Delay-Minimal Transmission for Average Power Constrained Multi-Access Communications

Jing Yang, *Student Member, IEEE,* and Sennur Ulukus, *Member, IEEE*

***Abstract*—We investigate the problem of minimizing the overall transmission delay of packets in a multi-access wireless communication system, where the transmitters have average power constraints. We use a multi-dimensional Markov chain to model the medium access control layer behavior. The state of the Markov chain represents current queue lengths. Our goal is to minimize the average packet delay through controlling the probability of departure at each state, while satisfying the average power constraint for each queue. We consider a general asymmetric system, where the arrival rates to the queues, channel gains and average power constraints of the two users are arbitrary. We formulate the problem as a constrained optimization problem, and then transform it to a linear programming problem. We analyze the linear programming problem, and develop a procedure by which we determine the optimal solution analytically. We show that the optimal policy has a threshold structure: when the sum of the queue lengths is larger than a threshold, both users should transmit a packet during the current slot; when the sum of the queue lengths is smaller than a threshold, only one of the users, the one with the longer queue, should transmit a packet during the current slot. We provide numerical examples for both symmetric and asymmetric settings.**

***Index Terms*—Delay minimization, multi-access communication, medium-access control, queue control, power allocation, cross-layer design.**

I. Introduction

In many applications, the average delay packets experience is an important quality of service criterion. Therefore, it is important to allocate the given resources, e.g., average power, energy, etc., in a way to minimize the average delay packets experience. Since power and energy are physical layer resources, and the delay is a medium access control (MAC) layer issue, such resource allocation problems require close collaboration of physical and MAC layers, and yield cross-layer solutions. Our goal in this paper is to combine information theory and queueing theory to devise a transmission protocol which minimizes the average delay experienced by packets, subject to an average power constraint at each transmitter.

Manuscript received September 8, 2008; revised November 2, 2009 and March 21, 2010; accepted March 22, 2010. The associate editor coordinating the review of this paper and approving it for publication was D. Gesbert.

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311, and CCF 07-29127, and presented in part at the 42nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, October 2008 [1].

J. Yang is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA (e-mail: yangjing@umd.edu).

S. Ulukus is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA (e-mail: ulukus@umd.edu).

Digital Object Identifier 10.1109/TWC.2010.062910.081209

Similar goals have been undertaken by various authors in recent years. Reference [2] considers a time-slotted system with N queues and one server. The length of the slot is equal to the transmission time of a packet in the queue. In each slot, the controller allocates the server to one of the connected queues, such that the average delay in the system is minimized. The authors develop an algorithm named longest connected queue (LCQ), where the server is allocated to the longest of all connected queues at any given slot. The authors prove that in a symmetric system, LCQ algorithm minimizes the average delay. Reference [2] does not consider the issue of power consumption during transmissions.

Reference [3] combines information theory and queueing theory in a multi-access communication over an additive Gaussian noise channel. Authors consider a continuous time system, where the arrival of packets is a Poisson process, and the service time required for each packet is random. Once a packet arrives, it is transmitted immediately with a fixed power, i.e, there are no queues at the transmitters. Each transmitter-receiver pair treats the other active pairs as noise. Because of the interference from the other transmitters, the service rate for each packet is a function of the number of active users in the system. Reference [3] derives a relationship between the average delay and a fixed probability of error requirement.

References [4]–[12] consider the data transmission problem from both information theory and queueing theory perspectives. Reference [4] (see also [5], [6]) aims to minimize the average delay through rate allocation in a multi-access scenario in additive Gaussian noise. Unlike [3], in the setting of [4], packets arrive randomly into the buffers of the transmitters. When the queue at a transmitter is not empty, it transmits a packet with a fixed power. Simultaneously achievable rates are characterized by the capacity region of a multiple-access channel, which, for the non-fading Gaussian case, is a pentagon. The purpose of [4] is to find an operating point on the capacity region of the corresponding multiple-access channel such that the average delay is minimized. The author develops the longer-queue-higher-rate (LQHR) allocation strategy in the symmetric multi-access case, which is shown to minimize the average delay of the packets. The LQHR allocation scheme selects an extreme point (i.e., a corner point) in the multi-access capacity region.

Reference [7] (see also [8]) considers the problem of rate/power control in a single-user communication over a fading channel. It considers a discrete-time model, and investigates adapting rate/power in each slot according to the queue length, source state and channel state. The objective is to minimize the average power with delay constraints. It discusses two transmission models. In the first model, the transmission time of a codeword is fixed, while the rate varies from block to block. In the second model, the transmission time for each codeword varies. It formulates the problem into a dynamic programming problem and develops a delay-power tradeoff curve.

References [9]–[12] have similar formulations. Reference [9] uses dynamic programming to numerically calculate the optimal power/rate control policies that minimize the average delay in a single-user system under an average power constraint. Reference [10] derives bounds on the average delay in a system with a single queue concatenated with a multi-layer encoder. Reference [11] formulates the power constrained average delay minimization problem into a Markov decision problem and analyzes the structure of the optimal solution for a single-user fading channel. Reference [12] proposes a dynamic programming formulation to find optimal power, channel coding and source coding policies with a delay constraint. As in [7], in these papers as well, because of the large number of possible rate/power choices at each stage, it is almost impossible to get analytical optimal solutions.

Reference [13] considers a cognitive multiple access system. In the model of [13], the primary user (PU) always transmits a packet during a slot whenever its queue is not empty. The secondary user (SU) always transmits when the PU is idle, and it transmits with some probability (which is a function of its own queue length) when the PU is active. The receiver operates at the corner point of the multiple access channel capacity region where the SU is decoded first and the PU is decoded next, so that even though the SU experiences interference from the PU, the PU is always decoded interference-free. Reference [13] aims to minimize the average delay through controlling the transmission probability of the SU. It formulates the problem as a one-dimensional Markov chain and derives an analytical result to minimize the average delay of the SU under an average power constraint.

In this paper, we generalize [13] to a two-user multi-access system, where both users have equal priority. Our goal is to minimize the average delay of the packets in the system under an average power constraint for each user. As in [7], [9], [11], [13], we consider a discrete-time model. We divide the transmission time into time slots. Packets arriving at the transmitters are stored in the queues at each transmitter. In each slot, each user decides on a transmission rate based on the current lengths of both queues. Unlike [7], [9], [11], where the rate per slot is a continuous variable, we restrict the transmission rate for each user in a slot to be either zero or one packet per slot. We define the probabilities of choosing each transmission rate pair, which can be $(0,0)$, $(0,1)$, $(1,0)$ or $(1,1)$, for each given pair of queue lengths.

Our objective is to find a set of transmission probabilities that minimizes the average delay while satisfying the average power constraints for both users. As in [13], there are two main reasons that we introduce transmission probabilities: First, a randomized policy is more general than a deterministic policy; probability selections of 0 and 1 correspond to a deterministic policy, which is a special case of the randomized policy. Secondly, since we cannot choose arbitrary departure rates in each slot, the use of transmission probabilities enables us to control the average rate per slot at a finer scale. Compared to [7], [9], [11], our model has a more restricted policy space at each stage, however, this model enables us to construct a two-dimensional discrete-time Markov chain and eventually gives us a closed-form optimal solution.

In the rest of this paper, we first express the average delay and the average power consumed for each user as functions of the transmission probabilities and steady state distribution of the queue lengths. We then transform our problem into a linear programming problem, and derive the optimal transmission scheme analytically. We show that the optimal transmission policy has a threshold structure, i.e., if the sum of the queue lengths exceeds a threshold, both users transmit a packet from their queues, and if the sum of the queue lengths is smaller than a threshold, only one user, which has the larger queue length, transmits a packet from its queue, while the other user remains silent (equal queue length case is resolved by flip of a potentially biased coin). We provide a rigorous mathematical proof for the optimality of the solution. We also provide numerical examples for both symmetric and asymmetric settings.

II. System Model

A. Physical Layer Model

We consider a discrete-time additive Gaussian noise multiple-access system with two transmitters and one receiver. The received signal is

$$Y = \sqrt{h_1}X_1 + \sqrt{h_2}X_2 + Z \tag{1}$$

where X_i is the signal of user i, $\sqrt{h_i}$ is the channel gain of user i, and Z is a Gaussian noise with zero-mean and variance σ^2. Here, h_1 and h_2 are real constants, with $h_1 \neq h_2$ in general.

In this two-user system, since the multiple-access capacity region is given as [14]

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{h_1 P_1}{\sigma^2}\right) \tag{2}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{h_2 P_2}{\sigma^2}\right) \tag{3}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{h_1 P_1 + h_2 P_2}{\sigma^2}\right) \tag{4}$$

the region of feasible received powers is given by [15]

$$h_1 P_1 \geq \sigma^2(2^{2R_1} - 1) \tag{5}$$

$$h_2 P_2 \geq \sigma^2(2^{2R_2} - 1) \tag{6}$$

$$h_1 P_1 + h_2 P_2 \geq \sigma^2(2^{2(R_1+R_2)} - 1) \tag{7}$$

In each slot, the transmitters adjust their transmitted powers to achieve the desired transmission rates. We assume that for each user, the average transmitted power over all of the slots must satisfy a constraint. We denote the average power constraints for the first and second user as P_{1avg} and P_{2avg}, respectively.

B. Medium Access Control (MAC) Layer Model

In the MAC layer, we assume that packets arrive at the transmitters at a uniform size of B bits per packet. We partition

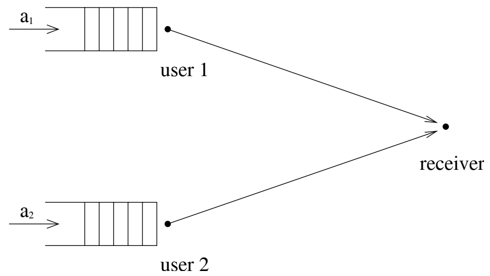

Fig. 1. System model.

the time into small slots such that we have at most one packet arrive and/or depart during each slot. Let $a_1[n]$ and $a_2[n]$ denote the number of packets arriving at the first and second transmitters, respectively, during time slot n; see Figure 1. We assume that the packet arrivals are i.i.d. from slot to slot, and the probabilities of arrivals are

$$Pr\{a_i[n] = 1\} = \theta_i \tag{8}$$
$$Pr\{a_i[n] = 0\} = 1 - \theta_i \tag{9}$$

where θ_i is the arrival rate for user i, $i = 1, 2$.

There is a buffer with capacity N at each transmitter to store the packets, where N is a finite positive integer. Once the buffer is not empty, the transmitters decide to transmit one packet in the slot with some probability based on the current lengths of both queues. Let $d_1[n]$ and $d_2[n]$ denote the number of packets transmitted in time slot n. The queue length in each buffer evolves according to

$$q_1[n+1] = (q_1[n] - d_1[n])^+ + a_1[n] \tag{10}$$
$$q_2[n+1] = (q_2[n] - d_2[n])^+ + a_2[n] \tag{11}$$

where $(x)^+$ denotes $\max(0, x)$.

The departure rate for each queue in each slot is either zero or one packet per slot, and the decision whether it should be zero or one packet per slot depends on the current queue lengths. When both queues are empty, the departure rates for both queues should be zero packet per slot. In all other situations where there is at least one packet in at least one of the queues, the departure rates for both queues should not be zero packet per slot simultaneously. This is because, keeping both transmitters idle does not save any power, but causes unnecessary delay. Therefore, in these situations, there are three possible departure rate pairs: $(d_1, d_2) = (1, 0), (0, 1)$ or $(1, 1)$, i.e., one packet is transmitted from queue 1 and no packet is transmitted from queue 2; no packet is transmitted from queue 1 and one packet is transmitted from queue 2; or, one packet is transmitted from each queue. We enumerate them as d^1, d^2, d^3. When the first queue length is i and the second queue length is j, we define the probabilities of choosing each pair of these departure rates as g_{ij}^1, g_{ij}^2, g_{ij}^3, respectively. Note that $g_{ij}^1 + g_{ij}^2 + g_{ij}^3 = 1$. We also note that g_{ij}^1, g_{ij}^2, g_{ij}^3, for $i = 0, 1, \ldots, N$ and $j = 0, 1, \ldots, N$ are the main parameters we aim to choose optimally in this paper.

The state space of the Markov chain consists of all possible pairs of queue lengths. We denote the state as $\mathbf{q} \triangleq (q_1, q_2)$. When both of the queues are empty, i.e., $\mathbf{q}[n] = (0, 0)$, transmitters have no packet to send, and from (10)-(11), $\mathbf{q}[n+1] = \mathbf{a}[n]$. The corresponding transition probabilities in this case are:

$$\begin{aligned}
Pr\{\mathbf{q}[n+1] = (0,0)|\mathbf{q}[n] = (0,0)\} &= (1-\theta_1)(1-\theta_2)\\
Pr\{\mathbf{q}[n+1] = (1,0)|\mathbf{q}[n] = (0,0)\} &= \theta_1(1-\theta_2)\\
Pr\{\mathbf{q}[n+1] = (0,1)|\mathbf{q}[n] = (0,0)\} &= \theta_2(1-\theta_1)\\
Pr\{\mathbf{q}[n+1] = (1,1)|\mathbf{q}[n] = (0,0)\} &= \theta_1\theta_2
\end{aligned} \tag{12}$$

When one of the queues is empty, there is only one possible departure rate pair, which is either $(0, 1)$ or $(1, 0)$, depending on which queue is empty. Therefore, from our argument above, the departure probabilities should not be free parameters, but must be chosen as $g_{i0}^1 = g_{0j}^2 = 1$. The corresponding transition probabilities are:

$$\begin{aligned}
Pr\{\mathbf{q}[n+1] = (i-1,0)|\mathbf{q}[n] = (i,0)\} &= (1-\theta_1)(1-\theta_2)\\
Pr\{\mathbf{q}[n+1] = (i-1,1)|\mathbf{q}[n] = (i,0)\} &= \theta_2(1-\theta_1)\\
Pr\{\mathbf{q}[n+1] = (i,0)|\mathbf{q}[n] = (i,0)\} &= \theta_1(1-\theta_2)\\
Pr\{\mathbf{q}[n+1] = (i,1)|\mathbf{q}[n] = (i,0)\} &= \theta_1\theta_2
\end{aligned} \tag{13}$$

A similar argument is valid when the first queue is empty, i.e., $\mathbf{q}[n] = (0, j)$. Transition probabilities in this case can be written similar to (13).

When neither of the queues is empty, i.e., for $\mathbf{q}[n] = (i, j)$, where $1 \leq i, j \leq N$, the transition probabilities are:

$$\begin{aligned}
Pr\{(i-1,j-1)|(i,j)\} &= g_{ij}^3(1-\theta_1)(1-\theta_2)\\
Pr\{(i-1,j+1)|(i,j)\} &= g_{ij}^1\theta_2(1-\theta_1)\\
Pr\{(i+1,j-1)|(i,j)\} &= g_{ij}^2\theta_1(1-\theta_2)\\
Pr\{(i,j+1)|(i,j)\} &= g_{ij}^1\theta_1\theta_2\\
Pr\{(i+1,j)|(i,j)\} &= g_{ij}^2\theta_1\theta_2\\
Pr\{(i-1,j)|(i,j)\} &= g_{ij}^3\theta_2(1-\theta_1)+g_{ij}^1(1-\theta_1)(1-\theta_2)\\
Pr\{(i,j-1)|(i,j)\} &= g_{ij}^3\theta_1(1-\theta_2)+g_{ij}^2(1-\theta_1)(1-\theta_2)\\
Pr\{(i,j)|(i,j)\} &= g_{ij}^1\theta_1(1-\theta_2)+g_{ij}^2\theta_2(1-\theta_1)+g_{ij}^3\theta_1\theta_2
\end{aligned} \tag{14}$$

For example, the first equation in (14) is obtained by noting that, for the next queue state to be $(i-1, j-1)$, we need to transmit one packet from each queue and we should have no arrivals to either of the queues. The probability of this event is g_{ij}^3, probability of transmitting one packet from each queue, multiplied by $(1-\theta_1)$, probability of having no arrivals to queue 1, and $(1-\theta_2)$, probability of having no arrivals to queue 2.

In this paper, we assume that the average power constraints are large enough to prevent any packet losses. In order to prevent overflows, we always choose to transmit one packet from a queue whenever its length reaches N. Therefore, we have $g_{iN}^1 = g_{Nj}^2 = g_{NN}^3 = 1$. The two-dimensional Markov chain is shown in Figure 2.

In [16], it is proven that, for all irreducible, positive recurrent discrete-time Markov chains with state space S, there exists a stationary distribution $\{\pi_s, s \in S\}$, which is given by the unique solution to

$$\sum_{s\in S}\pi_s p_{sr} = \pi_r, \qquad \sum_{s\in S}\pi_s = 1 \tag{15}$$

It is also stated that for a reducible Markov chain with a single closed positive recurrent aperiodic class and a nonempty set T, where for any $i \in T$, the probability of getting absorbed in the closed class starting from state i is 1, and the steady

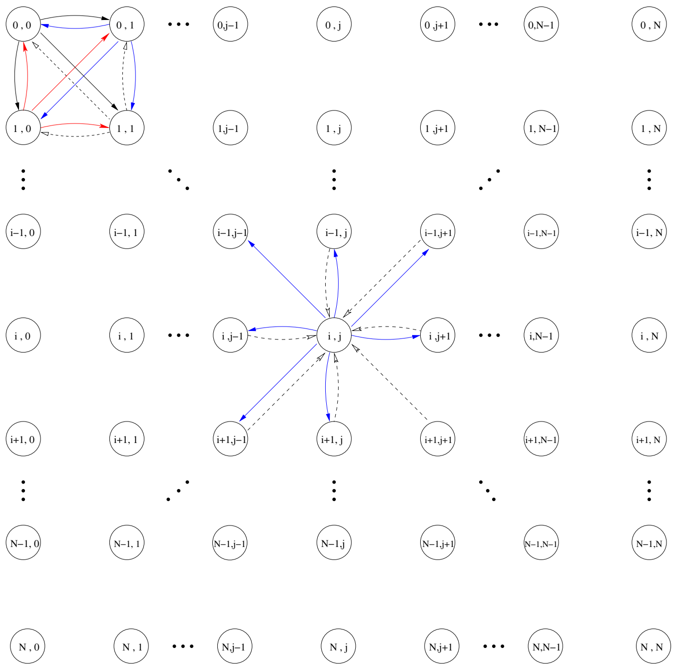

Fig. 2. Two-dimensional Markov chain.

state distribution exists. In our problem, we first assume that the stationary distribution exists for the optimal solution. Once we determine the solution, we verify that the corresponding Markov chain has a unique stationary distribution.

Let us define the steady state distribution of this Markov chain as $\boldsymbol{\pi} = [\pi_{00}, \pi_{01}, \cdots, \pi_{0N}, \pi_{10}, \cdots, \pi_{NN}]$. Then, the steady state distribution must satisfy

$$\boldsymbol{\pi}\mathbb{P} = \boldsymbol{\pi}, \qquad \boldsymbol{\pi}\mathbf{1} = 1 \tag{16}$$

where $\mathbb{P}$ is the transition matrix defined by the transition probabilities (12)-(14). We can express the average number of packets in the system as $\sum_{i,j} \pi_{ij}(i+j)$. According to Little's law [16], for a fixed sample path in a queueing system, if the limits of average waiting time W and average arrival rate λ exist as time goes to infinity, then the limit of average queue length L exists and they are related as $L = \lambda W$. For our problem, in a system without overflow, these limits exist and the average delay D is equal to

$$D = \frac{1}{\theta_1 + \theta_2} \sum_{i,j} \pi_{ij}(i+j) \tag{17}$$

where $\theta_1 + \theta_2$ is the average arrival rate for the system.

III. Problem Formulation

The transmission rate for both transmitters during a slot is either one packet per slot or zero packet per slot. Equivalently, the transmission rate is either B/τ bits/channel use or 0 bits/channel use, where τ is the number of channel uses in each slot. We assume that in each slot we can use codewords with finite block length to get arbitrarily close to the boundary of feasible powers and achieve a satisfactory level of reliability.

Next, let us consider the power consumptions during each slot. When only one user transmits, since there is no interference from the other transmitter, the transmitted power for the active user needs to satisfy

$$h_i P_i \geq \sigma^2(2^{2R} - 1) \triangleq \alpha \tag{18}$$

where $R = B/\tau$. In order to minimize the power, the transmitted power for the active user should be α/h_i, depending on which user is transmitting. When both users transmit simultaneously, the received powers should additionally satisfy

$$h_1 P_1 + h_2 P_2 \geq \sigma^2(2^{4R} - 1) \triangleq \beta \tag{19}$$

The feasible transmitted power region is shown in Figure 3. Let us denote the received power pair as (β_1, β_2). In order to minimize the transmit power, this pair should be on the dominant face of the feasible power region, i.e., $\beta_1 + \beta_2 = \beta$. Then, the corresponding transmit power pair is $(\beta_1/h_1, \beta_2/h_2)$. Note that different operating points need different sum of transmit powers.

Thus, for any state $(i,j) \neq (0,0)$, the average power consumed for the first queue is $\frac{1}{h_1}(g_{ij}^1\alpha + g_{ij}^3\beta_1)$, while the average power consumed for the second queue is $\frac{1}{h_2}(g_{ij}^2\alpha + g_{ij}^3\beta_2)$. Our goal is to find the transmission policy, i.e., the probabilities g_{ij}^k, $k = 1,2,3$, $i = 0,1,\ldots,N$, $j = 0,1,\ldots,N$ along with the operating point (β_1, β_2), such that the average delay is minimized, subject to an average power constraint for each user. Therefore, our problem can be expressed as:

$$\min_{\mathbf{g},\beta_1,\beta_2} \quad \frac{1}{\theta_1 + \theta_2} \sum_{i,j} \pi_{ij}(i+j) \tag{20}$$

$$\text{s.t.} \quad \frac{1}{h_1} \sum_{i,j} \pi_{ij}(g_{ij}^1\alpha + g_{ij}^3\beta_1) \leq P_{1avg} \tag{21}$$

$$\frac{1}{h_2} \sum_{i,j} \pi_{ij}(g_{ij}^2\alpha + g_{ij}^3\beta_2) \leq P_{2avg} \tag{22}$$

$$\boldsymbol{\pi}\mathbb{P} = \boldsymbol{\pi}, \quad \boldsymbol{\pi}\mathbf{1} = 1 \tag{23}$$

$$g_{ij}^1 + g_{ij}^2 + g_{ij}^3 = 1, \quad i,j = 0,1,\ldots,N \tag{24}$$

$$g_{ij}^k \geq 0, \quad i,j = 0,1,\ldots,N, \quad k = 1,2,3 \tag{25}$$

We note that the state transition matrix $\mathbb{P}$ is filled with variables in (12)-(14) which depend on g_{ij}^ks. Also, through (23), π_{ij}s depend on g_{ij}^ks, as well. Unlike [13], we have a two-dimensional Markov chain, and it does not admit closed-form expressions for the steady state distribution π_{ij}s in terms of g_{ij}^ks. Therefore, solving the above optimization problem becomes rather difficult. Our methodology will be to transform our optimization problem into a linear programming problem, and exploit its special structure to obtain the globally optimal solution analytically.

IV. Analysis of the Problem

Note that $g_{ij}^1 + g_{ij}^2 + g_{ij}^3 = 1$ for any $(i,j) \neq (0,0)$, therefore $\pi_{ij} = \pi_{ij}(g_{ij}^1 + g_{ij}^2 + g_{ij}^3)$. Define $x_{00} = \pi_{00}$, $x_{ij}^k = \pi_{ij} g_{ij}^k$, $k = 1,2,3$, $i = 0,1,\ldots,N$, $j = 0,1,\ldots,N$. Since g_{ij}^k is the conditional probability of choosing policy k when the system is in state (i,j), x_{ij}^k can be interpreted as the unconditional probability of staying in state (i,j) and choosing policy k. Our aim is to find optimal g_{ij}^ks. However, as we will see, our analysis will be more tractable with variables x_{ij}^ks. Once we

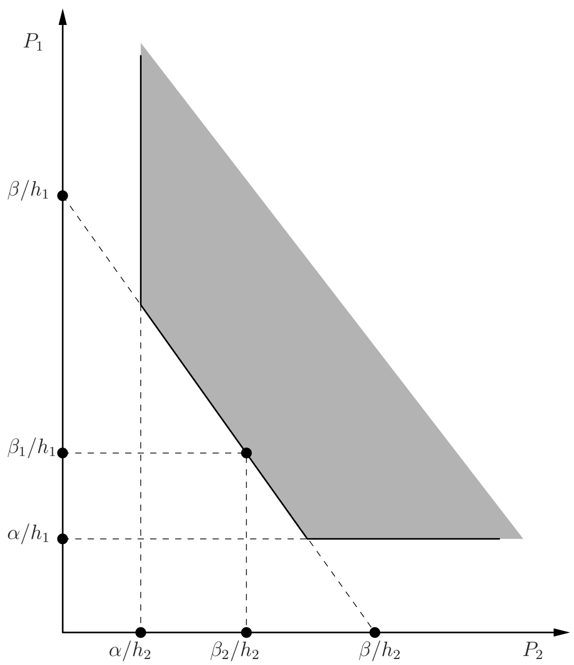

Fig. 3. Feasible power region.

find optimal x_{ij}^ks, we can obtain optimal g_{ij}^ks through

$$g_{ij}^k = \frac{x_{ij}^k}{\sum_{k=1}^3 x_{ij}^k} \tag{26}$$

Let us construct a vector of all of our unknowns $\mathbf{x} = [x_{00}, x_{01}^1, x_{01}^2, x_{01}^3, \ldots, x_{NN}^3]^T$.

First, we consider the average power consumption when average power constraints for both users are large enough such that each user is able to transmit a packet during a slot whenever its queue is not empty. In this scenario, the corresponding Markov chain has four non-transient states, (0,0), (0,1), (1,0), (1,1), and the stationary distribution is

$$\begin{aligned}\pi_{01} &= \theta_2(1-\theta_1), \quad \pi_{00} = (1-\theta_1)(1-\theta_2),\\ \pi_{10} &= \theta_1(1-\theta_2), \quad \pi_{11} = \theta_1\theta_2\end{aligned} \tag{27}$$

The average power consumption for each queue is

$$\begin{aligned}P_{1csmp} &= \frac{1}{h_1}(\pi_{10}\alpha + \pi_{11}\beta_1)\\ &= \frac{1}{h_1}\left(\theta_1(1-\theta_2)\alpha + \theta_1\theta_2\beta_1\right)\end{aligned} \tag{28}$$

$$\begin{aligned}P_{2csmp} &= \frac{1}{h_2}(\pi_{01}\alpha + \pi_{11}\beta_2)\\ &= \frac{1}{h_2}\left(\theta_2(1-\theta_1)\alpha + \theta_1\theta_2\beta_2\right)\end{aligned} \tag{29}$$

We note that

$$P_{1csmp}h_1 + P_{2csmp}h_2 = (\theta_1 + \theta_2 - 2\theta_1\theta_2)\alpha + \theta_1\theta_2\beta \tag{30}$$

From Figure 3, we note that $\beta_1, \beta_2 \geq \alpha$, therefore, each individual term in (30) must additionally satisfy

$$P_{1csmp} \geq \frac{1}{h_1}\theta_1\alpha \tag{31}$$

$$P_{2csmp} \geq \frac{1}{h_2}\theta_2\alpha \tag{32}$$

Therefore, if the average power constraints P_{1avg} and P_{2avg} satisfy the following inequalities

$$P_{1avg}h_1 + P_{2avg}h_2 \geq (\theta_1 + \theta_2 - 2\theta_1\theta_2)\alpha + \theta_1\theta_2\beta \tag{33}$$

$$P_{1avg} \geq \frac{1}{h_1}\theta_1\alpha \tag{34}$$

$$P_{2avg} \geq \frac{1}{h_2}\theta_2\alpha \tag{35}$$

then we can always find an operating point (β_1, β_2) such that $P_{1csmp} \leq P_{1avg}$ and $P_{2csmp} \leq P_{2avg}$, and we achieve the minimal possible delay in the system, which is one slot. The available power in this case is so large that the solution is trivial. If

$$P_{1avg}h_1 + P_{2avg}h_2 < (\theta_1 + \theta_2 - 2\theta_1\theta_2)\alpha + \theta_1\theta_2\beta \tag{36}$$

and P_{1avg} and P_{2avg} are large enough to prevent any overflows, both power constraints should be tight. Therefore, from (21)-(22), we have two equality power constraints,

$$\frac{1}{h_1}\sum_{i,j}(x_{ij}^1\alpha + x_{ij}^3\beta_1) = P_{1avg} \tag{37}$$

$$\frac{1}{h_2}\sum_{i,j}(x_{ij}^2\alpha + x_{ij}^3\beta_2) = P_{2avg} \tag{38}$$

Because the average arrival rate must be equal to the average departure rate when there is no overflow, we also have

$$\sum_{i,j}(x_{ij}^1 + x_{ij}^3) = \theta_1 \tag{39}$$

$$\sum_{i,j}(x_{ij}^2 + x_{ij}^3) = \theta_2 \tag{40}$$

Solving (37)-(40), we obtain

$$\beta_1 = \alpha + \frac{(\beta - 2\alpha)(P_{1avg}h_1 - \theta_1\alpha)}{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha} \tag{41}$$

$$\beta_2 = \alpha + \frac{(\beta - 2\alpha)(P_{2avg}h_2 - \theta_2\alpha)}{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha} \tag{42}$$

$$\sum_{i,j} x_{ij}^1 = \theta_1 - \frac{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha}{\beta - 2\alpha} \tag{43}$$

$$\sum_{i,j} x_{ij}^2 = \theta_2 - \frac{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha}{\beta - 2\alpha} \tag{44}$$

$$\sum_{i,j} x_{ij}^3 = \frac{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha}{\beta - 2\alpha} \tag{45}$$

By jointly considering the normalization equation in (23), we also have

$$x_{00} = 1 - \frac{(\theta_1+\theta_2)(\beta-\alpha) - (P_{1avg}h_1 + P_{2avg}h_2)}{\beta - 2\alpha} \tag{46}$$

Thus, we transform our optimization problem in (20)-(24) into

$$\min_{\mathbf{x}} \quad \sum_{i,j}\left(\sum_{k=1}^3 x_{ij}^k(i+j)\right) \tag{47}$$

$$\text{s.t.} \quad x_{00} = 1 - \frac{(\theta_1+\theta_2)(\beta-\alpha) - (P_{1avg}h_1 + P_{2avg}h_2)}{\beta - 2\alpha} \tag{48}$$

$$\sum_{i,j} x_{ij}^1 = \theta_1 - \frac{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha}{\beta - 2\alpha} \tag{49}$$

$$\sum_{i,j} x_{ij}^2 = \theta_2 - \frac{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha}{\beta - 2\alpha} \tag{50}$$

$$\sum_{i,j} x_{ij}^3 = \frac{P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha}{\beta - 2\alpha} \tag{51}$$

$$\mathbb{Q}\mathbf{x} = \mathbf{0},\ x_{ij}^k \geq 0,\ i,j = 0,1,\ldots,N,\ k=1,2,3 \tag{52}$$

which is in terms of x_{ij}^ks. Here, $\mathbb{Q}$ is a $(N+1)^2 \times (4(N+1)^2-3)$ matrix defined by matrix $\mathbb{P}$. We get the equations in (52) from (23) by substituting $\pi_{ij}g_{ij}^k$ for x_{ij}^k.

The optimization problem in (47)-(52) is a linear programming problem. In addition, we observe that, in the objective function, all of the x_{ij}^ks with the same sum of indices share the same weight $i+j$. If we look into the two-dimensional Markov chain, this corresponds to the states on the diagonals running from the upper right corner to the lower left corner. This motivates us to group the x_{ij}^ks along the diagonals of the two-dimensional Markov chain in Figure 2 and define their sum, for the nth diagonal, as

$$y_n = \sum_{i=0}^{n} (x_{i,n-i}^1 + x_{i,n-i}^2) \tag{53}$$

$$t_n = \sum_{i=0}^{n} x_{i,n-i}^3 \tag{54}$$

Then, $y_n \geq 0$, $t_n \geq 0$, and the objective function in (47) is equivalent to

$$\sum_{n=1}^{2N} (y_n + t_n) n \tag{55}$$

We also get $2N$ flow-in-flow-out equations between the diagonal groups. For $n = 0, 1$, we have

$$x_{00}(\theta_1+\theta_2-\theta_1\theta_2) = (y_1+t_2)(1-\theta_1)(1-\theta_2) \tag{56}$$

$$(x_{00}+y_1)\theta_1\theta_2 = (y_2+t_3)(1-\theta_1)(1-\theta_2) + t_2(1-\theta_1\theta_2) \tag{57}$$

and for $n = 2, 3, \ldots, 2N-2$, we have

$$y_n\theta_1\theta_2 = (y_{n+1}+t_{n+2})(1-\theta_1)(1-\theta_2) + t_{n+1}(1-\theta_1\theta_2) \tag{58}$$

$$y_{2N-1}\theta_1\theta_2 = t_{2N}(1-\theta_1\theta_2) \tag{59}$$

Figure 4 shows the transitions between diagonal groups for a system with $N = 3$; we use different colors to distinguish the transitions caused by different departure rate pairs.

We multiply both sides of the nth equation in (56)-(59) with z^n and sum with respect to n to obtain

$$\begin{aligned} & x_{00}(\theta_1+\theta_2-\theta_1\theta_2+\theta_1\theta_2 z) \\ & + \big(\theta_1\theta_2 - (1-\theta_1)(1-\theta_2)z^{-1}\big) \sum_{n=1}^{2N} y_n z^n \\ & - \big((1-\theta_1\theta_2)z^{-1} + (1-\theta_1)(1-\theta_2)z^{-2}\big) \sum_{n=1}^{2N} t_n z^n = 0 \end{aligned} \tag{60}$$

Taking the derivative of (60) with respect to z and letting

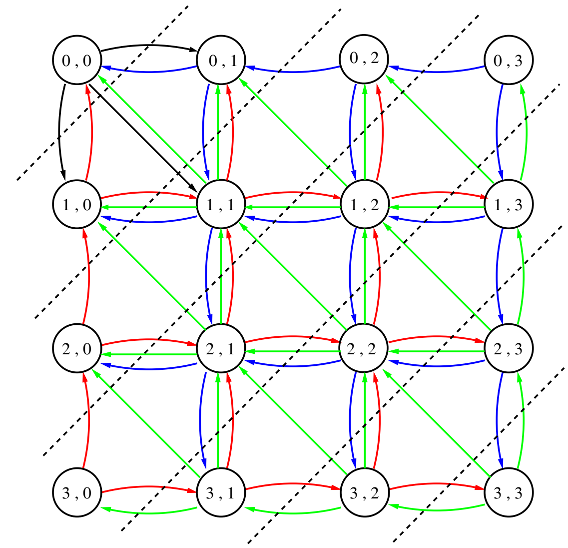

Fig. 4. The transitions between diagonal groups when $N = 3$.

$z = 1$, we have

$$\begin{aligned} \sum_{n=1}^{2N} t_n n = & \frac{1}{2-\theta_1-\theta_2}\Bigg((\theta_1+\theta_2-1)\left(\sum_{n=1}^{2N} y_n n\right) \\ & + (1-\theta_1)(1-\theta_2)\left(\sum_{n=1}^{2N} y_n\right) \\ & + \big(1-\theta_1\theta_2+2(1-\theta_1)(1-\theta_2)\big)\left(\sum_{n=1}^{2N} t_n\right) + x_{00}\theta_1\theta_2 \Bigg) \end{aligned} \tag{61}$$

From the definition of y_n and t_n in (53) and (54), we note

$$\sum_{n=1}^{2N} y_n = \sum_{n=1}^{2N}\sum_{i=0}^{n} (x_{i,n-i}^1 + x_{i,n-i}^2) = \sum_{i,j}(x_{ij}^1 + x_{ij}^2) \tag{62}$$

$$\sum_{n=1}^{2N} t_n = \sum_{n=1}^{2N}\sum_{i=0}^{n} x_{i,n-i}^3 = \sum_{i,j} x_{ij}^3 \tag{63}$$

From (62) and (63), and using (49)-(51), we conclude that $\sum_{n=1}^{2N} y_n$ and $\sum_{n=1}^{2N} t_n$ are constants that depend on system parameters α, β, θ_1, θ_2 and P_{1avg}, P_{2avg}. Using (62) and (49)-(50), for future reference, let us define

$$\begin{aligned} \sum_{n=1}^{2N} y_n &= \sum_{i,j} x_{ij}^1 + \sum_{i,j} x_{ij}^2 \\ &= \theta_1+\theta_2 - \frac{2(P_{1avg}h_1 + P_{2avg}h_2 - (\theta_1+\theta_2)\alpha)}{\beta-2\alpha} \\ &\triangleq \Psi \end{aligned} \tag{64}$$

Using the definition of y_n, t_n and (61), the objective function in (47) becomes

$$\sum_{n=1}^{2N} (y_n+t_n) n = \frac{1}{2-\theta_1-\theta_2}\left(\sum_{n=1}^{2N} y_n n\right) + C \tag{65}$$

where C is a constant, and $\frac{1}{2-\theta_1-\theta_2}$ is positive. Therefore, minimizing the original objective function in (47) is equivalent to minimizing $\sum_{n=1}^{2N} y_n n$. Since from (64) the sum of y_ns is fixed, and y_ns are positive, intuitively, the optimization problem requires us to assign larger values to y_ns with smaller indices n, without conflicting with the transition equation constraints.

V. The Modified Optimization Problem and a Two-Step Solution

In this section, we will prove the following main result of our paper: If the average power constraints P_{1avg}, P_{2avg} are large enough to prevent any packet losses, the delay-optimal policy has a threshold structure. When the sum of the queue lengths is larger than the threshold, both users should transmit; when the sum of the queue lengths is smaller than the threshold, only the user with the longer queue should transmit; the equal queue length case can be resolved through flip of a potentially biased coin.

We propose to solve our original optimization problem in two steps. In the first step, we will consider the optimization problem in terms of y_ns and t_ns, where the objective function is $\sum_{n=1}^{2N} y_n n$, and the constraints are (64), (48), (56)-(59), and positivity constraints on y_ns and t_ns. The objective function of this optimization problem is exactly the same as that of our original optimization problem in (47)-(52), however, our constraints are more lenient than those of (47)-(52). First, (64) is weaker than (49)-(51), as it imposes a constraint on the sum while (49)-(51) impose constraints on individual terms. Secondly, the transition equations in (52) are between all of the states in the two-dimensional Markov chain, while the transition equations in (56)-(59) are only between the diagonal groups in the Markov chain. Finally, we do not explicitly impose the sum constraint on t_n on the new problem. These imply that, the result we obtain in the first step, in principle, may not be feasible for the original problem.

Therefore, in the second step we allocate y_ns and t_ns we obtain from the first step to x_{ij}^ks in such a way that the remaining independent transition equations in (52) are satisfied. We note that (39) and (40) can be derived from (52), therefore, once (52) is satisfied, (39) and (40) will be satisfied. Together with (64), we can make sure that (49)-(51) are all satisfied. Therefore, if we can find a valid allocation in the second step, we will conclude that the solution found in the first step is a feasible solution to our original problem. Since the problem we solve in the first step has the cost function of our problem, but is subject to more lenient constraints, its solution, in principle, may be better than the solution of our original problem. However, when we prove that the solution we obtain in our first step is within the feasible set of our original problem, we will have solved our original problem. In addition, once we prove the optimality of the solution in the first step, it will be globally optimal for the original problem.

First, we will minimize $\sum_{n=1}^{2N} y_n n$ subject to (64), (48), (56)-(59), and y_n, $t_n \geq 0$. This means that we will allocate Ψ to y_ns in a way to minimize $\sum_{n=1}^{2N} y_n n$. This will require us to allocate larger values to y_ns with smaller n, while making sure that (64), (48), and (56)-(59) are satisfied. We state the result of our first step in the following theorem.

Theorem 1 *The optimal solution of the problem*

$$\begin{aligned} \min \quad & \sum_{n=1}^{2N} y_n n \\ \textit{s. t.} \quad & (64), (48), (56)-(59), \textit{and } y_n \geq 0, t_n \geq 0, \forall n \end{aligned} \tag{66}$$

has a threshold structure. In particular, there exists a threshold $\bar{n}$ such that for $n < \bar{n}$, $t_n = 0$ and for $n > \bar{n}$, $y_n = 0$.

The proof of this theorem is given in Appendix A.

In the following, we consider the transition equations within groups for each state. Since adding more constraints cannot improve the optimization result, if we can find a way to allocate y_ns and t_ns to x_{ij}^ks, such that all of the remaining transition equations are satisfied, then we will conclude that the assignments we obtained in the first step are actually feasible for the original problem. Therefore, next, in our second step, we focus on the assignment of the y_ns and t_ns found in the first step to x_{ij}^ks.

First, we use a simple example to illustrate the procedure of allocation within each group, then, we generalize the procedure to arbitrary cases. In this simple example, we assume that $N = 4$.

Assume that after the group allocation, we obtained $y_1, \ldots, y_5$ and $t_5, t_6 \neq 0$, and the rest of the y_ns and t_ns are equal to zero. In order to keep the allocation simple, when we assign y_3, y_5, t_5 in each group, we assign them only to two states: $(1, 2), (2, 1)$ and $(2, 3), (3, 2)$, respectively; while we assign y_4 to three states: $(1, 3), (2, 2), (3, 1)$, and we assign t_6 to a single state $(3, 3)$. Figure 5 illustrates the allocation pattern within groups. We do not assign any values to the states with dotted circles. The dotted states will be transient states after the allocation. We need to guarantee that the nonzero-valued states only transit to other nonzero-valued states. This requires us to set $x_{12}^1 = x_{21}^2 = x_{23}^1 = x_{32}^2 = 0$, and $x_{13}^1 = x_{13}^3 = x_{31}^2 = x_{31}^3 = 0$. The valid transitions are represented as arrows in Figure 5. We can see that the transitions are within the positive recurrent class.

Then, let us examine each group and find transition equations to be satisfied for each state. For states $(0, 1), (0, 2), (1, 2), (1, 3), (2, 3)$, the transition equations to be satisfied are

$$\begin{aligned} x_{01}^2(1-\theta_2(1-\theta_1)) =& (x_{00} + x_{10}^1)\theta_2(1-\theta_1) \\ & + (x_{02}^2 + x_{11}^1)(1-\theta_1)(1-\theta_2) \\ x_{02}^2(1-\theta_2(1-\theta_1)) =& x_{11}^1\theta_2(1-\theta_1) \\ x_{12}^2(1-\theta_2(1-\theta_1)) =& (x_{02}^2 + x_{11}^1)\theta_1\theta_2 + x_{21}^1\theta_2(1-\theta_1) \\ & + (x_{13}^2 + x_{22}^1 + x_{23}^3)(1-\theta_1)(1-\theta_2) \\ x_{13}^2(1-\theta_2(1-\theta_1)) =& (x_{11}^1 + x_{23}^3)\theta_2(1-\theta_1) \\ x_{23}^2(1-\theta_2(1-\theta_1)) +& x_{23}^3(1-\theta_1\theta_2) = (x_{13}^2 + x_{22}^1)\theta_1\theta_2 \\ & + (x_{32}^1 + x_{33}^3)\theta_2(1-\theta_1) \end{aligned} \tag{67}$$

We have five more similar transition equations for states $(0, 1), (0, 2), (1, 2), (1, 3), (2, 3)$. All the unknown variables are interacting with each other through these equations. How to find an allocation satisfying all of these equations simultaneously becomes rather difficult. After simple manipulations, equations in (67) become equivalent to

$$\begin{aligned}
x_{01}^2 =& (x_{00}+x_{10}^1+x_{01}^2)\theta_2(1-\theta_1)+(x_{02}^2+x_{11}^1)(1-\theta_1)(1-\theta_2)\\
x_{02}^2 =& (x_{11}^1+x_{02}^2)\theta_2(1-\theta_1)\\
x_{12}^2 =& (x_{02}^2+x_{11}^1)\theta_1\theta_2+(x_{12}^2+x_{21}^1)\theta_2(1-\theta_1)\\
&+(x_{13}^2+x_{22}^1+x_{23}^3)(1-\theta_1)(1-\theta_2)\\
x_{13}^2 =& (x_{22}^1+x_{13}^2+x_{23}^3)\theta_2(1-\theta_1)\\
x_{23}^2 =& (x_{13}^2+x_{22}^1)\theta_1\theta_2+(x_{32}^1+x_{23}^2+x_{33}^3)\theta_2(1-\theta_1)\\
&-x_{23}^3(1-\theta_1\theta_2)
\end{aligned} \tag{68}$$

Observing the right hand sides of (68), we note that, x_{00}, $x_{10}^1+x_{10}^2$, $x_{12}^2+x_{21}^1$, $x_{32}^1+x_{23}^2$, x_{33}^3 are known, therefore, the allocation for states $(0,1),(0,2),(1,2),(1,3),(2,3)$ depends only on the values of $x_{02}^2+x_{11}^1$, $x_{22}^1+x_{13}^2$, and x_{23}^3. Similarly, the allocation for states $(1,0),(2,0),(2,1),(3,1),(3,2)$ also depends on the values of $x_{20}^1+x_{11}^2$, $x_{22}^2+x_{31}^1$, and x_{32}^3 only. Since

$$y_2=(x_{02}^2+x_{11}^1)+(x_{20}^1+x_{11}^2) \tag{69}$$
$$y_4=(x_{22}^1+x_{13}^2)+(x_{22}^2+x_{31}^1) \tag{70}$$
$$t_5=x_{23}^3+x_{32}^3 \tag{71}$$

the allocation actually depends on how we split y_2, y_4 and t_5 between $(x_{02}^2+x_{11}^1)$ and $(x_{20}^1+x_{11}^2)$, $(x_{22}^1+x_{13}^2)$ and $(x_{22}^2+x_{31}^1)$, x_{23}^3 and x_{32}^3, respectively. Once we fix the values of $x_{02}^2+x_{11}^1$, $x_{22}^1+x_{13}^2$, and x_{23}^3, we obtain the values of all of the states, completing the allocation. We note that there is more than one feasible allocation within groups, and for each feasible allocation, all of the transition equations are satisfied, and the power constraints are satisfied as well. In order to keep the solution simple, we let

$$x_{02}^2+x_{11}^1=y_2/2 \tag{72}$$
$$x_{22}^1+x_{13}^2=y_4/2 \tag{73}$$
$$x_{23}^3=t_5/2 \tag{74}$$

Plugging these into (68), we get

$$\begin{aligned}
x_{01}^2 =& (x_{00}+y_1)\theta_2(1-\theta_1)+\frac{1}{2}y_2(1-\theta_1)(1-\theta_2)\\
x_{02}^2 =& \frac{1}{2}y_2\theta_2(1-\theta_1)\\
x_{12}^2 =& \frac{1}{2}y_2\theta_1\theta_2+y_3\theta_2(1-\theta_1)+\frac{1}{2}(y_4+t_5)(1-\theta_1)(1-\theta_2)\\
x_{13}^2 =& \frac{1}{2}(y_4+t_5)\theta_2(1-\theta_1)\\
x_{23}^2 =& \frac{1}{2}y_4\theta_1\theta_2+(y_5+t_6)\theta_2(1-\theta_1)-\frac{1}{2}t_5(1-\theta_1\theta_2)
\end{aligned} \tag{75}$$

Going back to (72)-(73), we obtain

$$\begin{aligned}
x_{11}^1 =& \frac{1}{2}y_2(1-\theta_2(1-\theta_1))\\
x_{22}^1 =& \frac{1}{2}y_4-\frac{1}{2}(y_4+t_5)\theta_2(1-\theta_1)
\end{aligned} \tag{76}$$

Since $y_n \geq t_{n+1}\rho/\delta$, we can easily verify that $x_{23}^2 \geq 0$, $x_{22}^1 \geq 0$. The allocation for the remaining half of the states has a similar structure. Thus, each state has a positive value, and the allocation is feasible.

Once we obtain the values of x_{ij}^ks, we can compute the transmission probabilities using $g_{ij}^k = \frac{x_{ij}^k}{\sum_{k=1}^3 x_{ij}^k}$. Here, we have

$$g_{11}^1=\frac{1-\theta_2(1-\theta_1)}{2-\theta_2(1-\theta_1)-\theta_1(1-\theta_2)} \tag{77}$$
$$g_{11}^2=\frac{1-\theta_1(1-\theta_2)}{2-\theta_2(1-\theta_1)-\theta_1(1-\theta_2)} \tag{78}$$
$$g_{22}^1=\frac{y_4-(y_4+t_5)\theta_2(1-\theta_1)}{2y_4-(y_4+t_5)(\theta_2(1-\theta_1)+\theta_1(1-\theta_2))} \tag{79}$$
$$g_{22}^2=\frac{y_4-(y_4+t_5)\theta_1(1-\theta_2)}{2y_4-(y_4+t_5)(\theta_2(1-\theta_1)+\theta_1(1-\theta_2))} \tag{80}$$

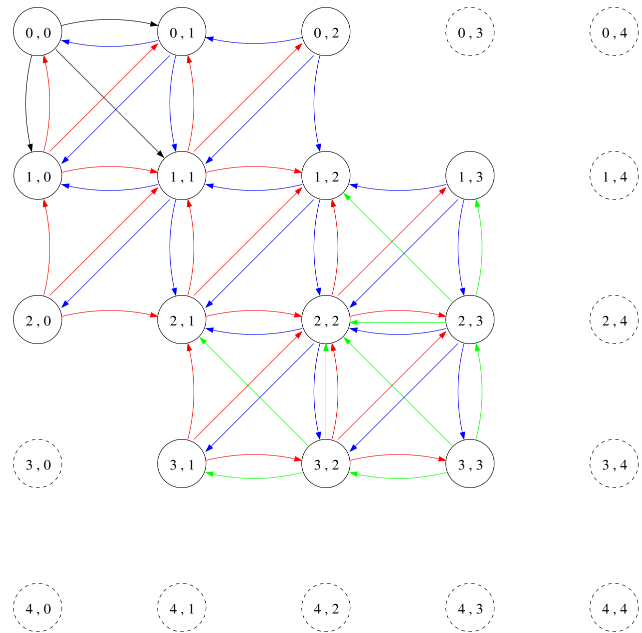

Fig. 5. Example: allocation within groups when $N = 4$.

We observe that a threshold structure exists. In this example, the threshold is 5. When the sum of the two queue lengths is greater than 5, both users transmit during a slot. When the sum of the two queue lengths is less than 5, only one user with longer queue transmits during a slot; in this case, if both queue lengths are the same, users transmit according to probabilities in (77)-(80).

Following steps similar to those in the example above, we can assign y_ns and t_ns to x_{ij}^ks and obtain a feasible allocation for general settings. The following theorem states this fact formally.

Theorem 2 *For the y_ns and t_ns obtained in the first step, there always exists a feasible x_{ij}^k assignment, such that x_{ij}^ks are positive and satisfy all of the transition equations.*

The proof of this theorem is given in Appendix B. Since this is a constructive proof, it also gives the exact method by which y_ns and t_ns are assigned to x_{ij}^ks.

Therefore, in order to prove the optimality of the x_{ij}^k assignment, it suffices to prove the optimality of the solution of the first step. The following theorem proves the optimality of the first step.

Theorem 3 *The allocation scheme in Theorem 1 minimizes the average delay in the system.*

The proof of this theorem is given in Appendix C.

In summary, the two-step allocation scheme is feasible and optimal for our original problem. The transition probabilities can be computed once we determine the allocation for each sate. From our allocation, we note that there exists a threshold $\bar{n}$, where $\bar{n}$ is the largest group index n such that $y_n \neq 0$. We have $t_n > 0$ only when $n \geq \bar{n}$. Since $g_{ij}^k = \frac{x_{ij}^k}{\sum_{k=1}^3 x_{ij}^k}$, we have $g_{ij}^3 = 1$ when $n > \bar{n}$. When $n < \bar{n}$, we have $g_{ij}^1 = 1$ if $i > j$ and $g_{ij}^2 = 1$ if $i < j$. Then, for $n \leq \bar{n}$, and n is even, we have

$$g_{n/2,n/2}^1 = \frac{y_n - (y_n + t_{n+1})\theta_2(1-\theta_1)}{2y_n - (y_n + t_{n+1})(\theta_2(1-\theta_1) + \theta_1(1-\theta_2)) + t_n} \tag{81}$$

$$g_{n/2,n/2}^2 = \frac{y_n - (y_n + t_{n+1})\theta_1(1-\theta_2)}{2y_n - (y_n + t_{n+1})(\theta_2(1-\theta_1) + \theta_1(1-\theta_2)) + t_n} \tag{82}$$

$$g_{n/2,n/2}^3 = \frac{t_n}{2y_n - (y_n + t_{n+1})(\theta_2(1-\theta_1) + \theta_1(1-\theta_2)) + t_n} \tag{83}$$

If $t_n, t_{n+1} = 0$, which happens when $n < \bar{n} - 1$, (81)-(83) reduce to

$$g_{n/2,n/2}^1 = \frac{1 - \theta_2(1-\theta_1)}{2 - \theta_2(1-\theta_1) - \theta_1(1-\theta_2)} \tag{84}$$

$$g_{n/2,n/2}^2 = \frac{1 - \theta_1(1-\theta_2)}{2 - \theta_2(1-\theta_1) - \theta_1(1-\theta_2)} \tag{85}$$

Therefore, if the sum of the two queue lengths is greater than $\bar{n}$, both users should transmit one packet during the slot. If the sum of the two queue lengths is less than $\bar{n}$, only the user with the longer queue transmits one packet in the slot and the other user remains silent; if in this case both queues have the same length, then the probability that the first user transmits one packet while the second user keeps silent is $\frac{1-\theta_2(1-\theta_1)}{2-\theta_2(1-\theta_1)-\theta_1(1-\theta_2)}$, and the probability that the second user transmits one packet while the first user keeps silent is $\frac{1-\theta_1(1-\theta_2)}{2-\theta_2(1-\theta_1)-\theta_1(1-\theta_2)}$. When the system is symmetric, i.e., $\theta_1 = \theta_2$, these probabilities become $1/2$ and $1/2$.

VI. Numerical Examples

Here we give simple examples to show how our allocation scheme works. We choose $N = 10$, i.e., each queue has a buffer of size 10 packets. Therefore, the joint queue sates is represented by an 11×11 Markov chain.

First, we consider the symmetric scenario, where $\theta_1 = \theta_2 = \theta$, $h_1 = h_2 = h$ and $P_{1avg} = P_{2avg} = P_{avg}$. We assume the arrival rate $\theta = 1/2$, and the power levels $\alpha = 1$, $\beta = 3$. Therefore, we have $\eta = 3$, $\delta = 1$, $\rho = 3$. From the analysis, we know that if $P_{avg} \geq 5/8$, the average delay is one slot, which is the minimal possible delay in the system.

If $P_{avg} = 9/16$, we have $x_{00} = 1/8$, $\sum_{i,j} x_{ij}^1 = \sum_{i,j} x_{ij}^2 = 3/8$, $\sum_{i,j} x_{ij}^3 = 1/8$. Therefore, $\Psi = 3/4$. Following our allocation scheme, we have $y_1 = 3/8$, $y_2 = 3/8$, $t_3 = 1/8$. Then, we need to allocate these within groups.

We start with y_1. Because of the symmetry of the setting, we simply let $x_{10}^1 = x_{01}^2 = y_1/2 = 3/16$, $x_{12}^3 = x_{21}^3 = t_3/2 = 1/16$. Then, we consider y_2. We also let $x_{20}^1 = x_{02}^2$, $x_{11}^1 = x_{11}^2$. This symmetric allocation guarantees that the flow equations for states $(0,1)$ and $(1,0)$ are satisfied. The values of x_{20}^1 and x_{11}^1 also depend on the allocation of t_3. The state $(2,0)$ must satisfy the transition equation

$$x_{20}^1 \left(\theta(1-\theta) + \theta^2 + (1-\theta)^2\right) = (x_{11}^2 + x_{21}^3)\theta(1-\theta)$$

Together with the symmetric allocation, we have

$$x_{20}^1 + x_{11}^2 = y_2/2 = 3/16$$

Solving these equations, we get the allocation for the second group as

$$x_{20}^1 = x_{02}^2 = 1/16, \qquad x_{11}^2 = x_{11}^1 = 1/8$$

We see that the two values are positive, thus feasible. Then, the transmission probabilities are $g_{11}^1 = g_{11}^2 = 1/2$, $g_{12}^3 = g_{21}^3 = 1$. The threshold of the sum of the queue lengths is 2 in this case. If the sum of the queue lengths is greater than 2, both users transmit, if the sum of the queue lengths is less than or equal to 2, only the user with the longer queue transmits and the other user remains silent; if both queues have one packet in their queues, each queue transmits with probability 1/2 while the other queue remains silent.

If $P_{avg} = 17/32$, we have $x_{00} = 1/16$, $\sum_{i,j} x_{ij}^1 = \sum_{i,j} x_{ij}^2 = 7/16$, $\sum_{i,j} x_{ij}^3 = 1/16$. Therefore, $\Psi = 7/8$. Following our allocation scheme, we have $y_1 = 3/16$, $y_2 = y_3 = 1/4$, $y_4 = 3/16$, $t_5 = 1/16$. Then, we assign these within groups. For y_1, we simply let $x_{10}^1 = x_{01}^2 = y_1/2 = 1/32$. Then, considering to allocate y_2, we have $x_{20}^1 = x_{02}^2 = 1/32$, $x_{11}^2 = x_{11}^1 = 3/32$. After completing the allocation, we have $x_{21}^1 = x_{12}^2 = 1/8$, $x_{31}^1 = x_{13}^2 = 1/32$, $x_{22}^2 = x_{22}^1 = 1/16$, $x_{23}^3 = x_{32}^3 = 1/32$. The transmission probabilities are $g_{11}^1 = g_{11}^2 = g_{22}^1 = g_{22}^2 = 1/2$, $g_{10}^1 = g_{01}^2 = g_{20}^1 = g_{02}^2 = g_{21}^1 = g_{12}^2 = g_{13}^1 = g_{31}^2 = g_{32}^3 = g_{23}^3 = 1$. The threshold of the sum of the queue lengths is 4 in this case. If the sum of the queue lengths is greater than 4, both users transmit, if the sum of the queue lengths is less than or equal to 4, only the user with the longer queue transmits and the other user remains silent; if both queues have equal length, which is either 1 or 2 in this case, each queue transmits with probability 1/2 while the other queue remains silent.

We compute the average delay as a function of average power for $\theta = 0.5$, $\theta = 0.48$ and $\theta = 0.46$, and plot them in Figure 6. We observe that it is a piecewise linear function, and each linear segment corresponds to the same threshold value. This is because based on our optimal allocation scheme, for a fixed threshold value, the objective function is a linear function in x_{00}, thus it is linear in P_{avg}. If P_{avg} increases, D_{avg} decreases, and the threshold decreases as well. The minimum value of P_{avg} on each curve corresponds to the maximum threshold, which is 19 in this example. This is also the minimum amount of average power required to prevent any overflows. We also observe that the delay-power tradeoff curve is convex, which is consistent with the result in [7]. We note that although these three values of θ are close to each other, the average delay varies significantly. This is because the average delay is not a linear function of θ.

For the asymmetric scenario, we assume $\theta_1 = 1/2$, $\theta_2 =$

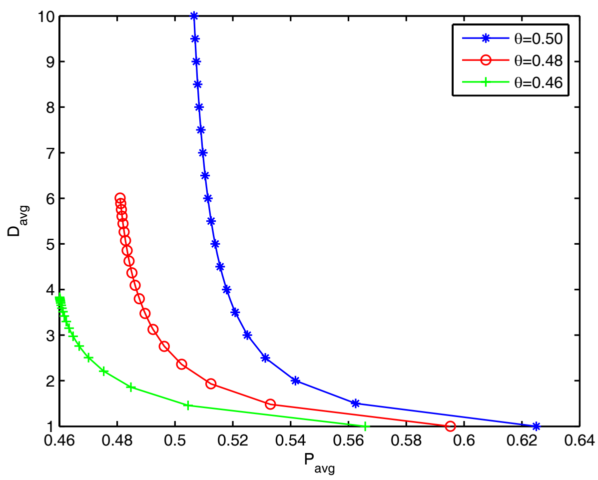

Fig. 6. The average delay versus average power in the symmetric scenario.

1/3, then $\eta = 2$, $\delta = 1/2$, $\rho = 5/2$. We assume $h_1 = 1$, $h_2 = 2$. From (33), we know that if $P_{1avg}h_1 + P_{2avg}h_2 \geq 1$, $P_{1avg} \geq 1/2$, $P_{2avg} \geq 2/3$, then each user can always transmit a packet whenever its queue is not empty, and the average delay is one slot.

If $P_{1avg} = 19/36$, $P_{2avg} = 13/18$, then $P_{1avg}h_1 + P_{2avg}h_2 = 8/9$. Plugging these into (41)-(48), we have $\beta_1 = 1/2$, $\beta_2 = 1/2$, $\sum_{i,j}^{1} x_{ij}^1 = 4/9$, $\sum_{i,j}^{2} x_{ij}^1 = 5/18$, $\sum_{i,j}^{3} x_{ij}^1 = 1/18$, $x_{00} = 2/9$. Then, $\Psi = 13/18$. Following the group allocation scheme, we have $y_1 = 4/9$, $y_2 = 5/18$, $t_3 = 1/18$. Then, we need to assign them within groups. From (118)-(124), we get $x_{01}^2 = 1/6$, $x_{10}^1 = 5/18$, $x_{02}^2 = 1/36$, $x_{11}^1 = 4/36$, $x_{11}^2 = 3/36$, $x_{20}^1 = 2/36$, and $x_{12}^3 = x_{21}^3 = 1/18$. The transmission probabilities are $g_{11}^1 = 4/7$, $g_{11}^2 = 3/7$, $g_{10}^1 = g_{01}^2 = g_{20}^1 = g_{02}^2 = g_{12}^3 = g_{21}^3 = 1$. The threshold is 2. If the sum of the queue lengths is greater than 2, both users transmit, if the sum of the queue lengths is less than or equal to 2, only the user with the longer queue transmits and the other user remains silent; if both queues have one packet in their queues, the first queue transmits with probability 4/7, and the second queue transmits with probability 3/7.

VII. Conclusions

We investigated the average delay minimization problem for a two-user multiple-access system with average power constraints for the general asymmetric scenario, where users have arbitrary powers, channel gains, and arrival rates. We considered a discrete-time model. In each slot, the arrivals at each queue follow a Bernoulli distribution, and we transmit at most one packet from each queue with some probability. Our objective is to find the optimal set of departure probabilities. We modeled the problem as a two-dimensional Markov chain, and minimized the average delay through controlling the departure probabilities in each time slot. We transformed the problem into a linear programming problem and found the optimal solution analytically. The optimal policy has a threshold structure. Whenever the sum of the queue lengths exceeds a threshold, both queues transmit one packet during the slot, otherwise, only one of the queues, which is longer, transmits one packet during the slot and the other queue remains silent; if both queues have the same length, only one of the queues transmits with a probability which depends on the arrival rates to both queues while the other queue remains silent.

Appendix

A. The Proof of Theorem 1

Let us define

$$\eta = \frac{\theta_1 + \theta_2 - \theta_1\theta_2}{(1-\theta_1)(1-\theta_2)} \tag{86}$$

$$\delta = \frac{\theta_1\theta_2}{(1-\theta_1)(1-\theta_2)} \tag{87}$$

$$\rho = \frac{1-\theta_1\theta_2}{(1-\theta_1)(1-\theta_2)} \tag{88}$$

Then, (56)-(59) are equivalent to

$$x_{00}\eta = y_1 + t_2 \tag{89}$$

$$(x_{00} + y_1)\delta = (y_2 + t_3) + t_2\rho \tag{90}$$

and for $n = 2, 3, \ldots, 2N-2$,

$$y_n\delta = (y_{n+1} + t_{n+2}) + t_{n+1}\rho \tag{91}$$

$$y_{2N-1}\delta = t_{2N}\rho \tag{92}$$

The optimization requires us to assign larger values to y_ns with smaller indices n as much as possible. Examining (89)-(92), we note that for fixed x_{00}, maximizing y_1, y_2, ... requires us to set t_2, t_3, ... to zero. Therefore, we choose

$$y_1 = x_{00}\eta \tag{93}$$

$$y_2 = (x_{00} + y_1)\delta \tag{94}$$

$$y_n = y_{n-1}\delta, \quad t_n = 0, \quad n = 1, 2, \ldots, n^* \tag{95}$$

where n^* is the largest integer satisfying $\sum_{n=1}^{n^*} y_n < \Psi$.

Let $\Delta = \Psi - \sum_{n=1}^{n^*} y_n$. We need to check that all of the group transition equations are satisfied.

Assume that $n^* > 2$. If $\Delta = y_{n^*}\delta\rho/(\delta+\rho)$, then let

$$y_{n^*+1} = \Delta, \quad \text{and} \quad y_n = 0, \quad n = n^*+2, \ldots, N-1 \tag{96}$$

$$t_{n^*+2} = y_{n^*+1}\delta/\rho, \quad \text{and} \quad t_n = 0, \quad n \neq n^*+2 \tag{97}$$

We can verify that after this allocation, group transition equations (56)-(59) are satisfied. We also note that Ψ is allocated to $\{y_n\}_{n=1}^{n^*+1}$, among which, $\{y_n\}_{n=1}^{n^*}$ attain their maximum possible values. Therefore, the objective function achieves its minimal possible value for the first step.

If $\Delta \neq y_{n^*}\delta\rho/(\delta+\rho)$, if we assign it to y_{n^*+1} directly, the group transition equations are not satisfied automatically. In order to satisfy the group transition equations, we need to do some adjustments.

If $\Delta > y_{n^*}\delta\rho/(\delta+\rho)$, we assign Δ to y_{n^*+1} and y_{n^*+2} proportionally. Specifically, we let

$$y_{n^*+1} = \frac{\Delta(\rho+\delta) + y_{n^*}\delta\rho^2}{\rho^2 + \delta\rho + \delta + \rho} \tag{98}$$

$$y_{n^*+2} = \frac{\Delta(\rho+\delta)\rho - y_{n^*}\delta\rho^2}{\rho^2 + \delta\rho + \delta + \rho} \tag{99}$$

$$t_{n^*+2} = \frac{y_{n^*}\delta(\delta\rho + \delta + \rho) - \Delta(\rho+\delta)}{\rho^2 + \delta\rho + \delta + \rho} \tag{100}$$

$$t_{n^*+3} = \frac{\Delta(\rho+\delta)\delta - y_{n^*}\delta^2\rho}{\rho^2+\delta\rho+\delta+\rho} \tag{101}$$

Since $y_{n^*}\delta > \Delta > y_{n^*}\delta\rho/(\delta+\rho)$, we can verify that each value above is positive, and the sum constraint and the group transition equations are satisfied. Among the non-zero $\{y_n\}_{n=1}^{n^*+2}$, although $\{y_n\}_{n=1}^{n^*}$ attain their maximum, y_{n^*+1} does not. Therefore, different from the first scenario, in this case, we cannot immediately claim that the result is optimal. We will give the mathematical proof for the optimality of this assignment later.

If $\Delta < y_{n^*}\delta\rho/(\delta+\rho)$, we need to remove some value from y_{n^*} and assign it to y_{n^*+1} to satisfy the equations. Define $\Delta' = \Delta + y_{n^*}$ and assign Δ' to y_{n^*} and y_{n^*+1} as follows

$$y_{n^*} = \frac{\Delta'(\rho+\delta) + y_{n^*-1}\delta\rho^2}{\rho^2+\delta\rho+\delta+\rho} \tag{102}$$

$$y_{n^*+1} = \frac{\Delta'(\rho+\delta)\rho - y_{n^*-1}\delta\rho^2}{\rho^2+\delta\rho+\delta+\rho} \tag{103}$$

$$t_{n^*+1} = \frac{y_{n^*-1}\delta(\delta\rho+\delta+\rho) - \Delta'(\rho+\delta)}{\rho^2+\delta\rho+\delta+\rho} \tag{104}$$

$$t_{n^*+2} = \frac{\Delta'(\rho+\delta)\delta - y_{n^*-1}\delta^2\rho}{\rho^2+\delta\rho+\delta+\rho} \tag{105}$$

Since $y_{n^*-1}\delta < \Delta' < y_{n^*-1}\delta(\delta\rho/(\delta+\rho)+1)$, we can also verify that each value above is positive, and the sum constraint and the group transition equations are satisfied. Similar to the second case, we cannot immediately claim that this result is optimal because after the adjustment, y_{n^*} does not achieve its maximum value. We will give the proof of optimality later.

When $n^* = 1$, the allocation will be in a different form. If $\Delta \geq (x_{00}+y_1)\delta\rho/(\delta+\rho)$, then we need to use $(x_{00}+y_1)$ instead of y_{n^*} in (96)-(101). If $\Delta < (x_{00}+y_1)\delta\rho/(\delta+\rho)$, then

$$y_{n^*} = \frac{\Psi(\rho+\delta) + x_{00}(\eta-\delta)\rho}{\rho^2+\delta\rho+\delta+\rho} \tag{106}$$

$$y_{n^*+1} = \frac{\Psi(\rho+\delta)\rho - x_{00}(\eta-\delta)\rho}{\rho^2+\delta\rho+\delta+\rho} \tag{107}$$

$$t_{n^*+1} = \frac{x_{00}(\eta\delta\rho+\eta\delta+\eta\rho^2+\delta\rho) - \Psi(\rho+\delta)}{\rho^2+\delta\rho+\delta+\rho} \tag{108}$$

$$t_{n^*+2} = \frac{\Psi(\rho+\delta)\delta - x_{00}(\eta-\delta)\delta}{\rho^2+\delta\rho+\delta+\rho} \tag{109}$$

When $n^* = 2$, if $\Delta \geq y_{n^*}\delta\rho/(\delta+\rho)$, the allocation of Ψ has the same form as in (96)-(101). If $\Delta < y_{n^*}\delta\rho/(\delta+\rho)$, then we need to use $(x_{00}+y_1)$ instead of y_{n^*-1} in (102)-(105).

B. The Proof of Theorem 2

While we generalize the simple example to an arbitrary setting, we follow the same basic allocation pattern. If n is odd, we assign y_n and t_n only to two states $\left(\frac{n+1}{2}, \frac{n-1}{2}\right)$ and $\left(\frac{n-1}{2}, \frac{n+1}{2}\right)$; if n is even, we assign y_n to three states: $\left(\frac{n}{2}+1, \frac{n}{2}-1\right), \left(\frac{n}{2}, \frac{n}{2}\right), \left(\frac{n}{2}-1, \frac{n}{2}+1\right)$, and we assign t_n to a single state $\left(\frac{n}{2}, \frac{n}{2}\right)$. We illustrate the allocation pattern in Figure 7. We need to make sure that the transitions only happen within the positive recurrent class. Therefore, when n is odd, we let $x^1_{\frac{n-1}{2},\frac{n+1}{2}} = x^2_{\frac{n+1}{2},\frac{n-1}{2}} = 0$; when n is even, we let $x^1_{\frac{n}{2}-1,\frac{n}{2}+1} = x^2_{\frac{n}{2}+1,\frac{n}{2}-1} = 0$. Then, let us examine the

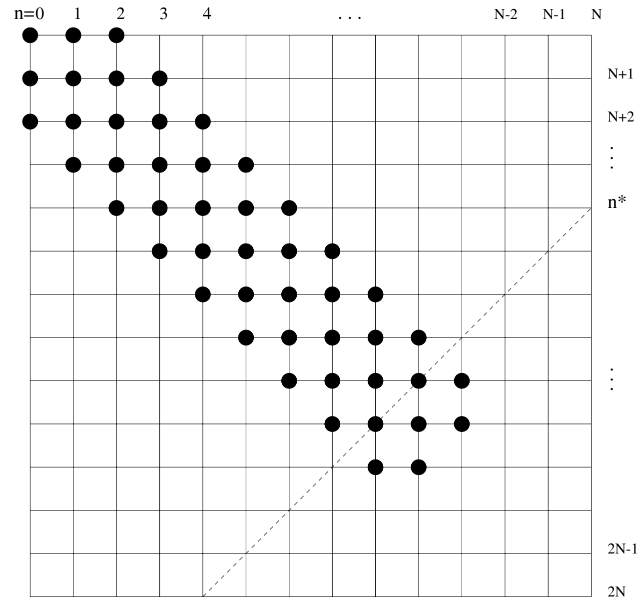

Fig. 7. Allocation pattern within groups.

transition equations for the states. For $n = 1$, we have

$$x^2_{01}(1-\theta_2(1-\theta_1)) = (x_{00} + x^1_{10} + x^3_{11})\theta_2(1-\theta_1) + (x^2_{02} + x^1_{11} + x^3_{12})(1-\theta_1)(1-\theta_2) \tag{110}$$

For $n = 2, 3, \ldots$, if n is even, the transitions between states are illustrated in Figure 8. The transition equation for state $\left(\frac{n}{2}-1, \frac{n}{2}+1\right)$ is

$$x^2_{\frac{n}{2}-1,\frac{n}{2}+1}(1-\theta_2(1-\theta_1)) = (x^1_{\frac{n}{2},\frac{n}{2}} + x^3_{\frac{n}{2},\frac{n}{2}+1})\theta_2(1-\theta_1) \tag{111}$$

If n is odd, the transitions between states are illustrated in Figure 9. The transition equation for state $\left(\frac{n-1}{2}, \frac{n+1}{2}\right)$ is

$$\begin{aligned} &x^2_{\frac{n-1}{2},\frac{n+1}{2}}(1-\theta_2(1-\theta_1)) + x^3_{\frac{n-1}{2},\frac{n+1}{2}}(1-\theta_1\theta_2) \\ &= (x^2_{\frac{n-3}{2},\frac{n+1}{2}} + x^1_{\frac{n-1}{2},\frac{n-1}{2}})\theta_1\theta_2 \\ &\quad + (x^1_{\frac{n+1}{2},\frac{n+1}{2}} + x^2_{\frac{n-1}{2},\frac{n+3}{2}} + x^3_{\frac{n+1}{2},\frac{n+3}{2}})(1-\theta_1)(1-\theta_2) \\ &\quad + (x^1_{\frac{n+1}{2},\frac{n-1}{2}} + x^3_{\frac{n+1}{2},\frac{n+1}{2}})\theta_2(1-\theta_1) \end{aligned} \tag{112}$$

After a transformation, (110) is equivalent to

$$x^2_{01} = (x_{00} + x^1_{10} + x^2_{01} + x^3_{11})\theta_2(1-\theta_1) + (x^2_{02} + x^1_{11} + x^3_{12})(1-\theta_1)(1-\theta_2) \tag{113}$$

where x_{00} is known, $x^1_{10} + x^2_{01} = y_1$, $x^3_{11} = t_2$.

For $n = 2, 3, \ldots$, when n is even, (111) is equivalent to

$$x^2_{\frac{n}{2}-1,\frac{n}{2}+1} = (x^1_{\frac{n}{2},\frac{n}{2}} + x^2_{\frac{n}{2}-1,\frac{n}{2}+1} + x^3_{\frac{n}{2},\frac{n}{2}+1})\theta_2(1-\theta_1) \tag{114}$$

and when n is odd, (112) is equivalent to

$$\begin{aligned} x^2_{\frac{n-1}{2},\frac{n+1}{2}} &= (x^2_{\frac{n-3}{2},\frac{n+1}{2}} + x^1_{\frac{n-1}{2},\frac{n-1}{2}})\theta_1\theta_2 - x^3_{\frac{n-1}{2},\frac{n+1}{2}}(1-\theta_1\theta_2) \\ &\quad + (x^1_{\frac{n+1}{2},\frac{n+1}{2}} + x^2_{\frac{n-1}{2},\frac{n+3}{2}} + x^3_{\frac{n+1}{2},\frac{n+3}{2}})(1-\theta_1)(1-\theta_2) \\ &\quad + (x^1_{\frac{n+1}{2},\frac{n-1}{2}} + x^2_{\frac{n-1}{2},\frac{n+1}{2}} + x^3_{\frac{n+1}{2},\frac{n+1}{2}})\theta_2(1-\theta_1) \end{aligned} \tag{115}$$

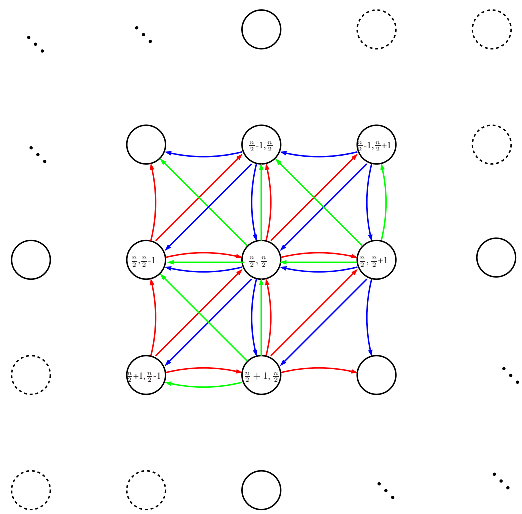

Fig. 8. The transitions between states when n is even.

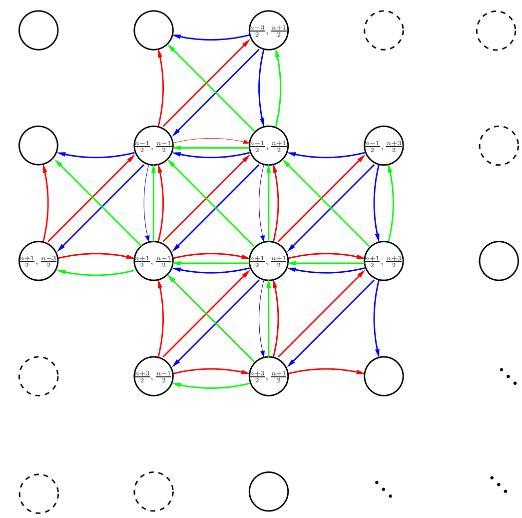

Fig. 9. The transitions between states when n is odd.

where $x^1_{\frac{n+1}{2},\frac{n-1}{2}} + x^2_{\frac{n-1}{2},\frac{n+1}{2}} = y_n$, $x^3_{\frac{n+1}{2},\frac{n+1}{2}} = t_{n+1}$.

The transition equations for the remaining half of the recurrent states can be expressed in a similar form. Therefore, the values of x^k_{ij}s are determined only by the allocation of y_n between $x^1_{\frac{n}{2}+1,\frac{n}{2}-1} + x^2_{\frac{n}{2},\frac{n}{2}}$ and $x^2_{\frac{n}{2}-1,\frac{n}{2}+1} + x^1_{\frac{n}{2},\frac{n}{2}}$ when n is even, and the allocation of t_n to $x^3_{\frac{n+1}{2},\frac{n-1}{2}}$ and $x^3_{\frac{n-1}{2},\frac{n+1}{2}}$ when n is odd. If we let

$$x^1_{\frac{n}{2},\frac{n}{2}} + x^2_{\frac{n}{2}-1,\frac{n}{2}+1} = y_n/2, \qquad \text{when } n \text{ is even} \tag{116}$$

$$x^3_{\frac{n-1}{2},\frac{n+1}{2}} = t_n/2, \qquad \text{when } n \text{ is odd} \tag{117}$$

and solve equations (113)-(115), then, for $n = 1$, we obtain

$$\begin{aligned} x^2_{01} =& (x_{00} + y_1 + t_2)\theta_2(1-\theta_1) \\ &+ \frac{1}{2}(y_2 + t_3)(1-\theta_1)(1-\theta_2) \\ x^1_{10} =& (x_{00} + y_1 + t_2)\theta_1(1-\theta_2) \\ &+ \frac{1}{2}(y_2 + t_3)(1-\theta_1)(1-\theta_2) \end{aligned} \tag{118}$$

For $n = 2, 3, \ldots$, if n is even, we get

$$x^2_{\frac{n}{2}-1,\frac{n}{2}+1} = \frac{1}{2}(y_n + t_{n+1})\theta_2(1-\theta_1) \tag{119}$$

$$x^1_{\frac{n}{2}+1,\frac{n}{2}-1} = \frac{1}{2}(y_n + t_{n+1})\theta_1(1-\theta_2) \tag{120}$$

$$x^1_{\frac{n}{2},\frac{n}{2}} = \frac{1}{2}y_n - \frac{1}{2}(y_n + t_{n+1})\theta_2(1-\theta_1) \tag{121}$$

$$x^2_{\frac{n}{2},\frac{n}{2}} = \frac{1}{2}y_n - \frac{1}{2}(y_n + t_{n+1})\theta_1(1-\theta_2) \tag{122}$$

and if n is odd, we have

$$\begin{aligned} x^2_{\frac{n-1}{2},\frac{n+1}{2}} =& \frac{1}{2}y_{n-1}\theta_1\theta_2 + (y_n + t_{n+1})\theta_2(1-\theta_1) \\ &+ \frac{1}{2}(y_{n+1} + t_{n+2})(1-\theta_1)(1-\theta_2) - \frac{1}{2}t_n(1-\theta_1\theta_2) \end{aligned} \tag{123}$$

$$\begin{aligned} x^1_{\frac{n+1}{2},\frac{n-1}{2}} =& \frac{1}{2}y_{n-1}\theta_1\theta_2 + (y_n + t_{n+1})\theta_1(1-\theta_2) \\ &+ \frac{1}{2}(y_{n+1} + t_{n+2})(1-\theta_1)(1-\theta_2) - \frac{1}{2}t_n(1-\theta_1\theta_2) \end{aligned} \tag{124}$$

This completes the allocation. Note that $t_n \neq 0$ only when n is equal to n^*+1, n^*+2, and/or n^*+3, depending on the value of Δ. When $t_{n+1} = 0$, it automatically disappears from the right hand sides of (118)-(124). From the group transition equations, we have $y_n \geq t_{n+1}\rho'/\delta'$, and it is easy to verify that all states have nonnegative assignments and the transition equations are also satisfied in this case. Therefore, there always exists a feasible allocation to satisfy all of the transition equations with y_ns defined through this allocation scheme.

C. The Proof of Theorem 3

In a convex optimization problem, where the inequality constraints are convex and the equality constraints are affine, if x^* is such that there exists a set of Lagrange multipliers which together with x^* satisfy the KKT conditions, then x^* is a global minimizer for the problem [17] [18]. In the first step, we have a linear objective function and linear constraints. Therefore, if we prove that the point achieved by the assignment satisfies the KKT conditions, then we can say that it is the global minimizer for our problem.

In the allocation scheme, if $\Delta = y_{n^*}\delta\rho/(\delta+\rho)$, then it is easy to prove that the resulting allocation is optimal, since every $y_n, n < n^*$ achieves its maximum possible value. However, this is not the case when $\Delta \neq y_{n^*}\delta\rho/(\delta+\rho)$, because the second to last nonzero y_n does not achieve its maximum. In the following, we prove that our allocation is optimal for this case as well. Define $\mathbf{y} = [y_1, y_2, \ldots, y_{2N-1}, t_2, \ldots, t_{N-1}, t_{2N}]$. Then, the linear equality constraints, including the $2N$ group transition equations and the sum constraint can be written as a $(2N+1)\times 2(2N-1)$ matrix form as follows

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\ -\delta & 1 & 0 & \cdots & 0 & \rho & 1 & 0 & \cdots & 0 \\ 0 & -\delta & 1 & \cdots & 0 & 0 & \rho & 1 & \cdots & 0 \\ & \vdots & & \ddots & \vdots & & & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & -\delta & 0 & 0 & 0 & \cdots & \rho \\ 1 & 1 & 1 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \mathbf{y}^T = \begin{pmatrix} x_{00}\eta \\ x_{00}\delta \\ 0 \\ \vdots \\ 0 \\ 0 \\ \Psi \end{pmatrix}$$

which we write equivalently as,

$$\mathbb{A}\mathbf{y}^T = \mathbf{b} \tag{125}$$

$$\mathbf{b}^T = \begin{pmatrix} x_{00}\eta & x_{00}\delta(1+\eta) & x_{00}\delta^2(1+\eta) & \cdots & x_{00}\delta^{2N-1}(1+\eta) & \Psi \end{pmatrix}^T \tag{126}$$

$$\mathbb{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & \rho+\delta & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & (\rho+\delta)\delta & \rho+\delta & 1 & \cdots & 0 \\ & \vdots & & \ddots & \vdots & & & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 & (\rho+\delta)\delta^{2N-3} & (\rho+\delta)\delta^{2N-4} & (\rho+\delta)\delta^{2N-5} & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 & (\rho+\delta)\delta^{2N-2} & (\rho+\delta)\delta^{2N-3} & (\rho+\delta)\delta^{2N-4} & \cdots & \rho+\delta \\ 1 & 1 & 1 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{127}$$

by defining $\mathbf{b}$, $\mathbb{A}$ in (126) and (127) at the top of the page.

The Lagrangian is expressed as

$$L(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{c}^T\mathbf{y} - \boldsymbol{\lambda}^T(\mathbb{A}\mathbf{y} - \mathbf{b}) - \boldsymbol{\mu}^T\mathbf{y} \tag{128}$$

where $\mathbf{c} = [1, 2, \cdots, 2N-1, 0, 0, \cdots, 0]$, $\boldsymbol{\lambda} \in \mathbf{R}^{2N+1}$ and $\boldsymbol{\mu} \in \mathbf{R}^{4N-2}$.

We need to prove that there exists a set of $\boldsymbol{\lambda}^*$, $\boldsymbol{\mu}^*$ associated with our allocation $\mathbf{y}^*$, such that they satisfy

$$\boldsymbol{\mu}^* \geq \mathbf{0}, \quad {\boldsymbol{\mu}^*}^T\mathbf{y}^* = 0 \tag{129}$$

$$\mathbf{y}^* \geq \mathbf{0}, \quad \mathbb{A}{\mathbf{y}^*}^T = \mathbf{b} \tag{130}$$

$$\mathbf{c} = \mathbb{A}^T\boldsymbol{\lambda}^* + \boldsymbol{\mu}^* \tag{131}$$

Consider the $\mathbf{y}$ we obtained with the algorithm. Let us consider the case when $\Delta < y_{n^*}\delta\rho/(\delta+\rho)$ first. The allocation indicates that $y_n > 0$ only when $n = 1, 2, \ldots, n^*+1$, and $t_n > 0$ only when $n = n^*+1, n^*+2$. Because of the complementary slackness in (129), we obtain

$$\mu_n = 0, \quad n = 1, 2, \ldots, n^*+1, n^*+2N-1, n^*+2N \tag{132}$$

Plugging this into (131), and solving the equations, we have

$$\begin{aligned} \lambda_n &= \frac{1}{\rho+1} + n - n^* - 1, \quad n = 1, 2, \ldots, n^*+1 \\ \lambda_{2N+1} &= \frac{\rho}{\rho+1} + n^* \\ \mu_{n+2N-2} &= -\left(\lambda_{n-1} + (\rho+\delta)\sum_{i=n}^{n^*-1}\lambda_i\delta^{i-n} + \rho\delta^{n^*-n}\lambda_{n^*}\right), \\ &\quad n = 2, 3, \ldots, n^* \end{aligned} \tag{133}$$

Thus, we have $\lambda_n < 0$ when $n \leq n^*$, which guarantees the positiveness of $\{\mu_n\}_{n=2N}^{n^*+2N-2}$. We also have

$$\sum_{i=n^*+2}^{2N}\lambda_i\delta^{i-n^*-2} = -\frac{1}{(\rho+\delta)(\rho+1)} \tag{134}$$

and

$$\begin{aligned} \mu_n &= \frac{1}{\rho+1} + n - n^* - 1 - \lambda_n, \quad n = n^*+2, \ldots, 2N-1 \\ \mu_n &= -\left(\lambda_{n-1} + (\rho+\delta)\sum_{i=n}^{2N}\lambda_i\delta^{i-n}\right), \\ &\quad n = n^*+2N+1, \ldots, 4N-2 \end{aligned} \tag{135}$$

We can always find a set of negative $\{\lambda_i\}_{i=n^*+2}^{2N}$ to satisfy (134). Since they are all negative, this guarantees that $\{\mu_n\}_{n=n^*+2}^{2N-1}$ and $\{\mu_n\}_{n=n^*+2N+1}^{4N-2}$ are positive. Therefore, at the point $\mathbf{y}^*$, we can always find a set of multipliers satisfying all of the KKT constraints. This proves that the allocation our algorithm gives is a global minimizer.

REFERENCES

[1] J. Yang and S. Ulukus, "Delay-minimal transmission for average power constrained multi-access communications," *Asilomar Conference*, 2008.

[2] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inf. Theory*, vol. 39, pp. 466–478, Mar. 1993.

[3] I. E. Telatar and R. G. Gallager, "Combining queueing theory with information theory for multiaccess," *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 963–969, Aug. 1995.

[4] E. Yeh, "Delay-optimal rate allocation in multiaccess communications: a cross-layer view," in *Proc. IEEE Workshop on Multimedia Signal Processing 2002*, pp. 404–407, Dec. 2002.

[5] E. Yeh and A. Cohen, "Throughput optimal power and rate control for queued multi-access and broadcast communications," in *Proc. IEEE International Symposium on Information Theory*, p. 112, June 2004.

[6] E. Yeh, "Minimum delay multi-access communication for general packet length distributions," in *Proc. Allerton Conference on Communication, Control, and Computing*, pp. 1536–1545, Sep. 2004.

[7] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1135–1149, May 2002.

[8] R. Berry, "Power and delay trade-offs in fading channels," Ph.D. dissertation, MIT, June 2000.

[9] I. Bettesh and S. Shamai, "Optimal power and rate control for minimal average delay: the single-user case," *IEEE Trans. Inf. Theory*, vol. 52, pp. 4115–4141, Sep. 2006.

[10] A. Steiner and S. Shamai, "On queueing and multi-layer coding," *Conference on Information Sciences and Systems*, Mar. 2005.

[11] M. Goyal, A. Kumar, and V. Sharma, "Power constrained and delay optimal policies for scheduling transmission over a fading channel," in *Proc. IEEE Infocom*, vol. 1, pp. 311–320, Mar. 2003.

[12] T. Holliday and A. Goldsmith, "Optimal power control and source-channel coding for delay constrained traffic over wireless channels," in *Proc. IEEE Int. Conf. Communications*, vol. 2, pp. 831–835, Apr. 2002.

[13] W. Chen, K. B. Letaief, and Z. Cao, "A joint coding and scheduling method for delay optimal cognitive multiple access," in *Proc. IEEE ICC*, 2008.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons, Inc., 1991.

[15] E. Uysal-Biyikoglu and A. El Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3081–3094, Dec. 2004.

[16] V. G. Kulkarni, *Modeling and Analysis of Stochastic Systems*. Chapman and Hall/CRC, 1995.

[17] A. Tits, *Optimal Control*, lecture notes, University of Maryland, 2002.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. United Kingdom: Cambridge University Press, 2004.

Jing Yang received the B.S. degree in electronic engineering and information science from University of Science and Technology of China, Hefei, China in 2004. Since then, she has been a PhD graduate student in the department of electrical and computer engineering at the University of Maryland, College Park. Her research interests are in wireless communication theory and networking, multi-user information theory, queueing theory and optimization in wireless networks, with particular focus on delay optimization in wireless communications and networks.

Sennur Ulukus received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 1991 and 1993, respectively, and the Ph.D. degree in electrical and computer engineering from Rutgers University, NJ, in 1998. During her Ph.D. studies, she was with the Wireless Information Network Laboratory (WINLAB), Rutgers University. From 1998 to 2001, she was a Senior Technical Staff Member at AT&T Labs-Research in NJ. In 2001, she joined the University of Maryland at College Park, where she is currently an Associate Professor in the Department of Electrical and Computer Engineering, with a joint appointment at the Institute for Systems Research (ISR). Her research interests are in wireless communication theory and networking, network information theory for wireless networks, signal processing for wireless communications and security for multi-user wireless communications.

Sennur Ulukus is a recipient of the 2005 NSF CAREER Award, and a co-recipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications. She serves/served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY since 2007, as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS between 2003-2007, as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2006-2008, as the co-chair of the Communication Theory Symposium at the 2007 IEEE Global Telecommunications Conference, as the co-chair of the Medium Access Control (MAC) Track at the 2008 IEEE Wireless Communications and Networking Conference, as the co-chair of the Wireless Communications Symposium at the 2010 IEEE International Conference on Communications, as the co-chair of the 2011 Communication Theory Workshop, and as the Secretary of the IEEE Communication Theory Technical Committee (CTTC) in 2007-2009.