Optimal Routing to Two Parallel Heterogeneous Servers with Resequencing

Serge Ayoun and Zvi Rosberg

Abstract—**Customers arrive to a single *service queue* according to a Poisson process with rate λ, from which they are routed to two parallel heterogeneous and exponential servers whose rates are $\mu_1 \geq \mu_2$. Customers are released from the system after service completion, according to their arrival order—a requirement introducing additional *resequencing delays.* Customers which are delayed due to resequencing are waiting in a *resequencing queue.* We consider the optimal routing problem under the class of *fixed-position* routing policies, that route customers to the faster server from the head of the service queue, and to the slower server from position J. The cost function is taken as the long-run average holding cost of the customers in the system. We show that an optimal stationary policy exists and is of the following type: the faster service is kept active as long as the service queue is not empty. The decision whether or not to route a customer to the slower server is independent of the state of the resequencing queue. If the position J is greater than $J_0 = \lceil \ln(1-\alpha)/\ln \alpha \rceil$, $\alpha = \mu_1/\mu_1 + \mu_2$, then customers are routed to the slower server if and only if the length of the service queue is at least m^* (a threshold policy). We also show that the routing position J_0 is "optimal" in the sense that every policy can be improved by dispatching a customer from position J_0 (if not empty), rather than from position J.**

I. Introduction

In this paper we consider a queueing system (Fig. 1) which is composed of an infinite capacity queue Q attended by two exponential servers operating at rates $\mu_1 > \mu_2$. Customers arrive into the system according to a Poisson process with rate λ, and are assigned consecutive integers which serve as their identifiers. Throughout we assume the stability condition $\lambda < \mu \overset{\text{def}}{=} \mu_1 + \mu_2$. Arriving customers join at the end of queue Q and are routed to one of the servers according to some given routing policy (to be defined below). Customers in service cannot be rerouted.

In many applications of routing in communication network, customers (messages) are released from the service system (the channel and receiver) according to the order of their arrivals. That is, customer i is not released from the system unless he and all customers whose numbers are smaller than i, have finished their service. The waiting time of a customer that has completed his service, for the release of customers with lower sequence numbers, is referred to as *resequencing delay*. Note that resequencing delays are possible since servers are operating at different rates. Moreover, a

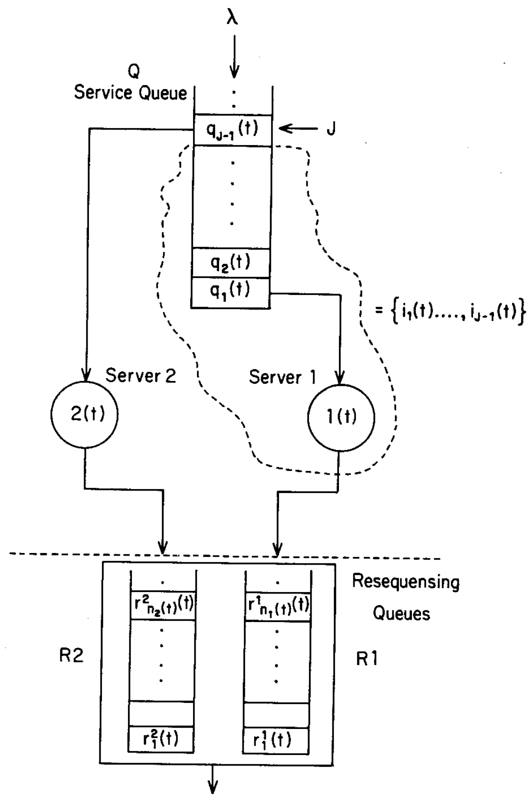


Fig. 1. Routing from position J when $2(t) = 0$ and $1(t) \neq 0$.

routing policy may assign customers from an arbitrary position of the queue. Customers which are being delayed due to resequencing, are waiting in one of two resequencing queues: $R1$ for customers which have been served by server 1, and $R2$ for those which have been served by server 2.

The positions in queue Q from which customers are being routed to the servers (which are perceived as two alternative routes), clearly affect the overall resequencing delays (see [3]). The optimal routing problem with variable positions turned out to be extremely difficult. Therefore, we restrict our attention to *fixed-position* routing policies which route customers to server 1 only from the head of queue Q, and to server 2 only from a fixed position J, $J \geq 2$. By position J we mean the Jth customer among those in server 1 and in queue Q. Beside tractability, this restriction is also motivated by the result in [2]. It has been shown there, that if routing positions are allowed to vary in time, then under light and heavy loads one can take the optimal policy within the class

Manuscript received February 23, 1990; revised May 16, 1991. Paper recommended by Associate Editor At Large, P. R. Kumar.

The authors are with IBM Israel, Science and Technology, Technion City, Haifa 32000, Israel.

IEEE Log Number 9103543.

0018-9286/91$01.00

of fixed-position routings. Also, as it will become apparent, it is not optimal to keep server 1 idle if queue Q is not empty, and therefore the requirement of $J \geq 2$ does not exclude the head of the line.

Let $X(t)$ be a tuple denoting the state of the system at time t (to be defined below) and $|X(t)|$ be the number of customers in the system at that state. A routing policy π is any rule that at every time $t \geq 0$ decides, on the basis of past states and of past decisions up to time t, which idle servers to activate. Policies may leave a server idle even when there is a customer in the corresponding position.

With a holding cost accrued at a fixed rate of 1, the long-run average cost associated with the policy π is then defined by

$$J_\pi(x) \stackrel{\text{def}}{=} \limsup_{T\to\infty} \frac{1}{T} E_x^\pi\left[\int_0^T |X(t)|\, dt\right], \quad \text{for every state } x \qquad (1)$$

where $E_x^\pi[\cdot]$ denotes the expectations with respect to the probability measure induced by the policy π on the process $X = \{X(t),\ t \geq 0\}$ starting in state x. A routing policy π^* is *optimal* if it minimizes (1), i.e., if

$$J_{\pi^*}(x) \leq J_\pi(x)$$

for every policy π and state x.

For the exponential system considered here, the optimization problem associated with (1) falls within the purview of continuous-time Markov decisions processes which are uniformizable, i.e., which are equivalent to uniformized discrete-time Markov decisions processes [6]. The reader is referred for details to [4], where the same problem without resequencing delays is studied. To define the discrete-time decision process, consider that at any given instant, each server is working either on a real customer, if activated, or on a dummy customer otherwise. Dummy customers always return to queue Q upon completing service and incur no contribution to the cost. Transitions are associated either with arrivals or service completions at one of the servers of a customer—either real or dummy—determine free transitions. These free transitions occur according to a Poisson process of rate $\lambda + \mu$. A (free) transition due to an arrival occurs with probability $\lambda/\lambda + \mu$, whereas a transition due to a service completion at server i occurs with probability $\mu_i/\lambda + \mu$. If in state x before a transition, the process will jump after this transition to a state which depends on the current state x and on the action taken under the policy π in use. The cost function for using policy π which corresponds to (1) is then given by

$$V_\pi(x) \stackrel{\text{def}}{=} \limsup_{N\to\infty} \frac{1}{N} E_x^\pi\left[\sum_{m=0}^{N} |X(m)|\right], \qquad x \in S \quad (2)$$

where $X(m)$ now denotes the state sampled at the mth transition. We also need the total β-discounted cost ($0 < \beta < 1$) associated with the policy π, which is defined by

$$V_\pi^\beta(x) \stackrel{\text{def}}{=} E_x^\pi\left[\sum_{m=0}^{\infty} \beta^m |X(m)|\right], \qquad x \in S. \quad (3)$$

The complex structure of the state space of X (see Section II) results in a complex class of stationary policies. A simpler subclass are the policies whose decisions are functions of the length of queue Q only. This subclass will be referred to as the *resequencing-invariant* class. A further simpler subclass are the *threshold policies*.

A policy t_m is a threshold policy with level $m \geq J$ if: i) the first customer in queue Q is routed to server 1 whenever he becomes free; ii) the customer from position J is routed to server 2 when and only when he is free and the number of customers in server 1 and queue Q is at least m.

One result of this study is that the optimal policy can be taken within the resequencing-invariant class. Another result is that for a certain range of positions J, the optimal policy can be taken within the threshold class. We also show that there is a preferable routing position J_0.

For the routing problem without resequencing delays, the routing position J is irrelevant since service requirements are identically distributed. This problem was first studied in [7], where it was conjectured that the optimal policy would be of threshold type. In [1], a version of the problem with N servers was considered under the assumptions that the system has an initial load of n customers and no new customers enter the system, i.e., $\lambda = 0$. A simple policy which minimizes the expected flow time has been determined. This optimal policy has the following simple form [1]:

for $1 < j \leq N$,

set

$$R_j \stackrel{\text{def}}{=} \frac{\mu_1 + \cdots + \mu_{j-1}}{\mu_j} - (j-1) \qquad (4)$$

and define $R_1 = 0$. If there are n customers that remain unprocessed and server j is the fastest server available (i.e., with the largest μ_j), then the idle server j is activated—and a customer dispatched to it—if and only if $n > R_j$.

The conjecture from [7] on the threshold form of the optimal policy was settled in the affirmative in [4] for $N = 2$. Using policy iteration, it has been shown that the optimal policy is of threshold type with threshold level $R(\lambda)$ (which depends on λ). It was also conjectured there that as $\lambda \downarrow 0$, $R(\lambda)$ increases and converges to R_2 given by (4). In [12], simple stochastic coupling arguments were used to prove the optimality of the threshold policy for $N = 2$. Motivated by the conjecture made in [4], it has been shown in [10] (for a general number of parallel servers) and in [8] (for two servers) that the threshold policy above for $\lambda = 0$, is also optimal for *small* enough values of the arrival rate λ.

In light of the results above, one is naturally led to explore the idea that when resequencing delays are introduced, the optimal policy would also be of threshold type. We settle this question in the affirmative only for $J > J_0$.

The issue of resequencing delays in this context has been first introduced in [3], where queueing statistics have been evaluated under the class of fixed-position threshold policies. It has been further shown there, that for a given threshold level m, there is an optimal position J^* from which one should route customers to server 2. This position is given by

$$J^* = \begin{cases} m, & \text{if } m < J_0; \\ J_0, & \text{if } m \geq J_0, \end{cases}$$

$$\text{where } J_0 = \left\lceil \frac{\ln(1-\alpha)}{\ln \alpha} \right\rceil \text{ and } \alpha = \frac{\mu_1}{\mu_1 + \mu_2}. \qquad (5)$$

In other words, when a customer has to be routed to server 2 according to the threshold policy t_m, then the best fixed position is the nearest to J_0. This property of J_0, will be referred to as its "optimality property."

Reviewing the optimality property of J_0 for a threshold policy, and considering the fact that threshold policies may not necessarily be optimal, we are intrigued by another question, whether J_0 has the optimality property for a more general class of policies. We will show that this is indeed the case.

The paper is organized as follows. In Section II, we define the state space and the transitions under fixed-position routings. Section III is subdivided into two parts. In Section III-A, we show that the faster server should be kept active as long as the service queue is not empty. In Section III-B which is further subdivided, we consider the optimal control of the slower server. In Section III-B-1), we show that the optimal control is independent of the state of the resequencing queues. In Section III-B-2), we show the "optimality property" of position J_0, and in Section III-B-3) we show that for $J > J_0$, the optimal policy is of threshold type.

II. The State Process Definitions and Basic Results

In this section, we define the states and the transitions of the Markov decision process that describes our routing problem and examine its state evolution.

A. States and Transitions

We start with the state definition. After every transition t, $t = 0, 1, \cdots$, in the discrete-time decision process, let $n(t)$ denote the number of customers in queue Q, and $e_i(t)$, $i = 1, 2$ denote the state of server i (with the understanding that $e_i(t) = 1$ if server i is busy, and $e_i(t) = 0$ otherwise). To describe the resequencing queues $R1$ and $R2$ we need the following notion.

We say that customer i in a resequencing queue is *being delayed by customer* k_0 if:

i) customer k_0 did not finish service;

ii) $k_0 < i$;

iii) k_0 is the maximal k that satisfies i) and ii).

Thus, customer i is released immediately after the service completion of customer k_0.

Let $l(t)$ be the number of customers in queue $R1$ (after the tth transition), that are being delayed by the customer which is being served by server 2. Here, $l(t) = 0$ if $e_2(t) = 0$. Also (see Fig. 1), denote by $i_1(t) < i_2(t) <, \cdots, < i_{J-1}(t)$, the $J - 1$ customers with the lowest sequence numbers among those in queue Q and server 1 after the tth transition. The number of customers in queue $R2$ that are being delayed by customer $i_m(t)$, $1 \le m \le J - 1$ is denoted by $l_m(t)$.

Observe that the customers in $R1$ can be delayed only by the customer which is being served by server 2, and those in $R2$ by one of the customers in $\{i_m(t) \mid 1 \le m \le J - 1\}$. (These are formally proven in Section II-B below.)

The lengths of the resequencing queues are determined by the tuple

$$R(t) = \big(l(t), \big(l_1(t), \cdots, l_{J-1}(t)\big)\big)$$

which will be referred to as the state of the resequencing queues. Finally, let $k(t)$ be the highest position of the customers in $\{i_m(t) \mid 1 \le m \le J - 1\}$ that would delay the customer being served by server 2 during the tth transition, if he completes his service immediately. If there is no such customer in $\{i_m(t) \mid 1 \le m \le J - 1\}$, or if server 2 is idle, then $k(t) = 0$.

The variable $X(t) = (n(t), e_1(t), e_2(t), R(t), k(t))$ is a natural state variable that may assume values in $S = \mathscr{N} \times \{0, 1\}^2 \times \mathscr{N}^J \times \{1, \cdots, J - 1\}$, where $\mathscr{N} = \{0, 1, \cdots, \}$.

To describe the transitions of the process X it is useful to define the transformations

$$A, D_1, D_2 : S \to S$$

that describe the states to which the process will jump from state x, when a free transition occurs. These transformations correspond to an arrival, a service completion at server 1 and a service completion at server 2, respectively. For the formal definition we need the following notations.

A state $x \in S$ stands for a tuple $x = (n, e_1, e_2, R, k)$, where $R = (l, (l_1, \cdots, l_{J-1}))$ with the understanding that l_m customers in queue $R2$, $1 \le m \le J - 1$, are being delayed by customer i_m. For every $0 \le k \le J - 1$ and $e_2 \in \{0, 1\}$ denote

$$S^1_{k, e_2}(R) = \begin{cases} \big(0, (l_2, \cdots, l_{J-1}, 0)\big), \\ \quad \text{if } k > 0, \text{ or } e_2 = 0 \text{ (implies } k = 0\text{)}; \\ \big(l + 1, (0, \cdots, 0)\big), \\ \quad \text{if } k = 0 \text{ and } e_2 = 1, \end{cases}$$

$$S^2_k(R) = \begin{cases} \big(l, (l_1, \cdots, l_{k-1}, l_k + 1, 0, \cdots, 0)\big), \\ \quad \text{if } k > 0; \\ \big(0, (0, \cdots, 0)\big), \\ \quad \text{if } k = 0. \end{cases}$$

The transformation S^1_{k, e_2} defines the state that queues $R1$ and $R2$ would jump to from state x, when server 1 would complete service of a real customer. Observe that by definition, if $k = 0$ and $e_2 = 1$, then the customer that is being served by server 2 is the "oldest" in the system. Otherwise, the customer that is being served by server 1 is the "oldest." Thus, if $k > 0$ or $e_2 = 0$, when server 1 would complete service of a real customer, this customer and those in queue $R2$ which are being delayed by him, would leave the system. In this case we necessarily have $l = 0$. If $k = 0$ and $e_2 = 1$, we necessarily have $R = (l, (0, \cdots, 0))$, and the customer that would finish service in server 1, would join queue $R1$. (These observations are proven in the next section.)

The transformation S^2_k defines the state that queues $R1$ and $R2$ would jump to from state x, when server 2 would complete service of a real customer. Recall that for $k = 0$ and $e_2 = 1$ we necessarily have $R = (l, (0, \cdots, 0))$. Therefore, when the customer that is being served by server 2 would finish service, he and the customers in queue $R1$ would leave the system. If $k > 0$ and $e_2 = 1$, then the customer that would finish service in server 2, would join queue $R2$ and would be delayed by customer i_k.

Now, the free transitions of process X from state $x \in S$ (when no routings are made), are as follows.

$$A(x) = (n+1, e_1, e_2, R, k),$$

$$D_1(x) = \begin{cases} x, & \text{if } e_1 = 0; \\ (n, 0, e_2, S^1_{k,1}(R), (k-1)^+), & \text{if } e_1 = 1 \end{cases}$$

$$D_2(x) = \begin{cases} x, & \text{if } e_2 = 0; \\ (n, e_1, 0, S^2_k(R), 0), & \text{if } e_2 = 1 \end{cases}$$

where $x^+ = \max\{0, x\}$. The probabilities that a free transition $A(x)$, $D_1(x)$, or $D_2(x)$ occurs, are $\lambda/\lambda + \mu$, $\mu_1/\lambda + \mu$, and $\mu_2/\lambda + \mu$, respectively.

Here, it is convenient to identify a stationary policy π with a function $\pi: S \to \{P_h, P_1, P_2, P_b\}$ as follows. Assume that a free transition—either an arrival or a service completion—occurs that would make the state jump to $x \in S$ if no action were taken. The policy π uses at state x an operator P_a, $a \in \{h, 1, 2, b\}$, that makes the state jump instantaneously from x to $P_a(x)$, where

$$P_h(x) = x;$$

$$P_1(n, 0, e_2, R, k) = (n-1, 1, e_2, R, k), \qquad n \geq 1;$$

$$P_2(n, e_1, 0, R, 0) = (n-1, e_1, 1, R, J-1) \qquad n \geq 1;$$

$$P_b(n, 0, 0, R, 0) = (n-2, 1, 1, R, J-1), \qquad n \geq 2.$$

The operator P_h does not route any customers, P_1 routes the customer from the head of the queue to server 1, P_2 routes the customer from position J to server 2, and P_b does P_1 and P_2. (Notice that from the way we define the position J, the order in P_b is irrelevant.)

B. Basic Results

Since the cost function is linear in the state variable and the total number of customers in the system changes by at most one at every transition, it is well known that an optimal policy exists for the β-discounted problem (associated with (3)), and that it can be taken in the class of Markov stationary policies [11]. One of the conclusions of this study is that the exact same result also holds for the long-run average cost criterion (2). Furthermore, for every stationary policy π, the limit in (2) exists and is independent of the initial state x.

Without loss of generality we may assume that $\lambda + \mu = 1$. Under any stationary policy π, the forward equations of $V^\beta_\pi(x)$ are

$$V^\beta_\pi(x) = |x| + \beta[\lambda V^\beta_\pi(\pi(A(x))) + \mu_1 V^\beta_\pi(\pi(D_1(x))) + \mu_2 V^\beta_\pi(\pi(D_2(x)))] \quad (6)$$

where $\pi(y) \in \{P_h(y), P_1(y), P_2(y), P_b(y)\}$.

In the following lemmas we present some basic properties of the state evolution. The first lemma resolves the order among the customers at any instant.

Denote (see Fig. 1):

- $r^1_s(t)$ The customer (i.e., its sequence number) in the sth position in queue $R1$ at the tth transition (time t).
- $r^2_s(t)$ The customer in the sth position in queue $R2$ at time t.
- $q_s(t)$ The customer in the sth position in queue Q at time t.
- $n_1(t)$ The number of customers in queue $R1$ at time t.
- $n_2(t)$ The number of customers in queue $R2$ at time t.
- $1(t)$ The customer which is being served by server 1 at time t, or 0 if the server is idle.
- $2(t)$ The customer which is being served by server 2 at time t, or 0 if the server is idle.

Lemma 2.1: At every time t and for every occupied positions s and p, or, respectively, 1 and J, in the corresponding queues

a) $q_p(t) < q_s(t)$, for $p < s$;
b) $r^1_p(t) < r^1_s(t)$, for $p < s$;
c) $r^2_p(t) < r^2_s(t)$, for $p < s$;
d) $q_s(t') \leq q_s(t)$, for $t' < t$;
e) $r^1_s(t) < 1(t) < q_1(t)$ for $1(t) > 0$, and $r^1_s(t) < q_1(t)$ for $1(t) = 0$;
f) $r^2_s(t) < 2(t) < q_J(t)$ for $2(t) > 0$, and $r^2_s(t) < q_J(t)$ for $2(t) = 0$;
g) $2(t) < r^1_s(t)$;
h) There exists a p, $1 \leq p \leq J-2$, such that $1(t) < r^2_s(t)$ or $q_p(t) < r^2_s(t)$.

Proof: Properties a)–f) are direct consequences from the facts that customers join at the end of the queues and are being dispatched from fixed positions.

Property g): Customer $r^1_s(t)$ is being delayed by a lower customer. From properties a), b), and e), it could only be customer $2(t)$. Thus, $2(t) < r^1_s(t)$.

Property h): Similarly, for customer $r^2_s(t)$. From properties f) and a) it could only be one of the customers in $\{1(t), q_1(t), \cdots q_{J-2}(t)\}$. □

In the next lemma we show that the two resequencing queues cannot be nonempty at the same time.

Lemma 2.2: At every time t, at least one of the queues $R1$ or $R2$ is empty.

Proof: Suppose that $n_1(t) > 0$ and $n_2(t) > 0$ for some t. As in the proof of Lemma 2.1 g), $n_1(t) > 0$ and a), b), and e) of Lemma 2.1 imply that $2(t) > 0$. Hence, from Lemma 2.1 f) and g)

$$r^2_1(t) < 2(t) < r^1_1(t). \quad (7)$$

However, from part e) of the Lemma, $r^1_1(t) < 1(t) < q_1(t)$, and from part h), $r^1_1(t) < r^2_1(t)$, which is in contradiction with (7). □

The following lemma asserts that at every time t, any customer in queue $R1$ is being delayed by customer $2(t)$. Furthermore, any customer in queue $R2$ is being delayed by one of the customers in $\{1(t), q_1(t), \cdots, q_{J-1}(t)\}$.

For every t denote by $I_1(t), I_2(t), \cdots, I_J(t)$, the set of customers in queue $R2$ that are being delayed by customers $1(t), q_1(t), \cdots, q_{J-1}(t)$, respectively, and $I(t)$ the set of customers in queue $R1$ that are being delayed by customer $2(t)$.

Lemma 2.3:

a) If queue $R1$ is not empty at time t, then $e_2(t) = 1$ and at the next service completion of server 2, queue $R1$ will become empty.

b) Any customer in queue $R2$ is being delayed by one of the customers in $\{1(t), q_1(t), \cdots, q_{J-1}(t)\}$.

c) The customers in every set $I_m(t)$ are consecutive, and those in $I_p(t)$ are smaller than those in $I_{p+1}(t)$.

Proof:

a) From Lemma 2.2, queue $R2$ is empty. From Lemma 2.1 b) and e), all customers in queue $R1$ are smaller than those in queue Q and server 1. Thus, they could be delayed only by customer $2(t)$.

b) If queue $R2$ is not empty, then by Lemma 2.2, queue $R1$ is empty. Thus, the customers in queue $R2$ could be delayed only by those in queue Q and in server 1. The result now follows from parts a) and f) of Lemma 2.1.

c) The result is an immediate consequence of the fact that for $2 \le m \le J$, $q \in I_m(t)$, if and only if $i_m(t) < q < i_{m-1}(t)$ (see Fig. 1 for notation clarification). Similarly, for $q \in I_1(t)$. □

Remark 2.1: If $1(t) = 0$ then $I_1(t) = \phi$. Otherwise, $I_J = \phi$. That is, there are at most $J - 1$ nonempty sets from $\{I_1(t), I_2(t), \cdots, I_J(t)\}$.

III. Optimal Routing

In this section, we consider the β-discounted and the average-cost Markov decision processes. The optimal control is split into two parts: routing to the faster server and routing to the slower server. In Section III-A, we show by probabilistic arguments, that the faster server should be utilized as long as queue Q is not empty. In Section III-B, which is further subdivided, we consider the optimal control of the slower server. In Section III-B-1), we show that the optimal control is independent of the state of the resequencing queues. In Section III-B-2), we show the "optimality property" of position J_0, and in Section III-B-3) we show that for $J > J_0$, the optimal policy is of threshold type.

A. Routing to the Faster Server

In this section we use arguments similar to those presented in [12] in order to show that server 1 is kept active if queue Q is not empty.

To fix the notation, all the proofs in this section are based on pathwise comparison arguments between an original state process X under a given policy π, and another state process $\tilde{X}$ under policy $\tilde{\pi}$ derived from π. The latter is referred to as the *tilde system*, and we use a tilde to denote all relevant quantities in the tilde system.

Lemma 3.1: For every $0 < \beta < 1$, the β-optimal policy has the property that whenever it activates a server, it activates the fastest available one.

Proof: Let π be any given policy and let $X(0) = x$ be an initial state at which π activates server 2 while leaving server 1 idle. By definition, server 2 is activated by the Jth customer from queue Q. We will show that π can be strictly improved.

To simplify notation we may assume without loss of generality, that the customers in queue Q have consecutive numbers starting from 1. (This is possible since only the order among them determine their departure times from the system. Also from the state definition of the resequencing queues, this assumption does not change the system state.)

Define a policy $\tilde{\pi}$ and a corresponding process $\tilde{X}$ as follows. The initial state $\tilde{X}(0) = X(0)$, and a time 0, $\tilde{\pi}$ takes the same action as π, except that it activates server 1 (with customer number 1) instead of server 2. From then on, the realizations of X and $\tilde{X}$ are coupled. This is done by feeding both systems with the same arrival process and assuming that the first service time at server 2 in X equals $T_2 = (\mu_1/\mu_2)\tilde{T}_1$. (Here, T_j is the service time of a customer at server j.) Observe that this coupling is made possible by the fact that $\tilde{T}_1$ is exponentially distributed with parameter μ_1 and therefore T_2 is exponentially distributed with parameter μ_2.

After time 0, policy $\tilde{\pi}$ mimics the actions of policy π (activates the same servers by customers from the appropriate positions) with one exception:

i) Let τ be the first time that π activates server 1. If $\tau < \tilde{T}_1$, then $\tilde{\pi}$ activates server 2 at time τ (instead of 1) with the customer from the Jth position. Since server 1 is busy at that time, this is customer number J. (Observe that in the tilde system, server 2 is available at time τ since $\tau < \tilde{T}_1 < T_2$, which in turn implies that 2 has not been activated under π until time τ, and therefore neither under $\tilde{\pi}$. Also, $\tilde{\pi}$ is feasible since π and the realizations of all r.v.'s in X are known in $\tilde{X}$.)

For all realizations where i) occurs we reach at time τ in both systems, to states which are the same except for the following. In $\tilde{X}$, the customer at server 1 (customer number 1) has been given some service while in X he has not. The converse holds for the customer at server 2 (whose number is J in both systems). To continue the coupling observe that at time τ, the residual service time of customer 1 in $\tilde{X}$ is exponentially distributed with parameter μ_1, which is the same as his service time in X. Moreover, the condition $\{\tilde{T}_1 > \tau\}$ implies the condition $\{T_2 > (\mu_1/\mu_2)\tau\}$ and therefore the residual service time of customer J in X from time $(\mu_1/\mu_2)\tau$ $(> \tau)$ is exponentially distributed with parameter μ_2.

Hence, we can couple the residual service time of customer 1 in $\tilde{X}$ from time τ with his service time in X which starts at time τ. Furthermore, we can also couple the residual service time of J in X from time $(\mu_1/\mu_2)\tau$ $(> \tau)$, with $\tilde{T}_2$ —the service time of J in $\tilde{X}$ from time τ. The latter implies that customer J completes service in X at time $(\mu_1/\mu_2)\tau + \tilde{T}_2$, while in $\tilde{X}$ at time $\tau + \tilde{T}_2$. After time τ, $\tilde{\pi}$ continue to mimic π's actions. From the coupling above and the definition of τ, it is clear that this is feasible. Hence, for all realizations in X where i) occurs we have

$$|\tilde{X}(t)| = \begin{cases} |X(t)|, & \\ \quad \text{for } \tau + \tilde{T}_2 \le t < (\mu_1/\mu_2)\tau + \tilde{T}_2; \\ |X(t)|, & \text{otherwise.} \end{cases}$$

For all other realizations in $\{\tau > \tilde{T}_1\}$, $\tilde{\pi}$ mimics π's actions and we obtain

$$|\tilde{X}(t)| \begin{cases} = |X(t)|, & \text{for } 0 \le t < \tilde{T}_1; \\ = |X(t)| - 1, & \text{for } \tilde{T}_1 \le t < \tau; \\ \le |X(t)|, & \text{for } t \ge \tau. \end{cases} \quad (8)$$

The first equality is straightforward. The second equality follows from the fact that customer 1 leaves $\tilde{X}$ at time $\tilde{T}_1$, while no one is leaving X before time τ (when customer number 1 starts his service). Note that even if customer J had finished his service before τ, he would be delayed by customer 1. The last inequality in (8) is based on the two following observations. From the definition of the fixed-position routing and the fact that at time τ server 1 is idle in $\tilde{X}$ (after customer 1 departs) we have the following.

i) Whenever π routes customer number i, $1 \le i \le J-1$, $\tilde{\pi}$ routes customer number $i+1$. For customer numbers i, $i \ge J+1$, both policies route the same customers.

ii) From time τ and on, all service completions (except for the first completion time of customer J at server 2 in X) are coupled in both systems.

From i) and ii), when customer i, $1 \le i \le J-1$, completes service in X, so does customer $i+1$ in $\tilde{X}$. Therefore, since customer 1 had left $\tilde{X}$ before time τ, customers $\{1, \cdots, J\}$ and those in $R1$ and $R2$ that have been delayed by them had left $\tilde{X}$ earlier than they left X. This holds also for all customers i, $i \ge J+1$, and the last inequality is satisfied.

Since the discount factor is less than 1 and the event where $|\tilde{X}(t)| < |X(t)|$ occurs with positive probability, it follows that $\tilde{\pi}$ strictly improves π. □

Hereafter, we may restrict attention to policies with the property given in Lemma 3.1. Another property of the β-optimal policies is given in the following Lemma.

Lemma 3.2: For every $0 < \beta < 1$, the β-optimal policy keeps server 1 active if queue Q is not empty.

Proof: Let $X(0) = x$ be an initial state such that $Q(0) > 0$ and $e_1(0) = 0$; and π a policy that does not activate server 1 at that state. We show that π can be strictly improved.

From Lemma 3.1 we may assume that π does not activate any server at time 0 (otherwise it is not optimal and we are done). Define the following policy $\tilde{\pi}$ and its corresponding state process $\tilde{X}$ which starts at the same initial state $\tilde{X}(0) = X(0)$, and where all arrivals are coupled to those in the original system. Let σ be an exponential r.v. with parameter μ_1 which is independent of anything else in the systems. At time 0, $\tilde{\pi}$ routes customer 1 to server 1 and his service time is taken as σ. Also, let τ be the first time that π activates a server. (From Lemma 3.1, it would be server 1.)

If $\{\tau < \sigma\}$ we couple at time τ, the residual service time of customer 1 in $\tilde{X}$ to the service time of that customer in X. (Under these realizations both are exponential with parameter μ_1.) We also couple all other service requirements in both systems. From time τ and on, $\tilde{\pi}$ mimics π's actions. By this coupling both systems start at time τ in the same state and therefore have the same state evolutions. Thus, for every realization where $\{\tau < \sigma\}$ we have $|\tilde{X}(t)| = |X(t)|$ for every t.

If $\{\tau > \sigma\}$ then by time $\tau-$, customer 1 and those that had been delayed by him, left the tilde system. In the original system they are still in the system at that time. In this case, the service coupling is done by coupling the services which are given by the servers in both systems, irrespective of the customer identities. Again, from time τ and on, $\tilde{\pi}$ mimics π's server activations with possibly one dummy activation.

The following is a consequence of the different states at time $\tau-$, and the definitions of the fixed-position routing, $\tilde{\pi}$ and the resequencing delay. If at any given time after τ, customer i and those that have been delayed by him leave X, then customers $\{1, \cdots, i-1\}$ and those that have been delayed by them, had left X earlier. Therefore, by the definition of $\tilde{\pi}$, customers $\{1, \cdots, i\}$ and possibly customer $i+1$ and those that have been delayed by them, had left $\tilde{X}$ by that time. Thus, $|\tilde{X}(t)| \le |X(t)| - 1$ for $\sigma < t \le \tau$, and $|\tilde{X}(t)| \le |X(t)|$ otherwise. Since there is a positive probability that $\{\tau > \sigma\}$ and $\beta < 1$, $\tilde{\pi}$ strictly improves π. □

Remark 3.1: Using the same argument as in the proof of Lemma 3.2, it is even simpler to show that keeping server 1 active whenever possible, is also optimal within the nonfixed position routing policies.

Hereafter, we may further restrict attention to policies with the additional property that server 1 be kept active whenever possible.

B. Routing to the Slower Server

In the previous section we proved that the β-optimal policy keeps the faster server active whenever queue Q is not empty. Thus, the optimality problem becomes a problem of routing customers to server 2. That is, at which set of states a customer should be routed to server 2 (when idle), given that the dispatching position is J. Hereafter, a routing decision will be understood as routing to the slower server only.

In the following sections we will derive some useful attributes of the optimal routing policy. The first attribute is that its decisions are independent of the states of the resequencing queues. Another attribute relates to the routing position J. It will be shown that J_0 has an "optimality property" in the sense that one would like to route from the nearest position to J_0. A third attribute is that for $J > J_0$, the optimal policy is of threshold type.

1) The Resequencing-Invariant Property: The next lemma is essential for the proof that state R does not play any role in the optimal routing decision. Observe that from Lemma 2.2, Remark 2.1, and Lemma 3.2, the feasible values of R are of the form $(0, (l_1, \cdots, l_{J-1}))$ or $(l, (0, \cdots, 0))$, where $l_1, \cdots, l_{J-1}$ correspond to the customers in server 1 and in the first $(J-2)$ positions of queue Q. For $R = (0, (0, \cdots, 0))$ we fix the notation [0].

Lemma 3.3: There is a function $h^\beta(R)$ such that for every routing policy π whose decisions are independent of R

$$V_\pi^\beta(n, e_1, e_2, R, k) = V_\pi^\beta(n, e_1, e_2, [0], k) + h^\beta(R). \quad (9)$$

Proof: Let $x_0 = (n, e_1, e_2, R, k)$ and $\tilde{x}_0 = (n, e_1, e_2, [0], k)$ be two initial states, and X and $\tilde{X}$ the processes that are governed by policy π and start at x_0 and $\tilde{x}_0$, respectively. Since π is independent of $R(t)$, we may couple the arrivals and service times in both systems. This is made possible by the same evolutions of $(n(t), e_1(t), e_2(t))$ and $(\tilde{n}(t), \tilde{e}_1(t), \tilde{e}_2(t))$. (Here, we use the tilde notation as in Section II.) There are two cases of R that have to be considered.

Case i): Assume that $R = (0, (l_1, \cdots, l_{J-1}))$. Set $\tau_0 = \tilde{\tau}_0 = 0$, and for every $1 \le j \le J-1$ let $\tau_j(\tilde{\tau}_j)$ be the instant that the customer present at time 0 in position j, leaves the system. By the coupling, τ_j and $\tilde{\tau}_j$ are identical.

Since π routes from position J, the customers that are present at time 0 in the first $(J-1)$ positions, will be routed to server 1. Thus, τ_j is distributed as the sum of j independent geometric r.v.'s with parameter μ_1. By the definition of the resequencing delay, we therefore have

$$|X(t)| \begin{cases} = |\tilde{X}(t)| + \sum_{i=j}^{J-1} l_i, \\ \quad \text{for } \tau_{j-1} \le t < \tau_j,\ 1 \le j \le J-1; \\ = |\tilde{X}(t)|, \qquad \text{for } t \ge \tau_{J-1}. \end{cases}$$

For this case the lemma follows by defining

$$h^\beta(R) = \sum_{i=1}^{J-1} l_i E\left[1 + \beta + \cdots + \beta^{\tau_i - 1}\right]. \tag{10}$$

Case ii): Assume that $R(l, (0, \cdots, 0))$. If $l = 0$ then the lemma is trivial. For $l > 0$, let τ be the instant that the customer present at time 0 in server 2, completes his service. Clearly, τ is geometrically distributed with parameter μ_2. We have

$$|X(t)| \begin{cases} = |\tilde{X}(t)| + l, & \text{for } 0 \le t < \tau; \\ = |\tilde{X}(t)|, & \text{for } t \ge \tau. \end{cases}$$

For this case the lemma follows by defining

$$h^\beta(R) = lE\left[1 + \beta + \cdots + \beta^{\tau - 1}\right]. \tag{11}$$

From (10) and (11), the function

$$h^\beta(R) = lE\left[1 + \beta + \cdots + \beta^{\tau-1}\right] + \sum_{i=1}^{J-1} l_i E\left[1 + \beta + \cdots + \beta^{\tau_i - 1}\right] \tag{12}$$

satisfies (9). Here the expectations are taken with respect to the geometric r.v.'s which are clearly independent of π. □

The function $h^\beta(R)$ represents the accrued discounted cost that is contributed by the customers present at time 0 in the resequencing queues. For later references denote $h^\beta(k) \stackrel{\text{def}}{=} h^\beta((0, (0, \cdots, 0, 1, 0, \cdots, 0)))$, where the 1 corresponds to position k. Observe that from (10)

$$h^\beta(k+1) > h^\beta(k). \tag{13}$$

By using Lemma 3.3 and the following value and policy iterations, we will show that the routing decisions of the β-optimal policy are independent of R.

Let $\mathscr{F}$ be the Banach space of all functions $f: S \to R$ with the norm $\|\cdot\|$ defined by $\|f\| = \sup_{x \in S} |f(x)/\max\{1, |x|\}|$. From (6) and Lemma 3.2 we may define for every stationary policy π, the dynamic programming operator $T_\pi: \mathscr{F} \to \mathscr{F}$, by

$$(T_\pi f)(x) = |x| + \beta\left[\lambda f(\pi(A(x))) + \mu_1 f(\pi(D_1(x))) + \mu_2 f(\pi(D_2(x)))\right] \tag{14}$$

where $\pi(y) \in S$ is the state to which the process jumps from state y after policy π takes the action whether or not to route a customer to server 2 at state y. Also, define the optimal dynamic programming operator $T: \mathscr{F} \to \mathscr{F}$, by

$$(Tf)(x) = |x| + \beta\left[\lambda \min_\pi f(\pi(A(x))) + \mu_1 \min_\pi f(\pi(D_1(x))) + \mu_2 \min_\pi f(\pi(D_2(x)))\right]. \tag{15}$$

Notice that if the decision at state y, $\pi(y)$, for which the $\min_\pi f(\pi(y))$ is attained, is consistently chosen for every y, then Tf defines a stationary policy π' which satisfies

$$(T_{\pi'} f)(x) = (Tf)(x) = \min_\pi (T_\pi f)(x). \tag{16}$$

The procedure by which a new value function is derived by using operator T is known as *value iteration*, and by which a new stationary policy is derived by using T, as *policy iteration*.

Theorem 3.1: The routing decisions of the β-optimal policy are independent of the state of the resequencing queues, R.

Proof: First, we show that if π's decisions are independent of R, so are the decisions of the policy derived by the policy iteration TV_π^β. Then we show that the optimal policy preserves the same property. For every $f \in \mathscr{F}$, define

$$g_f(n, R) = f(n, 1, 0, R, 0) - f(n-1, 1, 1, R, J-1), \quad \text{for } n \ge J-1. \tag{17}$$

let π_0 be a policy whose routing decisions are independent of R, and for every $m \ge 0$ define π_{m+1} as the policy that is derived by the policy iteration $TV_{\pi_m}^\beta$. That is, $T_{\pi_{m+1}} V_{\pi_m}^\beta = TV_{\pi_m}^\beta$. From (15) and (17), $\pi_{m+1}(y)$ is either 0 or 1, depending on whether $g_{V_{\pi_m}^\beta}(n, R)$ is negative or nonnegative, respectively. From Lemma 3.3 it follows that if π_m's decisions are independent of R, then $g_{V_{\pi_m}^\beta}(n, R) = g_{V_{\pi_m}^\beta}(n, [0])$, which implies that π_{m+1}'s decisions are also independent of R.

Since a limit point of $\{\pi_m\}$ does not necessarily exist, we cannot deduce the theorem by the policy iteration procedure. However, we can extract it by the value iteration procedure as follows. Consider the sign of g_{V^β}, where $V^\beta = \inf_\pi V_\pi^\beta$ is the β-value function.

Since π_0's decisions are independent of R, it follows by the argument above that so are π_m's decisions, and by Lemma 3.3, the sign of $g_{V_{\pi_m}^\beta}(n, R)$ is independent of R, $m \ge 0$. since the $\lim_{m \to \infty} V_{\pi_m}^\beta$ exits and equals to V^β (see, e.g., [4, Lemma 3]), the sign of $g_{V^\beta}(n, R)$ is also independent of R.

To conclude the proof, observe that the β-optimal policy π^*, is the solution to the optimality equations $V^\beta = TV^\beta$. Now, from (15), $\pi^*(y) = 1$ if and only if $g_{V^\beta}(n, R) = g_{V^\beta}(n, [0]) \geq 0$, and the solution is independent of R. □

Theorem 3.1 would also apply to the optimal policy with respect to the average cost, if one could guarantee the following limits

$$f(x) = \lim_{\beta_k \to 1} \left(V^{\beta_k}(x) - V^{\beta_k}(0) \right);$$

$$g^* = \lim_{\beta_k \to 1} (1 - \beta_k) V^{\beta_k}(0)$$

for some sequence $\{\beta_k\}$. Then, since $g_{V^\beta}(n, R)$ is independent of R, the result is a straightforward consequence of the following optimality equations for the average cost problem:

$$f(x) + g^* = |x| + \lambda \min_\pi f(\pi(A(x))) + \mu_1 \min_\pi f(\pi(D_1(x))) + \mu_2 \min_\pi f(\pi(D_2(x))).$$

By Lemma 3.7 and Remark 3.3 below, if $\lambda < \mu_1$ the limits above follow from [5, Theorem 3].

Hereafter, we may further restrict attention to policies whose routing decisions (to server 2) are functions of the length of queue Q only. Although this structure is the same as in the problem without resequencing delays, it does not imply that we have the same optimal policy. This is due to the different evolutions of the cost structures.

2) An Optimal Routing Position: In Section I, we described the optimality property of J_0 that has been derived in [3] for the class of fixed-position threshold policies t_m. In this section, we extend this property to a more general class of fixed-position policies.

From the results above, the β-optimal fixed-position routing policy can be taken in the class of stationary policies that are functions of two parameters. i) The set of lengths of queue Q at which a policy routes a customer to server 2 (if idle). ii) The position J from which customers are being dispatched. Note that since the position is fixed, the set in i) is restricted by the position in ii).

We say that a class of policies Π is *routing-invariant*, if the policies differ only by the positions from which customers are being dispatched. That is, for every $\pi \in \Pi$, the sets in i) above are identical. One example is the class of threshold policies with level m and routing positions J, $J \leq m$. Let $J(\Pi)$ be the set of routing positions that correspond to class Π. We will show that the optimality property of J_0 holds for every routing-invariant class.

To proceed, we first characterize J_0 in terms of the expected delay of a customer in position k under two alternative policies. One is a policy that routes customer k to server 2, and customers $\{1, 2, \cdots, k-1\}$ to server 1. The other policy routes customers $\{1, 2, \cdots, k\}$ to server 1.

Let $\{X_i\}$ be a sequence of independent geometric r.v.'s with parameter μ_1, and Y an independent geometric r.v. with parameter μ_2. For $k \geq 1$, denote $X_{(k)} = \sum_{i=1}^k X_i$ and $Z_k = \max\{Y, X_{(k-1)}\}$, where $X_{(0)} = 0$. For every $0 < \beta \leq 1$ define the function

$$\gamma_\beta(k) = E\left[1 + \beta + \cdots + \beta^{Z_k - 1}\right] - E\left[1 + \beta + \cdots + \beta^{X_{(k)} - 1}\right], \qquad k \geq 1. \quad (18)$$

Note that for $\beta = 1$, $\gamma(k) \overset{\text{def}}{=} \gamma_1(k) = E[Z_k] - E[X_{(k)}]$.

The function $\gamma_\beta(k)$ represents the difference in the accrued cost that is contributed by a customer present at time 0 in position k, under the two alternative routing policies above. recalling that $\alpha = \mu_1/\mu_1 + \mu_2$, we obtain by the forward equations.

$$\gamma_\beta(k) = \alpha E\left[\beta^{T_0}\right]\gamma_\beta(k-1) - (1-\alpha) \cdot E\left[\beta^{T_0 + X(k-1)}\left(1 + \beta + \cdots + \beta^{X_k - 1}\right)\right] \quad (19)$$

where T_0 is the time of the first service completion at one of the servers—either real or dummy. (Recall that without loss of generality we assumed that $\lambda + \mu_1 + \mu_2 = 1$. Hence, T_0 is geometrically distributed with parameter $\mu_1 + \mu_2$.)

From (19) it is clear that $\gamma_\beta(k)$ is a decreasing function. Furthermore, there are $\epsilon > 0$ and $\beta_0 < 1$, such that $\lim_{k \to \infty} \gamma_\beta(k) < -\epsilon$ for every $\beta \geq \beta_0$. The latter property is an immediate consequence of the facts that $\lim_{k\to\infty}([Z_k - X_{(k)}] - [X_{(k-1)} - X_{(k)}]) = 0$ with probability one, $E[X_{(k-1)} - X_{(k)}] = -1/\mu_1$ and $\lim_{\beta \to 1} \gamma_\beta(k) = E[Z_k] - E[X_{(k)}]$. Since $\gamma_\beta(1) > 0$, there exists an integer $J_0(\beta) \geq 2$, for which the function $\gamma_\beta(k)$ $(\beta \geq \beta_0)$ becomes strictly negative for the first time.

As our prime interest is the average cost criterion, we consider β's for which $J_0(\beta) = J_0(1)$. Since the $j_0(\beta)$'s are integers and $\gamma_\beta(k) \to \gamma(k)$, it is clear that there exists a $\beta_1 < 1$ such that $J_0(\beta) = J_0(1)$ for every $\beta \geq \beta_1$.

Finally, we show that $J_0(1) = J_0$ (which is defined in (5)). From the memoryless of the geometric distribution it is standard to show that $\gamma(k) = \alpha^k/\mu_2 - (1-\alpha)^k/\mu_1$, from which it follows that $\gamma(k) < 0$ if and only if $(\mu_2/\mu_1)(1 + \mu_2/\mu_1)^{k-1} > 1$. The latter relation implies that $J_0(1) = J_0$.

Now we are ready to prove the optimality property of J_0. We need the following policy transformations, which are applied also to nonfixed-position policies.

i) For every integer $l \geq 1$, every routing policy π and $k \geq 2$, define $T_k^+(l, \pi)$ as the nonstationary policy that differs from π only by the following action at the lth step. If π routes a customer from position k, then $T_k^+(l, \pi)$ routes a customer from position $k+1$, if not empty. Otherwise (at the lth step and other steps), it takes the same routing actions as π. That is, at step l, $T_k^+(l, \pi)$ routes customers to server 2, at the same length of queue Q that π routes, but possibly from a higher position. At other steps, $T_k^+(l, \pi)$ routes customers to server 2, at the same lengths and from the same positions as π routes.

ii) For every integer $l \geq 1$, every routing policy π and $k \geq 3$, define $T_k^-(l, \pi)$ as the policy that differs from π only by the following action at the lth step. If π routes a customer from position k, then $T_k^-(l, \pi)$ routes a customer from position $k-1$. For $l = 1$, these transformations will be denoted by $T_k^+(\pi)$ and $T_k^-(\pi)$, respectively.

In the following lemma, we consider policies that may route from variable positions (with some restrictions), and show that if position k, $k \neq J_0$, is feasible under π, then π can be improved by one of the transformations above.

Lemma 3.4: The following hold for every β, $\beta_1 \leq \beta < 1$:

i) if π routes customers to server 2 from positions larger than or equal to k, $k < J_0$, and k is a feasible position, then $T_k^+(l, \pi)$, $l \geq 1$, is at least as good as π;

ii) If π routes customers to server 2 from positions larger than or equal to $k - 1$, $k > J_0$, and k is a feasible position, then there is a $\beta_2 < 1$ such that $T_k^-(l, \pi)$, $l \geq 1$, strictly improves π for every $\beta_2 \leq \beta < 1$.

Proof: The proof is based on a pathwise comparison of the state process X under policy π to the state process $\tilde{X}$ under policy $T_k^+(\pi)$ (for part i)) or $T_k^-(\pi)$ (for part ii)). To compare realizations, we couple the arrivals and service completions in both systems. (Note that a service completion may correspond to different customers in X and $\tilde{X}$.) The process X and $\tilde{X}$ are identical until step l. Hence, for our pathwise comparison, we may assume without loss of generality that the processes start at step l. Therefore it suffices to prove the lemma for $T_k^+(\pi)$ and $T_k^-(\pi)$.

Part i): Let $x_0 = (n, 1, 0, R, 0)$, $n \geq k$, be an initial state at which π and $T_k^+(\pi)$ routes from different positions. Since for all other initial states $V_\pi^\beta(x) = V_{T_k^+(\pi)}^\beta(x)$, we have to show that

$$V_\pi^\beta(n-1,1,1,R,k-1) - V_\pi^\beta(n-1,1,1,R,k) \geq 0.$$

This is due to the fact that after the first action, X (respectively, $\tilde{X}$) instantaneously jumps to state $(n-1,1,1,R,k-1)$ (respectively, to $(n-1,1,1,R,k)$). From then on, both processes are governed by policy π.

As in the proof of Lemma 3.1, we may assume without loss of generality that the customers in server 1 and in queue Q are numbered by $1, 2, \cdots, n+1$. We will show that for every customer $i \neq k$, its departure times in both systems are the same, while the expected departure time of customer k is smaller in $\tilde{X}$.

Since π routes from positions larger than or equal to k, it follows from the coupling that every customer $i < k$ (and those that at time 0, are being delayed by him), would leave both systems at the same time. Furthermore, the departure times of customer $k+1$ (and those that at time 0, are being delayed by him) would also be the same. This is plain from the fact that $(k+1)$ leaves the system at the first instant at which customers $\{1, 2, \cdots, k+1\}$ have been released. Since states $(n-1,1,1,R,k-1)$ and $(n-1,1,1,R,k)$ differ only by the locations of customers k and $k+1$, it is apparent from the coupling that this instant is the same in both systems.

Every customer $i > k+1$ in both systems, is routed at the same time and to the same server, and its completion time is also the same. Since he would leave the system at the first instant at which he and all preceding customers would have been released, it follows by induction that its departure time must be the same in both systems. Hence, it is left to show that the expected accrued cost due to the delay of customer k, is smaller in $\tilde{X}$.

Let $\tau(\tilde{\tau})$ be the departure time of customer k from system $X(\tilde{X})$. From the identities for the rest of the departure times

$$\begin{aligned} V_\pi^\beta(n-1,1,1,R,k-1) &- V_\pi^\beta(n-1,1,1,R,k) \\ &= (l_k+1)\{E[1+\beta \cdots + \beta^{\tau-1}] \\ &\quad - E[1+\beta \cdots + \beta^{\tilde{\tau}-1}]\}. \end{aligned} \tag{20}$$

Thus, we have to show that the expression within the braces is nonnegative. To prove this, first note that customer k in system $\tilde{X}$ is routed to server 2, if and only if $(k+1)$ in X is routed to server 2. Also note, that since π routes from positions k or higher, customer k in $\tilde{X}$ would definitely be served by server 1, if this server would complete his first service before server 2 does. This event occurs with probability $\mu_1/\mu_1 + \mu_2$.

Let T_0 be as in (19) and T_1 be the number of steps after T_0 that it takes to route customer $(k+1)$ in X to server 2 (and infinite, if he is routed to server 1). Denote by η the conditional probability (conditioned on the state at time T_0) that $\{T_1 < \infty\}$. By using the forward equations from time 0 to T_0, it follows from the definitions of Z_k and $X_{(k)}$ that

$$\begin{aligned} E[(1+\beta &\cdots + \beta^{\tau-1}) - (1+\beta \cdots + \beta^{\tilde{\tau}-1})] \\ &= \frac{\mu_1}{\mu_1+\mu_2} E[\beta^{T_0}((1+\beta \cdots + \beta^{Z_{k-1}-1}) \\ &\quad - (1+\beta \cdots + \beta^{X_{(k-1)}-1}))] \\ &\quad + \frac{\mu_2}{\mu_1+\mu_2} E[\eta\beta^{T_0+T_1}((1+\beta \cdots + \beta^{X_{(k-1)}-1}) \\ &\quad - (1+\beta \cdots + \beta^{Z_k-1}))] \\ &\quad - \frac{\mu_2}{\mu_1+\mu_2} E[(1-\eta)\beta^{T_0+X_{(k-1)}} \\ &\quad \cdot (1+\beta \cdots + \beta^{X_k-1})]. \end{aligned}$$

Thus, from (18), (19) and the fact that

$$\begin{aligned} (1+\beta \cdots + \beta^{X_{(k)}-1}) &= (1+\beta \cdots + \beta^{X_{(k-1)}-1}) \\ &\quad + \beta^{X_{(k-1)}}(1+\beta \cdots + \beta^{X_k-1}) \end{aligned}$$

with probability one, we have

$$\begin{aligned} E[(1+\beta &\cdots + \beta^{\tau-1}) - (1+\beta \cdots + \beta^{\tilde{\tau}-1})] \\ &= \gamma_\beta(k) + \frac{\mu_2}{\mu_1+\mu_2} \\ &\quad \cdot E[\eta\beta^{T_0+X_{(k-1)}}(1+\beta \cdots + \beta^{X_k-1})] \\ &\quad - \frac{\mu_2}{\mu_1+\mu_2} E[\eta\beta^{T_0+T_1}]\gamma_\beta(k) - \frac{\mu_2}{\mu_1+\mu_2} \\ &\quad \cdot E[\eta\beta^{T_0+T_1+X_{(k-1)}}(1+\beta \cdots + \beta^{X_k-1})] \\ &\geq \left(1 - \frac{\mu_2}{\mu_1+\mu_2} E[\eta\beta^{T_0+T_1}]\right)\gamma_\beta(k) \geq 0. \end{aligned}$$

The last inequality follows from the monotonicity of $\gamma_\beta(k)$, the definition of J_0 and the fact that $k < J_0$. This completes the proof of Part i).

Part ii): Let $x_0 = (n, 1, 0, R, 0)$, $n \geq k - 1$, be an initial state at which π and $T_k^-(\pi)$ routes from different positions. Since for all other initial states $V_\pi^\beta(x) = V_{T_k^-(\pi)}^\beta(x)$, we have to show that

$$V_\pi^\beta(n-1,1,1,R,k-2) - V_\pi^\beta(n-1,1,1,R,k-1) < 0.$$

As in Part i), the departure times of every customer $i \neq (k-1)$ in both systems are the same and therefore it suffices to show that

$$E\left[1+\beta \cdots +\beta^{\tilde{\tau}-1}\right] - E\left[1+\beta \cdots +\beta^{\tau-1}\right] < 0. \quad (21)$$

Here, τ and $\tilde{\tau}$ relate to the departure times of customer $(k-1)$. Similarly, alter the definitions of T_1 and η in Part i) by relating them to customer $(k-1)$. Again, by the forward equations we have

$$\begin{aligned}
&E\left[\left(1+\beta \cdots +\beta^{\tilde{\tau}-1}\right) - \left(1+\beta \cdots +\beta^{\tau-1}\right)\right] \\
&= \frac{\mu_1}{\mu_1+\mu_2} E\left[\beta^{T_0}\left(\left(1+\beta \cdots +\beta^{Z_{k-2}-1}\right)\right.\right. \\
&\quad \left.\left. -\left(1+\beta \cdots +\beta^{X_{(k-2)}-1}\right)\right)\right] \\
&\quad + \frac{\mu_2}{\mu_1+\mu_2} E\left[\eta\beta^{T_0+T_1}\left(\left(1+\beta \cdots +\beta^{X_{(k-2)}-1}\right)\right.\right. \\
&\quad \left.\left. -\left(1+\beta \cdots +\beta^{Z_{k-1}-1}\right)\right)\right] \\
&\quad - \frac{\mu_2}{\mu_1+\mu_2} E\left[(1-\eta)\beta^{T_0+X_{(k-2)}}\right. \\
&\quad \left. \cdot\left(1+\beta \cdots +\beta^{X_{k-1}-1}\right)\right] \\
&= \left(1 - \frac{\mu_2}{\mu_1+\mu_2} E\left[\eta\beta^{T_0+T_1}\right]\right)\gamma_\beta(k-1) \\
&\quad + \frac{\mu_2}{\mu_1+\mu_2} E\left[\eta\beta^{T_0+X_{(k-2)}}\right. \\
&\quad \left. \cdot\left(1+\beta \cdots +\beta^{X_{k-1}-1}\right)\left(1-\beta^{T_1}\right)\right]. \qquad (22)
\end{aligned}$$

To complete the proof we have to reduce the last positive term above. Observe that given $\{T_1 < \infty\}$, T_1 is definitely smaller than the first service time that would be given by server 1 after T_0. Therefore, T_1 is stochastically smaller than X_1. Hence, by Jensen inequality the second summand in the right-hand side of (22), can be made arbitrarily close to zero, for β arbitrarily close to one.

Finally, by the definition of J_0 and the fact that $(k-1) > J_0$, it follows that $\gamma_\beta(k-1) < 0$. Thus, there is a β_2, $\beta_1 \le \beta_2 < 1$, such that the right-hand side of (22) is negative for every $\beta_2 \le \beta < 1$. This completes the proof of Part ii). □

Remark 3.2: If $\gamma_\beta(J_0 - 1) > 0$, then in Part i) of Lemma 3.4, $T_k^+(l, \pi)$ strictly improves π.

From Lemma 3.4, $T_k^+(l,\pi)$ and $T_k^-(l,\pi)$, $l \ge 1$, are improvement transformations of policies that route from positions other than J_0. Therefore, they could successively be used to obtain a limiting stationary policy.

Let π_0 be a fixed-position routing policy that routes from position $J \neq J_0$. For every $l \ge 0$, recursively define the nonstationary policy $\pi_{l+1} = T_k^+(l+1, \pi_l)$ (alternatively $\pi_{l+1} = T_k^-(l+1, \pi_l)$). Notice that for every l, π_l satisfies the conditions of Lemma 3.4 and therefore π_{l+1} improves π_l. Let π_∞ be the limiting policy. Policy π_∞ is stationary and routes customers at the same queue lengths (of queue Q) that π_0 does. However, if π_0 routes from position $J > J_0$, then π_∞ routes from position $(J-1)$. If π_0 routes from position $J < J_0$, then π_∞ routes either from position $(J+1)$ (if not empty) or from position J (otherwise). Furthermore, Lemma 3.4 and Remark 3.1 imply that for $J > J_0$, π_∞ is strictly better then π_0, and for $J < J_0$, π_∞ is at least as good as π_0. For $J < J_0 - 1$, and for $J = J_0 - 1$ with $\gamma_\beta(J_0 - 1) > 0$, π_∞ is strictly better then π_0.

The following theorem extends the optimality property of J_0 to any routing invariant class.

Theorem 3.2: The following hold with respect to the β-discounted cost, $\beta_2 \le \beta < 1$, and to the average cost criteria.

a) For every routing-invariant class $\mathbf{\Pi}$ which result in positive recurrent Markov chains:

a.1) if $J_0 \in \boldsymbol{J}(\mathbf{\Pi})$, then the policy $\pi \in \mathbf{\Pi}$ that routes from position J_0 is optimal within $\mathbf{\Pi}$;

a.2) if for every $J \in \boldsymbol{J}(\mathbf{\Pi})$, $J > J_0$, then the policy $\pi \in \mathbf{\Pi}$ that routes from position $J^* = \min\{J \mid J \in \boldsymbol{J}(\mathbf{\Pi})\}$, is optimal within $\mathbf{\Pi}$;

a.3) if for every $J \in \boldsymbol{J}(\mathbf{\Pi})$, $J < J_0$, then the policy $\pi \in \mathbf{\Pi}$ that routes from position $J^* = \max\{J \mid J \in \boldsymbol{J}(\mathbf{\Pi})\}$, is optimal within $\mathbf{\Pi}$.

b) For every fixed-position routing policy π that routes from position J and result in a positive recurrent Markov chain:

b.1) if $J > J_0$, then π is inferior to the policy that routes at the same lengths of queue Q, but from position J_0.

b.2) if $J < J_0$, then π is inferior to the non-fixed position policy that routes at the same lengths of queue Q, but from position $(J+1)$ if not empty, and from J otherwise.

Proof: The proof for the β-discounted cost criterion is immediate from Lemma 3.4 and the discussion that follows. Indeed, within a routing-invariant class $\mathbf{\Pi}$, one can successively improve a policy by gradually increasing (accordingly, decreasing) the routing position within $\boldsymbol{J}(\mathbf{\Pi})$, until one hits J_0, $\max\{J \mid J \in \boldsymbol{J}(\mathbf{\Pi})\}$ or $\min\{J \mid J \in \boldsymbol{J}(\mathbf{\Pi})\}$. Parts a.1), a.2), and a.3) follow, respectively. Furthermore, b.1) is an immediate consequence of a.1), and b.2) follows from Part i) of Lemma 3.4.

The results for the average cost criterion follows form [5] by using the convergence

$$\lim_{\beta \to 1}(1-\beta)V_\pi^\beta(x) = V_\pi(x)$$

which holds for problems with a linear cost structure and continuous state jumps as ours (see [5]). □

3) An Optimal Policy of Threshold Type: In this section we show that if the routing position J is greater than J_0, then an optimal policy with respect to the average cost criterion exists, and is of threshold type. Assume that $J > J_0$ and start with the β-discounted problem, $\beta_0 \le \beta < 1$. Recall that for such J's

$$\gamma_\beta(J-1) < 0, \quad \beta_2 \le \beta < 1. \tag{23}$$

The proof is based on policy iteration and develops along the same lines as the proof in [4], with some changes that are required from our different state space. Define a partial order "$\preceq$" on the states, as follows. Recall that a state x is a tuple $x = (n, e_1, e_2, R, k)$. We say that $x \preceq y$, $x, y \in S$, if at least one of the following conditions hold:

i) $x = y$ (component-wise);

ii) $x = D_1(y)$;

iii) $x = D_2(y)$;

iv) $A(x) = y$;

v) all components of x and y are equal except for one, which is smaller in x.

vi) there is a $z \in S$ such that, $x \preceq z$ and $z \preceq y$.

For every $f \in \mathscr{F}$ we also define the function:

$$\Delta_f(n,k) = \begin{cases} f(n-2,1,1,[0],k) - f(n-3,1,1,[0],k), \\ \quad n \ge 3, \quad 0 \le k \le \min\{n-2, J-1\}; \\ f(0,1,1,[0],1) - f(0,0,1,[0],0), \\ \quad n = 2, \quad k = 0. \end{cases} \tag{24}$$

$$\Delta_f^1(n) = \begin{cases} f(n-1,1,0,[0],0) - f(n-2,1,0,[0],0), \\ \quad n \ge 2; \\ f(0,1,0,[0],0) - f(0,0,0,[0],0), \\ \quad n = 1. \end{cases} \tag{25}$$

In the following lemma we list some properties of $f \in \mathscr{F}$ that propagates to $T_{t_m} f$, $m \ge J$. This will be used to show that under every threshold policy t_m, $V_{t_m}^\beta$ also satisfies the same properties.

Lemma 3.5: If $f \in \mathscr{F}$ satisfies the following properties, so does $T_{t_m} f$, $m \ge J$:

a) for every $x, y \in S$, if $x \preceq y$ then $f(x) \le f(y)$;

b) for every $n \ge 2$, $\Delta_f(n, J-1) \ge h^\beta(\min\{n-2, J-1\})$;

c) for every $n \ge 2$, $\Delta_f^1(n) \ge h^\beta(\min\{n-1, J-1\})$;

d) for every $n \ge 2$, $\Delta_f(n, s) = \Delta_f(n, 0)$, $0 < s < \min\{n-2, J-1\}$;

e) $f(n, e_1, e_2, R, k) = f(n, e_1, e_2, [0], k) + h^\beta(R)$;

f) $f(n-1, 1, 1, [0], k-1) - f(n-1, 1, 1, [0], k) = \gamma_\beta(k)$, $1 \le k \le \min\{n, J-1\}$.

The proof of this lemma is standard but extremely tedious and we do not present the details here. The main lines are as follows. The function $T_{t_m} f$ is represented via (14) and the properties are verified one by one. The full verification is given in [2] and the reader may reproduce it based on the following properties which are easily shown:

$$h^\beta(k+1) > h^\beta(k).$$

$$f(n+1,1,1,[0],k) \ge f(n,1,1,[0]+1_k, J-1), \quad k \le J-1$$

$$f(n,1,1,[0],J-1) - f(n,1,0,[0],0) \ge h^\beta(J-1), \quad n \le J-2.$$

Here, $1_k = (l, (l_1, \cdots, l_k, \cdots, l_{J-1})) = (0, (0, \cdots, 1, \cdots, 0))$.

From Lemma 3.5 one may show by successively using the operator T_{t_m}, that $V_{t_m}^\beta$ also satisfies properties a)–f) of the lemma. Indeed, it is easy to construct in a recursive manner a function f_0 that satisfies properties a)–f). From the lemma it follows that $T_{t_m}^{n+1} f_0 \overset{\text{def}}{=} T_{t_m}(T_{t_m}^n f_0)$, $n \ge 1$, also satisfies these properties. Now, since $\lim_{n\to\infty} T_{t_m}^n f_0 = V_{t_m}^\beta$, we obtain the following corollary.

Corollary 3.1: For every $m \ge J$, the β-discounted cost function under policy t_m satisfies properties a)–f) of Lemma 3.5

The next lemma is the basis of our final result and its proof is similar to that in [4, Lemma 4]. The assumption $J > J_0$ and the property in (23) are crucial for reproducing the proof. The lemma asserts that the new policy that is obtained from $V_{t_m}^\beta$ by the policy iteration procedure, is also of threshold type.

Lemma 3.6: For every m_0, $2 \le J_0 < J \le m_0 < \infty$, there exists an m_1, $J \le m_1 \le m_0 + 1$, such that $T_{t_{m_1}} V_{t_{m_0}}^\beta = T V_{t_{m_0}}^\beta$.

Proof: To prove the lemma we need to explore the properties of the function

$$g_{V_{t_{m_0}}^\beta}(n, R) = V_{t_{m_0}}^\beta(n,1,0,R,0) - V_{t_{m_0}}^\beta(n-1,1,1,R,J-1), \qquad \text{for } n \ge J-1. \tag{26}$$

From Lemma 3.3 it suffices to explore the function $g(n) \overset{\text{def}}{=} g_{V_{t_{m_0}}^\beta}(n, [0])$. This will be carried out by using the forward equations in (6) and representing $g(n)$ in a recursive form. The forward equations depend on the value n and we separately consider all possible cases.

Case i): $1 \le J-1 \le n < m_0 - 2$. (The policy t_{m_0} does not route a customer at queue lengths $n+1$ and below.)

From (6)

$$\begin{aligned} g(n) = \beta\lambda\Big[& V_{t_{m_0}}^\beta(n+1,1,0,[0],0) \\ & - V_{t_{m_0}}^\beta(n,1,1,[0],J-1)\Big] \\ + \beta\mu_1\Big[& V_{t_{m_0}}^\beta(n-1,1,0,[0],0) \\ & - V_{t_{m_0}}^\beta(n-2,1,1,[0],J-2)\Big] \\ + \beta\mu_2\Big[& V_{t_{m_0}}^\beta(n,1,0,[0],0) \\ & - V_{t_{m_0}}^\beta(n-1,1,0,[0]+1_{J-1},0)\Big]. \end{aligned} \tag{27}$$

First, note that the expression in the first braces is $g(n+1)$. Next, add and subtract $V_{t_{m_0}}^\beta(n-2,1,1,[0],J-1)$ within the second braces. From the proof of part ii) of Lemma 3.4 and the assumption $J > J_0$, we have

$$V_{t_{m_0}}^\beta(n-2,1,1,[0],J-1) - V_{t_{m_0}}^\beta(n-2,1,1,[0],J-2) > 0. \tag{28}$$

Thus, by (27) we obtain

$$\begin{aligned} g(n) &> \beta\lambda g(n+1) + \beta\mu_1 g(n-1) \\ &\quad + \beta\mu_2\Big[V^\beta_{t_{m_0}}(n,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-1,1,0,[0]+1_{J-1},0)\Big] \\ &\geq \beta\lambda g(n+1) + \beta\mu_1 g(n-1) \\ &\quad + \beta\mu_2\Big[V^\beta_{t_{m_0}}(n,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-1,1,1,[0],J-1)\Big] \\ &= \beta\lambda g(n+1) + \beta\mu_1 g(n-1) + \beta\mu_2 g(n). \end{aligned}$$

The last inequality follows from Corollary 3.1 and property a) of Lemma 3.5. Since $\lambda + \mu_1 + \mu_2 = 1$, we obtain for this case

$$\begin{aligned} &(1-\beta)g(n) - \beta\lambda(g(n+1) - g(n)) \\ &\quad \geq \beta\mu_1(g(n-1) - g(n)), \\ &\qquad 1 \leq J-1 \leq n \leq m_0 - 2. \end{aligned} \tag{29}$$

The same inequality is obtained for $1 \leq n < J-1$, by defining $g(n) = V^\beta_{t_{m_0}}(n, 1, 0, [0], 0) - V^\beta_{t_{m_0}}(n-1, 1, 1, [0], n)$, $1 \leq n < J-1$.

Case ii): $n = m_0 - 2 > 1$. (The policy t_{m_0} routes a customer at queue length $n+1$, but does not route at queue lengths n and below.)

From (6), the definition in (25), property e) of Lemma 3.5 and (28), we have

$$\begin{aligned} g(n) &= \beta\mu_1\Big[V^\beta_{t_{m_0}}(n-1,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-2,1,1,[0],J-2)\Big] \\ &\quad + \beta\mu_2\Big[V^\beta_{t_{m_0}}(n,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-1,1,0,[0]+1_{J-1},0)\Big] \\ &= \beta\mu_1 g(n-1) + \beta\mu_2\Big[V^\beta_{t_{m_0}}(n,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-1,1,0,[0],0) \\ &\quad + V^\beta_{t_{m_0}}(n-1,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-1,1,0,[0]+1_{J-1},0)\Big] \\ &\geq \beta\mu_1 g(n-1) + \beta\mu_2\Big(\Delta^1_{V^\beta_{t_{m0}}}(n+1) - h^\beta(J-1)\Big) \\ &\geq \beta\mu_1 g(n-1). \end{aligned} \tag{30}$$

The last inequality follows from property c) of Lemma 3.5. The same inequality is obtained for the case $n = m_0 - 2 = 1$. Observe that the assumption $J > J_0 > 2$ and the requirement $m_0 \geq J$, implies that $m_0 \geq 3$.

Case iii): $n = m_0 - 1 \geq 2$. (The policy t_{m_0} routes a customer at queue length n and above.)

From (6)

$$\begin{aligned} g(n) &= \beta\mu_1\Big[V^\beta_{t_{m_0}}(n-1,1,0,[0],0) \\ &\quad - V^\beta_{t_{m_0}}(n-2,1,1,[0],J-2)\Big] \\ &\quad + \beta\mu_2\Big[V^\beta_{t_{m_0}}(n-1,1,1,0,[0],J-1) \\ &\quad - V^\beta_{t_{m_0}}(n-1,1,0,[0]+1_{J-1},0)\Big]. \end{aligned}$$

Again, from (28), the expression in the first braces is greater than $g(n-1)$. Furthermore, from Corollary 3.1 and property a) of Lemma 3.5

$$g(n) \geq \beta\mu_1 g(n-1). \tag{31}$$

Case iv): $n \geq m_0 \geq 3$. (The policy t_{m_0} routes a customer at queue length $n-1$ and above.)

From (6)

$$\begin{aligned} g(n) &= \beta\mu_1\Big[V^\beta_{t_{m_0}}(n-2,1,1,[0],J-1) \\ &\quad - V^\beta_{t_{m_0}}(n-2,1,1,[0],J-2)\Big] \\ &\quad + \beta\mu_2\Big[V^\beta_{t_{m_0}}(n-1,1,1,[0],J-1) \\ &\quad - V^\beta_{t_{m_0}}(n-2,1,1,[0]+1_{J-1},J-1)\Big]. \end{aligned}$$

From (28), the expression in the first braces is positive. From Corollary 3.1 and property a) of Lemma 3.5, the expression in the second braces is also nonnegative. Thus

$$g(n) \geq 0, \quad n \geq m_0. \tag{32}$$

To complete the proof note that from (29)–(32), $g(n)$ satisfies the conditions of the corresponding function in [4, Eq. (10)]. As a consequence, the rest of the proof is identical to the proof of [4, Lemma 4], and our lemma follows. □

The assertion of the next theorem and its proof are identical to [4, Theorem 5]. The proof applies the convergence of the policy iteration to the β-optimal policy.

Theorem 3.3: For every $J > J_0$ and $\beta_2 \leq \beta < 1$:

i) there exists a stationary policy of threshold type, with threshold $m^*(\beta) \leq \infty$;

ii) if $V^\beta_{t_m}(x) < V^\beta_{t_{m+1}}(x)$, for some state x, then $m^*(\beta) \leq m$.

In our final theorem we show by applying [5, Theorem 3], that the optimal policy with respect to the average cost is of threshold type. Here, we cannot reproduce the results from [4, Section IV] since a close form for V_{t_m} is intractable. We will show instead, that Assumptions 1–5 of [5, Theorem 3] hold for our problem. The main assumption requires the following lemma.

Under every threshold policy t_m, $m \leq \infty$, define τ_m (respectively, C_m) as the number of steps (respectively, the accrued cost) until the first return to an empty system. Also, let $E_x(\tau_m)$ and $E_x(C_m)$ be their expected values given that the system starts at state x.

Lemma 3.7: If $\lambda < \mu_1$, then for every state x, $\sup_m E_x(\tau_m) < \infty$ and $\sup_m E_x(C_m) < \infty$.

Proof: For $m = \infty$, τ_m is distributed as the first return time to state 0 in an $M/M/1$ queue with arrival and service rates λ and μ_1, respectively. Since $\lambda < \mu_1$, $E_x(\tau_\infty) < \infty$. We will show that this implies a uniform bound on $E_x(\tau_m)$.

For every $m < \infty$, consider the systems that operate under t_∞ (i.e., $M/M/1$) and under t_m. To compare their paths, we feed them with the same arrival process and couple the service completion times—either real or dummy—in both systems. (To clarify the coupling, imagine the servers producing completion events at rates μ_1 and μ_2, irrespective whether or not a customer is being served. When a completion event occurs in a server that is serving a real customer, this customer would complete his service. Our coupling is referred to these completion events, irrespective of the customer identities that are being served. It is quite clear that for exponential systems, this view is statistically the same as identifying the services with the customers.)

Under t_∞, define σ as the first instant that server 2 completes a dummy service immediately after the system becomes empty. That is, at time $\sigma - 1$ the system just became empty and the next jump was due to a service completion in server 2. Observe that from our coupling, at time σ, both systems are empty. Hence

$$\sup_m E_x(\tau_m) \leq E_x(\sigma). \tag{33}$$

From the renewal property of state 0 in an $M/M/1$ queue and that of the residual completion time, σ can be represented as follows. Let θ_i, $i \geq 1$ be the ith time that the system (under t_∞) is empty, and K be the number of returns to an empty system until server 2 completes a service immediately after the system becomes empty. We have

$$\sigma = (\theta_1 + 1) + (\theta_2 + 1) + \cdots + (\theta_K + 1). \tag{34}$$

Since at every step, the probability of a service completion at server 2 is $\mu_2 > 0$, K is geometrically distributed with parameter μ_2. Furthermore, θ_i, $i \geq 2$ are i.i.d and independent of K. By definition

$$E(\theta_1) = E_x(\tau_\infty) < \infty;$$
$$E(\theta_2) < \max\{E_0(\tau_\infty); E_1(\tau_\infty)\} < \infty \tag{35}$$

where the indexes in $E_0(\cdot)$ correspond to states 0 and 1 in the $M/M/1$ queue.

From (33)-(35) and Wald's lemma

$$\sup_m E_x(\tau_m) \leq E_x(\tau_\infty) + E(K)(1 + E(\theta_2)) < \infty. \tag{36}$$

To prove that $\sup_m E_x(C_m) < \infty$, denote by $A(t)$ the number of arrivals until time t. Given that the system under t_m, $m \leq \infty$, starts at state x

$$C_m \leq |x| + \sigma \cdot A(\sigma).$$

From the Poisson arrivals

$$E_x(C_m) \leq |x| + \lambda \cdot E_x(\sigma^2), \qquad m \leq \infty.$$

The finiteness of $E(\sigma^2)$ follows from the fact that all moments of τ_∞ and K are finite. □

Remark 3.3: Since at most one customer could be served by server 2 at every instant, the same proof implies that the expected return time and accrued cost, given that $X(0) = x$, is uniformly bounded over all admissible policies.

Now we are ready for our final theorem.

Theorem 3.4: For every $J > J_0$, there exists a stationary policy of threshold type t_{m^*} whose level is a limit point $m^* = \lim_{\beta_k \to 1} m^*(\beta_k)$ (m^* could be infinite).

Proof: We consider two cases.

Case i): $\mu_1 \leq \lambda < \mu_1 + \mu_2$. The proof of this case is identical to the proof of the corresponding case in [4, Lemma 7, Theorem 8]. In this case, due to the instability of t_∞, it is also shown that $m^* < \infty$.

Case ii): $\lambda < \mu_1$. In this case we apply Theorem 3 from [5]. Assumptions 1-3 there, trivially hold in our problem. From Theorem 3.3, the policies $t_{m^*(\beta)}$ are β-optimal for every $\beta_2 \leq \beta < 1$. Therefore, Lemma 3.7 implies that Assumptions 4 and 5 there also hold, for a subsequence $\beta_k \to 1$ for which $m^*(\beta_k) \to m^*$. Hence, our theorem is a direct consequence of [5, Theorem 3]. □

To combine this result with the results of Theorem 3.2 about the optimal position, let $t_{m^*(J)}$, $J > J_0$, be the optimal policy with respect to the average cost, given that the routing position is J. From the proof of Lemma 3.4, Part ii), it is clear that $t_{m^*(J_0+1)}$ is at least as good as $t_{m^*(J)}$. Hence, $t_{m^*(J_0+1)}$ is at least as good as any other fixed-position policy that routes customers from $J > J_0$. Furthermore, from Theorem 3.2, part b.1), the policy $t^{J_0}_{m^*(J_0+1)}$ that routes customers whenever $t_{m^*(J_0+1)}$ does, but from position J_0, is at least as good as $t_{m^*(J_0+1)}$. Thus, the following corollary is obtained.

Corollary 3.2: The policy $t^{J_0}_{m^*(J_0+1)}$ is at least as good as any other policy that routes from position $J > J_0$.

References

[1] A. K. Agrawala, E. G. Coffman, Jr., M. R. Garey, and S. K. Tripathi, "A stochastic optimization algorithm minimizing expected flow times on uniform processors," *IEEE Trans. Computers*, vol. C-33, no. 4, pp. 351-356, Apr. 1984.

[2] S. Ayoun, "Optimal control of a queueing system with two heterogeneous servers with resequencing," M.S. thesis, Dep. Electr. Eng., Technion, Haifa, Israel, Feb. 1989.

[3] I. Illiadis and Y. C. Lien, "Resequencing delay for a queueing system with two heterogeneous servers under a threshold-type scheduling," *IEEE Trans. Commun.*, vol. COM-36, pp. 692-702, 1988.

[4] W. Lin and P. R. Kumar, "Optimal control of queueing systems with two heterogeneous servers," *IEEE Trans. Automat. Contr.*, vol. AC-84, pp. 696-703, Aug. 1984.

[5] S. A. Lippman, "Semi-Markov decision processes with unbounded rewards," *Management Sci.*, vol. 19, pp. 717-731, 1973.

[6] —, "Applying a new device in the optimization of exponential queueing systems," *Operations Res.*, vol. 23, pp. 687-710, 1975.

[7] R. L. Larsen, "Control of multiple exponential servers with application to computer systems," Ph.D. dissertation, Tech. Rep. 1041, Univ. Maryland, College Park, MD, 1981.

[8] M. I. Reiman, "Optimal control of a heterogeneous two server queue in light traffic," AT&T Bell Lab., Murray Hill, NJ, 1989.

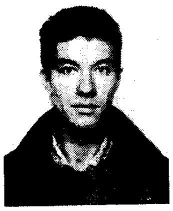
[9] M. I. Reiman and B. Simon, "Open queueing systems in light traffic," *Math. Operations Res.*, 1989.

[10] Z. Rosberg and A. Makowski, "Optimal routing to parallel heterogeneous servers—Small arrival rates," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 789-796, July 1990.

[11] M. Schäl, "Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal," *Z.*

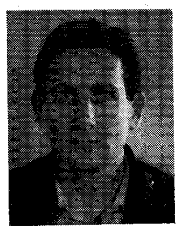
Warscheinlichkeitstheorie Verw. Gebiete, vol. 32, pp. 179-196, 1975.

[12] J. Walrand, "A note on the optimal control of a queueing system with two heterogeneous servers," *Syst. Contr. Lett.*, vol. 4, pp. 131-134, 1984.



Serge Ayoun received the B.Sc. and the M.Sc. degrees in computer engineering from the Technion Institute of Technology, Haifa, Israel, in 1986 and 1989, respectively.

Since 1989 he has been with IBM Israel Science and Technology working in the area of image processing.



Zvi Rosberg received the B.Sc., M.A., and Ph.D. degrees from the Hebrew University, Jerusalem, Israel, in 1971, 1974, and 1978, respectively.

From 1972 to 1978 he was a Senior System Analyst in the General Computer Bureau of the Israeli Government. From 1978 to 1979 he had a Research Fellowship at the Center of Operation Research and Econometrics (CORE), University of Louvain, Belgium. From 1979 to 1980, he was a Visiting Assistant Professor at the University of Illinois. From 1980 to 1989, he was with the Department of Computer Science, Technion, Israel. Since 1990, he has been with IBM Israel, the Science and Technology Center. During 1985-1987 he was on leave at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. His main research interest include probabilistic models of communication networks and computer systems, performance evaluation, queueing theory, and applied probability.