Age-Optimal Low-Power Status Update over Time-Correlated Fading Channel

Guidan Yao*, Ahmed M. Bedewy*, and Ness B. Shroff *†
* Department of Electrical and Computer Engineering, Ohio State University,
†Department of Computer Science and Engineering, Ohio State University
yao.539@osu.edu, bedewy.2@osu.edu, shroff.11@osu.edu

***Abstract*—In this paper, we consider transmission scheduling in a status update system, where updates are generated periodically and transmitted over a Gilbert-Elliott fading channel. The goal is to minimize the long-run average *age of information* (AoI) at the destination under an average energy constraint. The channel state is revealed by the feedback (Ack/Nack) of a transmission; while it remains unknown if there is no transmission. Thus, we have to design a scheduling policy that balances tradeoffs across energy, AoI, channel exploration, and channel exploitation. The problem is formulated as a constrained partially observable Markov decision process problem (POMDP). We show that the optimal policy is a randomized mixture of no more than two stationary deterministic policies each of which is of a *threshold-type* in the belief on the channel. We propose a *finite-state approximation* for our infinite-state belief MDP and show convergence. Based on the theoretical insights gained from studying this problem, we develop an optimal algorithm using the structure of the problem.**

I. Introduction

For status update systems, where time-sensitive status updates of certain underlying physical process are sent to a remote destination, it is important that the destination receives fresh updates. The *age of information* (AoI) is a performance metric that is a good measure of the freshness of the data at the destination. In particular, AoI is defined as the time elapsed since the generation of the recently received status update.

The problem of minimizing the AoI in status update systems has attracted significant recent attention (e.g., [1]–[10]). Due to the fact that sensors in the status update system are usually battery-powered and thus have limited energy supply, the problem of minimizing the long-run average AoI has to take energy constraints into account. Moreover, communication over a wireless channel is subject to multiple impairments such as fading, path loss and interference, which may lead to status updating failure. Since each failed transmission consumes unnecessary energy, there is a strong motivation for designing intelligent transmission scheduling algorithms i.e., retransmission or suspension of transmission to increase channel utilization as well as prolong battery life.

Many existing works that deal with the AoI minimization problem under energy constraints in status update systems assume either perfect knowledge of the channel state or noiseless channel to guarantee successful transmission. In [11], [12], the authors assume that the channel is noiseless, and propose offline or online status updating policies. In [13], the authors jointly design sampling and updating processes over a channel with perfect channel state information. However, in many practical scenarios, the channel state may not be known a priori. Thus, more recent works have also considered unreliable transmissions with imperfect knowledge of wireless channels. For example, in [14], the authors consider a block fading channel, where the channel is assumed to vary independently and identically over time slots. In [15], the authors consider an error-prone channel, where decoding error depends only on the number of retransmissions.

However, these works neglect an important characteristics of the wireless fading channel: The *channel memory* or *time correlation* [16] when studying unreliable transmissions with imperfect knowledge of channel states. Indeed, the memory can be intelligently exploited to predict the channel state and thus to design efficient scheduling policies in the presence of transmission cost. A finite state Markov chain is an often used and appropriate model for fading channel [17]. A somewhat simplified but often-used abstraction is a two-state Markovian model known as the Gilbert-Elliot channel [18]. The model assumes that the channel can be either in a good or bad state, and captures the essence of the fading process. In [19], the authors consider status updating in cognitive radio networks. The occupation of primary user's channel is modeled as a two-state Markov chain. Although a Markov chain is used to model occupation of primary channel, their threshold-type structural result is built on perfect knowledge of the channel state since update decisions are made based on perfect sensing results. In contrast, in our work, we do not assume that the channel state is known a priori at the time of making updating decisions.

Motivated by the time-correlation in a fading channel and the fact that sensors in practice are typically configured to generate status updates periodically [20], in this paper, we consider a status update system where the status update is generated periodically and transmitted over a Gilbert-Elliot channel. We do not assume that the channel state is known a priori. In particular, we consider the case *(without channel sensing)* that the channel state is revealed by the ACK/NACK feedback of a transmission, so its state remains unknown if there is no transmission. In our technical report [21], we also study the case *(with delayed channel sensing)* that delayed CSI is always available via delayed channel sensing regardless of transmission decisions. To increase the reliability of received status updates, retransmissions are allowed. With these, we study the problem of how to minimize the average AoI under a long-run average energy constraint, which is formulated as a constrained partially observable Markov decision process problem (POMDP). Our key contributions are as follows:

(i) We show that the optimal transmission scheduling policy

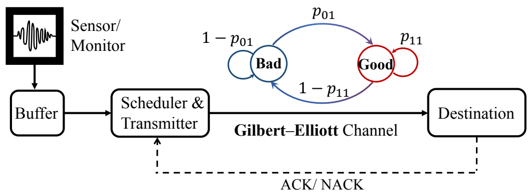

Fig. 1: System Model

is a randomized mixture of no more than two *stationary deterministic threshold-type* policies (Theorem 1 and Corollary 2). Note that although there are some works that deal with showing optimality of threshold-type policies in POMDPs [22]–[26], the techniques in these papers cannot be applied to our problem. This is because, given hidden state and action, the one-stage cost in these papers is constant and bounded, while the one-stage cost in our paper depends on varying and unbounded AoI.

(ii) We propose a finite-state approximation for our infinite-state (unbounded AoI and belief on channel state) belief MDP and show that the optimal policy for the approximated belief MDP converges to the original one (Theorem 2). Based on this, we propose an optimal efficient structure-aware transmission scheduling algorithm (Algorithm 1) for the approximate belief MDP.

II. System Model

We consider a status update system where status updates are generated periodically and transmitted to a remote destination over a time-correlated fading channel as shown in Fig. 1. We consider a time-slotted system, where a time slot corresponds to the time duration of the packet transmission time and feedback period. Every K consecutive time slots form *a frame*. Updates are generated at the beginning of each frame. In any frame, if the generated status update is not delivered by the end of the frame, then it gets replaced by a new one in the next frame. Define $\mathcal{K}$ as the set of *relative slot index* within a frame, $\mathcal{K} \triangleq \{1, 2, \cdots, K\}$. Use $t \in \{1, 2, \cdots\}$ as an *absolute index* for the time slot count, which increments indefinitely with time. For any time slot t, the corresponding frame index $l_t \in \{1, 2, \cdots\}$ is determined by $l_t = \lceil \frac{t}{K} \rceil$ and relative slot index $k_t \in \mathcal{K}$ is determined by $k_t = ((t-1) \bmod K) + 1$, where $\lceil \cdot \rceil$ is the ceiling function.

A. Channel Model

The time-correlated fading channel for transmission is assumed to evolve as a two-state Gilbert-Elliot model [18]. Let h_t denote the channel state at time slot t. Then, $h_t = 1$ ($h_t = 0$) denotes that channel is in a "good" ("bad") state. In the "bad" state, the channel is assumed to be in a deep fade such that transmission fails with probability one; while in the "good" state, a transmission attempt is always successful. This assumption conforms with the signal-to-noise ratio (SNR) threshold model for reception where successful decoding of a packet at the destination occurs if and only if the SNR exceeds a certain threshold value. The channel transition probabilities are given by $\mathbb{P}(h_{t+1} = 1 | h_t = 1) = p_{11}$ and $\mathbb{P}(h_{t+1} = 1 | h_t = 0) = p_{01}$. We assume that the channel transitions occur at the end of each time slot, and that p_{11} and p_{01} are known.

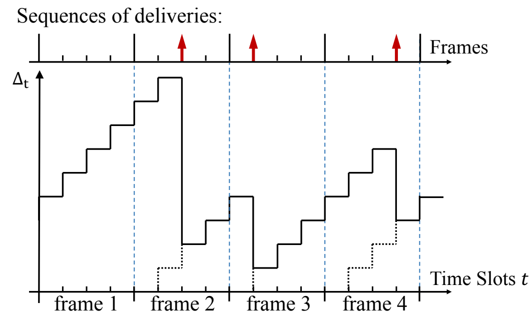

Fig. 2: On the top, a sample sequence of deliveries during four frames. Each frame consists of 4 time slots. The upward arrows represent the times of deliveries. On the bottom, the associated evolution of AoI.

The presence of channel memory (time correlation) makes it possible to predict the channel state. Define Markovian channel memory as $\mu = p_{11} - p_{01}$ [27], [28]. In this paper, we assume that $p_{11} \geq p_{01}$ (positively correlated channel) (similar assumptions have been used in [24], [26]).

B. Transmission Scheduler

At the beginning of each slot t, the scheduler takes a decision $u_t \in \mathcal{U} \triangleq \{0, 1\}$, where $u_t = 1$ means transmitting (retransmitting) the undelivered statues update, and $u_t = 0$ denotes suspension of transmission (retransmission). In each frame, if the generated update is delivered at the k_t-th slot of the frame, then we have $u_t = 0$ for the remaining slots in the frame. For simplicity, we use transmission to refer to both transmission and retransmission in the remaining content. If a transmission is attempted, then the scheduler receives an error-free ACK/NACK feedback from the destination specifying whether the status update was delivered or not before the end of the time slot. We use Θ to denote the set of observations, $\Theta \triangleq \{0, 1\}$. Let $\theta_t \in \Theta$ be the observation at time slot t. Then, $\theta_t = 1$ denotes a successful transmission, while $\theta_t = 0$ denotes either a failed transmission or the transmission is suspended.

C. Age of Information

Age of information (AoI) reflects the timeliness of the information at the destination. Let Δ_t denote the AoI at the beginning of the time slot t. Let $U(t)$ denote the generation time of the last successfully received status update for time slot t. Then, Δ_t is given by $\Delta_t \triangleq t - U(t)$.

If a status update is not successfully delivered in a given slot, then the AoI increases by one, otherwise, the AoI drops to the time elapsed since the beginning of the frame. Then, the evolution of AoI is as follows:

$$\Delta_{t+1} = \begin{cases} k_t & \text{if } u_t = 1, \theta_t = 1, \\ \Delta_t + 1 & \text{otherwise.} \end{cases} \tag{1}$$

Let $\mathcal{A}_k$ denote the set of all possible AoI values at the k-th slot of a frame. By (1), $\mathcal{A}_k = \{\Delta : \Delta = mK + (k)_-, m \in \{0, 1, 2, \cdots\}\}$, where $(k)_- \triangleq ((K + k - 2) \bmod K) + 1$ denotes the relative slot index before k. An example of the AoI evolution with $K = 4$ is illustrated in Fig. 2.

We aim to design an energy efficient scheduler, where each transmission consumes one unit energy. Therefore, the long-run average energy consumption cannot exceed a certain limit $E_{\max} \in (0, 1]$. Observe that $E_{\max} = 1$ means that we have enough energy to support a transmission in every time slot. Although a failed transmission does not decrease AoI, it provides channel state information at the cost of energy. Thus, the transmission scheduler has to balance tradeoffs across energy, AoI, channel exploration, and channel exploitation.

III. Constrained POMDP Formulation and Lagrangian relaxation

A. Constrained POMDP Formulation

Recall that when a transmission is suspended in a time slot, the channel state cannot be revealed in that time slot. Together with the average energy constraint, the problem we consider in the paper turns out to be a constrained partially observable Markov decision problem (POMDP). It has been shown in [29] that for any slot t, a belief state ω_t is a *sufficient statistic* to describe the knowledge of underlying channel state and thus can be used for making optimal decisions at time slot t.

Definition 1. *The belief state ω_t is the conditional probability (given observation and action history) that channel is in a good state at the beginning of the time slot t.*

Thus, adding the belief to the system state, the constrained POMDP can be written as constrained belief MDP [30]. We describe the components of the framework as follows:

States: The system state consists of completely observable states and the belief state, i.e., the system state at slot t is defined by $\mathbf{s}_t{=}(\Delta_t, k_t, \omega_t)$, where $\Delta_t{\in}\mathcal{A}_{k_t}$ is the AoI state that evolves as (1); $k_t \in \mathcal{K}$ is the relative slot index in the frame l_t that evolves as $k_{t+1} = (k_t)_+$, where $(y)_+ \triangleq (y \mod K) + 1$; ω_t is the belief state whose evolution is defined in the following paragraph.

Belief Update: Given u_t and θ_t, the belief state in time slot $t+1$ is updated by $\omega_{t+1} = \Lambda(\omega_t, u_t, \theta_t)$, where $\Lambda(\omega_t, u_t, \theta_t)$ is given by

$$\omega_{t+1} = \Lambda(\omega_t, u_t, \theta_t) = \begin{cases} p_{11} & \text{if } u_t = 1, \theta_t = 1, \\ p_{01} & \text{if } u_t = 1, \theta_t = 0, \\ \mathcal{T}(\omega_t) & \text{if } u_t = 0, \end{cases} \tag{2}$$

where $\mathcal{T}(\omega_t) = \omega_t p_{11} + (1-\omega_t)p_{01}$ denotes the one-step belief update. Observe that, if $u_t = 0$, then the scheduler will not learn the channel state and the belief is updated only according to the Markov chain. If $u_t = 1$, the observation θ_t after the transmission provides the true channel state before the state transition, which occurs at the end of the time slot.

Let $\mathcal{T}^m(\omega_t) \triangleq \mathbb{P}(h_{t+m} = 1|\omega_t)$ denote m-step belief update when the channel is unobserved for m consecutive slots, where $m \in \{0, 1, \cdots\}$ and $\mathcal{T}^0(\omega) = \omega$. Note that by (2), after a transmission ($u_t = 1$), ω_{t+1} is either p_{01} or p_{11}. The belief state ω is, hereafter, updated by $\mathcal{T}$ upon each suspension until next transmission attempt. Thus, the belief state ω is in the form of $\mathcal{T}^m(p_{01})$ or $\mathcal{T}^m(p_{11})$, where $m \geq 0$. Moreover, an increase in AoI by one results from either a failed transmission or suspension. Thus, given AoI state Δ_t, the maximum suspension time after last transmission is no longer than $\Delta_t - 1$. By this, given AoI state Δ, the belief state belongs to the following set $\Omega_\Delta \triangleq \{\omega : \omega = \mathcal{T}^m(p_{01}) \text{ or } \mathcal{T}^m(p_{11}), 0 \leq m < \Delta\}$. As a result, the state space is given by $\mathcal{S} \triangleq \{(\Delta, k, \omega) : k \in K, \Delta \in \mathcal{A}_k, \omega \in \Omega_\Delta\}$.

Actions: Action set is $\mathcal{U} = \{0, 1\}$ defined in Section II-B.

Transition probabilities: Given the current state $\mathbf{s}_t = (\Delta_t, k_t, \omega_t)$ and action u_t at time slot t, the transition probability to the state $\mathbf{s}_{t+1} = (\Delta_{t+1}, k_{t+1}, \omega_{t+1})$ at the next time slot $t+1$, which is denoted by $P_{\mathbf{s}_t\mathbf{s}_{t+1}}(u_t)$, is defined as

$$\begin{aligned} P_{\mathbf{s}_t\mathbf{s}_{t+1}}(u_t) &\triangleq \mathbb{P}(\mathbf{s}_{t+1}|\mathbf{s}_t, u_t) \\ &= \sum_{\theta_t \in \Theta} \mathbb{P}(\theta_t|\mathbf{s}_t, u_t)\mathbb{P}(\mathbf{s}_{t+1}|\mathbf{s}_t, u_t, \theta_t), \end{aligned} \tag{3}$$

where

$$\mathbb{P}(\theta_t|\mathbf{s}_t, u_t) = \begin{cases} \omega_t & \text{if } u_t = 1, \theta_t = 1, \\ 1 - \omega_t & \text{if } u_t = 1, \theta_t = 0, \\ 1 & \text{if } u_t = 0, \theta_t = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$\begin{aligned} &\mathbb{P}(\mathbf{s}_{t+1}|\mathbf{s}_t, u_t, \theta_t) \\ &= \begin{cases} 1 & \text{if } \mathbf{s}_{t+1} = (k_t, (k_t)_+, \Lambda(\omega_t, u_t, \theta_t)), u_t = 1, \theta_t = 1, \\ 1 & \text{if } \mathbf{s}_{t+1} = (\Delta_t + 1, (k_t)_+, \Lambda(\omega_t, u_t, \theta_t)), \theta_t = 0, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{5}$$

Costs: Given a state $\mathbf{s}_t = (\Delta_t, k_t, \omega_t)$ and an action choice u_t at slot t, the cost of one slot is the AoI at the beginning of this slot, i.e., we have $C_\Delta(\mathbf{s}_t, u_t) = \Delta_t$. Moreover, the energy consumption of one slot is $C_E(\mathbf{s}_t, u_t) = u_t$.

A transmission scheduling policy $\pi = \{d_1, d_2, \cdots\}$ specifies the decision rule for each time slot, where a decision rule d_t maps the history of states and actions, and the current state to an action. A policy is *stationary* if the decision rule is independent of time, i.e., $d_t = d$, for all t. Moreover, a policy is *randomized* if $d_t : \mathcal{S} \to \mathcal{P}(\mathcal{U})$ specifies a probability distribution on the set of actions. The policy is *deterministic* if $d_t : \mathcal{S} \to \mathcal{U}$ chooses an action with certainty. For any policy π, we assume that the resulted Markov chain is a unichain (same assumptions are also made in [13], [31]). Our objective is to design a policy π that minimizes the long-run average AoI $\bar{A}(\pi)$ while the long-run average energy consumption $\bar{E}(\pi)$ does not exceed E_{max}, which is formulated as

Problem 1 (Constrained average-AoI belief MDP):

$$\begin{aligned} \bar{A}^\star \triangleq \min_\pi \quad & \bar{A}(\pi) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_\pi\Big[\sum_{t=1}^{T} C_\Delta(\mathbf{s}_t, u_t)\Big] \\ \text{s.t.} \quad & \bar{E}(\pi) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_\pi\Big[\sum_{t=1}^{T} C_E(\mathbf{s}_t, u_t)\Big] \leq E_{\text{max}}. \end{aligned} \tag{6}$$

We use $\bar{A}^\star$ to denote the optimal average AoI, which is the solution to the problem (6). We show in Section IV that there exists a stationary policy which is a randomized mixture of no more than two deterministic policies that achieves $\bar{A}^\star$.

B. Lagrange Formulation of the Constrained POMDP

To obtain the optimal transmission scheduling policy, we reformulate the constrained average-AoI belief MDP in (6) as a parameterized unconstrained average cost belief MDP using the Lagrangian approach. Given Lagrange multiplier λ, the instantaneous Lagrangian cost at time slot t is defined by

$$C(\mathbf{s}_t, u_t; \lambda) = C_\Delta(\mathbf{s}_t, u_t) + \lambda C_E(\mathbf{s}_t, u_t). \tag{7}$$

Then, we have an unconstrained average cost belief MDP which aims at minimizing the average Lagrangian cost:

Problem 2 (Unconstrained average cost belief MDP):

$$\bar{L}^{\star}(\lambda) \triangleq \min_{\pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{\pi}\big[\sum_{t=1}^{T} C(\mathbf{s}_t, u_t; \lambda)\big], \tag{8}$$

where $\bar{L}^{\star}(\lambda)$ is the optimal average Lagrangian cost with regard to λ. A policy is said to be *average cost optimal* if it minimizes the average Lagrangian cost.

The relation between the optimal solutions of the problems (6) and (8) is provided in the following corollary.

Corollary 1. *The optimal average AoI of problem* (6) *and the optimal average Lagrangian cost of problem* (8) *satisfy*

$$\bar{A}^{\star} = \sup_{\lambda\geq 0} \bar{L}^{\star}(\lambda) - \lambda E_{max}. \tag{9}$$

Proof. Please see our technical report [21]. □

IV. Structure Based Algorithm Design

A. Structure of the Constrained Average-AoI Optimal Policy

We begin by showing that there exists a stationary deterministic *threshold-type* scheduling policy that solves the unconstrained average cost belief MDP in (8).

Theorem 1. *Given λ, there exists a stationary deterministic unconstrained average cost optimal policy that is of threshold-type in belief. Specifically,* (8) *can be minimized by a policy of the form $\pi^{\star}_{\lambda} = (d^{\star}_{\lambda}, d^{\star}_{\lambda}, \cdots)$, where*

$$d^{\star}_{\lambda}(\Delta, k, \omega) = \begin{cases} 0 & \text{if } 0 \leq \omega < \omega^{\star}(\Delta, k; \lambda), \\ 1 & \text{if } \omega^{\star}(\Delta, k; \lambda) \leq \omega, \end{cases} \tag{10}$$

where $\omega^{\star}(\Delta, k; \lambda)$ denotes the threshold given pair of AoI and relative slot index (Δ, k) and Lagrange multiplier λ.

Proof Sketch: We obtain the result by relating the average cost belief MDP to a discounted belief MDP. For the discounted belief MDP, we show that the relation between the Q-functions corresponding to the available actions, either u=0 or u=1, can be one of two possible cases. For one case, we obtain threshold property by concavity and monotonicity of the value function. In the other case, we show that it is optimal to always take $u = 0$ via value iteration. Moreover, the properties of the value function and threshold structure need to be proved jointly. For details, please see our technical report [21]. □

Note that the techniques in papers dealing with threshold property in POMDP [22]–[26] cannot be applied to our problem. This is because, given hidden state and action, the one-stage cost in these papers is *constant and bounded*, while the one-stage cost in our paper depends on *varying and unbounded* AoI. Next, we show that the optimal policy for the original problem (6) is a mixture of no more than two stationary deterministic threshold-type policies.

Corollary 2. *There exists a stationary randomized policy $\pi^{\star}$ that is the optimal solution to the constrained average-AoI belief MDP in* (6)*, where $\pi^{\star}$ is a randomized mixture of threshold-type policies as follows:*

$$\pi^{\star} = q\pi^{\star}_{\lambda_1} + (1-q)\pi^{\star}_{\lambda_2}, \tag{11}$$

where $q \in [0, 1]$ is a randomization factor, and $\pi^{\star}_{\lambda_1}$ and $\pi^{\star}_{\lambda_2}$ are the optimal threshold-type policies (10) *for some Lagrange multipliers λ_1 and λ_2, respectively.*

Proof. Please see our technical report [21]. □

The method to determine λ_1, λ_2 and q will be discussed in Section IV-B2.

B. Structure-Aware Algorithm Design

We exploit Corollary 2 to design a structure-aware algorithm for (6) in two steps: We first design a structure-aware algorithm for (8), and then construct a way to determine parameters λ_1, λ_2 and q.

1) Structure-Aware Algorithm for the approximate unconstrained average cost belief MDP: In practice, classic value iteration cannot work if state space is infinite. To deal with this, we first propose a finite-state approximation for infinite-state belief MDP in (8) and show the convergence of our approximate belief MDPs to the original one.

Let N be an upper bound for the AoI and the number of Markov transitions from p_{01} or p_{11}. Since $\mathcal{T}^{i}(p_{01}) \leq \mathcal{T}^{i+1}(p_{01})$ and $\mathcal{T}^{i}(p_{11}) \geq \mathcal{T}^{i+1}(p_{11})$ for $i \in \mathbb{N}$, we have that with bound N, the state space of the approximate belief MDP is given by $\mathcal{S}^N \triangleq \{(\Delta, k, \omega) \in \mathcal{S} : \Delta \leq N, p_{01} \leq \omega \leq \mathcal{T}^{N}(p_{01}) \text{ or } \mathcal{T}^{N}(p_{11}) \leq \omega \leq p_{11}\}$. Without loss of generality, we assume $N > K$.

Given the state $(\Delta_t, k_t, \omega_t) \in \mathcal{S}^N$, the state $\mathbf{s}_{t+1} = (\Delta_{t+1}, k_{t+1}, \omega_{t+1}) \in \mathcal{S}^N$ is updated as follows:

$$\mathbf{s}_{t+1} = \begin{cases} \big(k_t, (k_t)_+, p_{11}\big) & \text{if } u_t = 1, \theta_t = 1, \\ \big(\phi(\Delta_t + 1), (k_t)_+, p_{01}\big) & \text{if } u_t = 1, \theta_t = 0, \\ \big(\phi(\Delta_t + 1), (k_t)_+, \varphi(\mathcal{T}(\omega_t))\big) & \text{if } u_t = 0, \end{cases} \tag{12}$$

where $\phi(x) = \min\{x, N\}$, and $\varphi(y)$ is given by[1]

$$\varphi(y) = \begin{cases} \mathcal{T}^{N}(p_{11}) & \text{if } \mathcal{T}^{N}(p_{01}) < y < \mathcal{T}^{N}(p_{11}), \\ y & \text{otherwise.} \end{cases} \tag{13}$$

Given action u, the transition probability from s to s$'$ on state space $\mathcal{S}^N$, denoted by $P^{N}_{\mathbf{ss}'}(u)$, is expressed as

$$P^{N}_{\mathbf{ss}'}(u) = P_{\mathbf{ss}'}(u) + \sum_{\mathbf{r}\in\mathcal{S}-\mathcal{S}^N} P_{\mathbf{sr}}(u)\mathbb{1}_{\{\nu(\mathbf{r})=\mathbf{s}'\}}, \tag{14}$$

where $P_{\mathbf{ss}'}(u)$ and $P_{\mathbf{sr}}(u)$ are the transition probabilities on $\mathcal{S}$ defined in (3), $\mathbb{1}_{\{.\}}$ is the indicator function, and approximation operation to state is $\nu\left((z1, z2, z3)\right) \triangleq (\phi(z1), z2, \varphi(z3))$.

In general, a sequence of approximate MDPs may not converge to the original MDP [32]. In Theorem 2, we show the convergence of our approximate MDPs to the original MDP.

Theorem 2. *Let $\bar{L}^{N\star}(\lambda)$ be the minimum average Lagrangian cost for the approximate MDP with regard to bound N and Lagrange multiplier λ. Then, $\bar{L}^{N\star}(\lambda) \to \bar{L}^{\star}(\lambda)$ as $N \to \infty$.*

Proof. Please see our technical report [21]. □

The Relative Value Iteration (RVI) algorithm can be utilized to obtain an optimal stationary deterministic policy for the

[1] We upper bound the belief state by $\mathcal{T}^{N}(p_{11})$. This ensures that the optimal policy for the approximate unconstrained belief MDP is of threshold-type.

approximate MDP. In particular, RVI starts with $V_0^N(\mathbf{s}) = 0$, $\forall \mathbf{s} \in \mathcal{S}^N$ and updates $V_{n+1}^N(\mathbf{s})$ by minimizing the RHS of equation (15) in the $(n+1)$-th iteration, $n \in \{0, 1, 2, \cdots\}$.

$$V_{n+1}^N(\mathbf{s}) = \min_u \Big\{ C(\mathbf{s}, u; \lambda) + \sum_{\mathbf{s}' \in \mathcal{S}^N} P_{\mathbf{ss}'}^N(u) h_n^N(\mathbf{s}') - h_n^N(\mathbf{0}) \Big\}, \quad (15)$$

where $\mathbf{0}$ is the reference state and $h_n^N(\mathbf{s}) = V_n^N(\mathbf{s}) - V_n^N(\mathbf{0})$. Note that similar to the proof in Section IV-A, it can be shown that the optimal policy for the approximate MDP is still of threshold-type. Thus, we utilize the threshold property in RVI algorithm and propose a threshold-type RVI to reduce the complexity in Algorithm 1 (Line 4-24). For each iteration, we update the threshold $\omega^\star(\Delta, k; \lambda)$ (Line 16) in addition to $V^N(\mathbf{s})$. If certain state satisfies the threshold condition (Line 11), then the optimal action for the state in this iteration is determined immediately without doing the optimization operation (Line 12), which reduces the algorithm complexity.

Algorithm 1: Structure-Aware Scheduling without Channel Sensing

1 given tolerance $\epsilon > 0, \epsilon_\lambda > 0$, $\lambda^{\star -}$, $\lambda^{\star +}$, N ;
2 **while** $|\lambda^{\star +} - \lambda^{\star -}| > \epsilon_\lambda$ **do**
3     $\lambda = (\lambda^{\star +} + \lambda^{\star -})/2$;
4     $V^N(\mathbf{s}) = 0, h^N(\mathbf{s}) = 0, h_{prev}^N(\mathbf{s}) = \infty$, for all $\mathbf{s} \in \mathcal{S}^N$;
5   **while** $\max_{\mathbf{s} \in \mathcal{S}^N} |h^N(\mathbf{s}) - h_{prev}^N(\mathbf{s})| > \epsilon$ **do**
6         $\omega^\star(\Delta, k; \lambda) = \infty$ for all $\mathbf{s} = (\Delta, k, \omega) \in \mathcal{S}^N$;
7   **foreach** $\mathbf{s} = (\Delta, k, \omega) \in \mathcal{S}^N$ **do**
8   **if** $\Delta < K$ **then**
9                 $u^\star = 0$;
10   **else**
11   **if** $\omega \geq \omega^\star(\Delta, k; \lambda)$ **then**
12                     $u^\star = 1$;
13   **else**
14                     $u^\star = \arg\min_{u \in \{0,1\}} \{C(\mathbf{s}, u; \lambda) + \sum_{\mathbf{s}' \in \mathcal{S}^N} P_{\mathbf{ss}'}^N(u) h^N(\mathbf{s}')\}$;
15   **if** $u^\star = 1$ **then**
16                         $\omega^\star(\Delta, k; \lambda) = \omega$;
17   **end**
18   **end**
19                 $V^N(\mathbf{s}) = C(\mathbf{s}, u^\star; \lambda) + \sum_{\mathbf{s}' \in \mathcal{S}^N} P_{\mathbf{ss}'}^N(u^\star) h^N(\mathbf{s}') - h^N(\mathbf{0})$;
20   **end**
21             $h_{prev}^N(\mathbf{s}) = h^N(\mathbf{s})$;
22             $h^N(\mathbf{s}) = V^N(\mathbf{s}) - V^N(\mathbf{0})$;
23   **end**
24   **end**
25   Compute the average energy cost $\bar{E}(\lambda)$;
26   **if** $\bar{E}(\lambda) > E_{max}$ **then**
27         $\lambda^{\star -} = \lambda$;
28   **else**
29         $\lambda^{\star +} = \lambda$;
30   **end**
31 **end**

2) Lagrange Multiplier Estimation: By Lemma 3.4 of [33], for $\lambda_1 < \lambda_2$, we have $\bar{A}(\pi_{\lambda_1}^\star) \leq \bar{A}(\pi_{\lambda_2}^\star)$ and $\bar{E}(\pi_{\lambda_1}^\star) \geq \bar{E}(\pi_{\lambda_2}^\star)$. Thus, the optimal Lagrangian multiplier $\lambda^\star$ is defined as $\lambda^\star \triangleq \inf\{\lambda > 0 : \bar{E}(\pi_\lambda^\star) \leq E_{\max}\}$. If there exists $\lambda^\star$ such that $\bar{E}(\pi_{\lambda^\star}^\star) = E_{\max}$, then the constrained average-AoI optimal policy is a stationary deterministic policy where q in Corollary 2 is either 0 or 1. Otherwise, the optimal policy $\pi^\star$ chooses policy $\pi_{\lambda^{\star -}}^\star$ with probability q and policy $\pi_{\lambda^{\star +}}^\star$ with probability $1-q$. The randomization factor q can be computed by

$$q = \frac{E_{\max} - \bar{E}(\pi_{\lambda^{\star +}}^\star)}{\bar{E}(\pi_{\lambda^{\star -}}^\star) - \bar{E}(\pi_{\lambda^{\star +}}^\star)}. \quad (16)$$

The bisection method is used to compute $\lambda^{\star -}$, $\lambda^{\star +}$ and thus q (Line 2-3 and Line 26-30 in Algorithm 1). The algorithm starts with $\lambda^{\star -} = 0$ and sufficiently large $\lambda^{\star +}$.

V. Scheduling with Delayed Channel Sensing

With delayed channel sensing, the CSI of the last time slot is always available at the beginning of each time slot. Thus, the problem in this case reduces to a constrained MDP. We show that the optimal policy in this case is also a randomized mixture of no more than two stationary deterministic threshold-type policies. However, due to the simplification in the state, the threshold here is on AoI. This further reduces the complexity of the structure-aware algorithm. The details are provided in our technical report [21]

VI. Numerical Results

We assume the approximation bound $N = 1000$. Let e_t denote total energy consumption before time slot t. Then, $\bar{e}_t \triangleq e_t/(t-1)$ denotes the average energy consumed before time slot t. We compare the proposed transmission scheduling policies with a greedy policy that transmits when the update is not delivered and $\bar{e}_t < E_{\max}$. We set $K = 3$, $p_{11} = 0.7$, $p_{01} = 0.3$, in which case the optimal AoI with no energy constraint is achieved with 0.6167 units energy on average. Thus, the comparison is conducted with energy constraint ranging from 0.1 to 0.6. In Fig. 3, we can observe that the proposed policies outperform the greedy policy in both cases (with delayed channel sensing and without channel sensing). Due to space limitation, we provide more numerical studies and discussions of the proposed algorithms for the two cases in our technical report [21].

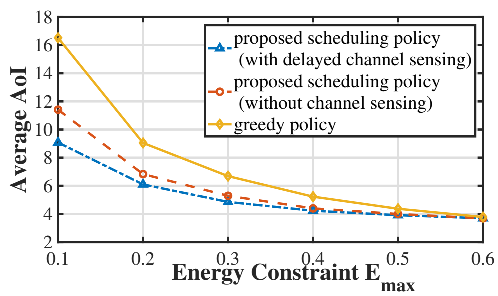

Fig. 3: Comparison with greedy policy

VII. Conclusion

We studied scheduling transmission of periodically generated updates over a Gilbert-Elliott fading channel. The problem is a constrained POMDP and is rewritten as a constrained belief MDP by introducing the belief state. We show that the optimal policy for the constrained belief MDP is a randomization of no more than two stationary deterministic policies, each of which is of a threshold-type in the belief on the channel. We propose a finite-state approximation for our infinite belief MDP and show convergence. Based on this, we propose an optimal structure-aware algorithm.

Acknowledgement

This work was funded in part through NSF grants: CNS-1901057, CNS- 2007231, CNS-1618520, and CNS-1409336, and an Office of Naval Research under Grant N00014-17-1-241.

REFERENCES

[1] Igor Kadota, Elif Uysal-Biyikoglu, Rahul Singh, and Eytan Modiano. Minimizing the age of information in broadcast wireless networks. In *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 844–851. IEEE, 2016.

[2] Igor Kadota, Abhishek Sinha, Elif Uysal-Biyikoglu, Rahul Singh, and Eytan Modiano. Scheduling policies for minimizing age of information in broadcast wireless networks. *IEEE/ACM Transactions on Networking*, 26(6):2637–2650, 2018.

[3] Ahmed M Bedewy, Yin Sun, Rahul Singh, and Ness B Shroff. Optimizing information freshness using low-power status updates via sleep-wake scheduling. In *Proceedings of the Twenty-First International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, pages 51–60, 2020.

[4] Deli Qiao and M Cenk Gursoy. Age minimization for status update systems with packet based transmissions over fading channels. In *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–6. IEEE, 2019.

[5] Ahmed M Bedewy, Yin Sun, Sastry Kompella, and Ness B Shroff. Optimal sampling and scheduling for timely status updates in multi-source networks. *arXiv preprint arXiv:2001.09863*, 2020.

[6] Parisa Rafiee, Peng Zou, Omur Ozel, and Suresh Subramaniam. Maintaining information freshness in power-efficient status update systems. *arXiv preprint arXiv:2003.13577*, 2020.

[7] Ahmed M Bedewy, Yin Sun, and Ness B Shroff. The age of information in multihop networks. *IEEE/ACM Transactions on Networking*, 27(3):1248–1257, 2019.

[8] Ahmed M Bedewy, Yin Sun, and Ness B Shroff. Minimizing the age of information through queues. *IEEE Transactions on Information Theory*, 65(8):5215–5232, 2019.

[9] Guidan Yao, Ahmed M Bedewy, and Ness B Shroff. Battle between rate and error in minimizing age of information. *arXiv preprint arXiv:2012.09351*, 2020.

[10] Ahmed M Bedewy, Yin Sun, Rahul Singh, and Ness B Shroff. Low-power status updates via sleep-wake scheduling. *accepted by IEEE/ACM Transactions on Networking*, 2021.

[11] Baran Tan Bacinoglu, Elif Tugce Ceran, and Elif Uysal-Biyikoglu. Age of information under energy replenishment constraints. In *2015 Information Theory and Applications Workshop (ITA)*, pages 25–31. IEEE, 2015.

[12] Xianwen Wu, Jing Yang, and Jingxian Wu. Optimal status update for age of information minimization with an energy harvesting source. *IEEE Transactions on Green Communications and Networking*, 2(1):193–204, 2017.

[13] Bo Zhou and Walid Saad. Joint status sampling and updating for minimizing age of information in the internet of things. *IEEE Transactions on Communications*, 67(11):7468–7482, 2019.

[14] Haitao Huang, Deli Qiao, and M Cenk Gursoy. Age-energy tradeoff in fading channels with packet-based transmissions. *arXiv preprint arXiv:2005.05610*, 2020.

[15] Elif Tuğçe Ceran, Deniz Gündüz, and András György. Average age of information with hybrid arq under a resource constraint. *IEEE Transactions on Wireless Communications*, 18(3):1900–1913, 2019.

[16] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.

[17] Qinqing Zhang and Saleem A Kassam. Finite-state markov model for rayleigh fading channels. *IEEE Transactions on communications*, 47(11):1688–1692, 1999.

[18] Edgar N Gilbert. Capacity of a burst-noise channel. *Bell system technical journal*, 39(5):1253–1265, 1960.

[19] Shiyang Leng and Aylin Yener. Age of information minimization for an energy harvesting cognitive radio. *IEEE Transactions on Cognitive Communications and Networking*, 5(2):427–439, 2019.

[20] Jaya Prakash Champati, Hussein Al-Zubaidy, and James Gross. Statistical guarantee optimization for aoi in single-hop and two-hop systems with periodic arrivals. *arXiv preprint arXiv:1910.09949*, 2019.

[21] Guidan Yao, Ahmed M Bedewy, and Ness B Shroff. Age-optimal low-power status update over time-correlated fading channel. *arXiv preprint arXiv:2012.02958*, 2020.

[22] William S Lovejoy. Some monotonicity results for partially observed markov decision processes. *Operations Research*, 35(5):736–743, 1987.

[23] S Christian Albright. Structural results for partially observable markov decision processes. *Operations Research*, 27(5):1041–1053, 1979.

[24] Amine Laourine and Lang Tong. Betting on gilbert-elliot channels. *IEEE Transactions on Wireless communications*, 9(2):723–733, 2010.

[25] Keqin Liu and Qing Zhao. Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access. *IEEE Transactions on Information Theory*, 56(11):5547–5567, 2010.

[26] Mehdi Salehi Heydar Abad, Ozgur Ercetin, and Deniz Gündüz. Channel sensing and communication over a time-correlated channel with an energy harvesting transmitter. *IEEE Transactions on Green Communications and Networking*, 2(1):114–126, 2017.

[27] Hong Shen Wang and Nader Moayeri. Finite-state markov channel-a useful model for radio communication channels. *IEEE transactions on vehicular technology*, 44(1):163–171, 1995.

[28] Leigh A Johnston and Vikram Krishnamurthy. Opportunistic file transfer over a fading channel: A pomdp search theory formulation with optimal threshold policies. *IEEE Transactions on Wireless Communications*, 5(2):394–405, 2006.

[29] Richard D Smallwood and Edward J Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations research*, 21(5):1071–1088, 1973.

[30] Yoshikazu Sawaragi and Tsuneo Yoshikawa. Discrete-time markovian decision processes with incomplete state observation. *The Annals of Mathematical Statistics*, 41(1):78–86, 1970.

[31] Dejan V Djonin and Vikram Krishnamurthy. Mimo transmission control in fading channels- a constrained markov decision process formulation with monotone randomized policies. *IEEE Transactions on Signal processing*, 55(10):5069–5083, 2007.

[32] Linn I Sennott. *Stochastic dynamic programming and the control of queueing systems*, volume 504. John Wiley & Sons, 2009.

[33] Linn I Sennott. Constrained average cost markov decision chains. *Probability in the Engineering and Informational Sciences*, 7(1):69–83, 1993.