Service Rate Control of Tandem Queues With Power Constraints

Li Xia, *Senior Member, IEEE*, Daniel Miller, Zhengyuan Zhou, and Nicholas Bambos

***Abstract*—In this paper, we study the optimal control of service rates of a tandem queue with power constraints. The service rate of a server is determined by the power allocated to that server. The total power of the system is fixed. The system cost is comprised of two parts, the holding cost reflecting the congestion of queues and the operating cost reflecting the power consumed at servers. The optimization objective is to find the optimal power allocation policy among servers, which can minimize the system average cost. We formulate this problem as a Markov decision process with a constrained action space. Sensitivity-based optimization theory is applied to study this problem. The necessary and sufficient condition of optimal service rates, and the optimality of the vertexes of the feasible domain are derived when the operating cost has a linear or concave form. An iterative algorithm is further developed to find the optimal service rates. This algorithm may work well even when the cost function has a general form. The extension to general tandem queues with many servers is also studied. Finally, we conduct numerical experiments under different parameter settings to demonstrate the main idea of this paper.**

***Index Terms*—Service rate control, tandem queue, power constraint, sensitivity-based optimization, convex optimization.**

I. Introduction

Service rate control is a classical optimization problem of queueing systems. The objective is to identify the optimal policy that determines the optimal service rates of every server at every state such that the system average performance can be optimum. The service rate control problem can be applied to study many practical problems, such as the resource allocation in communication, computer, transportation, inventory, healthcare, etc. [11], [12], [24].

The service rate control problem has been extensively studied in the literature, from a single-server model to a multi-server model. Crabill [9] studies the service rate control problem of an M/M/1 queue with K possible service rates and it is shown that the optimal service rate is monotone (nondecreasing) in the queue length under the assumption of nondecreasing cost structure. For an M/G/1 queue with a linear cost structure, it is proved that the N-policy is optimal: When the queue length is smaller than N, shutdown the server, otherwise run the server at the maximal service rate [14], [30]. Other similar policies, such as D-policy (D is the cumulative service time) [4] and T-policy (T is the length of idle period) [15], are also studied. The service rate control problem of a two-node cyclic network is studied under the background of flow control in computer networks [18]. The monotone structure of optimal service rates in a general cyclic queue is further obtained [26]. For a general Jackson network, it is proved that when the cost function has a linear structure, the system average performance is monotone in the queue length and the bang–bang control is optimal [20], [31]. In a recent study, it is further derived that the structure of the cost function in Jackson networks can be extended to a concave form while the optimality of bang–bang control remains valid [29]. There are also many other works that study the service rate control problem from the viewpoint of game theory and economics [6], [10], [13], which is motivated by the pioneering work [21]. The optimization of state-dependent service rate and arrival rate is investigated by [3], where the optimal pricing problem is also studied for an M/M/1 queue. Furthermore, the structure of equilibrium and the price of anarchy of a game of service rate control are studied for closed Jackson networks [28].

There are also many studies about the service rate control for a specific tandem queue (or called series queue). For a tandem queue with two servers, Rosberg *et al.* [23] study a special case where the optimization variable is the service rate of the first server and the service rate of the second server is a constant. If the system only has the holding cost, it is shown that the optimal state-dependent service rate of server 1 has a structure of bang–bang control determined by a switch function of the queue length at server 2. For a tandem queue with two servers with a nondecreasing cost function, Liu *et al.* [19] prove that the optimal service rates also have a nondecreasing structure with respect to the queue length. Server allocation problem can be viewed as a special case of service rate control, where the servers

Manuscript received November 24, 2016; revised February 22, 2017; accepted February 28, 2017. Date of publication March 3, 2017; date of current version September 25, 2017. This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB0901902, in part by the National Natural Science Foundation of China under Grant 61573206, Grant 61203039, and Grant U1301254, and in part by the Suzhou-Tsinghua Innovation Leading Action Project. Recommended by Associate Editor L. H. Lee.

L. Xia is with the Center for Intelligent and Networked Systems, Department of Automation, TNList, Tsinghua University, Beijing 100084, China (e-mail: xial@tsinghua.edu.cn).

D. Miller, Z. Zhou, and N. Bambos are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: danielrm@stanford.edu; zyzhou@stanford.edu; bambos@stanford.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2017.2678109

are dynamically allocated among service stations [2], [25]. For a two-stage tandem queue with flexible servers, Kaufman *et al.* [16] also show that the optimal number of servers to have in the system is nondecreasing in the queue length. The optimality of a server allocation policy similar to the bang–bang control is also studied when the flexible servers are collaborative or noncollaborative [1]. Of course, there are certainly numerous other studies about the service rate control of queueing systems in the literature. However, in most of the prior studies, it is assumed that the service rate can be adjusted in a certain value domain and there is no resource constraint during the allocation process of service rates.

In this paper, we study the service rate control problem of tandem queues with a resource constraint. The contributions of this paper are threefold. First, different from most of the unconstrained service rate control problems in the literature, one key novel feature of our problem formulated in Section II is the total resource constraint for service rate allocation: When we adjust the service rate for each individual server, the sum of service rates must not exceed the total resource available. Consequently, the control policy is required to satisfy the condition that the sum of all the service resources (or service rates equivalently) allocated to every server is less than or equal to a given resource constraint. This is an important problem with wide applicability. For example, in a computer facility, the total computation resource is fixed and a commonly encountered task is to optimally allocate the resources among different jobs running in parallel. Such problems are called the scheduling of *pipelines* in computer systems or *service function chaining* recently proposed in computer networks [17]. Such service chaining is an important paradigm in cloud computing where the rich computational resources are provided by machines in a data center (or data centers) and the jobs to be processed come from external paying customers. These jobs tend to be computationally intensive and in general require multiple stages of processing before being completed. Google Cloud Dataflow provides one such example, where it is a fully managed pipelined service that takes an incoming data processing job (either in stream or in batch) and breaks it into multiple stages of the pipeline. More specifically, we consider a tandem queue with a Poisson arrival and exponential service times. The service rate of each server is *state-dependent*, i.e., we can adjust the service rate when the system state changes. The service resource (such as the power) allocated to each server determines its service rate. The total resource of the whole system is fixed. The system cost is comprised of two parts. One is the *holding cost* that reflects the cost related to the system state, including a cost for congestion or a fixed cost of the system. The other is the *operating cost* that reflects the cost for resource consumption at that server.

Second, inspired by [7] and [8], we take a novel approach of sensitivity-based optimization to identify the optimal service rate control policy such that the long-run average cost of the tandem queue is minimized. Although the traditional method of dynamic programming [5], [22] can also be used to study this problem, it is difficult to derive optimality properties. The dynamic programming is based on the Bellman optimality equation, which, despite an elegant conceptual framework, suffers from the well-known curse of dimensionality: It is difficult to apply the Bellman optimality equation to study the problem structure because the equation has high dimensions and it becomes too complicated to handle when the system dimension grows. As a comparison, we use the sensitivity-based optimization method and derive interesting optimality properties in Section III. Specifically, we derive a difference formula that describes the relation between the system average cost and the service rates adopted. With the difference formula, we can clearly see how the system average cost changes when the service rates change. We prove that when the operating cost has a concave structure, the optimal service rate can be reached at a vertex of multidimensional polyhedron of feasible domain. When the operating cost has a linear structure, we prove that the system average cost is monotone in the service rate and the optimal service rates also can be reached at a vertex of feasible domain. When the operating cost has a convex structure, we can decompose the original problem into a series of subproblems that are convex and easy to solve. A selection rule of optimal service rates is also derived, which is simply based on the value distribution of variables about the difference of performance potentials at neighbor states.

Third, leveraging the above analytical results, we further propose an iterative optimization algorithm in Section IV to find the optimal service rate control policy. This algorithm is of policy iteration type and it decomposes the optimization problem into a series of subproblems that are easy to solve in different scenarios. We then conduct simulation experiments in Section V to demonstrate the interesting behaviors of this optimization problem under different parameter settings. It is verified that our approach works well for various general forms of cost functions, while most of the studies in the literature require the cost function to have a special form, such as nondecreasing or convex, etc. Our approach gives a promising way to study such category of optimization problems in queueing systems.

Fig. 1. Tandem queue with two servers.

II. Problem Formulation

Consider a tandem queue with two servers, which is illustrated in Fig. 1. The customer arrival is a Poisson process with rate λ. The service time of servers is exponentially distributed. We denote the system state as $\boldsymbol{n} = (n_1, n_2)$, where n_1 and n_2 are the number of customers at server 1 and 2, respectively. The state space is denoted as $\mathcal{S}$, i.e., $\mathcal{S} := \{\text{all } \boldsymbol{n} | \ n_1, n_2 = 0, 1, 2, \ldots\}$. The service rate of server i is state-dependent and it is denoted as $\mu_{i,\boldsymbol{n}}, i = 1, 2, \boldsymbol{n} \in \mathcal{S}$. We further denote the vector of service rates at state $\boldsymbol{n}$ as

$$\boldsymbol{\mu_n} := (\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}}). \tag{1}$$

The service discipline is first come first serve and the waiting room is assumed infinite.

The system cost is comprised of two parts, the holding cost and the operating cost. The holding cost reflects the penalty for the congestion in the queueing system and possible fixed costs of the system. We denote the holding cost rate function as $\varphi(\boldsymbol{n})$, $\boldsymbol{n} \in \mathcal{S}$. In some literature, $\varphi(\boldsymbol{n})$ is further assumed nondecreasing and decomposable in queue length, such as $\varphi(\boldsymbol{n}) = \sum_{i=1}^{2} n_i + c_s$, where the first part reflects the cost for congestion and the second part c_s is a fixed cost of the system. In this paper, we do not need such assumption, and $\varphi(\boldsymbol{n})$ can be any function. The operating cost reflects the power consumption at the server. The service rate $\mu_{i,\boldsymbol{n}}$ is determined by the power allocated to server i. If more power is allocated to server i, then the service rate $\mu_{i,\boldsymbol{n}}$ is bigger, and so is the operating cost. We denote the operating cost rate function as $\phi(\boldsymbol{\mu_n})$ that is increasing in $\mu_{i,\boldsymbol{n}}$, according to common rationales. Thus, the total cost rate function of the tandem queue is as below:

$$f(\boldsymbol{n}, \boldsymbol{\mu_n}) = \varphi(\boldsymbol{n}) + \phi(\boldsymbol{\mu_n}). \tag{2}$$

In some special cases, we will further specify the form of $\phi(\boldsymbol{\mu_n})$ and assume that the service rate and the operating cost are both linear to the power allocated at the server. That is, we have

$$\phi(\boldsymbol{\mu_n}) = b_1\mu_{1,\boldsymbol{n}} + b_2\mu_{2,\boldsymbol{n}} + c_o \tag{3}$$

where c_o reflects a fixed cost to keep the system operating at a minimal level. Therefore, the total cost rate function of the system is further specified as below:

$$f(\boldsymbol{n}, \boldsymbol{\mu_n}) = \varphi(\boldsymbol{n}) + b_1\mu_{1,\boldsymbol{n}} + b_2\mu_{2,\boldsymbol{n}} + c_o. \tag{4}$$

For most cases considered in this paper, we will first conduct the analysis for the general cost function (2), and then for the specific form (4).

The total power to be allocated is limited. That is, the sum of power at server 1 and 2 should be smaller or equal to a constant. We assume that the service rate is linear to the power allocated to that server. Therefore, a constraint on the total power is equivalent to a constraint on the total service rates. That is, we have the following constraint:

$$\mu_{1,\boldsymbol{n}} + \mu_{2,\boldsymbol{n}} \le U, \qquad \forall\, \boldsymbol{n} \in \mathcal{S} \tag{5}$$

where U is the maximal total service rates reflecting the total power constraint. For the stability of queueing systems, the value of U has to satisfy a *necessary condition* $U > 2\lambda$. On the other hand, it is common to require that the server should run at least with a minimal service rate when the queue is not empty. That is, $\mu_{i,\boldsymbol{n}} \ge u_i$, where u_i is the minimal service rate of server i. Clearly, we have $0 \le u_i \ll U$. Therefore, the value domain of the service rate vector $\boldsymbol{\mu_n}$ is a two-dimensional triangle area illustrated in Fig. 2. We denote the value domain as $\mathcal{D}_U$ and we have

$$\mathcal{D}_U := \{(\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}}) \big| \mu_{1,\boldsymbol{n}} + \mu_{2,\boldsymbol{n}} \le U,\ \mu_{1,\boldsymbol{n}} \ge u_1,\ \mu_{2,\boldsymbol{n}} \ge u_2\}. \tag{6}$$

Note that if $n_i = 0$, we always have $\mu_{i,\boldsymbol{n}} = 0$, where $i = 1, 2$.

We denote the allocation policy of service rates as $\boldsymbol{\mu}$, which is a mapping from the state space $\mathcal{S}$ to the value domain $\mathcal{D}_U$. That is, $\boldsymbol{\mu}$ is a vector composed of elements $\boldsymbol{\mu_n}$, with each element $\boldsymbol{\mu_n} \in \mathcal{D}_U$, $\boldsymbol{n} \in \mathcal{S}$. Denote $\mathfrak{n}(t)$ as the system state at time t. Since the interarrival time and the service times are exponential, $\mathfrak{n}(t)$ is a continuous time Markov process. It is difficult to figure out a necessary and sufficient condition for the stability of the tandem queue. Here we give a sufficient condition as follows. If there exists a constant k and for any state $\boldsymbol{n}$ with $n_i > k$, we always have $\mu_{i,\boldsymbol{n}} > \lambda$, then the tandem queue is stable and the steady-state distribution exists. We denote the steady-state distribution as an infinite dimensional row vector $\boldsymbol{\pi}$, where $\pi(\boldsymbol{n})$ is the element of $\boldsymbol{\pi}$ and it represents the probability that the system stays at $\boldsymbol{n}$ after the system reaches steady state. We further denote $\boldsymbol{f}$ as an infinite dimensional column vector, and its element $f(\boldsymbol{n}, \boldsymbol{\mu_n})$ as the cost function at state $\boldsymbol{n}$. For an ergodic Markov process (even a unichain), we have the long-run average cost of the system as below:

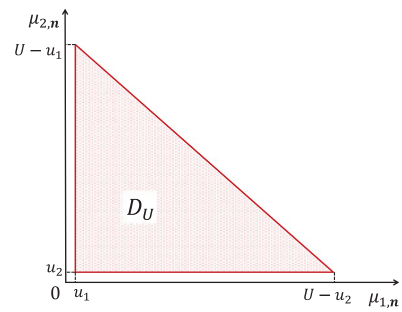

Fig. 2. Value domain (triangle area) $\mathcal{D}_U$ of service rates.

$$\eta = \boldsymbol{\pi f} = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(\mathfrak{n}(t), \boldsymbol{\mu}_{\mathfrak{n}(t)})dt. \tag{7}$$

Note that the values of η, $\boldsymbol{\pi}$, and $\boldsymbol{f}$ are related to the service rate allocation policy $\boldsymbol{\mu}$ and we should denote them as $\eta_{\boldsymbol{\mu}}$, $\boldsymbol{\pi_\mu}$, and $\boldsymbol{f_\mu}$ for the rigorousness of notations. However, to keep the notation concise, we omit the subscript "$\boldsymbol{\mu}$." Similar issues exist for other notations in the rest of the paper.

The infinitesimal generator of the Markov process is denoted as $\boldsymbol{B}$. For a particular state, say $\boldsymbol{n}$, it is easy to verify that the row vector $\boldsymbol{B}(\boldsymbol{n}, :)$ is as follows:

- $B(\boldsymbol{n}, \boldsymbol{n}) = -\lambda - \mu_{1,\boldsymbol{n}} - \mu_{2,\boldsymbol{n}}$;
- $B(\boldsymbol{n}, (n_1+1, n_2)) = \lambda$;
- $B(\boldsymbol{n}, (n_1-1, n_2+1)) = \mu_{1,\boldsymbol{n}}$;
- $B(\boldsymbol{n}, (n_1, n_2-1)) = \mu_{2,\boldsymbol{n}}$;
- $B(\boldsymbol{n}, \boldsymbol{n}') = 0$ for all the other states $\boldsymbol{n}'$ with $n_1, n_2 > 0$.

For the case that n_1 or n_2 equals 0, it is easy to obtain the element of $\boldsymbol{B}$ in a similar way. For example, if $n_1 > 0$ and $n_2 = 0$, we have $B(\boldsymbol{n}, \boldsymbol{n}) = -\lambda - \mu_{1,\boldsymbol{n}}$; $B(\boldsymbol{n}, (n_1+1, n_2)) = \lambda$; $B(\boldsymbol{n}, (n_1-1, n_2+1)) = \mu_{1,\boldsymbol{n}}$; $B(\boldsymbol{n}, \boldsymbol{n}') = 0$ for all the other states $\boldsymbol{n}'$. It is known that the steady-state distribution $\boldsymbol{\pi}$ can be determined by the following equations:

$$\boldsymbol{\pi B} = \boldsymbol{0}, \qquad \boldsymbol{\pi e} = 1 \tag{8}$$

where $\boldsymbol{e}$ is an infinite dimensional column vector with all the elements as 1.

The optimization objective is to find the optimal policy $\boldsymbol{\mu}^*$ that minimizes the long-run average cost of the tandem queue.

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu}_{\boldsymbol{n}} \in \mathcal{D}_U\,,\ \forall \boldsymbol{n} \in \mathcal{S}}{\operatorname{argmin}} \{\eta\}. \tag{9}$$

Since the arrival process and the service time both have the Markovian property, the above optimization problem can be viewed as a Markov decision process (MDP). The action is the value of service rates at the current state and the action space is the continuous multidimensional value domain D_U. Therefore, both the action space and the state space are infinite, and so is the policy space. The control policy $\boldsymbol{\mu}$ can be viewed as a parameterized policy. The traditional approaches of the MDP theory are difficult to apply because of the continuous and high-dimensional policy space. In the next section, we use the sensitivity-based optimization theory to study the optimality structure of this problem and develop an efficient algorithm to find the optimal policy.

III. Main Results

In this section, we derive some optimality properties of this constrained service rate control problem. These properties can simplify the continuous and infinite action space to a discrete and finite space while retaining the optimality.

First, we introduce the sensitivity-based optimization theory for Markov systems [7], [8]. The key result is the performance difference formula that quantifies the relation of the performance of Markov systems and adopted policies. For the continuous time Markov process defined in Section II, we have a column vector *performance potential* $\boldsymbol{g}$ whose element $g(\boldsymbol{n})$, $\boldsymbol{n} \in \mathcal{S}$, is defined as below:

$$g(\boldsymbol{n}) = \lim_{T\to\infty} \mathbb{E}\left\{ \int_0^T [f(\mathfrak{n}(t), \boldsymbol{\mu}_{\mathfrak{n}(t)}) - \eta]dt \middle| \mathfrak{n}(0) = \boldsymbol{n} \right\}. \tag{10}$$

The performance potential $\boldsymbol{g}$ is also called the *relative value function* in the MDP theory [22]. We denote $\tau_{\boldsymbol{n}}$ as the sojourn time before the Markov process encounters the first state transition. According to the Markovian property, (10) can be rewritten as

$$g(\boldsymbol{n}) = [f(\boldsymbol{n}, \boldsymbol{\mu}_{\boldsymbol{n}}) - \eta]\mathbb{E}\{\tau_{\boldsymbol{n}}\} - \sum_{\boldsymbol{n}' \neq \boldsymbol{n}} \frac{B(\boldsymbol{n}, \boldsymbol{n}')}{B(\boldsymbol{n}, \boldsymbol{n})} \lim_{T\to\infty} \mathbb{E} \left\{ \int_{\tau_{\boldsymbol{n}}}^T [f(\mathfrak{n}(t), \boldsymbol{\mu}_{\mathfrak{n}(t)}) - \eta]dt \middle| \mathfrak{n}(\tau_{\boldsymbol{n}}) = \boldsymbol{n}' \right\}. \tag{11}$$

Since the sojourn time of a Markov process is exponentially distributed and its mean is

$$\mathbb{E}\{\tau_{\boldsymbol{n}}\} = -\frac{1}{B(\boldsymbol{n}, \boldsymbol{n})} \tag{12}$$

we substitute the above equation and (10) into (11) and obtain the following recursive equation:

$$B(\boldsymbol{n}, \boldsymbol{n})g(\boldsymbol{n}) = -[f(\boldsymbol{n}, \boldsymbol{\mu}_{\boldsymbol{n}}) - \eta] - \sum_{\boldsymbol{n}' \neq \boldsymbol{n}} B(\boldsymbol{n}, \boldsymbol{n}')g(\boldsymbol{n}'). \tag{13}$$

We can further rewrite the above equation in a matrix form and obtain the following *Poisson equation*:

$$-\boldsymbol{B}\boldsymbol{g} = \boldsymbol{f} - \eta \boldsymbol{e}. \tag{14}$$

The above equation can be further rewritten as below since $\boldsymbol{g}$ plus any constant is still a solution to the above equation

$$(\boldsymbol{B} - \boldsymbol{e}\boldsymbol{\pi})\boldsymbol{g} = -\boldsymbol{f} \tag{15}$$

where we let $\boldsymbol{\pi}\boldsymbol{g} = \eta$ and the matrix $(\boldsymbol{B} - \boldsymbol{e}\boldsymbol{\pi})$ is invertible when the Markov process is a unichain.

Suppose the allocation policy is changed from $\boldsymbol{\mu}$ to $\boldsymbol{\mu}'$, and the two-element tuple of the Markov system is also changed from $(\boldsymbol{B}, \boldsymbol{f})$ to $(\boldsymbol{B}', \boldsymbol{f}')$. Denote $\boldsymbol{\pi}'$ and η' as the steady-state distribution and the system average cost under the new policy $\boldsymbol{\mu}'$, respectively. Left-multiplying $\boldsymbol{\pi}'$ on both sides of (14), we have

$$-\boldsymbol{\pi}'\boldsymbol{B}\boldsymbol{g} = \boldsymbol{\pi}'\boldsymbol{f} - \eta. \tag{16}$$

Since $\boldsymbol{\pi}'\boldsymbol{B}' = \boldsymbol{0}$ and $\boldsymbol{\pi}'\boldsymbol{f}' = \eta'$, the above equation can be further rewritten as below:

$$\eta' - \eta = \boldsymbol{\pi}'\left[(\boldsymbol{B}' - \boldsymbol{B})\boldsymbol{g} + (\boldsymbol{f}' - \boldsymbol{f})\right]. \tag{17}$$

This equation is called the *performance difference formula* that quantifies the change of the Markov system performance when the policy or parameters change. In (17), the value of $\boldsymbol{g}$ can be computed based on (15) or estimated from the sample path based on (10). Although we do not know the exact value of $\boldsymbol{\pi}'$ under every possible new policy $\boldsymbol{\mu}'$, the element of $\boldsymbol{\pi}'$ is always positive for ergodic Markov systems. If we choose a proper $(\boldsymbol{B}', \boldsymbol{f}')$ such that the element of the column vector represented by the square bracket in (17) is always nonpositive, then we have $\eta' - \eta \leq 0$ and the system average cost is reduced.

Since the values of $\boldsymbol{B}$ and $\boldsymbol{f}$ have been determined in Section II, we substitute the value of $\boldsymbol{B}$ and (2) into (17) and obtain

$$\eta' - \eta = \sum_{\boldsymbol{n} \in \mathcal{S}} \pi'(\boldsymbol{n}) \Big\{ (\mu'_{1,\boldsymbol{n}} - \mu_{1,\boldsymbol{n}})[g(n_1 - 1, n_2 + 1) - g(n_1, n_2)] + (\mu'_{2,\boldsymbol{n}} - \mu_{2,\boldsymbol{n}})[g(n_1, n_2 - 1) - g(n_1, n_2)] + \phi(\boldsymbol{\mu}'_{\boldsymbol{n}}) - \phi(\boldsymbol{\mu}_{\boldsymbol{n}}) \Big\} \tag{18}$$

where we note that if $n_i = 0$, then we always have $\mu'_{i,\boldsymbol{n}} = \mu_{i,\boldsymbol{n}} = 0$ for any $i = 1, 2$ and the corresponding part in the above equation should be removed.

For simplicity, we further define another two quantities as below:

$$G(\boldsymbol{n}, 1) := g(n_1 - 1, n_2 + 1) - g(n_1, n_2);$$
$$G(\boldsymbol{n}, 2) := g(n_1, n_2 - 1) - g(n_1, n_2) \tag{19}$$

where we set $G(\boldsymbol{n}, i) = 0$ if $n_i = 0$, for any $i = 1, 2$. We substitute (19) into (18) and obtain

$$\eta' - \eta = \sum_{\boldsymbol{n} \in \mathcal{S}} \pi'(\boldsymbol{n}) \Big\{ (\mu'_{1,\boldsymbol{n}} - \mu_{1,\boldsymbol{n}})G(\boldsymbol{n}, 1) + (\mu'_{2,\boldsymbol{n}} - \mu_{2,\boldsymbol{n}})G(\boldsymbol{n}, 2) + \phi(\boldsymbol{\mu}'_{\boldsymbol{n}}) - \phi(\boldsymbol{\mu}_{\boldsymbol{n}}) \Big\}. \tag{20}$$

The above difference formula describes the relation between the average system cost and the service rates. It gives a clear and efficient way to study the optimality properties and optimization algorithms for this constrained service rate control problem.

With (20), we can derive a *policy improvement* step that is equivalent to the policy iteration in the MDP theory. That is, based on the current policy $\boldsymbol{\mu}$, we can construct a new policy as follows:

$$\boldsymbol{\mu}'_{\boldsymbol{n}} = \operatorname*{argmin}_{\boldsymbol{\mu}_{\boldsymbol{n}} \in \mathcal{D}_U} \{\mu_{1,\boldsymbol{n}} G(\boldsymbol{n},1) + \mu_{2,\boldsymbol{n}} G(\boldsymbol{n},2) + \phi(\boldsymbol{\mu}_{\boldsymbol{n}})\} \quad (21)$$

for all $\boldsymbol{n} \in \mathcal{S}$. Substituting the above equation into (20), we can find $\eta' - \eta \leq 0$ since $\pi'(\boldsymbol{n})$ is always positive. If at certain state $\boldsymbol{n}$, the minimal value obtained in (21) is different from that associated to the current action $\boldsymbol{\mu}_{\boldsymbol{n}}$, we can find $\eta' - \eta < 0$ and the average cost is strictly reduced. This is also the key idea of the policy iteration, explained from the perspective of the difference formula (20). Based on this idea, we can directly obtain the following *necessary and sufficient condition* of the optimal service rates.

Theorem 1: $\boldsymbol{\mu}$ is the optimal service rates, if and only if it satisfies

$$\begin{aligned} &\mu_{1,\boldsymbol{n}} G(\boldsymbol{n},1) + \mu_{2,\boldsymbol{n}} G(\boldsymbol{n},2) + \phi(\boldsymbol{\mu}_{\boldsymbol{n}}) \leq \\ &\mu'_{1,\boldsymbol{n}} G(\boldsymbol{n},1) + \mu'_{2,\boldsymbol{n}} G(\boldsymbol{n},2) + \phi(\boldsymbol{\mu}'_{\boldsymbol{n}}) \end{aligned} \quad (22)$$

for any $\boldsymbol{\mu}'_{\boldsymbol{n}} \in \mathcal{D}_U$ and $\boldsymbol{n} \in \mathcal{S}$.

The proof of this theorem is straightforward based on the difference formula (20) and we omit it for simplicity. If the structure of the cost function is more specific, we can further derive the following theorem about the property of optimal service rates.

Theorem 2: If the operating cost rate function $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is linear to the service rate $\mu_{i,\boldsymbol{n}}$, then the system average cost η is monotonic w.r.t. $\mu_{i,\boldsymbol{n}}$, $i = 1, 2$ and $\boldsymbol{n} \in \mathcal{S}$.

Proof: Since $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is linear to $\mu_{i,\boldsymbol{n}}$, we know that the cost function has the form as (4). Substituting (4) into (20), we obtain

$$\begin{aligned} \eta' - \eta = &\sum_{\boldsymbol{n} \in \mathcal{S}} \pi'(\boldsymbol{n}) \Big\{ (\mu'_{1,\boldsymbol{n}} - \mu_{1,\boldsymbol{n}}) [G(\boldsymbol{n},1) + b_1] \\ &+ (\mu'_{2,\boldsymbol{n}} - \mu_{2,\boldsymbol{n}}) [G(\boldsymbol{n},2) + b_2] \Big\}. \end{aligned} \quad (23)$$

Without loss of generality, we assume that at a particular state $\boldsymbol{n}$, the service rate $\mu_{1,\boldsymbol{n}}$ is changed to $\mu'_{1,\boldsymbol{n}}$ and other service rates remain unvaried. Therefore, the difference formula becomes

$$\eta' - \eta = \pi'(\boldsymbol{n})(\mu'_{1,\boldsymbol{n}} - \mu_{1,\boldsymbol{n}})[G(\boldsymbol{n},1) + b_1]. \quad (24)$$

On the other hand, we reversely assume that the current service rate of server 1 at the particular state $\boldsymbol{n}$ is $\mu'_{1,n}$, while all the other service rates are the same as the above. If $\mu'_{1,n}$ is changed to $\mu_{1,n}$ and other service rates remain unvaried, we have the following difference formula:

$$\eta - \eta' = \pi(\boldsymbol{n})(\mu_{1,\boldsymbol{n}} - \mu'_{1,\boldsymbol{n}})[G'(\boldsymbol{n},1) + b_1]. \quad (25)$$

Comparing (24) and (25), we have

$$\frac{G'(\boldsymbol{n},1) + b_1}{G(\boldsymbol{n},1) + b_1} = \frac{\pi'(\boldsymbol{n})}{\pi(\boldsymbol{n})} > 0. \quad (26)$$

That is, the sign of $G'(\boldsymbol{n},1) + b_1$ and $G(\boldsymbol{n},1) + b_1$ are the same if we only change the value of $\mu_{1,\boldsymbol{n}}$. Based on (24), we can easily derive the performance derivative formula as below:

$$\frac{\partial \eta}{\partial \mu_{1,\boldsymbol{n}}} = \pi(\boldsymbol{n})[G(\boldsymbol{n},1) + b_1]. \quad (27)$$

If we change the service rate $\mu_{1,\boldsymbol{n}}$ to $\mu'_{1,\boldsymbol{n}}$ at the particular state $\boldsymbol{n}$ while keeping the service rates at other states unvaried, the corresponding derivative formula becomes

$$\frac{\partial \eta}{\partial \mu_{1,\boldsymbol{n}}} = \pi'(\boldsymbol{n})[G'(\boldsymbol{n},1) + b_1]. \quad (28)$$

Since the sign of $G'(\boldsymbol{n},1) + b_1$ and $G(\boldsymbol{n},1) + b_1$ are the same, the sign of derivative $\frac{\partial \eta}{\partial \mu_{1,\boldsymbol{n}}}$ will also remain unvaried if we only change the value of $\mu_{1,\boldsymbol{n}}$. That is, η is monotonic w.r.t. $\mu_{1,\boldsymbol{n}}$. The theorem is proved. ■

With Theorem 2, we can see that the optimal service rates can be reached at the boundary of the feasible domain. If the operating cost is not linear, the monotonicity property may not hold. With further study, we derive the following theorem to describe the optimality property of optimal service rates.

Theorem 3: Assume $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is concave in $\mu_{i,\boldsymbol{n}}$. If the value of $\mu_{j,\boldsymbol{n}}$ is given, then the optimal value of $\mu_{i,\boldsymbol{n}}$ can be either u_i or $U - \mu_{j,\boldsymbol{n}}$, where $i, j = 1, 2$ and $j \neq i$, $\boldsymbol{n} \in \mathcal{S}$, $n_i > 0$.

Proof: Without loss of generality, we choose $i = 1$ and $j = 2$. The following analysis is also valid for the case $i = 2$ and $j = 1$ and we omit it for simplicity. Thus, $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is a concave function of $\mu_{1,\boldsymbol{n}}$. Since the value of $\mu_{2,\boldsymbol{n}}$ is given, we can rewrite (21) as below:

$$\begin{aligned} \mu'_{1,\boldsymbol{n}} &= \operatorname*{argmin}_{\mu_{1,\boldsymbol{n}} \in [u_1, U - \mu_{2,\boldsymbol{n}}]} \{\mu_{1,\boldsymbol{n}} G(\boldsymbol{n},1) + \mu_{2,\boldsymbol{n}} G(\boldsymbol{n},2) + \phi(\boldsymbol{\mu}_{\boldsymbol{n}})\} \\ &= \operatorname*{argmin}_{\mu_{1,\boldsymbol{n}} \in [u_1, U - \mu_{2,\boldsymbol{n}}]} \{\mu_{1,\boldsymbol{n}} G(\boldsymbol{n},1) + \phi(\boldsymbol{\mu}_{\boldsymbol{n}})\}, \quad \boldsymbol{n} \in \mathcal{S}. \end{aligned} \quad (29)$$

Obviously, the function $\mu_{1,\boldsymbol{n}} G(\boldsymbol{n},1) + \phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is also concave in $\mu_{1,\boldsymbol{n}}$ since $G(\boldsymbol{n},1)$ is a constant. Therefore, the minimum of the above problem can be achieved at the boundary of its value domain, i.e., u_1 or $U - \mu_{2,\boldsymbol{n}}$. With Theorem 1, we know that the optimal service rate of $\mu_{1,\boldsymbol{n}}$ can be either u_1 or $U - \mu_{2,\boldsymbol{n}}$. The theorem is proved. ■

Remark 1: Theorems 2 and 3 indicate that the optimal service rates of this problem have a structure of *bang–bang control* when $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is linear or concave.

Remark 2: The result in Theorem 3 is also valid for the model in Theorem 2, since a linear function is also concave.

With Theorem 3 and Remark 2, we can further derive the following theorem to reduce the feasible domain from $\mathcal{D}_U$ to its vertexes.

Theorem 4: If $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is differentiable and concave (including linear) w.r.t. $\mu_{i,\boldsymbol{n}}$, then the optimal service rates $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}})$ can be selected from a three-element set $\mathcal{D}_v := \{(u_1, u_2), (U - u_2, u_2), (u_1, U - u_1)\}$, instead of the triangle domain $\mathcal{D}_U$, where $n_1, n_2 > 0$.

Proof: Since $\phi(\boldsymbol{\mu}_{\boldsymbol{n}})$ is concave, Theorem 3 and Remark 2 are valid. Therefore, the interior area of the triangle can be removed from the value domain. Below, we prove that the edges of the

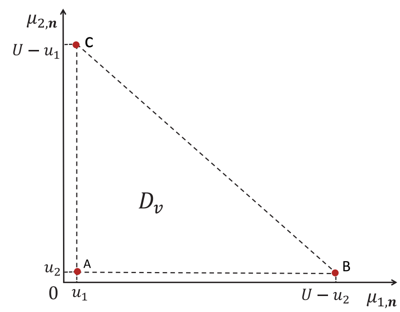

Fig. 3. Refined value domain $\mathcal{D}_v$ (three points A,B,C) of service rates.

triangle $\triangle(A, B, C)$ in Fig. 3 can be further reduced to its vertexes $\{A, B, C\}$.

Assume the optimal service rate is located at an interior point of line AB. With Theorem 3, we know that the optimal value of $\mu_{1,\boldsymbol{n}}$ can be u_1 or $U - u_2$. That is, the optimal service rate can be reduced to the end points A or B. For the case of line AC, the proof is similar. Now assuming that the optimal service rate is located at an interior point of line BC, we use the derivative comparison to study this case. Similar to (27), we have the derivative formula as below:

$$\begin{aligned} \frac{\partial \eta}{\partial \mu_{1,\boldsymbol{n}}} &= \pi(\boldsymbol{n})[G(\boldsymbol{n},1) + \frac{\partial \phi}{\partial \mu_{1,\boldsymbol{n}}}], \\ \frac{\partial \eta}{\partial \mu_{2,\boldsymbol{n}}} &= \pi(\boldsymbol{n})[G(\boldsymbol{n},2) + \frac{\partial \phi}{\partial \mu_{2,\boldsymbol{n}}}] \end{aligned} \tag{30}$$

where $\phi(\boldsymbol{\mu_n})$ is differentiable w.r.t. $\mu_{1,\boldsymbol{n}}$ and $\mu_{2,\boldsymbol{n}}$. At the current point, if $\frac{\partial \eta}{\partial \mu_{1,\boldsymbol{n}}} \leq \frac{\partial \eta}{\partial \mu_{2,\boldsymbol{n}}}$, we can move downward along line BC to obtain a reduced η; otherwise, we can move upward along line BC. This contradicts with the assumption that the current interior point is the optimal service rates. Therefore, the optimal value can be found at the end points B or C. In summary, the optimal value of service rates can be found in the vertexes set $\mathcal{D}_v$, as illustrated in Fig. 3. The theorem is proved. ■

Therefore, if $\phi(\boldsymbol{\mu_n})$ is concave, the searching space of optimal service rates can be reduced from the triangle area $\mathcal{D}_U$ in Fig. 2 to the three-element set $\mathcal{D}_v$ in Fig. 3. This is a significant saving of computation resources for optimization algorithms.

With Theorem 4 and the derivative formula (30), we can further derive the following corollary to specify the optimal value of service rates.

Corollary 1: Assume $\phi(\boldsymbol{\mu_n})$ is differentiable and concave w.r.t. $\mu_{i,\boldsymbol{n}}$. If the current service rate control policy is optimal, then we have the following.

1) If $G(\boldsymbol{n},1) + \frac{\partial \phi}{\partial \mu_{1,\boldsymbol{n}}} \leq G(\boldsymbol{n},2) + \frac{\partial \phi}{\partial \mu_{2,\boldsymbol{n}}}$ and $G(\boldsymbol{n},1) + \frac{\partial \phi}{\partial \mu_{1,\boldsymbol{n}}} < 0$, then $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}}) = (U - u_2, u_2)$.
2) If $G(\boldsymbol{n},1) + \frac{\partial \phi}{\partial \mu_{1,\boldsymbol{n}}} \geq G(\boldsymbol{n},2) + \frac{\partial \phi}{\partial \mu_{2,\boldsymbol{n}}}$ and $G(\boldsymbol{n},2) + \frac{\partial \phi}{\partial \mu_{2,\boldsymbol{n}}} < 0$, then $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}}) = (u_1, U - u_1)$.
3) If $G(\boldsymbol{n},1) + \frac{\partial \phi}{\partial \mu_{1,\boldsymbol{n}}} \geq 0$ and $G(\boldsymbol{n},2) + \frac{\partial \phi}{\partial \mu_{2,\boldsymbol{n}}} \geq 0$, then $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}}) = (u_1, u_2)$.

The proof of the above corollary is straightforward based on the proof procedure of Theorem 4. For simplicity, we omit the details.

With Corollary 1, we see that the distribution of the optimal value of service rates $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}})$ is illustrated in Fig. 4, where

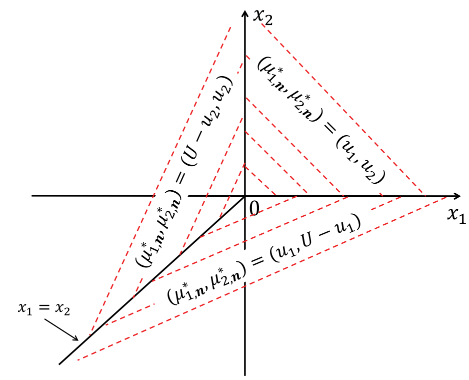

Fig. 4. Optimal value of service rates $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}})$ in different zones of (x_1, x_2).

we define auxiliary variables x_1 and x_2 as follows:

$$\begin{aligned} x_1 &:= G(\boldsymbol{n},1) + \frac{\partial \phi}{\partial \mu_{1,\boldsymbol{n}}}, \\ x_2 &:= G(\boldsymbol{n},2) + \frac{\partial \phi}{\partial \mu_{2,\boldsymbol{n}}}. \end{aligned} \tag{31}$$

Remark 3: If $n_i = 0$ and $n_j > 0$, then $\mu_{i,\boldsymbol{n}} = 0$ and the value domain of $\mu_{j,\boldsymbol{n}}$ is $[u_j, U]$; the result in Theorem 3 is still valid and the value of $\mu^*_{j,\boldsymbol{n}}$ in Theorem 4 is either u_j or U, where $i, j = 1, 2$ and $i \neq j$.

Remark 4: When $\phi(\boldsymbol{\mu_n})$ is convex w.r.t. $\boldsymbol{\mu_n}$, the monotonicity and the optimality of bang–bang control in Theorems 2 and 3 may not hold. The optimal service rate may be found in the interior area of $\mathcal{D}_U$, not necessarily on the boundary. Since the feasible domain $\mathcal{D}_U$ is a convex set, the policy improvement (21) becomes a convex optimization problem. We can easily solve (21) to conduct the policy improvement iteratively. For example, if $\phi(\boldsymbol{\mu_n}) = c_0 + b_1\mu^2_{1,\boldsymbol{n}} + b_2\mu^2_{2,\boldsymbol{n}}$, (21) becomes a quadratic problem that has a closed-form solution. We may also study the optimality property with further analysis.

We have now completed a comprehensive study for this resource-constrained service rate control problem, including the cases where the operating cost rate function $\phi(\boldsymbol{\mu_n})$ is linear, concave, or convex. If $\phi(\boldsymbol{\mu_n})$ is linear or concave, we have all the above optimality properties that help greatly reduce the optimization complexity. If $\phi(\boldsymbol{\mu_n})$ is convex, although these properties may not hold, the original-constrained service rate control problem may be decomposed into a series of convex optimization problems (21) that are easy to solve.

IV. Algorithm and Extension

In the previous section, we have a detailed study on this power-constrained service rate control problem in the two-server tandem queue. In this section, we give an iterative algorithm to find the optimal control policy for service rates. We also further discuss some extension issues that are related to practical applications.

A. Optimization Algorithm

Based on the main results derived in Section III, especially the difference formula (20) and Theorem 1, we can develop the following Algorithm 1 to find the optimal control policy of

Algorithm 1: An iterative algorithm to find the optimal service rates.

- Initialize: Choose any initial service rate policy $\boldsymbol{\mu}^{(0)}$, and set $k=0$.
- **while** stopping criterion $\boldsymbol{\mu}^{(k)} = \boldsymbol{\mu}^{(k-1)}$ or $\| \boldsymbol{\mu}^{(k)} - \boldsymbol{\mu}^{(k-1)} \|_{\infty} < \epsilon$ is unsatisfied, **do**

  Numerically compute or estimate the value of $G(\boldsymbol{n}, i)$'s based on (19), $\forall \boldsymbol{n} \in \mathcal{S}$ and $i = 1, 2$.

  for every state $\boldsymbol{n}$, **do**

  if $n_1 > 0$ and $n_2 > 0$, **then**

$$(\mu_{1,\boldsymbol{n}}^{(k+1)}, \mu_{2,\boldsymbol{n}}^{(k+1)}) = \underset{(\mu_{1,\boldsymbol{n}},\mu_{2,\boldsymbol{n}}) \in \mathcal{D}_U}{\text{argmin}} \Big\{ \mu_{1,\boldsymbol{n}} G(\boldsymbol{n}, 1) + \mu_{2,\boldsymbol{n}} G(\boldsymbol{n}, 2) + \phi(\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}}) \Big\}, \tag{32}$$

  where $\mathcal{D}_U$ becomes $\mathcal{D}_v = \{(u_1, u_2), (U - u_2, u_2), (u_1, U - u_1)\}$ if $\phi(\cdot)$ is concave;

  else if $n_1 = 0$ and $n_2 > 0$, **then**

$$\begin{aligned} \mu_{1,\boldsymbol{n}}^{(k+1)} &= 0, \\ \mu_{2,\boldsymbol{n}}^{(k+1)} &= \underset{\mu_{2,\boldsymbol{n}} \in [u_2, U]}{\text{argmin}} \Big\{ \mu_{2,\boldsymbol{n}} G(\boldsymbol{n}, 2) + \phi(0, \mu_{2,\boldsymbol{n}}) \Big\}, \end{aligned} \tag{33}$$

  where $[u_2, U]$ becomes $\{u_2, U\}$ if $\phi(\cdot)$ is concave;

  else if $n_1 > 0$ and $n_2 = 0$, **then**

$$\begin{aligned} \mu_{1,\boldsymbol{n}}^{(k+1)} &= \underset{\mu_{1,\boldsymbol{n}} \in [u_1, U]}{\text{argmin}} \Big\{ \mu_{1,\boldsymbol{n}} G(\boldsymbol{n}, 1) + \phi(\mu_{1,\boldsymbol{n}}, 0) \Big\}, \\ \mu_{2,\boldsymbol{n}}^{(k+1)} &= 0, \end{aligned} \tag{34}$$

  where $[u_1, U]$ becomes $\{u_1, U\}$ if $\phi(\cdot)$ is concave;

  else if $n_1 = 0$ and $n_2 = 0$, **then**

$$\mu_{1,\boldsymbol{n}}^{(k+1)} = \mu_{2,\boldsymbol{n}}^{(k+1)} = 0. \tag{35}$$

  end if

  end for

  $k \leftarrow k + 1$;

  end while
- Output $\boldsymbol{\mu}^{(k)}$ as the optimal service rates of the tandem queue.

service rates for this resource-constrained tandem queue. Note that the subproblems (32) in Algorithm 1 are easy to handle even if $\phi(\cdot)$ is not linear or concave, but a general differentiable function. One of the direct ways is by computing all the *stationary points* for the objective function in (32) and selecting the one with the minimal value. The stationary point is computed by letting the derivatives in (30) equal 0, or simply letting x_1 and x_2 in (31) equal 0. In (31), the value of $G(\boldsymbol{n}, i)$, $\boldsymbol{n} \in \mathcal{S}$ and $i = 1, 2$, can be numerically computed or online estimated based on the sample path, which is the same as the algorithms to obtain the performance potentials or value functions in Markov systems (details can be referred to the book [7, Chapters 3 and 5]).

The convergence of Algorithm 1 can be studied based on the necessary and sufficient condition of optimal policy in Theorem 1. Actually, Algorithm 1 can be viewed as a specific version of the policy iteration in the traditional MDP theory, except that the policy in the above algorithm is a parametric one. Therefore, the detailed discussion on algorithmic properties of Algorithm 1 is omitted. Interested audience can refer to [22].

B. Truncation of Infinite Buffer

When we want to apply the main results in Section III and Algorithm 1 to real problems, we may encounter some practical issues. The first issue is about the infinite size of the state space. All the results in Section III are derived under the assumption of infinite buffer size. That is, the state space $\mathcal{S}$ is infinite. Although the structure analysis of an infinite-dimensional optimization problem is desirable, we have a computation issue when we compute the value of $G(\boldsymbol{n}, i)$ for all $\boldsymbol{n} \in \mathcal{S}$, which is a fundamental quantity in the policy improvement such as (21). It is always difficult to compute the value of a variable with infinite dimensions. We have to resort to some approximation methods to handle this issue. In this paper, we use the *truncation technique* to approximate the infinite buffer as a finite buffer. That is, we truncate the system state when the queue length is large enough. The reasonability of such truncation lays in the fact that the steady-state probability usually decreases exponentially in queue lengths for a stable queueing system. Thus, the effect of these system states with large queue lengths is ignorable. More specifically, the truncation used in our problem is defined as below. We assume that the buffer sizes of servers 1 and 2 are N_1 and N_2, respectively. When the buffer of server 2 is full, its preceding node, server 1, will stop service. We call this a *blocking scheme*. When the buffer of server 1 is full, any newly arriving customer will be rejected by server 1. With such truncation, the system state space is finite and we can numerically compute all the values of $G(\boldsymbol{n}, i)$ in this truncated state space. Obviously, when N_i is large, the approximation error caused by truncation will be small enough to ignore. In Section V, we conduct a comparison of optimal performance under different values of N_i, which demonstrates the reasonability of truncation with large N_i, $i = 1, 2$.

Since the buffer is finite, people may further consider a *rejection cost* caused by customer losses, besides the holding and operating costs defined in (2). The rejection cost happens when $n_1 = N_1$ and any new arrival will be rejected. We may define the rejection cost rate function as below:

$$\psi(\boldsymbol{n}) = r_1 \lambda 1_{(n_1 = N_1)} \tag{36}$$

where $1_{(\cdot)}$ is the indicator function and r_1 is the unit rejection cost per customer rejected. The total cost rate function in (2) should be updated as

$$f(\boldsymbol{n}, \boldsymbol{\mu_n}) = \varphi(\boldsymbol{n}) + \psi(\boldsymbol{n}) + \phi(\boldsymbol{\mu_n}). \tag{37}$$

We can see that the rejection cost $\psi(\boldsymbol{n})$ is similar to the holding cost $\varphi(\boldsymbol{n})$, both have no relation to service rates. We can view $\psi(\boldsymbol{n})$ as a special part of $\varphi(\boldsymbol{n})$ and define a new holding cost $\varphi'(\boldsymbol{n})$. Introducing the rejection cost $\psi(\boldsymbol{n})$ does not bring difficulties to our optimization problem. All the results in Section III still hold and we only need to replace $\varphi(\boldsymbol{n})$ with a new defined $\varphi'(\boldsymbol{n}) := \varphi(\boldsymbol{n}) + \psi(\boldsymbol{n})$ throughout the analysis.

Moreover, we have to pay attention to another issue caused by the blocking scheme. When $n_2 = N_2$ and server 1 is blocked, the term $(\mu'_{1,\boldsymbol{n}} - \mu_{1,\boldsymbol{n}})[g(n_1 - 1, n_2 + 1) - g(n_1, n_2)]$ in (18) should be removed since $\mu_{1,\boldsymbol{n}}$ is fixed at 0 and $n_2 + 1$ is out of the range of buffer size. In other words, we define

$$G(\boldsymbol{n}, 1) := 0 \text{ and } \mu_{1,\boldsymbol{n}} := 0, \qquad \text{if } n_2 = N_2. \tag{38}$$

We should substitute (38) into all the analysis in Section III when the buffer size is finite. All the theorems and remarks in Section III still hold.

C. System Ergodicity

Another issue is about the ergodicity of the system. When the minimal service rate $u_i > 0$ and the queueing system is stable, the underlined Markov system is always ergodic. When $u_i = 0$ and some service rates may reach 0, it is possible that the Markov system is not ergodic. When the Markov system is not ergodic but a unichain, all the analysis in Section III still hold. When the Markov system is a multichain, the correctness of the analysis in Section III may need further investigation, which is out of the scope of this paper. In order to simplify the difficulty caused by the nonergodicity, we assume $u_i > 0$ during the simulation experiments in the next section. We can choose a very small u_i to approximate the case that the service rate may reach 0. This approximation is reasonable since most of the average performance of common systems in real world will change continuously when the minimal service rate is approaching to 0.

D. Tandem Queue With Many Servers

In this paper, we study the tandem queue with only two servers. It is desirable to extend our analysis to a general tandem queue with more than two servers. In this section, we give a preliminary study on this extension.

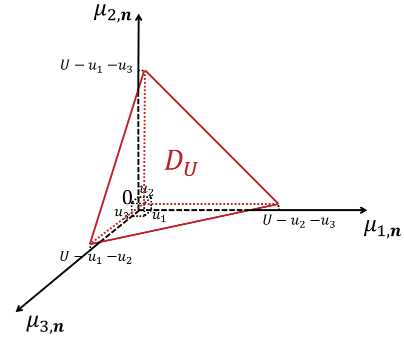

Fig. 5. Value domain (tetrahedron space covered by red lines) $\mathcal{D}_U$ of service rates in a tandem queue with three servers.

Consider a tandem queue with M servers. All the parameters are similar to those in Section III. The service rates of servers satisfy the constraint $\mu_{i,\boldsymbol{n}} \geq u_i$ and $\sum_{i=1}^{M} \mu_{i,\boldsymbol{n}} \leq U$, $\forall \boldsymbol{n} \in \mathcal{S}$ and $i = 1, 2, \ldots, M$. We can verify that the value domain of service rates $\mathcal{D}_U$ is an M-dimensional polyhedron, which can be viewed as a high-dimensional analog of the triangle area depicted in Fig. 2. For example, if $M = 3$, we can find that $\mathcal{D}_U$ is a tetrahedron, as illustrated in Fig. 5.

All the analysis in Section III can be conducted for this M-server tandem queue in a similar way. Without rigorous proof, we give the updated version of difference formula (20) as below:

$$\eta' - \eta = \sum_{\boldsymbol{n} \in \mathcal{S}} \pi'(\boldsymbol{n}) \left\{ \sum_{i=1}^{M} (\mu'_{i,\boldsymbol{n}} - \mu_{i,\boldsymbol{n}}) G(\boldsymbol{n}, i) + \phi(\boldsymbol{\mu}'_{\boldsymbol{n}}) - \phi(\boldsymbol{\mu_n}) \right\} \tag{39}$$

where $G(\boldsymbol{n}, i)$ is defined similarly as below:

$$G(\boldsymbol{n}, i) := g(\boldsymbol{n}_{-i}) - g(\boldsymbol{n}) \tag{40}$$

where $\boldsymbol{n}_{-i}$ is called a neighboring state of $\boldsymbol{n}$ and it is defined as below:

$$\begin{aligned} &\boldsymbol{n}_{-i} := (n_1, \ldots, n_i - 1, n_{i+1} + 1, \ldots, n_M), \quad i \leq M - 1; \\ &\boldsymbol{n}_{-i} := (n_1, \ldots, n_M - 1), \quad i = M \end{aligned} \tag{41}$$

where $n_i > 0$, $i = 1, \ldots, M$. If $n_i = 0$, we define $\mu_{i,\boldsymbol{n}} = 0$ and $G(\boldsymbol{n}, i) = 0$.

Based on the new difference formula (39), we can derive the results similar to Theorems 1 and 2 after necessary modifications. Theorem 3 is also similar except that we have to modify the statement as, given all $\mu_{j,\boldsymbol{n}}$ with $j \neq i$, the optimal value of $\mu_{i,\boldsymbol{n}}$ is either u_i or $U - \sum_{j \neq i}^{M} \mu_{j,\boldsymbol{n}}$. Theorem 4 is also similar except $D_v := \{(u_1, \ldots, u_M), (U - \sum_{j \neq 1}^{M} u_j, u_2, \ldots, u_M), \ldots, (u_1, \cdots, u_{M-1}, U - \sum_{j \neq M}^{M} u_j)\}$ which represents the vertexes of the polyhedron $\mathcal{D}_U$. For Corollary 1, the main idea should be modified as follows, let the server i with the minimal negative value of $G(\boldsymbol{n}, i) + \frac{\partial \phi}{\partial \mu_{i,\boldsymbol{n}}}$ choose $U - \sum_{j \neq i}^{M} u_j$ and let other servers choose minimal service rates; if all $G(\boldsymbol{n}, i) + \frac{\partial \phi}{\partial \mu_{i,\boldsymbol{n}}}$ are nonnegative, let all servers choose minimal service rates.

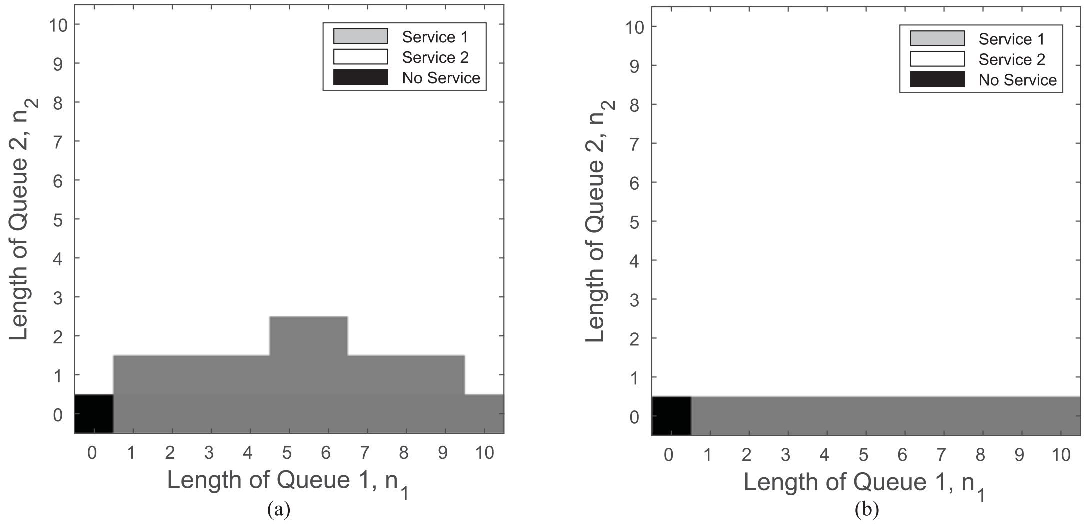

Fig. 6. Output policy of Algorithm 1 when the cost function is $\phi(\boldsymbol{\mu_n}) = \mu_{1,\boldsymbol{n}} + \mu_{2,\boldsymbol{n}}$ and $\varphi(\boldsymbol{n}) = n_1 + n_2$. The algorithm converges in three iterations. (a) Iteration 1, $\eta^{(1)} = 3.8678$. (b) Iteration 2, $\eta^* = 3.6304$.

Note that all the above statements are based on the extension of our analysis on the tandem queue with two servers. Such extension seems natural and we ignore the rigorous proof due to limited space. It may motivate our future study on more general cases, such as cyclic queues or even Jackson networks. In the next section, we still focus our numerical study on the tandem queue with two servers, since it is more intuitive to present the optimization data in a two-dimensional plane.

V. Simulation Study

In this section, we conduct some numerical experiments to demonstrate the correctness of our results and the efficiency of our optimization algorithm. The experiments are categorized according to different forms of operating cost and holding cost.

A. Linear Operating Cost

In this section, we study the case where the operating cost rate function has a linear form as (3). That is, we use (4) as the total cost rate function. Therefore, the main results in Section III always hold for this linear operating cost. Substituting (3) into Algorithm 1, we can rewrite the formulas (32)–(34) as below, respectively:

$$(\mu_{1,\boldsymbol{n}}^{(k+1)}, \mu_{2,\boldsymbol{n}}^{(k+1)}) = \underset{(\mu_{1,\boldsymbol{n}},\mu_{2,\boldsymbol{n}})\in\mathcal{D}_v}{\operatorname{argmin}} \Big\{\mu_{1,\boldsymbol{n}}[G(\boldsymbol{n},1)+b_1]+\mu_{2,\boldsymbol{n}}[G(\boldsymbol{n},2)+b_2]\Big\}. \quad (42)$$

$$\mu_{1,\boldsymbol{n}}^{(k+1)} = 0,$$
$$\mu_{2,\boldsymbol{n}}^{(k+1)} = \underset{\mu_{2,\boldsymbol{n}}\in\{u_2,U\}}{\operatorname{argmin}} \Big\{\mu_{2,\boldsymbol{n}}[G(\boldsymbol{n},2)+b_2]\Big\}. \quad (43)$$

$$\mu_{1,\boldsymbol{n}}^{(k+1)} = \underset{\mu_{1,\boldsymbol{n}}\in\{u_1,U\}}{\operatorname{argmin}} \Big\{\mu_{1,\boldsymbol{n}}[G(\boldsymbol{n},1)+b_1]\Big\},$$
$$\mu_{2,\boldsymbol{n}}^{(k+1)} = 0. \quad (44)$$

Below, we conduct some numerical examples where the holding cost and the value of parameters are different.

Example 1 Linear Operating Cost and Linear Holding Cost: First, we consider an example where the holding cost rate function is simply the sum of queue lengths in these two servers. The parameter setting is as follows:

1) Operating cost rate: $\phi(\boldsymbol{\mu_n}) = \mu_{1,\boldsymbol{n}} + \mu_{2,\boldsymbol{n}}$.
2) Holding cost rate: $\varphi(\boldsymbol{n}) = n_1 + n_2$.
3) Arrival rate: $\lambda = 1$.
4) Buffer limit: $N_1 = N_2 = 10$.
5) Resource limit: $U = 3$, $u_1 = u_2 = 0.01$.

We initialize the policy iteration algorithm with a simple constant policy $\mu_{1,\boldsymbol{n}} = \mu_{2,\boldsymbol{n}} = U/2$, $\forall\ \boldsymbol{n} \in \mathcal{S}$. This yields an initial average cost $\eta^{(0)} = 5.7939$. The algorithm is executed after the specification with (42)–(44) and it converges in three iterations. The policy and the associated average cost of the last two iterations are shown in Fig. 6, where the gray color means $\boldsymbol{\mu_n} = (U - u_2, u_2)$, the white color means $\boldsymbol{\mu_n} = (u_1, U - u_1)$, and the black color means $\boldsymbol{\mu_n} = (u_1, u_2)$.

From Fig. 6(b), we see that the optimal policy for this example allocates no power when both queues are empty (the black block), always favors queue 2 when it is not empty (the white block), and favors queue 1 only when queue 2 is empty (the gray block). The form of this optimal policy is very simple and it is easy for participators to adopt in practice. Of course, such a form of the optimal policy may depend on the value of system parameters. By changing different values of b_1 and b_2, we find that the optimal policy has other two forms besides that in Fig. 6(b): one is always allocating the minimal power for both queues when b_1 and b_2 are large enough, i.e., the policy map is similar to Fig. 6(b) by replacing with all black color; the other is always allocating the minimal power to queue 1 and allocating the maximal power to queue 2 if it is not empty, when b_1 is much larger that b_2, i.e., the policy map is similar to Fig. 6(b) by replacing gray color with black color. In these cases, we see that queue 2 always has a higher priority than queue 1. This is reasonable if we notice the fact that if $\varphi(\boldsymbol{n}) = n_1 + n_2$, a service completion at queue 1

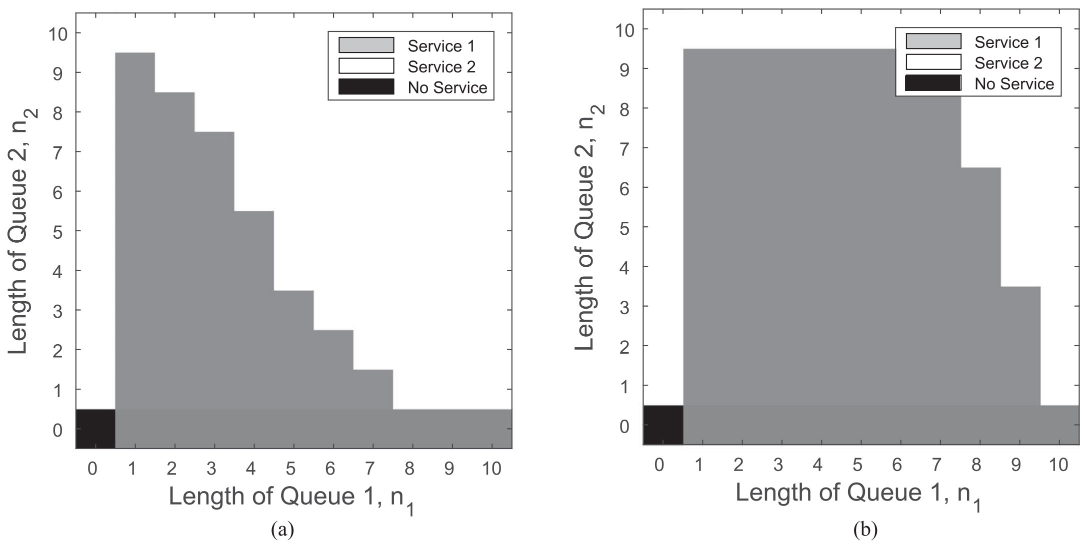

Fig. 7. Optimal policy found by Algorithm 1 using the same parameters as those in Fig. 6, except the holding cost $\varphi(\boldsymbol{n})$ is different. (a) Optimal policy for case $\varphi(\boldsymbol{n}) = 2.1n_1 + n_2$. (b) Optimal policy for case $\varphi(\boldsymbol{n}) = 10n_1 + n_2$.

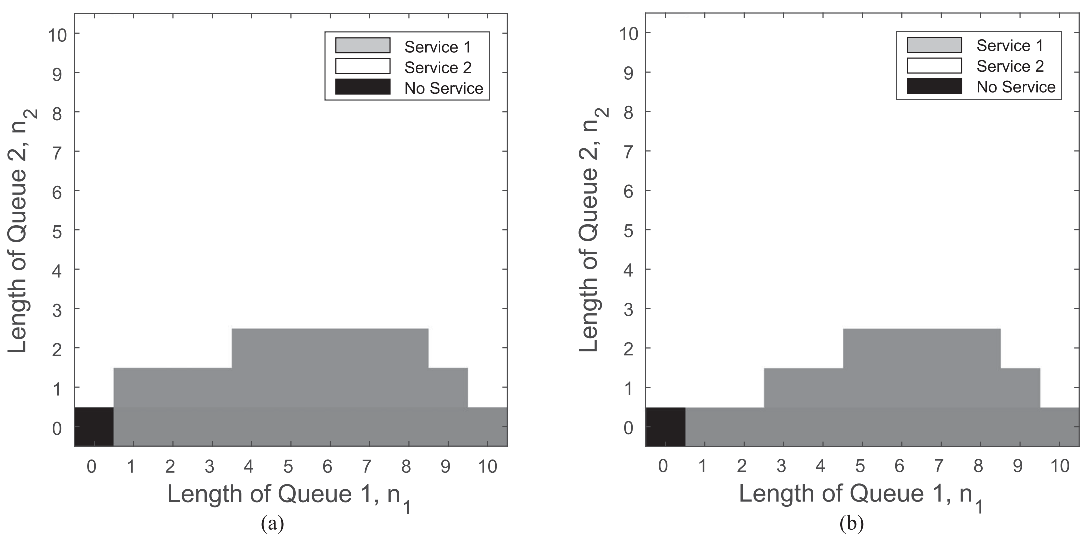

Fig. 8. Output policy of Algorithm 1 in the last 2 iterations when the cost function is $\phi(\boldsymbol{\mu_n}) = \mu_{1,\boldsymbol{n}} + \mu_{2,\boldsymbol{n}}$ and $\varphi(\boldsymbol{n}) = n_1^2 + n_2^2$. The algorithm converges in three iterations. (a) Iteration 1, $\eta^{(1)} = 6.8687$. (b) Iteration 2, $\eta^* = 6.6434$.

cannot reduce the current holding cost. Therefore, we prefer to allocate more power to queue 2 to reduce n_2, and therefore reduce $\varphi(\boldsymbol{n})$.

On the other hand, we study the truncation error by increasing the buffer size to $N_1 = N_2 = 20$. As discussed in Section IV-B, we use the truncation technique to avoid the computation difficulty caused by the infinite dimensional variable $G(\boldsymbol{n}, i)$. We hope such truncation error is small enough for practical applications. In this simulation experiment, we find that the optimal performance under this new truncation is $\eta^* = 3.6671$, which has only 1% relative error compared with the previous experiment listed above.

We further change the form of the holding cost rate function to see any possible change of the optimal policy. We conduct numerical experiments using the same parameter setting as above, except the holding cost is set as $\varphi(\boldsymbol{n}) = 2.1n_1 + n_2$ and $\varphi(\boldsymbol{n}) = 10n_1 + n_2$, respectively. We run Algorithm 1 and obtain the following optimal policies for these two cases. From Fig. 7, we see that we prefer to allocate more power to queue 1 since the cost weight of n_1 is larger than that of n_2. By comparing (a) and (b) in Fig. 7, we further find that when the cost weight of queue 1 is higher, the optimal power allocation policy favors queue 1 at more states.

Example 2 Linear Operating Cost and Convex Holding Cost: We study another example where the holding cost rate function is a convex function as below:

1) Holding cost rate: $\varphi(\boldsymbol{n}) = n_1^2 + n_2^2$;
2) Other parameters are the same as those in Example 1.

We also initialize Algorithm 1 with the simple constant policy $\mu_{1,\boldsymbol{n}} = \mu_{2,\boldsymbol{n}} = U/2, \forall\, \boldsymbol{n} \in \mathcal{S}$, and the associated initial average cost is $\eta^{(0)} = 18.5997$. Algorithm 1 is executed using (42)–(44) and it converges in three iterations. The output policies of the last two iterations are illustrated in Fig. 8. We see that the optimal policy is different from that in Fig. 6. In Fig. 8, the optimal policy

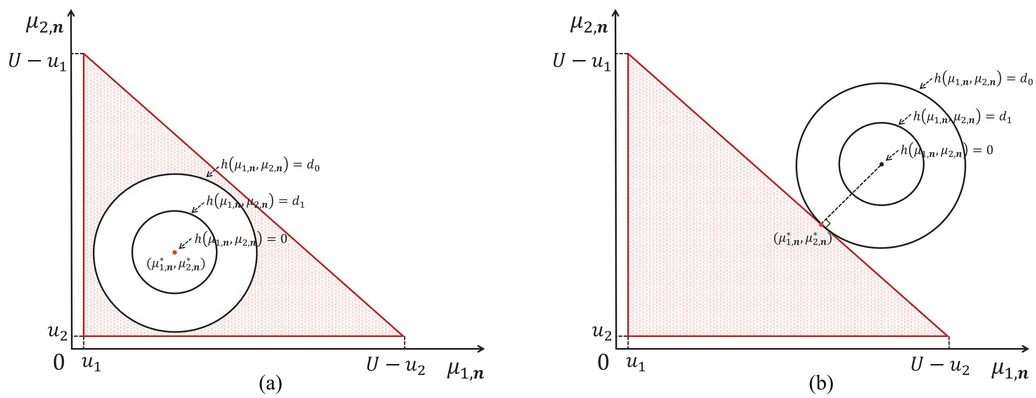

Fig. 9. Solution to the convex optimization problem (45). (a) Concentric point is in the feasible domain $\mathcal{D}_U$. (b) Concentric point is not in $\mathcal{D}_U$.

favors queue 1 at some states because a customer departure from queue 1 can reduce the holding cost $\varphi(\boldsymbol{n}) = n_1^2 + n_2^2$ more than a customer departure from queue 2 at these states.

B. Convex Operating Cost

When the operating cost is concave, we have similar results to those of linear case in Section V-A, and we omit it for simplicity. In this section, we study another case where the operating cost rate function $\phi(\boldsymbol{\mu_n})$ is convex. Although the results in Theorems 2–4 may not hold, we can still use Algorithm 1 to find the optimal policy. The subproblems (32)–(34) in Algorithm 1 are convex optimization problems that are easy to solve.

Example 3 (Convex Operating Cost and Linear Holding Cost): We choose the system parameters of the numerical example as follows:

1) Operating cost rate: $\phi(\boldsymbol{\mu_n}) = \mu_{1,\boldsymbol{n}}^2 + \mu_{2,\boldsymbol{n}}^2$.
2) Other parameters are the same as those in Example 1.

With the above operating cost, we see that the subproblem (32) in Algorithm 1 becomes

$$(\mu_{1,\boldsymbol{n}}^{(k+1)}, \mu_{2,\boldsymbol{n}}^{(k+1)}) = \underset{(\mu_{1,\boldsymbol{n}},\mu_{2,\boldsymbol{n}})\in\mathcal{D}_U}{\operatorname{argmin}} \left\{ h(\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}}) - \frac{G^2(\boldsymbol{n},1)}{4} - \frac{G^2(\boldsymbol{n},2)}{4} \right\}$$
$$= \underset{(\mu_{1,\boldsymbol{n}},\mu_{2,\boldsymbol{n}})\in\mathcal{D}_U}{\operatorname{argmin}} \left\{ h(\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}}) \right\} \quad (45)$$

where the objective function is defined as

$$h(\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}}) := \left(\mu_{1,\boldsymbol{n}} + \frac{G(\boldsymbol{n},1)}{2}\right)^2 + \left(\mu_{2,\boldsymbol{n}} + \frac{G(\boldsymbol{n},2)}{2}\right)^2. \quad (46)$$

Obviously, (45) is a convex optimization problem since the feasible domain $\mathcal{D}_U$ in Fig. 2 is a convex set. It is easy to derive the solution to (45) as follows. As illustrated in Fig. 9, the concentric circles indicate the contour of objective function $h(\mu_{1,\boldsymbol{n}}, \mu_{2,\boldsymbol{n}})$. In Fig. 9(a), when the concentric point belongs to the feasible domain, i.e., $(-\frac{G(\boldsymbol{n},1)}{2}, -\frac{G(\boldsymbol{n},2)}{2}) \in \mathcal{D}_U$, we directly have the solution to (45) as

$$(\mu_{1,\boldsymbol{n}}^*, \mu_{2,\boldsymbol{n}}^*) = \left(-\frac{G(\boldsymbol{n},1)}{2}, -\frac{G(\boldsymbol{n},2)}{2}\right). \quad (47)$$

In Fig. 9(b), when the concentric point is not in the feasible domain, i.e., $(-\frac{G(\boldsymbol{n},1)}{2}, -\frac{G(\boldsymbol{n},2)}{2}) \notin \mathcal{D}_U$, we can find a concentric circle with the minimal value of d_0, which has a cross point with the boundary of $\mathcal{D}_U$. As illustrated in Fig. 9(b), the cross point is also a tangent point and it is the solution to (45).

Similarly, we apply the above analysis to specify the form of (33) and (34) in Algorithm 1. We also initialize Algorithm 1 with the simple constant policy $\mu_{1,\boldsymbol{n}} = \mu_{2,\boldsymbol{n}} = U/2, \forall\, \boldsymbol{n} \in \mathcal{S}$. This yields an initial average cost $\eta^{(0)} = 6.7919$. We run Algorithm 1 and find the optimal policy is obtained within five iterations. The optimal average cost is $\eta^* = 5.8932$. Since the operating cost is convex, the optimal service rates may not always lie on the vertexes of the feasible domain $\mathcal{D}_U$. We cannot use the gray-scale map such as Fig. 6 to illustrate the structure of the optimal policy. Therefore, we use the three-dimensional curves in Fig. 10 to illustrate the optimal policy. In Fig. 10, these two curved surfaces represent the values of $\mu_{1,\boldsymbol{n}}^*$ and $\mu_{2,\boldsymbol{n}}^*$ at every system state, respectively. The specific values of $(\mu_{1,\boldsymbol{n}}^*, \mu_{2,\boldsymbol{n}}^*)$ are further shown in Table I. From Fig. 10 and Table I, we observe that most of the optimal service rates $(\mu_{1,\boldsymbol{n}}^*, \mu_{2,\boldsymbol{n}}^*)$ are on the boundary edge of D_U. But, there do exist some $(\mu_{1,\boldsymbol{n}}^*, \mu_{2,\boldsymbol{n}}^*)$'s that are in the interior of D_U. It demonstrates that Algorithm 1 still works well even

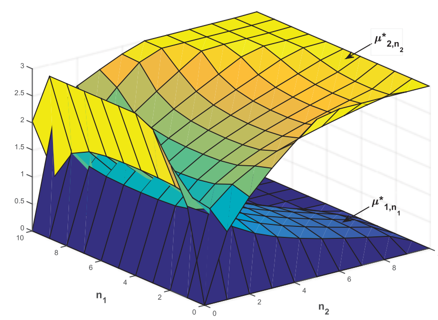

Fig. 10. Curved surface of optimal service rates $(\mu_{1,\boldsymbol{n}}^*, \mu_{2,\boldsymbol{n}}^*)$ in the state space when the cost function is $\phi(\boldsymbol{\mu_n}) = \mu_{1,\boldsymbol{n}}^2 + \mu_{2,\boldsymbol{n}}^2$ and $\varphi(\boldsymbol{n}) = n_1 + n_2$.

TABLE I
OPTIMAL SERVICE RATES $(\mu^*_{1,\boldsymbol{n}}, \mu^*_{2,\boldsymbol{n}})$ AT EVERY SYSTEM STATE WHEN THE COST FUNCTION IS $\phi(\boldsymbol{\mu_n}) = \mu^2_{1,\boldsymbol{n}} + \mu^2_{2,\boldsymbol{n}}$ AND $\varphi(\boldsymbol{n}) = n_1 + n_2$

n_1 \ n_2	0	1	2	3	4	5	6	7	8	9	10
0	(0.00,0)	(0.00,1.25)	(0.00,1.79)	(0.00,2.22)	(0.00,2.58)	(0.00,2.91)	(0.00,3.00)	(0.00,3.00)	(0.00,3.00)	(0.00,3.00)	(0,3)
1	(1.69,0)	(1.44,1.54)	(1.11,1.89)	(0.88,2.12)	(0.70,2.30)	(0.57,2.43)	(0.46,2.54)	(0.36,2.64)	(0.28,2.72)	(0.19,2.81)	(0,3)
2	(2.34,0)	(1.61,1.39)	(1.24,1.76)	(1.00,2.00)	(0.82,2.18)	(0.67,2.33)	(0.55,2.45)	(0.43,2.57)	(0.32,2.68)	(0.22,2.78)	(0,3)
3	(2.84,0)	(1.68,1.32)	(1.30,1.70)	(1.05,1.95)	(0.86,2.14)	(0.71,2.29)	(0.57,2.43)	(0.45,2.55)	(0.33,2.67)	(0.22,2.78)	(0,3)
4	(3.00,0)	(1.70,1.30)	(1.32,1.68)	(1.07,1.93)	(0.88,2.12)	(0.72,2.28)	(0.57,2.43)	(0.44,2.56)	(0.31,2.69)	(0.18,2.82)	(0,3)
5	(3.00,0)	(1.71,1.29)	(1.33,1.67)	(1.07,1.93)	(0.87,2.13)	(0.70,2.30)	(0.54,2.46)	(0.39,2.61)	(0.25,2.75)	(0.10,2.90)	(0,3)
6	(3.00,0)	(1.70,1.30)	(1.32,1.68)	(1.05,1.95)	(0.84,2.16)	(0.65,2.35)	(0.47,2.53)	(0.30,2.70)	(0.12,2.88)	(0.01,2.99)	(0,3)
7	(3.00,0)	(1.69,1.31)	(1.28,1.72)	(0.99,2.01)	(0.75,2.25)	(0.53,2.47)	(0.32,2.68)	(0.11,2.89)	(0.01,2.99)	(0.01,2.99)	(0,3)
8	(3.00,0)	(1.64,1.36)	(1.19,1.81)	(0.85,2.15)	(0.56,2.44)	(0.29,2.71)	(0.03,2.97)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0,3)
9	(3.00,0)	(1.47,1.53)	(0.93,2.07)	(0.52,2.48)	(0.17,2.83)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0,3)
10	(2.03,0)	(0.91,2.07)	(0.05,2.47)	(0.01,2.67)	(0.01,2.85)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0.01,2.99)	(0,3)

when the cost function is not linear and the special properties in Theorems 2–4 may not hold.

VI. DISCUSSION AND CONCLUSION

In this paper, we study the service rate control of a tandem queue with two servers. The system has a power constraint that limits the sum of total service rates. We formulate this problem as an MDP with a constrained action space. The necessary and sufficient condition of the optimal policy is derived. When the operating cost is a linear or concave function, we prove that the optimal service rates can be found on a vertex of the triangle feasible domain $\mathcal{D}_U$, which significantly reduces the optimization complexity. Some optimality properties are also obtained. Algorithm 1 is proposed to iteratively find the optimal service rates. This algorithm is still efficient even when the operating cost is not linear or concave. For example, when the operating cost is convex, we find that Algorithm 1 can decompose the original problem into a series of convex optimization problems that are easy to solve. The numerical examples demonstrate the above results.

Note that the study in this paper is based on the state-dependent service rate. When the service rate is load-dependent or state-independent, we need to do further investigation. For example, for the load-dependent service rate, the power constraint (5) limits the value selection of μ_{1,n_1} and μ_{2,n_2} at different states. The value selection of μ_{1,n_1} at one state may constrain the selection of μ_{2,n_2} at another state since they have to satisfy the common constraint $\mu_{1,n_1} + \mu_{2,n_2} \le U$. That means, the action selection at different states may be correlated, which violates the formulation of a standard MDP. Some new ideas may be utilized to study this problem, such as the event-based optimization or the parametric optimization of MDPs [7], [27]. For another example, when the service rate is static or state-independent, the power constraint can be $\mu_1 + \mu_2 \le U$. In this simple case, we can equivalently treat the tandem queue as two M/M/1 queues, based on Burke's theorem and the Jackson theorem. A closed-form solution of the system average cost can be derived and the optimal service rates can be computed analytically, since it involves only two optimization variables.

Another future research topic is to enhance our study for a general tandem queue with many servers, which is roughly studied in Section IV-D. It is also expected to extend our results to an even more general queueing networks, such as cyclic queues or Jackson networks.

ACKNOWLEDGMENT

The authors would like to thank the four anonymous reviewers and the associate editor for their comments which are very helpful to improve this paper.

REFERENCES

[1] H. S. Ahn, I. Duenyas, and M. E. Lewis, "Optimal control of a two-stage tandem queuing system with flexible servers," *Probab. Eng. Inf. Sci.*, vol. 16, pp. 453–469, 2002.

[2] S. Andradottir and H. Ayhan, "Throughput maximization for tandem lines with two stations and flexible servers," *Oper. Res.*, vol. 53, pp. 516–531, 2005.

[3] B. Ata and S. Shneorson, "Dynamic control of an M/M/1 service system with adjustable arrival and service rates," *Manage. Sci.*, vol. 52, pp. 1778–1791, 2006.

[4] K. R. Balachandran and H. Tijms, "On the d-policy for the M/G/1 queue," *Manage. Sci.*, vol. 21, pp. 1073–1076, 1975.

[5] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA, USA: Athena Scientific, 2012.

[6] A. N. Burnetas, "Customer equilibrium and optimal strategies in Markovian queues in series," *Ann. Oper. Res.*, vol. 208, pp. 515–529, 2013.

[7] X. R. Cao, *Stochastic Learning and Optimization—A Sensitivity-Based Approach*. New York, NY, USA: Springer, 2007.

[8] X. R. Cao and H. F. Chen, "Perturbation realization, potentials, and sensitivity analysis of Markov processes," *IEEE Trans. Autom. Control*, vol. 42, no. 10, pp. 1382–1393, Oct. 1997.

[9] T. B. Crabill, "Optimal control of a service facility with variable exponential service times and constant arrival rate," *Manage. Sci.*, vol. 18, pp. 560–566, 1972.

[10] Y. Dimitrakopoulos and A. Burnetas, "Customer equilibrium and optimal strategies in an M/M/1 queue with dynamic service control," 2011, arXiv:1112.1372.

[11] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*, 4th ed. Hoboken, NJ, USA: Wiley, 2008.

[12] M. Harchol-Balter, *Performance Modeling and Design of Computer Systems—Queueing Theory in Action*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[13] R. Hassin and M. Haviv, *To Queue or Not To Queue: Equilibrium Behavior in Queueing Systems*. New York, NY, USA: Springer, 2003.

[14] D. P. Heyman, "Optimal operating policies for M/G/1 queuing systems," *Oper. Res.*, vol. 16, pp. 362–382, 1968.

[15] D. P. Heyman, "The T-Policy for the M/G/1 queue," *Manage. Sci.*, vol. 23, pp. 775–778, 1977.

[16] D. L. Kaufman, H. S. Ahn, and M. E. Lewis, "On the introduction of an agile, temporary workforce into a tandem queueing system," *Queueing Syst.*, vol. 51, pp. 135–171, 2005.

[17] S. Kumar, N. Bagepalli, A. Patra, P. Quinn, J. Guichard, and J. Iyer, "Service-function chaining," U.S. Patent 14 168 447, Jan. 2014.
[18] A. Lazar, "Optimal flow control of a class of queueing networks in equilibrium," *IEEE Trans. Autom. Control*, vol. 28, no. 11, pp. 1001–1007, Nov. 1983.
[19] Z. Liu, G. Chen, and J. Wu, "Analysis of the optimal resource allocation for a tandem queueing system," 2015, arXiv: 1504.01248.
[20] D. J. Ma and X. R. Cao, "A direct approach to decentralized control of service rates in a closed Jackson network," *IEEE Trans. Autom. Control*, vol. 39, no. 7, pp. 1460–1463, Jul. 1994.
[21] P. Naor, "The regulation of queue size by levying tolls," *Econometrica*, vol. 37, pp. 15–24, 1969.
[22] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 1994.
[23] Z. Rosberg, P. Varaiya, and J. Walrand, "Optimal control of service in tandem queues," *IEEE Trans. Autom. Control*, vol. 27, no. 3, pp. 600–610, Jun. 1982.
[24] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. Hoboken, NJ, USA: Wiley, 1998.
[25] X. Wang, S. Andradottir, and H. Ayhan, "Dynamic server assignment with task-dependent server synergy," *IEEE Trans. Autom. Control*, vol. 60, no. 2, pp. 570–575, Feb. 2015.
[26] R. R. Weber and S. J. Stidham, "Optimal control of service rates in networks of queues," *Adv. Appl. Probab.*, vol. 19, pp. 202–218, 1987.
[27] L. Xia and Q. S. Jia, "Parameterized Markov decision process and its application to service rate control," *Automatica*, vol. 54, pp. 29–35, 2015.
[28] L. Xia, "Service rate control of closed Jackson networks from game theoretic perspective," *Eur. J. Oper. Res.*, vol. 237, no. 2, pp. 546–554, 2014.
[29] L. Xia and B. Shihada, "Max-Min optimality of service rate control in closed queueing networks," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 1051–1056, Apr. 2013.
[30] M. Yadin and P. Naor, "Queuing systems with a removable service station," *Oper. Res. Q.*, vol. 14, pp. 393–405, 1963.
[31] D. D. Yao and Z. Schechner, "Decentralized control of service rates in a closed Jackson network," *IEEE Trans. Autom. Control*, vol. 34, no. 2, pp. 236–240, Feb. 1989.

Li Xia (SM'16) received the B.S. degree and the Ph.D. degree in control theory from Tsinghua University, Beijing, China, in 2002 and 2007, respectively.

He is an Associate Professor in the Center for Intelligent and Networked Systems (CFINS), Department of Automation, Tsinghua University. After graduation, he worked at IBM Research China as a Research Staff Member (2007–2009) and at the King Abdullah University of Science and Technology (KAUST) Saudi Arabia as a Postdoctoral Research Fellow (2009–2011). Then he returned to Tsinghua University in 2011. He was a Visiting Scholar at Stanford University, the Hong Kong University of Science and Technology, etc. He serves/served as an Associate Editor and Program Committee Member of a number of international journals and conferences. His research interests include the methodology research in stochastic learning and optimization, queueing theory, Markov decision processes, reinforcement learning, and the application research in building energy, energy Internet, industrial Internet, Internet of things, etc.

Daniel Miller received the B.S. degree in electrical and computer engineering, and in applied mathematical sciences from Texas A&M University, College Station, TX, USA, in 2014. He is currently working toward the Ph.D. degree in the Department of Electrical Engineering, Stanford University, Stanford, CA, USA.

His industry experience includes work at Silicon Laboratories, and MIT Lincoln Laboratories. His main research interests include the application of data science to medical analytics, renewable energy, and IoT analytics.

Zhengyuan Zhou received the B.S. degree in electrical engineering and computer sciences, in 2012 and the B.A. degree in mathematics from the University of California, Berkeley, Berkeley, CA, USA, in 2012. He is currently working toward the Ph.D. degree in the Department of Electrical Engineering, Stanford University, Stanford, CA, USA.

He has substantial industry experience at Google, Microsoft Research, and Oracle. His main research interests include learning, optimization, game theory, and stochastic systems.

Nicholas Bambos received the M.S. degree in electrical engineering and computer sciences, the M.A. degree in mathematics, and the Ph.D. degree in electrical engineering and computer sciences, all from the University of California, Berkeley, in 1987, 1989, and 1989, respectively.

He is a Professor at Stanford University, Stanford, CA, USA, having a joint appointment in the Department of Electrical Engineering and the Department of Management Science & Engineering. He is the Chair of the Department of Management Science & Engineering, Stanford University. He heads the Network Architecture and Performance Engineering research group at Stanford, conducting research in wireless network architectures, the Internet infrastructure, packet switching, network management, and information service engineering, engaged in various projects of his Network Architecture Laboratory (NetLab). His current technology research interests include high-performance networking, autonomic computing, and service engineering. His methodological interests are in network control, online task scheduling, queueing systems, and stochastic processing networks.