Minimizing Age of Information with Power Constraints: Opportunistic Scheduling in Multi-State Time-Varying Networks

Haoyue Tang, *Student Member, IEEE*, Jintao Wang, *Senior Member, IEEE*, Linqi Song, *Member, IEEE*, and Jian Song, *Fellow, IEEE*

Abstract—This work is motivated by the need of collecting fresh data from power-constrained sensors in the industrial Internet of Things (IIoT) network. A recently proposed metric, the *Age of Information* (AoI) is adopted to measure data freshness from the perspective of the central controller in the IIoT network. We wonder what is the minimum AoI performance the network can achieve and how to design scheduling algorithms to approach it. To answer these questions when the channel states of the network are Markov time-varying and scheduling decisions are restricted to bandwidth constraint, we first decouple the multi-sensor scheduling problem into a single-sensor constrained Markov decision process (CMDP) through relaxation of the hard bandwidth constraint. Next we exploit the threshold structure of the optimal policy for the decoupled single sensor CMDP and obtain the optimum solution through linear programming (LP). Finally, an asymptotically optimal truncated policy that can satisfy the hard bandwidth constraint is built upon the optimal solution to each of the decoupled single-sensor. Our investigation shows that to obtain a small AoI performance, the scheduler exploits good channels to schedule sensors supported by limited power. Sensors equipped with enough transmission power are updated in a timely manner such that the bandwidth constraint can be satisfied.

Index Terms—Age of Information, Cross-layer Design, Opportunistic Scheduling, Constrained Markov Decision Process

I. INTRODUCTION

The forthcoming Industrial 4.0 revolution brings more stringent data freshness requirement to support the higher level automated applications such as industrial manufacturing and factory automation [2]. In many of these applications, the monitor or the central controller collects data from sensors tracking real-time processes via time-varying wireless links [3]. The finite battery capacity, limited recharge resources [4] and wireless interference constraints cast restrictions on real time data sampling process and communications between the

H. Tang, J. Wang and J. Song are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China and Beijing National Research Center for Information Science and Technology (BNRist). J. Wang and J. Song are also with Research Institute of Tsinghua University in Shenzhen, Shenzhen, 518057. (e-mail: tanghaoyue13@tsinghua.org.cn; wangjintao@tsinghua.edu.cn; jsong@tsinghua.edu.cn).

L. Song is with the City University of Hong Kong, 83 Tat Chee Ave, Kowloon Tong, Hong Kong. (e-mail: linqi.song@cityu.edu.hk).

sensor and the monitor. In addition, data freshness requirement is different from traditional quality of service (QoS) guarantees such as communication latency and throughput. Thus, it is of great importance to revisit sampling and scheduling strategies in wireless networks in order to obtain more fresh information.

Previous techniques on minimizing communication latency and maximizing throughput may not be applied directly to data freshness optimization, since low latency and high throughput may not fulfill a good data freshness requirement. For example, in an automated manufacturing system illustrated in Fig. 1, the central controller makes decisions based on the received information packets from sensors. The metric the central controller concerns about most is neither the transmission delay of each data packet, nor the time average throughput, but the available data freshness from each sensor, i.e., what is the temperature and pressure of the environment at each producer now? A more relevant metric that captures data freshness, the *Age of Information* (AoI) [5], namely the time elapsed since the generation time-stamp of the freshest information stored at the receiver, has received increasing attention. As have been shown in [6]–[8], analyzing AoI performance and guaranteeing low AoI requirement are especially challenging, since the performance is determined by fundamental tradeoffs between communication latency, sampling interval and throughput.

Moreover, combating the time-varying characteristic of wireless fading channels with limited communication resources such as power consumption and bandwidth is important but challenging in stochastic networks, since these constraints and randomness appear at different layers of the communication networks [9] and require a joint design of physical and data link layer. In addition, the exponential growth of the cardinality of system states and action spaces imposes a curse of dimension in searching for the optimal policy. Although cross-layer control strategies have been studied in delay minimization, throughput and utility maximization, the design to optimize Age of Information has not been very well studied.

To address these challenges, in our paper, we consider a single controller multi-sensor IIoT network where each sensor is scheduled to transmit update packet by the central controller, as depicted in Fig. 1. The goal is to understand the how to design AoI minimization strategies in time-varying wireless networks with power constrained sensors. This scenario can be used to model the following applications in Industrial 4.0:

- *Factory Automation:* This application requires the central

controller supervising all rounds of the production process in order to guarantee efficient and safe operation. Each sensor is charged by different amount of power and tracks different servers during the manufacturing process. The central controller designs efficient load balancing algorithm for parallel servers based on the current manufacturing process reported by each sensor.

- *Intelligent Logistic:* The design of efficient utility maximization scheduling strategy in the intelligent logistic system requires precise observation and estimation of user demands. In this scenario, sensors can be viewed as power constrained wireless hot spots that collect time-varying user preferences and requirements, while the central controller makes real-time scheduling decision in the logistic network based on these demands.

A main feature of the model is that the channel states are multi-state time-varying and information collected by the sensors are time sensitive. We generalize our previous work [1] by assuming the channels have Markov evolving characteristics, which is more suitable to capture real-time fading effect. To ensure successful transmission, different level of transmission power is used in different channel state, while each sensor has an average power consumption constraint. The overall objective is to design scheduling policy such that the expected average AoI performance over the entire bandwidth constrained network can be minimized. Based on a single sensor level decomposition through a relaxation of the hard bandwidth constraint, we propose a truncated scheduling policy that can achieve an asymptotic optimal average AoI performance over the entire network.

The main contributions of the paper are summarized as follows:

- We propose a cross-layer framework to study AoI minimization scheduling in time-varying networks with power constrained sensors. The channel state is modeled to be a finite-state ergodic Markov chain but remains constant in each slot. Different level of transmit power is adopted in different channel state to ensure successful packet transmission. This model captures key features of practical cross-layer network optimization problem and facilitates analysis.
- We decouple the multi-sensor power constrained scheduling problem into a single-sensor constrained Markov decision process (CMDP) through relaxing the hard bandwidth constraint with the Lagrange multiplier. The threshold structure of the optimal policy for the decoupled single-sensor CMDP is revealed, and the search for the optimal policy is converted into a Linear Programming (LP).
- We adopt a dual-method to search for the Lagrange multiplier such that the relaxed bandwidth constraint can be satisfied. Then, we propose an asymptotic optimum truncated scheduling policy so that the hard bandwidth constraint of the network can be satisfied. The performance of the algorithm is analyzed and verified through simulations.

The remainder of this paper is organized as follows. We review some related work in Sec. II. The network model and the data freshness metric, AoI, are introduced in Section III. In Section IV, we decouple the multi-sensor scheduling problem into single-sensor level CMDP and search for the optimal policy through LP. In Section V, a truncated multi-sensor scheduling policy is proposed. Section VI evaluates and analyzes the performance of the proposed algorithm. Section VII draws the conclusion.

Notations: Vectors and matrices are written in boldface lower and upper letters. The probability of event $\mathcal{A}$ given condition $\mathcal{B}$ is denoted as $\Pr(\mathcal{A}|\mathcal{B})$. The expectation operation with regard to random variable X is denoted as $\mathbb{E}_X[\cdot]$. The cardinality of a set Ω is denoted as $|\Omega|$.

II. RELATED WORK

The analysis and optimization of AoI performance in average power constrained point to point communication systems have been studied [10]–[16]. It is revealed that the optimal sampling policy with power constrained transmitter in the presence of queueing delay [12] and transmission failure [15] possesses a threshold structure, i.e., sampling and update transmission occur when information at the receiver is no longer fresh while the update packets, if successfully received, can significantly reduce data staleness.

Another line of work focuses on designing scheduling strategies to minimize AoI performance in multi-user wireless networks [17]–[25]. When all the users in the network are identical and update packets can be generated at will, a greedy policy that schedules the user with the largest AoI is shown to be optimal [17]. When there is no packet-loss in the network, this greedy policy is equivalent to the round robin strategy, which is shown to be order optimal when update packets can not be generated at will and arrive randomly [23]. In [18], it is revealed that users with relatively bad channel states are updated less frequently. Scheduling in networks with time-varying channels are studied in [20], [21], where centralized and decentralized policies to minimize the average peak age of information (PAoI) are proposed.

Cross-layer control strategy to minimize communication latency under transmit power constraints have been studied in [26]–[33]. In [31], a Lazy scheduling policy that assigns scheduling decision based on the queue backlog is proposed. Considering the time-varying fading nature of wireless channels, rate and power adaptation strategy is proposed in [32]. To minimize queueing delay in a point to point time-varying channel with average power constraint on the transmitter, a probabilistic scheduling strategy is proposed in [28], [29]. However the above work consider wireless fading to be an i.i.d process. When channel state evolution has Markov properties, scheduling to minimize delay performance and maximize throughput have been studied in [26], [27], [33]. Scheduling policy based on value iteration is proposed in [27] and a Whittle-like index policy to achieve delay-power tradeoff is studied in [26]. In [33], the multi-user power and bandwidth constrained scheduling problem is solved by packet level decomposition, and an asymptotically optimum truncated scheduling policy is proposed.
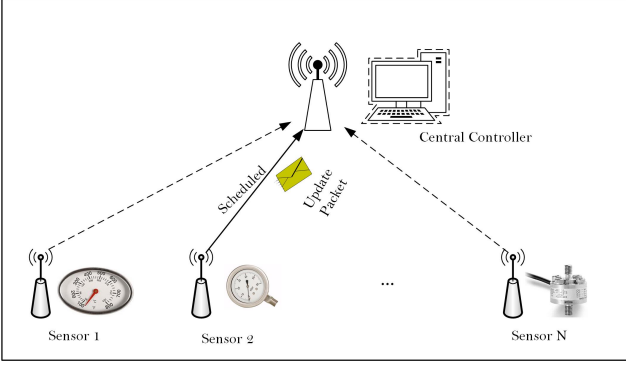
Fig. 1. Illustration of a typical industrial Internet of Things (IIoT) network.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. Network Model

We consider an industrial Internet of Things (IIoT) network as depicted in Fig. 1, where a central controller collects time-sensitive data from N sensors via wireless links. Let the time be slotted and use $t \in \{1, \cdots, T\}$ to denote the index of slot. Let the indicator function $u_n(t) = \{0, 1\}$ be a scheduling decision made by the central controller at the beginning of slot t. If $u_n(t) = 1$, then sensor n is scheduled to transmit update packet about his observation in slot t and the packet will be received by the end of the slot. Due to limited bandwidth constraint, no more than M sensors can be scheduled in slot t. Thus, we have the following constraint on $u_n(t)$:

$$\sum_{n=1}^{N} u_n(t) \leq M, \text{for all } t. \tag{1}$$

We assume that the communication channel between the central controller and each sensor is an ergodic Q-state Markov chain, where Q is a positive integer. Let $q_n(t) \in \{1, \cdots, Q\}$ be a random variable that captures the channel state of sensor n in slot t, which is assumed to remain constant during the slot. Large $q_n(t)$ indicates that communication channel of sensor n is more noisy and goes through stronger fading. Let $\mathbf{P}_n = \{p_{i,j}^{(n)}\}$ be the channel state transition matrix of sensor n, where each item $p_{i,j}^{(n)} \in [0, 1]$ is the probability that sensor n's channel state evolves from i to j in next slot, i.e.,

$$p_{i,j}^{(n)} = \Pr(q_n(t+1) = j | q_n(t) = i). \tag{2}$$

We assume that the channel states evolution is positively correlated, if the current channel state is good, it is more likely to evolve into good channel states in the next slot than from a bad state moving into good ones. Thus, for any channel state $i < j$, we have the following majorization relationship of $p_{i,k}^{(n)}$ and $p_{j,k}^{(n)}$ [34]:

$$p_{i,:}^{(n)} \succeq p_{j,:}^{(n)} : \sum_{k=1}^{K} p_{i,k}^{(n)} \geq \sum_{k=1}^{K} p_{j,k}^{(n)}, \forall K \leq Q. \tag{3}$$

When sensor n is scheduled to transmit updates in a slot and the corresponding channel state is q, in order to guarantee successful transmission, it will consume $\omega(q)$ units

of power. To combat the effect of channel fading, more power will be consumed when the channel is more noisy, thus $\omega(1) < \cdots < \omega(Q)$ is an increasing sequence. The transmitted packet will be successfully received by the central controller at the end of the slot. For a typical scheduling decision $\mathbf{u}_n(\pi) = [u_n(1), \cdots, u_n(T)]$ of sensor n, the average power consumed in T consecutive slots is:

$$E_n(\mathbf{u}_n(\pi)) = \frac{1}{T} \sum_{t=1}^{T} u_n(t) \omega(q_n(t)). \tag{4}$$

B. Age of Information

We measure data freshness of the central controller by using the metric *Age of Information*(AoI) [5]. By definition, the AoI is the time elapsed since the generation time-stamp of the freshest information at the receiver. An illustration of AoI evolution for a specific sensor is plotted in Fig. 2:
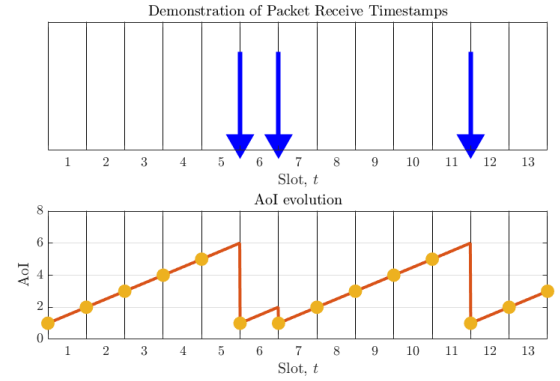

Fig. 2. Illustration of AoI evolution of a specific sensor. On the top, sample sequence representing the receiving time-stamps of the generate-at-will update packets. On the bottom, sample paths of AoI (red). The yellow dots depict the AoI at the beginning of each slot. Upon receiving a new packet, the AoI will drop to 1 at the beginning of next slot.

Let $x_n(t)$ be the AoI, i.e., the number of slots elapsed since the last delivery from sensor n at the beginning of slot t. If $u_n(t) = 1$, sensor n is scheduled in slot t and an update containing the freshest information tracked by sensor n will be received by the central controller, then $x_n(t+1) = 1$; otherwise, since there is no update packet received from sensor n during slot t, $x_n(t)$ increases linearly and $x_n(t+1) = x_n(t) + 1$. The evolution of $x_n(t)$ is organized as follows:

$$x_n(t+1) = \begin{cases} 1, & u_n(t) = 1; \\ x_n(t) + 1, & u_n(t) = 0. \end{cases} \tag{5}$$

C. Problem Formulation

For a given network setup with N sensors and channel states evolution $\{\mathbf{P}^{(n)}\}$, we measure the data freshness of the IIoT network by following policy π in terms of the expected average AoI of all sensors at the beginning of each slot for a total of consecutive $T \to \infty$ slots, which can be computed as follows:

$$J(\pi) = \lim_{T \to \infty} \{ \frac{1}{NT} \mathbb{E}_\pi \left[\sum_{t=1}^{T} \sum_{n=1}^{N} x_n(t) | \mathbf{x}(0) \right] \} \tag{6}$$

where the vector $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_N(t)] \in \mathbb{N}^N$ denotes the AoI of all sensors at the beginning of slot t. In this work, we assume that all the sources have been synchronized initially, i.e., $\mathbf{x}(0) = \mathbf{1}$ and omit it henceforth.

Let Π_{NA} denote the class of non-anticipated policies, i.e., scheduling decisions are made based on past, current AoI, channel states and their evolving probabilities, while no information about the future AoI or channel states are used. The average power constraint of each sensor is also known by the central controller. In this research, we aim at designing policy $\pi \in \Pi_{\mathrm{NA}}$ in order to minimize the average expected AoI of all the sensors, while each sensor has a time average power consumption constraint. The original bandwidth and power constrained AoI minimization problem (B&P-Constrained AoI) is organized as follows:

Problem 1 (B&P-Constrained AoI):

$$\pi^* = \arg\min_{\pi \in \Pi_{\mathrm{NA}}} \lim_{F \to \infty} \{ \frac{1}{NT} \mathbb{E}_\pi \left[\sum_{t=1}^{T} \sum_{n=1}^{N} x_n(t) \right] \}, \tag{7a}$$

$$\text{s.t.} \sum_{n=1}^{N} u_n(t) \leq M, \forall t, \tag{7b}$$

$$\lim_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=1}^{T} u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \forall n. \tag{7c}$$

Notice that the hard bandwidth constraint (7b) in every slot t suggests that the *B&P-Constrained AoI problem* is an intractable integer programming, whose computational complexity remains open. We tackle with this challenge through the following approaches:

- Inspired by [27], [33], [35], in Section IV-A, we first relax the hard bandwidth constraint (7b) and adopt a sensor level decomposition by using Lagrange multiplier. After relaxation, multiple sensors can be scheduled simultaneously.
- In Section V, we propose a truncated scheduling policy to satisfy the hard bandwidth constraint (7b) based on the solution to each of the decoupled single sensor.

IV. SCHEDULING BY SENSOR-LEVEL DECOMPOSITION

In this section, we start by relaxing and decoupling the *B&P-Constrained AoI*, then formulate the decoupled single sensor scheduling problem into a constrained Markov decision process (CMDP). We exploit the threshold structure of the optimal stationary randomized policy and the optimal solution is solved through linear programming (LP).

A. Sensor Level Decomposition

Let us first relax the hard constraint (7b) into an time-average constraint, the relaxed bandwidth and power constrained AoI minimization problem (RB&P-Constrained AoI) can be organized as follows:

Problem 2 (RB&P-Constrained AoI):

$$\pi_R^* = \arg\min_{\pi \in \Pi_{\mathrm{NA}}} \lim_{T \to \infty} \{ \frac{1}{NT} \mathbb{E}_\pi \left[\sum_{t=1}^{T} \sum_{n=1}^{N} x_n(t) \right] \}, \tag{8a}$$

$$\text{s.t.} \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{N} u_n(t) \leq M, \tag{8b}$$

$$\lim_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=1}^{T} u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \forall n. \tag{8c}$$

Next, let us place the relaxed constraint into the objective function in order to carry out the sensor-level decomposition afterwards:

$$L(\pi, W) = \lim_{T \to \infty} \{ \frac{1}{NT} \mathbb{E}_\pi \left[\sum_{n=1}^{N} \sum_{t=1}^{T} \left(x_n(t) + W u_n(t) - \frac{WM}{N} \right) \right] \}. \tag{9}$$

The Lagrange multiplier $W \geq 0$ associates with the relaxed constraint and can be viewed as a penalty incurred by policies that want to schedule more users than the relaxed constraint. For fixed W, the optimization problem (9) can then be decoupled into N single sensor AoI and scheduling penalty minimization problem with average power consumption constraint (7c), then the decoupled single sensor power constrained cost minimization problem (Decoupled P-Constrained Cost) can be written out as follows:

Problem 3 (Decoupled P-Constrained Cost):

$$\mathcal{L}_n(\pi_n, W) = \lim_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi_n} \left[\sum_{t=1}^{T} x_n(t) + W u_n(t) \right], \tag{10a}$$

$$\text{s.t} \lim_{T \to \infty} \mathbb{E}_{\pi_n} \left[\sum_{t=1}^{T} u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n. \tag{10b}$$

Since the primal relaxed problem (9) gets decoupled, we omit the subscript n henceforth. We formulate the *Decoupled P-Constrained Cost* minimization problem into an CMDP in Sec. III(B) and analyze the optimum structure in Sec. III(C). In Sec. III(D), we convert the single-sensor optimization problem with fixed W into a Linear Programming (LP).

B. Constrained Markov Decision Process Formulation

The decoupled single-sensor scheduling problem can be formulated into a CMDP that consists of a quadruplet $(\mathbb{S}, \mathbb{A}, \Pr(\cdot|\cdot), C(\cdot, \cdot))$, each item is explained as follows:

- **State Space:** The state of a sensor in slot t is the current AoI and the channel state $(x(t), q(t))$. The state space $\mathbb{S} = \{x \times q\}$ is thus countable but infinite.
- **Action Space:** There are two possible actions $s \in \mathbb{A} = \{0, 1\}$, while $s(t) = 1$ denotes the sensor is scheduled to deliver updates to the central controller in slot t, while $s(t) = 0$ represents that the sensor keeps idle and is not scheduled. Notice that $s(t)$ is different from scheduling decision $u(t)$, which has strict bandwidth constraint.
- **Probability Transfer Function:** If the sensor is not scheduled during slot t, i.e., $s(t) = 0$, then $x(t+1) = x(t)+1$, otherwise if the sensor is scheduled, then the AoI drops to $x(t+1) = 1$. The channel state $q(t+1)$ evolves independently of $x(t)$ and only relies on $q(t)$ due to its Markov property, hence the probability transfer function from state (x, q) is organized as follows:

$$\Pr((x, q) \to (x+1, q')) = p_{q,q'}, \qquad s = 0; \tag{11a}$$

$$\Pr((x, q) \to (1, q')) = p_{q,q'}, \qquad s = 1. \qquad (11b)$$

- **One-Step Cost:** For given state (x, q), the one-step cost by taking action s contains AoI growth and scheduling penalty, which can be computed as follows:

$$C_X(x, q, s) = x + Ws, \qquad (12a)$$

while the one-step power consumption is:

$$C_Q(x, q, s) = \omega(q)s. \qquad (12b)$$

The objective of the decoupled CMDP is to design a scheduling policy π such that the following average cost over infinite horizon can be minimized:

$$\frac{1}{T}\mathbb{E}_\pi\left[\sum_{t=1}^{T} C_X(x(t), q(t), s(t))\right],$$

while the average power constraint is satisfied,

$$\frac{1}{T}\mathbb{E}_\pi\left[\sum_{t=1}^{T} C_Q(x(t), q(t), s(t))\right] \leq \mathcal{E}.$$

C. Characterization of the Optimal Policy

In this part, we focus on exploiting the threshold structure of the optimal policy. First we provide the formal definition of stationary randomized policies and stationary deterministic policies:

Definition 1: Let Π_{SR} and Π_{SD} denote the class of stationary randomized and stationary deterministic policies, respectively. Given observation $(x(t) = x, q(t) = q)$, a stationary randomized policy $\pi_{\mathrm{SR}} \in \Pi_{\mathrm{SR}}$ chooses action $s(t) = 1$ with probability measure $\xi_{x,q} \in [0, 1]$ for all t. A stationary deterministic policy $\pi_{\mathrm{SD}} \in \Pi_{\mathrm{SD}}$ selects action $s(t) = a(x, q)$, where $a(\cdot) : (x, q) \to \{0, 1\}$ is a deterministic mapping from state space to action space.

According to [36, Theorem 4.4], the optimal policy to the above CMDP has the following property:

Corollary 1: An optimal stationary randomized policy $\pi^* \in \Pi_{\mathrm{SR}}$ exists for the decoupled single sensor power constrained scheduling problem (10a), and it is a mixture of no more than two stationary deterministic policies $\pi_{\mathrm{SD1}}, \pi_{\mathrm{SD2}} \in \Pi_{\mathrm{SD}}$. Let λ be the weight of following stationary deterministic policy π_{SD1} and $1 - \lambda$ be the weight of following π_{SD2}. Then the optimum policy is:

$$\pi^* = \lambda\pi_{\mathrm{SD1}} + (1 - \lambda)\pi_{\mathrm{SD2}}. \qquad (13)$$

Each of the deterministic policy can be obtained through the Lagrangian primal-dual method [36]. Let $\lambda \geq 0$ be the Lagrange multiplier related to the average power constraint, then the single sensor CMDP can be converted into an unconstrained MDP, the objective is to minimize the following overall cost:

$$\frac{1}{T}\mathbb{E}_\pi\left[\sum_{t=1}^{T} [C_X(x(t), q(t), s(t)) + \lambda C_Q(x(t), q(t), s(t))]\right] - \lambda\mathcal{E} \qquad (14)$$

For given Lagrange multiplier λ, a stationary deterministic policy to minimize the above unconstrained cost exists. Moreover, there exits a differential cost-to-go function $V(x, q)$ that satisfies the following Bellman equation:

$$V(x, q) + \gamma = \min\{C_X(x, q, 0) + \sum_{q'=1}^{Q} p_{q,q'}V(x+1, q'),$$

$$C_X(x, q, 1) + \sum_{q'=1}^{Q} p_{q,q'}V(1, q') + \lambda C_Q(x, q, 1)\}, \qquad (15)$$

where γ is the average cost by following the optimal policy. Next, we will prove the threshold structure of the stationary deterministic policy for given λ, which will present insight for the structure of the optimal stationary randomized policy to solve the *Decoupled P-Constrained Cost* minimization problem.

Lemma 1: With fixed λ, the optimal stationary deterministic policy for solving the *Decoupled P-Constrained Cost* problem (14) possesses a threshold structure for each channel state q, i.e., when $x \geq \tau_q$, the optimal action $s^*(x, q) = 1$ and when $x < \tau_q$, $s^*(x, q) = 0$.

Proof: The proof is provided in Appendix A. Here we provide an intuitive analysis. Since communication between the sensor and the controller is power constrained, we will only schedule when the information is no longer fresh or the channel state is good, i.e., x is large or q is small. This behavior characterizes a threshold structure. ∎

D. Probabilistic Scheduling Policy for Single Sensor Case

Let us now investigate into the class of stationary randomized policies. Denote $\xi_{x,q}$ to be the probability that the sensor is scheduled to send updates in state (x, q). We aim at finding a set of optimal transmission probability $\{\xi_{x,q}^*\}$ to solve the *Decoupled P-Constrained Cost* problem, i.e., total cost of AoI performance and scheduling penalty for a single decoupled sensor can be minimized. From Sec. IV(C), the optimal stationary randomized policy to the CMDP (10a) is a randomization between two stationary deterministic policies [36], each of them can be obtained by solving the unconstrained MDP (14). Considering the threshold structure of them and the stationary randomized strategy Eqn. (13), it can be concluded there exists a set of thresholds τ_q, for each state (x, q), if $x \geq \max_q \tau_q$, the stationary randomized policy is to schedule the sensor. Thus, $\xi_{x,q}^* = 1, \forall(x, q), x \geq \max_q \tau_q$. Therefore, for each of the decoupled single sensor problem, the AoI x cannot be larger than the largest threshold $\max_q \tau_q$. To find the optimal policy, we choose a large bound $X_{\max}$ for x that can guarantee $X_{\max} \geq \max_q \tau_q$. We only consider policy that satisfy $\xi_{x,q} = 1, \forall x \geq X_{\max}$ in the following analysis, since other policies that do not have the optimum structure can be excluded from the discussions.

Let $\mu_{x,q}$ denote the probability that the sensor's AoI is x and the current channel state is q. To illustrate the state transition relationship, we provide transfer graph for $Q = 2$ as an example in Fig. 3. Let $\alpha_{q,q'}^x$ denote the one step forward state transition probability from (x, q) to $(x + 1, q')$ and let $\beta_{q,q'}^x$ be the backward transition probability from (x, q) to $(1, q')$, respectively. From the discussed threshold structure of the stationary deterministic policies, with properly selected $X_{\max}$, under the optimal scheduling policy, the steady state

distribution $\mu_{X_{\max},q} = 0, \forall q$. According to the probability transfer graph Fig. 3, the forward and backward transition probability for a scheduling policy $\xi_{x,q}$ can be computed as follows:

$$\alpha_{q,q'}^x = \Pr((x,q) \to (x+1,q')) = (1-\xi_{x,q})p_{q,q'}, \quad (16a)$$
$$\beta_{q,q'}^x = \Pr((x,q) \to (1,q')) = \xi_{x,q}p_{q,q'}. \quad (16b)$$

Let $\boldsymbol{\mu} = [\mu_{1,1}, \cdots, \mu_{1,Q}, \cdots, \mu_{X_{\max},1}, \cdots, \mu_{X_{\max},Q}]^T$ be the steady state distribution. Let $\mathbf{Q}$ be the probability transfer matrix between the states, according to Fig. 3, $\mathbf{Q}$ is constructed as follows:

$$\mathbf{Q} = \begin{bmatrix} \boldsymbol{\beta}^1 & \boldsymbol{\beta}^2 & \cdots & \boldsymbol{\beta}^{X_{\max}-1} & \boldsymbol{\beta}^{X_{\max}} \\ \boldsymbol{\alpha}^1 & \mathbf{0}_Q & \cdots & \mathbf{0}_Q & \mathbf{0}_Q \\ \mathbf{0}_Q & \boldsymbol{\alpha}^2 & \cdots & \mathbf{0}_Q & \mathbf{0}_Q \\ \cdots & \cdots & \cdots & \cdots\cdots \\ \mathbf{0}_Q & \mathbf{0}_Q & \cdots & \boldsymbol{\alpha}^{X_{\max}-1} & \mathbf{0}_Q \end{bmatrix}, \quad (17)$$

where vector $\mathbf{0}_Q$ is a Q-dimension vector with all the elements being 0. Matrices $\boldsymbol{\alpha}^x$ and $\boldsymbol{\beta}^x$ are the forward and backward transition matrix from state x, respectively, which can be computed as follows:

$$\boldsymbol{\alpha}^x = \begin{bmatrix} \alpha_{1,1}^x & \alpha_{2,1}^x & \cdots & \alpha_{Q,1}^x \\ \alpha_{1,2}^x & \alpha_{2,2}^x & \cdots & \alpha_{Q,2}^x \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1,Q}^x & \alpha_{2,Q}^x & \cdots & \alpha_{Q,Q}^x \end{bmatrix}, \quad (18a)$$

$$\boldsymbol{\beta}^x = \begin{bmatrix} \beta_{1,1}^x & \beta_{2,1}^x & \cdots & \beta_{Q,1}^x \\ \beta_{1,2}^x & \beta_{2,2}^x & \cdots & \beta_{Q,2}^x \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{1,Q}^x & \beta_{2,Q}^x & \cdots & \beta_{Q,Q}^x \end{bmatrix}. \quad (18b)$$

According to property of the steady state distribution, we have $\mathbf{Q}\boldsymbol{\mu} = \boldsymbol{\mu}$. In addition, considering that $\forall x \geq X_{\max}$, the steady state distribution $\mu_{x,q} = 0$. We then have $\sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} \mu_{x,q} = 1$. Thus, the steady distribution $\boldsymbol{\mu}$ relates to strategy $\{\xi_{x,q}\}$ is the solution to the following linear equations:

$$\begin{bmatrix} \mathbf{Q} - \mathbf{I}_{Q \times X_{\max}} \\ \mathbf{1}_{Q \times X_{\max}}^T \end{bmatrix} \boldsymbol{\mu} = \begin{bmatrix} \mathbf{0}_{Q \times X_{\max}} \\ 1 \end{bmatrix}, \quad (19)$$

where $\mathbf{1}_{Q \times X_{\max}}$ is a $(Q \times X_{\max})$-dimension column vector with all the elements being 1 and $\mathbf{I}_{Q \times X_{\max}}$ is an identity matrix.

Next, we will convert the search for the optimal stationary randomized scheduling strategy into an LP. We introduce a new set of variables $y_{x,q} = \mu_{x,q}\xi_{x,q}$, each denotes the probability of the sensor being in state (x,q) and is scheduled to transmit an update. With this set of variables, we present the following theorem:

Theorem 1: Solving the *Decoupled P-Consrained Cost* minimization problem is equivalent to solve the following LP problem:

$$\{\mu_{x,q}^*, y_{x,q}^*\} = \arg \min_{\{\mu_{x,q}, y_{x,q}\}} \sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} (Wy_{x,q} + x\mu_{x,q}),$$
$$(20a)$$

$$\text{s.t. } \mu_{1,q} = \sum_{x=1}^{X} \sum_{q'=1}^{Q} y_{x,q'}p_{q',q}, \quad (20b)$$

$$\mu_{x,q} = \sum_{q'=1}^{Q} (\mu_{x-1,q'} - y_{x-1,q'})p_{q',q}, \quad (20c)$$

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} \mu_{x,q} = 1, \quad (20d)$$

$$y_{x,q} \leq \mu_{x,q}, \quad (20e)$$

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} y_{x,q}\omega(q) \leq \mathcal{E} \quad (20f)$$

$$0 \leq \mu_{x,q} \leq 1, 0 \leq y_{x,q} \leq 1, \forall x, q. \quad (20g)$$

Proof: Let us compute the equivalent time average cost to Eqn. (10a) by using variables $\{\mu_{x,q}\}$ and $\{y_{x,q}\}$. The probability that the sensor is in state (x,q) is $\mu_{x,q}$. With probability $\xi_{x,q}$, the sensor is selected to be scheduled and incurs a cost of $C_X(x,q,1) = x + W$, and the sensor is selected to keep idle with probability $1 - \xi_{x,q}$ and incurs a cost of $C_X(x,q,0) = x$. Then the time average cost by following policy $\{\xi_{x,q}\}$ can be computed by:

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} \mu_{x,q}(\xi_{x,q}(x+W) + (1-\xi_{x,q})x)$$
$$= \sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} (Wy_{x,q} + x\mu_{x,q}). \quad (21)$$

If the sensor is scheduled to transmit in state (x,q), the power consumed is $\omega(q)$. Then, the time-average power consumed by employing policy $\{\xi_{x,q}\}$ is:

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} \mu_{x,q}\xi_{x,q}\omega(q) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} y_{x,q}\omega(q), \quad (22)$$

with this equation the power constraint (7c) can be converted in the linear constraint (20f). The constraint Eqn, (20b)-(20d) can be obtained by substituting $\xi_{x,q}$ with $y_{x,q}$ and $\mu_{x,q}$ with relationship (19). Notice that $\xi_{x,q} \leq 1$, the inequality constraint (20e) can be obtained. ∎

Till now, we construct an LP problem to obtain the optimum stationary randomized policy that can minimize the total cost of a single sensor with fixed Lagrange multiplier W. Next, we can construct the optimal stationary randomized scheduling policy to minimize Lagrange function (10a) for a single user. According to the threshold structure of each deterministic policy and Eqn. (13), we will have the following property on $\xi_{x,q}^*$:

Corollary 2: For any channel state q, the optimal scheduling decisions $\xi_{\cdot,q}^*$ is monotonically increasing, i.e.,

$$\xi_{x_1,q}^* \leq \xi_{x_2,q}^*, \forall 1 \leq x_1 < x_2. \quad (23)$$

V. Multi-sensor Opportunistic Scheduling

In this section, we will provide an algorithm to determine the multiplier W such that relaxed bandwidth constraint can be satisfied and *RB&P-Constrained AoI* problem can be solved. Then, we propose a truncated scheduling algorithm for the multi-sensor case that satisfies the original hard bandwidth constraint Eqn. (7b).
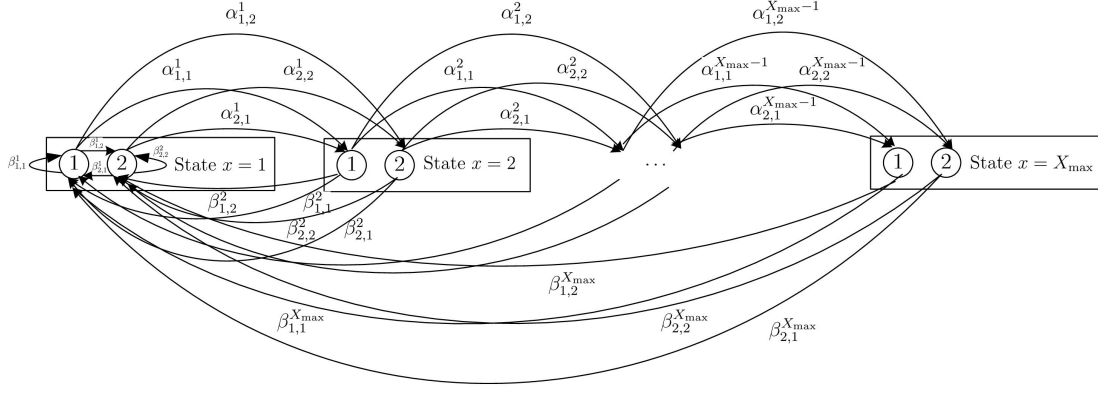
Fig. 3. Illustrative of the probability transfer graph for a stationary randomized policy with $Q = 2$ channel states. The circles denote channel state q and the rectangles denote the sensor's AoI x. The forward state transmission probability (x, q) to $(x+1, q')$ is $\alpha_{q,q'}^x$ and the backward state transmission probability from (x, q) to $(1, q')$ is $\beta_{q,q'}^x$.

A. Determination of Lagrange Multiplier

Let $g(W)$ denote the Lagrange dual function, i.e.,

$$g(W) = \min_{\pi \in \Pi_{NA}} \mathcal{L}(\pi, W). \tag{24}$$

Since the relaxed problem gets decoupled into N single user CMDP, the dual function can be computed by:

$$g(W) = \frac{1}{N} \sum_{n=1}^{N} g_n(W) - WM, \text{ where}$$

$$g_n(W) = \min_{\pi_n \in \Pi_{NA}} \left(\mathcal{L}_n(\pi_n, W) \right), \text{ s.t. Eqn. (7c)}. \tag{25}$$

By Theorem 1, the CMDP that minimizes $\mathcal{L}_n(\pi, W)$ is equivalent to an LP, then $g_n(W)$ is just the average cost of the CMDP. Let $\overline{X}_n(W)$ and $\overline{A}_n(W)$ denote the average AoI and the average scheduling probability of sensor n, respectively. By computing the optimum resource allocation vector $\{y_{x,q}^*\}$ through solving LP (20), function $g_n(W)$ can be computed as follows:

$$g_n(W) = \overline{X}_n(W) + W \overline{A}_n(W) \tag{26a}$$

$$\text{where } \overline{X}_n(W) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} x \mu_{x,q}^*, \tag{26b}$$

$$\overline{A}_n(W) = \sum_{x=1}^{X} \sum_{q=1}^{Q} y_{x,q}^*. \tag{26c}$$

Finally, we apply the subgradient descent method to search for the Lagrange optimizer. Let $W^{(k)}$ be the Lagrange multiplier used in the kth iteration. According to [37, Eqn. 6.1.1], the subgradient at $W^{(k)}$ can be computed by:

$$\mathrm{d}_W g(W^{(k)}) = \sum_{n=1}^{N} \overline{A}_n(W^{(k)}) - M. \tag{27}$$

We start with $W^{(0)} = 0$, if $\sum_{n=1}^{N} \overline{A}_n(W^{(0)}) - M \leq 0$, then scheduling does not have to consider the relaxed bandwidth constraint. The minimum AoI performance to the *RB&P-Constrained AoI* problem and the lower bound on the AoI performance to the primal *B&P-Constrained AoI* can be computed simply through:

$$\text{AoI}_{\text{LB}} = \text{AoI}_R^* = g(0). \tag{28}$$

Otherwise, we adopt an iterative algorithm update. By choosing a set of stepsizes γ_k similar to [15], the multiplier for the next iteration can be computed by:

$$W^{(k+1)} = W^{(k)} + \gamma_k \mathrm{d}_W g(W^{(k)}). \tag{29}$$

The iteration ends until $|W^{(k)} - W^{(k+1)}| < \varepsilon$.

However, since the *RB&P-Constrained AoI* is a constrained Markov decision process, the optimum scheduling policy of which should be a randomization between no more than two policies, each is the solution to minimize the Lagrange function Eqn. (9). The randomization between the two policies will enable us to satisfy the relaxed bandwidth constraint Eqn. (8b) in the *RB&P-Constrained AoI*. Next, we will talk about how to obtain the optimum randomized strategy from the obtained Lagrange multipliers sequence $\{W^{(k)}\}$.

Let W_l and W_u be two Lagrange multipliers chosen from sequence $W^{(k)}$,

$$W_l = \arg\max_{W^{(k)}} \sum_{n=1}^{N} \overline{A}_n(W^{(k)}), \text{s.t.} \sum_{n=1}^{N} \overline{A}_n(W^{(k)}) \leq M, \tag{30a}$$

$$W_u = \arg\min_{W^{(k)}} \sum_{n=1}^{N} \overline{A}_n(W^{(k)}), \text{s.t.} \sum_{n=1}^{N} \overline{A}_n(W^{(k)}) \geq M. \tag{30b}$$

Let $M_l = \sum_{n=1}^{N} \overline{A}_n(W_l)$ and $M_u = \sum_{n=1}^{N} \overline{A}_n(W_u)$ be the total bandwidth used with respect to minimize the function Eqn. (9). Suppose $\{\boldsymbol{\mu}^{n,l}, \mathbf{y}^{n,l}\}$ is the optimizer to sensor n's LP problem (21) with multiplier W_l and $\{\boldsymbol{\mu}^{n,u}, \mathbf{y}^{n,u}\}$ is the solution with multiplier W_u. To satisfy the relaxed bandwidth constraint, the optimum distribution $\{\boldsymbol{\mu}^{n,*}, \mathbf{y}^{n,*}\}$ of the relaxed problem is a linear combination of $\{\boldsymbol{\mu}^{n,l}, \mathbf{y}^{n,l}\}$ and $\{\boldsymbol{\mu}^{n,r}, \mathbf{y}^{n,r}\}$, which can be computed as follows:

$$\{\boldsymbol{\mu}^{n,*}, \mathbf{y}^{n,*}\} = \lambda \{\boldsymbol{\mu}^{n,l}, \mathbf{y}^{n,l}\} + (1 - \lambda) \{\boldsymbol{\mu}^{n,u}, \mathbf{y}^{n,u}\}, \tag{31}$$

where the mixing coefficient can be computed in a similar manner to [15]:

$$\lambda = \frac{M_u - M}{M_u - M_r}.$$

Notice that $\{\boldsymbol{\mu}^{n,*}, \mathbf{y}^{n,*}\}$ still satisfy the constraint of the LP problem for sensor n. Consider the structure of each *Decoupled P-Constrained Cost* problem, the optimum scheduling strategy π_R^* for the *RB&P-Constrained* is then constructed as follows:

In each slot t, the central controller observe the current AoI $x_n(t)$ and channel state $q_n(t)$ of sensor n, a scheduling decision $s_n(t) = 1$ is then made with probability $\xi_{x_n(t),q_n(t)}^{n,*}$ is can be computed as follows:

$$\xi_{x,q}^n = \begin{cases} 1, & \xi_{x-1,q}^n = 1 \text{ or } \mu_{x,q}^{n,*} = 0 \text{ or } x \geq X_{\max}; \\ \frac{y_{x,q}^{n,*}}{\mu_{x,q}^{n,*}}, & \text{otherwise.} \end{cases} \quad (32)$$

The algorithm flow chart to obtain $\xi_{x,q}^{n,*}$ is finally provided as follows:

Algorithm 1 Determination of the optimum scheduling probabilities $\xi_{x,q}^{n,*}$ to the *RB&P-Constrained AoI* Problem

---

1: **initialization**: start with $W^{(0)} = 0$, solve the corresponding LP (20) for each sensor n and compute $\overline{A}_n(W^{(0)})$, denote the optimizer as $\{\boldsymbol{\mu}^n, \mathbf{y}^n\}$.
2: $k \leftarrow 0, M_l \leftarrow 0, M_u \leftarrow 2M, \{\boldsymbol{\mu}^{n,l}, \mathbf{y}^{n,l}\} = \{\boldsymbol{\mu}^{n,u}, \mathbf{y}^{n,u}\} = \{\boldsymbol{\mu}^n, \mathbf{y}^n\}$
3: **if** $\sum_{n=1}^N \overline{A}_n(W^{(0)}) - M \leq 0$ **then** ▷ Relaxed Bandwidth Constraint is satisfied
4:     $\{\boldsymbol{\mu}^{n,*}, \mathbf{y}^{n,*}\} = \{\boldsymbol{\mu}^n, \mathbf{y}^n\}, \forall n$
5: **else**   ▷ Search for the Lagrange Multiplier
6:     **repeat**
7:         $k \leftarrow k + 1$
8:         $\mathrm{d}_W g(W^{(k)}) = \sum_{n=1}^N \overline{A}_n(W^{(k)}) - M$
9:         $W^{(k+1)} \leftarrow W^{(k)} + \gamma_k d_W g(W^{(k)})$
10:        Solve the corresponding LP (20) for each sensor n and compute $\overline{A}_n(W^{(k)})$, denote the optimizer as $\{\boldsymbol{\mu}^n, \mathbf{y}^n\}$
11:        **if** $M_l < \sum_{n=1}^N \overline{A}_n(W^{(0)}) \leq M$ **then**
12:            $M_l \leftarrow \sum_{n=1}^N \overline{A}_n(W^{(0)})$
13:            $\{\boldsymbol{\mu}^{n,l}, \mathbf{y}^{n,l}\} \leftarrow \{\boldsymbol{\mu}^n, \mathbf{y}^n\}$
14:        **else if** $M < \sum_{n=1}^N \overline{A}_n(W^{(0)}) \leq M_u$ **then**
15:            $M_u \leftarrow \sum_{n=1}^N \overline{A}_n(W^{(0)})$
16:            $\{\boldsymbol{\mu}^{n,u}, \mathbf{y}^{n,u}\} \leftarrow \{\boldsymbol{\mu}^n, \mathbf{y}^n\}$
17:     **until** $|W^{(k+1)} - W^{(k)}| < \varepsilon$
18:     $\lambda \leftarrow \frac{M_u - M}{M_u - M_r}$ ▷ Strategy Randomization
19:     $\{\boldsymbol{\mu}^{n,*}, \mathbf{y}^{n,*}\} = \lambda\{\boldsymbol{\mu}^{n,l}, \mathbf{y}^{n,l}\} + (1-\lambda)\{\boldsymbol{\mu}^{n,u}, \mathbf{y}^{n,u}\}$
20: Compute $\{\xi_{x,q}^n\}$ according to Eqn. (32)

Finally, the minimum AoI performance to the *RB&P-Constrained AoI* problem can be computed through according to the optimizer $\{\boldsymbol{\mu}^{n,*}, \mathbf{y}^{n,*}\}$, which also formulates the lower bound on the AoI performance to the primal *B&P-Constrained AoI*:

$$\text{AoI}_{\text{LB}} = \text{AoI}_R^* = \sum_{n=1}^N \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q x\mu_{x,q}^{n,*}. \quad (33)$$

B. Multi-sensor opportunistic scheduling with hard bandwidth constraint

In this part we construct a truncated policy π based on optimal scheduling policy for each of the decoupled sensor and solve the primal *B&P-Constrained AoI* problem. Let π_R^* be the optimum scheduling policy obtained in Sec. IV(A), where $s_n(t)$ is the scheduling decision under the relaxed constraint, which measures if sensor n is eager be scheduled. Denote $\Omega(t) = \{n | s_n(t) = 1\}$ as the set of sensors that are eager to be scheduled. The scheduling decision $u_n(t)$ under hard bandwidth constraint is then carried out as follows:

- If $|\Omega(t)| \leq M$, i.e., the total number of sensors that are eager to send updates is less than or equal to the bandwidth resource available, then the scheduling decision $u_n(t) = 1, \forall s_n(t) = 1$.
- Otherwise if $|\Omega(t)| > M$, the central controller selects a subset of $\mathcal{M}(t) \in \Omega(t), |\mathcal{M}(t)| = M$ sensors from $\Omega(t)$ randomly and schedules them to send updates. Those sensors that are in set $\Omega(t)$ but not selected in $\mathcal{M}(t)$ is not scheduled because of limited bandwidth constraint.

Theorem 2: With the proportion of scheduling resources $\frac{M}{N} = \theta$ keeps a constant, the deviation from the optimal scheduling policy for a network with N sensors under the proposed truncated policy $\tilde{\pi}$ is $\mathcal{O}(\frac{1}{\sqrt{N}})$. Thus, with $N \to \infty$ and $\frac{M}{N} = \theta$, the proposed truncated policy is shown to be asymptotically optimal for the primal *B&P-Constrained AoI* problem with hard bandwidth constraint.

Proof: The detailed proof will be provided in Appendix C.   ∎

VI. SIMULATIONS

In this section, we provide simulation results to demonstrate the performance of the proposed scheduling policy. We consider a $Q = 4$ states channel with identical evolution matrix:

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.25 & 0.3 & 0.25 & 0.2 \\ 0.2 & 0.25 & 0.3 & 0.25 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{bmatrix}.$$

We assume all the sensors have the same above evolving channels and the steady state distribution of channel states is $\boldsymbol{\eta} = [0.25, 0.25, 0.25, 0.25]$. The following simulation results are obtained over a consecutive of $T = 10^6$ slots.

Notice that from [18], the optimal policy to minimize AoI performance when all the sensors are identical is a greedy policy that selects the sensor with the largest AoI. If there is no packet-loss in the network, the greedy policy is equivalent to round robin, which requires a minimum power consumption of $\mathcal{E}^{\text{RR}} = \frac{M}{N} \sum_{q=1}^Q \eta_q \omega(q)$ for each sensor. In the following simulations, we measure power consumption constraint through ratio $\rho_n = \mathcal{E}_n / \mathcal{E}^{\text{RR}}$. Small ρ indicates that the corresponding sensor has a smaller amount of average power budget.

A. Average AoI performance

Fig. 4 studies average AoI performance as a number of sensors with fixed bandwidth $M = \{2, 5\}$. The power constraint

factor is taken from $[0.2, 1.6]$ and $\rho_n = 0.2 + \frac{1.4}{N}(n-1)$. Denote $C_n(t)$ as the total power consumed by sensor n until slot t and let $\mathcal{R}(t) = \{n | \mathcal{E}_n t - C_n(t) \geq 0\}$ be the set of sensors that has enough power to support transmission in slot t. We compare the proposed policy with a naive greedy policy that selects no more than M sensors with the largest AoI from set $\mathcal{R}(t)$ for scheduling. As can be seen from the figure, the proposed truncated scheduling achieves a close average AoI performance to the lower bound. While the available bandwidth keeps a constant but the number of sensors increases, the proposed truncated policy achieves nearly 40% average AoI decrease for $M = \{2, 5\}$ in a network with $N = 50$ sensors.
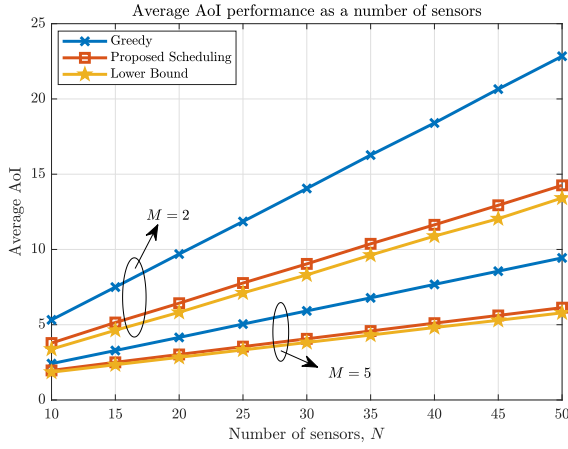

Fig. 4. Average AoI performance as a number of sensors N, $M = \{1, 2\}$.

Fig. 5 studies the asymptotic average AoI performance as a number of sensors, with $\frac{M}{N} = \{\frac{1}{5}, \frac{1}{8}\}$. The power constraint of each sensor is selected by $\rho_n = 0.2 + \frac{1.4}{N}(n-1)$. As can be observed from the figure, the difference between the proposed strategy and the lower bound decreases with N. The asymptotic performance is also verified in simulation results.
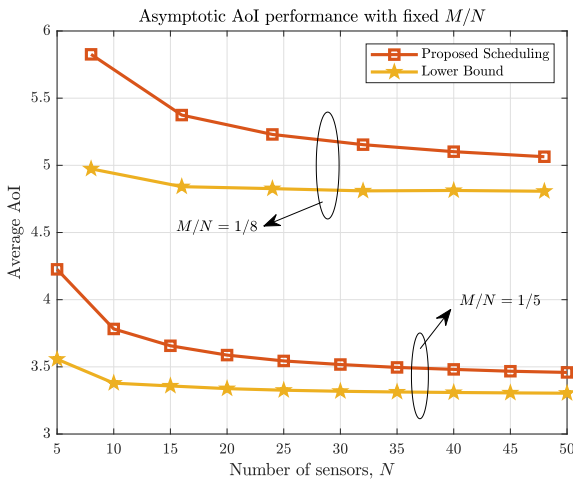

Fig. 5. Asymptotic average AoI performance as a number of sensors N, available bandwidth is chosen by $M/N = \{\frac{1}{5}, \frac{1}{8}\}$.

B. AoI-power trade-off and threshold structure

Fig. 6 plots the average AoI-power tradeoff curves for different number of sensors $N = \{4, 8, 16\}$, each sensor has identical channel fading characteristic and the same power constraint factor ρ. We assume $M = 1$, i.e., only one sensor can be scheduled in each slot. Since all the sensors are identical, it can be concluded that the average scheduling probability of each sensor is smaller than $\frac{1}{N}$. Hence, we can fix $W = 0$ and add another constraint on the activation probability to the LP (20),

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^{Q} y_{x,q} \leq \frac{1}{N}.$$

By solving this LP problem, we can obtain an lower bound on AoI performance for scheduling multiple identical power constrained sensors. The optimal average AoI performance with no power consumption constraint is plotted in green dashed lines. The yellow solid lines depict AoI obtained by solving the relaxed scheduling problem and red squares represent the AoI performance obtained through the proposed truncated scheduling policy. From the figure, average AoI by following the proposed truncated scheduling policy is close to the AoI lower bound. The average AoI performance decreases monotonically with the power consumption constraint. When ρ is near 1, indicating each sensor tends to have enough power to carry out a round robin strategy, AoI performance obtained by the proposed truncated scheduling policy and the AoI lower bound also approach the optimal performance by round robin where there is no power constraint. When ρ approaches zero, the average AoI increases dramatically and approaches infinity.
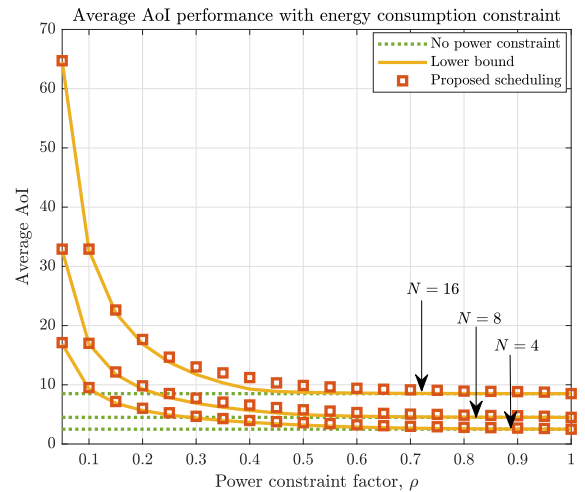

Fig. 6. Average AoI-power tradeoff curves for different number of identical sensors.

Inspired by the AoI decrease observed in Fig. 4, we then study the average AoI performance of different power constrained sensor in Fig. 7 and visualize the scheduling decisions in Fig. 8. We consider a network with $N = 8$ sensors and $M = 2$, each sensor has a power constraint factor $\rho_n = 0.2n$. As is observed from Fig. 7, the proposed algorithm brings

about 40% AoI decrease for the first two sensors, which have very limited power for transmission ($\rho_1 = 0.2, \rho_2 = 0.4$). The AoI deduction of the proposed algorithm is achieved partly through a more reasonable transmission opportunity allocation to sensors with very limited power. For sensors that have enough power, i.e., sensor 7 and 8 with $\rho_7 = 1.4, \rho_8 = 1.6$, our proposed policy guarantees timely updates from those sensors and thus they show similar AoI performance in simulations.
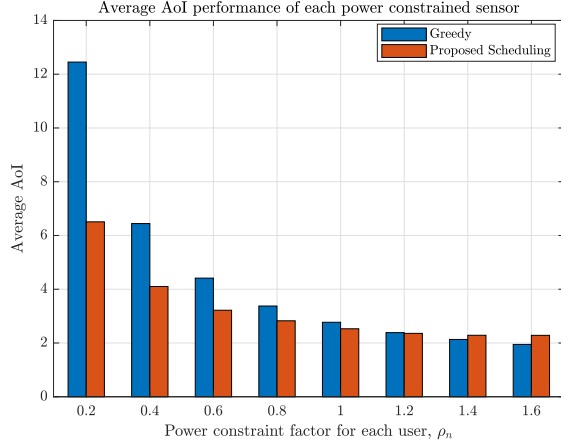

Fig. 7. Average AoI performance of each power constrained sensor in a network with $N = 8$ sensors and $M = 2$, $\rho_n = 0.2n$ and $M = 2$.

We visualize the scheduling policy for some representative sensors in Fig. 8, where (a)-(d) demonstrate sensor $\{1, 2, 7, 8\}$ with power constraint $\rho = \{0.2, 0.4, 1.4, 1.6\}$, respectively. The optimal scheduling decision for single sensor with power consumption constraint $\rho = \{0.2, 0.4, 1.4, 1.6\}$ but no bandwidth constraint are plotted in (e)-(h). In Fig. 8(a) and (b), the transmission power for each sensor is limited, the scheduling threshold τ_q is a increasing sequence as channel state q. Moreover, the threshold of each channel stated in Fig. 8(b) is smaller than corresponding threshold in Fig. 8(a), indicating that transmission is more likely to happen as a result of more available transmission power. In (a) and (b), the difference between the activation thresholds τ_q for each sensor is smaller compared with the difference between thresholds illustrated in (e) and (f), indicating the scheduler tries to maintain total probability of sensor scheduling small in order to satisfy the bandwidth constraint of the entire network. Thus, scheduling strategy for a single power constrained sensor seeks to exploit a good channel state, while trying to keep AoI small and use less bandwidth. If unfortunately the channel state is always bad, he will keep waiting until data staleness cannot be bare anymore or the channel state turns good. By comparing Fig. 8(a) and (b), the scheduler tries to make full use of the transmission power through a refinement of activation thresholds. When $\rho = 1.4$ and $\rho = 1.6$, power consumption is not a problem, all the channel states share the same activation threshold. The threshold is set in order to satisfy the relaxed bandwidth constraint. Considering the AoI performance depicted in Fig. 7, although greedy algorithm attempts to use up the power and brings smaller AoI performance to sensors equipped with enough power, it fails to exploit good chan-

nels and opportunistically schedule those power constrained sensors, hence lead to a much higher AoI performance. Thus, for a network with different power constrained sensors, the scheduling strategy for different sensors varies according to their power constraints. The scheduler seeks good channels to carry out scheduling decisions for those power constrained sensors, while sensors supported by enough power are updated in a timely manner that can satisfy bandwidth constraint.

VII. CONCLUSIONS

In this work, we investigate into the problem of age minimization scheduling in power constrained wireless networks, where communication channels are modeled to be an ergodic Markov chain and different level of transmission power is adopted to ensure successful transmission. We decouple the multi-sensor scheduling problem into a single sensor level constrained Markov decision process. We reveal the threshold structure of the optimal stationary randomized policy for the single sensor and convert the optimal scheduling problem into a linear programming. A truncated scheduling policy that satisfies the hard bandwidth constraint is proposed based on the solution to each decoupled sensor. It is revealed that when power of the sensor is very limited, the scheduler seeks to exploit a good channel state while keeping the information fresh. Sensors equipped with enough power are updated in a timely manner that can satisfy the hard bandwidth constraint.

As far as we know, this is the first work to explore the cross-layer scheduling policy to minimize AoI in bandwidth limited networks with average power constrained sensors. In the future, we will extend the work to more general scenarios. For example, when the update packet of each sensor arrive stochastically or each of the sensors are supported by random power charges.

APPENDIX A
PROOF OF LEMMA 1

Proof: The threshold structure of the optimal policy that minimizes the average cost of (14) is proved by insights from the α-discounted cost problems, where $0 < \alpha < 1$ is a discount factor. Given state (x, q), the expected α-discounted cost starting from the state over infinite horizons by following policy π can be computed:

$$J_{\alpha,\pi}(x, q) = \lim_{T \to \infty} \mathbb{E}_\pi \{ \sum_{t=0}^{T} \alpha^t [C_X(x(t), q(t), s(t))$$
$$+ \lambda C_Q(x(t), q(t), s(t))] | (x(0) = x, q(0) = q) \}. \quad (34)$$

Let $V_\alpha(x, q) = \min_{\pi \in \Pi_{\text{NA}}} J_{\alpha,\pi}(x, q)$ be the minimum expected total discounted cost starting from state (x, q). Then, the minimum total discounted cost will satisfy the following equation:

$$V_\alpha(x, q) = \min\{C_X(x, q, 0) + \alpha \sum_{q'=1}^{Q} p_{q,q'} V_\alpha(x+1, q'),$$

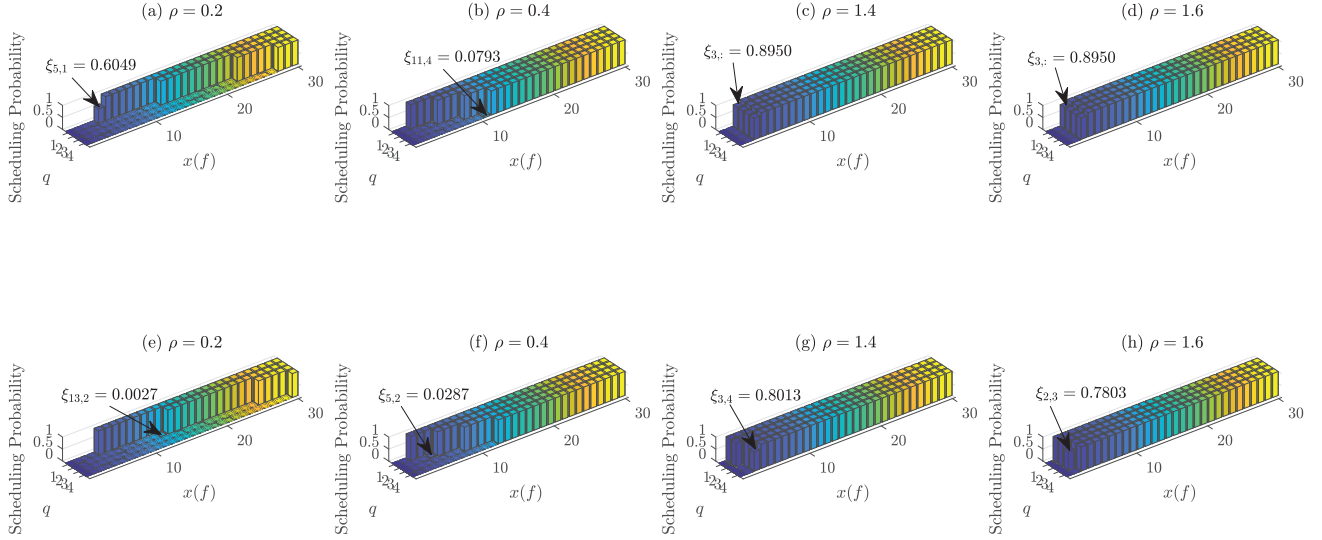$$C_X(x, q, 1) + \lambda C_Q(x, q, 1) + \alpha \sum_{q'=1}^{Q} p_{q,q'} V_\alpha(1, q')\}. \quad (35)$$

Fig. 8. (a)-(d): Scheduling decisions for sensors with different power constraint ρ_n in a network with $N = 8$ sensors and $M = 2$. (e)-(h): Scheduling decisions for single sensor with different power constraint ρ_n and with no bandwidth constraint.

To verify the threshold structure of the optimal policy to the total discounted cost problem, we will introduce the following characteristic of $V_\alpha(x, q)$:

Lemma 2: For given discount factor α and fixed channel state q, the value function $V_\alpha(\cdot, q)$ increases monotonically with x.

The details of the proof will be given in Appendix B. With this lemma, let us now verify the threshold structure. Denote $\Delta(x, q)$ to be the difference in value function by taking $a = \{0, 1\}$, i.e.,

$$\Delta(x, q) = C_X(x, q, 0) + \alpha \sum_{q'=1}^{Q} p_{q,q'} V_\alpha(x+1, q')$$

$$- C_X(x, q, 1) - \lambda C_Q(x, q, 1) - \alpha \sum_{q'=1}^{Q} p_{q,q'} V_\alpha(1, q')$$

$$= \alpha \sum_{q'=1}^{Q} p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda\omega(q) + W)$$

$$\tag{36}$$

If the optimal policy $s_\alpha^*(x, q) = 1$, i.e, it is better to schedule the sensor at state (x, q), by substituting Eqn. (12a) into $\Delta(x, q) \geq 0$, we can obtain the following inequality:

$$\alpha \sum_{q'=1}^{Q} p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda\omega(q) + W) \geq 0. \tag{37}$$

According to Lemma 2, the value function $V_\alpha(\cdot, q)$ is monotonic increasing. Hence, for any $x' > x$, $\Delta(x', q)$ can be lower bounded by:

$$\Delta(x', q)$$

$$= \alpha \sum_{q'=1}^{Q} p_{q,q'} (V_\alpha(x'+1, q') - V_\alpha(1, q')) - (\lambda\omega(q) + W)$$

$$\overset{(a)}{\geq} \alpha \sum_{q'=1}^{Q} p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda\omega(q) + W) \geq 0,$$

$$\tag{38}$$

where inequality (a) is obtained because $V(\cdot, q)$ is increasing. The positivity of $\Delta(x', q)$ implies that for state $x' > x$, the optimal policy for state (x', q) is to schedule the sensor. If at state (x, q) the optimal policy is to be passive, then for state $x' < x$, the optimal policy satisfies $s^*(x', q) = 0$ can be verified similarly.

Moreover, for any state q, according to the Bellman equation, the difference between the expected total discounted cost for keeping idle and being scheduled can be computed by

$$(C_X(x, q, 0) + \alpha\mathbb{E}[V_\alpha(x+1, q')])$$

$$- (C_X(x, q, 0) + \alpha\mathbb{E}[V_\alpha(1, q')])$$

$$\geq x + \alpha(x+1) - (x + \alpha\mathbb{E}[V_\alpha(1, q')])$$

$$= \alpha x + \alpha - \alpha\mathbb{E}[V_\alpha(1, q')] \tag{39}$$

which increases linearly with x. Hence for any channel state, there must be some state such that inequality (37) is satisfied. This suggests that the optimal solution cannot keep passive all the time. Thus, there exists a threshold τ_q for any state $x > \tau_q$, the optimal policy $s_\alpha^*(x, q) = 1$ and for state $x < \tau_q$, $s_\alpha^*(x, q) = 0$.

Finally, we present the generation of the threshold structure for total discounted cost to establish the structure of the average cost. Take a sequence of discount factors such that $\lim_{k\to\infty} \alpha_k = 1$. Then according to [38], the optimal policy $s_{\alpha_k}^*$ for minimizing the total α_k-discounted cost converges to the policy for minimizing the time-average cost, which verifies the threshold structure of the optimal policy s^* as stated in Lemma 1. ∎

APPENDIX B
PROOF OF LEMMA 2

Proof: In this section, we aim at verifying the monotonic characteristic of the discounted value function. The value of $V_\alpha(x, q)$ can be computed through value iteration regarding the Eqn. (35). Denote $V_\alpha^{(k)}(x, q)$ to be the value function obtained after the kth iteration, the monotonic characteristic is proved by induction.

Suppose $V_\alpha^{(k)}(\cdot, q)$ and $V_\alpha^{(k)}(x, \cdot)$ are non-decreasing. With no loss of generality, suppose $x_1 < x_2$. According to the one step cost, we have:

$$C_X(x_1, q, s) < C_X(x_2, q, s), C_Q(x_1, q, s) = C_Q(x_2, q, s). \tag{40}$$

Denote $J_{\alpha,s}^{(k)}(x, q)$ to be the expected total discounted cost if take action s. Then we have the following inequality:

$$
\begin{aligned}
&J_{\alpha,0}^{(k)}(x_1, q) \\
&= C_X(x_1, q, 0) + \alpha \sum_{q'=1}^{Q} p_{q,q'} V_\alpha^{(k)}(x_1, q) \\
&\overset{(a)}{<} C_X(x_2, q, 0) + \alpha \sum_{q'=1}^{Q} p_{q,q'} V_\alpha^{(k)}(x_2, q) \\
&= J_{\alpha,0}^{(k)}(x_2, q),
\end{aligned} \tag{41}
$$

where inequality (a) is obtained because of the monotonic characteristic of $V_\alpha^{(k)}(\cdot, q)$. Similarly, we will have the conclusion that $J_{\alpha,1}^{(k)}(x_1, q) < J_{\alpha,1}^{(k)}(x_2, q)$. Notice that the value function obtained in the $(k+1)^{\text{th}}$ iteration is obtained by:

$$V_\alpha^{(k+1)}(x, q) = \min_s J_{\alpha,s}^{(k)}(x, q),$$

and for any s, $J_{\alpha,s}^{(k)}(x_1, q) < J_{\alpha,s}^{(k)}(x_2, q)$. Thus, the value function $V_\alpha^{(k+1)}(x_1, q) < V_\alpha^{(k+1)}(x_2, q)$. By letting $k \to \infty$, the value function $V_\alpha^{(k)}(x, q) \to V_\alpha(x, q)$. Hence, $V_\alpha(\cdot, q)$ is monotonic increasing. ∎

APPENDIX C
PROOF OF THEOREM 2

Proof: Denote π_R^* be the policy that in each slot, schedule all the sensors with $s_n(t) = 1$ and let $\tilde{\pi}$ be the truncated policy described in Sec. V(B). Since π_R^* is the optimum performance to the *RB&P-Constrained AoI* problem, which formulates the lower bound on the primal *B&P-Constrained AoI* problem. We verify the asymptotic optimality of the proposed scheduling algorithm by computing the expected AoI difference obtained by π_R^* and $\tilde{\pi}$.

First, considering that π_R^* satisfy the relaxed constraint, the average number of sensors that are eager to send updates by following policy π_R^* can then be bounded:

$$\overline{\Omega} = \mathbb{E}[|\Omega(t)|] \leq M. \tag{42}$$

According to Lemma 1 and Corollary 2, the optimum policy to each decoupled single-sensor optimization problem possesses a threshold structure. Let $\Gamma_n = \max_q \tau_{n,q} - \min_q \tau_{n,q}$

be the difference between the largest and the smallest scheduling thresholds of sensor n in different channel states. Suppose in slot t, $s_n(t) = 1$ but sensor n is not scheduled. If in the next slot, the channel evolves into channel with higher scheduling threshold, then the sensor is likely not to be scheduled; or in the next slot there are still more than M sensors that eager to be scheduled and only M sensors can be selected randomly. Due to the majorization characteristic of channel states evolution in Eqn. (3), let $\gamma_n = \sum_{k=1}^{Q-1} P_{1,k}^{(n)} = 1 - P_{1,Q}^{(n)}$ to be the maximum probability that the sensor n's channel state evolves into a better one. Then with probability less than $z_n = (1 - \gamma_n) + \gamma_n \frac{N-M}{N}$, sensor n is still not scheduled. Let $z = \max_n z_n$. Thus, the probability that sensor n is still not scheduled in the next t' slots is upper bounded by $z^{(t'-\Gamma_n)^+}$, where $(\cdot)^+ = \max\{\cdot, 0\}$.

Next, we upper bound the effect of truncating in each slot by introducing a modified version of the truncated strategy $\hat{\pi}_R^*$. Based on the relaxed scheduling strategy π_R^*, when $|\Omega(t)| > M$, the new truncated strategy $\hat{\pi}_R^*$ is designed by: instead of not scheduling a sensor because of limited bandwidth constraint, schedule it as π_R^*, but add a penalty $\sum_{t'=0}^{\infty} z^{(t'-\Gamma_n)^+} x_n(t) = (\Gamma_n + \frac{1}{1-z}) x_n(t)$ on the total AoI. The AoI obtained by $\hat{\pi}_R^*$ will not decrease compared with $\tilde{\pi}$. Let $x_n(t)$ be the AoI obtained by π_R^* and $\mathbb{1}_{(\cdot)}$ be the indicator function, then the difference between $J(\tilde{\pi})$ and $J(\pi_R^*)$ can be upper bounded as follows:

$$
\begin{aligned}
&(J(\tilde{\pi}) - J(\pi_R^*)) \\
&\leq (J(\hat{\pi}_R^*) - J(\pi_R^*)) \\
&= \frac{1}{NT} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{n=1}^{N} (\Gamma_n + \frac{1}{1-z}) x_n(t) \mathbb{1}_{(|\Omega(t)|-M)>0}\right] \\
&\overset{(a)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{NT} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{n=1}^{N} x_n(t) (|\Omega(t)| - M)^+\right] \\
&\overset{(b)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{NT} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{n=1}^{N} x_n(t) (|\Omega(t)| - \overline{\Omega})^+\right] \\
&\overset{(c)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{NT} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{n=1}^{N} x_n(t) \left||\Omega(t)| - \overline{\Omega}\right|\right] \\
&\overset{(d)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{1-z}}{NT} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{n=1}^{N} \max_q \tau_{n,q} ||\Omega(t)| - \overline{\Omega}|\right] \\
&= \frac{(\max_n \Gamma_n + \frac{1}{1-z}) \sum_{n=1}^{N} \max_q \tau_{n,q}}{N} \mathbb{E}\left[\frac{1}{T} \sum_{t=1}^{T} ||\Omega(t)| - \overline{\Omega}|\right],
\end{aligned} \tag{43}
$$

where inequality (a) is obtained because $\mathbb{1}_{|\Omega(t)|-M} \leq (|\Omega(t)|-M)^+$, (b) is because inequality (42) and (c) is because $(\cdot)^+ \leq |\cdot|$. Inequality (d) is obtained because following the relaxed strategy π_R^*, each decoupled sensor has a set of activation thresholds, hence the AoI $x_n(t)$ cannot exceeds the largest thresholds $\max_q \tau_{n,q}$.

Finally, according to [39], the expectation of $||\Omega(t)| - \overline{\Omega}|$ satisfies:

$$\mathbb{E}[||\Omega(t)| - \overline{\Omega}|] = \mathcal{O}(\frac{1}{\sqrt{N}}). \tag{44}$$

Notice that the for sensors with fixed power constraint $\mathcal{E}_n$, the difference of threshold structure Γ_n does not grow with the number of sensors in the network N. Moreover, $\frac{1}{1-z} = \max_n \frac{1}{\gamma_n \theta}$, which only depends on θ and each sensor's channel state evolution characteristic, and hence do not grow with the number of sensors N. In addition, $\frac{N}{M} = \theta$ suggests the available bandwidth M will grow with the number of sensors N, thus the thresholds $\max_q \tau_{n,q}$ will not grow with N. Hence, the average of all thresholds:

$$\frac{1}{N} \sum_{n=1}^{N} \max_q \tau_{n,q} = \mathcal{O}(1). \tag{45}$$

By combining inequalities Eqn. (43)-(45), we will have the following upper bound:

$$J(\tilde{\pi}) - J(\pi_R^*) = \mathcal{O}(\frac{1}{\sqrt{N}}). \tag{46}$$

Considering that $J(\pi_R^*)$ is lower bounded by the performance of round robin policy $J(\pi^{\mathrm{RR}}) \geq \frac{1}{2}(\frac{N}{M}+1)$, which has no power consumption constraint. With $\frac{N}{M} = 1/\theta$ is a constant and let $N \to \infty$, we can lower bound $J(\pi_R^*)$ by:

$$J(\pi_R^*) \geq J(\pi^{\mathrm{RR}}) = \frac{1}{2}(\frac{1}{\theta}+1). \tag{47}$$

Finally, the asymptotic optimum performance of the proposed policy $\tilde{\pi}$ can be verified:

$$\frac{J(\tilde{\pi}) - J(\pi_R^*)}{J(\pi_R^*)} = \mathcal{O}(\frac{1}{\sqrt{N}}). \tag{48}$$

∎

REFERENCES

[1] H. Tang, J. Wang, L. Song, and J. Song, "Scheduling to minimize age of information in multi-state time-varying networks with power constraints," in *submitted to 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*.

[2] X. Jiang, H. Shokri-Ghadikolaei, G. Fodor, E. Modiano, Z. Pang, M. Zorzi, and C. Fischione, "Low-latency networking: Where latency lurks and how to tame it," *Proceedings of the IEEE*, vol. 107, no. 2, pp. 280–306, Feb 2019.

[3] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 766–773, 2016.

[4] C. Chau, F. Qin, S. Sayed, M. H. Wahab, and Y. Yang, "Harnessing battery recovery effect in wireless sensor networks: Experiments and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1222–1232, Sep. 2010.

[5] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *2012 Proceedings IEEE INFOCOM*, March 2012, pp. 2731–2735.

[6] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, March 2019.

[7] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "On the age of information with packet deadlines," *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6419–6428, Sep. 2018.

[8] R. Devassy, G. Durisi, G. C. Ferrante, O. Simeone, and E. Uysal, "Reliable transmission of short packets through queues and noisy channels under latency and peak-age violation guarantees," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 4, pp. 721–734, April 2019.

[9] A. Ephremides and B. Hajek, "Information theory and communication networks: an unconsummated union," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2416–2434, Oct 1998.

[10] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *2015 IEEE International Symposium on Information Theory (ISIT)*, June 2015, pp. 3008–3012.

[11] Y. Sun, E. Uysal-Biyikoglu, R. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, April 2016, pp. 1–9.

[12] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, Nov 2017.

[13] A. Arafa, J. Yang, and S. Ulukus, "Age-minimal online policies for energy harvesting sensors with random battery recharges," in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–6.

[14] J. Yang and J. Wu, "Optimal transmission for energy harvesting nodes under battery size and usage constraints," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 819–823.

[15] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid arq under a resource constraint," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018, pp. 1–6.

[16] A. Baknina, S. Ulukus, O. Oze, J. Yang, and A. Yener, "Sening information through status updates," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 2271–2275.

[17] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2016, pp. 844–851.

[18] 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, Dec 2018.

[19] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2017, pp. 486–493.

[20] R. Talak, I. Kadota, S. Karaman, and E. Modiano, "Scheduling policies for age minimization in wireless networks with unknown channel state," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 2564–2568.

[21] R. Talak, S. Karaman, and E. Modiano, "Optimizing age of information in wireless networks with perfect channel state information," in *2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2018, pp. 1–8.

[22] Y. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 561–565.

[23] Z. Jiang, B. Krishnamachari, X. Zheng, S. Zhou, and Z. Niu, "Decentralized status update for age-of-information optimization in wireless multiaccess channels," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 2276–2280.

[24] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. Mobihoc '18. New York, NY, USA: ACM, 2018, pp. 191–200.

[25] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, April 2018, pp. 1–9.

[26] V. S. Borkar, G. S. Kasbekar, S. Pattathil, and P. Y. Shetty, "Opportunistic scheduling as restless bandits," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1952–1961, Dec 2018.

[27] K. Chen and L. Huang, "Timely-throughput optimal scheduling with prediction," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2457–2470, Dec 2018.

[28] M. Wang, J. Liu, W. Chen, and A. Ephremides, "On delay-power tradeoff of rate adaptive wireless communications with random arrivals," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Dec 2017, pp. 1–6.

[29] ——, "Joint queue-aware and channel-aware delay optimal scheduling of arbitrarily bursty traffic over multi-state time-varying channels," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 503–517, Jan 2019.

[30] J. Yang and S. Ulukus, "Delay-minimal transmission for average power constrained multi-access communications," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2754–2767, Sep. 2010.

[31] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487–499, Aug 2002.

[32] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.

[33] R. Singh and P. R. Kumar, "Throughput optimal decentralized scheduling of multihop networks with end-to-end deadline constraints: Unreliable links," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 127–142, Jan 2019.

[34] E. Jorswieck, H. Boche *et al.*, "Majorization and matrix-monotone functions in wireless communications," *Foundations and Trends® in Communications and Information Theory*, vol. 3, no. 6, pp. 553–701, 2007.

[35] R. D. Yates, P. Ciblat, A. Yener, and M. Wigger, "Age-optimal constrained cache updating," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 141–145.

[36] E. Altman, *Constrained Markov decision processes*. CRC Press, 1999, vol. 7.

[37] D. P. Bertsekas and A. Scientific, *Convex optimization algorithms*. Athena Scientific Belmont, 2015.

[38] L. I. Sennott, "Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs," *Operations Research*, vol. 37, no. 4, pp. 626–633, 1989.

[39] P. Diaconis and S. Zabell, "Closed form summation for classical distributions: variations on a theme of de moivre," *Statistical Science*, pp. 284–302, 1991.