

# Delay-Optimal Probabilistic Scheduling in Green Communications with Arbitrary Arrival and Adaptive Transmission

Xiang Chen\*, Wei Chen\*, Joohyun Lee<sup>†</sup>, and Ness B. Shroff<sup>‡</sup>

\* Tsinghua National Laboratory for Information Science and Technology (TNList)  
Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China  
Email: chen-xiang12@mails.tsinghua.edu.cn, wchen@tsinghua.edu.cn

<sup>†</sup> Department of ECE, The Ohio State University, E-mail: lee.7119@osu.edu

<sup>‡</sup> Department of ECE and Department of CSE, The Ohio State University, E-mail: shroff@ece.osu.edu

**Abstract**—In this paper, we aim to obtain the optimal delay-power tradeoff and the corresponding optimal scheduling policy for arbitrary i.i.d. arrival process and adaptive transmissions. The number of backlogged packets at the transmitter is known to a scheduler, who has to determine how many backlogged packets to transmit during each time slot. The power consumption is assumed to be convex in transmission rates. Hence, if the scheduler transmits faster, the delay will be reduced but with higher power consumption. To obtain the optimal delay-power tradeoff and the corresponding optimal policy, we model the problem as a Constrained Markov Decision Process (CMDP), where we minimize the average delay given an average power constraint. By steady-state analysis and Lagrangian relaxation, we can show that the optimal tradeoff curve is decreasing, convex, and piecewise linear, and the optimal policy is threshold-based. Based on the revealed properties of the optimal policy, we develop an algorithm to efficiently obtain the optimal tradeoff curve and the optimal policy. The complexity of our proposed algorithm is much lower than a general algorithm based on Linear Programming. We validate the derived results and the proposed algorithm through Linear Programming and simulations.

## I. INTRODUCTION

In this paper, we study an important problem of how to schedule the number of packets to transmit over a link taking into account both the delay and the power cost. This is an important problem because delay is a vital metric for many emerging applications (e.g., instant messenger, social network service, streaming media, and so on), and power consumption is critical to battery life of various mobile devices. In other words, we are studying the tradeoff between the timeliness and greenness of the communication service.

Such a delay-power scheduling problem can be formulated using a Markov Decision Process (MDP). The authors in [1] were among the earliest who studied this type of scheduling problem. Specifically, they considered a two-state channel and finite time horizon. The dual problem was solved based

on results derived by Dynamic Programming and induction. Follow-up papers [2–5] extended this study in various directions. The optimal delay-power tradeoff curve is proven to be nonincreasing and convex in [2]. The existence of stationary optimal policy and the structure of the optimal policy are further investigated in [3]. Different types of power/rate control policies are studied in [4]. In [5], the asymptotic small-delay regime is investigated. In [6], a piecewise linear delay-power tradeoff curve was obtained along with an approximate closed form expression.

If one can show monotonicity or a threshold type of structure to the optimal policy for MDPs, it helps to substantially reduce the computation complexity in finding the optimal policy. Indeed, the optimal scheduling policies are shown to be threshold-based or monotone in [1, 3, 5, 7–10], proven by studying the convexity, superadditivity / subadditivity, or supermodularity / submodularity of expected cost functions by induction using dynamic programming. However, most of these results are limited to the unconstrained Lagrangian Relaxation problem. In [3, 10], some properties of the optimal policy for the constrained problem are described based on the results for the unconstrained problem. Detailed analysis on the optimal policy for the constrained problem is conducted in [8, 9]. In [8], properties such as unichain policies and multimodularity of costs are assumed to be true so that monotone optimal policies can be proven. In [9], the transmission action is either 1 or 0, i.e. to transmit or not. In order to obtain the detailed structure of the solution to the constrained problem, we believe that the analysis of the Lagrangian relaxation problem and the analysis of the structure of the delay-power tradeoff curve should be combined together.

In [11], we study the optimal delay-power tradeoff problem. In particular, we minimize the average delay given an average power constraint, considering Bernoulli arrivals and adaptive transmissions. Some technical details are given in [12], where we proved that the optimal tradeoff curve is convex and piecewise linear, and the optimal policies are threshold-based, by Constrained Markov Decision Process formulation and

This work has been supported in part by the National Science Foundation of China under Grant Nos. 61671269 and 61621091, US NSF grants CNS-1618566, CNS-1421576, CNS-1409336, the National 973 Program of China under Project No. 2013CB336600, and a grant from the Army Research Office ARO-W911NF-14-1-0368.

steady-state analysis. In this paper, we substantially generalize the Bernoulli arrival process to an arbitrary i.i.d. distribution. We show that the optimal policies for this generalized model are still threshold-based. Furthermore, we develop an efficient algorithm to find the optimal policy and the optimal delay-power tradeoff curve.

The remainder of this paper is organized as follows. The system model and the constrained problem are introduced in Section II. We show that the optimal policy is threshold-based in Section III by using steady-state analysis and Lagrangian relaxation. Based on theoretical results, we propose an efficient algorithm in Section IV to obtain the optimal tradeoff curve and the corresponding policies. In Section V, theoretical results and the proposed algorithm are verified by simulations. Section VI concludes the paper.

## II. SYSTEM MODEL

The system model is shown in Fig. 1. We assume there are  $a[n]$  data packet(s) arriving at the end of the  $n$ th timeslot. The number  $a[n]$  is i.i.d. for different values of  $n$  and its distribution is given by  $\Pr\{a[n] = a\} = \alpha_a$ , where  $\alpha_a \geq 0$ ,  $a \in \{0, 1, \dots, A\}$ , and  $\sum_{a=0}^A \alpha_a = 1$ . Therefore the expected number of packets arrived in each timeslot  $n$  is given by  $E_a = \sum_{a=0}^A a\alpha_a$ .

Let  $s[n]$  denote the number of data packets transmitted in timeslot  $n$ . Assume that at most  $S$  packets can be transmitted in each timeslot because of the constraints of the transmitter, and  $S \geq A$ . Let  $\tau[n]$  denote the transmission power consumed in timeslot  $n$ . Assume transmitting  $s$  packet(s) will cost power  $P_s$ , where  $s \in \{0, 1, \dots, S\}$ , therefore  $\tau[n] = P_{s[n]}$ . Transmitting 0 packet will cost no power, hence  $P_0 = 0$ . In typical communications, the power efficiency decreases as the transmission rate increases, hence we assume that  $P_s$  is convex in  $s$ . Detailed explanations can be found in the Introduction section in [12]. The convexity of the power consumption function will be utilized in Theorem 2 to prove that the optimal policy for the unconstrained problem is threshold-based.

Backlog packets are stored in a buffer with size  $Q$ . Let  $q[n] \in \{0, 1, \dots, Q\}$  denote the queue length at the beginning of timeslot  $n$ . Since data arrive at the end of the timeslot, in order to avoid buffer overflow (i.e.  $q[n] > Q$ ) and underflow (i.e.  $q[n] < 0$ ), we should have  $0 \leq q[n] - s[n] \leq Q - A$ . Therefore the dynamics of the buffer is given as

$$q[n+1] = q[n] - s[n] + a[n]. \quad (1)$$

In timeslot  $n$ , we can decide how many packets to be transmitted based on the buffer state  $q[n]$ . It can be seen that this is a Markov Decision Process (MDP), where the queue length  $q[n]$  is the state of the MDP, and the number of packets transmitted in each timeslot  $s[n]$  is the action we take in each timeslot  $n$ . The probability distribution of the next state  $q[n+1]$  is given by

$$\Pr\{q[n+1] = j | q[n] = q, s[n] = s\} = \begin{cases} \alpha_{j-q+s} & 0 \leq j - q + s \leq A, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

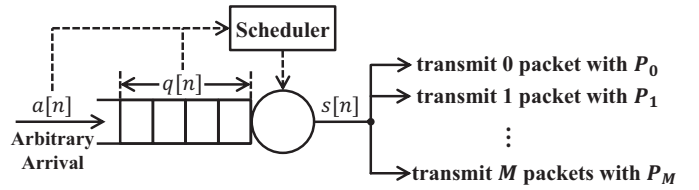


Fig. 1. System Model

We minimize the average queueing delay given an average power constraint, which makes it a Constrained Markov Decision Process (CMDP). For an infinite-horizon CMDP with stationary parameters, according to [13, Theorem 11.3], stationary policies are complete, which means stationary policies can achieve the optimal performance. Therefore we only need to consider stationary policies in this problem. Let  $f_{q,s}$  denote the probability to transmit  $s$  packet(s) when  $q[n] = q$ , i.e.,

$$f_{q,s} = \Pr\{s[n] = s | q[n] = q\}. \quad (3)$$

Then we have  $\sum_{s=0}^S f_{q,s} = 1$  for  $q = 0, \dots, Q$ . Since we guarantee that the transmission strategy will avoid overflow or underflow, we set

$$f_{q,s} = 0 \text{ if } q - s < 0 \text{ or } q - s > Q - A. \quad (4)$$

Let  $\mathbf{F}$  denote a  $(Q+1) \times (S+1)$  matrix whose element in the  $(q+1)$ th row and the  $(s+1)$ th column is  $f_{q,s}$ . Therefore matrix  $\mathbf{F}$  can represent a stationary transmission policy. Let  $P_{\mathbf{F}}$  and  $D_{\mathbf{F}}$  denote the average power consumption and the average queueing delay under policy  $\mathbf{F}$ . Let  $\mathcal{F}$  denote the set of all feasible stationary policies that guarantee no queue overflow or underflow. Let  $\mathcal{F}_D$  denote the set of all stationary and deterministic policies which can guarantee no overflow or underflow. Thus to obtain the optimal tradeoff curve, we can minimize the average delay given an average power constraint  $P_{\text{th}}$  shown as

$$\min_{\mathbf{F} \in \mathcal{F}} D_{\mathbf{F}} \quad (5a)$$

$$\text{s.t. } P_{\mathbf{F}} \leq P_{\text{th}}. \quad (5b)$$

From another perspective, policy  $\mathbf{F}$  will determine a point  $Z_{\mathbf{F}} = (P_{\mathbf{F}}, D_{\mathbf{F}})$  in the delay-power plane. Define  $\mathcal{R} = \{Z_{\mathbf{F}} | \mathbf{F} \in \mathcal{F}\}$  as the set of all feasible points in the delay-power plane. Intuitively, since the power consumption for each data packet increases if we want to transmit faster, there is a tradeoff between the average queueing delay and the average power consumption. Thus the optimal delay-power tradeoff curve can be presented as  $\mathcal{L} = \{(P, D) \in \mathcal{R} | \forall (P', D') \in \mathcal{R}, \text{ either } P' \geq P \text{ or } D' \geq D\}$ .

If we fix a stationary policy for a Markov Decision Process, the Markov Decision Process will degenerate to a Markov Reward Process (MRP). Let  $\lambda_{i,j}$  denote the transition probability from state  $i$  to state  $j$ . According to the system model, because of the constraints of transmission and arrival processes, the state transition probability can be derived as  $\lambda_{i,j} = \sum_{s=\max\{0, i+A-Q, i-j\}}^{\min\{S, i, i-j+A\}} \alpha_{j-i+s} f_{i,s}$ . An example of the transition diagram is shown in Fig. 2, where  $\lambda_{i,i}$  for  $i = 0, \dots, Q$  are omitted to keep the diagram legible.

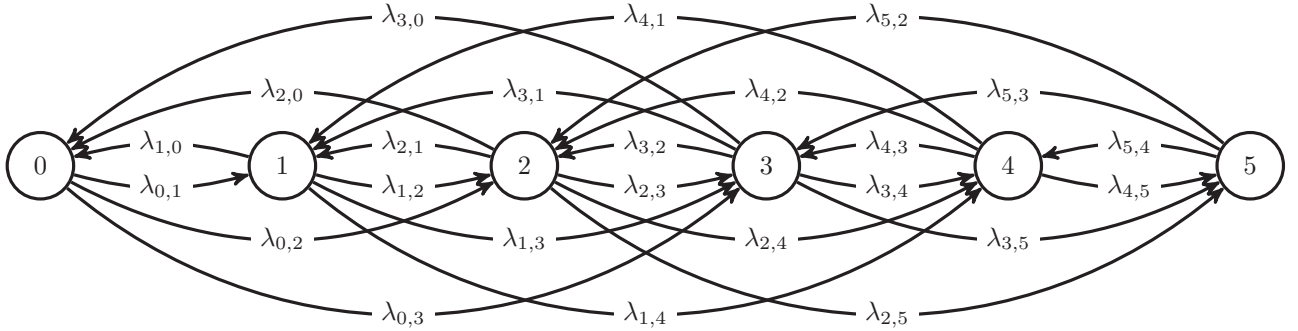


Fig. 2. Markov Chain of  $t[n]$  ( $Q = 5$ ,  $A = 3$ ,  $M = 3$ ,  $\lambda_{i,i}$  for all  $i$  are omitted to keep the diagram legible)

The Markov chain could have more than one closed communication classes under certain transmission policies. Under this circumstance, the limiting probability distribution and the average cost are dependent on the initial state and the sample paths. We can show that we only need to consider the cases where the Markov chain has only one closed communication class, which is called a unichain. Because of space limitation, we omit the proof details, which can be found in [14]. Because of this key result, we focus only on the unichain cases in the following.

### III. OPTIMAL THRESHOLD-BASED POLICY FOR THE CONSTRAINED MARKOV DECISION PROCESS

In this section, we will demonstrate that the optimal policy for the Constrained MDP problem is threshold-based. In other words, for an optimal policy, more data will be transmitted if the queue is longer. We give the rigorous definition of a stationary threshold-based policy  $\mathbf{F}$  that, there exist  $(S + 1)$  thresholds  $0 \leq q_{\mathbf{F}}(0) \leq q_{\mathbf{F}}(1) \leq \dots \leq q_{\mathbf{F}}(S) \leq Q$ , such that  $f_{q,s} > 0$  only when  $q_{\mathbf{F}}(s-1) \leq q \leq q_{\mathbf{F}}(s)$  (set  $q_{\mathbf{F}}(-1) = -1$  for simplicity of notation). According to this definition, under policy  $\mathbf{F}$ , when the queue state is larger than threshold  $q_{\mathbf{F}}(s-1)$  and smaller than  $q_{\mathbf{F}}(s)$ , it transmits  $s$  packet(s). When the queue state is equal to threshold  $q_{\mathbf{F}}(s)$ , it transmits  $s$  or  $(s+1)$  packet(s). Note that under this definition, probabilistic policies can also be threshold-based.

In the following, we will first conduct the steady-state analysis of the Markov process, based on which we can show the properties of the feasible delay-power region and the optimal delay-power tradeoff, and then by proving that the Lagrangian relaxation problem has a deterministic and threshold-based optimal policy, we can finally show that the optimal policy for the constrained problem is threshold-based.

#### A. Steady State Analysis

Since we can focus on unichain cases, which contain a single recurrent class plus possibly some transient states, the steady-state probability distribution exists for the Markov process. Let  $\pi_{\mathbf{F}}(q)$  denote the steady-state probability for state  $q$  when applying policy  $\mathbf{F}$ . Set  $\boldsymbol{\pi}_{\mathbf{F}} = [\pi_{\mathbf{F}}(0), \dots, \pi_{\mathbf{F}}(Q)]^T$ . Define  $\boldsymbol{\Lambda}_{\mathbf{F}}$  as a  $(Q + 1) \times (Q + 1)$  matrix whose element in the  $(i + 1)$ th column and the  $(j + 1)$ th row is  $\lambda_{i,j}$ , which is

determined by policy  $\mathbf{F}$ . Set  $\mathbf{I}$  as the identity matrix. Define  $\mathbf{1} = [1, \dots, 1]^T$ , and  $\mathbf{0} = [0, \dots, 0]^T$ . Set  $\mathbf{G}_{\mathbf{F}} = \boldsymbol{\Lambda}_{\mathbf{F}} - \mathbf{I}$ . Set  $\mathbf{H}_{\mathbf{F}} = \begin{bmatrix} \mathbf{1}^T \\ \mathbf{G}_{\mathbf{F}}(0 : (Q-1), :) \end{bmatrix}$  and  $\mathbf{c} = \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}$ .

According to the definition of the steady-state distribution, we have  $\mathbf{G}_{\mathbf{F}}\boldsymbol{\pi}_{\mathbf{F}} = \mathbf{0}$  and  $\mathbf{1}^T\boldsymbol{\pi}_{\mathbf{F}} = 1$ . For a unichain, the rank of  $\mathbf{G}_{\mathbf{F}}$  is  $Q$ . Therefore, we have  $\mathbf{H}_{\mathbf{F}}$  is invertible and

$$\mathbf{H}_{\mathbf{F}}\boldsymbol{\pi}_{\mathbf{F}} = \mathbf{c}. \quad (6)$$

For state  $q$ , transmitting  $s$  packet(s) will cost  $P_s$  with probability  $f_{q,s}$ . Define  $\mathbf{p}_{\mathbf{F}} = [\sum_{s=0}^S P_s f_{0,s}, \dots, \sum_{s=0}^S P_s f_{Q,s}]^T$ , which is a function of  $\mathbf{F}$ . The average power consumption  $P_{\mathbf{F}}$  can be expressed as

$$P_{\mathbf{F}} = \sum_{q=0}^Q \pi_{\mathbf{F}}(q) \sum_{s=0}^S P_s f_{q,s} = \mathbf{p}_{\mathbf{F}}^T \boldsymbol{\pi}_{\mathbf{F}}. \quad (7)$$

Similarly, define  $\mathbf{d} = [0, 1, \dots, Q]^T$ . According to Little's Law, the average delay  $D_{\mathbf{F}}$  under policy  $\mathbf{F}$  is

$$D_{\mathbf{F}} = \frac{1}{E_a} \sum_{q=0}^Q q \pi_{\mathbf{F}}(q) = \frac{1}{E_a} \mathbf{d}^T \boldsymbol{\pi}_{\mathbf{F}}. \quad (8)$$

The following theorem describes the structure of the feasible delay-power region and the optimal delay-power tradeoff curve.

**Theorem 1.** The set of all feasible points in the delay-power plane,  $\mathcal{R}$ , and the optimal delay-power tradeoff curve  $\mathcal{L}$ , satisfy that

- 1) The set  $\mathcal{R}$  is a convex polygon.
- 2) The curve  $\mathcal{L}$  is piecewise linear, decreasing, and convex.
- 3) Vertices of  $\mathcal{R}$  and  $\mathcal{L}$  are all obtained by deterministic scheduling policies.
- 4) The policies corresponding to adjacent vertices of  $\mathcal{R}$  and  $\mathcal{L}$  take different actions in only one state.

**Remark 1.** Due to space limitation, the proof details can be referred to in [14]. The sketch of the proof is similar to [12]. We can first prove that, if  $\mathbf{F}'$  and  $\mathbf{F}''$  are two policies differing in only one state, and  $\mathbf{F}$  is their convex combination, then  $Z_{\mathbf{F}}$  is the convex combination of  $Z_{\mathbf{F}'}$  and  $Z_{\mathbf{F}''}$ . Then step by step, it can be proven that the feasible delay-power region

for two policies differing in multiple states and the policies in between is a convex polygon. At last, it is discovered that  $\mathcal{R}$  is a convex polygon whose adjacent vertices are obtained by policies differing in only one state. The properties of  $\mathcal{L}$  can then be proven based on properties of  $\mathcal{R}$ .

### B. Optimal Deterministic Threshold-Based Policy for the Lagrangian Relaxation Problem

In (5), we formulate the optimization problem as a Constrained MDP, which is difficult to solve in general. Let  $\mu \geq 0$  denote the Lagrange multiplier. Consider the Lagrangian relaxation of (5)

$$\min_{\mathbf{F} \in \mathcal{F}} D_{\mathbf{F}} + \mu P_{\mathbf{F}} - \mu P_{\text{th}}. \quad (9)$$

In (9), the term  $-\mu P_{\text{th}}$  is constant. Therefore, the Lagrangian relaxation problem is minimizing the weighted average cost  $D_{\mathbf{F}} + \mu P_{\mathbf{F}}$ , which becomes an unconstrained infinite-horizon Markov Decision Process with an average cost criterion. It is proven in [15, Theorem 9.1.8] that, there exists an optimal stationary deterministic policy. Moreover, the optimal policy for the relaxation problem has the following property.

**Theorem 2.** An optimal policy  $\mathbf{F}$  for the unconstrained Markov Decision Process is threshold-based. That is to say, there exists  $(S+1)$  thresholds  $q_{\mathbf{F}}(0) \leq q_{\mathbf{F}}(1) \leq \dots \leq q_{\mathbf{F}}(S)$ , such that

$$\begin{cases} f_{q,s} = 1 & q_{\mathbf{F}}(s-1) < q \leq q_{\mathbf{F}}(s), s = 0, \dots, S \\ f_{q,s} = 0 & \text{otherwise} \end{cases} \quad (10)$$

where  $q_{\mathbf{F}}(-1) = -1$ .

**Remark 2.** The proof of Theorem 2 is also omitted because of space limitation. Detailed proof can be found in [14]. As an extension of [12], the proof uses induction on the policy iteration algorithm. The policy iteration algorithm is proven to converge to the optimal policy in finite steps by using the potential function for the Markov Decision Process. In each iteration in our proof, the potential function will be kept convex, and the policy will be kept threshold based. Theorem 2 indicates that to minimize the weighted average cost, more data should be transmitted if the queue is longer.

### C. Optimal Threshold-Based Policy for the Constrained Problem

From another perspective,  $D_{\mathbf{F}} + \mu P_{\mathbf{F}} = \langle (\mu, 1), (P_{\mathbf{F}}, D_{\mathbf{F}}) \rangle$  can be seen as the inner product of vector  $(\mu, 1)$  and  $Z_{\mathbf{F}}$ . Since  $\mathcal{L}$  is piecewise linear, decreasing and convex, the corresponding  $Z_{\mathbf{F}}$  minimizing the inner product will be obtained by the vertices of  $\mathcal{L}$ , as can be observed in Fig. 3. Since the conclusion in Theorem 2 holds for any  $\mu$ , the vertices of the optimal tradeoff curve can all be obtained by optimal policies for the Lagrangian relaxation problem, which are deterministic and threshold-based. Moreover, from Theorem 1, the adjacent vertices of  $\mathcal{L}$  are obtained by policies which take different actions in only one state. Therefore, we can have the following theorem.

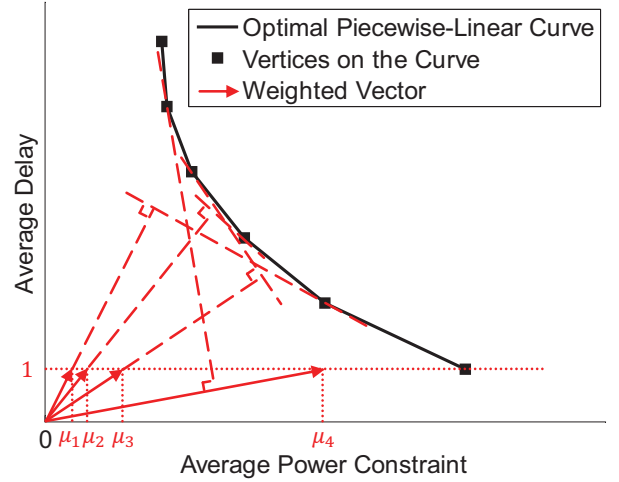


Fig. 3. The minimum inner product of points on  $\mathcal{L}$  and the weighted vector can always be obtained by vertices of  $\mathcal{L}$

**Theorem 3.** Given an average power constraint, the scheduling policy  $\mathbf{F}$  to minimize the average delay takes the following form: there exists  $(S+1)$  thresholds  $q_{\mathbf{F}}(0) \leq q_{\mathbf{F}}(1) \leq \dots \leq q_{\mathbf{F}}(S)$ , one of which we denote as  $q_{\mathbf{F}}(s^*)$ , such that

$$\begin{cases} f_{q,s} = 1 & q_{\mathbf{F}}(s-1) < q \leq q_{\mathbf{F}}(s), s \neq s^* \\ f_{q,s^*} = 1 & q_{\mathbf{F}}(s^*-1) < q < q_{\mathbf{F}}(s^*) \\ f_{q_{\mathbf{F}}(s^*),s^*} + f_{q_{\mathbf{F}}(s^*),s^*+1} = 1 \\ f_{q,s} = 0 & \text{otherwise} \end{cases} \quad (11)$$

where  $q_{\mathbf{F}}(-1) = -1$ .

*Proof:* Since the optimal tradeoff curve is piecewise linear, assume  $Z_{\mathbf{F}}$  is on the line segment between vertices  $Z_{\mathbf{F}'}$  and  $Z_{\mathbf{F}''}$ . According to Theorem 2, the form of optimal policies  $\mathbf{F}'$  and  $\mathbf{F}''$ , which are corresponding to vertices of the optimal tradeoff curve, satisfies (10). Moreover, according to Theorem 1, the policies corresponding to adjacent vertices of  $\mathcal{L}$  take different actions in only one state. Denote the thresholds for  $\mathbf{F}'$  as  $q_{\mathbf{F}'}(0), q_{\mathbf{F}'}(1), \dots, q_{\mathbf{F}'}(s^*), \dots, q_{\mathbf{F}'}(S)$ , then the thresholds for  $\mathbf{F}''$  can be expressed as  $q_{\mathbf{F}''}(0), q_{\mathbf{F}''}(1), \dots, q_{\mathbf{F}''}(s^*) - 1, \dots, q_{\mathbf{F}''}(S)$ , where the two policies take different actions only in state  $q_{\mathbf{F}'}(s^*)$ . Since  $Z_{\mathbf{F}}$ , the policy to obtain a point on the line segment between  $Z_{\mathbf{F}'}$  and  $Z_{\mathbf{F}''}$  is the convex combination of  $\mathbf{F}'$  and  $\mathbf{F}''$ , it should have the form shown in (11). ■

We can see that the optimal policy for the Constrained Markov Decision Process may not be deterministic. At most two elements in the policy matrix  $\mathbf{F}$ , i.e.  $f_{q_{\mathbf{F}}(s^*),s^*}$  and  $f_{q_{\mathbf{F}}(s^*),s^*+1}$ , can be decimal, while the other elements are either 0 or 1. Policies in this form also satisfy our definition of stationary threshold-based policy at the beginning of Section III.

## IV. ALGORITHM TO EFFICIENTLY OBTAIN THE OPTIMAL TRADEOFF CURVE

We design Algorithm 1 to efficiently obtain the optimal delay-power tradeoff curve and the corresponding optimal



---

**Algorithm 1** Constructing the Optimal Delay-Power Tradeoff

```

1: Construct  $\mathbf{F}$  whose thresholds  $q_{\mathbf{F}}(s) = s$  for  $s < A$  and
    $q_{\mathbf{F}}(s) = Q$  for  $s \geq A$ 
2: Calculate  $D_{\mathbf{F}}$  and  $P_{\mathbf{F}}$ 
3:  $\mathcal{F}_c \leftarrow [\mathbf{F}]$ ,  $D_c \leftarrow D_{\mathbf{F}}$ ,  $P_c \leftarrow P_{\mathbf{F}}$ 
4: while  $\mathcal{F}_c \neq \emptyset$  do
5:    $\mathcal{F}_p \leftarrow \mathcal{F}_c$ ,  $D_p \leftarrow D_c$ ,  $P_p \leftarrow P_c$ 
6:    $\mathcal{F}_c \leftarrow \emptyset$ ,  $slope \leftarrow +\infty$ 
7:   while  $\mathcal{F}_p \neq \emptyset$  do
8:      $\mathbf{F} = \mathcal{F}_p.pop(0)$ 
9:     for all  $0 < s^* \leq A$  do
10:      Construct  $\mathbf{F}'$  where  $q_{\mathbf{F}'}(s^*) = q_{\mathbf{F}}(s^*) + 1$ 
11:      and  $q_{\mathbf{F}'}(s) = q_{\mathbf{F}}(s)$  for  $s \neq s^*$ 
12:      if  $\mathbf{F}'$  is feasible and threshold-based then
13:        Calculate  $D_{\mathbf{F}'}$  and  $P_{\mathbf{F}'}$ 
14:        if  $D_{\mathbf{F}'} = D_p$  and  $P_{\mathbf{F}'} = P_p$  then
15:           $\mathcal{F}_p.append(\mathbf{F}')$ 
16:        else if  $D_{\mathbf{F}'} \geq D_p$  and  $P_{\mathbf{F}'} < P_p$  then
17:          if  $\frac{D_{\mathbf{F}'} - D_p}{P_p - P_{\mathbf{F}'}} < slope$  then
18:             $\mathcal{F}_c \leftarrow [\mathbf{F}']$ ,  $slope \leftarrow \frac{D_{\mathbf{F}'} - D_p}{P_p - P_{\mathbf{F}'}}$ 
19:             $D_c \leftarrow D_{\mathbf{F}'}$ ,  $P_c \leftarrow P_{\mathbf{F}'}$ 
20:          else if  $\frac{D_{\mathbf{F}'} - D_p}{P_p - P_{\mathbf{F}'}} = slope$  then
21:            if  $P_{\mathbf{F}'} = P_c$  then
22:               $\mathcal{F}_c.append(\mathbf{F}')$ 
23:            else if  $P_{\mathbf{F}'} > P_c$  then
24:               $\mathcal{F}_c \leftarrow [\mathbf{F}']$ ,  $slope \leftarrow \frac{D_{\mathbf{F}'} - D_p}{P_p - P_{\mathbf{F}'}}$ 
25:               $D_c \leftarrow D_{\mathbf{F}'}$ ,  $P_c \leftarrow P_{\mathbf{F}'}$ 
26:            end if
27:          end if
28:        end if
29:      end for
30:    end while
31:    Draw the line segment connecting  $(P_p, D_p)$  and
     $(P_c, D_c)$ 
32: end while

```

---

policies. Similar to [12], this algorithm takes advantage of the properties we have shown, i.e., the optimal delay-power tradeoff curve is piecewise linear, the vertices are obtained by deterministic threshold-based policies, and policies corresponding to two adjacent vertices take different actions in only one state. Therefore given the optimal policy for a certain vertex, we can narrow down the alternatives of optimal policies for its adjacent vertex. The policies corresponding to points between two adjacent vertices can also be easily generated.

Our proposed iterative algorithm starts from the bottom-right vertex of the optimal tradeoff curve, whose corresponding policy is known to transmit as much as possible. Then for each vertex we have determined, we enumerate the candidates for the next vertex. According to the properties we have obtained, we only need to search for deterministic threshold-based policies which take different actions in only one threshold. By comparing all the candidates, the next vertex will be determined by the policy candidate whose connecting line with

the current vertex has the minimum absolute slope and the minimum length. Note that a vertex can be obtained by more than one policy, therefore we use lists  $\mathcal{F}_p$  and  $\mathcal{F}_c$  to restore all policies corresponding to the previous and the current vertices.

The complexity of this algorithm is much smaller than using general methods. Since during each iteration, one of the thresholds of the optimal policy will be decreased by 1, the maximum iteration times are  $AQ$ . Within each iteration, we have  $A$  thresholds to try. For each candidate, the most time consuming operation, i.e. the matrix inversion, costs  $\mathcal{O}(Q^3)$ . Therefore the complexity of the algorithm is  $\mathcal{O}(A^2Q^4)$ .

In comparison, we also formulate a Linear Programming (LP) to obtain the optimal tradeoff curve. As demonstrated in [13, Chapter 11.5], all CMDP problems with infinite horizon and average cost can be formulated as Linear Programming. In our case, by taking  $x_{q,s} = \pi(q)f_{q,s}$  as variables, we can formulate an LP with  $QS$  variables to minimize the average delay given a certain power constraint. Due to space limitation, we provide the LP without explanations.

$$\min \frac{1}{E_a} \sum_{q=0}^Q q \sum_{s=0}^S x_{q,s} \quad (12a)$$

$$\text{s.t.} \quad \sum_{q=0}^Q \sum_{s=0}^S P_s x_{q,s} \leq P_{\text{th}} \quad (12b)$$

$$\sum_{l=\max\{0, q-A\}}^{q-1} \sum_{a=0}^A \sum_{s=0}^{l+a-q} \alpha_a x_{l,s} = \sum_{r=q}^{\min\{q+S-1, Q\}} \sum_{a=0}^A \sum_{s=r+a-q+1}^S \alpha_a x_{r,s} \quad q = 1, \dots, Q \quad (12c)$$

$$\sum_{q=0}^Q \sum_{s=0}^S x_{q,s} = 1 \quad (12d)$$

$$x_{q,s} = 0 \quad \forall q - s < 0 \text{ or } q - s > Q - A \quad (12e)$$

$$x_{q,s} \geq 0 \quad \forall 0 \leq q - s \leq Q - A. \quad (12f)$$

By solving the LP, we can obtain a point on the optimal tradeoff curve. If we apply the ellipsoid algorithm to solve the LP problem, the computational complexity is  $\mathcal{O}(S^4Q^4)$ . It means that, the computation to obtain one point on the optimal tradeoff curve by applying LP is larger than obtaining the entire curve with our proposed algorithm. This demonstrates the inherent advantage of using the revealed properties of the optimal tradeoff curve and the optimal policies.

## V. NUMERICAL RESULTS

In this section, we validate our theoretical results and the proposed algorithm by conducting LP numerical computation and simulations. We consider a practical scenario with adaptive M-PSK transmissions. The optional modulations are BPSK, QPSK, and 8-PSK. Assume the bandwidth = 1 MHz, the length of a timeslot = 10 ms, and the target bit error rate  $\text{ber} = 10^{-5}$ . Assume a data packet contains 10,000 bits, and in each timeslot the number of arriving packet could be 0,

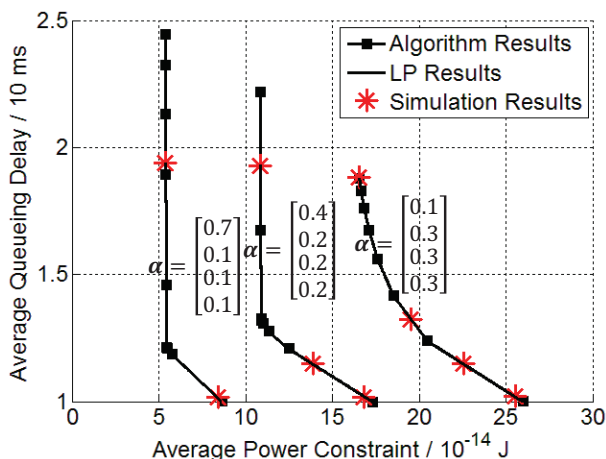


Fig. 4. Optimal Delay-Power Tradeoff Curves

1, 2 or 3. Then by adaptively applying BPSK, QPSK, or 8-PSK, we can respectively transmit 1, 2, or 3 packets in a timeslot, which means  $S = 3$ . Assume the one-sided noise power spectral density  $N_0 = -150$  dBm/Hz. The transmission power for different transmission rates can be calculated as  $P_0 = 0$  J,  $P_1 = 9.0 \times 10^{-14}$  J,  $P_2 = 18.2 \times 10^{-14}$  J, and  $P_3 = 59.5 \times 10^{-14}$  J. Set the buffer size as  $Q = 100$ .

The optimal delay-power tradeoff curves are shown in Fig. 4 and Fig. 5. In each figure, we vary the arrival process to get different tradeoff curves. As can be observed, the tradeoff curves generated by Algorithm 1 perfectly match the LP and simulation results. As proven in Theorem 1, the optimal tradeoff curves are piecewise linear, decreasing, and convex. The vertices of the curves obtained by Algorithm 1 are marked by squares. The corresponding optimal policies are threshold-based. The minimum average delay is 1 for all curves, because when we transmit as much as we can, all data packets will stay in the queue for exactly one timeslot. In Fig. 4, with the average arrival rate increasing, the curve gets higher because of the heavier workload. In Fig. 5, the three arrival processes have the same average arrival rate and different variance. When the variance gets larger, it is more likely that the queue size gets long in a short time duration, which leads to higher delay. It is interesting to characterize the effect of the variance in the arrival process, which we leave as a future work.

## VI. CONCLUSION

In this paper, we extend our previous work to obtain the optimal delay-power tradeoff and the corresponding optimal scheduling policy considering arbitrary i.i.d. arrival and adaptive transmissions. The scheduler optimizes the transmission in each timeslot according to the buffer state. We formulate this problem as a CMDP, and minimize the average delay to obtain the optimal tradeoff curve. By studying the steady-state properties and the Lagrangian relaxation of the CMDP problem, we can prove that the optimal delay-power tradeoff curve is convex and piecewise linear, on which the adjacent vertices are obtained by policies taking different actions in only one state. Based on this, the optimal policies are proven

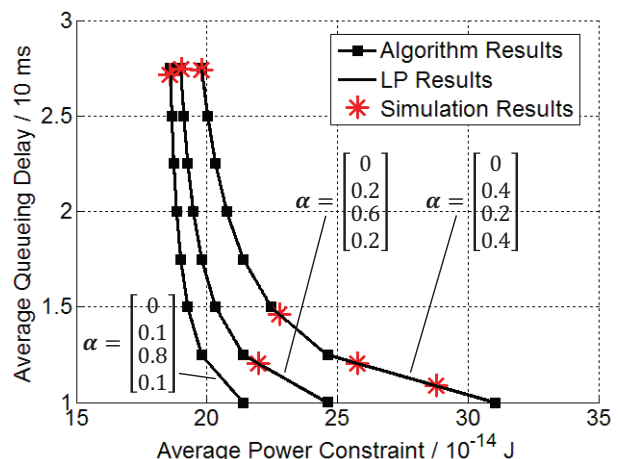


Fig. 5. Optimal Delay-Power Tradeoff Curves with the Same Arrival Rate

to be threshold-based. We also design an efficient algorithm to obtain the optimal tradeoff curve and the optimal policies. Linear Programming and simulations are conducted to confirm the theoretical results and the proposed algorithm.

## REFERENCES

- [1] B. Collins and R. L. Cruz, "Transmission policies for time varying channels with average delay constraints," in *Proc. 37th Allerton Conf. Commun. Control, Comput., Monticello, IL*, 1999, pp. 709–717.
- [2] 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, 2002.
- [3] M. Goyal, A. Kumar, and V. Sharma, "Power constrained and delay optimal policies for scheduling transmission over a fading channel," in *Proc. IEEE INFOCOM*, 2003, pp. 311–320.
- [4] I. Betteh and S. Shamai, "Optimal power and rate control for minimal average delay: The single-user case," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4115–4141, 2006.
- [5] R. Berry, "Optimal power-delay tradeoffs in fading channels—small-delay asymptotics," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3939–3952, June 2013.
- [6] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 125–144, 2004.
- [7] M. Agarwal, V. S. Borkar, and A. Karandikar, "Structural properties of optimal transmission policies over a randomly varying channel," *IEEE Trans. Autom. Control*, vol. 53, no. 6, pp. 1476–1491, 2008.
- [8] D. V. Djonin and V. Krishnamurthy, "Mimo transmission control in fading channels—a constrained markov decision process formulation with monotone randomized policies," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5069–5083, 2007.
- [9] M. H. Ngo and V. Krishnamurthy, "Monotonicity of constrained optimal transmission policies in correlated fading channels with arq," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 438–451, 2010.
- [10] B. Ata, "Dynamic power control in a wireless static channel subject to a quality-of-service constraint," *Oper. Res.*, vol. 53, no. 5, pp. 842–851, 2005.
- [11] X. Chen and W. Chen, "Delay-optimal buffer-aware probabilistic scheduling with adaptive transmission," in *Proc. IEEE/CIC ICC*, 2015, pp. 1–6.
- [12] X. Chen, W. Chen, J. Lee, and N. B. Shroff, "Delay-optimal buffer-aware scheduling with adaptive transmission," *arXiv preprint arXiv:1609.03260*, 2016.
- [13] E. Altman, *Constrained Markov decision processes*. CRC Press, 1999, vol. 7.
- [14] X. Chen, W. Chen, J. Lee, and N. B. Shroff, "Delay-optimal probabilistic scheduling in green communications with arbitrary arrival and adaptive transmission," *arXiv preprint arXiv:1702.08052*, 2017.
- [15] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.