Delay Minimization with a General Pentagon Rate Region

Jing Yang Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
yangjing@umd.edu *ulukus@umd.edu*

***Abstract*—We consider a communication channel with two transmitters and one receiver, with an underlying rate region which is approximated as a general pentagon. Different from the Gaussian multiple access channel (MAC) capacity region, the sum-rate on the dominant face of this pentagon is not a constant. We allocate rates from this rate region to users according to their current queue lengths in order to minimize the average delay in the system. We formulate the problem as a Markov decision problem (MDP), and derive the structural properties of the corresponding discounted-cost MDP. We show that the delay-optimal policy has a switch curve structure. For the discounted-cost problem, we prove that the switch curve has a limit along one of the dimensions.**

I. Introduction

Traditional information theory ignores the burstiness of arrivals and the associated issue of delay by assuming that all of the bits have already arrived and are available at the transmitter before the transmission starts. This is necessary to invoke asymptotics (e.g., large block sizes), which is needed to prove reliability of communication. Network and queueing theory, on the other hand, give sophisticated analysis for delay and related issues, but, assume simplified models for the underlying communication rates, which serve as the server rates of the queues. Network theory typically assumes slotted or time-divided communications in order to minimize the interactions between the queues, as the analysis of interacting queues is known to be notoriously difficult. Many authors have pointed to the need to bring information and network theory together to jointly address the goals of reliability, high rates and low delay, e.g., [1]. The goal of this paper is to use a general pentagon shaped underlying rate region (hence, non-time-divided transmissions) and determine the optimal rate allocation policy from this available rate region, as a function of the current queue sizes of the users, to minimize the delay.

Delay minimization for a single-user communication channel has been investigated in [2]–[4], where the structural properties of the optimum power/rate allocation policies, and relationships between average power and delay have been determined for fading channels, using dynamic programming and Markov decision process (MDP) formulations. In these works, due to the large number of possible rate/power choices at each channel state, it has been almost impossible to get analytical closed-form optimal solutions. For multi-user systems, even the properties of the optimum rate allocation have been impossible to obtain, except for special rate regions.

Reference [5] considers a symmetric Gaussian multiple access channel (MAC), whose capacity region for two-users is a symmetric pentagon. Reference [5] proves that in order to minimize the *packet* delay, the system should operate at an extreme point of the MAC capacity region, i.e., at one of the two corner points of the symmetric pentagon. In particular, [5] determines explicitly the corner point the system should operate at as a function of the queue sizes, by proving that the larger rate should be given to the user with the larger queue size, hence the name of the proposed policy: longer-queue-higher-rate (LQHR). Reference [6] generalizes [5] to a potentially asymmetric setting, and proves that the system should again operate at one of the two corner points of the capacity region, which in this case is a potentially asymmetric pentagon. This proves that the delay-optimal policy has a switch structure, i.e., that the queue state space should be divided into two, and in each region, the system should operate at one of the two corner points. However, unlike the symmetric case in [5], the explicit form of the switch curve is unknown. Reference [7] develops a policy named “modified LQHR” which works at a corner point of the pentagon when the queue lengths are different, and switches to the mid-point of the dominant face of the pentagon when the queue lengths become equal. The “modified LQHR” algorithm is shown to minimize the average *bit* delay in a symmetric system. The third chapter of [8] extends “modified LQHR” to a symmetric M-user scenario. In [9], we consider a *discrete-time* symmetric Gaussian MAC, and prove that the *queue length balancing* policy, which minimizes the queue length difference while working on the dominant face of the capacity region in each slot, minimizes the average bit delay in the system.

From the literature above, we observe that the explicit solution of the queue-length based delay-minimization problem is known only for the symmetric Gaussian MAC, where the underlying rate region is a symmetric pentagon. Even for the asymmetric pentagon, the delay-minimizing policy is not known. The reason for this is that delay-minimization requires maximizing the throughput at the current time as well as maximizing the throughput in the future. These are often conflicting objectives. The first objective requires maximizing the sum-rate while the second objective requires balancing the queue lengths. Unbalanced queue lengths increases the

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311 and CCF 07-29127.

likelihood of one of the queues becoming empty, which results in inefficiency of transmission, as it decreases the future achievable sum-rates. Thanks to the special properties of the capacity region of the symmetric Gaussian MAC, these two objectives can be achieved simultaneously.

However, having a symmetric pentagon as a capacity region is a peculiarity of the symmetric Gaussian MAC. The capacity region of a general (non-Gaussian) MAC is not a pentagon, it is a union of pentagons. Likewise, the capacity regions of the fading Gaussian MAC, the Gaussian MAC with multiple antennas, or the Gaussian MAC with user cooperation are not pentagons. In this paper, we will consider a two-user communication channel with a general pentagon rate region. Different from the Gaussian MAC capacity region, the pentagon we assume does not have a $45°$ dominant face. The motivation to study such a rate region is two-fold: First, it is the simplest extension of the rate regions studied so far, that changes a characteristic in a fundamental way. This characteristic is that the two corner points on the dominant face do not have equal sum-rates. Therefore, in this example rate region, we are able to observe the tension between throughput optimality, i.e., the desire to work at the point that yields the largest sum-rate, and balancing the queue lengths, i.e., the desire to favor the longer queue over the shorter one, more explicitly. Secondly, this asymmetric pentagon with a non-$45°$ dominant face can be seen as a crude approximation of a general rate region, as shown in Fig. 1. That is, we can imagine this asymmetric pentagon to be the largest such shape fitting in a general rate region.

Our goal in this paper is to assign rate pairs to users from the underlying rate region based on their current queue lengths in order to minimize the average delay in the system. We formulate the problem as an MDP and prove that the delay-optimal policy should operate at one of the two corner points of the rate region. Through value iteration, we prove that a switch curve structure exists in the queue state space. Next, we prove that for the discounted-cost MDP, the switch curve has a limit on one of the queue lengths, i.e., when one of the queue lengths exceeds a threshold, the transmitters always operate at the corner point which has the larger sum-rate (see Fig. 2). That is, the delay-optimal policy favors throughput-optimality (i.e., larger sum-rate) unless the other queue gets very large, in which case, the policy favors balancing queue lengths. Our result has two practical implications: First, it gives a *partial analytical characterization* for the delay-optimal switch curve. Secondly, it implies that we can operate the queues *partially distributedly*, in that, if the current queue length of the first user is larger than the limit, then this user does not need to know the current queue length of the other user in order to decide about the rate point at which it should operate on the rate region.

II. System Model and Problem Formulation

We consider a communication system with two transmitters, and one receiver. The underlying rate region is a general pentagon as shown in Fig. 1. We denote the two corner points as points 1 and 2, with rate pairs (a_1, b_1) and (a_2, b_2), respectively. Without loss of generality, we assume that $a_1 + b_1 < a_2 + b_2$, i.e., that point 2 has the larger sum-rate. We denote the difference between the two sum-rates $\delta = a_2 + b_2 - (a_1 + b_1)$.

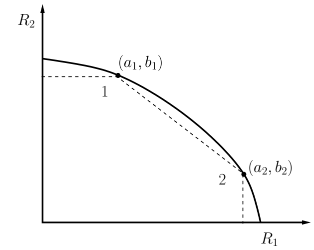

Fig. 1. The asymmetric pentagon rate region with non-$45°$ dominant face.

In the medium access control layer, we assume that packets arrive at the source nodes according to independent Poisson processes with parameters λ_1 and λ_2. We also assume that the packet lengths are independent and identically distributed exponential random variables with unit mean. Therefore, for a given transmission rate r, the transmission time for a packet is an exponential random variable with parameter r. There is a buffer with infinite capacity at each transmitter, storing the packets until they are transmitted. Let $q_1(t)$, $q_2(t)$ denote the number of packets in the two buffers at time t. The transmitters determine their transmission rates, which are the components of the rate vector $\mathbf{r}$, where $\mathbf{r}$ is in the rate region, based on the current queue length vector $\mathbf{q}(t) = (q_1(t), q_2(t))$. Therefore, on the medium access control layer, the queue lengths evolve according to a continuous-time Markov chain, whose transition rates are determined by the arrival and transmission rates.

According to Little's law [10], minimizing the average delay in the system is equivalent to minimizing the average number of packets in the system. If the system starts from state $\mathbf{q}(0)$, the delay minimization problem is to obtain an optimal policy $\mathbf{u}$, to minimize the long-term average cost:

$$\limsup_{t\to\infty} \frac{1}{t} E\left[\int_0^t \mathbf{q}(s)^T \mathbf{e} ds | q(0)\right] \tag{1}$$

where $\mathbf{e}$ is the vector of all ones.

Sampling the system at certain epoches, we can convert the original continuous-time problem into a discrete-time problem [11]. Intuitively, we intend to sample the system at any epoch when an arrival or departure occurs. However, because the transition rates are different at different operating points, the sampling frequency may be different for different states. In order to sample the system at a uniform frequency, we adopt the normalized method in [12]. Since $a_2 + b_2$ is the maximum sum of transmission rates, the maximum total transition rate of the system is $\lambda_1 + \lambda_2 + a_2 + b_2$, which we define as γ. Let us denote the transmission rates of the users as r_1 and r_2. If $r_1 + r_2 < a_2 + b_2$, we assume that there is a third transmitter transmitting a dummy packet with rate $a_2 + b_2 - (r_1 + r_2)$. Then, we sample at the epoches when either a packet arrives, or a packet (dummy or real) departs. Therefore, the sampling frequency for all of the states will be the same, and the corresponding discrete-time Markov chain will precisely represent the original system.

After sampling and discretizing the continuous-time system, our goal will be to choose $\mathbf{r}$ at every transition epoch to minimize the average delay. Let us denote the indices of the transition epoches as n, $n = 1, 2, \ldots$. Given the initial queue lengths $\mathbf{q}_0$, the delay minimization problem is to determine the optimal policy that minimizes:

$$\limsup_{N\to\infty} \frac{1}{N} E\left[\sum_{n=0}^{N-1} \mathbf{q}[n]^T \mathbf{e} | \mathbf{q}[0] = \mathbf{q}_0\right] \quad (2)$$

Let us define A_i and D_i to be an arrival or (potential) departure at the ith queue, $i = 1, 2$. For example, $A_1\mathbf{q} = (q_1+1, q_2)$, $D_1\mathbf{q} = ((q_1 - 1)^+, q_2)$. We first define the corresponding discounted-cost problem with a discount factor β, and obtain the dynamic programming formulation:

$$\begin{aligned} V_N^\beta(\mathbf{q}) =& \mathbf{q}^T\mathbf{e} + \beta\gamma^{-1}\Big[\lambda_1 V_{N-1}^\beta(A_1\mathbf{q}) + \lambda_2 V_{N-1}^\beta(A_2\mathbf{q}) \\ &+ \min_{\mathbf{r}\in C}\Big\{r_1 V_{N-1}^\beta(D_1\mathbf{q}) + r_2 V_{N-1}^\beta(D_2\mathbf{q}) \\ &+ (a_2 + b_2 - r_1 - r_2) V_{N-1}^\beta(\mathbf{q})\Big\}\Big] \end{aligned} \quad (3)$$

As $N \to +\infty$, $V_N^\beta(\mathbf{q}) \to V^\beta(\mathbf{q})$, which is the unique solution of the optimality equation:

$$\begin{aligned} V^\beta(\mathbf{q}) =& \mathbf{q}^T\mathbf{e} + \beta\gamma^{-1}\Big[\lambda_1 V^\beta(A_1\mathbf{q}) + \lambda_2 V^\beta(A_2\mathbf{q}) \\ &+ \min_{\mathbf{r}\in C}\Big\{r_1 V^\beta(D_1\mathbf{q}) + r_2 V^\beta(D_2\mathbf{q}) \\ &+ (a_2 + b_2 - r_1 - r_2) V^\beta(\mathbf{q})\Big\}\Big] \end{aligned} \quad (4)$$

This is a two-dimensional MDP, which is difficult to solve in general. We first determine some structural properties of the optimal policy.

Lemma 1 *$V^\beta(\mathbf{q})$ is monotonically increasing in q_i, $i = 1, 2$.*

Lemma 2 *The optimal operating point must lie on the boundary of the capacity region. In addition, it must be one of the two corner points.*

Lemma 2 can be proved based on Lemma 1, which in turn can be proved using induction.

Let T be an operator defined on real-valued functions by:

$$\begin{aligned} Tf(\mathbf{q}) =& \mathbf{q}^T\mathbf{e} + \beta\gamma^{-1}\Big[\lambda_1 f(A_1\mathbf{q}) + \lambda_2 f(A_2\mathbf{q}) \\ &+ \min\Big\{a_1 f(D_1\mathbf{q}) + b_1 f(D_2\mathbf{q}) + \delta f(\mathbf{q}), \\ &a_2 f(D_1\mathbf{q}) + b_2 f(D_2\mathbf{q})\Big\}\Big] \end{aligned} \quad (5)$$

Therefore, the dynamic programming optimality equation can be written as

$$V_{N+1}^\beta(\mathbf{q}) = TV_N^\beta(\mathbf{q}) \quad (6)$$

III. An Inductive Proof of the Switch Structure

In this section, we prove that the delay-optimal policy has a switch structure. In order to prove that, we first define a set of functions with properties which are sufficient to have a switch structure. We show that these properties are preserved under the operator T. Since $V_0^\beta = 0$ is within this set, using induction, we will show that V^β will be within this set.

Let us define $\mathcal{F}$ to be the set of real-valued functions such that:

1) $f(\mathbf{q})$ is increasing in q_1 and q_2.
2) $f(\mathbf{q}+\mathbf{x}) - f(\mathbf{q})$ is increasing in q_1 and q_2 for any fixed $\mathbf{x}$.
3) $(a_1-a_2)f(D_1\mathbf{q})+(b_1-b_2)f(D_2\mathbf{q})+\delta f(\mathbf{q})$ is increasing in q_1.
4) $(a_1-a_2)f(D_1\mathbf{q})+(b_1-b_2)f(D_2\mathbf{q})+\delta f(\mathbf{q})$ is decreasing in q_2.

Then, we have the following lemma.

Lemma 3 *If $f \in \mathcal{F}$, then $Tf \in \mathcal{F}$.*

The proof of Lemma 3, when $\delta = 0$, can be found in [13]. When $\delta \neq 0$, the proof is different, but can be carried out similarly. Due to space limitations, the proof is omitted here.

Lemma 4 *$V_n^\beta(\mathbf{q}) \in \mathcal{F}$ for all n.*

This lemma can be verified as follows. Since $V_0^\beta = 0$, V_0^β is in $\mathcal{F}$. Using Lemma 3 recursively, we have $V_n^\beta(\mathbf{q}) \in \mathcal{F}$ for $n = 0, 1, 2, \ldots$. We now define the switch function:

$$\begin{aligned} s_n(q_1) = \min\big\{q_2 :& (a_1 - a_2)f(D_1\mathbf{q}) + (b_1 - b_2)f(D_2\mathbf{q}) \\ &+ \delta f(\mathbf{q}) \leq 0\big\} \end{aligned} \quad (7)$$

Theorem 1 *The optimal policy for the discounted-cost MDP has a switch structure, i.e., $s_n(q_1)$ is increasing for every n.*

This theorem can be proved using properties 3) and 4) of $V_n^\beta(\mathbf{q})$. The switch curve partitions the queue state space into two parts, each corresponding to one of the two operating points (corner points of the pentagon). Following the arguments in [6], [13], we can prove that the switch structure still exists when $\beta \to 1$.

IV. The Limit on the Switch Curve

Although we have shown that the delay optimal policy has a switch structure, it is difficult to obtain the exact switch curve analytically. Numerical techniques, such as value iteration method, can be used to obtain the optimal policy. In this section, we will show that the switch curve is bounded in the q_1-dimension. In other words, we can find a threshold N, such that, for all q_1 greater than this threshold, the optimal operating point is the second corner point of the pentagon. In order to prove this, we start from an initial function f_0, which is linear in $q_1 + q_2$. We will use f_0 to approximate V^β over a large portion of the state space. Specifically, this region includes states $\mathbf{q}$ with $q_1, q_2 > N$, where N is a large enough number. Let us define:

$$f_0(\mathbf{q}) = \frac{1}{1-\beta}(q_1+q_2) + \frac{\beta}{(1-\beta)^2}\frac{\lambda_1+\lambda_2-a_2-b_2}{\lambda_1+\lambda_2+a_2+b_2} \quad (8)$$

Clearly, $f_0 \in \mathcal{F}$. It is easy to verify that

$$Tf_0(\mathbf{q}) - f_0(\mathbf{q}) = \begin{cases} 0 & q_1, q_2 \neq 0 \\ \frac{\beta(a_2+b_2)}{\gamma(1-\beta)} & \mathbf{q} = \mathbf{0} \\ \frac{\beta(a_1+\delta)}{\gamma(1-\beta)} & q_1 = 0 \\ \frac{\beta b_2}{\gamma(1-\beta)} & q_2 = 0 \end{cases} \quad (9)$$

that is, Tf_0 and f_0 differ only on the boundary, and for all states away from the boundary, these two functions have the same value. This is a key property that will be essential in this section. Note that under the operator T, the difference caused by the boundary only propagates into the interior region of the state space by one layer in each iteration; rest of the states are not affected by the operator.

Let us define:

$$|f|_k = \max\{f(\mathbf{q}) : q_1, q_2 \geq 0, q_1+q_2 \leq k\} \quad (10)$$

which is the maximum value of the function f in the region where the sum of the queue lengths is less than k. Similarly, let us define

$$|f|_\infty = \sup\{f(\mathbf{q}) : q_1, q_2 \geq 0\} \quad (11)$$

which is allowed to be infinity. Then, we have the following property.

Lemma 5 *For $\forall f, g \in \mathcal{F}$, $|Tf - Tg|_k \leq \beta|f-g|_{k+1}$.*

Proof:

$$\begin{aligned} &Tf(\mathbf{q}) - Tg(\mathbf{q}) \\ &= \beta\gamma^{-1}\Big[\lambda_1 f(A_1\mathbf{q}) + \lambda_2 f(A_2\mathbf{q}) - \lambda_1 g(A_1\mathbf{q}) - \lambda_2 g(A_2\mathbf{q}) \\ &\quad + \min\Big\{a_1 f(D_1\mathbf{q}) + b_1 f(D_2\mathbf{q}) + \delta f(\mathbf{q}), a_2 f(D_1\mathbf{q}) + b_2 f(D_2\mathbf{q})\Big\} \\ &\quad - \min\Big\{a_1 g(D_1\mathbf{q}) + b_1 g(D_2\mathbf{q}) + \delta g(\mathbf{q}), a_2 g(D_1\mathbf{q}) + b_2 g(D_2\mathbf{q})\Big\}\Big] \end{aligned}$$

Since $|\min\{a,b\} - \min\{c,d\}| \leq \max\{|a-c|, |b-d|\}$, we have

$$\begin{aligned} &|Tf - Tg|_k \\ &\leq \beta\gamma^{-1}\Big[\lambda_1|f-g|_{k+1} + \lambda_2|f-g|_{k+1} && (12) \\ &\quad + \max\Big\{a_1|f-g|_{k-1} + b_1|f-g|_{k-1} + \delta|f-g|_k, \\ &\quad\quad a_2|f-g|_{k-1} + b_2|f-g|_{k-1}\Big\}\Big] \\ &\leq \beta\gamma^{-1}(\lambda_1+\lambda_2+a_2+b_2)|f-g|_{k+1} && (13) \\ &= \beta|f-g|_{k+1} && (14) \end{aligned}$$

completing the proof. ■

Lemma 6 *$T^n f_0$ converges to a function f as $n \to +\infty$, and $Tf = f$.*

Proof: Since $f_0 \in \mathcal{F}$, $T^n f_0 \in \mathcal{F}$ for any $n > 0$.

$$\begin{aligned} |T^{n+1}f_0 - T^n f_0|_k &\leq \beta|T^n f_0 - T^{n-1}f_0|_{k+1} && (15) \\ &\leq \beta^n|Tf_0 - f_0|_{k+n} && (16) \\ &\leq \frac{\beta^{n+1}(a_2+b_2)}{\gamma(1-\beta)} && (17) \end{aligned}$$

where (17) follows from (9). We observe that (17) does not depend on k, thus, $|T^{n+1}f_0 - T^n f_0|_\infty$ is uniformly bounded by (17). Since $\beta < 1$, the right hand side of (17) forms a Cauchy sequence, therefore, $T^n f_0$ converges to a function f pointwise. In other words, for any ϵ, we can find an $N_1(\epsilon)$ such that when $n \geq N_1(\epsilon)$, we have $|f - T^{n-1}f_0|_\infty \leq \epsilon$. Thus, for such n, we have

$$\begin{aligned} |Tf - f|_\infty &\leq |Tf - T^n f_0|_\infty + |T^n f_0 - f|_\infty && (18) \\ &\leq \beta|f - T^{n-1}f_0|_\infty + |T^n f_0 - f|_\infty && (19) \\ &\leq (\beta+1)\epsilon = \epsilon' && (20) \end{aligned}$$

Therefore, for any ϵ', we can find a $n > N_1(\frac{\epsilon'}{\beta+1})$, such that $|Tf - f|_\infty \leq \epsilon'$. In other words, Tf and f are arbitrarily close. Thus, $Tf = f$. ■

Lemma 7 *Let $V_0^\beta(\mathbf{q}) = 0$, then, $V_n^\beta(\mathbf{q}) = T^n V_0^\beta(\mathbf{q})$ converges to $V^\beta(\mathbf{q})$, and $f(\mathbf{q}) = V^\beta(\mathbf{q})$.*

Proof: In order to prove that $f(\mathbf{q}) = V^\beta(\mathbf{q})$ pointwise, we start from the following:

$$\begin{aligned} &|f - V^\beta|_k && (21) \\ &\leq |f - T^n f_0|_k + |T^n f_0 - V_n^\beta|_k + |V_n^\beta - V^\beta|_k && (22) \\ &\leq |f - T^n f_0|_k + \beta|T^{n-1}f_0 - V_{n-1}^\beta|_{k+1} + |V_n^\beta - V^\beta|_k && (23) \\ &\leq |f - T^n f_0|_k + |V_n^\beta - V^\beta|_k + \beta^n|f_0 - V_0^\beta|_{k+n} && (24) \\ &= |f - T^n f_0|_k + |V_n^\beta - V^\beta|_k \\ &\quad + \beta^n\left(\frac{n+k}{1-\beta} + \frac{\beta}{(1-\beta)^2}\frac{\lambda_1+\lambda_2-a_2-b_2}{\lambda_1+\lambda_2+a_2+b_2}\right) && (25) \\ &\leq \epsilon_1 + \epsilon_2 + \epsilon_3 && (26) \end{aligned}$$

where (23) follows from Lemma 5, (25) follows from the definition of f_0, and (26) follows from the fact that $T^n f_0$ converges to f_0, V_n^β converges to V^β, and $\beta^n n \to 0$. Therefore, when n is large enough, we have the difference bounded by (26). We note that (26) does not depend on k, thus $f(\mathbf{q}) = V^\beta(\mathbf{q})$ for any point $\mathbf{q}$. ■

Lemma 5 means that starting from f_0 and performing the iterations, V^β converges to the same function if we started from $V_0^\beta = 0$. The convergence point is the unique solution of the optimality equation (4). Next, we will prove that $f(\mathbf{q})$ gets arbitrarily close to $f_0(\mathbf{q})$ when $q_1, q_2 \to +\infty$.

Lemma 8 $|f - T^n f_0|_\infty \leq \frac{\beta^{n+1}(a_2+b_2)}{\gamma(1-\beta)^2}$.

Proof:

$$\begin{aligned}
&|T^{n+p}f_0 - T^n f_0|_k \\
&\le |T^{n+p}f_0 - T^{n+p-1}f_0|_k + |T^{n+p-1}f_0 - T^{n+p-2}f_0|_k \\
&\quad + \cdots + |T^{n+1}f_0 - T^n f_0|_k \\
&\le \left(\beta^{n+p-1} + \beta^{n+p-2} + \cdots + \beta^n\right)|Tf_0 - f_0|_{k+n+p} \quad (27)\\
&\le \frac{\beta^n(1-\beta^p)}{1-\beta}\frac{\beta(a_2+b_2)}{\gamma(1-\beta)} \quad (28)
\end{aligned}$$

Note that (28) does not depend on k, therefore, $|T^{n+p}f_0 - T^n f_0|_\infty$ is uniformly bounded, and we have

$$\begin{aligned}
|f - T^n f_0|_\infty &= \lim_{p\to\infty}|T^{n+p}f_0 - T^n f_0|_\infty \quad (29)\\
&= \frac{\beta^{n+1}(a_2+b_2)}{\gamma(1-\beta)^2} \quad (30)
\end{aligned}$$

■

Theorem 2 *$f(\mathbf{q})$ gets arbitrarily close to $f_0(\mathbf{q})$ when $q_1, q_2 \to +\infty$. Therefore, the switch curve has a limit on q_1.*

Proof: For any fixed state $\mathbf{q}$, we have

$$|f(\mathbf{q}) - f_0(\mathbf{q})| \le |f(\mathbf{q}) - T^n f_0(\mathbf{q})| + |T^n f_0(\mathbf{q}) - f_0(\mathbf{q})| \quad (31)$$

Based on Lemma 8, we can see that for $\forall\epsilon$, there exists $N(\epsilon)$, such that $|f - T^{N(\epsilon)}f_0|_\infty \le \epsilon$. From the definition in (11),

$$|f(\mathbf{q}) - T^{N(\epsilon)}f_0(\mathbf{q})| \le |f - T^{N(\epsilon)}f_0|_\infty \le \epsilon \quad (32)$$

At the same time, from (9), we know that $T^{N(\epsilon)}f_0(\mathbf{q})$ only differs from $f_0(\mathbf{q})$ over the states which are within $N(\epsilon)$ layers away from the boundary. Thus, for all $q_1 > N(\epsilon), q_2 > N(\epsilon)$,

$$T^{N(\epsilon)}f_0(\mathbf{q}) - f_0(\mathbf{q}) = 0 \quad (33)$$

Therefore, combining (31)-(33), for any $\mathbf{q}$, $q_1 > N(\epsilon), q_2 > N(\epsilon)$, (31) is bounded by

$$|f(\mathbf{q}) - f_0(\mathbf{q})| \le |f - f_0|_\infty + 0 = \epsilon \quad (34)$$

i.e., $-\epsilon \le f(\mathbf{q}) - f_0(\mathbf{q}) \le \epsilon$. Thus, in this region, as shown in Fig. 2, we have

$$\begin{aligned}
&a_1 f(D_1\mathbf{q}) + b_1 f(D_2\mathbf{q}) + \delta f(\mathbf{q}) - a_2 f(D_1\mathbf{q}) - b_2 f(D_2\mathbf{q}) \\
&= (b_1 - b_2)f(D_2\mathbf{q}) + \delta f(\mathbf{q}) - (a_2 - a_1)f(D_1\mathbf{q}) \quad (35)\\
&\ge (b_1 - b_2)\left(f_0(D_2\mathbf{q}) - \epsilon\right) + \delta(f_0(\mathbf{q}) - \epsilon) \\
&\quad - (a_2 - a_1)\left(f_0(D_1\mathbf{q}) + \epsilon\right) \quad (36)\\
&= \frac{\delta}{1-\beta} - 2(a_2 - a_1)\epsilon \quad (37)
\end{aligned}$$

where the inequality follows from (34). Therefore, when

$$\epsilon \le \frac{\delta}{2(a_2 - a_1)(1-\beta)} \quad (38)$$

(37) is always greater than zero, thus point 2 is always better than point 1. From Lemma 8, let

$$\epsilon = \frac{\beta^{n+1}(a_2+b_2)}{\gamma(1-\beta)^2} = \frac{\delta}{2(a_2-a_1)(1-\beta)} \quad (39)$$

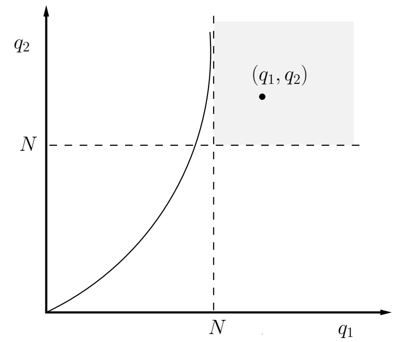

Fig. 2. The switch curve of the discounted-cost MDP.

from which, we have

$$N(\epsilon) = \left\lceil \log_\beta \frac{\delta\gamma(1-\beta)}{2(a_2+b_2)(a_2-a_1)} \right\rceil - 1 \quad (40)$$

Since we have proved in the previous section that the optimal policy must have a switch curve structure, for any $\mathbf{q}$, such that $q_1 \ge N(\epsilon)$, the optimal policy is always to operate the system at point 2. Thus, the switch curve has a limit. ■

The result implies that when both q_1, q_2 are large, the objective of maximizing the sum-rate is more important than balancing the queue lengths in order to minimize the average delay. Thus, in this scenario, operating at point 2 is optimal. When one queue (q_1 in this paper) becomes close to empty, the objective of balancing the queue lengths becomes more important, and the operating point must be switched from point 2 to point 1.

References

[1] A. Ephremides and B. Hajek, "Information theory and communication networks: an unconsummated union," *IEEE Transactions on Information Theory*, vol. 44, pp. 2416–2434, Oct 1998.

[2] I. Bettesh and S. Shamai, "Optimal power and rate control for minimal average delay: The single-user case," *IEEE Transactions on Information Theory*, vol. 52, pp. 4115–4141, Sep 2006.

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

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

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

[6] N. Ehsan and T. Javidi, "Delay optimal transmission policy in a wireless multi-access channel," *IEEE Transactions on Information Theory*, no. 8, pp. 3745–3751, Aug 2008.

[7] E. Yeh, "Minimum delay multi-access communication for general packet length distributions," *Allerton Conf.*, pp. 1536–1545, Sep/Oct 2004.

[8] S. Musy, "Delay and coding in multiple-user communications," Ph.D. dissertation, EPFL, 2007.

[9] J. Yang and S. Ulukus, "Delay minimization in multiple access channels," *International Symposium on Information Theory*, 2009.

[10] V. G. Kulkarni, *Modeling and analysis of stochastic systems*. Chapman and Hall/CRC, 1995.

[11] J. Walrand, *An introduction to queueing networks*. Prentice Hall, 1988.

[12] W. Lin and P. R. Kumar, "Optimal control of a queueing system with two heterogenous servers," *IEEE Transactions on Automatic Control*, vol. 29, pp. 696–703, Aug 1984.

[13] B. Hajek, "Optimal control of two interacting service stations," *IEEE Transactions on Automatic Control*, vol. 29, no. 6, pp. 491–499, June 1984.