_ϕ^*.

Proposition 3. *Suppose that Assumption 1 holds. Then, the MFE controller ξ_ϕ^*, $\forall \phi \in \Phi$ has the form*

$$U_k^* = \xi_\phi^*(Z_k) = -K_1(\phi)Z_k - K_2(\phi)g_{k+1}, \tag{12}$$

where matrices $K_1(\cdot)$ and $K_2(\cdot)$ are functions of model parameters and g_{k+1} is computed using the MFE trajectory μ^*. Furthermore, the cost under ξ_ϕ^* is bounded from above.

We remark here that another major advantage of this work is that, under the MATB-P scheduling policy, we can guarantee boundedness of cost without any assumption on the $\|A\|_F$ (in contrast to our earlier work [2]) with $p = 0$. The MFE trajectory μ^* can be computed using a fixed point operation under the contraction mapping assumption [15].

Assumption 2. $\|A_{cl}(\phi)\| + \sum_{\phi\in\Phi} \|Q(\phi)\| \|B(\phi)K_2(\phi)\|(1 - \|A_{cl}(\phi)\|)^{-2}\mathbb{P}(\phi) < 1$, $\forall \phi \in \Phi$.

Details of the computation of g and μ^* can be found in [19]. Having characterized the MFE, we now state the ϵ–Nash guarantee for this MFE.

Theorem 5. *Under Assumptions 1-2 the MFE constitutes an ϵ–Nash equilibrium for the N–agent game* (1)-(3)*, i.e.,*

$$J_i(\xi^{*,i}, \xi^{*,-i}) \leq \inf_{\pi_c^i \in \Pi_i} J_i(\pi_c^i, \xi^{*,-i}) + \mathcal{O}\Big(1/\sqrt{\min_{\phi\in\Phi} N_\phi}\Big),$$

where N_ϕ denotes the number of agents of type ϕ.

This guarantee shows that the MFE is an ϵ-NE with centralized information structure I^{C_i} (Section II) and ϵ approaches 0 at the rate $1/\sqrt{\min_{\phi\in\Phi} N_\phi}$.

VI. An Illustrative Example

In this section, we validate the theoretical results using a numerical example. We first demonstrate the asymptotic optimality of MATB-P. For this purpose, we consider values for N from 5 till 100, a time horizon of 5000 seconds, a low capacity $\mathcal{C} = 0.25N$, and an erasure probability $p = 0.2$. We plot the average weighted AoI of the system as a function of N in Fig. 2, for both the relaxed policy and MATB-P. We can see that the difference in the average cost under the above decays to 0, which shows the asymptotic optimality of the MATB policy.

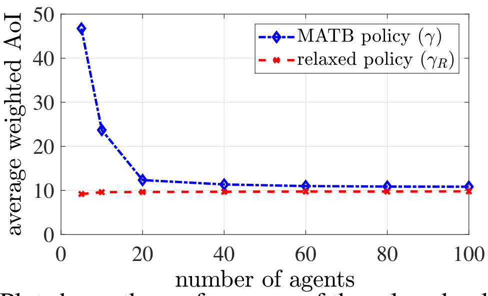

Fig. 2: Plot shows the performance of the relaxed policy (γ_R) and the MATB policy, converging to each other.

Next, we simulate the behavior of a 900–agent system, with 3 types of (scalar) agents, namely, with $A = 0.5, 1.0, 1.15$, under the MATB scheduling protocol and the MFE policy ξ^*. We take $B = 0.1269, C_W = 5$, $Q = R = 2$, and a horizon of 500 seconds. In Fig. 3, in the left plot, we show the variation of the average cost per agent as a function of the available capacity for a fixed erasure probability $p = 0.2$. Next, in the right, we show the variation of the average cost per agent as a function of the channel erasure for a fixed capacity ratio $\alpha = 0.45$. The figures show 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, and p, respectively. We can easily see that the average cost varies inversely with the available downlink capacity and in direct proportion to the erasure probability, aligned with intuition.

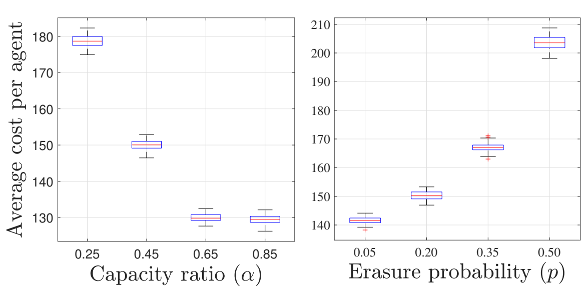

Fig. 3: Plots show the variation of aggregate cost per agent with (a) capacity ratio α, and (b) erasure probability p.

VII. Conclusion

In this paper, we have formulated a large population game problem involving information transmission over unreliable networks, thereby extending the setting of [2] and improving the guarantees in the special case considered in [2]. The network is regulated by a BS, for which we have constructed an asymptotically optimal scheduling policy. We have provided a tail analysis of the AoI under the same, first, for the case when the channel is free of any erasure and then for the case with erasure. Next, by using this policy, we have solved the consensus problem between the non-cooperative agents using the MFG framework by showing the ϵ–Nash property of the MFE to the finite-agent game problem. Finally, we have simulated a numerical example, which corroborates the theoretical developments.

References

[1] A. Osseiran, J. F. Monserrat, and P. Marsch, *5G Mobile and Wireless Communications Technology*. Cambridge University Press, 2016.

[2] 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.

[3] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.

[4] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

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

[6] 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.

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

[8] Y. Chen, H. Tang, J. Wang, and J. Song, "Optimizing age penalty in time-varying networks with Markovian and error-prone channel state," *Entropy*, vol. 23, no. 1, p. 91, January 2021.

[9] B. Sombabu, B. Dedhia, and S. Moharir, "Whittle index based age-of-information aware scheduling for Markovian channels," *Computer Networks and Comm.*, vol. 1, no. 1, pp. 59–84, December 2022.

[10] S. Wu, X. Ren, Q.-S. Jia, K. H. Johansson, and L. Shi, "Towards efficient dynamic uplink scheduling over multiple unknown channels," *arXiv:2212.06633*, December 2022.

[11] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2637–2650, December 2018.

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

[13] 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.

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

[15] 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 Apps.*, vol. 13, pp. 118–164, 2023.

[16] 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.

[17] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018.

[18] S. Aggarwal, M. A. uz Zaman, M. Bastopcu, and T. Başar, "Large population games with timely scheduling over constrained networks," in *IEEE ACC*, May-June 2023, pp. 4772–4778.

[19] ——, "Large population games on constrained unreliable networks," *arXiv: 2303.09515*, March 2023.

[20] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. John Wiley & Sons, 2012.

[21] E. Altman, *Constrained Markov Decision Processes: Stochastic Modeling*. Routledge, 1999.

[22] 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.

[23] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, 1998.