Stochastic Optimal Controls for Parallel-Server Channels with Zero Waiting Buffer Capacity and Multi-Class Customers

Wanyang Dai[1] and Youyi Feng[2]

[1]*Department of Mathematics and State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China. Email: nan5lu8@netra.nju.edu.cn*

[2]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong*

Abstract *We study the stochastic optimal admission and routing controls for a system consisting of two parallel-server channels with zero waiting buffer capacity and multi-class customers via Markov decision processes moving in discrete triangles. In each channel, there are two classes of customers to be served, which are varying in their arrival rates, payments and penalty costs. The class of structured value functions is found and justified to guarantee the existence of optimal value function and optimal admission control policy. The policy for each class of customers is characterized by a state-dependent threshold function that is non-increasing in the number of the other class of customers being served in the channel and decreases at most one unit when the number of the other class of customers increases a unit. Between channels, some classes of customers can be selected to be served in either of the channels according to certain routing probabilities, which are obtained by combining the optimal value functions with simulation. Numerical and application examples are presented to illustrate the usage and efficiency of our algorithm and optimal policy.*

Keywords *Dynamic programming Optimal control Markov decision process State-dependent threshold policy Queue Optimization Parallel-server system Zero waiting buffer capacity Multi-class customer*

1. Introduction

Parallel-server systems have become a common tool in the studies of many communication networks, manufacturing and service systems, see, for examples, Dai[4] and Gans et al.[7]. Motivated from some practical communication project, in this paper, we consider a system that consists of two parallel-server channels with zero waiting buffer capacity, which differs from widely studied parallel-server systems with infinite waiting buffer capacity (see, e.g., Harrison[8], Bell and Williams[1]). In each channel, there are two classes of customers to be served, which are varying in their arrival rates, payments and penalty costs. Payments are random and depend on the lengths of service times, which are different from assumptions in existing studies such as Ormeci et al[12] and Ormeci and Wal[13], where constant payments are assumed. Between channels, some classes of customers can be selected to be served in either of the channels according to certain routing probabilities. Since the waiting buffer capacity is zero for both channels, optimal admission and routing controls for such a system are of particular interests. The purpose of the controller is to determine how the system should route an incoming customer to a particular channel and whether the channel should accept or reject the customer into service in order to maximize the expected infinite horizon discounted profit for the service provider. We will restrict our study to Markovian control policies since the optimal policy belongs to this class (see, e.g., Sennott[16]). Some related discussions about network optimization and control can be found in Cheng et al[3], Dai and Jiang[5], Ding et al[6] and Ma et al[11].

Optimal control problems for queueing systems arise in different contexts such as Lippman[10]. The unified approach developed in [10] allows one to transform a continuous time

Markov decision process into an equivalent discrete process and to establish associated optimality equations. Moreover, following the approach, one can investigate the structure of an optimal control policy by identifying a class of structured value functions that is preserved under certain sense and find a state-dependent threshold policy that is easy to implement, see, e.g., Benjaafar and ElHafsi[2], and etc. for a number of recent applications in production and inventory controls. However, the classes of structured value functions are varying for different applications and they are difficult to be identified for the corresponding decision processes moving in multi-dimensional domains. In addition, the justification of preservation of properties for each application is complicated and involves heavy calculation.

In this paper, firstly, we adopt the approach to study the optimal admission control problem for the above parallel-server system in each channel. The class of structured value functions is found to satisfy the concave and certain submodular properties and the preservation of these properties under some policy related transformation is justified to guarantee the existence of optimal value function and optimal admission control policy. The obtained optimal policy for each class of customers is characterized by a state-dependent threshold function that is non-increasing in the number of the other class of customers being served in the channel and decreases at most one unit when the number of the other class of customers increases a unit. Because the waiting buffer capacity is zero and there are multiple classes of customers in the system, our Markov decision process is confined to move within a 2-dimensional discrete triangle that is not a good shape to our study. Due to this reason, our function class is different from those such as in [2] and the justification of the structure preservation for our problem is particular complex along the boundary of the triangle, moreover, due to the difference of model formulations, our admission policy and associated analysis are different from those as in [12] and [13]. Secondly, by combining the optimal value functions in different channels with simulation, we obtain the optimal routing parameters. Thirdly, we implement a number of numerical examples to support our theoretical findings. From these examples, we see that our algorithm stably converges and is efficient, and furthermore, our policy is better than some intuitively reasonable policy that is based on two 1-dimensional Markov decision processes and is designed for the purpose of comparison.

The rest of the paper is organized as follows. In Section 2, we describe our model formulation. In Section 3, we derive the optimality equations and study the characterization of the optimal policy. Numerical examples are given in Section 4. In Section 5, we provide the technical and lengthy proofs for three lemmas. Concluding remarks and future research are presented in Section 6.

2. Model Formulation

We consider a communication system consisting of two channels as pictured in the following Figure 1. There are n independent and identical servers in Channel A and $\bar{n}$ servers in Channel $\bar{A}$. Both channels have no additional buffering capability for an incoming customer except being served immediately. There are two classes of external arriving customers to Channel A and the arrival stream for each class $i \in \{1, 2\}$ follows a Poisson process with rate l_i. Similarly, there are two classes of external arriving customers to Channel $\bar{A}$ and the arrival stream for each class $i \in \{1, 2\}$ follows a Poisson process with rate $\bar{l}_i$. For each $i \in \{1, \mathbf{K}, m\}$ with some given integer $1 \leq m \leq 2$, the arrival streams corresponding to l_i and $\bar{l}_i$ are split from a common Poisson process with rate b_i, that is, $l_i + \bar{l}_i = b_i$, and moreover, it will be discussed later about how to determine the optimal values of l_i and $\bar{l}_i$ for a given b_i according to routing decision. For each class $i \in \{1, 2\}$, an arrival customer to Channel A is either rejected into service with penalty cost l_i or accepted into service and experiences exponentially distributed amount of service time with rate m according to some admission control policy, where we assume that service times among different customers are independent. Similarly, for each class $i \in \{1,2\}$, an arrival customer to Channel $\overline{A}$ is either rejected into service with penalty cost $\overline{l_i}$ or accepted into service and experiences exponentially distributed amount of service time with rate $\overline{m}$. Furthermore, for each $i \in \{1, \mathbf{K}, m\}$, we suppose that $l_i = \overline{l_i}$.

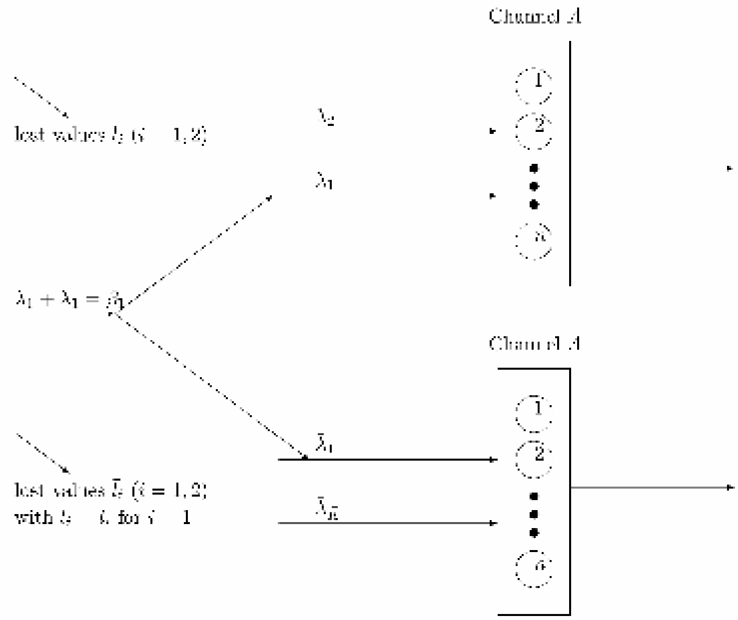

Figure 1 A communication system consisting of two channels with parallel servers is under admission and routing controls.

For each $i \in \{1,2\}$, let $X_i(t)$ denote the number of class *i* customers being served in Channel *A* at time *t*, which takes values in the state space $S_i = \{0,1,\mathbf{K},n\}$, and moreover, let $X(t) \equiv (X_1(t), X_2(t))$, which takes values in the following 2-dimensional discrete triangle

$$S \equiv \left\{ (x_1, x_2) : \sum_{i=1}^{2} x_i \le n, x_i \in S_i, i = 1, 2 \right\}. \tag{1}$$

Let $h(X(t))$ denote the payment function defined on *S*, which satisfies certain conditions that will be elaborated later. Associated with Channel *A*, there is a fixed cost *c* per unit of time, and let $N_i(t)$ denote the total number of rejected class *i* customers at Channel *A* by time *t*. Then we can write the value function (expected infinite horizon discounted profit) for a given discount factor a, a given initial state $x \in S$ and a given admission control policy p as follows,

$$v^p(x) \equiv E_x^p \left[\int_0^\infty e^{-at} h(X(t))dt - \int_0^\infty e^{-at} \sum_{i=1}^{2} l_i dN_i(t) - \int_0^\infty e^{-at} c dt \right]. \tag{2}$$

Similarly, for each $i \in \{1,2\}$, let $\overline{X}_i(t)$ denote the number of class *i* customers being served in Channel $\overline{A}$ at time *t*, which takes values in the state space $\overline{S}_i = \{0,1,\mathbf{K},\overline{n}\}$, and moreover, let $\bar{X}(t)=(\bar{X}_1(t),\bar{X}_2(t))$, which takes values in the following 2-dimensional discrete triangle

$$S\equiv\left\{(x_1,x_2):\sum_{i=1}^{2}x_i\le\bar{n},x_i\in\bar{S}_i,i=1,2\right\}. \quad (3)$$

Let $\bar{h}(\bar{X}(t))$ denote the payment function defined on $\bar{S}$. Associated with Channel $\bar{A}$, there is a fixed cost $\bar{c}$ per unit of time, and let $\bar{N}_i(t)$ denote the total number of rejected class *i* customers at Channel $\bar{A}$ by time *t*. Then we can write the value function for the given discount factor a, a given initial state $\bar{x}\notin\bar{S}$ and a given admission control policy $\bar{p}$ as follows,

$$\bar{v}^{\bar{p}}(\bar{x})\equiv E_{\bar{x}}^{\bar{p}}\left[\int_0^\infty e^{-at}\bar{h}(\bar{X}(t))dt-\int_0^\infty e^{-at}\sum_{i=1}^{2}\bar{l}_i d\bar{N}_i(t)-\int_0^\infty e^{-at}\bar{c}dt\right]. \quad (4)$$

An admission control policy p (or $\bar{p}$) specifies at each time instant if an arrived customer to Channel *A* (or Channel $\bar{A}$) is accepted into service or is rejected, and furthermore, we restrict our analysis to Markovian policies since the optimal policy belongs to this class (e.g., Sennott[16]). Our objective is to look for optimal policies p^* and $\bar{p}^*$ such that

$$v^{p^*}(x,l)=\max_{p\in\Pi}v^p(x,l),\quad \bar{v}^{\bar{p}^*}(\bar{x},\bar{l})=\max_{\bar{p}\in\bar{\Pi}}\bar{v}^{\bar{p}}(\bar{x},\bar{l}) \quad (5)$$

where $l=(l_1,\mathbf{K},l_m),\bar{l}=(\bar{l}_1,\mathbf{K},\bar{l}_m)$ with $m\in\{1,2\}$, and the sets Π and $\bar{\Pi}$ of Markov decision policies are associated with some action spaces which will be elaborated later. Moreover, we are to find the optimal values l^* and $\bar{l}^*$, for a given vector $b=(b_1,\mathbf{K},b_m)$ such that

$$v^{p^*}(x,l^*)+\bar{v}^{\bar{p}^*}(\bar{x},\bar{l}^*)=\max_{(l,\bar{l})\in\Lambda}\left(v^{p^*}(x,l)+\bar{v}^{\bar{p}^*}(\bar{x},\bar{l})\right) \quad (6)$$

where the set Λ is defined to be $\Lambda\equiv\{(l,\bar{l}):l_i+\bar{l}_i=b_i,i=1,\mathbf{K},m\}$. Moreover, we remark that the determinations of l^* and $\bar{l}^*$ are essentially the choices of the optimal routing probabilities to two channels for the common Poisson input streams corresponding to the rate vector b.

3. The Optimality Equations and Characterization of the Optimal Policy

We first focus our discussions on Channel *A*. Let $h^p(x)$ be the control associated with a Markovian policy p when the system is in state $x\in S$. It follows from Lemma 5.1 in Section 5 that there is at most one customer arriving at the channel at a time point. Thus we take $h^p(x)=h_i^p(x)$ when the arrived customer belongs to class *i*, where $h_i^p(x)$ corresponds to the admission action such that

$$h_i^p(x)\equiv\begin{cases}0 & \text{if the action is to accept the arrived class } i \text{ customer into service,}\\ 1 & \text{if the action is to reject the arrived class } i \text{ customer into service.}\end{cases}$$

Thus the action space can be taken as {0, 1} that is independent of state variable *x*. Let

$$g\equiv\sum_{i=1}^{2}l_i+nm=\sup_{(x,h(x))\in S\times\{0,1\}}l(x,h(x)). \quad (7)$$

where $h(x)$ is an action and $l(x,h(x))$ is the rate of changing of the system when it is in state *x*. Then by the studies in Lippman[10], we can transform the continuous time decision process into an equivalent discrete time decision process. Moreover, define $v^*(x) \equiv v^{p^*}(x)$ and without loss of generality, suppose that $a+g=1$, then we have the following lemma.

Lemma 3.1 For each $x \in S$, the following optimality equation holds

$$v^*(x) = Tv^*(x) \tag{8}$$

where the operator T is defined as follows,

$$Tv(x) = (h(x)-c) + \sum_{i=1}^{2} l_i T_i v(x) + m\left(\sum_{i=1}^{2} x_i v(x-e_i) + \left(n - \sum_{i=1}^{2} x_i\right)v(x)\right),$$

e_i is the 2-dimensional unit vector whose ith component takes value one and all other components take values 0, and furthermore,

$$T_i v(x) \equiv \begin{cases} \max\{v(x+e_i), v(x)-l_i\} & if \sum_{i=1}^{2} x_i < n, \\ v(x)-l_i & if \sum_{i=1}^{2} x_i = n. \end{cases}$$

The proof of this lemma can be found in Section 5. To characterize the optimal policy, for each $i \in \{1,2\}$, we define

$$\Delta_i v(x) \equiv v(x+e_i) - v(x). \tag{9}$$

Next, suppose that $v(x)$ is a function satisfying the following conditions, that is, for each fixed $i \in \{1,2\}$ and a $j \in \{1,2\}$,

$$\Delta_i \Delta_i v(x) \equiv \Delta_i v(x+e_i) - \Delta_i v(x) \le 0, \tag{10}$$

$$\Delta_j \Delta_i v(x) \equiv \Delta_i v(x+e_j) - \Delta_i v(x) \le 0 \quad for\ any\ j \ne i, \tag{11}$$

$$\Delta_{ij} \Delta_i v(x) \equiv \Delta_i v(x+e_i) - \Delta_i v(x+e_j) \le 0 \quad for\ any\ j \ne i. \tag{12}$$

Remark 3.1 Condition (10) indicates that $\Delta_i v$ is non-increasing in the component x_i or equivalently v is concave in x_i; Condition (11) states that v is submodular in the direction of e_i and e_j ; We call that v is submodular in the direction of e_i and $e_i - e_j$ under condition (12) since this condition is equivalent to $\Delta_i v(x+e_i-e_j) - \Delta_i v(x) \le 0\ for\ 0 < x \in S.$ When $\sum_{j=1, j\ne i}^{2} x_j = n-1$, the concavity property of v in terms of x_i is always true since there are only two points involved; When one dimensional problem is concerned, the only required condition is the concavity.

The following examples are given to show the existence of the class of functions that satisfy conditions (10)-(12).

Example 3.1 $v(x) = \sum_{i=1}^{2} f_i(x_i)\ for\ x \in S$, where $f_i(x_i)$ is a concave function of x_i for each $i \in \{1,2\}$. For examples, (1) v(x) is linear in x, that is, $v(x) = \sum_{i=1}^{2} r_i x_i$ where $r_i > 0$ is a constant for each i; (2) $v(x) = \sum_{i=1}^{2} (a_i - b_i g_i(x_i))$ where a_i and b_i are positive constants, $g_i(x_i)$ is convex in x_i, e.g., $g_i(x_i) = e^{\pm r_i x_i}$ for some constant $r_i > 0$, or $g_i(x_i) = (x_i - r_i)^2$ and etc.

Example 3.2 Let v be a concave function on R and define $f(x) \equiv \sum_{i=1}^{2} x_i$. Notice that f is an affine mapping from R^2 to R (see, e.g., page 23 of Hiriart-Urruty and Lemarechal[9]). Then it is easy to check that *v(f(x))* satisfies condition (12). Moreover, the conditions (10) and (11) are reduced to the only concavity condition (10) and this property can be easily shown by the similar method used in Proposition 2.1.4 in page 88 of [9].

Here we remark that more such examples can be constructed, for example, by applying Proposition 2.1.7 in page 88 of [9]. So, basing on the above discussions, we can introduce the following set of functions

$$A \equiv \{v(x) : v(x) \text{ satisfies the conditions } (10)-(12), x \in S\}. \qquad (13)$$

If associated with the value function of a policy, the conditions (10)-(12) imply some particular properties for that policy. Concretely, for each $i \in \{1,2\}$ and each $x \in S$ with $\sum_{i=1}^{2} x_i \le n-1$, let $\hat{x}_i = (x_j, j \ne i)$, moreover let $\hat{x}_i$ fixed and x_i vary, then for each $v \in A$, we can define

$$a_i(\hat{x}_i) \equiv \max\{x_i : \Delta_i v(x) + l_i > 0\}. \qquad (14)$$

From (14), we see that $a_i(\hat{x}_i)$ depends only on $\hat{x}_i$ as long as v and l_i is given. In the sequel, we also use $a_i(x)$ to denote $a_i(\hat{x}_i)$ for convenience. Then we have the following lemma.

Lemma 3.2 If $v \in A$, then $x_i \le a_i(\hat{x}_i)$ if and only if $\Delta_i v(x) + l_i > 0$.

Proof. If $x_i \le a_i(\hat{x}_i)$, then by the definition of $a_i(\hat{x}_i)$ in (14), we have that $\Delta_i v(x) + l_i > 0$. Conversely, let $x = (x_i, \hat{x}_i) \in S$ satisfying $\Delta_i v(x) + l_i > 0$ (here we take $i = 1$ without loss of generality). Due to the concavity property (10) of v, we have that $\Delta_i v(x') + l_i \ge \Delta_i v(x) + l_i > 0$ with $x' = (x_i', \hat{x}_i)$ and $x_i' \le x_i$. Then by the definition given in (14), we can conclude that $x_i \le a_i(\hat{x}_i)$. □

Remark 3.2 Lemma 3.2 implies that the admission control associated with a policy p that has value function $v^p \in A$ can be described with 2-dimensional switching surfaces. More specifically, the policy can be expressed by state-dependent thresholds $a_i(x)$ for all $i \in \{1,2\}$ such that it is optimal to accept a class i customer into service if $x_i \le a_i(x)$ (which means that accepting one is to get the bigger profit than rejecting one) and otherwise to reject.

Lemma 3.3 For a fixed $i \in \{1,2\}$, a $j \in \{1,2\}$ $(j \ne i)$, a $v \in A$ and a $x \in S$ with $\sum_{i=1}^{2} x_i < n-1$, we have

$$a_i(x+e_j) \equiv \begin{cases} a_i(x), \\ a_i(x)-1. \end{cases}$$

Proof. Without loss of generality, we take i = 1. Then let $\tilde{x} = (x_i, a_i(x+e_j))$. Due to the submodular property (11) of v in the direction of e_i and e_j and the definition of $a_i(x+e_j)$, we have that

$$\Delta_i v(\tilde{x}) \ge \Delta_i v(\tilde{x}+e_j) > -l_i \qquad (15)$$

Thus, by (15) and the definition of $a_i(\tilde{x})$, we can conclude that

$$a_i(x+e_j) = a_i(\tilde{x}+e_j) \le a_i(\tilde{x}) = a_i(x).$$

Secondly, if $a_i(x)$ = 0, it is obvious that $a_i(x+e_j) > a_i(x)-1$. If $a_i(x) \ge 1$, we take $\bar{x} = (x_i, a_i(x))$, then by using $\bar{x} - e_i$ to replace x in condition (12), we have

$$\Delta_i v(\bar{x} - e_i + e_j) \ge \Delta_i v(\bar{x}) > -l_i.$$

Hence we have

$$a_i(x)-1=a_i(\tilde{x})-1\le a_i(\tilde{x}+e_j)=a_i(x+e_j).$$

Therefore it follows from the above discussions that the lemma is true. □

The following lemma shows that *T* defined in Theorem 3.1 preserves the properties of *v* and hence preserves the switching surface structure of the associated control policy.

Lemma 3.4 If $h,v\in A$, then $Tv\in A$, where the operator *T* is defined in Proposition 3.1 and

$$T_i v(x)\equiv\begin{cases}v(x+e_i) & if\ \sum_{i=1}^{2}x_i<n, x_i\le a_i(x),\\ v(x)-l_i & if\ \sum_{i=1}^{2}x_i<n, x_i>a_i(x),\\ v(x)-l_i & if\ \sum_{i=1}^{2}x_i=n.\end{cases}$$

The lengthy proof of the above lemma will be given in Section 5. Instead, we present our main result as follows.

Theorem 3.1 There is a unique optimal value function v^* which belongs to *A* if $h\in A$. Consequently, the optimal policy is a threshold admission control policy with state-dependent thresholds for each class. Specifically, the policy has the following properties.

1. For each $\hat{x}_i=(x_j, j\ne i)$, there is a corresponding threshold $a_i^*(\hat{x}_i)$ such that it is optimal to accept a class *i* customer into service if $x_i\le a_i^*(\hat{x}_i)$ and to reject otherwise.
2. The threshold $a_i^*(\hat{x}_i)$ is non-increasing in each of the variables $x_j(j\ne i)$ with $a_i^*(x)-1\le a_i^*(x+e_j)\le a_i^*(x)$ where one can use the same explanation in (14) for $a_i^*(x)$.

Proof. It follows from Lemma 3.4, Theorem 5.1 of Porteus[14] and Theorem 6.10.4 of Puterman[15] that $v^*=\lim_{m\to\infty}T^m v\in A$ and it is the unique solution of *v* = *Tv*, where T^m refers to *m* compositions of operator *T*. Therefore, v^* satisfies conditions (10)-(12). Hence the remain results of the theorem follow from Lemmas 3.2 and 3.3. □

4. Numerical Examples

In this section, we use numerical examples to illustrate the usage and efficiency of our optimal threshold policy. Firstly, we present several examples with different payment functions to show the dynamic evolving of the optimal thresholds in terms of the number of the other class of customers in a channel. Secondly, we conduct the numerical comparison between our optimal value function and another intuitively reasonable one to exhibit the optimality of our policy. Thirdly, we obtain optimal routing parameters by integrating our optimal policy with simulation. Fourthly, we extend our system to a more general one where different classes of customers may have different service rates and we show that our optimal policy is still reasonably good by a numerical example. All of the numerical results given here are obtained by solving the dynamic programs iteratively and the value iteration algorithm is run enough times in order that at least four-digit accuracy is achieved for each problem instance.

Example 4.1 Here we take the payment function *h(x)* to be a linear function of *x*, that is, for some constants q_1 and q_2, $h(x) = q_1x_1 + q_2x_2$. The corresponding numerical results about the evolving of thresholds are described in Figure 2 and Figure 3.

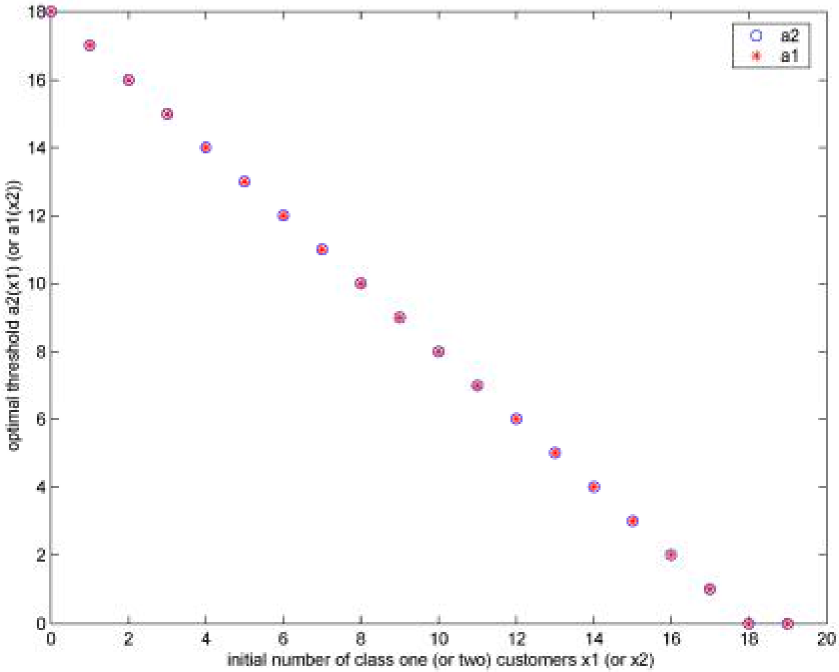

Figure 2 Optimal thresholds with n = 19, a = 0.1, l_1 = 0.15, l_2 = 0.25, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q_1 = 1000 and q_2 = 2000.

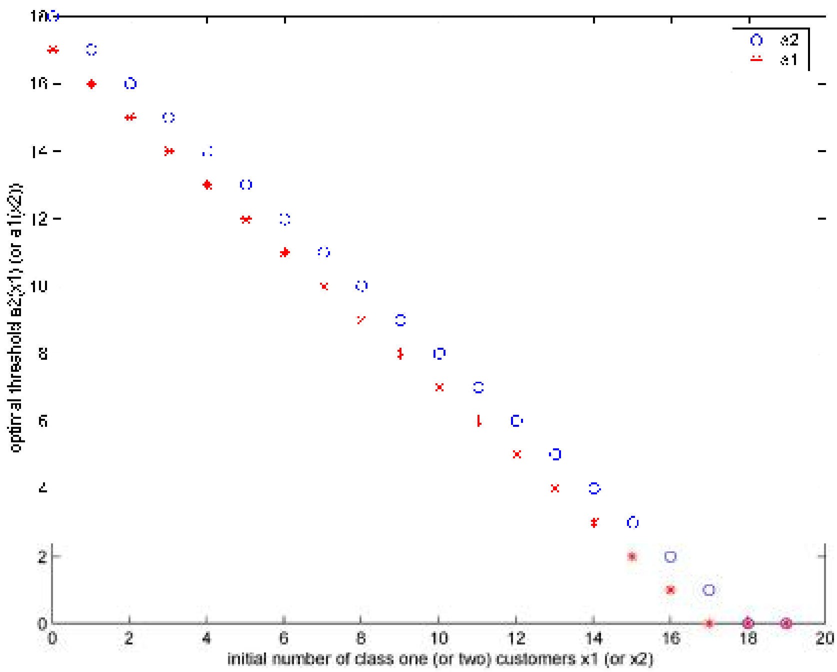

Figure 3 Optimal thresholds with n = 19, a = 0.1, l_1 = 0.03, l_2 = 0.37, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q_1 = 1000 and q_2 = 2000.

Example 4.2 For some constant q, we take the payment function $h(x) = q(2500 - 6(x_1 - 4)^2 - 10(x_2 - 11)^2)$. The corresponding numerical results about the evolving of thresholds are shown in Figure 4.

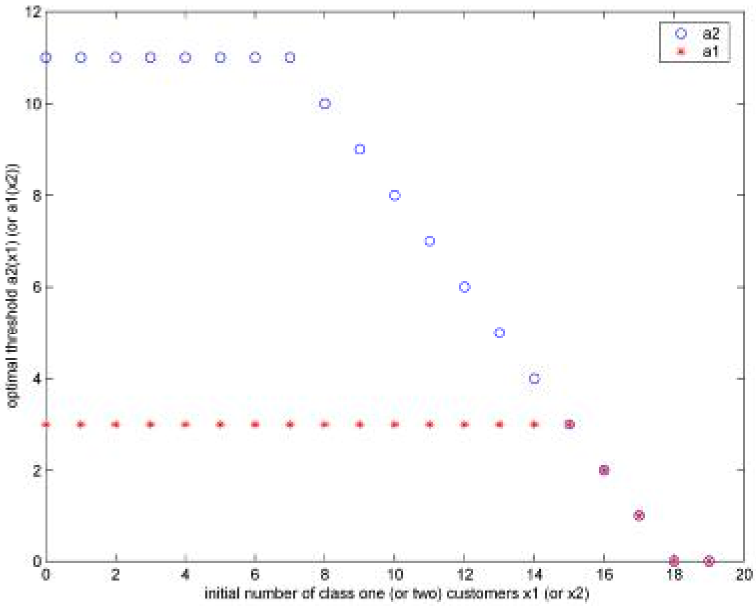

Figure 4 Optimal thresholds with n = 19, a = 0.1, l_1 = 0.15, l_2 = 0.25, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000 and q = 2000.

Example 4.3 For some constants q, p_1 and p_2, we take the payment function $h(x) = q(2 - e^{-p_1 x_1} - e^{-p_2 x_2})$. The corresponding numerical results about the evolving of thresholds are displayed in Figure 5.

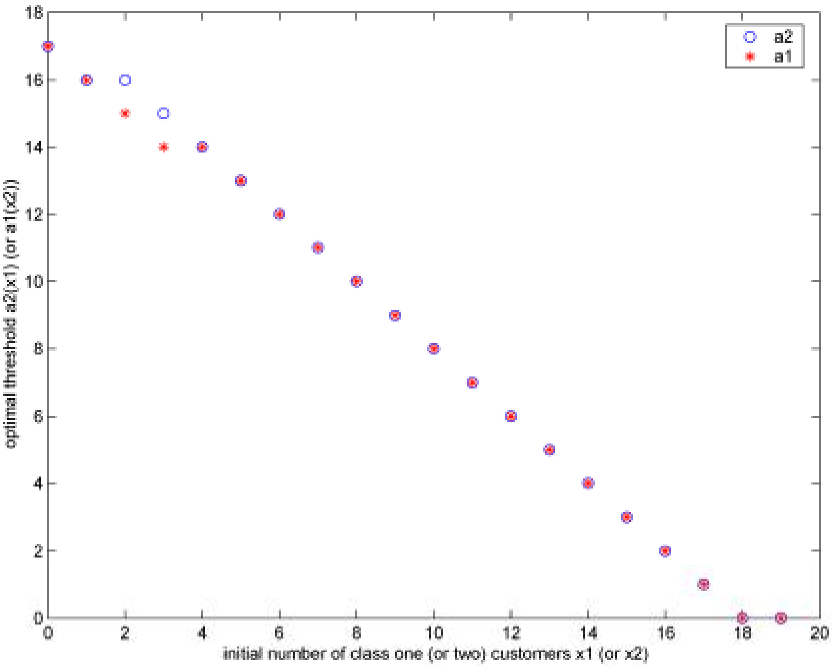

Figure 5 Optimal thresholds with n = 19, a = 0.1, l_1 = 0.15, l_2 = 0.25, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q = 2000, p_1 = 0.1 and p_2 = 0.3.

Example 4.4 In this example, firstly, we take the payment function $h(x) = q_1x_1 + q_2x_2$ for some constants q_1 and q_2 and use our algorithm to obtain the values of the optimal value function defined in (8) at all possible initial states. Secondly, we separate the servers into two groups, each is assigned to serve only for a particular customer class. Then we use two 1-dimensional Markov decision processes on a finite set to generate a 2-dimensional optimal value function. Concretely, suppose that there are b number of servers assigned for class one customers and $(n-b)$ number of servers assigned for class two customers. Then the iteration algorithms for the associated 1-dimensional Markov decision processes are given by

$$T^1 v_1(x_1) = \frac{1}{a + g_1}\Big(h_1(x_1) - c_1 + l_1 T_1^1 v_1(x_1) + m(x_1 v_1(x_1 - 1) + (b - x_1) v_1(x_1))\Big)$$

$$T^2 v_2(x_2) = \frac{1}{a + g_2}\Big(h_2(x_2) - c_2 + l_2 T_2^2 v_2(x_2) + m(x_2 v_2(x_2 - 1) + (n - b - x_2) v_2(x_2))\Big)$$

where $g_1 = l_1 + bm$ and $g_2 = l_2 + (n-b)m$, $h_1(x_1) = q_1 x_1$ and $h_2(x_2) = q_2 x_2$ for some constants q_1 and q_2, $c_1 = cb/n$ and $c_2 = c - c_1$, moreover,

$$T_1^1 v_1(x_1) = \begin{cases} v_1(x_1+1) & if\ x_1 < b, x_1 \le a_1, \\ v_1(x_1) - l_1 & if\ x_1 < b, x_1 > a_1, \\ v_1(x_1) - l_1 & if\ x_1 = b, \end{cases}$$

$$T_2^2 v_2(x_2) = \begin{cases} v_2(x_2+1) & if\ x_2 < n-b, x_2 \le a_2, \\ v_2(x_2) - l_2 & if\ x_2 < b, x_2 > a_2, \\ v_2(x_2) - l_2 & if\ x_2 = n-b \end{cases}$$

and $a_i = \max\{x_i : \Delta_i v_i(x_i) + l_i > 0\}$ with $\Delta_i v_i(x_i) = v_i(x_i+1) - v_i(x_i)$ for $i = 1, 2$. Then we can obtain the value differences $v^*(x_1, x_2) - (v_1^*(x_1) + v_2^*(x_2))$ for $0 \le x_1 \le b$ and $0 \le x_2 \le n-b$. In our numerical implementation, for any $b \in \{0, 1, \mathbf{K}, n\}$, the corresponding value differences are positive and see for one particular example given in Figure 6 where $b = 3$.

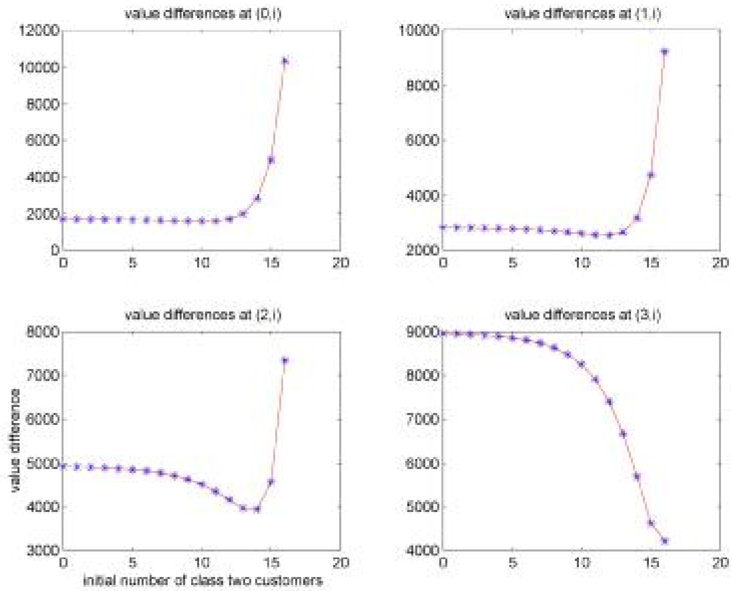

Figure 6 Comparison of value functions with $n = 19$, $b = 3$, $a = 0.1$, $l_1 = 0.15$, $l_2 = 0.25$, $m = 0.5/19$, l_1

= 200, l_2 = 400, c = 1000, q_1 = 1000, q_2 = 2000.

Example 4.5 In this example, we consider two channels A and $\overline{A}$. Note that all notations related to Channel $\overline{A}$ will be put a bar above the corresponding notations for Channel A. The payment functions $h(x)$ and $\overline{h}(\overline{x})$ are taken to be linear in terms of x and $\overline{x}$ respectively, i.e., $h(x) = q_1x_1 + q_2x_2$ and $\overline{h}(\overline{x}) = \overline{q}_1\overline{x}_1 + \overline{q}_2\overline{x}_2$ for some constants q_1, q_2, $\overline{q}_1$ and $\overline{q}_2$. Basing on the iterative algorithm given in Proposition 3.1 for Markov decision process, we add one more simulation step to obtain the optimal routing parameters l_1 and $\overline{l}_1$ that satisfy $l_1 + \overline{l}_1 = 0.3$ and solve the optimization problem (6). In the simulation, we take $l_1 = 0.3(i/N)$ for $i = 0,1,\ldots,N$. Firstly, we handle the case that Channel $\overline{A}$ is configured exactly the same as Channel A with n = $\overline{n}$ = 6. Due to the symmetry property, the optimal routing parameters should coincide for the initial states (i,2,2,2) and (2,2,i,2) with $i = 0,1,\ldots 4$, which are shown in Figure 7 and Figure 8. In Figure 8, we see some difference between optimal parameters at points (1, 2, 2, 2) and (2, 2, 1, 2). This phenomenon can be interpreted as follows: when N takes odd number (for example N = 49), l_1 can never take the real optimal value 0.15, and moreover, due to the fact that, $l_1 + \overline{l}_1 = 0.3$, the difference appears. Secondly, we handle the asymmetry case with n = 7 and $\overline{n}$ = 5. The associated numerical results about the optimal routing parameters are given in Figure 9.

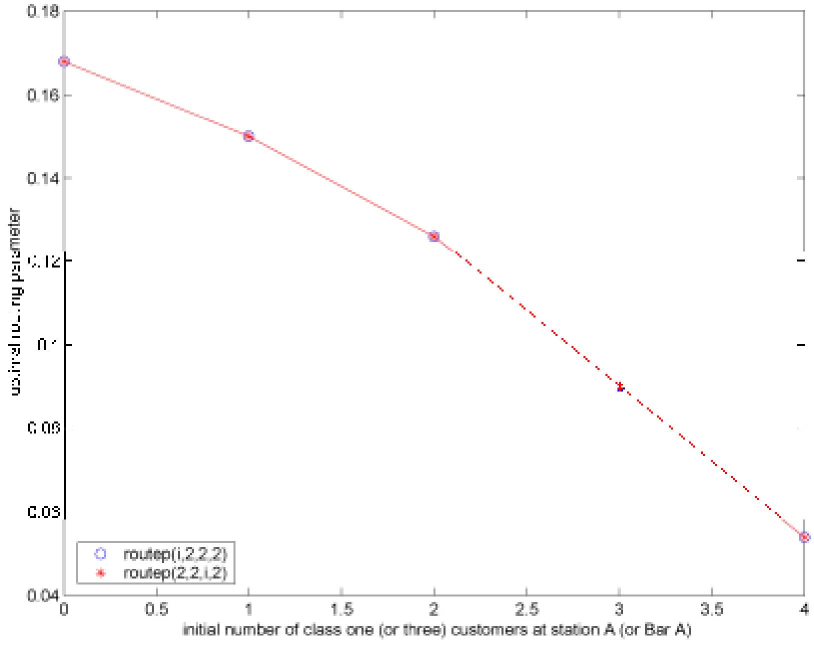

Figure 7 Optimal routing parameters with n = 6, a = 0.1, l_1 = 0.15, l_2 = 0.25, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q_1 = 1000, q_2 = 2000; $\overline{n}$ = 6, $\overline{a}$ = 0.1, $\overline{l}_1$ = 0.15, $\overline{l}_2$ = 0.25, $\overline{m}$ = 0.5/19, $\overline{l}_1$ = 200, $\overline{l}_2$ = 400, $\overline{c}$ = 1000, $\overline{q}_1$ = 1000, $\overline{q}_2$ = 2000, N = 50.

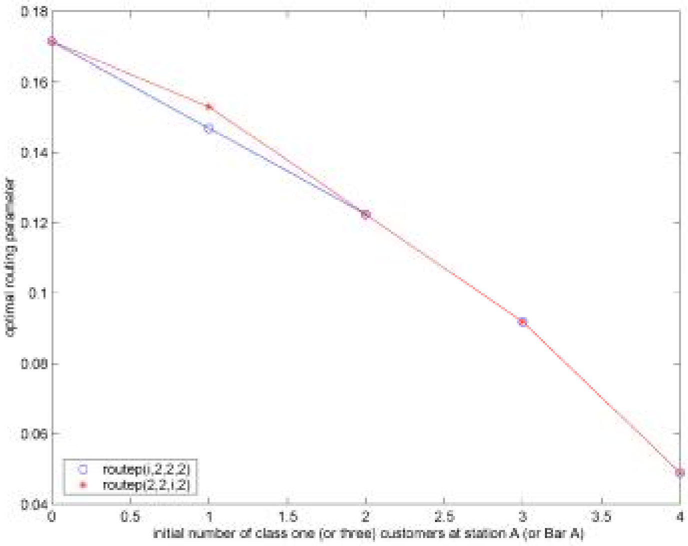

Figure 8 Optimal routing parameters with n = 6, a = 0.1, l_1 = 0.15, l_2 = 0.25, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q_1 = 1000, q_2 = 2000; $\overline{n}$ = 6, $\overline{a}$ = 0.1, $\overline{l}_1$ = 0.15, $\overline{l}_2$ = 0.25, $\overline{m}$ = 0.5/19, $\overline{l}_1$ = 200, $\overline{l}_2$ = 400, $\overline{c}$ = 1000, $\overline{q}_1$ = 1000, $\overline{q}_2$ = 2000, N = 49.

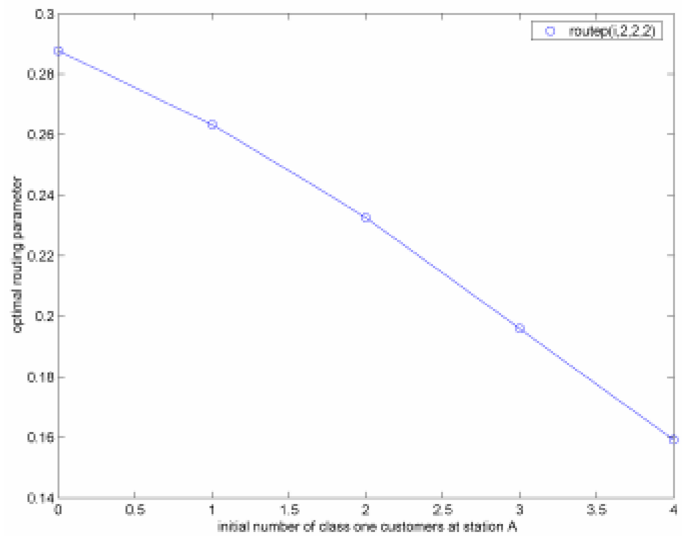

Figure 9 Optimal routing parameters with n = 7, a = 0.1, l_1 = 0.15, l_2 = 0.25, m = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q_1 = 1000, q_2 = 2000; $\overline{n}$ = 5, $\overline{a}$ = 0.1, $\overline{l}_1$ = 0.15, $\overline{l}_2$ = 0.25, $\overline{m}$ = 0.5/19, $\overline{l}_1$ = 200, $\overline{l}_2$ = 400, $\overline{c}$ = 1000, $\overline{q}_1$ = 1000, $\overline{q}_2$ = 2000, N = 49.

Example 4.6 In this example, we extend the iterative algorithm presented in Proposition 3.1 to allow different service rates $m_i (i = 1,2)$ for different customer classes. Concretely, the new algorithm can be designed as follows,

$$Tv(x) = (h(x) - c) + \sum_{i=1}^{2} l_i T_i v(x) + \sum_{i=1}^{2} m_i x_i v(x - e_i) + \left(n \max\{m_1, m_2\} - \sum_{i=1}^{2} m_i x_i \right) v(x)$$

where the threshold policy related to $T_i v(x)$ for i = 1,2 is the same as the one given in

Proposition 3.4. In our numerical implementation, we see that the new algorithm still converges and the limiting thresholds keep the properties as stated in Theorem 3.1, see for an example, numerical results in Figure 10 where $h(x)=q(2-e^{-p_1x_1}-e^{-p_2x_2})$ for some constants q, p_1 and p_2. Moreover, let $v^*(x_1,x_2)$ denote the limiting value function of our new algorithm. Then as in Example 4.4, we can obtain the value differences $v^*(x_1,x_2)-(v_1^*(x_1)+v_2^*(x_2))$ for $0\le x_1\le b$ and $0\le x_2\le n-b$, and for any $b\in\{0,1,\ldots,n\}$, the corresponding value differences are positive as shown, for a particular example, in Figure 11. From these numerical results, we reach a reasonable conjecture that the limiting value function $v^*(x)$ is also an optimal value function in certain function class.

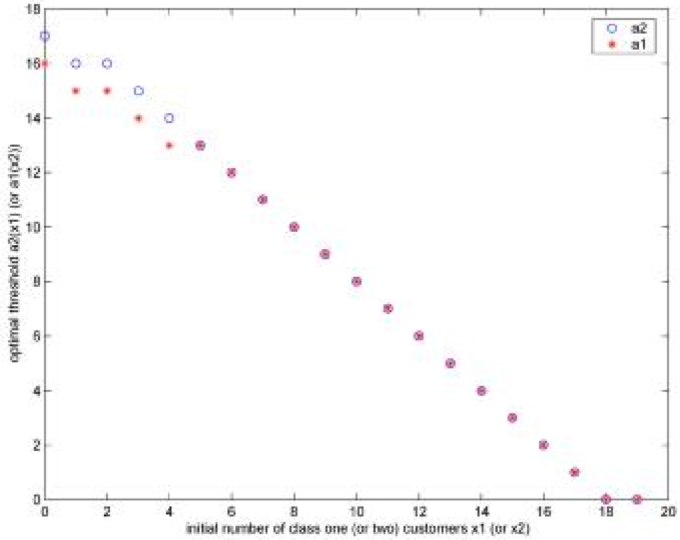

Figure 10 Optimal thresholds with n = 19, a = 0.1, l_1 = 0.15, l_2 = 0.25, m_1 = 0.3/19, m_2 = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q = 2000, p_1 = 0.1 and p_2 = 0.3.

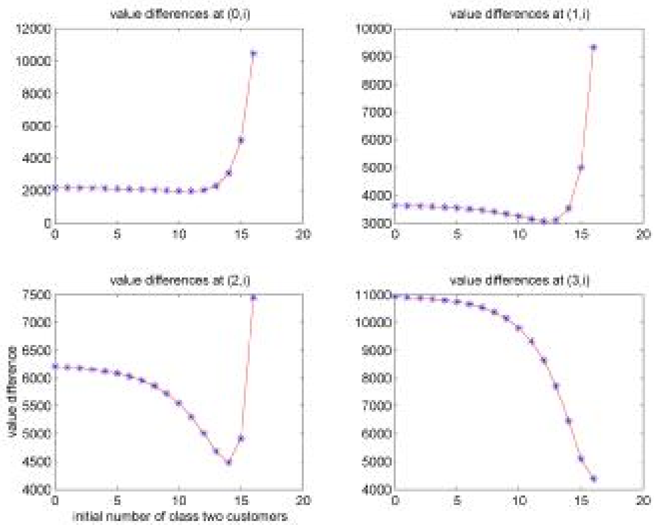

Figure 11: Comparison of value functions with n = 19, b = 3, a = 0.1, l_1 = 0.15, l_2 = 0.25, m_1 = 0.3/19, m_2 = 0.5/19, l_1 = 200, l_2 = 400, c = 1000, q_1 = 1000, q_2 = 2000.

5. Justifications of Some Previously Given Lemmas

5.1 Lemma 5.1 and Its Proof

Lemma 5.1 Let $A_i(\cdot)$ for each $i \in \{1, 2\}$ be a Poisson process with rate l_i and let $I_i(\cdot)$ denote the process satisfying that $I_i(t) = 1$ if there is an arrival of class *i* customer at the time point *t* and $I_i(t) = 0$ otherwise. Moreover, let $I(\cdot) \equiv \sum_{i=1}^{2} I_i(\cdot)$, then

$$P\{I(t) > 1 \ \textit{for all} \ t \in [0, \infty)\} = 0 \ . \tag{16}$$

Proof. Define $A(t) \equiv \sum_{i=1}^{2} A_i(t)$ for each $t \geq 0$, then the superposed process $A(\cdot)$ is still a Poisson process with rate $l = \sum_{i=1}^{2} l_i$. Thus, due to the independent increment property and by using the Poisson distribution, we have

$$P\{I(t) > 1\} \leq P\{A(t + \Delta t) - A(t) > 1\} = o(\Delta t).$$

Let $\Delta t \to 0$, we have that, for each $t \in [0, \infty)$,

$$P\{I(t) > 1\} = 0 \, . \tag{17}$$

Notice that $[0, \infty)$ is an uncountable set, therefore we take a sequence of time points as $0 < t_1 < t_2 < \mathbf{L} < t_m < \mathbf{L}$ satisfying $t_m \to \infty$ when $m \to \infty$, then the associated sequence of sets $\{I(t) > 1 \text{ for all } t \in [0, t_m]\}$ decreases to the set $\{I(t) > 1 \text{ for all } t \in [0, \infty)\}$ as $m \to \infty$. Moreover, for each t_m, let $t_m^{p,q} = \frac{q}{2^p} t_m$ for positive integers *p* and *q* that satisfy $q = 0, 1, \mathbf{K}, 2^p$. Then it follows from the right continuous property of $I(\cdot)$ and equation (17) that

$$\begin{aligned}
& P\{I(t) > 1 \text{ for all } t \in [0, \infty)\} \\
& = \lim_{m \to \infty} P\{I(t) > 1 \text{ for all } t \in [0, t_m]\} \\
& = \lim_{m \to \infty} P\left\{ \mathbf{U}_{p=1}^{\infty} \{I(t_m^{p,q}) > 1\} \right\} \\
& \leq \lim_{m \to \infty} \sum_{p=1}^{\infty} P\{I(t_m^{p,q}) > 1\} \\
& = 0.
\end{aligned}$$

Hence we complete the proof of the lemma. □

Remark 5.1 One can employ the same idea to prove that almost surely along any path, there is at most one customer who finishes service at a time point in $[0, \infty)$ since the service time distributions are assumed to be exponential.

5.2 Proof of Lemma 3.1

The proof of the lemma is the evaluation of equations (1) and (2) in Lippman[10]. First, we calculate the following expected a -discounted reward earned during one transition when starting from state *x* and choosing action *a*(*x*) through formula (2) in [10]. Following equation (3) and its expalnations in [10], we substitute g given in equation (7) into equation (2) in [10]. Then we have

$$
\begin{aligned}
r_a(x,a(x)) &= \int_S \left\{ \int_0^\infty \left[\int_0^t e^{-at} d((h(x)-c)t, t/x, a(x), x') \right] g e^{-gt} dt \right\} dq'(x'/x, a(x)) \\
&= (h(x)-c) \int_S \left\{ \int_0^\infty \left[\int_0^t e^{at} dt \right] g e^{-gt} \right\} dq'(x'/x, a(x)) \\
&= (h(x)-c) \int_S \left\{ \int_0^\infty \left(-\frac{1}{a} \right) e^{-at} \Big|_0^t g e^{-gt} dt \right\} dq'(x'/x, a(x)) \\
&= (h(x)-c) \int_0^\infty \left(-\frac{1}{a} \right) e^{-at} \Big|_0^t g e^{-gt} dt \\
&= \frac{1}{a+g} (h(x)-c) \\
&= h(x)-c
\end{aligned}
\tag{18}
$$

Next, notice that associated with a class i customer's arrival, the transition probability $q(x'/x, a(x))$ appeared in equation (3) of [10] is $l_i / l_{x,a(x)}$, and associated with a class i customer's departure, the probability is $x_i m_i / l_{x,a(x)}$. Therefore, by substituting g in equation (7), the above equation (18) and equation (3) in [10] into equation (1) in [10], we can conclude that the lemma is true. □

5.3 Proof of Lemma 3.4

Step One. We prove the concave property (10) for $Tv(x)$. In doing so, we first show that

$$
\Delta_j \Delta_j T_i v(x) \le 0 \quad \text{for all } i, j \in \{1, 2\}. \tag{19}
$$

As a matter of fact, for $j = i$, we have

$$
\Delta_i \Delta_i T_i v(x) \equiv \begin{cases}
\Delta_i \Delta_i v(x+e_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i < a_i(x)-1 \\
-(\Delta_i v(x+e_i) + l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x)-1 \\
\Delta_i v(x+e_i) + l_i \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x) \\
\Delta_i \Delta_i v(x) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i > a_i(x) \\
-(\Delta_i v(x+e_i) + l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i \le a_i(x)-1 \\
\Delta_i v(x+e_i) + l_i \le 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i = a_i(x) \\
\Delta_i \Delta_i v(x) \le 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i > a_i(x).
\end{cases}
$$

Then, it follows from Lemma 3.3 that $a_i(x) - 1 \le a_i(x+e_j) \le a_i(x)$ and $a_i(x+e_j) - 1 \le a_i(x+2e_j) \le a_i(x+e_j)$. Thus, for $j \ne i$, we have

$$
\Delta_j \Delta_j T_i v(x) \equiv \begin{cases}
\Delta_j \Delta_j v(x+e_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i \le a_i(x)-2 \\
\Delta_{ji} \Delta_j v(x+e_i) - (\Delta_i v(x+e_j) + l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x)-1, \\
& \quad a_i(x+e_j) = a_i(x)-1 \text{ or} \\
& \quad a_i(x+e_j) = a_i(x), \\
& \quad a_i(x+2e_j) = a_i(x)-2
\end{cases}
$$

$$\Delta_j\Delta_j T_i v(x) \equiv \begin{cases} \Delta_j\Delta_j v(x+e_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x)-1 \\ & a_i(x+e_j) = a_i(x)-1 \text{ or} \\ & a_i(x+e_j) = a_i(x), \\ & a_i(x+2e_j) = a_i(x+e_j) \\ \Delta_{ji}\Delta_j v(x+e_i) + (\Delta_i v(x+e_j)+l_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x)-1 \\ \Delta_{ji}\Delta_j v(x+e_i) - (\Delta_i v(x+e_j)+l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x), \\ & a_i(x+2e_j) = a_i(x+e_j)-1 \\ \Delta_j\Delta_j v(x+e_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x), \\ & a_i(x) = a_i(x+e_j) = a_i(x+2e_j) \\ \Delta_j\Delta_j v(x) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i > a_i(x), \\ & a_i(x) \ge a_i(x+e_j) \ge a_i(x+2e_j) \\ \Delta_{ji}\Delta_j v(x+e_i) - (\Delta_i v(x+e_j)+l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i \le a_i(x)-1 \\ \Delta_{ji}\Delta_j v(x+e_i) + (\Delta_i v(x+e_j)+l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x)-1 \\ \Delta_{ji}\Delta_j v(x+e_i) - (\Delta_i v(x+e_j)+l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i = a_i(x), \\ & a_i(x) = a_i(x+e_j) \\ \Delta_j\Delta_j v(x) \le 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i > a_i(x). \end{cases}$$

Step Two. We prove submodular property (11) for $Tv(x)$ by showing the following claim,

$$\Delta_j\Delta_i T_i v(x) = \Delta_i\Delta_j T_i v(x) \le 0 \quad \text{for} \quad j \ne i \tag{20}$$

As a matter of fact,

$$\Delta_j\Delta_i T_i v(x) \equiv \begin{cases} \Delta_j\Delta_i v(x+e_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i < a_i(x)-2 \\ -(\Delta_i v(x+e_i)+l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x)-1, \\ & a_i(x+e_i) = a_i(x)-1 \\ \Delta_j\Delta_i v(x+e_i) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x)-1 \\ & a_i(x+e_j) = a_i(x) \\ \Delta_i v(x+e_j)+l_i \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x)-1 \\ 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x) \\ \Delta_j\Delta_i v(x) \le 0 & \text{if } \Sigma_{i=1}^2 x_i < n-2, x_i > a_i(x) \\ -(\Delta_i v(x+e_i)+l_i) < 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i \le a_i(x)-1 \end{cases}$$

$$\Delta_j\Delta_i T_i v(x) \equiv \begin{cases} \Delta_i v(x+e_j)+l_i \le 0 & \text{if } \Sigma_{i=1}^2 x_i = n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x)-1 \end{cases}$$

$$\Delta_j \Delta_i T_i v(x) \equiv \begin{cases} 0 & if \sum_{i=1}^{2} x_i = n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x) \\ \Delta_j \Delta_i v(x) \le 0, & if \sum_{i=1}^{2} x_i = n-2, x_i > a_i(x). \end{cases}$$

Step Three. We prove property (12) for $Tv(x)$. We firstly show the following claim,

$$\Delta_{ij}\Delta_i T_i v(x) \le 0 \quad \text{for} \quad i \ne j. \tag{21}$$

As a matter of fact, we have

$$\Delta_{ij}\Delta_i T_i v(x) \equiv \begin{cases} \Delta_{ij}\Delta_i v(x+e_i) \le 0 & if \sum_{i=1}^{2} x_i < n-2, x_i \le a_i(x)-2 \\ 0 & if \sum_{i=1}^{2} x_i < n-2, x_i \le a_i(x)-1, \\ & a_i(x+e_j) = a_i(x)-1 \\ -(\Delta_i v(x+e_i+e_j)+l_i) < 0 & if \sum_{i=1}^{2} x_i < n-2, x_i = a_i(x)-1, \\ & a_i(x+e_j) = a_i(x) \\ \Delta_{ij}\Delta_i v(x) \le 0 & if \sum_{i=1}^{2} x_i < n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x)-1 \\ \Delta_i v(x+e_i)+l_i \le 0 & if \sum_{i=1}^{2} x_i < n-2, x_i = a_i(x) \\ & a_i(x+e_j) = a_i(x) \\ \Delta_{ij}\Delta_i v(x) \le 0 & if \sum_{i=1}^{2} x_i < n-2, x_i > a_i(x), \\ 0 & if \sum_{i=1}^{2} x_i = n-2, x_i \le a_i(x)-2, \\ 0 & if \sum_{i=1}^{2} x_i = n-2, x_i = a_i(x)-1, \\ & a_i(x+e_j) = a_i(x)-1 \\ 0 & if \sum_{i=1}^{2} x_i = n-2, x_i = a_i(x)-1, \\ & a_i(x+e_j) = a_i(x) \\ \Delta_{ij}\Delta_i v(x) \le 0 & if \sum_{i=1}^{2} x_i = n-2, x_i = a_i(x) \\ & a_i(x+e_j) = a_i(x)-1 \\ \Delta_i v(x+e_i)+l_i \le 0 & if \sum_{i=1}^{2} x_i = n-2, x_i = a_i(x), \\ & a_i(x+e_j) = a_i(x) \\ \Delta_{ij}\Delta_i v(x) \le 0 & if \sum_{i=1}^{2} x_i = n-2, x_i > a_i(x). \end{cases}$$

We secondly show that

$$\Delta_{ji}\Delta_j T_i v(x) \le 0 \text{ for } i \ne j. \tag{22}$$

As a matter of fact, $\Delta_{ji}\Delta_j T_i v(x) \equiv$

$$\begin{cases} \Delta_{ji}\Delta_j v(x+e_i) \le 0 & if \sum_{i=1}^{2} x_i < n-2, x_i \le a_i(x)-2 \\ \Delta_{ji}\Delta_j v(x) + \Delta_{ij}\Delta_i v(x) \le 0 & if \sum_{i=1}^{2} x_i < n-2, x_i = a_i(x)-1 \\ & a_i(x+e_j) = a_i(x)-1, \\ & a_i(x+2e_j) = a_i(x+e_j)-1 \end{cases}$$

$$
\begin{cases}
\Delta_{ji}\Delta_j v(x+e_i)+(\Delta_i v(x+e_i+e_j)+l_i)\le 0 & if\ \sum_{i=1}^2 x_i<n-2, x_i=a_i(x)-1,\\
 & a_i(x+e_j)=a_i(x)-1,\\
 & a_i(x+2e_j)=a_i(x+e_j)\\
\Delta_{ji}\Delta_j v(x+e_i)\le 0 & if\ \sum_{i=1}^2 x_i<n-2, x_i=a_i(x)-1,\\
 & a_i(x+e_j)=a_i(x)\\
 & a_i(x+2e_j)=a_i(x+e_j)-1\\
\Delta_{ji}\Delta_j v(x)\le 0 & if\ \sum_{i=1}^2 x_i<n-2, x_i=a_i(x),\\
 & a_i(x+e_j)=a_i(x-1),\\
 & a_i(x+2e_j)=a_i(x+e_j)-1\\
\Delta_{ji}\Delta_j v(x)\le 0 & if\ \sum_{i=1}^2 x_i<n-2, x_i=a_i(x),\\
 & a_i(x+e_j)=a_i(x)-1,\\
 & a_i(x+2e_j)=a_i(x+e_j)\\
\Delta_{ji}\Delta_j v(x)-(\Delta_i v(x+e_j)+l_i)<0 & if\ \sum_{i=1}^2 x_i<n-2, x_i=a_i(x),\\
 & a_i(x+e_j)=a_i(x),\\
 & a_i(x+2e_j)=a_i(x+e_j)-1\\
\Delta_j\Delta_j v(x+e_i)\le 0 & if\ \sum_{i=1}^2 x_i<n-2, x_i=a_i(x),\\
 & a_i(x+e_j)=a_i(x),\\
 & a_i(x+2e_j)=a_i(x+e_j)-1\\
\Delta_{ji}\Delta_j v(x)\le 0 & if\ \sum_{i=1}^2 x_i<n-2, x_i>a_i(x)\\
\Delta_{ji}\Delta_j v(x)+\Delta_{ij}\Delta_i v(x)\le 0 & if\ \sum_{i=1}^2 x_i=n-2, x_i\le a_i(x)-1\\
\Delta_{ji}\Delta_j v(x)\le 0 & if\ \sum_{i=1}^2 x_i=n-2, x_i=a_i(x),\\
 & a_i(x+e_j)=a_i(x)-1\\
\Delta_{ji}\Delta_j v(x)-(\Delta_i v(x+e_j)+l_i)<0 & if\ \sum_{i=1}^2 x_i=n-2, x_i=a_i(x),\\
 & a_i(x+e_j)=a_i(x)\\
\Delta_{ji}\Delta_j v(x)\le 0 & if\ \sum_{i=1}^2 x_i=n-2, x_i>a_i(x).
\end{cases}
$$

Step Four. Let

$$f(x)=\sum_{i=1}^{2}x_i v(x-e_i)+(n-\sum_{i=1}^{2}x_i)v(x) \tag{23}$$

Then we have the claim that $f(x)$ satisfies (10)-(12) for $\sum_{i=1}^{2}x_i\le n-2$. Firstly, we show that (10) is satisfied. As a matter of fact, for $i,j\in\{1,2\}$,

$$
\begin{aligned}
&\Delta_j\Delta_j f(x)\\
=&\ \sum_{i=1}^{2}\Delta_j\Delta_j(x_i v(x-e_i))+n\Delta_j\Delta_j v(x)-\sum_{i=1}^{2}\Delta_j\Delta_j(x_i v(x))\\
=&\ (x_i\Delta_j\Delta_j v(x-e_i)+x_j\Delta_j\Delta_j v(x-e_j))+(n-x_i-x_j-2)\Delta_j\Delta_j v(x)\ \ (for\ \ i\ne j)\\
\le&\ \ 0.
\end{aligned} \tag{24}
$$

Secondly, we show that (11) is satisfied. As a matter of fact, for $j,k \in \{1,2\}$ and $j \neq k$,

$$\begin{aligned} & \Delta_j \Delta_k f(x) \\ = & \sum_{i=1}^{2} \Delta_j \Delta_k (x_i v(x-e_i)) + n\Delta_j \Delta_k v(x) - \sum_{i=1}^{2} \Delta_j \Delta_k (x_i v(x)) \\ = & (x_k \Delta_j \Delta_k v(x-e_k) + x_j \Delta_j \Delta_k v(x-e_j)) + (n - x_k - x_j - 2)\Delta_j \Delta_k v(x) \\ \leq & \ 0. \end{aligned} \tag{25}$$

Thirdly, we show that (12) is satisfied. As a matter of fact, for $j,k \in \{1,2\}$ and $j \neq k$,

$$\begin{aligned} & \Delta_{jk} \Delta_j f(x) \\ = & \sum_{i=1}^{2} \Delta_{jk} \Delta_j (x_i v(x-e_i)) + n\Delta_{jk} \Delta_j v(x) - \sum_{i=1}^{2} \Delta_{jk} \Delta_j (x_i v(x)) \\ = & (x_k \Delta_{jk} \Delta_j v(x-e_k) + x_j \Delta_{jk} \Delta_j v(x-e_j)) + (n - x_k - x_j - 2)\Delta_{jk} \Delta_j v(x) \\ \leq & \ 0. \end{aligned} \tag{26}$$

Finally, by steps one to four and by the fact that $h(x)$ satisfies conditions (10)-(12), we know that $Tv(x)$ satisfies conditions (10)-(12). Therefore we complete the proof of the lemma. □

6. Concluding Remarks and Future Research

In this paper, we studied the stochastic optimal admission and routing controls for a system consisting of two parallel-server channels with zero waiting buffer capacity and multi-class customers via Markov decision processes moving in discrete triangles. Our findings can possibly be extended to more general cases in two directions. Firstly, in the current paper, the classes of customers are classified according to the arrival rates, payments and penalty costs. We conjecture that our results should be true when the service rates for different classes are also different. This conjecture is partially justified by the numerical implementation summarized in Example 4.6. Secondly, when there are more than two classes of customers who are served in each channel, the corresponding Markov decision processes move in high-dimensional discrete trihedrons. The research to find the suitable class of structured value functions that is preserved under certain sense is also interesting.

7. Acknowledgements

Wanyang Dai is supported by National Natural Science Foundation of China under Contract 10371053. The main part of the project was done while the first author was visiting the Department of Systems Engineering and Engineering Management at The Chinese University of Hong Kong in 2007. The author expresses sincere thanks for the kind invitation and financial support from the department and the university. Youyi Feng is supported by a grant from the Research Council of the Hong Kong Special Administrative Region (Project No. CUHK 4320/02E).

References

[1] Bell, S.L., Williams, R.J. Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: asymptotic optimality of a threshold policy. *Annals of Applied probability*, 2001, 11(3): 608-649.

[2] Benjaafar, S., ElHafsi, M. Production and inventory control of a single product Assemble-to-Order system with multiple customer classes. *Management Science*, 2006, 52

(12): 1896-1912.

[3] Cheng, C., Huang, S., Peng, S., Kang, J. The optimization of manufacturing system based on material flow and capital flow. *Advances in Systems Science and Applications*, 2005, 5(2): 266-270.

[4] Dai, W. Product-form solutions for integrated services packet networks. *Proceedings of MICAI'06*. IEEE Computer Society Press, Los Alamitos. 2006: 296-308.

[5] Dai, W., Jiang, Q. Stochastic optimal control of ATO systems with batch arrivals via diffusion approximation. *Probability in the Engineering and Informational Sciences*, 2007, 21(3): 477-495.

[6] Ding, L.Y., Zhou, Y., Luo, H.B., Gu, X. Study on construction cost control system based on multiple dimension data model. *Advances in Systems Science and Applications*, 2007, 7(4): 541-548.

[7] Gans, N., Koole, G., Mandelbaum, A. Telephone call centers: Tutorial, review and research prospects. *Manufacturing and Service Operations Management*, 2003, 5(2): 79-141.

[8] Harrison, J.M. Heavy traffic analysis of a system with parallel servers: asymptotic optimality of discrete-review policies. *Annals of Applied Probability*, 1998, 8(3): 822-848.

[9] Hiriart-Urruty, J., Lemarechal, C. Fundamentals of Convex Analysis. Springer, Berlin. 2001: 23-88.

[10]Lippman, S. Applying a new device in the optimization of exponential queueing systems. *Operations Research*, 1975, 23(4): 687-710.

[11]Ma, C., Fang, H., Kong, D., Huo, Z. Stochastic optimal control of networked control systems with multi-step delay. *Advances in Systems Science and Applications*, 2005, 5(4): 566-574.

[12]Ormeci, E.L. Burnetas, A., Wal, J.V.D. Admission policies for a two class loss system. *Stochastic Models*, 2001, 17(4): 513-539.

[13]Ormeci, E.L., Wal, J.V.D. Admission policies for a two class loss system with general interarrival times. *Stochastic Models*, 2006, 22(1): 37-53.

[14]Porteus, E.L. Conditions for characterizing the structure of optimal strategies in infinite horizon dynamic programs. *Journal of Optimization Theory and Applications*, 1982, 36(4): 419-432.

[15]Puterman, M. “Markov Decision Processes: Discrete Stochastic Dynamic Programming” (Chapetr 6). John Wiley and Sons, Inc. New York. 1994: 142-276.

[16]Sennott, L. Average reward optimization theory for denumerable state spaces. *In Handbook of Markov Decision Processes Methods and Applications: International Series in Operations Research and Management Science*. Editors: E.A. Feinberg and A. Schwartz. Kluwer Academic. 2002: 153-172.