J. Appl. Prob. **55,** 944–967 (2018)
*doi:*10.1017/jpr.2018.59

OPTIMAL ROUTEING IN TWO-QUEUE POLLING SYSTEMS

I. J. B. F. ADAN,* *Eindhoven University of Technology*
V. G. KULKARNI ** AND
N. LEE,** *University of North Carolina*
E. LEFEBER,*** *Eindhoven University of Technology*

Abstract

We consider a polling system with two queues, exhaustive service, no switchover times, and exponential service times with rate μ in each queue. The waiting cost depends on the position of the queue relative to the server: it costs a customer c per time unit to wait in the busy queue (where the server is) and d per time unit in the idle queue (where there is no server). Customers arrive according to a Poisson process with rate λ. We study the control problem of how arrivals should be routed to the two queues in order to minimize the expected waiting costs and characterize individually and socially optimal routeing policies under three scenarios of available information at decision epochs: no, partial, and complete information. In the complete information case, we develop a new iterative algorithm to determine individually optimal policies (which are symmetric Nash equilibria), and show that such policies can be described by a switching curve. We use Markov decision processes to compute the socially optimal policies. We observe numerically that the socially optimal policy is well approximated by a linear switching curve. We prove that the control policy described by this linear switching curve is indeed optimal for the fluid version of the two-queue polling system.

Keywords: Customer routeing; dynamic programming; fluid queue; linear quadratic regulator; Nash equilibrium; polling system; Ricatti equation; individual optimum, social optimum

2010 Mathematics Subject Classification: Primary 60K25
Secondary 90B22

1. Introduction

Polling systems have applications in diverse fields such as manufacturing, telecommunications, time-sharing computer systems, and wireless networks. There is a very large body of research devoted to polling systems, and we refer the reader to [4], [17], [22], and [25] for an overview of the full range of issues for such systems.

Takagi [22] considered a simple polling system consisting of a single server serving N queues in an exhaustive cyclic fashion, which means that it serves the customers in the ith queue until it becomes empty and then moves to queue $i + 1$ (or 1 if $i = N$). Results were

Received 16 August 2016; revision received 18 April 2018.
* Postal address: Department of Industrial Engineering, Eindhoven University of Technology, PO Box 513, 5600MB Eindhoven, The Netherlands. Email address: i.adan@tue.nl
** Postal address: Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC 27599, USA.
*** Postal address: Department of Mechanical Engineering, Eindhoven University of Technology, PO Box 513, 5600MB Eindhoven, The Netherlands.

obtained, for example, about the limiting distribution of the number of customers in the N queues, their means, and their waiting times. These results were quickly extended to service policies other than exhaustive, for example, gated, k-limited, and Bernoulli, as well as noncyclic server routeing, and nonzero switchover times. We refer the reader to the sources mentioned above for the detailed references.

The control issues of polling systems have received less attention than the performance analysis of polling systems. There are several possible control problems arising in polling systems. First, the order in which the queues are served can be determined to optimize system performance (such as weighted expected waiting times), assuming that the service discipline is fixed (such as exhaustive or gated); see [7], [24], and [29]. When the server can switch after every service, the optimal dynamic service order can be studied in greater detail, and may lead to simple rules such as the $c\mu$ rule; see, for example, [10], [12], and [13]. We refer the reader to [25] for many more references in this area.

Customer routeing in polling systems has received less attention. Takine *et al.* [23], Sidi *et al.* [21], and Boon *et al.* [5] studied a Jackson network style routeing of customers among N queues, served cyclically by a single server. The control of customer routeing is the focus of this paper. This subject is also less studied in comparison with the control of server routeing. To the best of the authors' knowledge, the only paper on this topic is [20]. The authors considered the problem where the customers arriving at one of the queues can be routed to any of the others, while customers arriving at the other queues have no flexibility. The authors studied static randomized routeing policies and considered the optimal fraction to be routed to each queue in order to minimize the weighted expected waiting cost. This approach comes somewhat close to our model that we will describe in the next section. However, our cost model is very different from the one in [20] and we consider optimal policies under several scenarios of availability of information, and who is controlling the system. We consider a cyclic exhaustive polling system where every arriving customer needs to be routed to one of the queues, and distinguish three levels of observability of the system. We use the terminology of [9].

- *Unobservable.* We do not know the queue lengths or the position of the server (that is, which queue the server is serving) at decision epochs.
- *Almost unobservable.* We know where the server is, but not the queue lengths, at decision epochs.
- *Fully observable.* We know the position of the server, and the queue lengths, at decision epochs.

If the cost of waiting in all the queues is identical, or joining the busy queue (where there is a server) is cheaper than joining the idle queue (where there is no server), the resulting routeing policies are fairly straightforward. Interesting routeing policies arise when we assume that the waiting cost in the idle queue is less than that in the busy queue. This may be the result of the fact that it is costlier to operate the busy queue as compared to the idle queue. It creates the interesting tradeoff that customers might be able to reduce the waiting cost by joining the idle queue, even though it may increase their total waiting time.

Finally, we consider the control problem from two different viewpoints: the customer (or individually optimal) and the system manager (socially optimal). Socially versus individually optimal policies have been well studied in the literature on queueing; see [11] and [18]. Computation of individually optimal policies becomes complicated when the decisions of later customers can influence a customer's waiting cost. For example, Altman and Shimkin [1] studied individually optimal policies in processor sharing queues, where the decisions by later customers influenced the earlier customer's waiting costs, since they affect the effective service rate available to any customer. The authors also introduced an iterative algorithm to determine Nash equilibria. In our case, the analysis of individually optimal policies is similarly complicated by the fact that a customer's total cost is affected by the behavior of the customers subsequently arriving. We provide a new iterative algorithm to derive Nash equilibria in such a case.

To keep the analysis simple, we consider an exponential system with only two queues and no switchover times. Even for such a simple system, the analysis provides interesting insights, and can be quite involved. We introduce the model and notation in Section 2. The case of the unobservable system is studied in Section 3, and the almost unobservable system in Section 4. In both cases, we study Nash equilibria in the individually optimal analysis, and minimize the long-run average cost rate in the socially optimal case. The case of a fully observable system is studied in Section 5. We present a new iterative algorithm to determine individually optimal policies, and show that such policies can be described by a switching curve. The socially optimal policies can be derived by using negative dynamic programming. We present a novel proof of the existence of average cost optimal policies, but we have not been able to derive structural results in this case. However, numerical experimentation suggests a simple approximate socially optimal routeing policy, which can be described by a linear switching curve. In Section 6 we formulate the problem as a control problem of a fluid polling queueing system, and prove that the approximate policy mentioned in Section 5 is, in fact, optimal in the fluid model. We conclude the paper with a numerical example and summary in Section 7.

2. Polling model

We consider a polling system with two queues; see Figure 1. Customers arrive at this system according to a Poisson process with rate λ. The service times are independent and exponentially distributed with rate μ at each queue. A single server serves the two queues in a cyclic fashion with exhaustive service. The switchover times are assumed to be 0. For stability, we assume that $\rho = \lambda/\mu < 1$. The only costs in the system are waiting costs: it costs a customer c dollars to wait in a queue that is being served (called the busy queue), while it costs him/her d per unit time to wait in a queue that is not being served (called the idle queue). We study how the arrivals should be routed to the two queues in order to minimize expected waiting costs. In the following sections we characterize individually optimal and socially optimal routeing policies

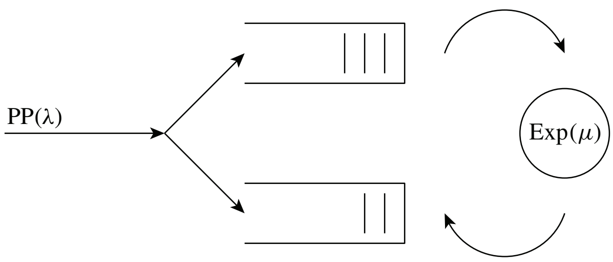

FIGURE 1: The two-queue exponential polling model with exhaustive service. Customers arrive at this system according to a Poisson process with rate λ (PP(λ)). The service times are independent and exponentially distributed with rate μ (Exp(μ)) at each queue.

under the various levels of observability of the system: unobservable, almost unobservable, and fully observable.

3. Unobservable system

Suppose that arriving customers have no information about the state of the system, that is, they do not know where the server is and what the queue lengths are. In this case, the most general policy is described by a single parameter $p \in [0, 1]$. Each customer joins queue 1 with probability $p_1 = p$ and queue 2 with probability $p_2 = 1 - p_1$. We use the notation $\rho_i = \rho p_i$ ($i = 1, 2$). Define L_{ij} as the expected number of customers in queue i given that the server is serving queue j. In the next theorem we state these quantities; see, for example, [3], [27], and [28].

Theorem 1. *Under the above routeing policy, we have, for all* $0 \le p \le 1$,

$$L_{11} = \frac{\rho_1(1-\rho+\rho_1\rho_2+\rho_2^2)}{(1-\rho)(1-\rho+2\rho_1\rho_2)} + 1, \qquad L_{12} = \frac{\rho_1(\rho_1\rho_2+(1-\rho_1)^2)}{(1-\rho)(1-\rho+2\rho_1\rho_2)},$$
$$L_{21} = \frac{\rho_2(\rho_1\rho_2+(1-\rho_2)^2)}{(1-\rho)(1-\rho+2\rho_1\rho_2)}, \qquad L_{22} = \frac{\rho_2(1-\rho+\rho_1\rho_2+\rho_1^2)}{(1-\rho)(1-\rho+2\rho_1\rho_2)} + 1.$$

Using the above theorem, we derive the socially optimal policies in the next theorem.

Theorem 2. (Socially optimal policies.) (i) *If $c > d$, there is a unique socially optimal policy $p = \frac{1}{2}$ (customers join either queue with probability $\frac{1}{2}$).*

(ii) *If $c = d$, all policies $p \in [0, 1]$ are socially optimal.*

(iii) *If $c < d$, there are two socially optimal policies $p = 0$ and $p = 1$ (everyone joins queue 1 or everyone joins queue 2).*

Proof. The socially optimal policy minimizes the expected cost of a customer in the steady state, which is given by

$$C = p_1C_1 + p_2C_2,$$

where C_1 and C_2 are the expected cost of joining queue 1 and 2, respectively. To compute C_1 we argue as follows. The expected number of customers in queue 1, as seen by an arrival, is given by (using PASTA),

$$L_1 = \rho_1 L_{11} + \rho_2 L_{12}.$$

This many customers will have to be served by the server before serving the new arrival at the start of serving queue 1. Hence, the cost incurred by the customer, once the service starts in the queue that he/she has joined, is $c(L_1 + 1)/\mu$. Now, let W be the expected time until the server starts serving queue 1. It is 0 if the server is already at queue 1, which happens with probability ρ_1. With probability ρ_2, the server is at queue 2. Hence, W is the expected busy period of queue 2, started by the number of customers there in the steady state. So we have

$$W = \rho_1 \cdot 0 + \rho_2 \frac{L_{22}}{\mu - \lambda p_2} = \frac{1}{\mu}\frac{\rho_2}{1-\rho_2}L_{22}$$

and, thus,

$$C_1 = \frac{c(L_1+1)}{\mu} + dW = \frac{1}{\mu}\left(c(\rho_1 L_{11} + \rho_2 L_{12}) + d\frac{\rho_2}{1-\rho_2}L_{22} + c\right).$$

A similar argument yields

$$C_2 = \frac{1}{\mu}\left(c(\rho_1 L_{21} + \rho_2 L_{22}) + d\frac{\rho_1}{1-\rho_1}L_{11} + c\right).$$

It then follows that

$$\frac{dC}{dp_1} = \frac{c-d}{\mu}\frac{\rho(\rho^2 + 2(1-\rho))(2p_1 - 1)}{(1-\rho+2\rho_1\rho_2)^2}.$$

Now if $c > d$, the above expression implies that C decreases as p_1 increases from 0 to $\frac{1}{2}$ and increases as p_1 increases from $\frac{1}{2}$ to 1. Hence, C is minimized at $p_1 = \frac{1}{2}$. If $c < d$, C increases as p_1 increases from 0 to $\frac{1}{2}$, and decreases as p_1 increases from $\frac{1}{2}$ to 1. From symmetry both these minima are identical. Thus, C is minimized at $p_1 = 0$ and $p_1 = 1$. When $c = d$, C does not depend on p_1. Hence, the result follows. □

In the next theorem we state the results concerning the individually optimal (Nash equilibrium) policies.

Theorem 3. (Nash equilibrium policies.) (i) *If $c(1-\rho) > d$, there is a unique Nash equilibrium policy $p = \frac{1}{2}$. This policy is socially optimal.*

(ii) *If $c(1-\rho) = d$, every policy $p \in [0, 1]$ is a Nash equilibrium policy, but only $p = \frac{1}{2}$ is socially optimal.*

(iii) *If $c(1-\rho) < d$, there are three Nash equilibrium policies $p = 0$, $p = \frac{1}{2}$, and $p = 1$. Only policy $p = \frac{1}{2}$ is socially optimal if $c > d$, each of them is socially optimal if $c = d$, and $p = 0$ and $p = 1$ are socially optimal if $c < d$.*

Proof. Suppose that arriving customers join queue 1 with probability $p_1 = p$ and queue 2 with probability $p_2 = 1 - p_1$. Now suppose that a smart customer knows how the other customers are behaving and decides to use this system to minimize his/her own waiting costs. If the customer joins queue 1, his/her expected cost is C_1; otherwise, if the customer joins queue 2, his/her expected cost is C_2. It then follows that

$$C_1 - C_2 = \frac{\rho(1-2p_1)(d - c(1-\rho))}{\mu(1-\rho)(1-\rho+2\rho_1\rho_2)}.$$

Consider the $c(1-\rho) > d$ case. If all customers use $p_1 > \frac{1}{2}$ then $C_1 > C_2$ and the smart customer will join queue 2, that is, use $p_1 = 0$; and if all customers use $p_1 < \frac{1}{2}$, he/she will use $p_1 = 1$. Thus, none of these policies is a Nash equilibrium. If all the customers follow policy $p_1 = \frac{1}{2}$, the smart customer is indifferent between the two options and can choose $p_1 = \frac{1}{2}$. Thus, $p_1 = \frac{1}{2}$ is a Nash equilibrium. If $c(1-\rho) = d$, the smart customer is also indifferent, so all policies are a Nash equilibrium. Next, in the $c(1-\rho) < d$ case, it is readily verified that there are three Nash equilibrium policies, $p_1 = 0$, $p_1 = \frac{1}{2}$, and $p_1 = 1$. Together with Theorem 2, this concludes the proof. □

Remark 1. From Theorem 3(iii), there is a follow-the-crowd policy that is individually and socially optimal. There is only one such policy ($p = \frac{1}{2}$) if $c > d$, two such policies ($p = 0$ and $p = 1$) if $c < d$, and three such policies ($p = 0$, $p = \frac{1}{2}$, and $p = 1$) if $c = d$. This phenomena is frequently observed in game theoretic models in queues; see [11].

4. Almost unobservable system

In this section we consider the case of partial information. Specifically, we assume that all customers know which queue is being served by the server, but the individual queue lengths at the two queues are not known. We call the queue that the server is at the busy queue and the other queue the idle queue. We assume that if both queues become empty after a service completion, the server stays at the queue it served last. Thus, the busy queue and idle queue are well defined at all times.

Now the most general policy that a customer can follow is described by a single parameter $p \in [0, 1]$ as follows: join the busy queue with probability $p_1 = p$ and join the idle queue with probability $p_2 = 1 - p_1$. Hence, under this policy, the Poisson arrival rates in the two queues depend on the server location. This system with 'smart customers' was analyzed by Boon *et al.* [6]. Let L_{B} be the expected number of customers in the busy queue, and L_{I} be the expected number of customers in the idle queue, in the steady state, under this policy. In the next theorem, we state these two quantities; see, for example, [6].

Theorem 4. *Under the above policy, we have, for all* $0 \le p \le 1$,

$$L_{\mathrm{B}} = \frac{\rho(1-\rho_1)}{(1-\rho_1)^2 - \rho_2^2}, \qquad L_{\mathrm{I}} = \frac{\rho\rho_2}{(1-\rho_1)^2 - \rho_2^2}.$$

Using the above theorem we derive the socially optimal policies in the next theorem.

Theorem 5. (Socially optimal policies.) (i) *If $c > d$, the socially optimal policy is for everyone to join the idle queue.*

(ii) *If $c = d$, all policies are socially optimal.*

(iii) *If $c < d$, the socially optimal policy is for everyone to join the busy queue.*

Proof. The socially optimal policy minimizes the expected cost of a customer in the steady state. The expected cost of an arriving customer in the steady state is

$$C = p_1 C_{\mathrm{B}} + p_2 C_{\mathrm{I}}$$

with

$$C_{\mathrm{B}} = \frac{1}{\mu}(cL_{\mathrm{B}} + c), \qquad C_{\mathrm{I}} = \frac{1}{\mu}\biggl(dL_{\mathrm{B}}\frac{1}{1-\rho_1} + cL_{\mathrm{I}} + c\biggr),$$

where C_{B} and C_{I} are the expected cost of joining the busy and idle queues, respectively. This can be simplified to

$$C = \frac{c}{\mu(1-\rho)} + \frac{(d-c)\rho_2}{\mu(1-\rho)(1+\rho_2-\rho_1)}.$$

Using $\rho_2 = \rho - \rho_1$, direct calculations yield

$$\frac{dC}{d\rho_1} = \frac{c-d}{\mu(1-\rho_1+\rho_2)^2}.$$

Thus, if $c > d$, C is an increasing function of ρ_1; hence, it is minimized at $\rho_1 = 0$. That is, the socially optimal policy is for everyone to join the idle queue. On the other hand, if $c < d$, C is a decreasing function of ρ_1; hence, it is minimized at $\rho_1 = 1$. Then the socially optimal policy is for everyone to join the busy queue. If $c = d$ then the cost does not depend on ρ_1, and all policies are optimal. $\square$

In the next theorem we state the results concerning the individually optimal (Nash equilibrium) policies.

Theorem 6. (Nash equilibrium policies.) (i) *If* $c(1-\rho) > d$*, the Nash equilibrium policy is to join the idle queue. It is socially optimal.*

(ii) *If* $c(1-\rho) \le d < c$*, the Nash equilibrium policy is to join the busy queue. It is not socially optimal.*

(iii) *If* $c \le d$*, the Nash equilibrium policy is to join the busy queue. It is socially optimal.*

Proof. Suppose that arriving customers join the busy queue with probability p_1 and the idle queue with probability $p_2 = 1 - p_1$. Now suppose that a smart customer knows how the other customers are behaving and decides to use this system to minimize his/her own waiting costs. If the customer joins the busy queue, his/her expected cost is C_{B}; otherwise, if the customer joins idle queue, his/her expected cost is C_{I}. Using the equations for L_{B} and L_{I} from Theorem 4, we obtain

$$C_{\mathrm{B}} - C_{\mathrm{I}} = \frac{\rho}{\mu(1-\rho)} \frac{c(1-\rho) - d}{1 - \rho_1 + \rho_2}. \tag{1}$$

Note that the sign of $C_{\mathrm{B}} - C_{\mathrm{I}}$ does not depend on p_1, the policy followed by all the other customers. We now consider three cases.

Case (i): $c(1-\rho) > d$. Equation (1) implies that $C_{\mathrm{B}} > C_{\mathrm{I}}$ and, hence, the smart customer will also join the idle queue, regardless of what the other customers are doing. Thus, joining the idle queue is a Nash equilibrium. In this case, we also have $c > d$. Hence, from Theorem 5, the socially optimal policy is to join the idle queue. Thus, Nash equilibrium is also the socially optimal policy.

Case (ii): $c(1-\rho) \le d < c$. In this case $C_{\mathrm{B}} \le C_{\mathrm{I}}$ and, hence, the smart customer will join the busy queue, regardless of what the other customers are doing. Hence, joining the busy queue is a Nash equilibrium policy. However, the socially optimal policy is for everyone to join the idle queue. Thus, the Nash equilibrium is to join the busy queue, but the socially optimal policy is to join the idle queue. Individual optimization in this case actually maximizes the social cost.

Case (iii): $c \le d$. The analysis is similar. □

Remark 2. We can write the condition $c(1-\rho) > d$ as

$$\frac{c}{\mu} > \frac{d}{\mu - \lambda}.$$

The left-hand side is the expected cost of waiting in the busy queue for one service time, while the right-hand side is the expected cost of waiting in the idle queue for a busy period initiated by a single customer. It makes sense that the smart customers consider these two costs in order to make a decision, while the social optimizer compares c and d. This results in the Nash equilibrium policies sending more customers to the busy queue than the socially optimal policies.

5. Fully observable system

Now suppose that every customer has complete knowledge of the state of the system; namely, the server location and the length of each queue. We consider how the customers would use this information to decide which queue to join.

5.1. Single smart customer

Suppose that the static routeing policy p is applicable (which does not have to be optimal) and that a special customer wants to use this information to minimize his/her own expected total waiting cost. The special customer can observe the number of customers in the two queues when he/she arrives at the system: i in the busy queue and j in the idle queue. We now consider the choice of queue to join.

If the special customer joins the busy queue, his/her total expected cost is ci/μ. If the special customer joins the idle queue, the total expected waiting cost is $di/(\mu - \lambda p) + cj/\mu$. Here the first term represents the i busy periods that he/she must wait before the server starts serving the idle queue (and making it the busy queue). Thus, it is optimal for the special customer to join the queue under service if

$$\frac{ci}{\mu} < \frac{di}{\mu - \lambda p} + \frac{cj}{\mu},$$

and to join the idle queue if

$$\frac{ci}{\mu} \geq \frac{di}{\mu - \lambda p} + \frac{cj}{\mu}.$$

Clearly, he/she could choose either queue if equality holds. If $d = 0$, the decision rule reduces to joining the shortest queue. Otherwise, the decision rule is a linear switching curve.

5.2. Smart customer population: symmetric Nash equilibrium policies

Now suppose that all customers are smart and each makes a decision to minimize his/her own total expected waiting cost, assuming that other customers will do the same. If $d = 0$, each customer will decide to join the shortest queue, and since this decision is independent of how the other customers behave, this produces a Nash equilibrium. The $d \geq c$ case is also obvious: each customer will decide to join the busy queue, which is a Nash equilibrium. However, the $0 < d < c$ case is not so obvious. In this case, the single smart customer's decision was made under the assumption that all other customers join the busy queue with probability p and the idle queue with probability $1 - p$. However, if every customer chooses the policy derived by the single smart customer then the single customer's analysis is no longer valid. We show that in this case the individually optimal policy is described by a switching curve $h(\cdot)$ such that it is optimal for every customer to join the busy queue in state (i, j) if $j > h(i)$, and that h is a nondecreasing function of i. Note that it is not clear *a priori* that a pure equilibrium policy exists in the fully observable case. There are instances (as described by Altman and Shimkin [1]) where mixed equilibrium policies exist, but no pure ones. In our case, we restrict our quest to the class of pure policies and provide a constructive proof that within this class there exists an equilibrium policy.

So suppose we are given a decision function $f\colon \{0, 1, 2, \ldots\} \times \{0, 1, 2, \ldots\} \to \{I, B\}$ such that $f(i, j) = B(f(i, j) = I)$ implies that an arriving customer that finds i customers in the busy queue and j customers in the idle queue joins the busy (idle) queue. Let $\tau_f(i, j)$ be the expected time until the busy queue empties if the system starts with i customers in the busy queue, and j in the idle queue, under decision function f. Note that $\tau_f(i, j)$ is bounded by $i/(\mu - \lambda)$, which is the expected time to empty the busy queue if all future arrivals are sent to the busy queue. It is individually optimal to join the busy queue if

$$\frac{ci}{\mu} < d\tau_f(i, j+1) + \frac{cj}{\mu},$$

and to join the idle queue if

$$\frac{ci}{\mu} \geq d\tau_f(i, j+1) + \frac{cj}{\mu}.$$

We say that f^* is an individually optimal decision function if

$$f^*(i, j) = \begin{cases} B & \text{if and only if } ci/\mu < d\tau_{f^*}(i, j+1) + cj/\mu, \\ I & \text{if and only if } ci/\mu \geq d\tau_{f^*}(i, j+1) + cj/\mu. \end{cases}$$

The function f^* also describes a Nash equilibrium policy.

We now present a recursive method to compute f^*. We consider a finite horizon system that operates as follows. Let $n \geq 0$ be a given integer (the 'horizon'). Let (i, j) be the initial state of the system ($i \geq 1$, $j \geq 0$). We assume that after n events (arrivals or departures), arrivals are turned off and only departures are allowed to occur, and the system ceases operation once it becomes empty. Let $\delta_n(i, j)$ represent the new state of the system if a customer arrives when the horizon is n, and the system is in state (i, j) and the customer chooses an action that minimizes his/her own cost. Let $\tau_n(i, j)$ be the expected time until the busy queue becomes empty if the system with horizon n starts in state (i, j), and all the arrivals behave in an individually optimal way. We have

$$\tau_0(i, j) = \frac{i}{\mu}, \qquad i \geq 1,\ j \geq 0.$$

This reflects that a zero-horizon system has no more arrivals and, hence, the server completes the work in the current queue after an expected time of i/μ. Now recursively define, for all $n \geq 0,\ i \geq 1,\ j \geq 0$,

$$\delta_n(i, j) = \begin{cases} (i+1, j) & \text{if } ci/\mu < d\tau_n(i, j+1) + cj/\mu, \\ (i, j+1) & \text{if } ci/\mu \geq d\tau_n(i, j+1) + cj/\mu, \end{cases}$$

$$\tau_{n+1}(i, j) = \frac{1}{\lambda + \mu} + \frac{\mu}{\lambda + \mu}\tau_n(i-1, j) + \frac{\lambda}{\lambda + \mu}\tau_n(\delta_n(i, j)),$$

where $\tau_n(0, j) = 0$.

In the next lemma we formulate the monotonicity properties of $\tau_n(i, j)$.

Lemma 1. *For all* $n \geq 0,\ i \geq 1,\ j \geq 0$,

$$\tau_n(i, j) \leq \tau_n(i, j+1), \tag{2}$$

$$\tau_n(i, j+1) \leq \tau_n(i+1, j), \tag{3}$$

$$\tau_n(i, j) \leq \frac{i}{\mu - \lambda}, \tag{4}$$

$$\tau_n(i, j) \leq \tau_{n+1}(i, j). \tag{5}$$

Proof. By induction. For $n = 0$, we have

$$\tau_0(i, j) = \tau_0(i, j+1) = \frac{i}{\mu} < \frac{i+1}{\mu} = \tau_0(i+1, j),$$

$$\tau_1(i, j) \geq \frac{1}{\lambda + \mu} + \frac{\mu}{\lambda + \mu}\frac{i-1}{\mu} + \frac{\lambda}{\lambda + \mu}\frac{i}{\mu} = \frac{i}{\mu} = \tau_0(i, j).$$

Hence, (2)–(5) hold for $n = 0$. Now assume that (2)–(5) hold for n. Then we show that these inequalities also hold for $n + 1$. To establish (2) for $n + 1$, consider

$$\begin{aligned}\tau_{n+1}(i, j + 1) - \tau_{n+1}(i, j) = {}& \frac{\mu}{\lambda + \mu}(\tau_n(i - 1, j + 1) - \tau_n(i - 1, j)) \\ &+ \frac{\lambda}{\lambda + \mu}(\tau_n(\delta_n(i, j + 1) - \tau_n(\delta_n(i, j))).\end{aligned}$$

The first term is nonnegative due to (2). If $\delta_n(i, j + 1) = (i + 1, j + 1)$ then for both $\delta_n(i, j) = (i + 1, j)$ and $\delta_n(i, j) = (i, j + 1)$, we can conclude that the second term is nonnegative by an application of (2) and (3). If $\delta_n(i, j+1) = (i, j+2)$ then $\delta_n(i, j) = (i, j+1)$ using (2) and, thus, we can again conclude that the second term is nonnegative by (2). For (3), we have

$$\begin{aligned}\tau_{n+1}(i + 1, j) - \tau_{n+1}(i, j + 1) = {}& \frac{\mu}{\lambda + \mu}(\tau_n(i, j) - \tau_n(i - 1, j + 1)) \\ &+ \frac{\lambda}{\lambda + \mu}(\tau_n(\delta_n(i + 1, j)) - \tau_n(\delta_n(i, j + 1))).\end{aligned}$$

The first term on the right-hand side is nonnegative due to (3). If $\delta_n(i+1, j) = (i+2, j)$ then for both $\delta_n(i, j + 1) = (i + 1, j + 1)$ and $\delta_n(i, j + 1) = (i, j + 2)$, we find that the second term is nonnegative by (repeated) application of (3). If $\delta_n(i + 1, j) = (i + 1, j + 1)$, we arrive at the same conclusion. From (2),

$$\tau_{n+1}(i, j) \le \frac{1}{\lambda + \mu} + \frac{\mu}{\lambda + \mu}\tau_n(i - 1, j) + \frac{\lambda}{\lambda + \mu}\tau_n(i + 1, j)$$

and, thus, using (4),

$$\tau_{n+1}(i, j) \le \frac{1}{\lambda + \mu} + \frac{\mu}{\lambda + \mu}\frac{i - 1}{\mu - \lambda} + \frac{\lambda}{\lambda + \mu}\frac{i + 1}{\mu - \lambda} = \frac{i}{\mu - \lambda}.$$

Finally, to prove (5) for $n + 1$,

$$\begin{aligned}\tau_{n+2}(i, j) - \tau_{n+1}(i, j) = {}& \frac{\mu}{\lambda + \mu}(\tau_{n+1}(i - 1, j) - \tau_n(i - 1, j)) \\ &+ \frac{\lambda}{\lambda + \mu}(\tau_{n+1}(\delta_{n+1}(i, j)) - \tau_n(\delta_n(i, j))).\end{aligned}$$

The first term is nonnegative due to (5). If $\delta_{n+1}(i, j) = (i + 1, j)$ then for both $\delta_n(i, j) = (i + 1, j)$ and $\delta_n(i, j) = (i, j + 1)$, it follows that the second term is nonnegative by an application of (3) for $n + 1$ and (5). If $\delta_{n+1}(i, j) = (i, j + 1)$ then also $\delta_n(i, j) = (i, j + 1)$ by (2), and, thus, the second term is nonnegative using (5). □

In the following theorem we state that this recursive procedure generates an individually optimal decision function f^*.

Theorem 7. *For all* $i \ge 1,\ j \ge 0$,

$$\lim_{n\to\infty} \tau_n(i, j) = \tau(i, j) = \tau_{f^*}(i, j), \qquad \lim_{n\to\infty} \delta_n(i, j) = \delta(i, j),$$

where f^* *is defined as*

$$f^*(i, j) = \begin{cases} B & \text{if and only if } \delta(i, j) = (i + 1, j), \\ I & \text{if and only if } \delta(i, j) = (i, j + 1). \end{cases}$$

Proof. By virtue of (4) and (5), the sequence $\tau_n(i, j)$ is nondecreasing in n and bounded. Hence, the limits of $\tau_n(i, j)$ and $\delta_n(i, j)$ exist and satisfy, for all $i \geq 1,\ j \geq 0$,

$$\delta(i, j) = \begin{cases} (i+1, j) & \text{if } ci/\mu < d\tau(i, j+1) + cj/\mu, \\ (i, j+1) & \text{if } ci/\mu \geq d\tau(i, j+1) + cj/\mu, \end{cases}$$

and

$$\tau(i, j) = \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\tau(i-1, j) + \frac{\lambda}{\lambda+\mu}\tau(\delta(i, j)),$$

where $\tau(0, j) = 0$. The equation for $\tau(i, j)$ follows from two properties:

(a) in case of a tie, we choose the action I, and

(b) τ_n is nondecreasing in n, from (5).

Without the latter monotonicity property, the strict inequality $ci/\mu < d\tau_n(i, j+1) + cj/\mu$ (even if it holds for all n) could turn into an inequality in the limit, thereby changing the limit policy. The expected values $\tau_{f^*}(i, j)$ satisfy, for all $i \geq 1,\ j \geq 0$,

$$\tau_{f^*}(i, j) = \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\tau_{f^*}(i-1, j) + \frac{\lambda}{\lambda+\mu}\tau_{f^*}(\delta(i, j)),$$

where $\tau_{f^*}(0, j) = 0$. To prove $\tau_{f^*}(i, j) = \tau(i, j)$, consider $v(i, j) = \tau_{f^*}(i, j) - \tau(i, j)$ satisfying

$$v(i, j) = \frac{\mu}{\lambda+\mu}v(i-1, j) + \frac{\lambda}{\lambda+\mu}v(\delta(i, j)), \qquad i \geq 1,\ j \geq 0,$$

or in vector-matrix notation

$$\boldsymbol{v} = \boldsymbol{P}\boldsymbol{v}, \tag{6}$$

where $\boldsymbol{P}$ is the (transient) transition probability matrix with

$$P_{(i,j),(i-1,j)} = 1 - P_{(i,j),\delta(i,j)} = \frac{\mu}{\lambda+\mu}, \qquad i \geq 1,\ j \geq 0.$$

Iterating (6) yields $\boldsymbol{v} = \boldsymbol{P}^n\boldsymbol{v}$. Since transitions are restricted to neighboring states, we have $P^n_{(i,j),(k,l)} = 0$ for all (k, l) with $k > i + n$. Hence, since $\tau(i, j)$ and $\tau_{f^*}(i, j)$ are bounded by $i/(\mu - \lambda)$,

$$|v(i, j)| = |(\boldsymbol{P}^n\boldsymbol{v})_{(i,j)}| \leq (\boldsymbol{P}^n\mathbf{1})_{(i,j)}\frac{2(n+i)}{\mu-\lambda}, \tag{7}$$

where $\mathbf{1}$ is the vector of all ones and $(\boldsymbol{P}^n\mathbf{1})_{(i,j)}$ is the probability that the Markov chain P does not reach the absorbing boundary $i = 0$ in n transitions when starting in (i, j). This probability is bounded by $\mathbb{P}(X_i > n)$, where X_i is the number of transitions to reach 0 of the random walk on the nonnegative integers with one-step probabilities $P_{j,j-1} = 1 - P_{j,j+1} = \mu/(\lambda + \mu)$ when it starts in state i. This random walk reflects that all future arrivals are sent to the busy queue. By Markov's inequality, $\mathbb{P}(X_i > n) \leq \mathbb{E}(X_i^2)/n^2$. Hence, from (7),

$$|v(i, j)| \leq \mathbb{P}(X_i > n)\frac{2(n+i)}{\mu-\lambda} \leq \frac{\mathbb{E}(X_i^2)}{n^2}\frac{2(n+i)}{\mu-\lambda}.$$

Letting $n \to \infty$, we conclude that $v(i, j) = 0$, which completes the proof. □

Next we describe the main structural properties of the policy f^*.

Theorem 8. *It holds that* $f^*(i, j) = B$ *for all* $1 \le i \le j$.

Proof. For $1 \le i \le j$,

$$\frac{ci}{\mu} < d\tau(i, j+1) + \frac{cj}{\mu},$$

since $\tau(i, j+1) > 0$. Hence, $f^*(i, j) = B$ by definition. □

From the above theorem we see that if the busy queue is no longer than the idle queue, then the individually optimal decision for any customer is to join the busy queue. In the theorem below we state the monotonicity of the individually optimal policy in j.

Theorem 9. *For all* $i \ge 1$, $j \ge 0$, *if* $f^*(i, j) = B$ *then* $f^*(i, j+1) = B$.

Proof. Suppose that $f^*(i, j) = B$ for some $i \ge 1$, $j \ge 0$. This implies that

$$\frac{ci}{\mu} < d\tau(i, j+1) + \frac{cj}{\mu} < d\tau(i, j+1) + \frac{c(j+1)}{\mu} \le d\tau(i, j+2) + \frac{c(j+1)}{\mu},$$

where the last inequality follows from (2) by taking $n \to \infty$. Hence, $f^*(i, j+1) = B$. □

To prove monotonicity in i, we first need a technical result.

Theorem 10. *Suppose that* $f^*(k, j) = B$ *for every* $1 \le k \le i$ *and fixed* $j \ge 0$. *Then* $\tau(k, j)$ *is concave for* $1 \le k \le i$.

Proof. Fix $j \ge 0$. First, we show by induction that, for all $1 \le k \le i-1$,

$$\tau(k+1, j) - \tau(k, j) \le \frac{1}{\mu - \lambda}. \tag{8}$$

For $k = 1$,

$$\begin{aligned}\tau(2, j) - \tau(1, j) &= \tau(2, j) - \frac{1}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}\tau(2, j) \\ &= \frac{\mu}{\lambda+\mu}\tau(2, j) - \frac{1}{\lambda+\mu} \\ &\le \frac{\mu}{\lambda+\mu}\frac{2}{\mu-\lambda} - \frac{1}{\lambda+\mu} \\ &= \frac{1}{\mu-\lambda},\end{aligned}$$

where the inequality follows from the bound $\tau(2, j) \le 2/(\mu - \lambda)$. Now we assume that (8) holds for $k \le i-2$ and then show that it also holds for $k+1$. Thus,

$$\begin{aligned}&\tau(k+2, j) - \tau(k+1, j) \\ &\quad = \tau(k+2, j) - \frac{1}{\lambda+\mu} - \frac{\mu}{\lambda+\mu}\tau(k, j) - \frac{\lambda}{\lambda+\mu}\tau(k+2, j) \\ &\quad = \frac{\mu}{\lambda+\mu}[\tau(k+2, j) - \tau(k, j)] - \frac{1}{\lambda+\mu} \\ &\quad = \frac{\mu}{\lambda+\mu}[\tau(k+2, j) - \tau(k+1, j)] + \frac{\mu}{\lambda+\mu}[\tau(k+1, j) - \tau(k, j)] - \frac{1}{\lambda+\mu}.\end{aligned}$$

Hence,

$$\tau(k+2,j)-\tau(k+1,j)=\frac{\mu}{\lambda}[\tau(k+1,j)-\tau(k,j)]-\frac{1}{\lambda}\leq\frac{\mu}{\lambda(\mu-\lambda)}-\frac{1}{\lambda}=\frac{1}{\mu-\lambda},$$

which concludes the proof of (8).

Next, to establish concavity, we have, for $1\leq k\leq i-2$,

$$\begin{aligned}
&[\tau(k+2,j)-\tau(k+1,j)]-[\tau(k+1,j)-\tau(k,j)]\\
&\quad=\tau(k+2,j)-2\left[\frac{1}{\lambda+\mu}+\frac{\mu}{\lambda+\mu}\tau(k,j)+\frac{\lambda}{\lambda+\mu}\tau(k+2,j)\right]+\tau(k,j)\\
&\quad=\frac{\mu-\lambda}{\mu+\lambda}[\tau(k+2,j)-\tau(k,j)]-\frac{2}{\lambda+\mu}\\
&\quad\leq\frac{\mu-\lambda}{\mu+\lambda}\frac{2}{\mu-\lambda}-\frac{2}{\lambda+\mu}\\
&\quad=0,
\end{aligned}$$

where the inequality follows by a repeated application of (8). □

With the above result we can prove monotonicity in i.

Theorem 11. *For all $i\geq 2$, $j\geq 0$, if $f^*(i,j)=B$ then $f^*(i-1,j)=B$.*

Proof. Fix $i\geq 2$. By downward induction we will prove for $j\geq 0$ that $f^*(i,j)=B$ implies that $f^*(k,j)=B$ for all $1\leq k\leq i$. By Theorem 8, this holds for $j\geq i$. Now we assume that it holds for j and then show that it also holds for $j-1$. Suppose that $f^*(i,j-1)=B$. If $j>1$ then $f^*(1,j-1)=B$ by Theorem 8. To show that this is also valid for $j=1$, first note that $f^*(i,0)=B$ implies that

$$\frac{ci}{\mu}<d\tau(i,1)$$

and, thus, using $\tau(i,1)\leq i/(\mu-\lambda)$,

$$\frac{d}{\mu-\lambda}>\frac{c}{\mu}.$$

Hence,

$$d\tau(1,1)=d\left[\frac{1}{\lambda+\mu}+\frac{\lambda}{\lambda+\mu}\tau(2,1)\right]\geq d\left[\frac{1}{\lambda+\mu}+\frac{\lambda}{\lambda+\mu}\frac{2}{\mu-\lambda}\right]=\frac{d}{\mu-\lambda}>\frac{c}{\mu},$$

so $f^*(1,0)=B$. Since $f^*(1,j-1)=f^*(i,j-1)=B$, we have

$$\frac{c}{\mu}<d\tau(1,j)+\frac{c(j-1)}{\mu},\qquad \frac{ci}{\mu}<d\tau(i,j)+\frac{c(j-1)}{\mu},$$

and, thus, for $1\leq k\leq i$,

$$\begin{aligned}
\frac{ck}{\mu}&=\frac{i-k}{i-1}\frac{c}{\mu}+\frac{k-1}{i-1}\frac{ci}{\mu}\\
&<d\left[\frac{i-k}{i-1}\tau(1,j)+\frac{k-1}{i-1}\tau(i,j)\right]+\frac{c(j-1)}{\mu}\\
&\leq d\tau(k,j)+\frac{c(j-1)}{\mu},
\end{aligned}$$

where the second inequality follows from Theorem 10. Hence, $f^*(k,j-1)=B$. □

Theorems 9 and 11 imply that the individually optimal policy is described by a switching curve $h(\cdot)$ such that it is optimal for a customer to join the busy queue in state (i, j) if $j > h(i)$, and that h is a nondecreasing function of i. Note that h depends on the costs c and d. This completes the discussion of the individually optimal policy.

5.3. Socially optimal policy

Finally, suppose there is a central controller who can route the customers so as to minimize the long-run expected waiting cost per unit time. Let $Z(t)$ be the total number of customers in the system (those in the busy queue plus those in the idle queue) at time t. We begin with a straightforward but important observation.

Lemma 2. *It holds that* $\{Z(t),\ t \geq 0\}$ *is the queue length process in an* M/M/1 *queue regardless of the routeing policy followed.*

Proof. The total arrival process to the system is a Poisson process with rate λ, the service times are independent and exponential with rate μ, and the polling service discipline is work conserving. Hence, the lemma follows. □

Now let $X(t)$ be the number of customers in the busy queue and $Y(t)$ be the number in the idle queue at time t. Then the total cost $C(t)$ over $(0, t]$ is

$$C(t) = \int_0^t (cX(u) + dY(u))\,\mathrm{d}u, \qquad t \geq 0.$$

The process $\{C(t),\ t \geq 0\}$ does depend on the routeing policy. Let T_n be the nth time when the system busy cycle ends, that is, when $Z(t)$ reaches 0. Let

$$C_n = \int_{T_n}^{T_{n+1}} (cX(u) + dY(u))\,\mathrm{d}u$$

be the total cost incurred over the interval $(T_n, T_{n+1}]$. An important implication of the above lemma is that $\{C(t),\ t \geq 0\}$ is a (delayed in the $Z(0) > 0$ case) renewal reward process, since $\{(C_n, T_{n+1} - T_n),\ n \geq 1\}$ is a sequence of independent and identically distributed bivariate random variables. Furthermore, $\{T_{n+1} - T_n,\ n \geq 1\}$ is a sequence of independent busy cycles in an M/M/1 queue. Hence, their common distribution does not depend on the routeing policy, and

$$\mathbb{E}(T_{n+1} - T_n) = \frac{\mu}{\lambda(\mu - \lambda)} < \infty.$$

Then, from the results on renewal reward processes (see [14]), we obtain

$$\lim_{t\to\infty} \frac{C(t)}{t} = \lim_{t\to\infty} \frac{\mathbb{E}(C(t))}{t} = \frac{\lambda(\mu - \lambda)}{\mu}\mathbb{E}\biggl(\int_0^{T_1} C(u)\,\mathrm{d}u \;\bigg|\; Z(0) = 1\biggr).$$

Also, we have the following bound:

$$\lim_{t\to\infty} \frac{C(t)}{t} = \lim_{t\to\infty} \frac{\mathbb{E}(C(t))}{t} \leq \max(c, d) \lim_{t\to\infty} \mathbb{E}(Z(t)) = \max(c, d)\frac{\lambda}{\mu - \lambda} < \infty.$$

Thus, the long-run average cost exists and is finite, and it is proportional to the total cost in the first busy cycle started with one customer in the system. Thus, the problem of finding an average cost optimal policy reduces to the problem of finding an optimal policy that minimizes the total expected cost C over a busy period T starting in state $(1, 0)$. This can be written as

$$C = \mathbb{E}\bigg(\int_0^T (cX(t) + dY(t))\,\mathrm{d}t\bigg) = d\mathbb{E}\bigg(\int_0^T Z(t)\,\mathrm{d}t\bigg) + (c-d)\mathbb{E}\bigg(\int_0^T X(t)\,\mathrm{d}t\bigg).$$

Clearly, the first term is independent of the routeing policy followed. Thus, to minimize C, we need to minimize the integral $\int X(t)\,\mathrm{d}t$ if $c > d$ and maximize it if $c < d$. If $c = d$, any policy is optimal. Clearly, when $c \le d$, it is optimal to send all traffic to the busy queue. The interesting case arises when $c > d$. Hence, we deal with this case below.

The above discussion implies that, without loss of generality, we can assume $c = 1$ and $d = 0$. Note that this is in stark contrast with the individually optimal policies that depend on both c and d. We can now formulate the cost minimization as a standard negative dynamic programming problem, see, for example, [19]. Below we make the details precise.

Let $v(i, j)$ be the minimum expected total cost starting in state $(X(0), Y(0)) = (i, j)$ over the time interval $[0, T)$, where

$$T = \min\{t \ge 0\colon Z(t) = 0 \mid Z(0) = i + j\}.$$

Without loss of generality, we can assume that $\lambda + \mu = 1$. From [19], it follows that v satisfies the optimality equations

$$\begin{aligned} v(i, j) &= i + \mu v(i-1, j) + \lambda \min(v(i+1, j), v(i, j+1)), \qquad i \ge 2,\ j \ge 0, \\ v(1, j) &= 1 + \mu v(j, 0) + \lambda \min(v(2, j), v(1, j+1)), \qquad j \ge 0, \end{aligned}$$

where $v(0, 0) = 0$. We are interested in the solution to the above equations, which can be obtained by the following value iteration for $n \ge 0$:

$$\begin{aligned} v_{n+1}(i, j) &= i + \mu v_n(i-1, j) + \lambda \min(v_n(i+1, j), v_n(i, j+1)), \qquad i \ge 2,\ j \ge 0, \\ v_{n+1}(1, j) &= 1 + \mu v_n(j, 0) + \lambda \min(v_n(2, j), v_n(1, j+1)), \qquad j \ge 0, \end{aligned}$$

with initially $v_0(i, j) = 0$ for all $i \ge 1,\ j \ge 0$, and $v_n(0, 0) = 0$ for all $n \ge 0$.

Note that the v_n in the above iteration is guaranteed to converge to v as $n \to \infty$, even though the costs are unbounded. Once v is computed, the theory of negative dynamic programming says that the optimal policy in state (i, j) is to route an incoming customer to the busy queue if $v(i+1, j) < v(i, j+1)$ and to the idle queue otherwise.

Unfortunately, we have been unable to formally derive any structural results for a socially optimal policy, the main stumbling block being the term $v_n(j, 0)$ on the right-hand side of the equation for $v_{n+1}(1, j)$. However, based on extensive numerical experimentation we see that a switching curve policy is optimal. That is, for each $i \ge 1$, there is a critical number $g(i)$ such that the optimal policy in state (i, j) is to route the incoming customer to the busy queue if $j > g(i)$ and to the idle queue otherwise. Furthermore, based on our numerical experimentation, we have the following linear approximation to the switching curve:

$$g(i) = \alpha i, \qquad i \ge 0, \tag{9}$$

where

$$\alpha = \frac{2\rho}{-1 + \rho + \sqrt{(1-\rho)(1+3\rho)}}. \tag{10}$$

It is straightforward to see that $\alpha > 1$. We also observe numerically that $h(i) \leq g(i)$ for all $i \geq 1$, where h is the switching curve for the individually optimal policy. That is, more customers join the busy queue under the individually optimal policy than under the socially optimal policy. We illustrate these comments with a numerical example in Section 7.

In the next section we develop a fluid model of this scenario and derive an optimal routeing policy.

6. Fluid model and optimal routeing policies

We see from the previous section that it is difficult to characterize the optimal policy for the fully observable system using the Markov decision process approach. Hence, in this section we study a fluid version of this problem. We first address the fluid model in the next subsection and then obtain the optimal policies for this fluid system in the following subsection.

6.1. Fluid scaling limit

We begin with a derivation of the fluid model. Consider $\{(X(t), Y(t)),\ t \geq 0\}$ as defined in the previous section, where $X(t)$ is the number of customers in the busy queue and $Y(t)$ is the number of customers in the idle queue. Suppose that we follow a stationary routeing policy that routes an incoming arrival in state (i, j) to the busy queue with probability $p(i, j) \in [0, 1]$ and to the idle queue with probability $1 - p(i, j)$ for some pre-specified function $p(\cdot, \cdot)$. Assume that $X(0) = x_0 > 0$, $Y(0) = 0$, and let $T = \min(t \geq 0\colon X(t^-) = 0\}$. That is, the servers switch from one queue to another at time T. Then we have $X(T) = Y(T^-)$ and $Y(T) = 0$, and the process repeats from time T onwards as before. Note that as long as $\lambda < \mu$ this is a stable system for all routeing functions $p(\cdot, \cdot)$.

Now create a sequence of stochastic processes $\{(X^n(t), Y^n(t)),\ t \geq 0\}$ indexed by a parameter $n = 1, 2, 3, \ldots$ with arrival and service rates

$$\lambda^n = n\lambda, \qquad \mu^n = n\mu,$$

and initial state

$$X^n(0) = nx_0, \qquad Y^n(0) = 0.$$

Furthermore, let $\alpha(x, y)$ $(x \geq 0,\ y \geq 0)$ be a function such that

$$\alpha(x, y) = \lim_{n\to\infty} p(\lfloor nx \rfloor, \lfloor ny \rfloor).$$

Now let T^n be the first time when the server switches queues in the nth system. We modify the (X^n, Y^n) process beyond T^n by assuming that

$$X^n(t) = 0, \qquad Y^n(t) = Y^n(T^{n-}) \quad \text{for } t \geq T^n.$$

Thus, the (X^n, Y^n) process does not change after T^n. Then a straightforward application of [15, Theorem 3.1] yields the following result.

Theorem 12. *Let $\tau = x_0/(\mu - \lambda)$, and $\{(x(t), y(t)),\ 0 \leq t \leq \tau\}$ be a solution to the following system of differential equations:*

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \begin{cases} \lambda\alpha(x(t), y(t)) - \mu & \text{if } x(t) > 0, \\ 0 & \text{if } x(t) = 0, \end{cases}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}y(t) = \begin{cases} \lambda(1-\alpha(x(t), y(t))) & \textit{if } x(t) > 0, \\ 0 & \textit{if } x(t) = 0, \end{cases}$$

with $(x(0), y(0)) = (x_0, 0)$*. Then, for every* $\delta > 0$*,*

$$\lim_{n\to\infty} P\left(\sup_{0\le t\le\tau} \max\left(\left|\frac{X^n(t)}{n} - x(t)\right|, \left|\frac{Y^n(t)}{n} - y(t)\right|\right) > \delta\right) = 0. \tag{11}$$

To understand the above theorem, define

$$\hat{T} = \min\{t \ge 0 \colon x(t) = 0\}.$$

Then from the above theorem, we see that as the arrival rate (λ^n) and departure rate (μ^n) in the nth system increase so that

$$\frac{\lambda^n}{n} \to \lambda, \qquad \frac{\mu^n}{n} \to \mu > \lambda, \qquad \frac{X^n(0)}{n} \to x_0,$$

then $\{(X^n(t)/n, Y^n(t)/n),\ 0 \le t < T^n\}$ converges to $\{(x(t), y(t)),\ 0 \le t < \hat{T}\}$ in the sense of (11). We say that $\{(x(t), y(t)),\ 0 \le t < \hat{T}\}$ is the fluid limit of $\{(X(t), Y(t)),\ 0 \le t < T\}$. Once the server switches to the other queue, we can use a similar fluid model over the second interval that finishes with the second switch, and so on. However, this cannot be made precise using the result of [15]; the basic problem being the jump in the $(X(t), Y(t))$ process at $t = T$.

Since $x(t)$ and $y(t)$ are deterministic functions of t (they do depend on x_0, but we ignore that to simplify the notation), we can write

$$r(t) = \lambda(1 - \alpha(x(t), y(t))), \qquad 0 \le t \le \hat{T}.$$

We can think of $r(t)$ as the rate at which fluid is being routed to the idle queue at time t and, thus, we can now talk about the optimal routeing control in the fluid system as the problem of picking the optimal routeing function $r(\cdot)$. However, to the best of the authors' knowledge, there is no theory that connects the optimal r function to the optimal p function in the original system. We establish such a connection using numerical experimentation in Section 7.

6.2. Optimal control of the fluid model

In this subsection we consider a fluid system $\{(x(t), y(t),\ t \ge 0\}$ that operates as follows. Denote by $x(t)$ the amount of fluid in the busy queue and $y(t)$ the amount of fluid in the idle queue at time t. Customers arrive as a fluid with deterministic rate λ per unit time. The rate at which the fluid is routed to the idle queue at time t is denoted by $r(t)$ and the rate at which it is routed to the busy queue is given by $\lambda - r(t)$. The cost structure remains the same. Once the server empties a queue, it switches to the other queue and continues to empty it. The fluid is removed at a deterministic rate $\mu > \lambda$ as long as there is fluid to be removed. Once the system becomes empty, the fluid is removed at rate λ and the system stays empty forever.

Suppose that the initial state is $x(0) = x_0 \ge 0$ and $y(0) = y_0 \ge 0$. Let $z(t) = x(t) + y(t)$ be the total fluid in the system at time t. Then $z(0) = x(0) + y(0)$ and regardless of the routeing policy followed, $z(t)$ decreases at rate $\lambda - \mu < 0$, until it hits 0 at time

$$T = \frac{x_0 + y_0}{\mu - \lambda},$$

and then $z(t)$ remains 0 for $t \geq T$. Note that this T is different than the T defined in the previous subsection, however, this should not cause any confusion. As in the previous section, the total cost incurred can be written as

$$\int_0^T (cx(t) + dy(t))\,\mathrm{d}t = d\int_0^T z(t)\,\mathrm{d}t + (c-d)\int_0^T x(t)\,\mathrm{d}t.$$

We want to determine the optimal routeing policy for the incoming fluid so as to minimize this cost. Since the routeing policy does not affect the trajectory of z, the optimal policy needs to minimize the second integral if $c > d$, and maximize it if $c < d$. Related work on the optimal control of service (instead of routeing) in fluid systems can be found in [16], [26], and [30].

The optimal routeing policy for $t \geq T$ is obvious: keep sending the incoming fluid to the busy queue, and both the queues will remain empty forever. Thus, we concentrate on the optimal policy for $0 \leq t \leq T$. If $c < d$, the optimal policy is to route all traffic to the busy queue. If $c = d$, all routeing policies are optimal. Hence, we further concentrate on the $c > d$ case in the rest of this section.

We assume the server has just switched to queue 1 at time 0 and the system is nonempty. Thus, $x_0 > 0$ and $y_0 = 0$, and queue 1 is the busy queue at time 0. Now let $t(0) = 0$ and t_k be the kth time at which the server switches from one queue to the other, called the kth switching time. These times are completely determined by the function $\{r(t),\ t \geq 0\}$ as follows:

$$t_1 = \min\left\{t \geq 0\colon x(t) = x_0 + \int_0^t (\lambda - r(u))\,\mathrm{d}u - \mu t = 0\right\}.$$

Thus, queue 2 becomes the busy queue at time t_1 and now has

$$x_1 = x(t_1) = \int_0^{t_1} r(u)\,\mathrm{d}u$$

amount of fluid in it. Queue 1 becomes the idle queue and has no fluid in it. Thus, we can recursively obtain, for $k \geq 1$,

$$t_{k+1} = \min\left\{t \geq t_k : x(t) = x_k + \int_{t_k}^t (\lambda - r(u))\,\mathrm{d}u - \mu(t - t_k) = 0\right\}, \tag{12}$$

$$x_{k+1} = x(t_{k+1}) = \int_{t_k}^{t_{k+1}} r(u)\,\mathrm{d}u. \tag{13}$$

We call $[t_k, t_{k+1})$ the kth cycle. Note that x_{k+1} also represents the total amount of fluid routed to the idle queue during the kth cycle. In the next theorem we state an important preliminary result on the optimal policy.

Theorem 13. *Let $\{r(t),\ t \geq 0\}$ be a given routeing policy, where $r(t)$ is the instantaneous rate at which incoming fluid is routed to the idle queue at time t. Let $\{t_k,\ k \geq 0\}$ and $\{x_k,\ k \geq 0\}$ be as given in (12) and (13). Let*

$$v_k = \frac{x_{k+1}}{\lambda}, \qquad k \geq 0,$$

and define a new routeing policy $\{s(t),\ t \geq 0\}$ (with switchover times $\{t_k,\ k \geq 0\}$) as follows:

$$s(t) = \begin{cases} \lambda & \text{for } t_k \leq t \leq t_k + v_k, \\ 0 & \text{for } t_k + v_k \leq t < t_{k+1}. \end{cases}$$

Then the total cost incurred by the routeing policy $\{s(t),\ t \geq 0\}$ *is no greater than that of policy* $\{r(t),\ t \geq 0\}$.

Proof. Let $x(t)$ and $y(t)$ be the fluid levels at time t under the routeing policy $\{r(t),\ t \geq 0\}$, and $x^s(t)$ and $y^s(t)$ be the fluid levels at time t under the routeing policy $\{s(t),\ t \geq 0\}$. First note that the amount of fluid routed to the idle queue under the $\{s(t),\ t \geq 0\}$ policy during the kth cycle is $\lambda v_k = x_{k+1}$, which is the same as under the $\{r(t),\ t \geq 0\}$ policy. However, this fluid is routed at the fastest rate possible; namely, λ. Hence,

$$y^s(t) \geq y(t), \qquad t_k \leq t < t_{k+1}, \quad k \geq 0.$$

Since $x(t) + y(t) = x^s(t) + y^s(t)$ for all t (since the total fluid content is independent of the routeing policy), it follows that

$$x^s(t) \leq x(t), \qquad t_k \leq t < t_{k+1}, \quad k \geq 0.$$

In fact, we have

$$x^s(t) = \begin{cases} x_k - \mu t & \text{for } t_k \leq t \leq t_k + v_k, \\ x_k - \mu v_k - (\mu - \lambda)(t - v_k) & \text{for } t_k + v_k \leq t < t_{k+1}. \end{cases}$$

Thus, $x^s(t) > 0$ for $t_k \leq t < t_{k+1}$, and $\{x^s(t),\ t \geq t_k\}$ reaches 0 for the first time at time t_{k+1}. Thus, the switching times under routeing policy s are the same as under policy r, and $x_k^s = x_k$ for all $k \geq 0$. Thus, we have

$$\begin{aligned} \int_0^{t_1} (cx^s(t) + dy^s(t))\,\mathrm{d}t &= \int_0^{t_1} (cx^s(t) + d(z(t) - x^s(t)))\,\mathrm{d}t \\ &= d\int_0^{t_1} z(t)\,\mathrm{d}t + (c - d)\int_0^{t_1} x^s(t)\,\mathrm{d}t \\ &\leq d\int_0^{t_1} z(t)\,\mathrm{d}t + (c - d)\int_0^{t_1} x(t)\,\mathrm{d}t \\ &= \int_0^{t_1} (cx(t) + dy(t))\,\mathrm{d}t. \end{aligned}$$

Thus, the cost under s is no greater than that under r over the first cycle. Since the state of the polling system under both policies is the same at time t_1, the above argument can be repeated to show that policy s performs at least as well as policy r over every cycle, and, hence, for all $0 \leq t \leq T$. □

The above theorem implies that a policy is equally well characterized by the switching times $\{t_k,\ k \geq 0\}$ it induces (with $t_0 = 0$) and, among all the policies with these switching times, a policy that sends all the traffic to the idle queue first as long as possible in each cycle is optimal. Thus, all that remains to be done is to identify the optimal switching times.

Since the system is deterministic, determining optimal $\{t_k,\ k \geq 0\}$ is equivalent to determining the optimal fluid levels $\{x_k,\ k \geq 1\}$ with a given initial level x_0. In the next theorem we show that this can be modeled and solved as a linear quadratic regulator (LQR) problem; see [2] and [8].

Theorem 14. *The optimal* $\{x_k,\ k \geq 1\}$ *are obtained by solving the following infinite horizon constrained LQR:*

$$\underset{u_k \in \mathbb{R}}{\text{minimize}} \quad \sum_{k=0}^{\infty} (1-\rho)x_k^2 + \rho u_k^2, \tag{14a}$$

$$\textit{subject to} \quad x_{k+1} = \rho(x_k - u_k), \tag{14b}$$

$$0 \leq u_k \leq x_k. \tag{14c}$$

Proof. Let $\{t_k,\ k \geq 0\}$ be the switching times of the optimal policy. From Theorem 13 we see that there exist $\{v_k,\ k \geq 0\}$ such that over the kth cycle $[t_k, t_{k+1})$ it is optimal to route all fluid to the idle queue over $[t_k, t_k + v_k)$ and then route all fluid to the busy queue over $[t_k + v_k, t_{k+1})$.

Thus, during the interval $[t_k, t_k + v_k)$, the fluid level in the busy queue decreases at rate μ from x_k to $x_k - \mu v_k$, after which it decreases at rate $\mu - \lambda$ from $x_k - \mu v_k$ to 0. For the idle queue, the fluid level increases at rate λ from 0 to λv_k over $[t_k, t_k + v_k)$ and then stays constant over $[t_k + v_k, t_{k+1})$. It is straightforward to show that

$$t_{k+1} - v_k - t_k = \frac{x_k - \mu v_k}{\mu - \lambda}.$$

Therefore, the cost during the kth cycle is

$$c\left(\frac{2x_k - \mu v_k}{2} v_k + \frac{(x_k - \mu v_k)^2}{2(\mu - \lambda)}\right) + d\left(\frac{\lambda v_k^2}{2} + \frac{\lambda v_k (x_k - \mu v_k)}{\mu - \lambda}\right). \tag{15}$$

Furthermore, we obtain the following dynamics:

$$x_{k+1} = \lambda v_k$$

subject to the constraint $0 \leq \mu v_k \leq x_k$. Note that using the dynamics, we have

$$\sum_{k=0}^{\infty} (x_k^2 - \lambda^2 v_k^2) = \sum_{k=0}^{\infty} (x_k^2 - x_{k+1}^2) = x_0^2. \tag{16}$$

So subtracting a multiple $d/2(\mu - \lambda)$ of the left-hand side of (16) (which is like subtracting a constant) from the sum of (15) over all k, we need to solve the following problem:

$$\underset{v_k \in \mathbb{R}}{\text{minimize}} \quad \frac{c-d}{2(\mu - \lambda)} \sum_{k=0}^{\infty} (x_k^2 - 2\lambda x_k v_k + \lambda \mu v_k^2),$$

$$\text{subject to} \quad x_{k+1} = \lambda v_k, \qquad 0 \leq \mu v_k \leq x_k.$$

To eliminate the product $\lambda x_k v_k$, we define the new variable

$$u_k = x_k - \mu v_k,$$

which allows us to write our problem as (14). □

In the next theorem we present the solution to the above problem.

Theorem 15. *The optimal solution to the constrained LQR in Theorem 14 is*

$$u_k = \frac{-(1-\rho) + \sqrt{(1-\rho)(1+3\rho)}}{1+\rho+\sqrt{(1-\rho)(1+3\rho)}} x_k, \qquad k \geq 0, \tag{17}$$

and

$$x_{k+1} = \beta x_k, \qquad k \geq 0, \tag{18}$$

where

$$\beta = \frac{2\rho}{1+\rho+\sqrt{(1-\rho)(1+3\rho)}}.$$

Proof. Without constraint (14c), problem (14) is the standard (infinite-horizon discrete-time) LQR problem. Initially, we ignore constraint (14c). Then the solution to the optimal control problem (14) is

$$u_k = -f x_k,$$

where

$$f = (r + bpb)^{-1} bpa$$

and p is the (unique) nonnegative solution of the discrete time algebraic Ricatti equation, that is,

$$p = q + a(p - pb(r + bpb)^{-1} bp)a,$$

where

$$a = \rho, \qquad b = -\rho, \qquad q = (1-\rho), \qquad r = \rho.$$

That is, the optimal solution is

$$u_k = \frac{-(1-\rho) + \sqrt{(1-\rho)^2 + 4\rho(1-\rho)}}{1+\rho+\sqrt{(1-\rho)(1+3\rho)}} x_k,$$

which can be written as (17). Recall that we ignored constraint (14c). However, solution (17) satisfies (14c), so (17) is also the optimal solution for problem (14) including constraint (14c). Using (17) in $x_{k+1} = \lambda v_k$ and $u_k = x_k - \mu v_k$, leads to (18). □

One can show that $0 \leq \beta < 1$. Thus, the amount of fluid at the switchover times decreases geometrically to 0. The optimal policy goes through an infinite number of switchovers before the system becomes empty. In the next theorem we specify the optimal policy implied by the above theorem.

Theorem 16. *Let α be as in (10). It is optimal to route all incoming fluid to the idle queue at time t if*

$$y(t) < \alpha x(t) \tag{19}$$

and all incoming fluid to the busy queue otherwise.

Proof. Let $y(0) = y_0$. We consider three cases.

Case 1: $y_0 = 0$. This was assumed in the proof of the previous theorems in this subsection. Note that at time v_k, the fluid level of the busy queue reduces to

$$q_1 = \frac{-(1-\rho) + \sqrt{(1-\rho)(1+3\rho)}}{1+\rho+\sqrt{(1-\rho)(1+3\rho)}} x_k$$

and that of the idle queue increases from 0 to

$$q_2 = \rho\left(x_k - \frac{-(1-\rho) + \sqrt{(1-\rho)(1+3\rho)}}{1+\rho+\sqrt{(1-\rho)(1+3\rho)}}x_k\right) = \frac{2\rho}{1+\rho+\sqrt{(1-\rho)(1+3\rho)}}x_k.$$

Thus, (19) is satisfied for $t \in [t_k + v_k, t_{k+1})$, and not satisfied for $t \in [t_k, t_k + v_k)$. At $t = t_k + v_k$ it is satisfied at equality.

Case 2: $0 < y_0 \le \alpha x_0$. Consider a system starting at time $\tau = -y(0)/\lambda$, in state $x(\tau) = x_0 + \mu\tau$, and $y(\tau) = 0$. Then following the policy dictated by the switching curve in (19) from time $t \ge \tau$ will bring the system to state $x(0) = x_0$ and $y(0) = y_0$. Hence, the same optimal policy will continue to hold for $t \ge 0$ from the principle of optimality.

Case 3: $0 < \alpha x_0 < y(0) = y_0$. In this case, define

$$\tau_1 = \frac{y(0) - x_0}{\alpha(\mu - \lambda)}, \qquad \tau_2 = \frac{y(0)}{\lambda}.$$

Now consider a system starting at time $\tau = -(\tau_1 + \tau_2)$, in state $x(\tau) = x_0 + \alpha(\mu-\lambda)\tau_1 + \mu\tau_2$, and $y(\tau) = 0$. Then following the optimal policy dictated by the switching curve in (19) from time $t \ge \tau$ will bring the system to state $x(0) = x_0$ and $y(0) = y_0$. Hence, the same optimal policy will continue to hold for $t \ge 0$ from the principle of optimality. □

Note that the switching curve in (19) matches the numerically observed curve in (9).

7. Numerical example and conclusions

In this section we present a numerical example with the following parameters:

$$\lambda = 0.3, \qquad \mu = 0.7, \qquad c = 6, \qquad d = 1.$$

In Figure 2 we present a plot of the three switching curves . The lower curve corresponds to the switching curve h of the individually optimal policy. It is optimal to join the busy queue in all states (i, j) that lie above this curve. We have numerically observed that as d approaches 0, the switching curve moves up and it reaches $h(i) = i$ when $d = 0$, that is, the individually optimal policy is to join the shortest queue when $d = 0$. On the other hand, as d increases, the switching curve moves down and reaches $h(i) = 0$ when $d \ge c$, that is, the optimal policy is to always join the busy queue.

The upper curve in Figure 2 corresponds to the switching curve g of the socially optimal policy. As we discussed before, this is independent of c and d as long as $c > d$. The middle curve corresponds to the fluid switching curve αi, where α is as in (10). The fluid curve is also independent of c and d as long as $c > d$. It is interesting to see that the two curves are quite close. For both policies, it is optimal to join the busy queue in all states (i, j) that lie above the curve.

Observe that both the socially optimal and the fluid switching curves are above the curve $j = i$, while the individually optimal curve is below it. This observation for the individually optimal policies was proved in Theorem 8. It follows for the fluid policy, since we know that $\alpha > 1$. We have not been able to prove it for the socially optimal policy.

In Figure 2 we also see that more customers join the busy queue under the individually optimal policy than under the socially optimal policy. This is consistent with the general observation in other queueing systems, and it is a result of externalities: in individually optimal

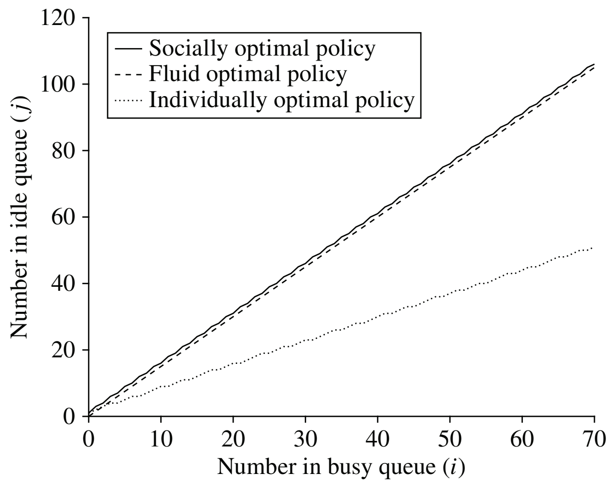

FIGURE 2: The switching curves for the three policies.

policies, customers are selfish and ignore the cost their decision imposes on other customers and, hence, tend to over-utilize the resources.

It would be interesting to formally prove these observations, but we leave that as future work. In this paper we considered a simple exhaustive polling system with two queues, identical exponential service times, and no switchover times and switchover costs. We considered three levels of observability: unobservable, almost unobservable, and fully observable, and derived individually and socially optimal policies. The almost unobservable case assumes that we know where the server is, but not the queue lengths. One can also consider the case where the queue lengths are known, but not where the server is (this could be called the almost observable case). Although it is possible to derive this expression, it becomes rather involved (since we need to know where the server is, given the queue lengths), and, hence, we have not considered this case here. We expect that the analysis will be necessarily of a numerical nature.

Clearly, several other extensions are possible: to more than two queues, nonidentical exponential service times, general service times, service policies other than exhaustive service, nonzero switchover times or costs, and so on. Each of these extensions makes the analysis more difficult, since the expressions for the expected queue lengths become more involved.

Acknowledgements

The authors would like to thank the anonymous reviewers for their insightful comments that helped to improve the paper.

References

[1] ALTMAN, E. AND SHIMKIN, N. (1998). Individual equilibrium and learning in processor sharing systems. *Operat. Res.* **46,** 776–784.

[2] BERTSEKAS, D. P. (2005). *Dynamic Programming and Optimal Control*, Vol. 1, 3rd edn. Athena Scientific, Belmont, MA.

[3] BOON, M. A. A. (2011). Polling models: from theory to traffic intersections. Doctoral Thesis, Eindhoven University of Technology.

[4] BOON, M. A. A., VAN DER MEI, R. D. AND WINANDS, E. M. M. (2011). Applications of polling systems. *Surveys Operat. Res. Manag. Sci.* **16,** 67–82.

[5] Boon, M. A. A., van der Mei, R. D. and Winands, E. M. M. (2013). Waiting times in queueing networks with a single shared server. *Queueing Systems* **74,** 403–429.

[6] Boon, M. A. A., van Wijk, A. C. C., Adan, I. J. B. F. and Boxma, O. J. (2010). A polling model with smart customers. *Queueing Systems* **66,** 239–274.

[7] Boxma, O. J., Levy, H. and Weststrate, J. A. (1991). Efficient visit frequencies for polling tables: minimization of waiting cost. *Queueing Systems* **9,** 133–162.

[8] Bryson, A. E., Jr. and Ho, Y. C. (1975). *Applied Optimal Control: Optimization, Estimation, and Control*. Hemisphere, Washington, DC.

[9] Burnetas, A. and Economou, A. (2007). Equilibrium customer strategies in a single server Markovian queue with setup times. *Queueing Systems* **56,** 213–228.

[10] Haijema, R. and van der Wal, J. (2008). An MDP decomposition approach for traffic control at isolated signalized intersections. *Prob. Eng. Inf. Sci.* **22,** 587–602.

[11] Hassin, R. and Haviv, M. (2003). *To Queue or not to Queue: Equilibrium Behavior in Queueing Systems*. Kluwer, Boston, MA.

[12] Klimov, G. P. (1975). Time-sharing service systems. I. *Theory Prob. Appl.* **19,** 532–551.

[13] Klimov, G. P. (1979). Time-sharing service systems. II. *Theory Prob. Appl.* **23,** 314–321.

[14] Kulkarni, V. G. (2016). *Modeling and Analyisis of Stochastic Systems*, 3rd edn. CRC, Boca Raton, FL.

[15] Kurtz, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* **7,** 49–58.

[16] Lefeber, E., Lämmer, S. and Rooda, J. E. (2011). Optimal control of a deterministic multiclass queuing system for which several queues can be served simultaneously. *Systems Control Lett.* **60,** 524–529.

[17] Levy, H. and Sidi, M. (1990). Polling systems: applications, modeling, and optimization. *IEEE Trans. Commun.* **38,** 1750–1760.

[18] Lippman, S. A. and Stidham, S., Jr. (1977). Individual versus social optimization in exponential congestion systems. *Operat. Res.* **25,** 233–247.

[19] Ross, S. (1983). *Introduction to Stochastic Dynamic Programming*. Academic Press, New York.

[20] Sharafali, M., Co, H. C. and Goh, M. (2004). Production scheduling in a flexible manufacturing system under random demand. *Europ. J. Operat. Res.* **158,** 89–102.

[21] Sidi, M., Levy, H. and Fuhrmann, S. W. (1992). A queueing network with a single cyclically roving server. *Queueing Systems* **11,** 121–144.

[22] Takagi, H. (1986). *Analysis of Polling Systems*. MIT Press, Cambridge, MA.

[23] Takine, T., Takagi, H. and Hasegawa, T. (1991). Sojourn times in vacation and polling systems with Bernoulli feedback. *J. Appl. Prob.* **28,** 422–432.

[24] Van der Wal, J. and Yechiali, U. (2003). Dynamic visit-order rules for batch-service polling. *Prob. Eng. Inf. Sci.* **17,** 351–367.

[25] Vishnevskiĭ, V. M. and Semenova, O. V. (2006). Mathematical methods for investigating polling systems. *Autom. Remote Control* **67,** 173–220.

[26] Weiss, G. (1999). Scheduling and control of manufacturing systems—a fluid approach. In *Proceedings of the 37th Allerton Conference*, pp. 577–586.

[27] Winands, E. M. M. (2007). *Polling, production & priorities*. Doctoral Thesis, Eindhoven University of Technology.

[28] Winands, E. M. M., Adan, I. J. B. F. and van Houtum, G. J. (2006). Mean value analysis for polling systems. *Queueing Systems* **54,** 35–44.

[29] Yechiali, U. (1993). Analysis and control of polling systems. In *Performance Evaluation of Computer and Communication Systems*, Springer, Berlin, pp. 630–650.

[30] Van Zwieten, D. A. J., Lefeber, E. and Adan, I. J. B. F. (2016). Optimal steady-state and transient trajectories of multi-queue switching servers with a fixed service order of queues. *Performance Evaluation* **97,** 16–35.