

Minimizing Age of Information With Power Constraints: Multi-User Opportunistic Scheduling in Multi-State Time-Varying Channels

Haoyue Tang¹, Student Member, IEEE, Jintao Wang², Senior Member, IEEE,
Linqi Song³, 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 average AoI 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 time-varying and scheduling decisions are restricted to both bandwidth and power consumption constraint, we first decouple the multi-sensor scheduling problem into a single-sensor constrained Markov decision process (CMDP) by relaxing 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 average AoI over the network: (1) The scheduler exploits good channels to schedule sensors supported by limited power; (2) 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.

Manuscript received July 16, 2019; revised January 14, 2020; accepted February 26, 2020. Date of publication March 16, 2020; date of current version May 7, 2020. This work was supported in part by the National Key Research and development Program of China under Grant 2017YFE0112300, in part by the Beijing National Research Center for Information Science and Technology under Grant BNR2019RC01014 and Grant BNR2019TD01001, in part by Tsinghua University Tutor Research Fund, in part by Shanghai Municipal Science and Technology Major Project (Grant No. 2018 SHZDZX04), and in part by Hong Kong RGC ECS under Grant 9048149. This work has been presented in part in 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton) [1]. (Corresponding author: Jintao Wang.)

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

Linqi Song is with the Department of Computer Science, City University of Hong Kong, Hong Kong (e-mail: linqi.song@cityu.edu.hk).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSAC.2020.2980911

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 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. A 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 are especially challenging since the performance is affected by fundamental trade-off between communication throughput and transmission delay.

Moreover, combating the time-varying nature of wireless 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, known as “the curse of dimension”, creates obstacles in searching for the optimal policy.

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,

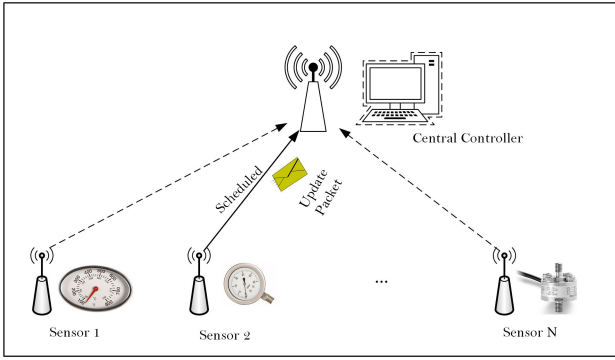


Fig. 1. Illustration of a typical industrial Internet of Things (IIoT) network.

as depicted in Fig. 1. The goal is to understand the how to design AoI minimization strategies in time-varying wireless channel for 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 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.

Main features of the model are: wireless channels connecting sensors and central receivers are time-varying, and information collected by the sensors is time-sensitive. We generalize our previous work [1] by assuming the channel evolution has Markov property, 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 that meets both power and bandwidth constraint, so that the expected average AoI over the entire network can be minimized. Based on 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:

- Consider that fresh update packet can be transmitted at every transmission, we propose a cross-layer framework to study AoI minimization scheduling in multi-user bandwidth limited network with power constrained sensors. The channel is modeled to be a finite-state ergodic Markov chain and remains constant in each slot. Different amount of transmit power is used in different channel state to ensure successful packet transmission. Unlike previous work, we consider both power and bandwidth constraint in a multi-user setup. This model captures

key features of practical cross-layer network optimization problem and facilitates analysis.

- We decouple the multi-sensor scheduling problem into a single-sensor constrained Markov decision process (CMDP) by relaxing the hard bandwidth constraint and 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). This approach has not been used in AoI problems before.
- Finally we search for the Lagrange multiplier so that the bandwidth constraint can be satisfied through dual method and obtain the optimum solution that satisfies the relaxed bandwidth constraint. Then, we propose an asymptotic optimum truncated scheduling policy based on the optimum solution to the relaxed problem so that the hard bandwidth constraint of the network can be satisfied. The performance of the algorithm is analyzed theoretically and verified through simulations.

The remainder of this paper is organized as follows. We review some related work in Section 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, respectively. 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 system 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]–[26]. 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 [24]. 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 channels with two states are considered, and centralized and decentralized policies to minimize AoI are proposed respectively.

Cross-layer control strategy to minimize communication latency under transmit power constraints have been studied in [27]–[34]. In [32], 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 [33]. 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 [29], [30]. 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 [27], [28], [34]. Scheduling policy based on value iteration is proposed in [28] and a Whittle-like index policy to achieve delay-power trade-off is studied in [27]. In [34], the multi-user power and bandwidth constrained scheduling problem is solved by packet level decomposition, and an asymptotically optimum truncated scheduling policy is proposed. Rajat *et. al* studied a joint rate control and scheduling problem for age minimization under general interference constraints [19], where joint rate control and scheduling policies are investigated for age optimality, and a separation principle policy is found to be approximately optimal. However, no power constraint is considered in that work.

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 $t \in \{1, \dots, T\}$ to denote the index of the current 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 . We assume each successful transmission takes one slot and the packet will be received successfully by the end of the slot. Due to limited bandwidth constraint, no more than M sensors can be scheduled in each slot. We consider a non-trivial case and assume the bandwidth $M < N$, thus we have the following constraint on $u_n(t)$:

$$\sum_{n=1}^N u_n(t) \leq M, \quad \forall t. \quad (1)$$

To model the time-varying characteristic of the channel between each sensor and the central controller, we class each channel into Q states and assume the channel state of sensor n , denoted by $\{q_n(t)\}$ is a Q -state ergodic Markov chain with transision probability $p_{i,j}^{(n)} \triangleq \Pr(q_n(t+1) = j | q_n(t) = i)$. If sensor n is scheduled to transmit updates when the current channel state is q , in order to guarantee the channel capacity is larger than the size of an update packet, it will consume $\omega(q)$ units of power. Similar to [27], [30], [32], we assume the transmitted packet will be successfully received by the central controller at the end of the slot. For a typical sheduling

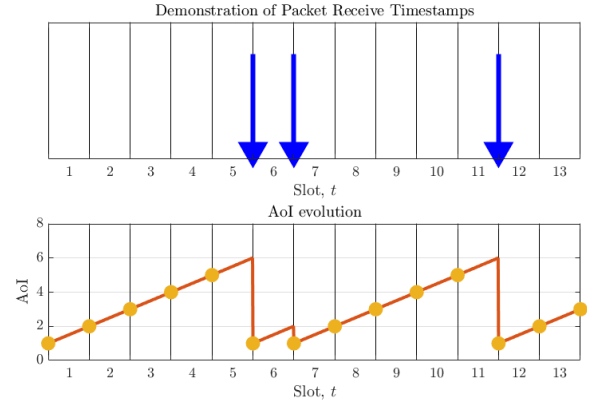


Fig. 2. Illustration of AoI evolution of a specific sensor. On the top, sample sequence representing the receiving time-stamps of the generate-at-will update packets. On the bottom, sample paths of AoI (red). The yellow dots depict the AoI at the beginning of each slot. Upon receiving a new packet, the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

strategy π that assigns $\mathbf{u}_n = [u_n(1), \dots, u_n(T)]$ to sensor n , the average power consumed in T consecutive slots is:

$$E_n(\mathbf{u}_n(\pi)) = \frac{1}{T} \sum_{t=1}^T u_n(t) \omega(q_n(t)). \quad (2)$$

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:

Let $x_n(t)$ be the AoI of sensor n at the beginning of slot t . By definition, it shows the number of slots elapsed since the latest delivery from sensor n is generated. We consider a generate at will model similar to [10], [18] and focus on minimizing the average AoI over the entire network. In this model, update packets generated before slot t will be discarded and the system experiences no queueing delay. Recall that 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 by definition $x_n(t+1) = 1$; 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 AoI $x_n(t)$ evolves as follows:

$$x_n(t+1) = \begin{cases} 1, & u_n(t) = 1; \\ x_n(t) + 1, & u_n(t) = 0. \end{cases} \quad (3)$$

C. Problem Formulation

For a given network setup with N sensors and channel states evolution $\{p_{i,j}^{(n)}\}$, 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 $T \rightarrow \infty$ slots, which can be computed as follows:

$$J(\pi) = \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_\pi \left[\sum_{t=1}^T \sum_{n=1}^N x_n(t) | \mathbf{x}(0) \right] \right\}, \quad (4)$$

where the vector $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)] \in \mathbb{N}^N$ 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 AoIs $\{x_n(t)\}$, channel states $\{q_n(t)\}$ and their evolving probabilities $\{p_{i,j}^{(n)}\}$. No information about the future AoI or channel states can be used. We assume the average power constraint of each sensor is known by the central controller. In this research, we aim at designing policy $\pi \in \Pi_{\text{NA}}$ to minimize the average expected AoI of all the sensors, while the time average power consumption constraint of each sensor can be satisfied. The original bandwidth and power constrained AoI minimization problem (B&P-Constrained AoI) is as follows:

Problem 1 (B&P-Constrained AoI):

$$\pi^* = \arg \min_{\pi \in \Pi_{\text{NA}}} \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_{\pi} \left[\sum_{t=1}^T \sum_{n=1}^N x_n(t) \right] \right\}, \quad (5a)$$

$$\text{s.t. } \mathbb{E}_{\pi} \left[\sum_{n=1}^N u_n(t) \right] \leq M, \quad \forall t, \quad (5b)$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi} \left[\sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \quad \forall n. \quad (5c)$$

Notice that the hard bandwidth constraint (5b) in every slot t suggests, there are $\binom{N}{1} + \dots + \binom{N}{M}$ possible scheduling decisions in each slot, it is hard to approach this problem through dynamic programming. We tackle this challenge through the following approaches:

- Inspired by [28], [34], [35], in Section IV-(A), we first relax the hard bandwidth constraint (5b) 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 (5b) 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 strategy is obtained through linear programming (LP).

A. Sensor Level Decomposition

Let us first relax the hard constraint (5b) 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):

$$\pi_R^* = \arg \min_{\pi \in \Pi_{\text{NA}}} \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_{\pi} \left[\sum_{t=1}^T \sum_{n=1}^N x_n(t) \right] \right\}, \quad (6a)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \mathbb{E}_{\pi} \left[\frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N u_n(t) \right] \leq M, \quad (6b)$$

$$\lim_{T \rightarrow \infty} \mathbb{E}_{\pi} \left[\frac{1}{T} \sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n, \quad \forall n. \quad (6c)$$

Notice that any policy π that satisfies the bandwidth constraint in the *B&P-Constrained AoI* satisfies the bandwidth constraint in *RB&P-Constrained AoI*, hence the expected average AoI following policy π_R^* formulates a lower bound on the expected average AoI obtained by π^* . To solve Problem 2, let us place the relaxed constraint into the objective function:

$$\begin{aligned} \mathcal{L}(\pi, W) &= \lim_{T \rightarrow \infty} \left\{ \frac{1}{NT} \mathbb{E}_{\pi} \left[\sum_{n=1}^N \sum_{t=1}^T \left(x_n(t) + W u_n(t) - \frac{WM}{N} \right) \right] \right\}. \end{aligned} \quad (7)$$

For fixed multiplier W , denote $\pi(W)$ be the optimum policy that minimizes the Lagrange function (7), i.e.,

$$\pi(W) = \arg \min_{\pi \in \Pi_{\text{NA}}} \mathcal{L}(\pi, W). \quad (8)$$

Notice that the optimum policy π_R^* to Problem 2 is a mixture of no more than two policies $\pi(W_1)$ and $\pi(W_2)$, which minimizes the Lagrange function under different multipliers W_1 and W_2 , respectively. Thus, in the following analysis, we will first solve $\pi(W)$ for fixed W and then provide how to obtain the two policies $\pi(W_1)$ and $\pi(W_2)$.

For fixed W , minimizing (7) can then be decoupled into N single sensor AoI and scheduling penalty minimization problem with average power consumption constraint (5c), 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):

$$\pi_{d,n}^* = \arg \min_{\pi \in \Pi_{\text{NA}}} \mathcal{L}(\pi_n, W), \text{ where} \quad (9a)$$

$$\mathcal{L}_n(\pi_n, W) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi_n} \left[\sum_{t=1}^T (x_n(t) + W u_n(t)) \right], \quad (9b)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{\pi_n} \left[\sum_{t=1}^T u_n(t) \omega(q_n(t)) \right] \leq \mathcal{E}_n. \quad (9c)$$

Since minimizing (7) gets decoupled, we omit the subscript n henceforth. We formulate the *Decoupled P-Constrained Cost* minimization problem into an CMDP in Section IV-(B) and analyze the structure of the optimum policy in Section IV-(C). In Section IV-(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 $(\mathcal{S}, \mathcal{A}, \Pr(\cdot|\cdot), C(\cdot, \cdot))$, 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:

$$\Pr((x, q) \rightarrow (x', q')) = \begin{cases} p_{q, q'}, & \{s = 0, x' = x + 1\}, \\ & \text{or } \{s = 1, x' = 1\}, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

- **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:

$$C_X(x, q, s) = x + Ws, \quad (11a)$$

while the one-step power consumption is:

$$C_Q(x, q, s) = \omega(q)s. \quad (11b)$$

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:

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

while the average power constraint is satisfied,

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

C. Characterization of the Optimal Policy

In this part, we focus on exploiting the threshold structure of the optimal policy. 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 $\pi_{\text{SR}} \in \Pi_{\text{SR}}$ chooses action $s(t) = 1$ with probability measure $\xi_{x,q} \in [0, 1]$ for all t . A stationary deterministic policy $\pi_{\text{SD}} \in \Pi_{\text{SD}}$ selects action $s(t) = a(x, q)$, where $a(\cdot) : (x, q) \rightarrow \{0, 1\}$ is a deterministic mapping from state space to action space.

According to [36, Theorem 4.4], the optimal policy to the above CMDP (*Decoupled P-Constrained Cost*) has the following property:

Corollary 1: An optimal stationary randomized policy $\pi_d^* \in \Pi_{\text{SR}}$ exists for the decoupled single sensor power constrained scheduling problem (9b), and it is a mixture of no more than two stationary deterministic policies $\pi_{\text{SD1}}, \pi_{\text{SD2}} \in \Pi_{\text{SD}}$. Let ρ be the weight of following stationary deterministic policy π_{SD1} and $(1 - \rho)$ be the weight of following π_{SD2} . Then the optimum policy is:

$$\pi_d^* = \rho\pi_{\text{SD1}} + (1 - \rho)\pi_{\text{SD2}}. \quad (12)$$

Proof: According to [36, Theorem 6.3] an optimum stationary randomized policy exists for constrained Markov decision process with infinite state and action space. Since the Lagrange relaxation removes only one constraint, according to [36, Theorem 4.4], the optimum policy is a mixture of two policies that minimize the Lagrange function with different multipliers λ_1 and λ_2 . Such derivations is used similarly in [15]. ■

To obtain the two deterministic policies π_{SD1} and π_{SD2} , next we establish an unconstrained MDP by placing the average power consumption constraint into the objective function. Let $\lambda \geq 0$ be the Lagrange multiplier related to the average power constraint, we write out the Lagrange function and the goal of the unconstrained MDP is to minimize the following overall average cost (we omit the constant item $-\lambda\mathcal{E}$):

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=1}^T [C_X(x(t), q(t), s(t)) + \lambda C_Q(x(t), q(t), s(t))] \right]. \quad (13)$$

For given Lagrange multiplier λ , a stationary deterministic policy to minimize the above unconstrained cost exists. Denote γ be the time-average cost by following the optimum strategy. Then, there exists a differential cost-to-go function $V(x, q)$ that satisfies the following Bellman equation:

$$V(x, q) + \gamma = \min \left\{ C_X(x, q, 0) + \sum_{q'=1}^Q p_{q,q'} V(x+1, q'), \right. \\ \left. C_X(x, q, 1) + \sum_{q'=1}^Q p_{q,q'} V(1, q') + \lambda C_Q(x, q, 1) \right\}, \quad (14)$$

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 provides 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 (13) possesses a threshold structure. That is there exists a sequence of threshold τ_q for each state, when $x \geq \tau_q$, the optimal action $s^*(x, q) = 1$ and when $x < \tau_q$, $s^*(x, q) = 0$.

Proof sketch: The proof is provided in Appendix A. Here we provide an intuitive analysis. Since communication between the sensor and the controller is power constrained,

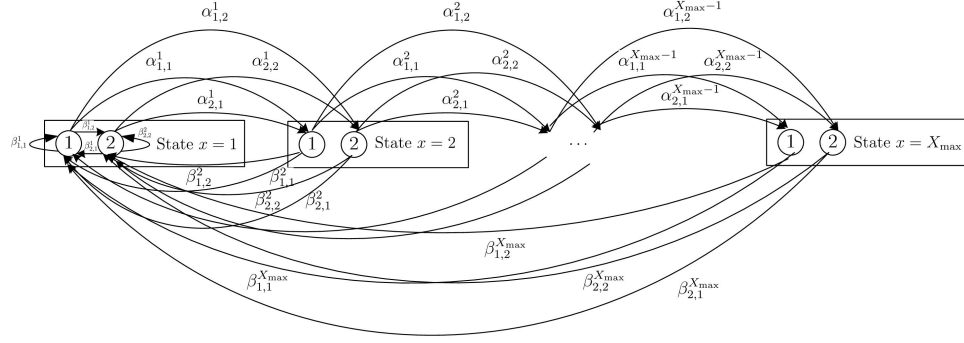


Fig. 3. Illustrative of the probability transfer graph for a stationary randomized policy with $Q = 2$ channel states. The circles denote channel state q and the rectangles denote the sensor's AoI x . The forward state transmission probability (x, q) to $(x + 1, q')$ is $\alpha^x_{q,q'}$, and the backward state transmission probability from (x, q) to $(1, q')$ is $\beta^x_{q,q'}$.

we 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.

Notice that optimal stationary randomized policy π_d^* to the CMDP (9b) is a randomization between no more than two stationary deterministic policies [36], each of them can be obtained by solving the unconstrained MDP (13) which possesses a threshold structure. Then 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 π_d^* schedules the sensor.

D. Probabilistic Scheduling Policy for Single Sensor Case

Let us now investigate 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. From Section IV-(C), since 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, the optimum policy π_d^* must satisfy $\xi_{x,q}^* = 1, \forall (x, q), x \geq \max_q \tau_q$. 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 satisfies $\xi_{x,q} = 1, \forall x \geq X_{\max}$ in the following analysis, since policies that do not have such properties are not optimum and thus 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 $\alpha^x_{q,q'}$ denote the one step forward state transition probability from (x, q) to $(x + 1, q')$ and let $\beta^x_{q,q'}$ 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 $\mu_{x,q} = 0, \forall x \geq X_{\max} + 1, q$. 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:

$$\alpha^x_{q,q'} = \Pr((x, q) \rightarrow (x + 1, q')) = (1 - \xi_{x,q})p_{q,q'}, \quad (15a)$$

$$\beta^x_{q,q'} = \Pr((x, q) \rightarrow (1, q')) = \xi_{x,q}p_{q,q'}. \quad (15b)$$

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

$$\mathbf{Q} = \begin{bmatrix} \beta^1 & \beta^2 & \dots & \beta^{X_{\max}-1} & \beta^{X_{\max}} \\ \alpha^1 & \mathbf{0}_Q & \dots & \mathbf{0}_Q & \mathbf{0}_Q \\ \mathbf{0}_Q & \alpha^2 & \dots & \mathbf{0}_Q & \mathbf{0}_Q \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}_Q & \mathbf{0}_Q & \dots & \alpha^{X_{\max}-1} & \mathbf{0}_Q \end{bmatrix}, \quad (16)$$

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

$$\alpha^x = \begin{bmatrix} \alpha^x_{1,1} & \alpha^x_{1,2} & \dots & \alpha^x_{1,Q} \\ \alpha^x_{2,1} & \alpha^x_{2,2} & \dots & \alpha^x_{2,Q} \\ \dots & \dots & \dots & \dots \\ \alpha^x_{Q,1} & \alpha^x_{Q,2} & \dots & \alpha^x_{Q,Q} \end{bmatrix}, \quad (17a)$$

$$\beta^x = \begin{bmatrix} \beta^x_{1,1} & \beta^x_{1,2} & \dots & \beta^x_{1,Q} \\ \beta^x_{2,1} & \beta^x_{2,2} & \dots & \beta^x_{2,Q} \\ \dots & \dots & \dots & \dots \\ \beta^x_{Q,1} & \beta^x_{Q,2} & \dots & \beta^x_{Q,Q} \end{bmatrix}. \quad (17b)$$

According to property of the steady state distribution, we have $\mathbf{Q}\boldsymbol{\mu} = \boldsymbol{\mu}$. In addition, considering that $\forall x \geq X_{\max} + 1$, the steady state distribution $\mu_{x,q} = 0, \forall q$. 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:

$$\begin{bmatrix} \mathbf{Q} - \mathbf{I}_{QX_{\max}} \\ \mathbf{1}_{QX_{\max}}^T \end{bmatrix} \boldsymbol{\mu} = \begin{bmatrix} \mathbf{0}_{QX_{\max}} \\ 1 \end{bmatrix}, \quad (18)$$

where $\mathbf{1}_{QX_{\max}}$ is a $(Q \times X_{\max})$ -dimension column vector with all the elements being 1 and $\mathbf{I}_{QX_{\max}}$ is a $(Q \times X_{\max})$ dimension 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-Constrained Cost* minimization problem is equivalent to solve the following LP problem:

$$\{\mu_{x,q}^*, y_{x,q}^*\} = \arg \min_{\{\mu_{x,q}, y_{x,q}\}} \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q (W y_{x,q} + x \mu_{x,q}), \quad (19a)$$

$$\text{s.t. } \mu_{1,q} = \sum_{x=1}^{X_{\max}} \sum_{q'=1}^Q y_{x,q'} p_{q',q}, \quad (19b)$$

$$\mu_{x,q} = \sum_{q'=1}^Q (\mu_{x-1,q'} - y_{x-1,q'}) p_{q',q}, \quad (19c)$$

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} = 1, \quad (19d)$$

$$y_{x,q} \leq \mu_{x,q}, \quad (19e)$$

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q y_{x,q} \omega(q) \leq \mathcal{E} \quad (19f)$$

$$0 \leq \mu_{x,q} \leq 1, \quad 0 \leq y_{x,q} \leq 1, \quad \forall x, q. \quad (19g)$$

Proof: Let us compute the equivalent time average cost to (9b) as a sum of $\{\mu_{x,q}\}$ 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 $C_X(x, q, 1) = x + W$; with probability $1 - \xi_{x,q}$ the sensor is selected to keep idle 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:

$$\begin{aligned} & \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} (\xi_{x,q}(x + W) + (1 - \xi_{x,q})x) \\ &= \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q (W y_{x,q} + x \mu_{x,q}). \end{aligned} \quad (20)$$

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:

$$\sum_{x=1}^{X_{\max}} \sum_{q=1}^Q \mu_{x,q} \xi_{x,q} \omega(q) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q y_{x,q} \omega(q). \quad (21)$$

With this equation the power constraint (5c) can be converted in the linear constraint (19f). The constraint (19b)-(19d) can be obtained by substituting $\xi_{x,q}$ with $y_{x,q}$ and $\mu_{x,q}$ with relationship (18). Notice that $\xi_{x,q} \leq 1$, the inequality constraint (19e) can be obtained. ■

Till now, we construct an LP problem to obtain $\mu_{x,q}$ and $y_{x,q}$ by following the optimum stationary randomized policy

that minimizes the total cost with fixed Lagrange multiplier W . Next, the optimal stationary randomized scheduling policy to minimize Lagrange function (9b) can be obtained through the relationship between $\xi_{x,q}$, $\mu_{x,q}$ and $y_{x,q}$. According to the threshold structure of each deterministic policy and (12), we will have the following property on $\xi_{x,q}^*$:

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

$$\xi_{x_1,q}^* \leq \xi_{x_2,q}^*, \quad \forall 1 \leq x_1 < x_2. \quad (22)$$

V. MULTI-SENSOR OPPORTUNISTIC SCHEDULING

In this section, we will provide an algorithm to determine the multiplier W so 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 (5b).

A. Determination of Lagrange Multiplier

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

$$g(W) = \min_{\pi \in \Pi_{\text{NA}}} \mathcal{L}(\pi, W). \quad (23)$$

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

$$\begin{aligned} g(W) &= \frac{1}{N} \sum_{n=1}^N g_n(W) - WM, \text{ where} \\ g_n(W) &= \min_{\pi_n \in \Pi_{\text{NA}}} (\mathcal{L}_n(\pi_n, W)), \text{ s.t. (7c) is satisfied.} \end{aligned} \quad (24)$$

By Theorem 1, the CMDP that minimizes $\mathcal{L}_n(\pi, W)$ is equivalent to an LP, then $g_n(W)$ equals the average cost of the CMDP. Let $\bar{X}_n(W)$ and $\bar{A}_n(W)$ denote the average AoI and the average scheduling probability of sensor n , respectively. Let $\{y_{x,q}^{n,W}\}$ be the solution of sensor n 's LP problem (19) with multiplier W , function $g_n(W)$ can be computed as follows:

$$g_n(W) = \bar{X}_n(W) + W \bar{A}_n(W), \quad (25a)$$

$$\text{where } \bar{X}_n(W) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q x \mu_{x,q}^{n,W}, \quad (25b)$$

$$\bar{A}_n(W) = \sum_{x=1}^{X_{\max}} \sum_{q=1}^Q y_{x,q}^{n,W}. \quad (25c)$$

According to [39], let W^* be the supreme Lagrange multiplier such that policy $\pi(W)$ that minimizes the Lagrange function (8) satisfies the relaxed bandwidth constraint, i.e.,

$$W^* = \sup\{W \mid \sum_{n=1}^N \bar{A}_n(W) \leq M\}.$$

If the bandwidth consumed by policy $\pi(W^*)$ satisfies $\sum_{n=1}^N \bar{A}_n(W^*) = M$, i.e., $\pi(W^*)$ consumes an average bandwidth M . Then the optimum solution π_R^* to problem 2 is just $\pi(W^*)$. Otherwise, π_R^* is a mixture of two policies π_1 and π_2 , which can be obtained by:

$$\pi_1 = \lim_{W \rightarrow W^* -} \pi(W), \quad \pi_2 = \lim_{W \rightarrow W^* +} \pi(W). \quad (26)$$

To search for policy $\pi(W^*)$, π_1 and π_2 , we apply the subgradient descent method. Let $W^{(k)}$ be the Lagrange multiplier used in the k^{th} iteration.

We start with $W^{(0)} = 0$, if $\sum_{n=1}^N \bar{A}_n(W^{(0)}) - M \leq 0$, 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:

$$\text{AoI}_{\text{LB}} = \text{AoI}_R^* = g(0). \quad (27)$$

Otherwise, we adopt an iterative algorithm update. According to [37, Eq. 6.1.1], the subgradient at $W^{(k)}$ can be computed by:

$$\text{d}_W g(W^{(k)}) = \sum_{n=1}^N \bar{A}_n(W^{(k)}) - M. \quad (28)$$

By choosing a set of stepsizes γ_k similar to [15], the multiplier for the k -th iteration can be computed by:

$$W^{(k)} = W^{(k-1)} + \gamma_k \text{d}_W g(W^{(k-1)}). \quad (29)$$

The iteration ends until both $|W^{(k)} - W^{(k-1)}| < \varepsilon$ and $\sum_{n=1}^N \bar{A}_n(W^{(k)}) \leq M$ are satisfied. Suppose the algorithm terminates at the K -th iteration. If $\sum_{n=1}^N \bar{A}_n(W^{(K)}) = M$, then $\pi_R^* = \pi(W^{(K)})$. Otherwise, we proceed to find two policies π_1 and π_2 that constitutes π_R^* in (26). Let W_l and W_u be two Lagrange multipliers chosen from sequence $W^{(k)}$,

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

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

Let $M_l = \sum_{n=1}^N \bar{A}_n(W_l)$ and $M_u = \sum_{n=1}^N \bar{A}_n(W_u)$ be the total bandwidth used with respect to minimize the function (7). Suppose $\{\mu^{n,l}, \mathbf{y}^{n,l}\}$ is the optimizer to sensor n 's LP problem (19a) with multiplier W_l and $\{\mu^{n,u}, \mathbf{y}^{n,u}\}$ is the solution with multiplier W_u . To satisfy the relaxed bandwidth constraint, the optimum distribution $\{\mu^{n,*}, \mathbf{y}^{n,*}\}$ of the relaxed problem is a linear combination of $\{\mu^{n,l}, \mathbf{y}^{n,l}\}$ and $\{\mu^{n,r}, \mathbf{y}^{n,r}\}$, which can be computed as follows:

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

where the mixing coefficient ν can be computed by:

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

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 observes 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 $\xi_{x_n(t), q_n(t)}^{n,*}$, which is 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 $\bar{A}_n(W^{(0)})$, denote the optimizer as $\{\mu^n, \mathbf{y}^n\}$.
- 2: $k \leftarrow 0, M_l \leftarrow 0, M_u \leftarrow 2M, \{\mu^{n,l}, \mathbf{y}^{n,l}\} = \{\mu^{n,u}, \mathbf{y}^{n,u}\} = \{\mu^n, \mathbf{y}^n\}$
- 3: **if** $\sum_{n=1}^N \bar{A}_n(W^{(0)}) - M \leq 0$ **then** \triangleright Relaxed Bandwidth Constraint is satisfied
- 4: $\{\mu^{n,*}, \mathbf{y}^{n,*}\} \leftarrow \{\mu^n, \mathbf{y}^n\}, \forall n$
- 5: **else** \triangleright Search for the Lagrange Multiplier
- 6: **repeat**
- 7: $k \leftarrow k + 1$
- 8: $\text{d}_W g(W^{(k-1)}) \leftarrow \sum_{n=1}^N \bar{A}_n(W^{(k-1)}) - M$
- 9: $W^{(k)} \leftarrow W^{(k-1)} + \gamma_k \text{d}_W g(W^{(k-1)})$
- 10: Solve the corresponding LP (19a)-(19g) for each sensor n and compute $\bar{A}_n(W^{(k)})$, denote the optimizer as $\{\mu^n, \mathbf{y}^n\}$
- 11: **if** $M_l < \sum_{n=1}^N \bar{A}_n(W^{(k)}) \leq M$ **then**
- 12: $M_l \leftarrow \sum_{n=1}^N \bar{A}_n(W^{(k)})$
- 13: $\{\mu^{n,l}, \mathbf{y}^{n,l}\} \leftarrow \{\mu^n, \mathbf{y}^n\}$
- 14: **else if** $M < \sum_{n=1}^N \bar{A}_n(W^{(k)}) \leq M_u$ **then**
- 15: $M_u \leftarrow \sum_{n=1}^N \bar{A}_n(W^{(k)})$
- 16: $\{\mu^{n,u}, \mathbf{y}^{n,u}\} \leftarrow \{\mu^n, \mathbf{y}^n\}$
- 17: **until** $|W^{(k)} - W^{(k-1)}| < \varepsilon$ and $\sum_{n=1}^N \bar{A}_n(W^{(k)}) \leq M$.
- 18: Compute W_l, W_u and M_l, M_u , compute $\nu \leftarrow \frac{M_u - M}{M_u - M_r}$
- 19: $\{\mu^{n,*}, \mathbf{y}^{n,*}\} \leftarrow \nu \{\mu^{n,l}, \mathbf{y}^{n,l}\} + (1 - \nu) \{\mu^{n,u}, \mathbf{y}^{n,u}\}$
- 20: Compute $\{\xi_{x,q}^{n,*}\}$ according to (32)

Denote AoI_{LB} be the AoI lower bound to the primal *B&P-Constrained AoI* and let AoI_R^* be the expected average AoI following policy π_R^* , which can be written out as a function of the optimizer $\{\mu^{n,*}, \mathbf{y}^{n,*}\}$. According to the discussion in Section IV-A, the average AoI by following π_R^* formulates the lower bound to Problem 1. Hence,

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

B. Multi-Sensor Opportunistic Scheduling With Hard Bandwidth Constraint

In this part we construct a truncated policy $\tilde{\pi}$ based on optimal scheduling policy for each of the decoupled sensor and solve the primal *B&P-Constrained AoI* problem. Recall π_R^* is the optimum scheduling strategy to the relaxed problem

and $s_n(t)$ is the corresponding scheduling decision for sensor n in slot t . Denote $\Omega(t) = \{n | s_n(t) = 1\}$ as the set of sensors that are scheduled under policy π_R^* . 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 currently wait 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 $\mathcal{M}(t) \subset \Omega(t), |\mathcal{M}(t)| = M$ 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 $N \rightarrow \infty$ 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.

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 the following evolution matrix, where the j -th element on the i -th row denotes p_{ij} , i.e., the probability that channel state evolves from i to j :

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

We assume all the sensors have the same above evolving channels and the steady state distribution of channel states is $\boldsymbol{\eta} = [0.2368, 0.2632, 0.2632, 0.2368]$. 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 $\rho_n = 0.2 + \frac{1.4}{N-1}(n-1)$. 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

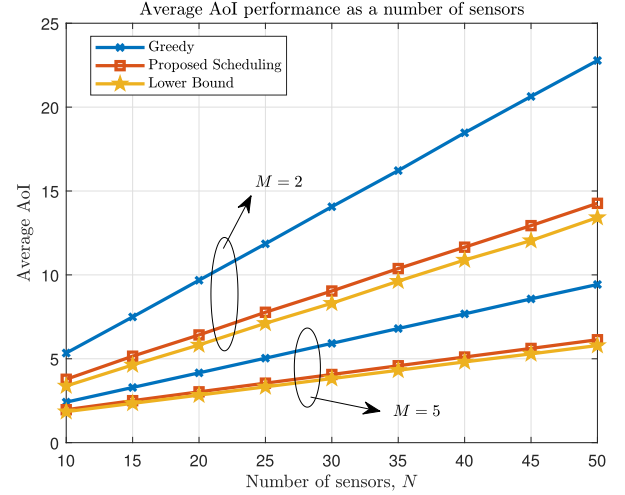


Fig. 4. Average AoI performance as a number of sensors N , $M = \{2, 5\}$.

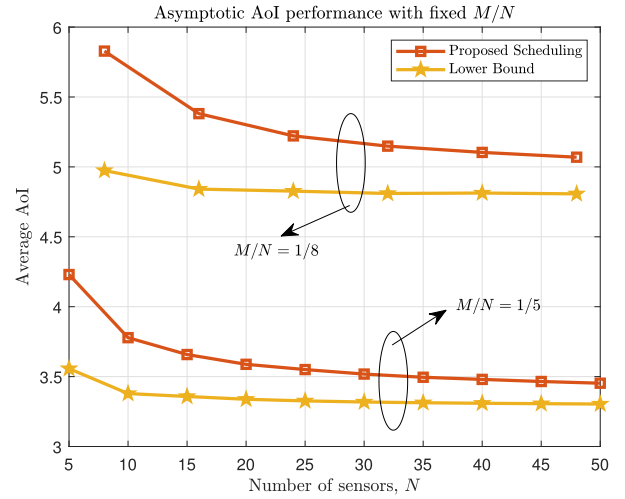


Fig. 5. Asymptotic average AoI performance as a number of sensors N , available bandwidth is chosen by $M/N = \{\frac{1}{5}, \frac{1}{8}\}$.

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. 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 $\rho_n = 0.2 + \frac{1.4}{N-1}(n-1)$. 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.

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

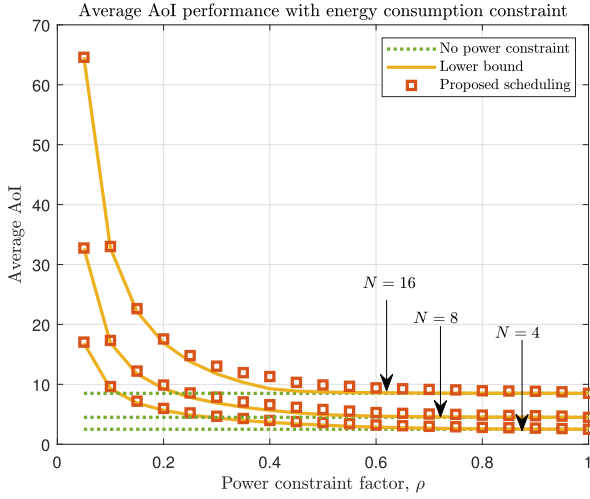


Fig. 6. Average AoI-power tradeoff curves for different number of identical sensors.

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 (19),

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

By solving this LP problem, we can obtain a 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 average AoI performance obtained following the proposed truncated scheduling policy. From the figure, the average AoI following the proposed truncated scheduling policy is close to the AoI lower bound. The average AoI performance decreases monotonically when the power constraint factor increases. 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.

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 $\rho_n = 0.2n$. The average AoI performance of sensor n obtained by the proposed algorithm is denoted by \bar{x}_n . As is observed from Fig. 7, the proposed algorithm brings about 40% AoI decrease to 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 wider bandwidth allocation and more reasonable transmission design

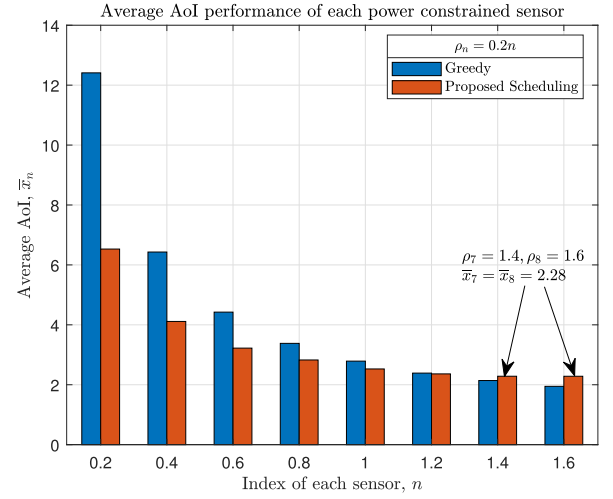


Fig. 7. Average AoI performance of each power constrained sensor in a network with $N = 8$ sensors and $M = 2$, $\rho_n = 0.2n$. The average AoI of sensor n is denoted by \bar{x}_n .

for sensors that have 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.

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 very limited, the scheduling threshold τ_q is an increasing sequence of channel state q . Moreover, the threshold of each channel in Fig. 8(b) is smaller than the 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 the AoI is large 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. By comparing Fig. 8(c) and (g), (d) and (h), when the sensor is equipped with enough power (e.g., $\rho = \{1.4, 1.6\}$), the proposed policy does not use up all the power and all the channel states have the same activation threshold. The threshold is set in order to satisfy the relaxed bandwidth constraint. The bandwidth saved compared with the greedy algorithm is then allocated properly to schedule power constrained sensors and hence achieved significant AoI decrease for those power constrained sensors. Thus, for a network with different power constrained sensors,

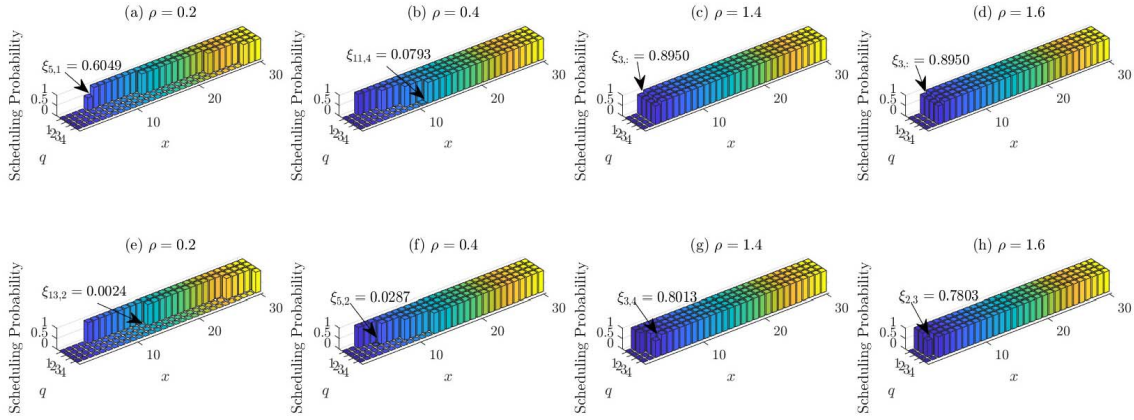


Fig. 8. (a)-(d): Scheduling decisions for sensors with different power constraint ρ_n in a network with $N = 8$ sensors and $M = 2$. (e)-(h): Scheduling decisions for single sensor with different power constraint ρ_n and with no bandwidth constraint.

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. CONCLUSION

In this work, we investigate the problem of age minimization scheduling in power constrained wireless networks, where communication channels are modeled to be ergodic Markov chains and different level of transmission power is used 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 the 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.

The network model considered in this work is a very simplified one. In the future, we will extend the work to more general scenarios. Our method generalizes well when the update packet of each sensor arrive stochastically [43] or packet transmission experiences random packet loss [42]. We will also study scheduling strategy under non-orthogonal multiple access scenario similar to [41].

APPENDIX A PROOF OF LEMMA 1

Proof: The threshold structure of the optimal policy that minimizes the average cost of (13) is proved by induction 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:

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

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

$$V_{\alpha}(x, q) = \min \{ C_X(x, q, 0) + \alpha \sum_{q'=1}^Q p_{q, q'} V_{\alpha}(x+1, q'), \\ C_X(x, q, 1) + \lambda C_Q(x, q, 1) + \alpha \sum_{q'=1}^Q p_{q, q'} V_{\alpha}(1, q') \}. \quad (35)$$

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.,

$$\Delta(x, q) = C_X(x, q, 0) + \alpha \sum_{q'=1}^Q p_{q, q'} V_{\alpha}(x+1, q') \\ - C_X(x, q, 1) - \lambda C_Q(x, q, 1) - \alpha \sum_{q'=1}^Q p_{q, q'} V_{\alpha}(1, q') \\ = \alpha \sum_{q'=1}^Q p_{q, q'} (V_{\alpha}(x+1, q') - V_{\alpha}(1, q')) \\ - (\lambda \omega(q) + W). \quad (36)$$

Denote $s_\alpha^*(x, q)$ be the optimum solution that achieves the minimum discounted cost $V_\alpha(x, q)$ at state (x, q) . If the optimal policy $s_\alpha^*(x, q) = 1$, i.e, it is better to schedule the sensor at state (x, q) , by substituting (11a) into $\Delta(x, q) \geq 0$, we can obtain the following inequality:

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

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:

$$\begin{aligned} \Delta(x', q) &= \alpha \sum_{q'=1}^Q p_{q,q'} (V_\alpha(x'+1, q') - V_\alpha(1, q')) - (\lambda\omega(q) + W) \\ &\stackrel{(a)}{\geq} \alpha \sum_{q'=1}^Q p_{q,q'} (V_\alpha(x+1, q') - V_\alpha(1, q')) - (\lambda\omega(q) + W) \geq 0, \end{aligned} \quad (38)$$

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

$$\begin{aligned} &(C_X(x, q, 0) + \alpha \mathbb{E}[V_\alpha(x+1, q')]) \\ &\quad - (C_X(x, q, 1) + \alpha \mathbb{E}[V_\alpha(1, q')]) \\ &\geq x + \alpha(x+1) - (x + \alpha \mathbb{E}[V_\alpha(1, q')] + W + \lambda\omega(q)) \\ &= \alpha x + \alpha - \alpha \mathbb{E}[V_\alpha(1, q')] - W - \lambda\omega(q), \end{aligned} \quad (39)$$

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 generalize the threshold structure of $s_\alpha^*(x, q)$ to the optimal strategy of the average cost minimization problem. Take a sequence of discount factors such that $\lim_{k \rightarrow \infty} \alpha_k = 1$. 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 (35). Denote $J_{\alpha,s}^{(k)}(x, q)$ to be the value function obtained

after the k^{th} iteration, the monotonic characteristic is proved by induction.

Suppose sequence $V_\alpha^{(k)}(1, q), V_\alpha^{(k)}(2, q), \dots$ is non-decreasing. With no loss of generality, suppose $x_1 < x_2$. According to the one step cost, we have:

$$C_X(x_1, q, s) < C_X(x_2, q, s), C_Q(x_1, q, s) = C_Q(x_2, q, s). \quad (40)$$

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

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

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 $k \rightarrow \infty$, the value function $V_\alpha^{(k)}(x, q) \rightarrow V_\alpha(x, q)$. Hence, $V_\alpha(\cdot, q)$ is monotonic increasing. ■

APPENDIX C PROOF OF THEOREM 2

Proof: Recall that π_R^* is the optimum policy to the *RB&P-Constrained* AoI minimization problem, whose expected AoI formulates the lower bound to the *Primal B&P-Constrained* problem. We verify the asymptotic optimal performance of the proposed truncated policy $\tilde{\pi}$ by comparing it with π_R^* .

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

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

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. This phenomenon implies $x_n(t) \geq \min_q \tau_{n,q}$. If the sensor is still not scheduled for Γ_n consecutive slots, then its AoI $x_n(t + \Gamma_n) \geq \max_q \tau_{n,q}$. Recall that $M/N = \theta$, and the probability that a sensor with $s_n(t) = 1$ is not scheduled by policy $\tilde{\pi}$ can be computed by $\frac{|\Omega(t)| - M}{|\Omega(t)|} \leq 1 - \theta$. Since for

$t' \geq t + \Gamma_n$, we have $s_n(t') = 1$ and with probability no more than $(1 - \theta)$ the sensor n is still not chosen to schedule in slot t' . Thus the probability that sensor n that should be scheduled in slot t but is not in the next consecutive t' slots is upper bounded by $(1 - \theta)^{(t' - \Gamma_n)^+}$, where $(\cdot)^+ = \max\{\cdot, 0\}$. Moreover, if the sensor is not scheduled in the consecutive t' slots, policy $\tilde{\pi}$ will cause an extra AoI growth of no more than $t' x_n(t)$ compared with policy π_R^* .

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 $\sum_{t'=0}^{\infty} (1 - \theta)^{(t' - \Gamma_n)^+} x_n(t) = (\Gamma_n + \frac{1}{\theta}) x_n(t)$ on the total AoI. Notice that the M sensors is chosen randomly, then in slot t , if $|\Omega(t)| > M$, the expected extra cost can be upper bounded by:

$$\begin{aligned} & \mathbb{1}_{|\Omega(t)| > M} \sum_{n=1}^N \left(\Gamma_n + \frac{1}{\theta} \right) x_n(t) \frac{|\Omega(t)| - M}{|\Omega(t)|} \\ & \leq \mathbb{1}_{|\Omega(t)| > M} \sum_{n=1}^N \left(\Gamma_n + \frac{1}{\theta} \right) x_n(t) \frac{|\Omega(t)| - M}{M}. \end{aligned} \quad (43)$$

otherwise if $|\Omega(t)| \leq M$ there is no extra cost.

Notice that 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:

$$\begin{aligned} & (J(\tilde{\pi}) - J(\pi_R^*)) \\ & \leq (J(\hat{\pi}_R^*) - J(\pi_R^*)) \\ & = \frac{1}{NT} \mathbb{E}_{\pi_R^*} \left[\sum_{t=1}^T \mathbb{1}_{|\Omega(t)| > M} \left(\sum_{n=1}^N \left(\Gamma_n + \frac{1}{\theta} \right) \right. \right. \\ & \quad \left. \left. \times x_n(t) \frac{|\Omega(t)| - M}{M} \right) \right] \\ & = \frac{1}{NT} \mathbb{E}_{\pi_R^*} \left[\sum_{t=1}^T \left(\sum_{n=1}^N \left(\Gamma_n + \frac{1}{\theta} \right) \right. \right. \\ & \quad \left. \left. \times x_n(t) \frac{(|\Omega(t)| - M)^+}{M} \right) \right] \\ & \leq \frac{\max_n \Gamma_n + \frac{1}{\theta}}{MNT} \mathbb{E}_{\pi_R^*} \left[\sum_{t=1}^T \sum_{n=1}^N x_n(t) (|\Omega(t)| - M)^+ \right] \\ & \stackrel{(a)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{\theta}}{MNT} \mathbb{E}_{\pi_R^*} \left[\sum_{t=1}^T \sum_{n=1}^N x_n(t) (|\Omega(t)| - \bar{\Omega})^+ \right] \\ & \stackrel{(b)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{\theta}}{MNT} \mathbb{E}_{\pi_R^*} \left[\sum_{t=1}^T \sum_{n=1}^N x_n(t) ||\Omega(t)| - \bar{\Omega}| \right] \\ & \stackrel{(c)}{\leq} \frac{\max_n \Gamma_n + \frac{1}{\theta}}{MNT} \mathbb{E}_{\pi_R^*} \left[\sum_{t=1}^T \sum_{n=1}^N \max_q \tau_{n,q} ||\Omega(t)| - \bar{\Omega}| \right] \\ & \stackrel{(d)}{=} \frac{(\max_n \Gamma_n + \frac{1}{\theta}) \sum_{n=1}^N \max_q \tau_{n,q}}{\theta N^2} \end{aligned}$$

$$\times \mathbb{E}_{\pi_R^*} \left[\frac{1}{T} \sum_{t=1}^T ||\Omega(t)| - \bar{\Omega}| \right], \quad (44)$$

where inequality (a) is because inequality (42) and (b) is because $(\cdot)^+ \leq |\cdot|$. Inequality (c) 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 $\max_q \tau_{n,q}$. Equality (d) is because $M = N\theta$.

Finally, according to [40], the expectation of $|\Omega(t) - \bar{\Omega}|$ satisfies:

$$\mathbb{E}_{\pi_R^*} [||\Omega(t)| - \bar{\Omega}|] = \mathcal{O}(\sqrt{N}),$$

which implies:

$$\mathbb{E}_{\pi_R^*} \left[\frac{1}{\theta NT} \sum_{t=1}^T ||\Omega(t)| - \bar{\Omega}| \right] = \mathcal{O} \left(\frac{1}{\sqrt{N}} \right). \quad (45)$$

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 . In addition, $\frac{M}{N} = \theta$ 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 . As a result, we will have the following upper bound:

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

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

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

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

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

ACKNOWLEDGEMENT

The authors are grateful to Prof. Philippe Ciblat, Prof. Michèle Wigger from Télécom Paris, Dr. Zhen Zhang, Mr. Yuchao Chen, Mr. Jingzhou Sun and Mr. Qining Zhang from Tsinghua University, Dr. Bo Zhou from Virginia Tech, Mr. Jiangwei Xu from Wuhan Tech and the anonymous reviewers for helpful suggestions and discussions that greatly improve the presentation and accuracy of the manuscript.

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 *Proc. 57th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Monticello, IL, USA, Sep. 2019, pp. 1198–1205.
- [2] X. Jiang *et al.*, "Low-latency networking: Where latency lurks and how to tame it," *Proc. 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.-K. Chau, F. Qin, S. Sayed, M. Wahab, and Y. Yang, "Harnessing battery recovery effect in wireless sensor networks: Experiments and analysis," *IEEE J. Sel. Areas Commun.*, 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 *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2731–2735.
- [6] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, Mar. 2019.
- [7] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "On the age of information with packet deadlines," *IEEE Trans. Inf. 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 J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 721–734, Apr. 2019.
- [9] A. Ephremides and B. Hajek, "Information theory and communication networks: An unconsummated union," *IEEE Trans. Inf. 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 *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 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 *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun.*, Apr. 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 Trans. Inf. 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 *Proc. IEEE Int. Conf. Commun. (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 *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 819–823.
- [15] E. T. Ceran, D. Gunduz, and A. Gyorgy, "Average age of information with hybrid ARQ under a resource constraint," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1900–1913, Mar. 2019.
- [16] A. Baknina, S. Ulukus, O. Oze, J. Yang, and A. Yener, "Sensing information through status updates," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 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 *Proc. 54th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 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 Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Dec. 2018.
- [19] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput. (MobiHoc)*, New York, NY, USA, pp. 61–70.
- [20] R. Talak, I. Kadota, S. Karaman, and E. Modiano, "Scheduling policies for age minimization in wireless networks with unknown channel state," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 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 *Proc. 16th Int. Symp. Model. Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, May 2018, pp. 1–8.
- [22] H. Tang, J. Wang, Z. Tang, and J. Song, "Scheduling to minimize age of synchronization in wireless broadcast networks with random updates," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 1027–1031.
- [23] Y.-P. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 561–565.
- [24] Z. Jiang, B. Krishnamachari, X. Zheng, S. Zhou, and Z. Niu, "Decentralized status update for age-of-information optimization in wireless multiaccess channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 2276–2280.
- [25] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, New York, NY, USA, 2018, pp. 191–200.
- [26] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2018, pp. 1–9.
- [27] V. S. Borkar, G. S. Kasbekar, S. Pattathil, and P. Y. Shetty, "Opportunistic scheduling as restless bandits," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 4, pp. 1952–1961, Dec. 2018.
- [28] K. Chen and L. Huang, "Timely-throughput optimal scheduling with prediction," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2457–2470, Dec. 2018.
- [29] M. Wang, J. Liu, W. Chen, and A. Ephremides, "On delay-power tradeoff of rate adaptive wireless communications with random arrivals," in *Proc. IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.
- [30] M. Wang, J. Liu, W. Chen, and A. Ephremides, "Joint queue-aware and channel-aware delay optimal scheduling of arbitrarily bursty traffic over multi-state time-varying channels," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 503–517, Jan. 2019.
- [31] J. Yang and S. Ulukus, "Delay-minimal transmission for average power constrained multi-access communications," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2754–2767, Sep. 2010.
- [32] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 487–499, Aug. 2002.
- [33] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.
- [34] R. Singh and P. R. Kumar, "Throughput optimal decentralized scheduling of multihop networks with end-to-end deadline constraints: Unreliable links," *IEEE Trans. Autom. Control*, vol. 64, no. 1, pp. 127–142, Jan. 2019.
- [35] R. D. Yates, P. Ciblat, A. Yener, and M. Wigger, "Age-optimal constrained cache updating," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 141–145.
- [36] E. Altman, *Constrained Markov Decision Processes*, vol. 7. Boca Raton, FL, USA: CRC Press, 1999, [Online]. Available: <https://www.sop.inria.fr/members/Eitan.Altman/TEMP/h.pdf>.
- [37] D. P. Bertsekas and A. Scientific, *Convex Optimization Algorithms*. Belmont, MA, USA: Athena Scientific, 2015.
- [38] L. I. Sennott, "Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs," *Oper. Res.*, vol. 37, no. 4, pp. 626–633, Aug. 1989.
- [39] F. J. Beutler and K. W. Ross, "Optimal policies for controlled Markov chains with a constraint," *J. Math. Anal. Appl.*, vol. 112, no. 1, pp. 236–252, Nov. 1985.
- [40] P. Diaconis and S. Zabell, "Closed form summation for classical distributions: Variations on a theme of de Moivre," *Stat. Sci.*, vol. 6, no. 3, pp. 284–302, Aug. 1991.
- [41] A. Maatouk, M. Assaad, and A. Ephremides, "Minimizing the age of information: NOMA or OMA?" in *Proc. IEEE Conf. Comput. Commun. Workshops*, Paris, France, Apr. 2019, pp. 102–108.
- [42] H. Tang, J. Wang, P. Ciblat, and J. Song, "Optimizing data freshness in time-varying wireless networks with imperfect channel state," 2019, *arXiv:1910.02353*. [Online]. Available: <https://arxiv.org/abs/1910.02353>
- [43] Y. Wang and W. Chen, "An AoI-optimal scheduling method for wireless transmissions with truncated channel inversion," to be published.

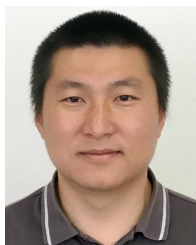


Haoyue Tang (Student Member, IEEE) received the bachelor's degree from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2017, where she is currently pursuing the Ph.D. degree. She was a Visiting Student with Technische Universität München from September 2015 to February 2016, and Télécom Paris from January 2019 to March 2019. Her research interests include age of information, stochastic network optimization, and statistical learning theory.



Jintao Wang (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively. He was an Assistant Professor with the Department of Electronic Engineering, Tsinghua University, from 2006 to 2009. Since 2020, he has been a tenured Full Professor. He has authored over 100 journal and conference papers and holds over 40 national invention patents. His current research interests include space-time coding, multiple-input multiple-output, and OFDM

systems. He is also a Standard Committee Member of the Chinese National Digital Terrestrial Television Broadcasting Standard.



Linqi Song (Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Tsinghua University, China, and the Ph.D. degree in electrical engineering from UCLA. He was a Post-Doctoral Scholar with the Electrical and Computer Engineering Department, University of California, Los Angeles. He is currently an Assistant Professor with the Computer Science Department, City University of Hong Kong. His research interests are in content-type coding, index coding, network coding, algorithms, big data, and machine learning.

He received a UCLA Fellowship for his graduate studies.



Jian Song (Fellow, IEEE) received the B.Eng. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1990 and 1995, respectively. He was with Tsinghua University, The Chinese University of Hong Kong, and the University of Waterloo, Waterloo, ON, Canada, in 1996 and 1997, respectively. In 2005, he joined as a Professor with the Faculty Team, Tsinghua University. He has been with Hughes Network Systems, Germantown, MD, USA. He is currently the Director of the DTV Technology R&D Center, Tsinghua

University. He has been working in quite different areas of fiber-optic, satellite and wireless communications, and the power-line communications. He has published more than 200 peer-reviewed journal and conference papers. He holds 2 U.S. and more than 50 Chinese patents. His current research interest is in the areas of digital TV broadcasting, network convergence, 5G, and the integration of powerline and visible light communications. He is also a Fellow of IET.