Joint Queue-Aware and Channel-Aware Delay Optimal Scheduling of Arbitrarily Bursty Traffic over Multi-State Time-Varying Channels

Meng Wang, Juan Liu, Wei Chen, *Senior Member, IEEE,* Anthony Ephremides, *Life Fellow, IEEE*

***Abstract*—This paper is motivated by the observation that the average queueing delay can be decreased by sacrificing power efficiency in wireless communications. In this sense, we naturally wonder what the minimum queueing delay is when the available power is limited and how to achieve the minimum queueing delay. To answer these two questions in the scenario where randomly arriving packets are transmitted over multi-state wireless fading channel, a probabilistic cross-layer scheduling policy is proposed in this paper, and characterized by a constrained Markov Decision Process (MDP). Using the steady-state probability of the underlying Markov chain, we are able to derive the mathematical expressions of the concerned metrics, namely, the average queueing delay and the average power consumption. To describe the delay-power tradeoff, we formulate a non-linear programming problem, which, however, is very challenging to solve. By analyzing its structure, this optimization problem can be converted into an equivalent Linear Programming (LP) problem via variable substitution, which allows us to derive the optimal delay-power tradeoff as well as the optimal scheduling policy. The optimal scheduling policy turns out to be dual-threshold-based, which means transmission decisions should be made based on the optimal thresholds imposed on the queue length and the channel state.**

***Index Terms*—Cross-layer design, delay-power tradeoff, quality of service, probabilistic scheduling, controllable queueing system, Markov Decision Process.**

I. Introduction

FUTURE wireless networks, such as the fifth Generation (5G) of mobile network, bring more stringent QoS (quality-of-service) to support emerging applications that involve explosive mobile devices [1]. Low latency is one of the most important QoS for URLLC (Ultra-Reliable Low Latency Communications) which is a features brought by 5G [2]. At the mean time, high energy efficiency is urgently required especially for these machine nodes that are usually powered by rechargeable batteries of finite capacity. Thus, it is of great importance to ensure the required latency with finite transmit power for these users in wireless communications [3], [4].

In general, it is very challenging to derive the delay-power tradeoff in such machine-based applications, considering the random behavior of the bursty traffic, and the time-varying characteristics of wireless channels [5]. These randomness occur in different layers of the transmitter, which increases the difficulty of characterizing the delay-power tradeoff [6]. We evaluate the latency and power efficiency performances under a point-to-point transmission scenario. In such case, the cross-layer design framework, first presented in [7], can be used for reference to capture the uncertainties occurring at different layers in the last decades [8]–[14].

Within the cross-layer architecture, many works have focused on revealing the delay-power tradeoff, which can be classified into two major categories. One line of the works attempt to find the analytical delay-power tradeoffs by considering some ideal or simplified assumptions on the system model [15]–[19]. In [15], the authors proposed a scheduling policy named Lazy scheduling which assigns transmission chances based on the backlog in the queue under the assumption that the arrival times of the packets are known in advance. In [16], the authors minimized the transmission power with QoS constraits by assuming that the data arrival is known ahead of schedule and the channel is static or slow fading. This line of works mainly provide theoretical value more than engineering value, since the assumptions are too ideal to be practical. However, they are able to provide deeper insights to guide for engineering applications such as protocol design.

The works in the other category consider more complex and practical system models [9]–[11], [20]–[22]. In [9], Berry and Gallager proposed adapting the users' transmission rate and power by regulating the average power and average buffer delay over a wireless fading channel. They also focused on studying the cross-layer resource allocation in wireless fading channels for [10] and deriving the optimal power-delay tradeoff for a single user in the regime of asymptotically small delays in [11]. Ata investigated the power minimization problem subject to the packet drop rate in [20], assuming the fixed channel state, Possion packet arrival and exponentially distributed packet size. In [21], [22], the authors studied the delay-bounded packet scheduling problem with bursty traffic arrival over wireless channels. This line of works studied the delay-power curve and analyzed its property under some circumstances. While it is difficult to derive theoretical solutions

Manuscript received March 8, 2018; revised May 30, 2018 and September 14, 2018; accepted September 30, 2018. The work of Meng Wang and Wei Chen is supported by the National Science Foundation of China under Grant No. 61671269 and the National Program for Special Support for Eminent Professionals (10000-Talent Program). The work of Juan Liu is supported by the National Science Foundation of China under Grant No. 61601255. The work of Anthony Ephremides is supported by the U.S. Office of Naval Research under Grant N000141812046 and the U.S. National Science Foundation under Grants CCF1420651, CNS1526309, and CCFR1813078.
The corresponding author is Wei Chen, wchen@tsinghua.edu.cn.

M. Wang and W. Chen are with the Department of Electrical Engineering and Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (email: mwang14@mails.tsinghua.edu.cn and wchen@tsinghua.edu.cn).

J. Liu is with the College of Information Science and Engineering, Ningbo University, Zhejiang 315211, China (email: eeliujuan@gmail.com).

A. Ephremides is with the Department of Electrical and Computer Engineering and the Institute for System Research, University of Maryland, College Park, MD 20740, USA (email: etony@ece.umd.edu).

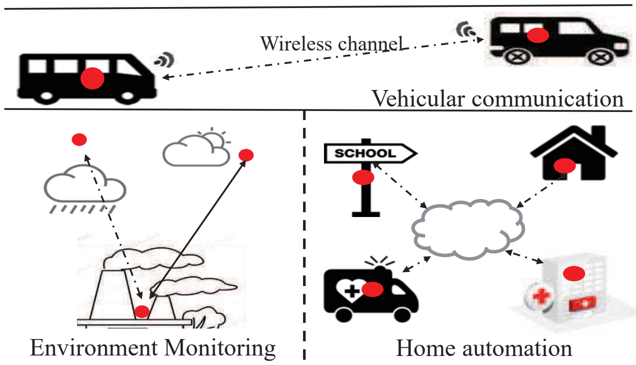


Fig. 1: Potential application scenarios in sensor network

in general cases. This line of works mainly focus on studying resource allocation solutions and designing efficient algorithms for practical usage, which is of great importance in designing delay/power-efficient wireless transmission strategies.

The cross-layer resource scheduling problem can be applied in many potential application scenarios. Dynamic resource management problems were studied under the IoT (Internet of Thing) in [23] and the Smart Grid scenarios in [24], respectively. The power constrained delay minimization problems were studied in [17] [25] for multi-access channel. The resource scheduling problem in energy harvesting was studied in [26]. Studying the point to point resource scheduling problem can provide useful insights for these important and emerging applications.

More recently, a simple probabilistic scheduling policy was proposed to achieve the minimum queueing delay under power constraint in our previous work [18], where Bernoulli packet arrivals and a two-state fading channel model were considered. Some potential application scenarios are shown in Fig.1. Further, arbitrarily random packet arrival patterns were considered to capture the impact of bursty network traffic in [27], [28] and adaptive transmission is considered in [29]. In these works, we proved that the optimal delay-power tradeoff can be achieved by applying the optimal scheduling polices which determine packet transmissions based on the threshold imposed on the queue length. The structured policy is appealing for the scheduler thanks to its ease of deployment. Hence, it inspires us to further dig into this topic. We naturally wonder if the optimal solution still has a special structure in more general scenarios and what kind of structure it may have.

In this paper, we study the delay-power tradeoff in wireless packet transmissions in a more realistic but complex communication system, where data packets are generated from an arbitrarily bursty traffic and a multi-state wireless fading channel is considered. The major challenges of this work lie in two aspects: 1) how to perform probabilistic scheduling jointly based on the randomness of the data packet arrival, the occupancy of the transmission data queue, and the time-varying characteristics of the wireless channel, and 2) how to reveal the structure of the optimal policy.

At the first task, the major challenge confronted is to build a proper cross-layer framework which includes all the system dynamics. Incorporating all these effects, our proposed scheduling policy performs joint scheduling based on the time-varying environment. Hence, it is very challenging to formulate the optimal cross-layer scheduling problem while facilitating theoretical analysis of its optimal solution. To deal with this difficulty, we propose a stochastic scheduling policy being aware of packet arrival, buffer and channel states. Then, we formulate a non-linear optimization problem to find the optimal probabilistic scheduling parameters. The challenge behind the second task is how to solve the optimal scheduling problem and derive the closed-form solution. This lies in the fact that the dimensionality of solving the optimal scheduling problem increases significantly due to the enlarged number of scheduling parameters that increases linearly with the number of channel and packet arrival states. By solving the obtained non-linear problem, we can surely obtain the optimal delay-power tradeoff. However, it is not trivial to search for the optimal solution to the non-linear optimization problem, let alone derive the optimal scheduling solution theoretically. To deal with this challenge, we first find a method to convert it to an LP problem, through which we can further analyze the structure of the optimal solution and reveal that the optimal scheduling policy has a dual-threshold-based structure step by step. By dual-threshold-based, we mean that packets should be transmitted based on the thresholds imposed on not only the queue state but also the on channel state.

The remainder of this paper is organized as follows. The system setting is introduced in Section II. In Section III, we propose the probabilistic scheduling policy to schedule packet transmissions based on the buffer and the channel states simultaneously. In Section IV, we formulate a non-linear power constrained delay minimization problem and then convert it to an equivalent LP problem. In Section V, we reveal that the optimal scheduling policy is dual-threshold-based with a rigorous mathematic proof and propose an algorithm to find simplified suboptimal policy. Simulation results are demonstrated in Section VI to validate the dual-threshold-based policy and concluding remarks are presented in Section VII. Some notations frequently used are explained as follows. Given a positive integer K, the notation $\mathbb{K}$ denotes an integer set $\{0,1,2,\cdots,K\}$ while $\mathbb{K}^+$ denotes integer set $\mathbb{K}/\{0\}$. Sets $\mathbb{W}$ and $\mathbb{W}^+$, $\mathbb{M}$ and $\mathbb{M}^+$ are defined in the same way[1].

II. System Model

We consider a wireless communication system where the source node transmits to the destination over a time-varying wireless link. As shown in Fig.2, packets of bursty traffic generated by higher-layer applications arrive at the network layer randomly, and are stored at the buffer in the data link layer. In the physical layer, the transmitter determines when to transmit the queued packets over a multi-state wireless channel, with the aid of efficient scheduling policies.

Let $a[n]$ denote the number of packets randomly arriving in the nth slot. To capture the burstiness and variability of

[1]Part of this work was published in [30], where main results were presented while most important derivations for some conclusions towards the dual-threshold-based structure were omitted due to the limited space.

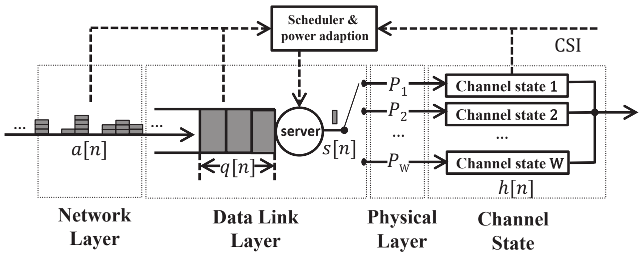


Fig. 2: System model

real-time applications, we assume an arbitrarily packet arrival pattern, i.e., the number of newly arriving packets could follow any distribution. Suppose that $a[n]$ is independent and identically distributed ($i.i.d.$). Thus, the mass probability function of $a[n]$ can be characterized by

$$\Pr\{a[n] = m\} = \theta_m, m = 0, 1, 2, \cdots \tag{1}$$

where $\theta_m \in [0, 1]$. Considering traffic shaping and admission control adopted in the system, the number of packets newly arriving in each time slot must be upper-bounded by a large integer M. In other words, there exists a positive integer M such that $\theta_m = 0$, for all $m > M$, and $\sum_{m=0}^{M} \theta_m = 1$. The average packet arrival rate $\bar{a}$ is obtained as

$$\bar{a} = \lim_{N\to\infty} sup \frac{1}{N} \sum_{n=0}^{N} a[n] = \sum_{m=0}^{M} m \cdot \theta_m. \tag{2}$$

At the source node, a buffer is employed to store the backlogged packets which cannot be sent immediately. The queue state, denoted by $q[n]$, is characterized by the number of packets in the buffer at the end of nth slot and updated as

$$q[n] = max\{min\{q[n-1] + a[n], K\} - s[n], 0\}, \tag{3}$$

where $s[n]$ is the transmitted packets in the nth time slot and K is the capacity of the buffer[2].

We adopt a W-state block fading channel model, where W is a positive integer. Let $h[n]$ denote the channel state in the nth time slot. By block fading, we mean that the channel state $h[n]$ stays invariant during each time slot and follows an $i.i.d.$ fading process across the time slots. Here, the discrete W channel states indicate different wireless channel qualities. Let $d_1 = 0 < d_2 < \cdots < d_W < d_{W+1} = \infty$ be the channel power gain levels. If the channel gain in the nth time slot ranges in interval $[d_w, d_{w+1})$, we say that the wireless channel is at state w. Since the channel quality becomes better with the increase of the index, $w = 1$ and $w = W$ represent the worst and the best channel condition, respectively. The mass probability function of $h[n]$ is described as

$$\Pr\{h[n] = w\} = \eta_w, \tag{4}$$

where $\eta_w \in [0, 1]$ and $w \in \mathbb{W}^+$.

Suppose that there exists a feedback channel through which the Channel State Information (CSI) is sent back from the receiver to the transmitter. Intuitively, the transmission power shall be adapted to the channel state to meet the requirement of successful packet delivery. Let P_w ($w \in \mathbb{W}^+$) denote the power needed to transmit one packet successfully in the channel sate w. Since more power is required to combat wireless channel fading when the channel condition is worse, it is reasonable to assume $P_1 > P_2 > \cdots > P_w > \cdots > P_W$.

In our model, we consider a fixed-rate transmission scheme which has been widely adopted in practice [31]. Without loss of generality, we assume the transmission rate is one packet per slot. Hence, at most one data packet can be delivered in each slot, namely, $s[n] \in \{0, 1\}$.

In the cross-layer design framework shown in Fig.2, the scheduler will schedule packets transmissions in each slot n based on the packet arrival state $a[n]$, the queueing state $q[n-1]$, and the channel state $h[n]$ subjected to a power constraint, as will be discussed in details in the next section where the scheduling problem is treated as a power constrained Markov Decision Process (MDP), and discussed in Section IV.

III. Probabilistic Scheduling Policy

In this section, we introduce a probabilistic scheduling policy based on which the transmitter decides whether or not to deliver one data packet to its receiver in each slot.

A. Probabilistic Scheduling

To improve the power efficiency, the transmitter should exploit a better channel state to deliver the packets to spend much less power. Thus, the source is more willing to keep silent till the channel state gets better. However, this may induce undesirable large latency waiting for good channel states, which is intolerable for serving delay-sensitive or time-critical traffics. To overcome this issue, some backlogged packets should be transmitted immediately at the cost of consuming higher power, even when the channel state may not be so good. Hence, the proposed scheduler must achieve a balance between the average delay and the power consumption.

In this work, a probabilistic cross-layer scheduling policy is proposed to schedule packet transmissions in each time slot. At the beginning of the nth time slot, the scheduler collects the current system state including the queueing state $q[n-1] = k$, the packet arrival state $a[n] = m$, and the channel state $h[n] = w$. Given $q[n-1] = k$, $a[n] = m$, and $h[n] = w$, it decides to transmit one packet with probability $f_{k+m,w}$ or keep silent with probability $1 - f_{k+m,w}$. By $f_{k+m,w}$, we mean that the scheduler can schedule packet transmissions based on the updated queue state $q[n-1] + a[n] = k+m$ after one packet arrival. The reason lies in the fact that one of the packets newly arriving at this slot can be delivered immediately. Hence, it is not necessary to distinguish between the backlogged packets and the newly arriving packets. Clearly, the transmission probability $f_{k+m,w}$ lies in the interval $[0, 1]$.

According to the above probabilistic scheduling policy, the number of transmitted packets $s[n]$ for the current slot is a random variable, the probability mass function of which is given by

$$s[n] = \begin{cases} 1 & w.p. \quad f_{k+m,w}, \\ 0 & w.p. \quad 1 - f_{k+m,w}, \end{cases} \tag{5}$$

[2]Packet overflow will occur if K is quite small. In this work, we assume that K is a sufficiently large constant such that no packet overflow will occur. In Section V, we give the conclusion that if K is greater than a threshold, the queueing length will never reach the capacity according to our proposed scheduling scheme. Thus, the max operation in (3) can be omitted.

where $k \in \mathbb{K}, m \in \mathbb{M}, w \in \mathbb{W}$ and the abbreviation $'w.p.'$ is short for $'with\ the\ probability\ of'$[3].

We aim to find the optimal policy with a set of optimal transmission probabilities $\{f^*_{k+m,w}\}$ that can minimize the average queueing delay under an average transmission power constraint.

B. Markov Decision Process

Based on the scheduling policy in section III-A, the scheduler makes decision of transmitting $s[n]$ packet(s) in every slot. The transmission decision affects the number of the packets queueing in the buffer as well as the transmission power. In this sense, we model the scheduling problem as a constrained MDP with the queue length $q[n]$ being the system state. The decision, either waiting or transmitting ($s[n] \in \{0,1\}$), is treated as one candidate action taken at the current state. Executing each action certainly causes some system costs, namely, the delay cost associated with the queue length and the power cost associated with the packet transmission. Let $\tau_{k,l}$ denote the one-step state transition probability from state $q[n-1] = k$ to state $q[n] = l$, i.e.,

$$\tau_{k,l} = \Pr\{q[n] = l \mid q[n-1] = k\}. \tag{6}$$

The transition probabilities of the underlying Markov chain are presented in *Lemma* 1.

Lemma 1. *The forward and backward state transition probabilities denoted by $\lambda_{k,m} = \tau_{k,k+m}$ and $\mu_k = \tau_{k,k-1}$ are obtained as*

$$\lambda_{k,m} = \theta_m \sum\nolimits_{w=1}^{W} \eta_w (1 - f_{k+m,w}) + \theta_{m+1} \sum\nolimits_{w=1}^{W} \eta_w f_{k+m+1,w}, \tag{7}$$

$$\mu_k = \theta_0 \sum\nolimits_{w=1}^{W} \eta_w f_{k,w}, \tag{8}$$

where $k \in \mathbb{K}$ and $m \in \mathbb{M}^+$. The state transition probability $\lambda_{k,0}$ is the probability that the queue length remains the same, given by

$$\lambda_{k,0} = \tau_{k,k} = \begin{cases} 1-\sum_{m=1}^{M} \lambda_{k,m}, & k = 0, \\ 1-\sum_{m=1}^{M} \lambda_{k,m} - \mu_k, & k \in \mathbb{K}^+. \end{cases} \tag{9}$$

Proof: Due to limited space, the detail is given in [32]. ∎

Notice that, $\tau_{k,l} = 0$ holds for $|l-k| > M$, since the queue length increases from k up to $l = k+M$ after one packet arrival. In Fig.3, we show an example of the MDP model with $M = 2$. In each time slot, $q[n]$ increases by no more than M due to one new data arrival, while decreases by one since at most one packet can be delivered. Let matrix $\mathbf{\Lambda}$ denote the $(K+1)$-by-$(K+1)$ transition probability matrix of the underlying Markov chain. The $(j+1, i+1)$-th element of $\mathbf{\Lambda}$ is transition probability $\tau_{i,j}$. The transition probability matrix $\mathbf{\Lambda}$ is a banded matrix, since the number of the newly arrival packets and departing packets are limited in one slot.

Let π_k denote the steady-state probability of the queue length being equal to k. The stationary distribution of the system state is denoted by the vector $\pi = [\pi_0, \pi_1, \cdots, \pi_K]^T$, where the superscript T denotes matrix transpose. Vectors $\mathbf{0}$ and $\mathbf{1}$ are used to denote the $(K+1)$-dimensional column vectors whose entries are zero and one, respectively. According to the property of the steady-state probability, we have $\mathbf{\Lambda}\pi = \pi$ and $\mathbf{1}^T \pi = 1$. Hence, the stationary distribution π is the solution to the following linear equations

$$\begin{bmatrix} \mathbf{Q}_K \\ \mathbf{1}^T \end{bmatrix} \pi = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \tag{10}$$

where $\mathbf{Q}_K$ is a matrix consisting of the first K rows of the generator matrix $\mathbf{Q} = \mathbf{\Lambda} - \mathbf{I}$. From Eq. (10) and *Lemma* 1, we can see that the steady-state probability π is determined by the scheduling policy with the parameters $\{f_{k+m,w}\}$.

[3]In Eq. (5), when $a[n] + q[n-1] = 0$, there is no packet waiting to be transmitted, and when $a[n] + q[n-1] > K$, packet loss will happen. Thus, $f_{0,w}$ ($w \in \mathbb{W}$) and $f_{k+m,w}$ ($k+m > K$, $w \in \mathbb{W}$) are set as zero for notational consistence.

IV. Delay and Power Tradeoff

In this section, we first analyze the two key performance metrics: the average queueing delay and the average power consumption. Then, we formulate optimization problems to describe the delay minimum power constrained scheduling problem, based on the stationary probability of the built Markov Decision Process.

A. Delay and Power Metrics

In accordance with every transmission action $s[n]$, the scheduler spends some system costs due to queue occupation and packet transmission. Given action $s[n]$, the queueing cost for buffer occupation is denoted by $C_q[n]$ and the power cost for packet transmission is denoted by $C_p[n]$, respectively, expressed as

$$C_q[n] = (q[n-1] + a[n] - s[n])^+ \text{ and } C_p[n] = P_w s[n] \tag{11}$$

As time goes by, the time-average costs can be built up as

$$Q_\Omega = \lim_{N\to\infty} \frac{1}{N} \sum\nolimits_{n=1}^{N} C_q[n]$$
$$\text{and} \quad P_\Omega = \lim_{N\to\infty} \frac{1}{N} \sum\nolimits_{n=1}^{N} C_p[n], \tag{12}$$

respectively. Considering the minus and connotative plus operators before $s[n]$ in Eq. (11), an action $s[n]$ exerts opposite influences on the buffer occupation and power consumption, which naturally leads to a tradeoff between the average delay (the Little's Law) and the average power.

The above analyses explain the average delay and the average power from the cost perspective of the scheduling policy. It's much easier to understand the tradeoff from the expressions of the two metrics given in Eq. (11). To mathematically derive the two metrics, we refer to the MDP model built in Section III-B. Once the stationary distribution π is obtained, the average queueing delay and power consumption can be derived and shown in the following theorem.

Theorem 1. *Given a probabilistic scheduling policy $\{f_{k+m,w}\}$, the average queueing delay D and power consumption P can be expressed as*

$$D = \frac{1}{\bar{a}} \sum\nolimits_{k=0}^{K} k\pi_k, \tag{13}$$

$$P = \sum\nolimits_{k=0}^{K} \pi_k \sum\nolimits_{w=1}^{W} \eta_w P_w \sum\nolimits_{m=0}^{M} \theta_m f_{k+m,w}. \tag{14}$$

Proof: Given the stationary probability distribution of the Markov chain, the average queue length can be expressed as

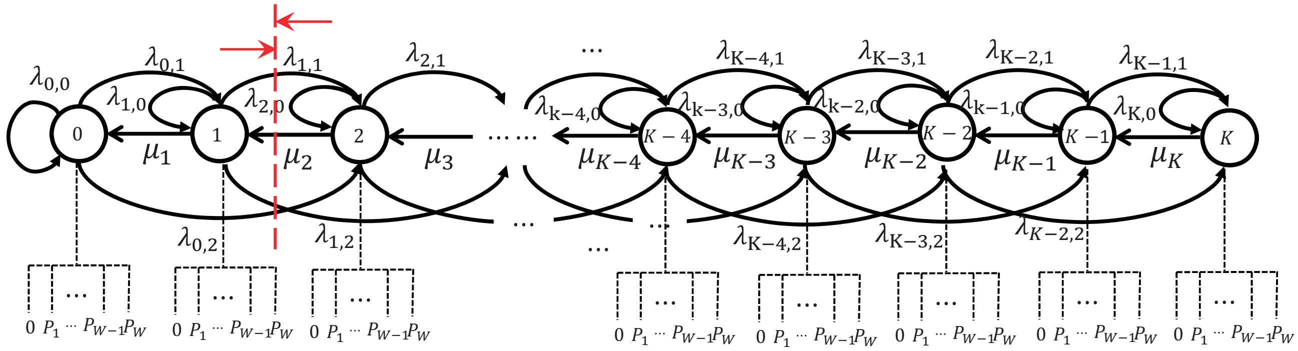


Fig. 3: Illustrative of the MDP model with $M = 2$: for each queue length, a transmission power P_w is consumed when transmitting one packet over the channel state w.

$Q = \mathbb{E}\{q[n]\} = \sum_{k=0}^{K} k\pi_k$. Then according to the Little's Law [33], the average queueing delay D can be derived as $Q/\bar{a}$ and shown in Eq. (13).

With $C_p[n] = P_w s[n]$, we have $C_p[n] = P_w$ and $C_p[n] = P_0 = 0$, respectively, when one packet is transmitted over the channel state w, i.e., $s[n] = 1$, and no transmission takes place, i.e., $s[n] = 0$. Let $\psi_{k,w}$ denote the conditional probability of $C_p[n] = P_w$ $\left(w \in \mathbb{W} = \{0 \cup \mathbb{W}^+\}\right)$ given the queue state $q[n-1] = k$ and channel state $h[n] = w$. It can be expressed as

$$\begin{aligned}\psi_{k,w} &= \Pr\{C_p[n] = P_w \big| q[n-1] = k, h[n] = w\} \\ &= \begin{cases} \sum_{m=0}^{M} \theta_m f_{k+m,w}, & w \in \mathbb{W}^+, \\ 1 - \sum_{w=1}^{W} \psi_{k,w}, & w = 0. \end{cases}\end{aligned} \tag{15}$$

By the law of total probability, the average power can be derived as

$$\begin{aligned} P = & \sum_{k=0}^{K} \sum_{w=1}^{W} \Pr\{q[n-1] = k\} \Pr\{h[n] = w\} \\ & \Pr\{C_p[n] = P_w \big| q[n-1] = k, h[n] = w\} \times P_w \\ = & \sum_{k=0}^{K} \sum_{w=1}^{W} \pi_k \eta_w \psi_{k,w} P_w \\ = & \sum_{k=0}^{K} \pi_k \sum_{w=1}^{W} \eta_w P_w \sum_{m=0}^{M} \theta_m f_{k+m,w}. \end{aligned} \tag{16}$$

■

We notice that, the steady state probability π is an implicit function of the transmission probabilities, since it is uniquely determined by the transmission probabilities $\{f_{k+m,w}\}$ based on the analyses in Section III-B. Thus, from *Theorem* 1, the average queueing delay and the average power consumption are both functions of transmission probabilities.

B. Delay-Power Tradeoff

To find the optimal scheduling policy with a set of transmission probabilities $\{f^*_{k+m,w}|, k \in \mathbb{K}, m \in \mathbb{M}, w \in \mathbb{W}\}$, we formulate an optimization problem to minimize the average queueing delay D under the power constraint P_{aver} as follows:

$$\begin{aligned} & \min_{\{f_{k+m,w}\}} \; D = \frac{1}{\bar{a}} \sum_{k=0}^{K} k\pi_k \\ & \text{s.t.} \begin{cases} P \leqslant P_{aver} & (a) \\ f_{k+m,w} \in [0,1], & (b) \\ \mathbf{Q}\pi = \mathbf{0} & (c) \\ \mathbf{1}^T \pi = 1 & (d) \\ \mathbf{0} \preceq \pi \preceq \mathbf{1} & (e) \end{cases} \end{aligned} \tag{17}$$

where $k \in \mathbb{K}$, $m \in \mathbb{M}$, $w \in \mathbb{W}$, and symbol '$\preceq$' represents the component-wise inequality between vectors. In problem (17), the objective is to minimize the average queueing delay. Constraint (17.a) denotes the maximum power constraint. Constraint (17.b) indicates the range of the optimization variables $\{f_{k+m,w}\}$. Constraints (17.c-17.e) are derived from the properties of the Markov chain. Constraint (17.e) specifies the range of the steady-state probabilities. Since problem (17) is a non-linear programming problem, it is rather difficult to obtain the optimal solution $\{f^*_{k+m,w}\}$ analytically. To make it tractable, we first convert problem (17) into an equivalent LP problem via variable substitution.

C. LP Problem Formulation

To formulate an LP problem, we introduce a set of new variables $\{y_{k,w}|k \in \mathbb{K}, w \in \mathbb{W}\}$ as

$$\begin{aligned} y_{k,w} &= \sum_{m=0}^{M} \pi_{k+1-m} \theta_m f_{(k+1-m)+m,w} \\ &= \sum_{m=0}^{M} \pi_{k+1-m} \theta_m f_{k+1,w}. \end{aligned} \tag{18}$$

In Eq. (18)[4], $\pi_{k+1-m}\theta_m f_{k+1,w}$ is the probability of transmitting one packet, i.e., $s[n] = 1$, when there are $q[n-1] = k+1-m$ data packets in the buffer and $a[n] = m$ data packets newly arriving at the transmitter. Thus, $y_{k,w}$ is the probability that there are k packets backlogged in the queue after one packet transmission over channel state w. This procedure allows us to express the objective function and the constraints of (17) as linear functions of $\{y_{k,m}\}$. Hence, we are able to convert the non-linear problem (17) into a more tractable LP problem, as shown below.

Theorem 2. *Let* $\xi = \sum_{m=1}^{M-1} \frac{m(m+1)}{2} \theta_{m+1}$ *be a constant. The optimization problem* (17) *is equivalent to the following LP*

[4] We assume the steady-state probability whose subscript is negative is zero for notation convenience. Otherwise, variable $y_{k,w}$ should be defined as $y_{k,w} = \sum_{m=0}^{min\{M,(k+1)\}} \pi_{k+1-m}\theta_m f_{k+1,w}$.

problem:

$$\min_{\{y_{k,w}\}} \; D=\frac{1}{\bar{a}^2}\Big(\sum_{k=0}^{K}\sum_{w=1}^{W}k\eta_w y_{k,w}-\xi\Big)$$
$$\text{s.t.}\begin{cases} P=\sum\limits_{k=0}^{K}\sum\limits_{w=1}^{W}\eta_w P_w y_{k,w}\leqslant P_{aver} & (a)\\ \sum\limits_{k=0}^{K}\sum\limits_{w=1}^{W}\eta_w y_{k,w}=\bar{a} & (b)\\ 0\leqslant y_{k,w}\leqslant \sum\limits_{m=0}^{M}\theta_m\sum\limits_{i=0}^{K}\sum\limits_{j=1}^{W}G_{(k+2-m,iW+j)}\cdot y_{i,j} & (c)\end{cases} \tag{19}$$

where $G_{(i,j)}$ *is the* (i,j)*-th element of* $(K+1)\times\left[W(K+1)\right]$ *matrix* $\mathbf{G}$ *which describes the relationship between the steady-state probabilities* $\{\pi_k\}$ *of the Markov chain and the variables* $\{y_{k,w}\}$*, as given by*

$$\pi_k=\sum\nolimits_{i=0}^{K}\sum\nolimits_{j=1}^{W}G_{(k+1,iW+j)}\cdot y_{i,j}, \tag{20}$$

Proof: The detail is given in *Appendix* A. ■

As shown in problem (19), there exists a minimum queueing delay for any feasible power constraint P_{aver}. Hence, the optimal queueing delay D^* can be expressed as a function of P_{aver}, i.e., $D^*=d(P_{aver})$. In the following theorem, we reveal the decreasing property of the delay-power function to discuss the structure of the optimal scheduling policy in the next section.

Theorem 3. *The delay function* $D^*=d(P_{aver})$ *monotonically decreases with the maximum transmission power* P_{aver}*.*

Proof: The detail is given in *Appendix* B. ■

Till now, we construct an LP problem to describe the delay-minimal scheduling problem under power constraint. After deriving the optimal solution $y^*_{k,w}$, we can then obtain the steady-state probability π^*_k by Eq. (20) and the optimal scheduling probability $\{f^*_{k,w}\}$ by Eq. (18). In the sequel, we show how to derive the optimal solution as well as the optimal probabilities.

V. Dual-threshold-based Policy

In this section, we focus on revealing the dual-threshold-based structure of the optimal scheduling policy. We first present the definition of the threshold-based structure.

Definition 1. *Let* $\mathbb{I}=\{0,1,2,\cdots\}$ *denote an integer set. A probability set* $\left\{\Upsilon_i|\ i\in\mathbb{I}\right\}$ *has a* i^****-threshold-based structure*** *if and only if there exists an optimal threshold* $i^*\in\mathbb{I}$ *such that* $\Upsilon_i=0,\ i<i^*$ *and* $\Upsilon_i=1,\ i>i^*$.

In what follows, we show that the optimal scheduling policy has such a structure on both the buffer state dimension and the channel state dimension, referred to as a dual-threshold-based policy. An example of the structure is illustrated in Fig.4, where positive scheduling probabilities with the indexes of buffer and channel states are plotted, and zero scheduling probabilities are omitted for briefness. In particular, given the queue state k, the optimal scheduling probabilities $\{f^*_{k,w}\}$ follows a threshold-based structure, i.e., $f^*_{k,w}=1$ for $w>T^*_k$ and $f^*_{k,w}=0$ for $w<T^*_k$, where T^*_k is the optimal threshold on the channel state dimension. Similarly, given the channel state w, the optimal scheduling probabilities $\{f^*_{k,w}\}$ has a threshold-based structure on the queue state dimension. That is, there exists an optimal threshold I^*_w on the queue state such that $f^*_{k,w}=1$ for $k>I^*_w$ and $f^*_{k,w}=0$ for $k<I^*_w$, respectively. The proof of the dual-threshold-based policy is presented in two steps in subsections A and B, in accordance with the two dimensions of the channel and buffer states. What's more, we show that among the threshold points, there is at most one joint state $(k^\star,w^\star)$ at which the optimal scheduling probability is non-zero in subsection C. A simplified threshold policy is proposed to achieve suboptimal performance in subsection D.

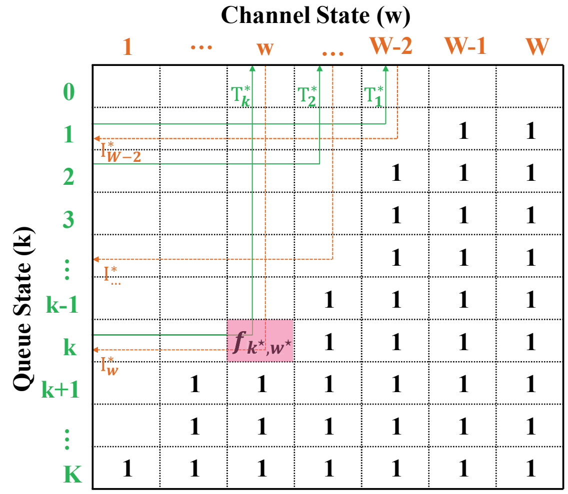


Fig. 4: The dual-threshold structure: 1) for any queue length k, the scheduling probabilities $\{f_{k,w}\}$ follows a threshold-based structure on the channel state dimension with T^*_k being the optimal threshold; 2) given the channel state w, the scheduling probabilities $\{f_{k,w}\}$ has a threshold-based structure on the queue state dimension with the optimal threshold I^*_w; 3) there is at most one threshold state at which the optimal scheduling probability is non-zero.

A. Threshold-based Structure on the Channel State Dimension

We firstly reveal the non-decreasing property of the optimal solution $\{y^*_{k,w}\}$ to problem (19). Then, we equivalently transform problem (19) into a new problem, which facilitates us to prove that $\{y^*_{k,w}\}$ has a T_k-threshold-based structure. By mapping $\{y^*_{k,w}\}$ back to $\{f^*_{k,w}\}$, the optimal scheduling policy is shown to have a threshold-based structure.

Lemma 2. *The optimal solution to problem (19)* $\{y^*_{k,w}\}$ *has the following property, for any queue length* k*,*

$$y^*_{k,w_1}\leqslant y^*_{k,w_2},\quad \forall \quad 0<w_1<w_2\leqslant W. \tag{21}$$

Proof: The detail is given in *Appendix* C. ■

Recall that, $y_{k,w}$ is the probability that there are k packets left in the queue after one packet transmission over channel w. Thus, the physical meaning of *Lemma* 2 is that, it reveals the tendency of exploiting a better channel state when one transmission has to be performed for the optimal policy.

Lemma 3. *The LP Problem* (19) *is equivalent to the following problem*

$$\begin{aligned} &\min_{\{y_{k,w}\}} \ D = \frac{1}{\bar{a}}\Big(\sum_{k=0}^{K} k\cdot \max_{w}\{y_{k,w}\} - p_0\pi_0 - \varsigma\Big) \\ &s.t. \begin{cases} \max_{w}\{y_{k,w}\} = \sum_{m=0}^{M} \theta_m \pi_{k+1-m} & (a) \\ (19.a) - (19.c), \end{cases} \end{aligned} \tag{22}$$

where $\varsigma = \sum_{i=1}^{M-1} i\theta_{i+1} - \theta_0$ *is a constant.*

Proof: The detail is given in *Appendix* D. ■

With above two lemmas, we derive the threshold structure imposed on the channel state for a given queue length of the optimal solution $\{y^*_{k,w}\}$ to problem (19) as follows:

Theorem 4. *For any queue length k, there exists an optimal integer threshold $T^*_k \in \mathbb{W}$ such that the variables $\{y^*_{k,w}\}$ has a T^*_k-threshold-based structure, i.e.,*

$$\begin{cases} y^*_{k,w} = 0, & 0 < w < T^*_k; \\ 0 \leqslant y^*_{k,w} \leqslant \sum_{m=0}^{M} \theta_m \pi^*_{k+1-m}, & w = T^*_k; \\ y^*_{k,w} = \sum_{m=0}^{M} \theta_m \pi^*_{k+1-m}, & w > T^*_k. \end{cases} \tag{23}$$

Proof: The detail is given in *Appendix* E. ■

On one hand, based on the results obtained in *Lemma* 2, *Theorem* 4 illustrates that one packet can only be transmitted if the channel state is better than a threshold. On the other hand, with the bond between $\{y^*_{k,w}\}$ and $\{f^*_{k,w}\}$, the threshold structure in *Theorem* 4 reflects the structure of the optimal scheduling policy $\{f^*_{k,w}\}$. Specifically, the optimal scheduling probability $f^*_{k,w}$ is derived according to Eq. (18) and given as 1) $f^*_{k,w} = 0$, if $w < T_k$; 2) $f^*_{k,w} = 1$, if $w > T_k$; 3) $f^*_{k,T_k} = y^*_{k-1,T_k}\Big(\sum_{m=0}^{M} \pi^*_{k+1-m}\theta_m\Big)^{-1}$. Thus, $\{f^*_{k,w}\}$ also satisfies *Definition* 1 and the optimal scheduling policy has a threshold-based structure on the channel state dimension for any given queue length k.

B. Threshold-based Structure on the Queue Length Dimension

It is not a trivial work to reveal the threshold structure on the queue state dimension straightforwardly due to the highly complicated relationship between the variables $\{y^*_{k,w}\}$. Thus, we turn to the scheduling action $s[n]$ taken by the optimal policy. Then, we map the transmission action $s[n]$ back to the scheduling probability $\{y^*_{k,w}\}$ and find that the optimal policy also has an I^*_w-threshold-based structure on the queue state dimension.

Lemma 4. *For a given channel state w, there exists an optimal integer threshold $I^*_w \in \mathbb{K}$ such that the optimal transmission action $s^*[n]$ has the I^*_w-threshold structure, namely*

$$s^*[n] = \begin{cases} 0, & t[n] < I^*_w; \\ 1, & t[n] > I^*_w, \end{cases} \tag{24}$$

where $t[n] = q[n-1] + a[n]$ denotes the updated queue state after one new packet arrival in the nth time slot.

Proof: The detail is given in *Appendix* F. ■

In *Lemma* 4, we show that the optimal transmission action $s^*[n]$ is determined based on the updated queue state $t[n]$ and the optimal threshold I^*_w. Together with Eq. (5), we can connect $s[n]$ to the scheduling probability $f^*_{k,w}$, and reveal that the probabilities $\{f^*_{k,w}\}$ also depend on the updated queue state $t[n] = k$ and the optimal threshold I^*_w: 1) $f^*_{k,w} = 0$, if $k < I_w$; 2) $f^*_{k,w} = 1$, if $k > I_w$. Thus, the optimal policy is proved to has a threshold-based structure on the queue length for any given channel state.

C. Dual-threshold-based Policy

The optimal scheduling policy turns out to be a dual-threshold-based policy, as illustrated in Fig.4. We further strengthen this result by specifying the values on the threshold points in what follows.

Theorem 5. *(1) The optimal scheduling policy corresponds to a dual-threshold policy: a) For any queue length k, there exists a threshold $T^*_k \in \mathbb{W}$, $f^*_{k,w} = 0$ for $w < T^*_k$ and $f^*_{k,w} = 1$ for $w > T^*_k$; b) There exists $T^*_1 \geqslant T^*_2 \geqslant \cdots \geqslant T^*_K$; (2) Among the threshold points, there is at most one joint state $(k^\star, w^\star)$ at which the optimal scheduling probability is non-zero.*

Proof: Conclusion (1-a) presents exactly the threshold structure obtained in subsection A. Based on the threshold structure imposed on the buffer states, we present in (1-b) the non-increasing property of $\{T^*_k\}$. Conclusion (2) comes from the fact that the optimal solution to an LP problem is always at a corner point of its feasible region. Its detailed proof can be seen in *Appendix* G in [32] due to space limitation. ■

According to our proposed scheduling scheme, once the queue length exceeds $\max_w\{I^*_w\}$, one packet will be transmitted whatever the channel state. Thus, if we set the buffer capacity $K > \max_w\{I^*_w\}$, the queueing length will never reach the capacity and no packet overflow will occur. The threshold structure is a tradeoff result of reducing the queueing delay and saving power resource. An intuition explanation that explains why the policy has such a structure can be found in *Appendix* H in [32] due to space limitation.

D. The Suboptimal scheduling Policy

It is not a trivial work to obtain closed-form expressions of the thresholds even if we have revealed their properties in *Theorem* 5. By solving the LP problem, we surely can obtain optimal thresholds and the non-zero scheduling parameter that might exist at one of the joint threshold points. Otherwise, we have to resort to some search methods to find these optimal thresholds directly. Searching the optimal policy by traversing all possible candidates of the threshold policy can be done by performing intensive computations. In what follows, we develop a structured search algorithm to find a suboptimal solution by fully exploiting the non-increasing properties of the optimal thresholds, as presented in *Theorem* 5. In other words, this property helps to reduce the search space of the candidate threshold points significantly.

In detail, combing the non-increasing property of the threshold points T_k, i.e., $T_{k_1} \geqslant T_{k_2}$ if $k_1 \leqslant k_2$, and the fact that

Algorithm 1 An algorithm to find the suboptimal scheduling policy

Input:
The average power constraint: P_{aver};
The dimension of the channel state: W;
The buffer capacity: K;
The table that with a sufficiently large capacity: $Table$;
Output:
The threshold on the queue states: k°;
The threshold on the channel states for $(0, k^{\circ}]$: w_1°;
The threshold on the channel states for $(k^{\circ}, K]$: w_2°;
if $Table == NULL$ **then**
  initialize $Table[K][W][W][2] = \infty$; # build up the table that stores the delay and power for all the deterministic policies
  for each $k^{\diamond} \in \{1, 2, \cdots, K\}$ **do**
    for each $w_1 \in \{1, 2, \cdots, W\}$ **do**
      for each $w_2 \in \{1, 2, \cdots, W\}$ **do**
        Set the scheduling parameters as:
        if $k \leqslant k^{\diamond}$ **then**
          $f_{k,w} = 0$, if $w \leqslant w_1^{\circ}$;
          $f_{k,w} = 1$, if $w > w_1^{\circ}$;
        else
          $f_{k,w} = 0$, if $w \leqslant w_2^{\circ}$;
          $f_{k,w} = 1$, if $w > w_2^{\circ}$;
        end if
        Calculate the average queueing delay by Eq. (13): $Delay$;
        Calculate the average power by Eq. (14): $Power$;
        $Table[k][w_1][w_2] = [Delay\ Power]$;
      end for
    end for
  end for
end if
$index \leftarrow \{Table : Power <= P_{aver}\}$ # look up the table, find the policies that consume less power than P_{aver}
$delay \leftarrow Min\{Table[index] : Delay\}$ # find the policy that generates the smallest delay
$[k^{\circ}, w_1^{\circ}, w_2^{\circ}] \leftarrow GetIndex\{Table : Delay == delay\}$ # return the threshold parameters of the suboptimal policy

the buffer capacity K is usually greater than the number of channel states W, we know some neighbor buffer states are likely to share a same threshold T_k. Based on this property, we can reduce the number of the thresholds points that need to be searched. In detail, the queue length range $[0, K]$ can be divided into several small intervals, each of which is assigned one threshold imposed on the channel states. Thus, we only need to determine how to divide the queue states and assign one threshold for each small interval. The simplified suboptimal policy is given in *Algorithm* 1, where the total queue states are divided into two sub-intervals. A table can be built up once for all to store the induced delay and power metrics for all the $\frac{1}{2}KW^2$ simple policies. Then, to obtain the suboptimal policy for a given power constraint, we only need to look up the table and return the thresholds. The performance can be further improved by assigning one scheduling probability to some threshold points.

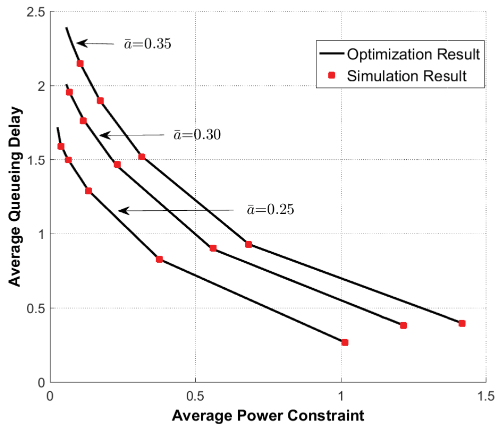
VI. Numerical Results

In this section, simulation results are given to validate the derived dual-threshold-based scheduling policy and to demonstrate its potential. For performance comparison, theoretical results of the optimal delay-power function $D^*(P_{aver})$ are obtained by solving the LP problem (19). Meanwhile, simulation results are obtained by applying the dual-threshold-based scheduling policy with the optimal transmission parameters. In simulations, data packets are generated following a given probabilistic distribution $\{\theta_m\}$. The W-state block fading channel model is adopted and follows with probability $\{\eta_m\}$. Each simulation runs over 10^6 time slots. As shown in Fig.5-8, the theoretical and simulation results are plotted by lines (solid or dashed) and marked by red square dots, respectively.

Fig.5 plots the delay-power tradeoff curves under different average packet arrival rates. The simulation results are in good agreement with the theoretical results, which validates the optimality of the derived dual-threshold-based policy. The delay-power tradeoff curve is piecewise linear since the threshold-based is obtained as the linear combinations of deterministic scheduling parameters. Besides, the average delay monotonically decreases with the maximum average power, as stated in *Theorem* 3. When the power constraint P_{aver} decreases to zero, the queueing delay increases dramatically to infinity, which implies that the queueing system is unstable. Given the same power constraint, the queueing delay increases with the packet arrival rate since more packets are detained in the buffer due to lack of transmission opportunities.

In Fig.6, we evaluate the effect of the burstiness of the packet arrival on the optimal delay-power tradeoff curves, considering different packet arrival patterns, namely, the Bernoulli arrival and the bursty arrival. We can see that the proposed scheduling policy has a better delay-power tradeoff performance when the packet arrivals follow the Bernoulli distribution rather than the more bursty probabilistic distribution (with larger variance), subject to the same average arrival rate. This is due to the fact that the bursty packet arrivals bring more randomness to the queueing system. The average queueing delay decreases with the increase of the power constraint and remains constant when the power constraint exceeds a constant P_{max}. In other words, the delay-power curve becomes flat after an inflection point (D^*_{min}, P_{max}), where D^*_{min} is the globally minimum delay and P_{max} denotes the power consumption that the source spends to keep transmitting packets as long as the buffer is not empty, regardless of the channel state. However, the value of P_{max} is identical for the two different patterns. The value of D_{min} is able to reach zero for the Bernoulli arrival since the transmission rate is fixed as one packet per slot and is greater than zero due to the burstiness.

Inspired by the observation in Fig.6, we further demonstrate the delay-power tradeoffs in Fig.7 for the packet arrivals have the same average arrival rate and different variances. It is observed that a higher queueing delay is induced when the

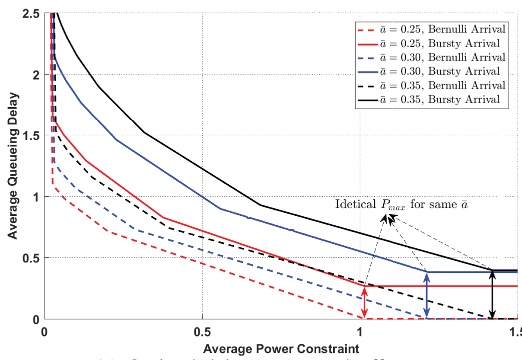


(a) Optimal delay-power tradeoff curves

$M = 2, \quad W = 4$		
$\bar{a} = 0.35$	$\theta_1 = 0.17$	$\theta_2 = 0.09$
$\bar{a} = 0.30$	$\theta_1 = 0.14$	$\theta_2 = 0.08$
$\bar{a} = 0.25$	$\theta_1 = 0.15$	$\theta_2 = 0.05$
$[\eta_1, \eta_2, \eta_3, \eta_4] = [0.135, 0.239, 0.232, 0.394]$		
$[P_1, P_2, P_3, P_4] = [10.14, 0.16, 0.08, 0.04]$		

(b) Simulation settings

Fig. 5: Optimal delay-power tradeoff curves under different arrival rates



(a) Optimal delay-power tradeoff curves

For Bernulli arrival:		
$\bar{a} = 0.25$	$\theta_1 = 0.25$	$\theta_0 = 0.75$
$\bar{a} = 0.30$	$\theta_1 = 0.30$	$\theta_0 = 0.70$
$\bar{a} = 0.35$	$\theta_1 = 0.35$	$\theta_0 = 0.65$
For Bursty arrival:		
See Fig. 4(b)		

(b) Simulation settings

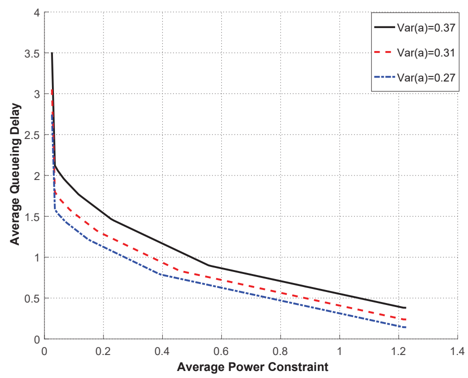
Fig. 6: The effect of the burstiness of the arrival

data arrival variance is larger. Due to higher bursty arrivals, some packets have to wait for a longer time before they are transmitted, which leads to a larger queueing delay.

In Fig.8, we demonstrate the theoretical results to validate the dual-threshold-based structure of the optimal scheduling policy, which are in agreement with the structure shown in Fig.4. The transmission probabilities reveal a threshold-based structure on both the channel state dimension and the queue length dimension. In Fig.8(a), the threshold is in channel state 1 and queue length 8 while in Fig.8(b), it is in channel state 1 and queue length 3. Thus, transmission is much easier to occur in Fig.8(b), which corresponds to a higher power consumption. That is, the scheduler makes use of the power resource mainly by adjusting the threshold point for quite different power constraints. In Fig.8(b) and Fig.8(c), it's calculated for both scenarios that the threshold is in channel state 1 and queue length 3. However the scheduler makes a decision of transmitting one packets with probability 0.2377 in Fig.8(b) and 0.4899 in Fig.8(c) on the threshold point, respectively. That is, the scheduler makes full use of the power resource mainly by adjusting the transmission probability on the threshold point for slight different power constraints.

In Fig.9, we plot the optimal delay-power curve of our proposed scheme and 1000 delay-power points of the deterministic policy with the binary transmission parameters $f_{k,w} \in \{0, 1\}$ randomly generated. As can be seen from this figure, the delay-power tradeoff curve is the lower boundary of the convex hull of the achievable delay-power region, which is in accordance with the conclusion proved in [21] that the optimal probabilistic policy can be constructed by the convex combination of deterministic scheduling policies. Hence, our proposed optimal scheduling policy outperforms any deterministic scheduling policies given the same power constraint. Meanwhile, our proposed stochastic scheduling policy with the optimal thresholds and scheduling parameters can achieve the same optimal delay-power tradeoff performance as the optimal scheduling policies found by the DP method.

In Fig.10, we plot the delay-power tradeoff curves of the optimal policy and suboptimal policy that applies *Algorithm* 1 to find the suboptimal thresholds $[k^\circ,\ w_1^\circ,\ w_2^\circ]$. We also plot the delay-power tradeoff of the improved suboptimal policy with the suboptimal thresholds $[k^\circ,\ w_1^\circ,\ w_2^\circ]$ and the scheduling parameter $f_{k,w}$. Surely, the optimal scheduling policy achieves the best delay-power tradeoff by exploiting the

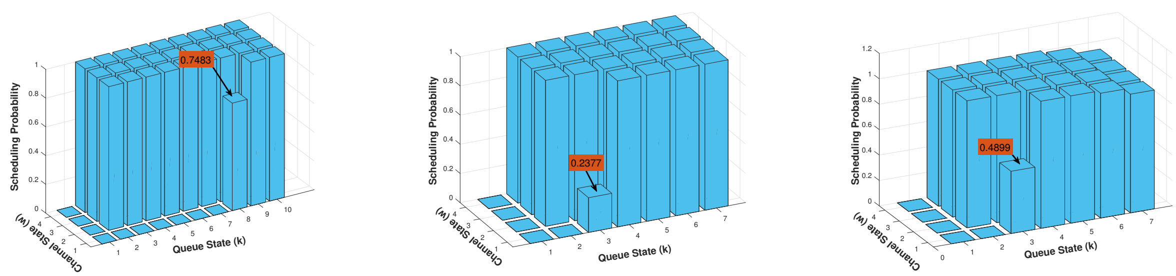


(a) Optimal delay-power tradeoff curves

$\bar{a} = 0.3$		
$Var(a) = 0.37$	$\theta_1 = 0.14$	$\theta_2 = 0.08$
$Var(a) = 0.31$	$\theta_1 = 0.20$	$\theta_2 = 0.05$
$Var(a) = 0.27$	$\theta_1 = 0.24$	$\theta_2 = 0.03$
Other settings:		
See Fig. 4(b)		

(b) Simulation settings

Fig. 7: Optimal delay-power tradeoff curve in different arrival variances



(a) Power constraint $P_{aver} = 0.10$, (b) Power constraint $P_{aver} = 0.53$, (c) Power constraint $P_{aver} = 0.63$

Fig. 8: The dual-threshold-based policy: $\bar{a} = 0.35$.

power resource to transmit in the most efficient way. Compared to the optimal policy, the suboptimal one achieves a zigzag delay-power tradeoff curve. This is because that it schedules transmissions based on the suboptimal thresholds which could remain the same for a range of power. Comparable with the optimal one, the improved suboptimal policy performs much better, since it can exploit the suboptimal thresholds together with the scheduling probability $f_{k,w}$ to schedule transmissions efficiently.

VII. Conclusion

In this paper, we studied the power-constrained delay-optimal scheduling problem in wireless systems, where arbitrary packet arrivals and multi-state block-fading channels were considered. A probabilistic queue-aware and channel-aware scheduling policy was proposed to schedule packet transmissions over a W-state wireless fading channel and investigated in the framework of constrained MDP. Through theoretical analysis, we reveal the dual-threshold-based structure of the optimal scheduling policy. It is found that the optimal scheduler always seeks to exploit a good channel while maintaining a relatively short queue as possible to reduce the latency. To this end, the scheduler should schedule packet transmissions based on the queue state and the channel state. Specifically, given a channel state, if the queue length exceeds the threshold, the transmitter should transmit to decrease the latency. Otherwise, it should keep silent to save power. In the future, we will extend this work to more general scenarios with adaptive-rate transmission and/or multi-user scheduling.

Appendix A
The proof of *Theorem* 2

In this appendix, we show that problem (17) can be equivalently converted into LP problem (19) with variables $\{y_{k,w}\}$ being the optimization variables. To make it clear, we explain the transformation procedure in the following five steps. We first present the equivalent expressions of the average queueing delay and the power constraint in Part A-1. Secondly, we specify the ranges of optimization variables $\{y_{k,w}\}$ corresponding to constraint (17.b) in Part A-2. Then, we reformulate constraints (17.c) and (17.d) in Part A-3 and Part A-4, respectively. Finally, we explain why constraints (17.c) and (17.e) are not shown in the LP problem (19) in Part A-5.

1) The average queueing delay and the power constraint can be re-expressed as :

$$\begin{cases} D = \frac{1}{\bar{a}^2}\Big(\sum\limits_{k=0}^{K}\sum\limits_{w=1}^{W} k\eta_w y_{k,w} - \xi\Big), \\ P = \sum\limits_{k=0}^{K}\sum\limits_{w=1}^{W} \eta_w P_w y_{k,w}, \end{cases} \tag{25}$$

where $\xi = \sum_{m=1}^{M-1} \frac{m(m+1)}{2}\theta_{m+1}$.

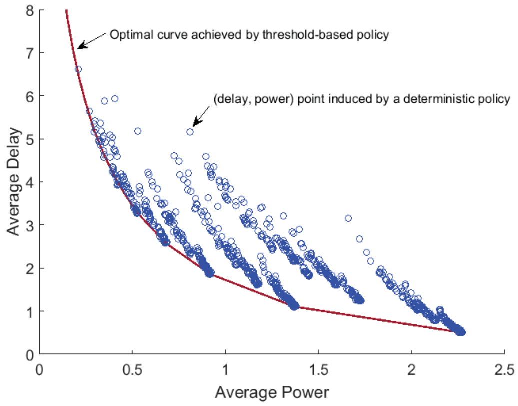


Fig. 9: The optimal delay-power tradeoff is the lower boundary of the achievable (D, P) region: the arrival distribution is $\bar{a} = 0.55$, $\theta_1 = 0.3$, $\theta_2 = 0.125$; the channel distribution is $[\eta_1 = 0.6, \eta_2 = 0.4]$ and $[P_1, P_2] = [10.14, 0.103]$.

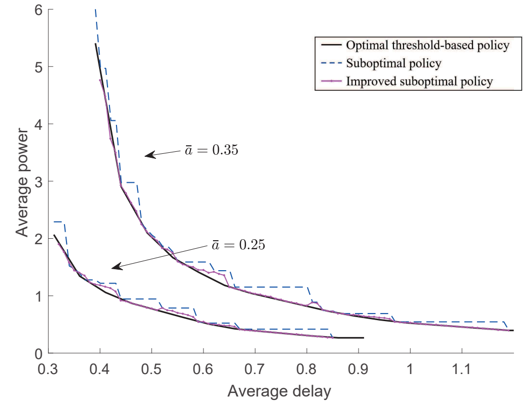


Fig. 10: The tradeoff curve induced by the suboptimal policy: the arrival distribution is $\bar{a} = 0.55$, $\theta_1 = 0.3$, $\theta_2 = 0.125$; the channel distribution is $[\eta_1 = 0.135, \eta_2 = 0.239, \eta_3 = 0.232, \eta_4 = 0.394,]$ and $[P_1, P_2, P_3, P_4] = [10, 5, 2, 1]$.

Firstly, we re-express the average queueing delay. Adding the weighted sum of the terms $\sum_{w=1}^{W} \eta_w y_{k,w}$ with k being the weight, we have

$$\begin{aligned}\sum_{k=0}^{K} k \sum_{w=1}^{W} \eta_w y_{k,w} &\overset{①}{=} \sum_{k=0}^{K} k \sum_{m=0}^{M-1} \pi_{k-m} \sum_{i=m+1}^{M} \theta_i \\ &\overset{②}{=} (\theta_1 + \theta_2 + \cdots + \theta_M) \sum_{k=0}^{K} k\pi_k \\ &\quad + (\theta_2 + \cdots + \theta_M) \sum_{k=0}^{K} k\pi_{k-1} + \cdots \\ &\quad + \theta_M \sum_{k=0}^{K} k\pi_{k-(M-1)} \\ &\overset{③}{=} (\theta_1 + \theta_2 + \cdots + \theta_M)Q + (\theta_2 + \cdots + \theta_M)(Q+1) \\ &\quad + \cdots + \theta_M(Q + (M-1)) \\ &\overset{④}{=} \bar{a}Q + \sum_{m=1}^{M-1} \sum_{i=1}^{m} i\theta_{m+1} \\ &\overset{⑤}{=} \bar{a}Q + \sum_{m=1}^{M-1} \frac{m(m+1)}{2}\theta_{m+1} = \bar{a}Q + \xi, \end{aligned} \tag{26}$$

where equality ① is derived by substituting (30), equality ② is obtained by expressing each term in $\pi_{k-m} \sum_{i=m+1}^{M} \theta_i$ separately for $m = 0, 1, 2, \cdots, M-1$, equality ③ comes from the definition of the average queue length, equality ④ is obtained by substituting $\bar{a} = \sum_{m=0}^{M} m \cdot \theta_m$, and equality ⑤ stems from $\sum_{i=1}^{m} i = m(m+1)/2$. Thus, the average queue length is:

$$Q = \frac{1}{\bar{a}}\Big(\sum_{k=0}^{K} \sum_{w=1}^{W} k\eta_w y_{k,w} - \xi\Big). \tag{27}$$

According to the Little's Law, we obtain the average queueing delay as given in Eq. (25).

Secondly, we express the average power as a function of variables $\{y_{k,m}\}$. From the definition of variable $y_{k,w}$ in Eq. (18), we have

$$\sum_{k=0}^{K} \pi_k \sum_{m=0}^{M} \theta_m f_{k+m,w} = \sum_{k=0}^{K} y_{k,w}. \tag{28}$$

By substituting Eq. (28) into Eq. (14), we obtain the average power as presented in Eq. (25).

2) The variable $y_{k,w}$ satisfies the following inequalities:

$$0 \leqslant y_{k,w} \leqslant \sum_{m=0}^{M} \theta_m \pi_{k+1-m}. \tag{29}$$

We know that probability $f_{k+1,w}$ takes its value from the interval $[0, 1]$. By substituting $f_{k+1,w} = 0$ and $f_{k+1,w} = 1$ in Eq.(18), we get the lower and upper bounds of variable $y_{k,w}$, respectively. In this way, the range of variable $y_{k,w}$ is specified by the inequalities in Eq. (29).

3) The constraint (17.c) can be equivalently expressed as:

$$\sum_{m=0}^{M-1} \pi_{k-m} \sum_{i=m+1}^{M} \theta_i = \sum_{w=1}^{W} \eta_w y_{k,w} = \sum_{i=0}^{M-1} \pi_{k-i} r_i. \tag{30}$$

where $r_i = \sum_{m=i+1}^{M} \theta_m$.

In fact, constraint (17.c) denotes the steady-state equilibrium equation of the underlying Markov chain:

$$\pi_{k+1}\mu_{k+1} = \sum_{m=0}^{M-1} \pi_{k-m} \sum_{i=m+1}^{M} \lambda_{k-m,i}, \quad k \in \mathbb{K}. \tag{31}$$

We can obtain the following conclusion

$$\sum_{i=m}^{M} \lambda_{k,i} = \sum_{i=m}^{M} \theta_i - \theta_m \sum_{w=1}^{W} \eta_w f_{k+m,w}, \quad m \in \mathbb{M} \tag{32}$$

by adding up the terms $\{\lambda_{k,i} \mid i = m, \cdots, M\}$ in Eq. (7).

By substituting Eq. (8) and Eq. (32) into Eq. (31), we have

$$\begin{aligned}&\pi_{k+1}(\theta_0 \sum_{w=1}^{W} \eta_w f_{k+1,w}) \\ &= \sum_{m=0}^{M-1} \pi_{k-m}(\sum_{i=m+1}^{M} \theta_i - \theta_{m+1} \sum_{w=1}^{W} \eta_w f_{k-m+m+1,w}).\end{aligned} \tag{33}$$

Then, we can further simplify this equation as

$$\begin{aligned}\sum_{m=0}^{M-1} \pi_{k-m} \sum_{i=m+1}^{M} \theta_i &= \sum_{m=-1}^{M-1} \pi_{k-m}\theta_{m+1} \sum_{w=1}^{W} \eta_w f_{k+1,w} \\ &= \sum_{w=1}^{W} \eta_w(\sum_{m=-1}^{M-1} \pi_{k-m}\theta_{m+1} f_{k+1,w}) \\ &= \sum_{w=1}^{W} \eta_w y_{k,w}.\end{aligned} \tag{34}$$

4) The constraint (17.d) can be re-expressed as $\sum_{k=0}^{K} \sum_{w=1}^{W} \eta_w y_{k,w} = \bar{a}$.: Summarizing all the k on both sides of Eq. (30), we have

$$\sum_{k=0}^{K} \sum_{w=1}^{W} \eta_w y_{k,w} = \sum_{k=0}^{K} \sum_{m=0}^{M-1} \pi_{k-m} \sum_{i=m+1}^{M} \theta_i$$

$$= \sum_{m=0}^{M-1} \sum_{i=m+1}^{M} \theta_i = \bar{a}, \tag{35}$$

where the second equality holds because the normalization of the steady-state probabilities, i.e., $\sum_{k=0}^{K} \pi_k = 1$.

To construct the LP problem (19), we need to express π_k as a linear function of variables $\{y_{k,m}\}$. For ease of exposition, we introduce a $(K+1)\times\left[W(K+1)\right]$ constant matrix $\mathbf{G}$ to describe the relationship between $\{\pi_k\}$ and $\{y_{k,m}\}$ based on Eq. (34). The $(k+1)$-th row vector of matrix $\mathbf{G}$, denoted by g_{k+1}, is given as

$$\begin{cases} \mathbf{g}_{k+1} &= \frac{1}{r_0}\mathbf{l}_1, \qquad k = 0, \\ \mathbf{g}_{k+1} &= \frac{1}{r_0}(\mathbf{l}_{k+1} - \sum_{i=1}^{M-1} r_i \mathbf{g}_{k-i}), \quad k \in \mathbb{K}. \end{cases} \tag{36}$$

In Eq. (36), $\mathbf{l}_{k+1}$ is a $W(K+1)$-dimensional vector whose $(Wk+w)$th element is η_w while the other elements are zero. In this way, the steady-state probabilities $\{\pi_k\}$ can be linearly expressed by the set of variables $\{y_{k,w}\}$, namely,

$$\pi_k = \sum_{i=0}^{K} \sum_{j=1}^{W} G_{(k+1,iW+j)} \cdot y_{i,j}, \tag{37}$$

where $G_{(i,j)}$ is the $(i \times j)$-th element of matrix $\mathbf{G}$.

5) Constraints (17.c) and (17.e) can be left out in the new equivalent LP problem: The constraint (17.c) is incorporated in the derivation of Eq. (37) which is the reason that it can be left out. As for constraint (17.e), since the elements in matrix $\mathbf{G}$ and variable $y_{i,j}$ are non-negative, as given by Eq. (37), π_k is clearly non-negative. On the other hand, since r_i (c.f. Eq. 30) is non-negative and $\sum_{w=1}^{W} \eta_w y_{k,w}$ is equal to or less than one, π_k is no more than one. Thus, constraint (17.e) is also implied in constraints (19.c).

In summary, we have shown that how to construct the LP problem (19) step by step equivalently from problem (17), which completes the proof of *Theorem* 2.

Appendix B
The proof of *Theorem* 3

Suppose $\{y_{k,w}\}$ is a set of variables that minimize the average delay D under constraints $\left[19.\text{(a-c)}\right]$. The corresponding transmission power P is equal to $\sum_{k=0}^{K}\sum_{w=1}^{W} \eta_w P_w y_{k,w}$. We show that if there exists another set of variables $\{\hat{y}_{k,w}\}$ which cost a higher power $\hat{P} > P$ to transmit, a smaller queueing delay $\hat{D} < D$ will be induced.

Let k_{th} be a positive integer. We construct variables $\{\hat{y}_{k,w}\}$ as follows:

$$\begin{cases} \hat{y}_{k,w} = y_{k,w} + \Delta y_{k,w}, & k \leqslant k_{th}; \\ \hat{y}_{k,w} = y_{k,w} - \Delta y_{k,w}, & k > k_{th}, \end{cases} \tag{38}$$

where $\{\Delta y_{k,w}\}$ are set as non-negative quantities to make sure that variables $\{\hat{y}_{k,w}\}$ meet constraints $\left[19.\text{(b-c)}\right]$ and the constraint $\hat{P} - P > 0$ is satisfied. Let $\Delta P = \hat{P} - P$, we then have

$$\Delta P = \sum_{w=1}^{W} \eta_w P_w \Big(\sum_{k=0}^{k_{th}} \Delta y_{k,w} - \sum_{k=k_{th}+1}^{K} \Delta y_{k,w} \Big) > 0, \tag{39}$$

which means that $\Delta y_{k,w}$ strictly stays positive for some $k \leqslant k_{th}$.

In this way, the average delay gap satisfies the following inequality

$$\begin{aligned} \hat{D} - D &= \frac{1}{\bar{a}^2} \sum_{k=0}^{K} \sum_{w=1}^{W} k\eta_w (\hat{y}_{k,w} - y_{k,w}) \\ &= \frac{1}{\bar{a}^2} \sum_{w=1}^{W} \eta_w \Big(\sum_{k=0}^{k_{th}} k\Delta y_{k,w} - \sum_{k=k_{th+1}}^{K} k\Delta y_{k,w} \Big) \\ &< \frac{1}{\bar{a}^2} \sum_{w=1}^{W} \eta_w \Big[k_{th} \sum_{k=0}^{k_{th}} \Delta y_{k,w} - (k_{th}+1) \sum_{k=k_{th}+1}^{K} \Delta y_{k,w} \Big]. \end{aligned} \tag{40}$$

Since both $\{\hat{y}_{k,w}\}$ and $\{y_{k,w}\}$ meet constraint (19.b), we have

$$\sum_{w=0}^{W} \eta_w \sum_{k=0}^{k_{th}} \Delta y_{k,w} = \sum_{w=0}^{W} \eta_w \sum_{k=k_{th}+1}^{K} \Delta y_{k,w}. \tag{41}$$

Combining Eq. (40) and Eq. (41), we know $\hat{D} < D$.

We have proven that if $P'_{aver} < P''_{aver}$, then $d(P''_{aver}) < d(P'_{aver})$. Thus, the delay-power tradeoff function $d(\cdot)$ is a monotonically decreasing function of the average power.

Appendix C
The proof of *Lemma* 2

We adopt the proof by contradiction, namely, if there exists a set of optimal variables $\{y_{k,w}\}$ that do not meet Eq. (21), we can find another set of variables $\{\hat{y}_{k,w}\}$ that can use less power to obtain the same queueing delay, which contradicts *Theorem* 3.

Suppose that, there exists a set of optimal variables $\{y_{k,w}\}$, in which there are $0 < w_1 < w_2 \leqslant W$ that make $y_{k,w_1} > y_{k,w_2}$. With this solution, the minimum delay D can be achieved at the power cost of Γ_k. We construct another set of $\{\hat{y}_{k,w}\}$ as

$$\hat{y}_{k,w} = \begin{cases} y_{k,w}, & w > w_2 \\ y_{k,w_1}, & w = w_2 \\ y_{k,w}, & w_1 < w < w_2 \\ y_{k,w} - \Delta y_{k,w}, & w \leqslant w_1, \end{cases} \tag{42}$$

where the quantities $\{\Delta y_{k,w}, w \leqslant w_2\}$ are all non-negative reals that meet the constraints that $\sum_{w=1}^{w_1} \eta_w \Delta y_{k,w} = \eta_{w_2}(y_{k,w_1} - y_{k,w_2})$ and $0 \leqslant \hat{y}_{k,w} \leqslant \hat{y}_{k,w+1}$. Thus, all the variables $\hat{y}_{k,w}$ satisfy Eq. (21) and $\sum_{w=1}^{W} \eta_w \hat{y}_{k,w} = \sum_{w=1}^{W} \eta_w y_{k,w}$. Also, we have $\max_w\{\hat{y}_{k,w}\} = \max_w\{y_{k,w}\}$. Thus, by introducing the new set of variables, the objective function in problem (19) does not change. The corresponding power consumption can be calculated as

$$\begin{aligned} \hat{\Gamma}_k &= \sum\nolimits_{w=1}^{W} \eta_w P_w \hat{y}_{k,w} \\ &= \Gamma_k + \eta_{w_2} P_{w_2} (y_{k,w_1} - y_{k,w_2}) - \sum\nolimits_{w=1}^{w_1} \eta_w P_w \Delta y_{k,w} \\ &< \Gamma_k + P_{w_2} \Big[\eta_{w_2} (y_{k,w_1} - y_{k,w_2}) - \sum\nolimits_{w=1}^{w_1} \eta_w \Delta y_{k,w} \Big], \end{aligned}$$

where the last inequality holds due to that less power is consumed for a better channel, namely, $P_1 > P_2 > \cdots > P_W$. Thus, we obtain that $\hat{\Gamma_k} < \Gamma_k$ since $\sum_{w=1}^{w_1} \eta_w \Delta y_{k,w} = \eta_{w_2}(y_{k,w_1} - y_{k,w_2})$. This means the new constructed variables lead to the same queueing delay while consuming less power. This contradicts the assumption that variables $\{y_{k,w}\}$ are the optimal solution.

Appendix D
The proof of *Lemma* 3

Let us denote $L = \sum_{k=0}^{K} k \sum_{m=0}^{M} \theta_m \pi_{k+1-m}$. With some transformation, we have

$$L = \theta_0 \sum\nolimits_{k=0}^{K} k\pi_{k+1} + \theta_1 \sum\nolimits_{k=0}^{K} k\pi_k$$

$$
\begin{aligned}
&+\cdots+\theta_M\sum\nolimits_{k=0}^{K} k\pi_{k+1-M}\\
=&\theta_0[Q-(1-\pi_0)]+\theta_1 Q+\theta_2(Q+1)\\
&+\cdots+\theta_M(Q+M-1) \qquad (43)\\
=&Q(\theta_0+\cdots\theta_M)-\theta_0(1-\pi_0)+\sum\nolimits_{m=1}^{M-1} m\theta_{m+1}\\
=&Q+\theta_0\pi_0-\theta_0+\sum\nolimits_{m=1}^{M-1} m\theta_{m+1}=Q+\theta_0\pi_0+\varsigma,
\end{aligned}
$$

where $\varsigma=\sum_{m=1}^{M-1} m\theta_{m+1}$. Thus, the average queue length Q can be expressed as $Q=L-\theta_0\pi_0-\varsigma$. From the Little's Law, we have

$$D=\frac{1}{\bar{a}}(L-\theta_0\pi_0-\varsigma). \qquad (44)$$

In Eq. (29), $y_{k,w}$ satisfies the inequalities $0 \leqslant y_{k,w} \leqslant \sum\limits_{m=0}^{M}\theta_m\pi_{k+1-m}$. Thus, we know $\max\limits_{w}\{y_{k,w}\} \leqslant \sum\limits_{m=0}^{M}\theta_m\pi_{k+1-m}$. Recalling that $L=\sum\limits_{k=0}^{K} k\sum\limits_{m=0}^{M}\theta_m\pi_{k+1-m}$, we have $L \geqslant \sum\limits_{k=0}^{K} k\cdot\max\limits_{w}\{y_{k,w}\}$. Thus,

$$D \geqslant \frac{1}{\bar{a}}(\sum_{k=0}^{K} k\cdot\max_{w}\{y_{k,w}\}-\theta_0\pi_0-\varsigma), \qquad (45)$$

the "=" holds if and only if $\max\limits_{w}\{y_{k,w}\}=\sum_{m=0}^{k+1}\theta_m\pi_{k+1-m}$.

Appendix E
The proof of Theorem 4

In *Theorem* 3, we have shown that problem (19) is equivalent to problem (22). To minimize the average queueing delay in (22), the weighted sum of $\max\limits_{w}\{y_{k,w}\}$ should be minimized while π_0 should be maximized. To minimize $\sum\limits_{k=1}^{K} k\max\limits_{w}\{y_{k,w}\}$, the term $\max\limits_{w}\{y_{k,w}\}$ should be assigned its maximum for smaller k and its minimum for larger k. Subject to constraint (19.c), we have

$$\max_{w}\{y^*_{k,w}\}=\begin{cases}\sum\limits_{m=0}^{M}\theta_m\pi^*_{k+1-m}, & k<k^*;\\ 0, & k>k^*,\end{cases} \qquad (46)$$

where k^* is a threshold imposed on the queue length. Once the queue length exceeds k^*, the maximum of variable $\{y^*_{k,w}\}$ is zero, and hence all of the variables $\{y^*_{k,w}\}$ are equal to zero. Thus, we only consider the situation when $k<k^*$.

Based on the result in *Lemma* 2, we know that for any k, there exists $y_{k,1} \leqslant y_{k,2} \leqslant \cdots \leqslant y_{k,W}$. Thus, $y_{k,W}=\max\limits_{w}\{y^*_{k,w}\}$, the threshold value T_k introduced in *Theorem* 4 meets the constraint that $T_k<W$. Suppose that there exists a set of optimal variables $\{y^*_{k,w}\}$ which violate the assignment described in Eq. (23), we can always find another set of variables $\{\hat{y}_{k,w}\}$ which consume less power while achieving to achieve the same delay.

Suppose that $\{y^*_{k,w}\}$ are the optimal variables and satisfy Eq. (23). Accordingly, the steady-state probabilities π^*_k are given by Eq. (20), and the minimum delay D^* ia achieved at the cost of power Γ_k. For a given queue state k, we reassign $\hat{\pi}_k=\pi^*_k$ to a new set of variables $\{\hat{y}_{k,w}\}$ as

$$\begin{cases}\hat{y}_{k,w}=0, \qquad 1\leqslant w<T_k-1,\\ 0\leqslant \hat{y}_{k,T_k-1}\leqslant \hat{y}_{k,T_k}\leqslant \sum\limits_{m=0}^{M}\theta_m\pi^*_{k+1-m},\\ \hat{y}_{k,w}=\sum\limits_{m=0}^{M}\theta_m\pi^*_{k+1-m}, \quad T_k<w\leqslant W.\end{cases} \qquad (47)$$

Notice that this set of variables violate the constraints in Eq. (23) since it has more than one term that lies between the maximum and the minimum of $\{\hat{y}_{k,w}\}$. Since the steady-state probabilities do not change, the minimum queueing delay remains the same. By comparing $\{\hat{y}_{k,w}\}$ with $\{y^*_{k,w}\}$, we get $y^*_{k,T_k-1}<\hat{y}_{k,T_k-1}\leqslant\hat{y}_{k,T_k}<y^*_{k,T_k}$. Since $\pi^*_k=\hat{\pi}_k$, from Eq. (30), we have $\eta_{T_k-1}(y^*_{k,T_k-1}-\hat{y}_{k,T_k-1})+\eta_{T_k}(y^*_{k,T_k}-\hat{y}_{k,T_k})=0$.The power consumption can be calculated as

$$
\begin{aligned}
\hat{\Gamma}_k&=\sum\nolimits_{w=1}^{W}\eta_w P_w\hat{y}_{k,w}\\
&=\Gamma_k+\eta_{T_k-1}P_{T_k-1}(\hat{y}_{k,T_k-1}-y^*_{k,T_k-1})\\
&\quad+\eta_{T_k}P_{T_k}(\hat{y}_{k,T_k}-y^*_{k,T_k}) \qquad (48)\\
&=\Gamma_k+P_{T_k-1}\eta_{T_k}(y^*_{k,T_k}-\hat{y}_{k,T_k})+\eta_{T_k}P_{T_k}(\hat{y}_{k,T_k}-y^*_{k,T_k})\\
&=\Gamma_k+\eta_{T_k}(y^*_{k,T_k}-\hat{y}_{k,T_k})(P_{T_k-1}-P_{T_k})>\Gamma_k.
\end{aligned}
$$

The last inequality holds since less power is consumed in a better channel, i.e., $P_{T_k-1}>P_{T_k}$ and $y^*_{k,T_k}>\hat{y}_{k,T_k}$. This means that $\{y_{k,w}\}$ violating Eq. (23) will cause a higher power consumption. The same conclusion can be obtained when more than two variables violate Eq. (23). In this way, we prove *Theorem* 4 by contradiction.

Appendix F
The proof of *Lemma* 4

Given channel state w, we define the delay minimization problem subject to the power constraint from the perspective of buffer occupation and power consumption costs discussed in Section IV as follows:

$$\min_{\{s[n]\}} \frac{1}{\bar{\alpha}}\mathbb{E}[(t[n]-s[n]+a[n+1])^+] \qquad (49)$$

$$\text{s.t.} \quad \mathbb{E}(P_w s[n])\leqslant P'_{aver}, \qquad (50)$$

where $t[n]=q[n-1]+a[n]$ is the queue state after a new datat arrival. The objective is derived based on the Little's Law and the average symbol $\mathbb{E}$ means the expectation taken over all the time slots. In the above optimization problem, the scheduling policy is described by the transmission variable $\{s[n]\}$. We mainly consider the optimal scheduling policy in the case when the equality (50) holds. If the power constraint is sufficiently large, i.e., the inequality in (50) holds, we only need to apply the scheduling policy to transmit packets as long as the queue is not empty, regardless of the channel state. Thus, using the method of Lagrangian multipliers, we only need to minimize the following Lagrangian function

$$
\begin{aligned}
\min_{\{s[n],\beta\}} \quad L(s[n],\beta')=&\frac{1}{\bar{\alpha}}\mathbb{E}\big[(t[n]-s[n]+a[n+1])^+\big]\\
&+\beta'\big[\mathbb{E}(P_w s[n])-P'_{aver}\big], \qquad (51)
\end{aligned}
$$

where β' is a Lagrangian multiplier. Equivalently, for each multiplier β', we should solve the following problem

$$\min_{\{s[n]\}} \quad \mathbb{E}[(t[n]-s[n]+a[n+1])^+]+\beta\mathbb{E}(s[n]), \qquad (52)$$

where β is defined as $\bar{\alpha}\beta' P_w$. In the sequel, we show that the optimal solution to Eq. (52) has a threshold structure. We define the expected total cost function in time slot n as

$$C_n(t[n],s[n])=\mathbb{E}[(t[n]-s[n]+a[n+1])^+]+\beta s[n], \qquad (53)$$

where the first term represents the queue length cost and the second term represents the power cost, respectively. Define

$V_n(t[n])$ as the total cost spent from slot n to slot N if we follow the optimal policy from slot n thereafter, namely,

$$V_n(t[n]) = \inf_{s[n'],n' \geqslant n} \mathbb{E} \sum_{n'=n}^{N} \gamma^{n'} C_{n'}\big(t[n'], X^{\Omega}(t[n'])\big). \quad (54)$$

The factor γ in Eq.(54) is the discount factor. Hence, the quantity V_n adds up all the costs spent across the slots n, $n+1$, $\cdots$, N and ignores the costs spent previously before slot n.

Let $z = t[n] - s[n]$. Define G_n as a function of z

$$G_n(z) = -\beta z + \mathbb{E}[(z + a[n+1])^+] + \gamma \mathbb{E}[V_{n+1}(z + a[n+1])]. \quad (55)$$

The cost V_n can be rewritten as

$$V_n(t[n]) = \beta t[n] + \min_z G_n(z). \quad (56)$$

Let $I_w = \operatorname{argmin}_z G_n(z)$. Since[5] $G_n(z)$ is a convex function of z among all the feasible z, we should choose a solution $z = I_w$ rather than any other $s[n]$. Otherwise, the transmission action $s[n]$ should always be chosen to make $z = t[n] - s[n]$ approach I_w. To this end, when $t[n]$ is greater than I_w, the transmission action $s[n] = 1$ should be made and otherwise $s[n] = 0$ is selected. Combining these two cases, we show that $s[n]$ is obtained as given by Eq. (24).

A. Function $G_n(z)$ is convex of z for all n

We use the mathematical induction to prove the convexity of $G_n(z)$. Firstly, considering the fact that term $(z + a[N+1])^+$ is convex in z for any value of $a[n+1]$,

$$G_N(z) = -\beta z + \mathbb{E}[(z + a[N+1])^+] \quad (57)$$

is convex in z. Secondly, we assume that $G_{n+1}(z)$ is convex. Then, we know

$$V_{n+1}(t[n+1]) = \beta t[n+1] + \min_z G_{(n+1)}(z) \quad (58)$$

is convex. Finally, We can derive $G_n(z)$ as

$$G_n(z) = -\beta z + \mathbb{E}[(z + a[n+1])^+] + \gamma \mathbb{E}[V_{n+1}(z + a[n+1])] \quad (59)$$

is convex in z.

References

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] P. Popovski, J. J. Nielsen, C. Stefanovic, E. de Carvalho, E. Strm, K. F. Trillingsgaard, A.-S. Bana, D. M. Kim, R. Kotaba, and J. Park, "Ultra-reliable low-latency communication (urllc): Principles and building blocks," 2017.

[3] J. Zhu, Y. Song, D. Jiang, and H. Song, "A new deep-q-learning-based transmission scheduling mechanism for the cognitive internet of things," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2375–2385, 2018.

[4] F. Qasmi, M. Shehab, H. Alves, and M. Latva-aho, "Optimum transmission rate in fading channels with markovian sources and qos constraints," *arXiv preprint arXiv:1808.06839*, 2018.

[5] C.-P. Li, J. Jiang, W. Chen, T. Ji, and J. Smee, "5g ultra-reliable and low-latency systems design," in *2017 European Conference on Networks and Communications (EuCNC)*, June 2017, pp. 1–5.

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

[7] B. Collins and R. L. Cruz, "Transmission policies for time varying channels with average delay constraints," in *Proc. Allerton Conf. Communication, Control, and Computing, Monticello, IL*, 1999, pp. 709–717.

[8] C. She, C. Yang, and T. Q. Quek, "Cross-layer optimization for ultra-reliable and low-latency radio access networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 127–141, 2018.

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

[10] R. A. Berry and E. M. Yeh, "Cross-layer wireless resource allocation," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 59–68, 2004.

[11] R. A. Berry, "Optimal power-delay tradeoffs in fading channels-small-delay asymptotics," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3939–3952, June 2013.

[12] G. Saleh, A. El-Keyi, and M. Nafie, "Cross-layer minimum-delay scheduling and maximum-throughput resource allocation for multiuser cognitive networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 4, pp. 761–773, April 2013.

[13] J. Choi, "Energy-delay tradeoff comparison of transmission schemes with limited csi feedback," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1762–1773, April 2013.

[14] H. Wang and N. B. Mandayam, "Opportunistic file transfer over a fading channel under energy and delay constraints," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 632–644, April 2005.

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

[16] M. A. Zafer and E. Modiano, "A calculus approach to energy-efficient data transmission with quality-of-service constraints," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 898–911, 2009.

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

[18] W. Chen, Z. Cao, and K. B. Letaief, "Optimal delay-power tradeoff in wireless transmission with fixed modulation," in *Proc. IEEE IWCLD*, 2007, pp. 60–64.

[19] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 2946–2957, August 2012.

[20] B. Ata, "Dynamic power control in a wireless static channel subject to a quality-of-service constraint," *Operation Research*, vol. 53, no. 5, pp. 842–851, 2005.

[21] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 125–144, Jan 2004.

[22] H. Wang and N. B. Mandayam, "A simple packet-transmission scheme for wireless data over fading channels," *IEEE Transactions on Communications*, vol. 52, no. 7, pp. 1055–1059, July 2004.

[23] M. M. Butt, E. A. Jorswieck, and N. Marchetti, "On optimizing power allocation for reliable communication over fading channels with uninformed transmitter," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2018.

[24] T. Qiu, K. Zheng, H. Song, M. Han, and B. Kantarci, "A local-optimization emergency scheduling scheme with self-recovery for a smart grid," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3195–3205, 2017.

[25] X. Zhao and W. Chen, "Delay optimal non-orthogonal multiple access with joint scheduling and superposition coding," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec 2017, pp. 1–6.

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

[27] M. Wang and W. Chen, "Achieving the optimal delay-power tradeoff in wireless transmission with arbitrarily random packet arrival: A cross-layer approach," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2014, pp. 1–6.

[28] ——, "Delay-power tradeoff of fixed-rate wireless transmission with arbitrarily bursty traffics," *IEEE Access*, vol. 5, pp. 1668–1681, 2017.

[29] X. Chen, W. Chen, J. Lee, and N. B. Shroff, "Delay-optimal buffer-aware scheduling with adaptive transmission," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2917–2930, July 2017.

[30] M. Wang, J. Liu, and W. Chen, "Delay optimal scheduling of arbitrarily bursty traffic over multi-state time-varying channels," in *2016 IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–6.

[31] D. Qiao, M. C. Gursoy, and S. Velipasalar, "The impact of QoS constraints on the energy efficiency of fixed-rate wireless transmissions," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5957–5969, 2009.

[32] http://arxiv.org/abs/1807.10128.

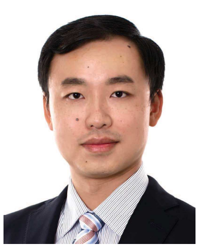
[5] See subsection F-A.

[33] L. Kleinrock, "Queueing systems, volume 1: Theory," *Lecture Notes in Computer Science*, 1975.

Meng Wang (S15-) received her BS degree from Xidian University in 2014. She is currently pursuing her Ph.D. degree in electronic engineering from Tsinghua University. In 2017 and 2018, she was a visiting Ph. D. student in the University of Maryland. Her research interests include cross-layer design, delay-power tradeoff, constraint MDP, and the AoI (age of information).

Juan Liu (M15-) received the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2011. From March 2012 to June 2014, she was a Postdoc Research Scholar in Department of Electrical and Computer Engineering, NC State University (NCSU), Raleigh, NC, USA. From February 2015 to February 2016, she was a Research Associate in Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong. Since March 2016, she has been with College of Information Science and Engineering, Ningbo University, Ningbo, China, where she is an Associate Professor.

Her research interest is in the wide range of wireless communications, and distributed signal processing and networking. Now she is focusing on the topics on wireless caching, mobile edge computing, and distributed inference for large-scale networks. She is an associate editor of IEEE Access. She has served as a TPC member for IEEE ICC 2010, ICC 2012, ICC 2014 and Globecom Workshop 2016. She has also served as a reviewer for IEEE journals including IEEE JSAC, TCOM, TSP, TVT, WCL, and IEEE conferences including IEEE INFOCOM, ICC, GLOBECOM, WCNC and VTC etc. She and her co-workers won the Best Paper Award in IEEE ICC 2016.



Wei Chen (S'05-M'07-SM'13) received his BS and Ph.D. degrees (both with the highest honors) from Tsinghua University in 2002 and 2007. From 2005 until 2007, he was also a visiting PhD student at the Hong Kong University of Science & Technology. Since 2007, he has been on the faculty at Tsinghua University, where he is a tenured full Professor, the director of degree office of Tsinghua University, and a university council member. During 2014 to 2016, he served as a deputy head of Department of Electronic Engineering. He visited Princeton University, Telecom ParisTech, and University of Southampton in 2016, 2014, and 2010 respectively. His research interests are in the areas of communication theory, stochastic optimization, and statistical learning.

Dr. Chen is a Cheung Kong young scholar and a member of national program for special support for eminent professionals, also known as 10000-talent program. He has also been supported by the national 973 youth project, the NSFC excellent young investigator project, the new century talent program of Ministry of Education, and the Beijing nova program. He received the IEEE Marconi Prize Paper Award and the IEEE Comsoc Asia Pacific Board Best Young Researcher Award in 2009 and 2011, respectively. He serves as an editor of the IEEE TRANSACTIONS COMMUNICATIONS and has served as a TPC co-chair of IEEE VTC-Spring, 2011, as well as, symposium co-chairs for IEEE ICC and Globecom. Dr. Chen is a receipt of the National May 1st Labor Medal and China Youth May 4th Medal.

Anthony Ephremides holds the Cynthia Kim Professorship of Information Technology at the Electrical and Computer Engineering Department of the University of Maryland in College Park where he is a Distinguished University Professor and has a joint appointment at the Institute for Systems Research, of which he was among the founding members in 1986. He obtained his PhD in Electrical Engineering from Princeton University in 1971 and has been with the University of Maryland ever since.

He is the author of several hundred papers, conference presentations, and patents, and his research interests lie in the areas of Communication Systems and Networks and all related disciplines, such as Information Theory, Control and Optimization, Satellite Systems, Queueing Models, Signal Processing, etc. He is especially interested in Wireless Networks, Energy Efficient Systems, and the new notion of Age of Information.