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-Elliott 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
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 $K$ as the set of relative slot index within a frame, $K \triangleq \{1, 2, \ldots, K\}$. Use $t \in \{1, 2, \ldots\}$ as an absolute index for the time slot count, which increments indefinitely with time. For any time slot $t$, the corresponding frame index $l_t \in \{1, 2, \ldots\}$ is determined by $l_t = \lceil \frac{t}{K} \rceil$ and relative slot index $k_t \in K$ is determined by $k_t = ((t-1) \mod K) + 1$, 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-Elliott 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 $P(h_{t+1} = 1|h_t = 1) = p_{11}$ and $P(h_{t+1} = 1|h_t = 0) = p_{01}$. We assume that the channel transitions occur at the end of each time slot, and that $p_{11}$ and $p_{01}$ are known.

![Fig. 1: System Model](image)

![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]).](image)

B. Transmission Scheduler

At the beginning of each slot $t$, the scheduler takes a decision $u_t \in U \triangleq \{0, 1\}$, where $u_t = 1$ means transmitting (retransmitting) the undelivered status 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 $\Theta$ 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 $\Delta_t$ 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, $\Delta_t$ is given by $\Delta_t = t - U(t)$.

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:

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

Let $A_k$ denote the set of all possible AoI values at the $k$-th slot of a frame. By (1), $A_k = \{\Delta : \Delta = mK + (k)_-, m \in \{0, 1, 2, \ldots\}\}$, where $(k)_- \triangleq ((K + k - 2) \mod K) + 1$ 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 \( \omega_t \) 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 \( \omega_t \) 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 \( s_t = (\Delta_t, k_t, \omega_t) \), where \( \Delta_t \in \mathcal{A}_t \) 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 \( (y) + \equiv (y \mod K) + 1 \); \( \omega_t \) is the belief state whose evolution is defined in the following paragraph.

**Belief Update:** Given \( u_t \) and \( \theta_t \), 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 \( \theta_t \) after the transmission provides the true channel state before the state transition, which occurs at the end of the time slot.

Let \( \mathcal{T}^m(\omega_t) \equiv \mathcal{P}(h_{t+m} = 1|\omega_t) \) 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 (\( u_t = 1 \), \( \omega_{t+1} \) is either \( p_{01} \) or \( p_{11} \). The belief state \( \omega_t \) is, hereafter, updated by \( \mathcal{T} \) upon each suspension until next transmission attempt. Thus, the belief state \( \omega_t \) 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 \( \Delta_t \), the maximum suspension time after last transmission is no longer than \( \Delta_t - 1 \). By this, given AoI state \( \Delta_t \), the belief state belongs to the following set \( \Omega_\Delta \equiv \{ \omega : \omega = \mathcal{T}^m(p_{01}) \) or \( \mathcal{T}^m(p_{11}), 0 \leq m < \Delta \} \). As a result, the state space is given by \( \mathcal{S} \equiv \{ (\Delta, k, \omega) : k \in \mathcal{K}, \Delta \in \mathcal{A}_t, \omega \in \Omega_\Delta \} \).

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

**Transition Probabilities:** Given the current state \( s_t = (\Delta_t, k_t, \omega_t) \) and action \( u_t \) at time slot \( t \), the transition probability to the state \( s_{t+1} = (\Delta_{t+1}, k_{t+1}, \omega_{t+1}) \) at the next time slot \( t + 1 \), which is denoted by \( P_{s_t,u_t}(s_{t+1}) \), is defined as

\[
P_{s_t,u_t}(s_{t+1}) = \sum_{\theta_t \in \Theta} \mathcal{P}(\theta_t|s_t, u_t) \mathcal{P}(s_{t+1}|s_t, u_t, \theta_t), \tag{3}
\]

where

\[
\mathcal{P}(\theta_t|s_t, u_t) = \begin{cases} 
  \omega_t & \text{if } u_t = 1, \theta_t = 1, \\
  1 - \omega_t & \text{if } u_t = 1, \theta_t = 0, \\
  1 & \text{if } u_t = 0, \theta_t = 0, \\
  0 & \text{otherwise},
\end{cases} \tag{4}
\]

**Costs:** Given a state \( s_t = (\Delta_t, k_t, \omega_t) \) 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 \( C_\Delta(s_t, u_t) = \Delta_t \). Moreover, the energy consumption of one slot is \( C_E(s_t, u_t) = u_t \).

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 \( d_t : \mathcal{S} \rightarrow \mathcal{P}(\mathcal{U}) \) specifies a probability distribution on the set of actions. The policy is deterministic if \( d_t : \mathcal{S} \rightarrow \mathcal{U} \) chooses an action with certainty. For any policy \( \pi \), 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 \( \pi \) that minimizes the long-run average AoI \( \bar{A}(\pi) \) while the long-run average energy consumption \( \bar{E}(\pi) \) does not exceed \( E_{\text{max}} \), which is formulated as

**Problem 1 (Constrained average-AoI belief MDP):**

\[
\bar{A}^* \equiv \min_{\pi} \bar{A}(\pi) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^{T} C_\Delta(s_t, u_t) \right] \tag{6}
\]

s.t. \( \bar{E}(\pi) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=1}^{T} C_E(s_t, u_t) \right] \leq E_{\text{max}}. \)

We use \( \bar{A}^* \) 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 \( \bar{A}^* \).

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 \( \lambda \), the instantaneous Lagrangian cost at time slot \( t \) is defined by

\[
C(s_t, u_t; \lambda) = C_\Delta(s_t, u_t) + \lambda C_E(s_t, u_t). \tag{7}
\]
Then, we have an unconstrained average cost belief MDP which aims at minimizing the average Lagrangian cost:

**Problem 2 (Unconstrained average cost belief MDP):**

\[
L^*(\lambda) \triangleq \min_{\pi} \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi}[\sum_{t=1}^{T} C(s_t, u_t; \lambda)],
\]

where \(L^*(\lambda)\) is the optimal average Lagrangian cost with regard to \(\lambda\). 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

\[
A^* = \sup_{\lambda \geq 0} L^*(\lambda) - \lambda E_{\text{max}},
\]

**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 \(\lambda\), 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 \(\pi^*_\lambda = (d^*_\lambda, d^*_\lambda, \cdots)\), where

\[
d^*_\lambda(\Delta, k, \omega) = \begin{cases} 0 & \text{if } 0 \leq \omega < \omega^*(\Delta, k; \lambda), \\ 1 & \text{if } \omega^*(\Delta, k; \lambda) \leq \omega, \end{cases}
\]

where \(\omega^*(\Delta, k; \lambda)\) denotes the threshold given pair of AoI and relative slot index \((\Delta, k)\) and Lagrange multiplier \(\lambda\).

**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 \(\pi^*\) that is the optimal solution to the constrained average-AoI belief MDP in (6), where \(\pi^*\) is a randomized mixture of threshold-type policies as follows:

\[
\pi^* = q\pi^*_{\lambda_1} + (1-q)\pi^*_{\lambda_2},
\]

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

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

The method to determine \(\lambda_1\), \(\lambda_2\) 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 \(\lambda_1\), \(\lambda_2\) 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 \(T^N(p_{01}) \leq T^{N+1}(p_{01})\) and \(T^N(p_{11}) \geq T^{N+1}(p_{11})\) for \(i \in \mathbb{N}\), we have that with bound \(N\), the state space of the approximate belief MDP is given by \(S^N \triangleq \{ (\Delta, k) \in \mathcal{S} : \Delta < N, p_{01} \leq \omega \leq T^N(p_{11}) \} \). Without loss of generality, we assume \(N > K\).

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

\[
s_{t+1} = \begin{cases} (k_t, (k_t)_+, p_{11}) & \text{if } u_t = 1, \theta_t = 1, \\ (\phi(\Delta_t+1), (k_t)_+, p_{01}) & \text{if } u_t = 1, \theta_t = 0, \\ (\phi(\Delta_t+1), (k_t)_+, \varphi(T(\omega_t))) & \text{if } u_t = 0, \end{cases}
\]

where \(\phi(x) = \min\{x, N\}\), and \(\varphi(y)\) is given by:

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

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

\[
P^N_{ss'}(u) = P^s_{ss'}(u) + \sum_{r \in S - S^N} P^r_{sr}(u) \mathbb{I}_{\{r(r) = s'\}},
\]

where \(P^s_{ss'}(u)\) and \(P^r_{sr}(u)\) are the transition probabilities on \(S\) defined in (3), \(\mathbb{I}_{\{r\}}\) is the indicator function, and approximation operation to state is \(\nu((z_1, z_2, z_3)) \triangleq (\phi(z_1), z_2, \varphi(z_3))\).

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

**Proof.** Please see our technical report [21].
approximate MDP. In particular, RVI starts with $V_0^N(s) = 0$, $\forall s \in S^N$ and updates $V_{n+1}^N(s)$ by minimizing the RHS of equation (15) in the $(n+1)$-th iteration, $n \in \{0,1,2,\cdots\}$.

$$V_{n+1}^N(s) = \min_c \left\{ C(s,u;\lambda) + \sum_{s' \in S^N} P_{w_s}^N(u) h_{n+1}^N(s') - h_{n}^N(0) \right\}, \tag{15}$$

where $0$ is the reference state and $h_{n}^N(s) = V_n^N(s) - V_n^N(0)$. 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 $\omega^*(\lambda, k; \lambda)$ (Line 16) in addition to $V_n^N(s)$. 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**

```plaintext
1  given tolerance $\epsilon > 0$, $\epsilon_1 > 0$, $\lambda^* = \lambda^+ + N$;
2  while $|\lambda^* - \lambda^+| > \epsilon_1$ do
3    $\lambda = (\lambda^* + \lambda^+)/2$;
4    $V^N(s) = 0, h_{prev}^N(s) = \infty$, for all $s \in S^N$;
5    while $\max_{s \in S^N} [h_n^N(s) - h_{prev}^N(s)] > \epsilon_2$ do
6      $\omega^*(\Delta, k; \lambda) = \infty$ for all $s = (\Delta, k, \omega) \in S^N$;
7      foreach $s = (\Delta, k, \omega) \in S^N$ do
8        if $\Delta < K$ then
9          $u^* = 0$;
10        else if $\omega \geq \omega^*(\Delta, k; \lambda)$ then
11          $u^* = 1$;
12          else $u^* = \arg\min_{u \in \{0,1\}} \{C(s,u;\lambda) + \sum_{s' \in S^N} P_{w_s}^N(u) h_n^N(s')\}$;
13        if $u^* = 1$ then
14          $\omega = \omega^*(\Delta, k; \lambda) = \omega$;
15      end
16    end
17    $V_{n+1}^N(s) = C(s,u^*;\lambda) + \sum_{s' \in S^N} P_{w_s}^N(u^*) h_n^N(s') - h_{n}^N(0)$;
18    $h_{prev}^N(s) = h_{n}^N(s)$;
19    $h_{n}^N(s) = V_{n+1}^N(s) - V_{n}^N(0)$;
20  end
21  Compute the average energy cost $\bar{E}(\lambda)$;
22  if $\bar{E}(\lambda) > E_{max}$ then
23    $\lambda^* = \lambda$;
24  else $\lambda^+ = \lambda$;
25  end
26 end
```

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

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

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

**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, $\tilde{e}_t = e_t/(t-1)$ 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 $\tilde{e}_t < E_{max}$. 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](image-url)

**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


