

## Research Article

# Analysis of the Optimal Resource Allocation for a Tandem Queueing System

Zaiming Liu, Wei Deng, and Gang Chen

*School of Mathematics and Statistics, Central South University, Changsha, Hunan 410083, China*

Correspondence should be addressed to Gang Chen; [chengmathcsu@163.com](mailto:chengmathcsu@163.com)

Received 1 May 2017; Accepted 2 November 2017; Published 28 November 2017

Academic Editor: Qingling Zhang

Copyright © 2017 Zaiming Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We study a controllable two-station tandem queueing system, where customers (jobs) must first be processed at upstream station and then the downstream station. A manager dynamically allocates the service resource to each station to adjust the service rate, leading to a tradeoff between the holding cost and resource cost. The goal of the manager is to find the optimal policy to minimize the long-run average costs. The problem is constructed as a Markov decision process (MDP). In this paper, we consider the model in which the resource cost and service rate functions are more general than linear. We derive the monotonicity of the optimal allocation policy by the quasiconvexity properties of the value function. Furthermore, we obtain the relationship between the two stations' optimal policy and conditions under which the optimal policy is unique and has the bang-bang control property. Finally, we provide some numerical experiments to illustrate these results.

## 1. Introduction

Queueing systems where customers must be processed at each station in series from upstream station to downstream station are called tandem queueing system. As we all know, tandem queueing models have widespread applications in both service organizations and production factory in the sense that the system performance measures and optimization are of primary concerns, such as the control of semiconductor fabrication processes and broadband wireless networks, appointment scheduling in hospital, and production inventory system (see [1–3] and therein). Recently, this issue has attracted much attention and vast literatures have been studied, especially the dynamic resource allocation problems of the tandem queueing system. Most of them are forced on the models in two directions: admission control type and server resource allocation type. The admission control of the tandem queues have been widely studied (e.g., [4, 5]), while little work has appeared concerning the structure of the optimal resource allocation policy in the tandem queues.

For many systems, service consists of two or more phases by one or more servers. A fundamental decision is how to

allocate the resource (servers or workforces) owned by system to each station. This problem is a classic topic, which roots from Rosberg et al. [6] where the service rate in station 1 can be selected from a compact set and constant in station 2. Optimal control of a two-stage tandem queues system with only two flexible servers was discussed in Ahn et al. [7]. Arumugam et al. [8] considered inventory based allocation policies for flexible servers in serial systems. Smith and Barnes [9] analyzed the optimal server allocation in closed finite queueing networks. This question has been considered by the above authors for different cost or reward criterion but without considering the structure of the optimal policy. They just make a numerical experiment to get the optimal policy. In fact, it is more complex for the manager to obtain the detailed optimal policy in the practical application. The managers prefer more to make basal insight for the structure of the optimal policy. While for the study of the structure of optimal policy in single queue, many papers have investigated this issue. Iravani et al. [10] studied the optimal service scheduling in nonpreemptive finite-population queueing systems. The single-queue systems of the optimal resource allocation policy were considered by Yang et al. [11], who investigated

the structural properties of the optimal resource allocation policy. Yang et al. [12] studied optimal resource allocation for parallel service facilities multiqueue systems with a shared server pool.

For the optimal control of the tandem queueing system, some relative works are discussed. Weber and Stidham [13] considered a problem of optimal service rate control in queueing networks where the optimal policy has a monotone structure. Veatch and Wein [14] generalize the monotonicity results of [13] where the control policies were studied under the full information and service rate functions are linear in the service resource. Mayorga et al. [15] studied the problem of allocating flexible servers for a firm that operates a make-to-order serial production system. The max-min optimality of service rate control problem in closed queueing networks was studied by Xia and Shihada [16] where the cost functions are strictly convex (concave). Few studies, even among the most recent, have considered the structure of the optimal resource allocation policy in the tandem queues with nonlinear service rate and cost functions. In many logistic environments, however, the assumption of the linear resource cost and service rate is not appropriate. It is well known that if the service cost is linear, these problems have all-or-nothing (bang-bang) optimal policy (see [14]). Different from the works quoted above, in our model, the service resource cost and service rate are more general than linear in service resource. Recently, Xia et al. [17] investigated the optimal control of service rates of a tandem queue with power constraints and general cost function. They mainly derived some structures of the optimal policy, such as the bang-bang control policy and 3-element set policy for some special cases, while, in this paper, we study the structure of the optimal control policy for the case with general cost and service rate functions. Moreover, some uninvestigated properties of the optimal policy are obtained in this paper. Using the theory of queueing system, we cast the optimal problem as a MDP. The theory of Markov, semi-Markov, and regenerative decision processes can be found in Morozov and Steyaert [18]. We mainly analyze the properties of the optimal policy under full information and partial information. Concretely, we first derived the properties (monotonicity and convexity property) of value function by the induction method and queueing theory (see [19]). Second, we provide insights into the optimal policy structure based on the properties of the value function and dynamic programming method (see [20–22]). Furthermore, we take Howard's iteration procedure to obtain numerical results.

The main contributions of this paper can be summarized as follows. First, to the best of our knowledge, our paper is the first to study the optimal resource allocation policy in the tandem queues with the general service rate and resource cost functions. Second, we get the monotone results of the optimal policy under the partial information based on the quasi-convexity property of the value function. Third, we derive the conditions under which the optimal policy is unique and the bang-bang control policy is established. This conclusion is totally new progress compared with all of the previous works in the literature. Furthermore, we derive the relationship between the two stations' optimal policy. As far as we know,

these are the most general results for the optimality of resource allocation in the tandem queueing system.

The rest of the paper is organized as follows. In Section 2, we introduce the model formulation in detail based on the controllable Markov decision problem. The characteristics of the optimization problem and the optimality equation are derived in Section 3. In Section 4, we present the structural properties of the optimal policy and main results of the paper. In Section 5, we give some numerical examples to provide the support for the results of the present model. Finally, some further discussions and conclusions are given in Section 6.

## 2. Model Description

We consider a tandem queueing system with two stations. Arrivals to the system at station 1 from outside follow a Poisson process with parameter  $\lambda$  and have exponentially distributed service requirement times at each station. After receiving service at station 1, customers join immediately to station 2 and receive service before leaving the system. A decision-maker can assign a number of service resources to each station. The service rate depends on the number of service resources assigned to the stations precisely. When a station has been allocated  $a$  resources, the service duration of the customer in station  $i$  is exponentially distributed with rate  $\mu_i(a)$ ,  $i = 1, 2$ , which is strictly increasing in  $a$ . Without loss of generality, we assume that  $\mu_i(0) = 0$ ,  $i = 1, 2$ . At any decision epoch, the decision-maker decides to choose the number of service resources to station 1 from a set  $A = [0, a_{\max}]$  and to station 2 from a set  $B = [0, b_{\max}]$  at the same time. Each station has a single infinite-size FCFS queue. The interarrival and service times are assumed to be mutually independent. We assume that the stability condition  $\lambda < \min\{\mu_1(a_{\max}), \mu_2(b_{\max})\}$  holds. Figure 1 gives an illustration of the system.

We consider the following cost structure in the system. Our objective is to obtain dynamic resource allocation policy that minimizes the long-run average costs.

(1) *Resources Cost.* When station  $i$  uses  $a$  resources, a cost of  $c_i(a)$ ,  $i = 1, 2$  is incurred by the system per unit time ( $c_i(a)$  is a continuous function and strictly increasing in  $a$ . Without loss of generality, we assume that  $c_i(0) = 0$ ).

(2) *Holding Cost.* Holding costs are incurred at rates  $h_1$  and  $h_2$  per unit time for each customer in stations 1 and 2, respectively.

Let  $X_i(t)$  denote the number of customers at station  $i$ ,  $i = 1, 2$ . The state of the system at time  $t$  can be described by  $X(t) = (X_1(t), X_2(t))$ . The system evolves as a continuous-time Markov process  $X(t) = \{(X_1(t), X_2(t)), t \geq 0\}$ . We define the notations  $x_i$  to classify the certain components of the vector state  $x \in E$ . Clearly, the system state space is  $E = \{(x_1, x_2) \mid x_1, x_2 \in N\}$  with  $N = \{0, 1, 2, \dots\}$ . We consider the stationary Markov policy under which the system evolves as a continuous-time Markov chain. Moreover, in order to study the optimal policy in the ergodic Markov process, we assumed that the model is stable and conservative. The transition rate under a control action  $(a, b)$  is given by

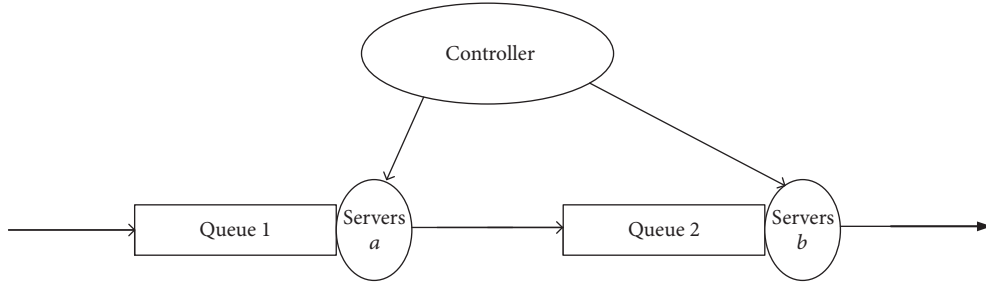


FIGURE 1: The controllable tandem queueing systems.

$$q_{xy}(a, b) = \begin{cases} \lambda & y = x + e_1; \\ \mu_1(a) & y = x - e_1 + e_2, x_1 > 0; \\ \mu_2(b) & y = x - e_2, x_2 > 0; \\ 0 & \text{else,} \end{cases} \quad (1)$$

where

$$\begin{aligned} q_{xy}(a, b) &\geq 0, \quad y \neq x, \\ q_{xx}(a, b) &= -q_x(a, b) = -\sum_{y \neq x} q_{xy}(a, b), \\ q_x(a, b) &< \infty. \end{aligned} \quad (2)$$

Here  $e_i$  is the 2-dimensional vector with 1 in the  $i$ th coordinate and 0 elsewhere,  $i = 1, 2$ .

The problem of the decision-maker is to choose an optimal dynamic policy based on the number of customers in each station that minimizes the long-run average costs. We formulate the service resource management problem as a Markov decision process. The set of decision epochs is composed of the set of all arrivals and service completions. The controllable system associated with a Markov process is a five-tuple

$$\{E, D = (A, B), Q(r), c_i(a), h_i\} \quad (i = 1, 2), \quad (3)$$

in which  $Q(r) = (q_{x,y}(r))_{x,y \in E}$  is the infinitesimal generator of the queueing system under the policy  $r$ . We consider the stationary Markov policy  $r : E \rightarrow D$  with  $r = (a, b)$ . Due to the Markov property of the queueing system, we know that the optimal policy depends only on the current state regardless of  $t$ . In our model we consider two situations: the decision with partial information and full information. Concretely, when the system state is  $x = (x_1, x_2)$ , the manager makes an action as follows:

- (i) Partial information: the action for station 1 (2) is  $a(x_1) \in A$  ( $b(x_2) \in B$ , resp.). That is the action of resource to station  $i$  only depends on the number of customers in station  $i$ .
- (ii) Full information: the action for station 1 (2) is  $a(x_1, x_2) \in A$  ( $b(x_1, x_2) \in B$ , resp.). That is the action of resource to station  $i$  depends on the number of customers in both stations.

### 3. Optimization Problem and Optimality Equation

It is obvious that, under the stability condition  $\lambda < \min\{\mu_1(a_{\max}), \mu_2(b_{\max})\}$ , the two-dimensional stochastic process  $X(t) = \{(X_1(t), X_2(t)), t \geq 0\}$  is an ergodic continuous-time Markov chain for any fixed stationary policy  $r$ . As it is known from Tijms [23], the long-run average cost per unit of time for the policy  $r$  in our ergodic Markov process can be written in the following form:

$$\begin{aligned} g(r) &= \lim_{t \rightarrow \infty} \frac{u^r(x, t)}{t} \\ &= \sum_{i=1}^2 \sum_{j=1}^2 [c_1(a) + c_2(b) + h_1 i + h_2 j] \pi_{ij}(r), \end{aligned} \quad (4)$$

in which  $u^r(x, t)$  denotes the total expected costs up to time  $t$  when the system starts in state  $x = (i, j)$  and  $\pi_{ij}(r)$  denotes a stationary probability of the process under policy  $r = (a, b)$ . The goal is to find a policy  $r^*$  that minimizes the long-term average costs:

$$g(r^*) = \min_r g(r). \quad (5)$$

Using the standard tools of uniformization and normalization, we construct a discrete-time equivalent of our original queueing system. Without loss of generality, we assume that  $\lambda + \mu_1(a_{\max}) + \mu_2(b_{\max}) = 1$ . Now we consider a real-valued function  $v(x)$  which is defined on the state space. The relative value function  $v(x)$  can be regarded as the asymptotic difference in total costs that results from starting the process in state  $x$  instead of some reference state. As is shown in Puterman [24], the optimal policy  $r$  and the optimal average cost  $g$  are the solutions of the optimality equation:

$$Tv(x) = v(x) + g, \quad (6)$$

where  $T$  is the dynamic programming operator acting on  $v$  defined as follows:

$$Tv(x) = \lambda v(x + e_1) + \sum_{i=1,2} T_i v(x) + \sum_{i=1,2} h_i x_i, \quad (7)$$

in which

$$\begin{aligned} T_1 v(x) &= \min_{a \in A} \{ \mu_1(a) v(x - e_1 + e_2) \\ &\quad + [\mu_1(a_{\max}) - \mu_1(a)] v(x) + c_1(a) \}, \\ T_2 v(x) &= \min_{b \in B} \{ \mu_2(b) v(x - e_2) \\ &\quad + [\mu_2(b_{\max}) - \mu_2(b)] v(x) + c_2(b) \}. \end{aligned} \quad (8)$$

The first term in the expression  $Tv(x)$  models the arrivals of customers to station 1 from outside the system and the last one the customer holding cost. Similarly the first term in the expression  $T_1 v(x)$  corresponds to a customer who finished his service in station 1 and into station 2 and the second one the uniformization constant. The last one in  $T_1 v(x)$  is the resources cost in station 1. The first term in the expression  $T_2 v(x)$  corresponds to a customer who finished his service in station 2 and the second one the uniformization constant. The last one in  $T_2 v(x)$  is the resources cost in station 2.

According to (4), we can solve another optimization problem: if  $c_i \equiv 0$ ,  $h_i = 1$ ,  $i = 1, 2$ , then (5) is equivalent to minimization of the mean number of customers in the queueing system. In this case, the optimal action would be always  $(a_{\max}, b_{\max})$  by intuition, which also satisfies the structure of the optimal policy in next section. In addition, the analysis method and structure in this section are held for both the partial and full information cases.

#### 4. Structural Properties of the Optimal Policy

In this section, we focus on deriving the optimal policy. The properties of the optimal policy will provide basal insight for us, and this also helps one to find the optimal policy with less computational effort due to a reduction of the solution search space.

In order to study the optimal policy, intuitively, the optimal equation  $Tv(x) = v(x) + g$  should be solved. However it is hard to solve analytically in practice. It can be obtained by recursively defining  $v_{n+1} = Tv_n$  for arbitrary  $v_0$ . We know that the actions converge to the optimal policy as  $n \rightarrow \infty$ . For existence and convergence of the solutions and optimal policy, we can see more details in the works of Aviv and Federgruen [25] and Sennott [26]. The backward recursion equation in our model is given by

$$v_{n+1}(x) = \lambda v_n(x + e_1) + \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} h_i x_i. \quad (9)$$

For ease of notation, let the  $\arg T_i v(x)$ ,  $i = 1, 2$  denote the set of optimal action for station  $i$  with state  $x = (x_1, x_2)$  in the partial information case in which the action is  $(a(x_1), b(x_2))$ , where  $a(x_1) \in \arg T_1 v(x)$ ,  $b(x_2) \in \arg T_2 v(x)$ .

Using the optimality equation and recursive method, we can get some properties of the relative value function in the following lemma which will be used in the proof of the main results and the proofs of these properties are given in Appendix A.

**Lemma 1.** For the optimal value function  $v(x)$  in this model, we have

- (i)  $v(x + e_i) \geq v(x)$ ,  $i = 1, 2$  for all  $x \in E$ ;
- (ii) if  $2h_2 \geq h_1$ , then  $v(x - e_1 + e_2) \geq v(x - e_2)$  for all  $x \in E$ ;
- (iii) if  $h_1 \geq h_2$ , then  $v(x) \geq v(x - e_1 + e_2)$  for all  $x \in E$ .

As we know, at the decision epochs if the manager gets the full information about the system, he will make a decision based on the number of the customers in both stations. Weber and Stidham [13] and Veatch and Wein [14] used submodularity of the value function  $v(x) - v(x + e_2 - e_1) - v(x - e_2) + v(x - e_1) \leq 0$  to prove the main conclusion transition monotonicity for the full information case  $a(x) \leq a(x - e_2)$  and  $b(x) \leq b(x + e_2 - e_1)$ . The optimal resource allocation policy has the switching function policy or region control policy type for the full information case [17]. However, the corresponding results for the partial information case are not studied. In this paper, we study the property of the optimal policy under partial information. Different from method in the full information case, we get some structure properties of the optimal policy by the quasiconvexity property of the relative value function and present the structure properties of the optimal policy in the following theorem.

**Theorem 2.** In our model under partial information, the optimal policy has the monotonicity properties, that is, for all  $x = (x_1, x_2) \in E$ :

- (i) If  $b(x_2 + 1) \in \arg T_2 v(x + e_2)$  and  $b(x_2) \in \arg T_2 v(x)$ , then  $b(x_2 + 1) \geq b(x_2)$ .
- (ii) If  $a(x_1 + 1) \in \arg T_1 v(x + e_1)$  and  $a(x_1) \in \arg T_1 v(x)$ , then  $a(x_1 + 1) \geq a(x_1)$ .

The proof of the above theorem is based on the following property which shows the quasiconvexity properties of the relative value function. The proofs of Theorem 2 and the following Lemma 3 are given in Appendix B.

**Lemma 3.** For the optimal value function  $v(x_1, x_2)$  under partial information, we have

- (i)  $v(x + e_2) - 2v(x) + v(x - e_2) \geq 0$ , for all  $x = (x_1, x_2) \in E$ ;
- (ii)  $v(x + e_1 - e_2) - 2v(x) + v(x - e_1 + e_2) \geq 0$ , for all  $x = (x_1, x_2) \in E$ .

Based on the above properties of the value functions, we derive the relationship between the two stations' optimal policy by analyzing the properties of the service rate and holding cost functions. The following theorem shows the conditions under which the optimal policy for station 1 is bigger than that in station 2.

**Theorem 4.** Assume that  $a_{\max} = b_{\max}$  and  $2h_2 \geq h_1$ , if the condition holds  $(c_1(x) - c_1(y)) / (c_2(x) - c_2(y)) \geq 1$  and  $(\mu_2(x) - \mu_2(y)) / (\mu_1(x) - \mu_1(y)) \geq 1$  when  $x \geq y$ . Then we have  $b \geq a$  where  $a \in \arg T_1 v(x)$ ,  $b \in \arg T_2 v(x)$ .

*Proof.* Let  $(a \in \arg T_1 v(x), b \in \arg T_2 v(x))$  be an arbitrary optimal policy for stations 1 and 2 in state  $x$ , respectively. The



proof is taken by contradiction method. Suppose that  $b < a$ ; then we compare the policy  $(a, b)$  with the policy  $(b, a)$  (the assumption  $a_{\max} = b_{\max}$  guarantees that we can always swap  $a$  and  $b$ ). Hence, we have

$$\begin{aligned}
T_{a,b}v_n(x) - T_{b,a}v_n(x) &= [\mu_1(a)v(x - e_1 + e_2) \\
&+ [\mu_1(a_{\max}) - \mu_1(a)]v(x) + c_1(a)] \\
&+ [\mu_2(b)v(x - e_2) + [\mu_2(b_{\max}) - \mu_2(b)]v(x) \\
&+ c_2(b)] - [\mu_1(b)v(x - e_1 + e_2) \\
&+ [\mu_1(b_{\max}) - \mu_1(b)]v(x) + c_1(b)] \\
&- [\mu_2(a)v(x - e_2) + [\mu_2(a_{\max}) - \mu_2(a)]v(x) \\
&+ c_2(a)] = [\mu_1(a) - \mu_1(b)][v(x - e_1 + e_2) \\
&- v(x)] - [\mu_2(a) - \mu_2(b)][v(x - e_2) - v(x)] \\
&+ c_1(a) - c_1(b) - c_2(a) + c_2(b) \geq [\mu_1(a) - \mu_1(b)] \\
&\cdot [v(x - e_1 + e_2) - v(x - e_2)] + c_1(a) - c_1(b) \\
&- c_2(a) + c_2(b) \geq 0.
\end{aligned} \tag{10}$$

The first equality is based on the definitions of operators  $T_1$  and  $T_2$ . The second equality follows by rearranging the terms. The first inequality follows the conditions  $(\mu_2(a) - \mu_2(b))/(\mu_1(a) - \mu_1(b)) \geq 1$  when  $a \geq b$  and Lemma 1 (i). The last inequality based on the conditions  $(c_1(a) - c_1(b))/(c_2(a) - c_2(b)) \geq 1$  when  $a \geq b$  and Lemma 1 (ii). So that we obtain  $T_{a,b}v_n(x) - T_{b,a}v_n(x) \geq 0$ , which implies that  $(a, b)$  is not an optimal policy for state  $x$ . Hence, we have  $b \geq a$ .  $\square$

*Remark 5.* From the above theorem we can conclude that under some conditions the optimal size of the service resources allocated to station 1 is less than that to station 2. We find that the optimal size of the resource allocated to each station depends on the resource cost variation  $c(a) - c(b)$  and the service rate variation  $\mu(a) - \mu(b)$  in each station. This condition seems to imply that when the same service resources are added to both station 1 and station 2, then the performance of station 2 is improved more than station 1 while the higher cost is incurred in station 1 than in station 2. So that it implies that the optimal policy satisfies the relationship  $b \geq a$ .

It is well known that if the service resource cost function is linear, then an all-or-nothing (bang-bang) control is optimal. Weber and Stidham [13] and Veatch and Wein [14] give a detailed conclusion for this issue. Actually it is not obvious whether the bang-bang control is also optimal, when the service resource cost and service rate functions are more general than linear in service resource. We are interested in the special structure of the optimal control policy in this model. In contrast to existing studies, the results in the following theorem are extension of the model with linear case. We are now ready to give some conditions under which the optimal policy is unique and has the bang-bang control property.

**Theorem 6.** (i) The optimal policy is unique if the following conditions hold: (1)  $m_1(a) = c'_1(a)/\mu'_1(a)$  and  $m_2(b) = c'_2(b)/\mu'_2(b)$  are monotonous on  $a \in A$ ,  $b \in B$ . (2)  $\exists a \in (0, a_{\max})$  and  $b \in (0, b_{\max})$  s.t.,  $c_1(a)/\mu_1(a) < c_1(a_{\max})/\mu_1(a_{\max})$  and  $c_2(b)/\mu_2(b) < c_2(b_{\max})/\mu_2(b_{\max})$ .

(ii) The optimal policy is a bang-bang control policy; that is,  $\arg T_1 v(x) = \{0, a_{\max}\}$ ,  $\arg T_2 v(x) = \{0, b_{\max}\}$  if the functions  $c_1(a)/\mu_1(a)$  and  $c_2(b)/\mu_2(b)$  are strict decreasing for all  $a \in (0, a_{\max})$ ,  $b \in (0, b_{\max})$ .

*Proof.* We prove only conclusion for station 1, and the same proof can be applied to get conclusion for station 2. To prove part (i), we consider the optimal policy  $a$  in station 1 service resource allocation. For the definition of the operator  $T_1$ , we have the following minimization problem:

$$\begin{aligned}
T_1 v(x) &= \min_{a \in A} \{ \mu_1(a)v(x - e_1 + e_2) \\
&+ [\mu_1(a_{\max}) - \mu_1(a)]v(x) + c_1(a) \}.
\end{aligned} \tag{11}$$

Rearranging the terms of the first-order optimality condition of the above problem, we obtain

$$\frac{c'_1(a)}{\mu'_1(a)} = v(x) - v(x - e_1 + e_2). \tag{12}$$

Because the allocation resource action  $a \in A = [0, a_{\max}]$ , the optimal policy  $a$  in station 1 can be 0 or  $a_{\max}$  or satisfies the above equation. Since the function  $m_1(a) = c'_1(a)/\mu'_1(a)$  is monotonous on  $a \in A$ , there is at most one solution  $a$  solving the above equation. Next if the optimal policy is 0 or  $a_{\max}$ , we show that the action 0 and  $a_{\max}$  cannot be the optimal policy simultaneous for station 1. We take the contradiction method. Assume that the action 0 and  $a_{\max}$  be the optimal policy simultaneous for station 1 at state  $x$ . Then we have  $T_1 v(x) = T_1^0 v(x) = \mu_1(a_{\max})v(x)$  and  $T_1 v(x) = T_1^{a_{\max}} v(x) = \mu_1(a_{\max})v(x - e_1 + e_2) + c_1(a_{\max})$ ; that is,

$$v(x) - v(x - e_1 + e_2) = \frac{c_1(a_{\max})}{\mu_1(a_{\max})}. \tag{13}$$

Because the action 0 is optimal policy, we have  $T_1^a v(x) \geq T_1^0 v(x)$  for every action  $a \in (0, a_{\max})$ , that is  $\mu_1(a)[v(x) - v(x - e_1 + e_2)] \leq c_1(a)$ . Taking the above equation into the inequality, we can get  $c_1(a)/\mu_1(a) \geq c_1(a_{\max})/\mu_1(a_{\max})$  for all  $a \in (0, a_{\max})$  which is contradicted against condition (2). Hence the optimal policy for station 1 is unique.

To prove part (ii), we consider the optimal policy  $a$  in station 1 service resource allocation. We use the contradiction method and assume that there exists a state  $x \in E$  for which the optimal policy  $a \in \arg T_1 v(x)$  in station 1 satisfies  $a \in (0, a_{\max})$ . For any  $\varepsilon > 0$ , we have

$$\begin{aligned}
T_1^{a+\varepsilon} v(x) - T_1^a v(x) &= [\mu_1(a + \varepsilon) - \mu_1(a)][v(x - e_1 + e_2) - v(x)] \\
&+ c_1(a + \varepsilon) - c_1(a) \geq 0,
\end{aligned} \tag{14}$$

which implies that

$$v(x) - v(x - e_1 + e_2) \leq \frac{c_1(a + \varepsilon) - c_1(a)}{\mu_1(a + \varepsilon) - \mu_1(a)}. \quad (15)$$

Since the function  $c_1(a)/\mu_1(a)$  is strict decreasing, we get  $(c_1(a + \varepsilon) - c_1(a))/(\mu_1(a + \varepsilon) - \mu_1(a)) < c_1(a)/\mu_1(a)$  so that  $v(x) - v(x - e_1 + e_2) < c_1(a)/\mu_1(a)$ . Because the action  $a$  is the optimal policy for station 1 in state  $x$ , we have  $T_1^a v(x) \leq T_1^0 v(x)$ ; that is,

$$\begin{aligned} \mu_1(a) v(x - e_1 + e_2) + [\mu_1(a_{\max}) - \mu_1(a)] v(x) \\ + c_1(a) \leq \mu_1(a_{\max}) v(x). \end{aligned} \quad (16)$$

So we have  $v(x) - v(x - e_1 + e_2) \geq c_1(a)/\mu_1(a)$  which is a contradiction with the above result  $v(x) - v(x - e_1 + e_2) < c_1(a)/\mu_1(a)$ . Hence the optimal policy for station 1 is  $\arg T_1 v(x) = \{0, a_{\max}\}$ ; that is, the optimal policy is a bang-bang control policy.  $\square$

*Remark 7.* From the above theorem we can conclude that, under some conditions, the optimal is a bang-bang control policy. We try to give intuitive interpretations to these conditions and results, which would help us to understand the theorem intuitively. For the conditions in Theorem 4 (ii), it is clear that  $c_1(a)/\mu_1(a)$  represents the expected service cost for one customer in station 1 under policy  $a$ . Since the function  $c_1(a)/\mu_1(a)$  is strict decreasing for all  $a \in (0, a_{\max})$  which yields that the policy  $a = 1$  is optimal for every customer service cost in station 1. For the total average cost of the system, we can regard it as the average cost per unit time since every customer must be processed in each station. While for the state  $x = (0, 0)$ , it is obvious that no service resource should be allocated in two stations.

*Remark 8.* By the proof of Theorem 6, we know that the results in Theorem 6 are held for both partial and full information cases. In addition, the conditions in Theorem 6 are a bit complex. We give the corresponding looser conditions for Theorem 6 (1) as follows. The optimal policy is unique if the functions  $c'_1(a)/\mu'_1(a)$  and  $c'_2(b)/\mu'_2(b)$  are monotonous and  $c_1(a)/\mu_1(a)$  and  $c_2(b)/\mu_2(b)$  are nondecreasing on  $a \in A$ ,  $b \in B$ , such as the case  $\mu_i(a) = 2a^2$ ,  $c_i(a) = 10a^3$ ,  $i = 1, 2$ .

## 5. Numerical Examples

For the full information case, the corresponding results and numerical example have been investigated in [14, 17]. In this section, we conduct numerical experiments under different parameter settings to demonstrate the main results obtained in this paper for the partial information case. On one hand, these examples provide direct insight into how the change of the system state may impact the optimal resource allocation parameters  $(a^*, b^*)$ . On the other hand, the numerical experiments and Figures 2, 3, 4, and 5 provide the direct support for the results about the structure of the optimal resource allocation policy obtained in the above section. The following experiments are made for the case of

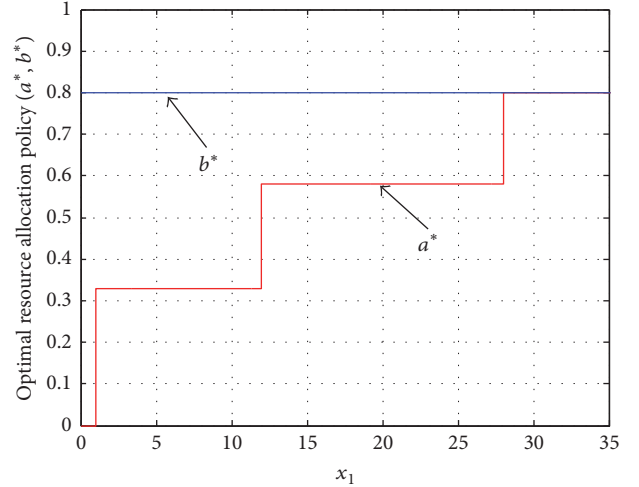


FIGURE 2: Optimal resource allocation policy  $(a^*, b^*)$  versus  $x_1$  for given  $x_2 = 12$ .

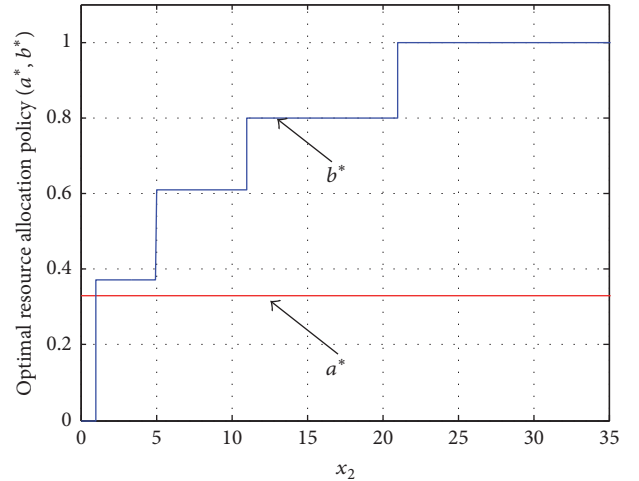


FIGURE 3: Optimal resource allocation policy  $(a^*, b^*)$  versus  $x_2$  for given  $x_1 = 10$ .

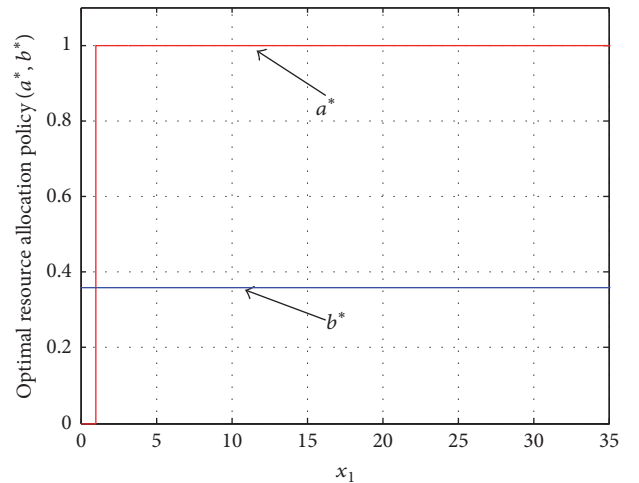


FIGURE 4: Optimal resource allocation policy  $(a^*, b^*)$  versus  $x_1$  for given  $x_2 = 8$ .

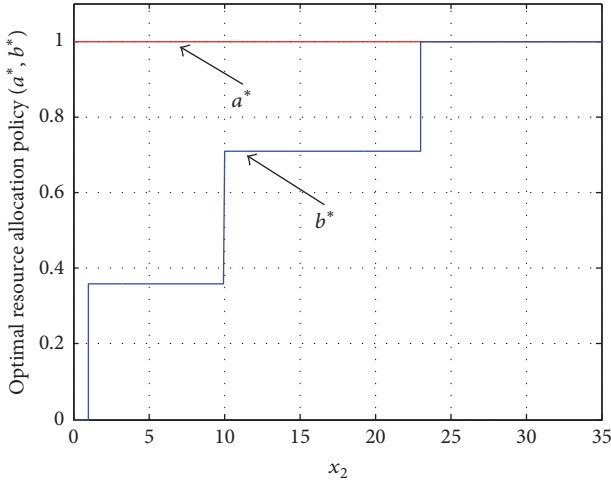


FIGURE 5: Optimal resource allocation policy  $(a^*, b^*)$  versus  $x_2$  for given  $x_1 = 20$ .

$A = B = [0, 1]$ . As is shown in the figures, we can make the following observations.

From Figures 2 and 3, we present numerical results of the optimal policy for the case with  $\lambda = 1.3$ ,  $h_1 = 1.5$ ,  $h_2 = 2$ ,  $\mu_i(a) = 2a^2$ ,  $c_i(a) = 10a^3$ ,  $i = 1, 2$ . As can be seen from Figure 2, the optimal resource allocation policy  $a^*$  increases as the number of customers in station 1 increases, which shows a staircase-like increasing pattern. This phenomenon is consistent with the results in Theorem 2, while the optimal policy  $b^*$  for station 2 remains constant for varied value of  $x_1$ . Meanwhile Figure 3 shows that the optimal resource allocation policy  $a^*$  remains constant and  $b^*$  also shows a staircase-like increasing pattern with the number of customers in station 2 increasing. Moreover, it is noted in these two figures that the line graph of the optimal policy  $a^*$  for station 1 is always under the line graph of the optimal policy  $b^*$  for station 2. This is easy to explain from the results in Theorem 4 that the conditions in this numerical experiment satisfy Theorem 4.

In Figures 4 and 5, we describe the characteristics of the optimal policy for the case with  $\lambda = 1.1$ ,  $h_1 = h_2 = 1$ ,  $\mu_1(a) = 2a^3$ ,  $c_1(a) = 2a^2$ ,  $\mu_2(b) = 3b^2$ ,  $c_2(b) = 4b^3$ . From Figure 4, we find that the optimal policy for station 1 is  $a^* = 0$  if  $x_1 = 0$  otherwise  $a^* = 1$ , which shows that the optimal policy for station 1 has a bang-bang control type, while the optimal policy  $b^*$  for station 2 remains constant. As it is observed from Figure 5, the optimal policy  $a^*$  for station 1 always equals 1, which also belongs to the bang-bang control policy. The optimal policy  $b^*$  for station 2 shows a staircase-like increasing pattern with the number of customers in station 2 increasing. These figures provide a direct support for the results in Theorem 6 since the functions  $\mu_1(a) = 2a^3$ ,  $c_1(a) = 2a^2$  in this numerical experiment follow the conditions in Theorem 6 (ii).

## 6. Conclusion

In this paper, we have analyzed the optimal resources allocation control policy of a tandem queueing system with the

general service cost and service rate functions. Applying the queueing system and MDP theories, we not only give some traditional properties of the relative value function and optimal policy but also derive the conditions under which the optimal policy is unique and has a bang-bang control property, which has not been studied before our work. In particular, we have provided the relationship between two stations' optimal policies, which can give the manager basal insight into the structure of optimal policy information to improve decision-making of the system.

From the above results, there arise some interesting extensions of the model which we may study in the near future. One possible change is to consider the tandem queueing system with retrial or feedback customers which will make the model more useful in practical system. Another way to extend the model is to apply the semi-Markov decision processes to consider the queueing system in which the service time of a customer is a general distribution. Furthermore, in practice, the production systems are often likely to be burdened by mixed uncertainties of both randomness and fuzziness; the study of the optimal control of the tandem queueing system with fuzziness may provide more precise information to managers, which is also an interesting topic for future research.

## Appendix

### A. Proof of Lemma 1

#### The Proof of Lemma 1

*Proof.* To prove Lemma 1 (i), the proof is done by induction on  $n$  in  $v_n$ . Define  $v_0(x) = 0$  for all state  $x \in E$ . This function obviously satisfies (i). Now, we assume that (i) holds for the function  $v_n(x)$ ,  $x \in E$ , and some  $n \in N$ . We should prove that  $v_{n+1}(x)$  satisfies the nondecreasing property as well. Then for  $i = 1$ , we can get

$$\begin{aligned} & v_{n+1}(x + e_1) - v_{n+1}(x) \\ &= \lambda [v_n(x + 2e_1) - v_n(x + e_1)] + h_1 \\ &+ \sum_{i=1,2} T_i v_n(x + e_1) - \sum_{i=1,2} T_i v_n(x). \end{aligned} \quad (\text{A.1})$$

The second term of the right-hand side is obviously positive.

Let  $(a \in \arg T_1 v(x + e_1), b \in \arg T_2 v_n(x + e_1))$  be an arbitrary optimal policy for two stations in state  $x + e_1$ . Then

$$\begin{aligned} & \sum_{i=1,2} T_i v_n(x + e_1) - \sum_{i=1,2} T_i v_n(x) \geq \mu_1(a) \\ & \cdot [v_n(x + e_2) - v_n(x + e_2 - e_1)] + \mu_2(b) \\ & \cdot [v_n(x - e_2 + e_1) - v_n(x - e_2)] \\ & + [\mu_1(a_{\max}) - \mu_1(a) + \mu_2(b_{\max}) - \mu_2(b)] \\ & \cdot [v_n(x + e_1) - v_n(x)] \geq 0. \end{aligned} \quad (\text{A.2})$$

Therefore, Lemma 1 (i) holds by induction for any  $n$ ,  $v(x)$  is a nondecreasing function. Lemma 1 (i) for  $i = 2$  can be proved in a similar manner.

To prove Lemma 1 (ii), the proof is similar to the proof of Lemma 1 (i). Define  $v_0(x) = 0$  for all state  $x \in E$ . This function obviously satisfies (ii). Now, we assume that (ii) holds for function  $v_n(x)$ ,  $x \in E$  and some  $n \in N$ . We should prove that  $v_{n+1}(x)$  satisfies Lemma 1 (ii):

$$\begin{aligned} & v_{n+1}(x - e_1 + e_2) - v_{n+1}(x - e_2) \\ &= \lambda [v_n(x + e_2) - v_n(x + e_1 - e_2)] + 2h_2 - h_1 \\ &+ \sum_{i=1,2} T_i v_n(x - e_1 + e_2) - \sum_{i=1,2} T_i v_n(x - e_2). \end{aligned} \quad (\text{A.3})$$

Since condition  $2h_2 \geq h_1$  holds, the second term of the right-hand side is obviously positive.

Let  $(a \in \arg T_1 v(x - e_1 + e_2), b \in \arg T_2 v(x - e_1 + e_2))$  be an arbitrary optimal policy for two stations in state  $x - e_1 + e_2$ . Then

$$\begin{aligned} & \sum_{i=1,2} T_i v_n(x - e_1 + e_2) - \sum_{i=1,2} T_i v_n(x - e_2) \geq \mu_1(a) \\ & \cdot [v_n(x - 2e_1 + 2e_2) - v_n(x - e_1)] + \mu_2(b) \\ & \cdot [v_n(x - e_1) - v_n(x - 2e_2)] \\ & + [\mu_1(a_{\max}) - \mu_1(a)] \\ & \cdot [v_n(x - e_1 + e_2) - v_n(x - e_2)] \\ & + [\mu_2(b_{\max}) - \mu_2(b)] \\ & \cdot [v_n(x - e_1 + e_2) - v_n(x - e_2)] \geq 0. \end{aligned} \quad (\text{A.4})$$

Therefore, Lemma 1 (ii) holds by induction for any  $n$ ; we have  $v(x - e_1 + e_2) \geq v(x - e_2)$  for all  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ . Lemma 1 (iii) can be proved in a similar manner.  $\square$

## B. Proof of Lemma 3 and Theorem 2

*The Proof of Lemma 3 (i) and Theorem 2 (i)*

*Proof.* To prove Lemma 3 (i), we assume that Lemma 3 (i) for function  $v_n(x)$ ,  $x \in E$ , and some  $n \in N$  holds. Then we need to prove that Lemma 3 (i) for  $n + 1$  also holds. We have

$$\begin{aligned} & v_{n+1}(x + e_2) - 2v_{n+1}(x) + v_{n+1}(x - e_2) \\ &= \lambda [v_n(x + e_2 + e_1) - 2v_n(x + e_1) \\ &+ v_n(x + e_1 - e_2)] + \sum_{i=1,2} T_i v_n(x + e_2) \\ &- 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2) \\ &\geq \sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) \\ &+ \sum_{i=1,2} T_i v_n(x - e_2). \end{aligned} \quad (\text{B.1})$$

The inequality holds by the induction hypothesis. The optimal policy of station 1 is only dependent on the number of customers in station 1 and the states  $x + e_2, x, x - e_2$  have same first entry  $x_1$ . Hence, they have the same optimal policy in station 1. We assume that  $a \in \arg T_1 v(x), b_1 \in \arg T_2 v(x + e_2), b_2 \in \arg T_2 v(x - e_2)$ :

$$\begin{aligned} & \sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2) \\ &\geq \mu_1(a) [v_n(x - e_1 + 2e_2) - 2v_n(x - e_1 + e_2) \\ &+ v_n(x - e_1)] + [\mu_1(a_{\max}) - \mu_1(a)] [v_n(x + e_2) \\ &- 2v_n(x) + v_n(x - e_2)] + [\mu_2(b_1) - \mu_2(b_2)] \\ &\cdot [v_n(x) - v_n(x - e_2)] + \mu_2(b_2) [v_n(x) \\ &- 2v_n(x - e_2) + v_n(x - 2e_2)] + [\mu_2(b_{\max}) \\ &- \mu_2(b_1)] [v_n(x + e_2) - v_n(x)] + [\mu_2(b_{\max}) \\ &- \mu_2(b_2)] [v_n(x - e_2) - v_n(x)] = \mu_1(a) \\ &\cdot [v_n(x - e_1 + 2e_2) - 2v_n(x - e_1 + e_2) \\ &+ v_n(x - e_1)] + [\mu_1(a_{\max}) - \mu_1(a)] [v_n(x + e_2) \\ &- 2v_n(x) + v_n(x - e_2)] + \mu_2(b_2) [v_n(x) \\ &- 2v_n(x - e_2) + v_n(x - 2e_2)] + [\mu_2(b_{\max}) \\ &- \mu_2(b_1)] [v_n(x + e_2) - 2v_n(x) + v_n(x - e_2)] \\ &\geq 0. \end{aligned} \quad (\text{B.2})$$

The first inequality follows by taking a potentially suboptimal action in the second term of  $\sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2)$ . The equality follows by rearranging the terms. The last inequality follows by the induction hypothesis. Hence, we have  $v(x + e_2) - 2v(x) + v(x - e_2) \geq 0$ .

To prove Theorem 2 (i), let  $(b_1 \in \arg T_2 v(x + e_2), b_2 \in \arg T_2 v(x))$  be an optimal policy for station 2 in states  $x + e_2, x$ , respectively. The proof is done by contradiction. Suppose that  $b_1 < b_2$ ; then

$$\begin{aligned} & T_2^{b_1} v(x) - T_2^{b_2} v(x) \\ &= [\mu_2(b_2) - \mu_2(b_1)] [v(x) - v(x - e_2)] \\ &- [c_2(b_2) - c_2(b_1)] \geq 0. \end{aligned} \quad (\text{B.3})$$

For Lemma 1 (i) and  $\mu_2(b_2) - \mu_2(b_1) > 0$ , we have

$$\begin{aligned} & T_2^{b_1} v(x + e_2) - T_2^{b_2} v(x + e_2) \\ &= [\mu_2(b_2) - \mu_2(b_1)] [v(x + e_2) - v(x)] \\ &- [c_2(b_2) - c_2(b_1)] \\ &> [\mu_2(b_2) - \mu_2(b_1)] [v(x) - v(x - e_2)] \\ &- [c_2(b_2) - c_2(b_1)] \geq 0. \end{aligned} \quad (\text{B.4})$$



However, this implies that  $b_1$  is not an optimal policy for station 2 in state  $x + e_2$ . Hence  $b_1 \geq b_2$ .  $\square$

*The Proof of Lemma 3 (ii) and Theorem 2 (ii)*

*Proof.* To prove Lemma 3 (ii), we assume that Lemma 3 (ii) holds for function  $v_n(x)$ ,  $x \in E$ , and some  $n \in N$ . Then we need to prove that Lemma 3 (ii) for  $n+1$  also holds. Using the optimality equation, we have

$$\begin{aligned}
& v_{n+1}(x + e_1 - e_2) - 2v_{n+1}(x) + v_{n+1}(x - e_1 + e_2) \\
&= \lambda [v_n(x + 2e_1 - e_2) - 2v_n(x + e_1) + v_n(x + e_2)] \\
&+ \sum_{i=1,2} T_i v_n(x + e_1 - e_2) - 2 \sum_{i=1,2} T_i v_n(x) \\
&+ \sum_{i=1,2} T_i v_n(x - e_1 + e_2) \\
&\geq \sum_{i=1,2} T_i v_n(x + e_1 - e_2) - 2 \sum_{i=1,2} T_i v_n(x) \\
&+ \sum_{i=1,2} T_i v_n(x - e_1 + e_2) \\
&= T_1 v_n(x + e_1 - e_2) - 2T_1 v_n(x) \\
&+ T_1 v_n(x - e_1 + e_2) + T_2 v_n(x + e_1 - e_2) \\
&- 2T_2 v_n(x) + T_2 v_n(x - e_1 + e_2).
\end{aligned} \tag{B.5}$$

The inequality above holds by the induction hypothesis. Now, we assume that  $a_1 \in \arg T_1 v(x + e_1 - e_2)$ ,  $b_1 \in \arg T_2 v(x + e_1 - e_2)$ ,  $a_2 \in \arg T_1 v(x - e_1 + e_2)$ ,  $b_2 \in \arg T_2 v(x - e_1 + e_2)$ . Then, we get

$$\begin{aligned}
& T_1 v_n(x + e_1 - e_2) - 2T_1 v_n(x) + T_1 v_n(x - e_1 + e_2) \\
&\geq \mu_1(a_1) [v_n(x) - v_n(x - e_1 + e_2)] + \mu_1(a_2) \\
&\cdot [v_n(x - 2e_1 + 2e_2) - v_n(x - e_1 + e_2)] \\
&+ [\mu_1(a_{\max}) - \mu_1(a_1)] [v_n(x + e_1 - e_2) - v_n(x)] \\
&+ [\mu_1(a_{\max}) - \mu_1(a_2)] [v_n(x - e_1 + e_2) - v_n(x)] \\
&= \mu_1(a_2) \\
&\cdot [v_n(x - 2e_1 + 2e_2) - 2v_n(x - e_1 + e_2) + v_n(x)] \\
&+ [\mu_1(a_{\max}) - \mu_1(a_1)] \\
&\cdot [v_n(x + e_1 - e_2) - 2v_n(x) + v_n(x - e_1 + e_2)] \\
&\geq 0.
\end{aligned} \tag{B.6}$$

The first inequality follows by taking a potentially suboptimal action in the second term of the operator  $T_1 v_n(x + e_1 - e_2) - 2T_1 v_n(x) + T_1 v_n(x - e_1 + e_2)$ . The equality follows by rearranging the terms. The last inequality follows by the induction hypothesis:

$$\begin{aligned}
& T_2 v_n(x + e_1 - e_2) - 2T_2 v_n(x) + T_2 v_n(x - e_1 + e_2) \\
&\geq \mu_2(b_1) [v_n(x + e_1 - 2e_2) - v_n(x - e_2)] \\
&+ \mu_2(b_2) [v_n(x - e_1) - v_n(x - e_2)] \\
&+ [\mu_2(b_{\max}) - \mu_2(b_1)] [v_n(x + e_1 + e_2) - v_n(x)] \\
&+ [\mu_2(b_{\max}) - \mu_2(b_2)] [v_n(x - e_1 + e_2) - v_n(x)] \\
&= \mu_2(b_2) \\
&\cdot [v_n(x + e_1 - 2e_2) - 2v_n(x - e_2) + v_n(x - e_1)] \\
&+ [\mu_2(b_{\max}) - \mu_2(b_2)] \\
&\cdot [v_n(x + e_1 - e_2) - 2v_n(x) + v_n(x - e_1 + e_2)] \\
&+ [\mu_2(b_1) - \mu_2(b_2)] \\
&\cdot [v_n(x + e_1 - 2e_2) - v_n(x + e_1 - e_2)] \geq 0.
\end{aligned} \tag{B.7}$$

The first inequality follows by taking a potentially suboptimal action in the second term of the operator above. The equality follows by rearranging the terms. The last one follows by the induction hypothesis and, because of Theorem 2 (i), we know that  $b_1 \leq b_2$ . So that we have  $\mu_2(b_1) - \mu_2(b_2) \leq 0$ . From Lemma 1, we know that  $v_n(x + e_1 - 2e_2) - v_n(x + e_1 - e_2) \leq 0$ . Thus, we derive that  $[\mu_2(b_1) - \mu_2(b_2)][v_n(x + e_1 - 2e_2) - v_n(x + e_1 - e_2)] \geq 0$ . Therefore, the last inequality is taken.

To prove Theorem 2 (ii), let  $(a_1 \in \arg T_1 v(x + e_1 - e_2))$ ,  $(a_2 \in \arg T_1 v(x))$  be an optimal policy for station 2 in states  $x + e_1 - e_2$ ,  $x$ , respectively. The proof is done by contradiction. Suppose that  $a_1 < a_2$ ; then

$$\begin{aligned}
& T_1^{a_1} v(x) - T_1^{a_2} v(x) \\
&= [\mu_1(a_2) - \mu_1(a_1)] [v(x - e_1 + e_2) - v(x)] \\
&- [c_1(a_2) - c_1(a_1)] \geq 0.
\end{aligned} \tag{B.8}$$

From Lemma 1 (ii) above and  $\mu_1(a_2) - \mu_1(a_1) > 0$ , we have

$$\begin{aligned}
& T_1^{a_1} v(x + e_1 - e_2) - T_1^{a_2} v(x + e_1 - e_2) \\
&= [\mu_1(a_2) - \mu_1(a_1)] [v(x) - v(x + e_1 - e_2)] \\
&- [c_1(a_2) - c_1(a_1)] \\
&\geq [\mu_1(a_2) - \mu_1(a_1)] [v(x - e_1 + e_2) - v(x)] \\
&- [c_1(a_2) - c_1(a_1)] \geq 0.
\end{aligned} \tag{B.9}$$

This implies that  $a_1$  is not an optimal policy for station 1 in state  $x + e_1 - e_2$ , which is with the assumption  $a_1 \in \arg T_1 v(x + e_1 - e_2)$ . Hence  $a_1 \geq a_2$ .

Since the optimal policy of station 1 is dependent only on the number of customers in station 1, and the states  $x + e_1$ ,  $x + e_1 - e_2$  have the same first entry  $x_1 + 1$ , they have the same optimal policy  $a_1$  in station 1, that is,  $a_1 \in \arg T_1 v(x + e_1)$ . Thus we get that if  $a_1 \in \arg T_1 v(x + e_1)$ ,  $a_2 \in \arg T_1 v(x)$  hold, then we have  $a_1 \geq a_2$  for all  $x = (x_1, x_2) \in E$ .  $\square$

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research is partially supported by the National Natural Science Foundation of China (11671404 and 11271373) and the Fundamental Research Funds for the Central Universities of Central South University (2017zzts061).

## References

- [1] I. Koutsopoulos, Resource allocation issues in broadband wireless networks with OFDM signaling, 2002.
- [2] A. Kuiper and M. Mandjes, "Appointment scheduling in tandem-type service systems," *OMEGA*, vol. 57, no. 2, pp. 145–156, 2015.
- [3] C.-H. Wu, J. T. Lin, and W.-C. Chien, "Dynamic production control in a serial line with process queue time constraint," *International Journal of Production Research*, vol. 48, no. 13, pp. 3823–3843, 2010.
- [4] B. Kim and J. Kim, "Optimal admission control for two station tandem queues with loss," *Operations Research Letters*, vol. 42, no. 4, pp. 257–262, 2014.
- [5] D. F. Silva, B. Zhang, and H. Ayhan, "Optimal admission control for tandem loss systems with two stations," *Operations Research Letters*, vol. 41, no. 4, pp. 351–356, 2013.
- [6] Z. Rosberg, P. P. Varaiya, and J. C. Walrand, "Optimal control of service in tandem queues," *IEEE Transactions on Automatic Control*, vol. 27, no. 3, pp. 600–610, 1982.
- [7] H.-S. Ahn, I. Duenyas, and M. E. Lewis, "Optimal control of a two-stage tandem queueing system with flexible servers," *Probability in the Engineering and Informational Sciences*, vol. 16, no. 4, pp. 453–469, 2002.
- [8] R. Arumugam, M. E. Mayorga, and K. M. Taaffe, "Inventory based allocation policies for flexible servers in serial systems," *Annals of Operations Research*, vol. 172, pp. 1–23, 2009.
- [9] J. M. Smith and R. Barnes, "Optimal server allocation in closed finite queueing networks," *Flexible Services and Manufacturing Journal*, vol. 27, no. 1, pp. 58–85, 2015.
- [10] S. M. Iravani, V. Krishnamurthy, and G. H. Chao, "Optimal server scheduling in nonpreemptive finite-population queueing systems," *Queueing Systems*, vol. 55, no. 2, pp. 95–105, 2007.
- [11] R. Yang, S. Bhulai, and R. van der Mei, "Structural properties of the optimal resource allocation policy for single-queue systems," *Annals of Operations Research*, vol. 202, pp. 211–233, 2013.
- [12] R. Yang, S. Bhulai, and R. van der Mei, "Optimal resource allocation for multiqueue systems with a shared server pool," *Queueing Systems*, vol. 68, no. 2, pp. 133–163, 2011.
- [13] R. R. Weber and J. Stidham, "Optimal control of service rates in networks of queues," *Advances in Applied Probability*, vol. 19, no. 1, pp. 202–218, 1987.
- [14] M. H. Veatch and L. M. Wein, "Monotone control of queueing networks," *Queueing Systems*, vol. 12, no. 3–4, pp. 391–408, 1992.
- [15] M. E. Mayorga, K. M. Taaffe, and R. Arumugam, "Allocating flexible servers in serial systems with switching costs," *Annals of Operations Research*, vol. 172, pp. 231–242, 2009.
- [16] L. Xia and B. Shihada, "Max-min optimality of service rate control in closed queueing networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 1051–1056, 2013.
- [17] L. Xia, D. Miller, Z. Zhou, and N. Bambos, "Service rate control of tandem queues with power constraints," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5111–5123, 2017.
- [18] E. Morozov and B. Steyaert, "Stability analysis of a two-station cascade queueing network," *Annals of Operations Research*, vol. 202, no. 1, pp. 135–160, 2013.
- [19] G. Koole, *Monotonicity in Markov Reward and Decision Chains: Theory and Applications*, vol. 1, Now Publishers Inc, Breda, Netherlands, 2007.
- [20] G. Koole, "Structural results for the control of queueing systems using event-based dynamic programming," *Queueing Systems*, vol. 30, no. 3–4, pp. 323–339, 1998.
- [21] D. L. Kaufman, H.-S. Ahn, and M. E. Lewis, "On the introduction of an agile, temporary workforce into a tandem queueing system," *Queueing Systems*, vol. 51, no. 1–2, pp. 135–171, 2005.
- [22] E. B. Çil, E. L. Örmeci, and F. Karaesmen, "Effects of system parameters on the optimal policy structure in a class of queueing control problems," *Queueing Systems*, vol. 61, no. 4, pp. 273–304, 2009.
- [23] H. C. Tijms, *Stochastic Models: An Algorithmic Approach*, vol. 303 of *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley & Sons Inc, Hoboken, NJ, USA, 1994.
- [24] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Hoboken, NJ, USA, 2014.
- [25] Y. Aviv and A. Federgruen, "The value iteration method for countable state Markov decision processes," *Operations Research Letters*, vol. 24, no. 5, pp. 223–234, 1999.
- [26] L. I. Sennott, *Stochastic Dynamic Programming and The Control of Queueing Systems*, Wiley Series in Probability and Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, NY, USA, 1999.



# Hindawi

Submit your manuscripts at  
<https://www.hindawi.com>

