

# Two-Queue Polling Models with a Patient Server

O.J. Boxma

*Department of Mathematics and Computer Science, Eindhoven University of  
Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

S. Schlegel

*EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

U. Yechiali

*Department of Statistics & Operations Research, School of Mathematical Sciences,  
Tel Aviv University, Tel Aviv 69978, Israel*

**Abstract.** We consider two-queue polling models with the special feature that a timer mechanism is employed at  $Q_1$ : whenever the server polls  $Q_1$  and finds it empty, it activates a timer and remains dormant, waiting for the first arrival. If such an arrival occurs before the timer expires, a busy period starts in accordance with  $Q_1$ 's service discipline. However, if the timer is shorter than the interarrival time to  $Q_1$ , the server does not wait any more and switches back to  $Q_2$ . We consider three configurations: (i)  $Q_1$  is controlled by the 1-limited protocol while  $Q_2$  is served exhaustively. (ii)  $Q_1$  employs the exhaustive regime while  $Q_2$  follows the 1-limited procedure. (iii) Both queues are served exhaustively. In all cases, we assume Poisson arrivals and allow general service and switchover time distributions. Our main results include the queue length distributions at polling instants, the waiting time distributions and the distribution of the total workload in the system.

**Keywords:** Two queues, alternating service, polling, 1-limited, exhaustive, timer, patient server.

## 1. Introduction

A single server attends two queues, denoted  $Q_1$  and  $Q_2$ , by alternating its service among them. The service discipline in each queue is either exhaustive or 1-limited. In the exhaustive regime, the server keeps serving a queue until it is empty. For the 1-limited policy, at most one customer is served. However, regardless of its specific regime,  $Q_1$  exercises an extra priority over  $Q_2$  by virtue of a timer mechanism, operating as follows. Whenever the server polls  $Q_1$  and finds it empty, it activates a timer and remains dormant, waiting for the first arrival. If such an arrival occurs before the timer expires, a busy period starts in accordance with  $Q_1$ 's service discipline. However, if the timer is shorter than the interarrival time to  $Q_1$ , the server does not wait any more and switches back to  $Q_2$ . This 'wait and see' policy is common in human behaviour and is employed in many real-life operations (road



© 2002 Kluwer Academic Publishers. Printed in the Netherlands.

traffic intersections; a machine that can process several classes of jobs, requiring change-over times between classes for tool switching; etc.). In spite of its importance, it has been studied only recently [9] in the context of a single queue with vacations. A variant of this 'wait and see' policy is studied in Peköz [19], where during a visit to a queue and after the queue becomes empty, the server always stays idle there for a deterministic amount of time.

In this work we extend the analysis of [9] to two-queue polling models in which the server exercises the wait option in  $Q_1$ . We consider three configurations: (i)  $Q_1$  is controlled by the 1-limited protocol while  $Q_2$  is served exhaustively. (ii)  $Q_1$  employs the exhaustive regime while  $Q_2$  follows the 1-limited procedure. (iii) Both queues are served exhaustively. In all cases, we assume that customers arrive at the queues according to independent Poisson processes, with service requests that are independent and follow general distributions. We consider both zero and nonzero switchover times; in the latter case, their distributions are general. Our main results include the queue length distributions at polling instants, the waiting time distributions and the distribution of the total workload in the system.

Let us briefly review the relevant literature; for extensive surveys on polling systems the reader is referred to Takagi [22, 23] and Yechiali [24]. Two-queue alternating-service systems without timers have been treated by many authors in the literature, under various assumptions on their operating schemes. Avi-Itzhak, Maxwell and Miller [1] were the first to study such a configuration, assuming the exhaustive service discipline in each queue and zero switchover times. They derived the mean queue size and expected waiting time, as well as the first two moments of the busy period, in each queue. Takács [21] studied the same model, obtaining Laplace-Stieltjes transforms (LST) and probability generating functions (PGF) of key variables. Eisenberg [12] investigated the same model but with nonzero switchover times.

The two-queue polling model with exhaustive service at one queue and 1-limited service at the other queue has been analysed in detail by Groenendijk [16] and Ibe [17]. The two-queue polling model with 1-limited service at both queues is intrinsically more difficult than those with exhaustive service at both queues or those with exhaustive service at one queue and 1-limited at the other. The joint queue length distribution at both 1-limited queues can be obtained via a translation to a boundary value problem (see e.g. Boxma and Groenendijk [6]), but extension of the results to more than two queues seems out of reach.

Instead of timers, additional priorities can also be implemented using thresholds. Threshold service disciplines, where  $Q_1$  is served exhaustively while  $Q_2$  is served only until either the work there is completed or

the queue size in the other ('primary') queue reaches a given threshold, were studied by Lee [18], Boxma, Koole and Mitrani [7, 8] and Boxma and Down [4]. In [7] the service times are exponentially distributed and services at  $Q_2$  are preemptively interrupted when the threshold at  $Q_1$  is reached, while in [8] the service process at  $Q_2$  is nonpreemptively interrupted when the threshold at  $Q_1$  is reached. [4] extends the analysis in [8] to the case where service times are generally distributed, and treats both cases of zero and nonzero switchover times. Exact expressions for the joint queue-length distributions at customer departure epochs and for the steady-state queue length and sojourn time distributions are derived. Lee [18] deals with a similar model and gives light and heavy traffic analyses.

Eliazar and Yechiali [14] recently studied a communication multiplexer problem, analyzing it as two alternating queues with dependent randomly-timed gated regime [13]. The primary queue is served exhaustively, whereas the duration of time the server resides in the secondary queue is determined by the dynamic evolution in  $Q_1$ . They derived numerous performance measures, each expressed as a function of an undetermined PGF of the number of messages at polling instants of  $Q_2$ , and obtained explicit approximated values for all performance measures that depend on the above PGF.

The paper is organized as follows. Section 2 contains a detailed model description. In Section 3 we study queue lengths and derive multi-dimensional PGFs of the system's state at polling instants. In Section 4 we calculate the LST of the workload in the system, derive decomposition results and obtain expressions for pseudoconservation laws, from which mean waiting times are determined. Waiting time distributions are considered in Section 5. Various possible extensions are mentioned in Section 6.

## 2. Model Description and Notation

We consider a polling system consisting of two queues  $Q_1$  and  $Q_2$  with infinite buffer capacity each, attended by a single server that alternates between the queues. Customers arrive at  $Q_i$ ,  $i = 1, 2$ , according to a Poisson process  $\{A_i(t), t \geq 0\}$  with intensity  $\lambda_i$ , and require a service time  $B_i$  with distribution  $B_i(\cdot)$ , mean  $\beta_i$ , second moment  $\beta_i^{(2)}$ , and LST  $\tilde{B}_i(\cdot)$ . Successive i.i.d. service times are denoted by  $B_{ik}$ ,  $k = 1, 2, \dots$ ,  $i = 1, 2$ . A similar notation is used for other random variables to be introduced below. Let  $\lambda = \lambda_1 + \lambda_2$  denote the total arrival rate,  $\rho_i = \lambda_i \beta_i$  the traffic intensity at  $Q_i$ , and  $\rho = \rho_1 + \rho_2$  the total traffic intensity. By  $B(\cdot)$  we denote the service time distribution of an arbitrary (arriving)

customer:

$$B(t) = \frac{\lambda_1}{\lambda_1 + \lambda_2} B_1(t) + \frac{\lambda_2}{\lambda_1 + \lambda_2} B_2(t).$$

Customers at  $Q_1$  have some priority in the sense that on finding  $Q_1$  empty the server waits there for a pre-specified duration of time  $T$  (which may be random or constant and which we call a timer), hoping for an arrival during that time. If the timer expires before an arrival occurs, the server switches to  $Q_2$ . We consider both zero and nonzero switchover times. In the latter case, switching from  $Q_i$  to the other queue,  $i = 1, 2$ , takes a random duration  $D_i$  with distribution  $D_i(\cdot)$ , mean  $d_i$  and LST  $\tilde{D}_i(\cdot)$ ;  $d = d_1 + d_2$ . We deal with two service disciplines at the two queues, namely, the *exhaustive* and *1-limited* regimes.

We restrict ourselves to the stationary situation. The stability condition depends on the chosen service disciplines. We will discuss them at the appropriate places; for extensive discussions of stability conditions in polling systems see Borovkov [2] and Fricker and Jaibi [15].

Let  $X_i^j$  be the number of customers at  $Q_j$  when  $Q_i$  is polled (i.e., is visited by the server), with joint PGF  $F_i(z_1, z_2) = \mathbb{E}[z_1^{X_i^1} z_2^{X_i^2}]$ . Let  $IA_1$  be the interarrival time at  $Q_1$ ,  $M_1 = \min\{IA_1, T\}$  with LST  $\tilde{M}_1(\cdot)$  and mean  $\mathbb{E}M_1 = a_1/\lambda_1$ , where  $a_1 = \mathbb{P}(IA_1 \leq T) = 1 - \tilde{T}(\lambda_1)$ , where  $\tilde{T}(\cdot)$  denotes the LST of  $T$ . Moreover,

$$\begin{aligned} B_i(z_1, z_2) &= \mathbb{E} \left[ z_1^{A_1(B_i)} z_2^{A_2(B_i)} \right] = \tilde{B}_i(\lambda_1(1 - z_1) + \lambda_2(1 - z_2)); \\ D_i(z_1, z_2) &= \mathbb{E} \left[ z_1^{A_1(D_i)} z_2^{A_2(D_i)} \right] = \tilde{D}_i(\lambda_1(1 - z_1) + \lambda_2(1 - z_2)); \\ f_i(z) &= \tilde{\theta}_i(\lambda_j(1 - z)), \quad i, j = 1, 2, i \neq j, z \geq 0, \end{aligned}$$

where  $\tilde{\theta}_i(\cdot)$  is the LST of a generic busy period  $\theta_i$  at  $Q_i$ ,  $i = 1, 2$ , with mean  $\mathbb{E}\theta_i = \beta_i/(1 - \rho_i)$  and  $\mathbb{E}(\theta_i^2) = \beta_i^{(2)}/(1 - \rho_i)^3$ . We only deal with busy periods at queues with exhaustive service. Note that  $B_i(\cdot, \cdot)$  ( $D_i(\cdot, \cdot)$ ) is the joint PGF of the number of arrivals at each of the queues during a service (switchover time) at (from)  $Q_i$ . Further,  $f_1(z)$  is the PGF of the number of arrivals at  $Q_2$  during one generic busy period at  $Q_1$ , and similar for  $f_2(z)$ :

$$f_1(z) = \mathbb{E}[z^{A_2(\theta_1)}], \quad f_2(z) = \mathbb{E}[z^{A_1(\theta_2)}]. \quad (1)$$

Finally, define the cycle time  $C$  as the time between two successive polling instants by the server of  $Q_1$ . By an easy balance argument, the mean cycle time is

$$\mathbb{E}C = (d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0))/(1 - \rho). \quad (2)$$

### 3. Queue Lengths

In this section we construct the evolution equations for the queue lengths at polling instants for various combinations of service disciplines at the two queues. We consider the following combinations: (i)  $Q_1$  follows the 1-limited rule while  $Q_2$  is controlled by the exhaustive regime. (ii)  $Q_1$  is served exhaustively while  $Q_2$  employs the 1-limited policy. (iii) Both queues operate under the exhaustive regime. Recall that the timer is initiated only if  $Q_1$  is empty at a polling instant. Based on the evolution equations we derive the PGF's of the queue lengths for each combination.

#### 3.1. $Q_1$ : 1-LIMITED; $Q_2$ : EXHAUSTIVE

In this model, though customers at  $Q_1$  have some priority reflected by the timer at  $Q_1$ , there is a trade-off for this preference by serving at most one customer during the course of a server's visit to  $Q_1$ .

The stability condition in this case must be the same as in the 1-limited/exhaustive polling model without a timer [16, 17], namely  $\rho + \lambda_1 d < 1$ . We refrain from a proof (for proof techniques, see [2, 15]). An intuitive argument is the following. Since  $Q_1$  can serve at most one customer per cycle, the bottleneck is at  $Q_1$ , and the stability condition is  $\lambda_1 \mathbb{E}C < 1$ , where  $\mathbb{E}C$  is given by (2). However, given that  $Q_1$  is in heavy traffic, the server never finds  $Q_1$  empty, and thus we get indeed  $\rho + \lambda_1 d < 1$ .

We derive the joint PGFs of the queue lengths at polling instants; cf. (9), (10). From these we obtain mean queue lengths at polling instants; cf. (13), (14).

The queue length  $X_1^1$  is given by the queue length at  $Q_1$  when  $Q_2$  was polled for the last time, i.e.  $X_2^1$ , plus the arrivals to  $Q_1$  during the server's stay at  $Q_2$  plus the arrivals to  $Q_1$  during the switching from  $Q_2$  to  $Q_1$ . Note that serving  $Q_2$  exhaustively means that if at the beginning of service at  $Q_2$  the number of jobs is  $X_2^2 > 0$ , the server stays there  $X_2^2$  regular busy periods of an  $M/G/1$  queue having Poisson arrival rate  $\lambda_2$  and service requirements  $B_2$ . Arguing in the same way, we get the evolution equations for the queue lengths at polling instants:

$$\begin{aligned} X_1^1 &= X_2^1 + A_1 \left( \sum_{k=1}^{X_2^2} \theta_{2k} \right) + A_1(D_2), & X_1^2 &= A_2(D_2), \\ X_2^1 &= \begin{cases} X_1^1 - 1 + A_1(B_1) + A_1(D_1), & \text{if } X_1^1 > 0, \\ A_1(B_1) \mathbb{1}(IA_1 \leq T) + A_1(D_1), & \text{if } X_1^1 = 0, \end{cases} \\ X_2^2 &= \begin{cases} X_1^2 + A_2(B_1) + A_2(D_1), & \text{if } X_1^2 > 0, \\ X_1^2 + A_2(M_1) + A_2(B_1) \mathbb{1}(IA_1 \leq T) + A_2(D_1), & \text{if } X_1^2 = 0, \end{cases} \end{aligned}$$

where  $\mathbb{I}(A)$  is the indicator function of the event  $A$ . Note that  $M_1$  and  $\mathbb{I}(IA_1 \leq T)$  are dependent. From this we obtain the generating functions

$$\begin{aligned} F_1(z_1, z_2) &= \mathbb{E}\left[z_1^{X_1^1} z_2^{X_1^2}\right] = \mathbb{E}\left[z_1^{X_2^1+A_1(\sum_{k=1}^{X_2^2} \theta_{2k})\mathbb{I}(X_2^2>0)+A_1(D_2)} z_2^{A_2(D_2)}\right] \\ &= D_2(z_1, z_2) \left( \mathbb{E}\left[z_1^{X_2^1+A_1(\sum_{k=1}^{X_2^2} \theta_{2k})} \mathbb{I}(X_2^2 > 0)\right] + \mathbb{E}\left[z_1^{X_2^1} \mathbb{I}(X_2^2 = 0)\right] \right) \\ &= D_2(z_1, z_2) F_2(z_1, f_2(z_1)), \end{aligned} \quad (3)$$

and

$$\begin{aligned} F_2(z_1, z_2) &= \mathbb{E}\left[z_1^{(X_1^1-1+A_1(B_1))\mathbb{I}(X_1^1>0)+A_1(B_1)\mathbb{I}(IA_1\leq T)\mathbb{I}(X_1^1=0)+A_1(D_1)}\right. \\ &\quad \left. \times z_2^{X_1^2+A_2(B_1)\mathbb{I}(X_1^1>0)+(A_2(M_1)+A_2(B_1)\mathbb{I}(IA_1\leq T))\mathbb{I}(X_1^1=0)+A_2(D_1)}\right] \\ &= D_1(z_1, z_2) \left( B_1(z_1, z_2) \frac{1}{z_1} \mathbb{E}\left[z_1^{X_1^1} z_2^{X_1^2} \mathbb{I}(X_1^1 > 0)\right] \right. \\ &\quad \left. + \mathbb{E}\left[z_1^{A_1(B_1)\mathbb{I}(IA_1\leq T)} z_2^{X_1^2+A_2(M_1)+A_2(B_1)\mathbb{I}(IA_1\leq T)} \mathbb{I}(X_1^1 = 0)\right] \right) \\ &= D_1(z_1, z_2) \left[ B_1(z_1, z_2) \frac{F_1(z_1, z_2) - F_1(0, z_2)}{z_1} + F_1(0, z_2) r(z_1, z_2) \right], \end{aligned} \quad (4)$$

where

$$r(z_1, z_2) = \mathbb{E}\left[z_1^{A_1(B_1)\mathbb{I}(IA_1\leq T)} z_2^{A_2(M_1)+A_2(B_1)\mathbb{I}(IA_1\leq T)}\right]$$

is a known function that can be specified explicitly for given distributions of  $B_1$  and  $T$ . With

$$c_1 = F_2(0, \tilde{\theta}_2(\lambda_1)) \quad (5)$$

we have from (3)

$$F_1(0, z_2) = c_1 D_2(0, z_2), \quad (6)$$

and thus substituting (3) into (4) yields

$$\begin{aligned} F_2(z_1, z_2) &= F_2(z_1, f_2(z_1)) \frac{D_1(z_1, z_2) D_2(z_1, z_2) B_1(z_1, z_2)}{z_1} \\ &\quad + c_1 D_2(0, z_2) D_1(z_1, z_2) \left( r(z_1, z_2) - \frac{B_1(z_1, z_2)}{z_1} \right). \end{aligned} \quad (7)$$

Putting  $z_2 = f_2(z_1)$  in (7) and solving for  $F_2(z_1, f_2(z_1))$  gives

$$\begin{aligned} F_2(z_1, f_2(z_1)) &= \\ &= \frac{c_1 D_2(0, f_2(z_1)) D_1(z_1, f_2(z_1)) \left( z_1 r(z_1, f_2(z_1)) - B_1(z_1, f_2(z_1)) \right)}{z_1 - D_1(z_1, f_2(z_1)) D_2(z_1, f_2(z_1)) B_1(z_1, f_2(z_1))}. \end{aligned} \quad (8)$$

Therefore, by plugging (8) into (3) and (7), respectively, we finally get

$$F_1(z_1, z_2) = c_1 D_1(z_1, f_2(z_1)) D_2(z_1, z_2) D_2(0, f_2(z_1)) \\ \times \frac{z_1 r(z_1, f_2(z_1)) - B_1(z_1, f_2(z_1))}{z_1 - D_1(z_1, f_2(z_1)) D_2(z_1, f_2(z_1)) B_1(z_1, f_2(z_1))} \quad (9)$$

and

$$F_2(z_1, z_2) = \\ = \frac{c_1 D_2(0, f_2(z_1)) D_1(z_1, f_2(z_1)) (z_1 r(z_1, f_2(z_1)) - B_1(z_1, f_2(z_1)))}{z_1 - D_1(z_1, f_2(z_1)) D_2(z_1, f_2(z_1)) B_1(z_1, f_2(z_1))} \\ \times \frac{D_1(z_1, z_2) D_2(z_1, z_2) B_1(z_1, z_2)}{z_1} \\ + c_1 D_2(0, z_2) D_1(z_1, z_2) \left( r(z_1, z_2) - \frac{B_1(z_1, z_2)}{z_1} \right). \quad (10)$$

To determine the constant  $c_1$ , we put  $z_1 = z_2 = 1$  in (9) to obtain

$$1 = F_1(1, 1) \\ = c_1 D_2(0, 1) \lim_{z_1 \rightarrow 1} \frac{z_1 r(z_1, f_2(z_1)) - B_1(z_1, f_2(z_1))}{z_1 - D_1(z_1, f_2(z_1)) D_2(z_1, f_2(z_1)) B_1(z_1, f_2(z_1))} \\ = c_1 \tilde{D}_2(\lambda_1) \frac{1 + \lambda_1(\lambda_2 \mathbb{E}\theta_2 \mathbb{E}M_1 - (1 - a_1)(\lambda_2 \mathbb{E}\theta_2 + 1)\beta_1)}{1 - \lambda_1(\lambda_2 \mathbb{E}\theta_2 + 1)(d_1 + d_2 + \beta_1)} \\ = c_1 \tilde{D}_2(\lambda_1) \frac{1 - \rho + \lambda_1 \rho_2 \mathbb{E}M_1 + a_1 \rho_1}{1 - \rho - \lambda_1 d} = c_1 \tilde{D}_2(\lambda_1) \frac{1 - \rho + \lambda_1 \rho \mathbb{E}M_1}{1 - \rho - \lambda_1 d},$$

where we employed l'Hospital's rule for the lim operation, used

$$\left. \frac{d}{dz_1} D_i(z_1, f_2(z_1)) \right|_{z_1=1} = \lambda_1 d_i / (1 - \rho_2), \\ \left. \frac{d}{dz_1} B_1(z_1, f_2(z_1)) \right|_{z_1=1} = \lambda_1 \beta_1 / (1 - \rho_2)$$

and  $a_1 = \lambda_1 \mathbb{E}M_1$ . Finally,

$$c_1 = \frac{1 - \rho - \lambda_1 d}{\tilde{D}_2(\lambda_1) (1 - \rho + \lambda_1 \rho \mathbb{E}M_1)}. \quad (11)$$

**Remark 3.1** Notice that  $c_1 > 0$  since the stability condition  $\rho + \lambda_1 d < 1$  holds. A more direct approach to determine  $c_1$ , which exploits the 1-limited protocol at  $Q_1$ , is the following. Since in steady-state the mean number of arrivals per cycle at one of the queues equals the mean number of services there, we have

$$\lambda_1 \mathbb{E}C = 0 \cdot \mathbb{P}(\text{no service at } Q_1) + 1 \cdot \mathbb{P}(\text{there is a service at } Q_1) \\ = \mathbb{P}(X_1^1 > 0) + \mathbb{P}(X_1^1 = 0, IA_1 \leq T) = 1 - \mathbb{P}(X_1^1 = 0)(1 - a_1).$$

Now, by substituting the value of  $\mathbb{E}C$  from (2) in the left-hand-side of the above, solving for  $\mathbb{P}(X_1^1 = 0)$ , using (6) with  $z_2 = 1$ , and recalling that  $D_2(0, 1) = \tilde{D}_2(\lambda_1)$ , we get (11). Notice also that  $F_1(0, 1) = \mathbb{P}(X_1^1 = 0)$  and thus

$$\mathbb{P}(X_1^1 = 0) = \frac{1 - \rho - \lambda_1 d}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1}. \quad (12)$$

Now, the PGF's of  $X_i^1$  and  $X_i^2$  ( $i = 1, 2$ ) are given by, respectively,

$$\mathbb{E}[z_1^{X_i^1}] = F_i(z_1, 1), \quad \mathbb{E}[z_2^{X_i^2}] = F_i(1, z_2),$$

from which, by differentiation, we obtain after a lengthy calculation

$$\begin{aligned} \mathbb{E}[X_1^1] &= \frac{\lambda_1^2(d + \mathbb{E}M_1(1 - \lambda_1 d))}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1} \frac{\lambda_1 \beta_1^{(2)} + \lambda_2 \beta_2^{(2)}}{2(1 - \rho_2)(1 - \rho - \lambda_1 d)} \\ &+ \frac{\lambda_1^2}{(1 - \rho_2)(1 - \rho - \lambda_1 d)} \left( \frac{d^{(2)}}{2} + \beta_1 d \right) - \lambda_1 \frac{\rho_2}{1 - \rho_2} \left( \frac{\tilde{D}'_2(\lambda_1)}{\tilde{D}_2(\lambda_1)} - d_1 \right) \\ &- \frac{\lambda_1 \rho(d + \mathbb{E}M_1(1 - \lambda_1 d))}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1} + \frac{\lambda_1 \rho \mathbb{E}M_1}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1} \\ &+ \frac{\lambda_1^2 \rho \rho_2 \mathbb{E}M_1^2}{2(1 - \rho_2)(1 - \rho + \lambda_1 \rho \mathbb{E}M_1)} + \frac{\lambda_1(d + \rho \mathbb{E}M_1)}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1}, \end{aligned} \quad (13)$$

and

$$\mathbb{E}[X_2^2] = \lambda_2 \frac{(d + \mathbb{E}M_1(1 - \lambda_1 d))(1 - \rho_2)}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1}, \quad (14)$$

where the latter one is also easily obtained directly from the evolution equations or by the following argument: due to exhaustive service at  $Q_2$ ,  $X_2^2$  is the number of arrivals to  $Q_2$  during the total switchover time and the server's stay at  $Q_1$ . Thus,

$$\mathbb{E}[X_2^2] = \lambda_2 \left( d + \beta_1 \mathbb{P}(X_1^1 > 0) + (\mathbb{E}M_1 + a_1 \beta_1) \mathbb{P}(X_1^1 = 0) \right).$$

Using the expression for  $\mathbb{P}(X_1^1 = 0)$  given in (12), this gives (14).

To obtain (13), we have used that  $\mathbb{E}M_1 = \int_0^\infty t e^{-\lambda_1 t} d\mathbb{P}(T \leq t) = \lambda_1 \mathbb{E}M_1^2/2$ , which follows from

$$\mathbb{E}M_1 = \int_0^\infty e^{-\lambda_1 t} \mathbb{P}(T > t) dt \quad \text{and} \quad \mathbb{E}M_1^2 = \int_0^\infty 2t e^{-\lambda_1 t} \mathbb{P}(T > t) dt.$$

**Remark 3.2** *The case of zero switchover times causes no difficulty as it does in some other polling models. This is due to the presence of the*



timer. Thus, for zero switchover times all expressions above simplify by setting  $\tilde{D}_i(\cdot) \equiv 1$  and  $d_i = 0$ . In particular,

$$c_1 = F_1(0, 1) = \mathbb{P}(X_1^1 = 0) = \frac{1 - \rho}{1 - \rho + \lambda_1 \rho \mathbb{E}M_1}.$$

Note that in this case  $X_1^2 = 0$  due to exhaustive service at  $Q_2$  and therefore  $F_1(z_1, z_2)$  is constant in  $z_2$ .

### 3.2. $Q_1$ : EXHAUSTIVE; $Q_2$ : 1-LIMITED

We now consider the case of exhaustive service at  $Q_1$  and 1-limited service at  $Q_2$  and again derive the joint PGFs  $F_i(z_1, z_2)$ ; cf. (20), (15). As before, the timer is at  $Q_1$ . That is,  $Q_1$  gets an extra priority over  $Q_2$  by exercising the timer procedure when  $Q_1$  is empty, in addition to its being served exhaustively.

Since  $Q_2$  can serve at most one customer per cycle, the stability condition in this case is  $\lambda_2 \mathbb{E}C < 1$ , where  $\mathbb{E}C$  is given by (2). When  $Q_2$  is in heavy traffic, i.e., there is one service at  $Q_2$  in each cycle, then the term  $\mathbb{P}(X_1^1 = 0)$  in (2) becomes  $\tilde{D}_1(\lambda_1)\tilde{D}_2(\lambda_1)\tilde{B}_2(\lambda_1)$ . Indeed,  $Q_1$  is left behind empty because of the exhaustive service discipline, and  $\tilde{D}_1(\lambda_1)\tilde{D}_2(\lambda_1)\tilde{B}_2(\lambda_1)$  is the probability of no arrival at  $Q_1$  in the subsequent switchovers and service at  $Q_2$ . Then the stability condition  $\lambda_2 \mathbb{E}C < 1$  reduces to  $\rho + \lambda_2 d + \lambda_2 \mathbb{E}M_1 \tilde{D}_1(\lambda_1)\tilde{D}_2(\lambda_1)\tilde{B}_2(\lambda_1) < 1$ .

The evolution equations of the queue lengths at polling instants for this model are given by

$$\begin{aligned} X_1^1 &= \begin{cases} X_2^1 + A_1(B_2) + A_1(D_2), & \text{if } X_2^2 > 0, \\ X_2^1 + A_1(D_2), & \text{if } X_2^2 = 0, \end{cases} \\ X_2^1 &= \begin{cases} X_2^2 - 1 + A_2(B_2) + A_2(D_2), & \text{if } X_2^2 > 0, \\ A_2(D_2), & \text{if } X_2^2 = 0, \end{cases} \\ X_2^1 &= A_1(D_1), \\ X_2^2 &= \begin{cases} X_1^2 + A_2\left(\sum_{k=1}^{X_1^1} \theta_{1k}\right) + A_2(D_1), & \text{if } X_1^1 > 0, \\ X_1^2 + A_2(M_1) + A_2(\theta_1)\mathbf{I}(IA_1 \leq T) + A_2(D_1), & \text{if } X_1^1 = 0. \end{cases} \end{aligned}$$

From this we obtain the generating functions

$$F_2(z_1, z_2) = D_1(z_1, z_2) \left[ F_1\left(f_1(z_2), z_2\right) + F_1(0, z_2)(h(z_2) - 1) \right], \quad (15)$$

where

$$h(z_2) = \mathbb{E}\left[z_2^{A_2(M_1) + A_2(\theta_1)\mathbf{I}(IA_1 \leq T)}\right], \quad (16)$$

and

$$F_1(z_1, z_2) = D_2(z_1, z_2) \left[ B_2(z_1, z_2) \frac{F_2(z_1, z_2) - F_2(z_1, 0)}{z_2} + F_2(z_1, 0) \right]. \quad (17)$$

Note that  $h(\cdot)$  cannot be factorized since  $M_1$  and  $\mathbb{I}(IA_1 \leq T)$  are dependent. Substituting (15) into (17) yields

$$\begin{aligned} F_1(z_1, z_2) &= F_1(f_1(z_2), z_2) \frac{D_1(z_1, z_2) D_2(z_1, z_2) B_2(z_1, z_2)}{z_2} \\ &+ F_1(0, z_2) (h(z_2) - 1) \frac{D_1(z_1, z_2) D_2(z_1, z_2) B_2(z_1, z_2)}{z_2} \\ &+ c_2 D_1(z_1, 0) D_2(z_1, z_2) \left( 1 - \frac{B_2(z_1, z_2)}{z_2} \right), \end{aligned} \quad (18)$$

where

$$c_2 = F_1(\tilde{\theta}_1(\lambda_2), 0) + F_1(0, 0)(h(0) - 1)$$

and

$$h(0) = \mathbb{P}(A_2(M_1) + A_2(\theta_1) \mathbb{I}(IA_1 \leq T) = 0).$$

Put  $z_1 = f_1(z_2)$  in (18) and solve for  $F_1(f_1(z_2), z_2)$  to get

$$\begin{aligned} F_1(f_1(z_2), z_2) &= \frac{D_2(f_1(z_2), z_2)}{z_2 - D_1(f_1(z_2), z_2) D_2(f_1(z_2), z_2) B_2(f_1(z_2), z_2)} \\ &\times \left\{ F_1(0, z_2) (h(z_2) - 1) D_1(f_1(z_2), z_2) B_2(f_1(z_2), z_2) \right. \\ &\quad \left. + c_2 D_1(f_1(z_2), 0) (z_2 - B_2(f_1(z_2), z_2)) \right\}. \end{aligned} \quad (19)$$

We use the following shorthand notation, for  $i = 1, 2$ :

$$\begin{aligned} \hat{D}_i &= D_i(z_1, z_2); & D_i^* &= D_i(f_1(z_2), z_2); \\ \hat{B}_2 &= B_2(z_1, z_2); & B_2^* &= B_2(f_1(z_2), z_2). \end{aligned}$$

Plugging (19) into (18) then yields

$$\begin{aligned} F_1(z_1, z_2) &= \frac{F_1(0, z_2) (h(z_2) - 1)}{z_2 - D_1^* D_2^* B_2^*} \hat{D}_1 \hat{D}_2 \hat{B}_2 + c_2 D_1(z_1, 0) \hat{D}_2 (1 - \hat{B}_2/z_2) \\ &+ \frac{c_2 D_1(f_1(z_2), 0) D_2^* (1 - B_2^*/z_2)}{z_2 - D_1^* D_2^* B_2^*} \hat{D}_1 \hat{D}_2 \hat{B}_2, \end{aligned} \quad (20)$$

and it remains to determine  $F_1(0, z_2)$  and  $c_2$ . Setting  $z_1 = 0$  in (20) gives, for  $z_2 \neq 0$ ,

$$\begin{aligned} F_1(0, z_2) &= \frac{c_2}{z_2 - D_1^* D_2^* B_2^* - (h(z_2) - 1) D_1^{**} D_2^{**} B_2^{**}} \\ &\times \left\{ D_1(0, 0) D_2^{**} (1 - B_2^{**}/z_2) (z_2 - D_1^* D_2^* B_2^*) \right. \\ &\quad \left. + D_1(f_1(z_2), 0) D_2^* (1 - B_2^*/z_2) D_1^{**} D_2^{**} B_2^{**} \right\}, \end{aligned} \quad (21)$$

where  $D_i^{**} = D_i(0, z_2)$ ,  $i = 1, 2$ , and  $B_2^{**} = B_2(0, z_2)$ . Now, (20) gives  $F_1(z_1, z_2)$ , while  $F_2(z_1, z_2)$  follows from (15) and (19). Further, from (15) we get  $c_2 = F_2(z_1, 0)/D_1(z_1, 0)$ . Put  $z_1 = z_2 = 1$  in (20) to determine this constant. Since  $D_i(1, 1) = B_2(1, 1) = 1$ ,  $i = 1, 2$ ,

$$1 = F_1(1, 1) = \lim_{z_2 \rightarrow 1} F_1(1, z_2) = \lim_{z_2 \rightarrow 1} F_1(0, z_2) \lim_{z_2 \rightarrow 1} \frac{h(z_2) - 1}{z_2 - D_1^* D_2^* B_2^*} \\ + c_2 D_1(f_1(1), 0) \lim_{z_2 \rightarrow 1} \frac{D_2^*}{z_2} \lim_{z_2 \rightarrow 1} \frac{z_2 - B_2^*}{z_2 - D_1^* D_2^* B_2^*}. \quad (22)$$

Remember that  $f_1(1) = 1$ . According to (21),

$$\lim_{z_2 \rightarrow 1} F_1(0, z_2) = c_2 \tilde{D}_1(\lambda) \tilde{D}_2(\lambda_1) \\ \times \lim_{z_2 \rightarrow 1} \frac{z_2 - D_1^* D_2^* B_2^*}{z_2 - D_1^* D_2^* B_2^* - (h(z_2) - 1) D_1^{**} D_2^{**} B_2^{**}} \lim_{z_2 \rightarrow 1} \frac{z_2 - B_2^{**}}{z_2} \\ + c_2 \tilde{D}_1(\lambda_2) \lim_{z_2 \rightarrow 1} \frac{(z_2 - B_2^*) \tilde{D}_1(\lambda_1) \tilde{D}_2(\lambda_1) \tilde{B}_2(\lambda_1)}{z_2 - D_1^* D_2^* B_2^* - (h(z_2) - 1) D_1^{**} D_2^{**} B_2^{**}}.$$

Thus, by using l'Hospital's rule,

$$F_1(0, 1) = c_2 \tilde{D}_2(\lambda_1) \quad (23) \\ \times \frac{\tilde{D}_1(\lambda)(1 - \rho - \lambda_2 d)(1 - \tilde{B}_2(\lambda_1)) + \tilde{D}_1(\lambda_1) \tilde{D}_1(\lambda_2) \tilde{B}_2(\lambda_1)(1 - \rho)}{1 - \rho - \lambda_2 d - \lambda_2 \mathbb{E} M_1 \tilde{D}_1(\lambda_1) \tilde{D}_2(\lambda_1) \tilde{B}_2(\lambda_1)}.$$

Further,

$$\lim_{z_2 \rightarrow 1} \frac{h(z_2) - 1}{z_2 - D_1^* D_2^* B_2^*} = \frac{\lambda_2 (\mathbb{E} M_1 + a_1 \mathbb{E} \theta_1)}{1 - \lambda_2 (\lambda_1 \mathbb{E} \theta_1 + 1)(d + \beta_2)} = \frac{\lambda_2 \mathbb{E} M_1}{1 - \rho - \lambda_2 d} \quad (24)$$

and

$$\lim_{z_2 \rightarrow 1} \frac{z_2 - B_2^*}{z_2 - D_1^* D_2^* B_2^*} = \frac{1 - \lambda_2 (\lambda_1 \mathbb{E} \theta_1 + 1) \beta_2}{1 - \lambda_2 (\lambda_1 \mathbb{E} \theta_1 + 1)(d + \beta_2)} = \frac{1 - \rho}{1 - \rho - \lambda_2 d}. \quad (25)$$

Thus, we finally get from (22), (23), (24) and (25):

$$c_2 = \frac{1 - \rho - \lambda_2 d - \lambda_2 \mathbb{E} M_1 \tilde{D}_1(\lambda_1) \tilde{D}_2(\lambda_1) \tilde{B}_2(\lambda_1)}{\lambda_2 \mathbb{E} M_1 \tilde{D}_1(\lambda) \tilde{D}_2(\lambda_1)(1 - \tilde{B}_2(\lambda_1)) + (1 - \rho) \tilde{D}_1(\lambda_2)}. \quad (26)$$

**Remark 3.3** Notice that  $c_2 > 0$  since the stability condition  $\lambda_2 \mathbb{E} C < 1$  holds. Again, exploiting the 1-limited service yields a more direct approach to determine  $c_2$ . We have

$$\lambda_2 \mathbb{E} C = \mathbb{E}[\text{number of services at } Q_2 \text{ per cycle}] = \mathbb{P}(X_2^2 > 0).$$

Substituting (2) and using (15) with  $z_1 = 1$ ,  $z_2 = 0$ , and (23) to express  $\mathbb{P}(X_2^2 = 0)$  and  $\mathbb{P}(X_1^1 = 0)$  in terms of  $c_2$ , respectively, yields (26).

$F_1(z_1, z_2)$  is determined by substituting the value of  $c_2$  into (20) and (21) and then substituting (21) into (20).  $F_2(z_1, z_2)$  is then determined from (15), (19) and (21). Queue length moments can now be obtained in a similar manner as in Subsection 3.1.

**Remark 3.4** *As before, for zero switchover times all expressions simplify by setting  $\tilde{D}_i(\cdot) \equiv 1$  and  $d_i = 0$ .*

### 3.3. EXHAUSTIVE REGIME IN BOTH QUEUES

In this section both  $Q_1$  and  $Q_2$  are assumed to be served exhaustively and, as before, the timer is at  $Q_1$ . Again we derive the joint PGFs  $F_i(z_1, z_2)$ ; cf. (31), (27). It is easily seen that the stability condition now is  $\rho_1 + \rho_2 < 1$ . For the system under consideration the evolution equations of the queue lengths at polling instants are given by

$$\begin{aligned} X_2^1 &= A_1(D_1), \quad X_1^2 = A_2(D_2), \quad X_1^1 = X_2^1 + A_1\left(\sum_{k=1}^{X_2^2} \theta_{2k}\right) + A_1(D_2), \\ X_2^2 &= \begin{cases} X_1^2 + A_2\left(\sum_{k=1}^{X_1^1} \theta_{1k}\right) + A_2(D_1), & X_1^1 > 0, \\ X_1^2 + A_2(M_1) + A_2(\theta_1)\mathbb{I}(IA_1 \leq T) + A_2(D_1), & X_1^1 = 0. \end{cases} \end{aligned}$$

From this we obtain the generating functions

$$F_1(z_1, z_2) = D_2(z_1, z_2) F_2\left(z_1, f_2(z_1)\right), \quad (27)$$

and

$$F_2(z_1, z_2) = D_1(z_1, z_2) \left[ F_1\left(f_1(z_2), z_2\right) - F_1(0, z_2) + F_1(0, z_2) h(z_2) \right], \quad (28)$$

where  $h(z_2)$  is given by (16). With the notation

$$k(z_2) = h(z_2) - 1, \quad g(z_2) = \tilde{\theta}_2\left(\lambda_1(1 - f_1(z_2))\right) = f_2(f_1(z_2)), \quad (29)$$

substituting (27) into (28) yields

$$\begin{aligned} F_2(z_1, z_2) &= D_1(z_1, z_2) D_2(f_1(z_2), z_2) F_2\left(f_1(z_2), g(z_2)\right) \\ &\quad + D_1(z_1, z_2) D_2(0, z_2) F_2\left(0, \tilde{\theta}_2(\lambda_1)\right) k(z_2). \end{aligned}$$

With

$$\begin{aligned} D(z_1, z_2) &= D_1(z_1, z_2) D_2(f_1(z_2), z_2), \\ E(z_1, z_2) &= D_1(z_1, z_2) D_2(0, z_2) k(z_2), \\ c_3 &= F_2\left(0, \tilde{\theta}_2(\lambda_1)\right), \end{aligned} \quad (30)$$

we get by iteration, after  $K$  steps,

$$\begin{aligned}
F_2(z_1, z_2) &= D(z_1, z_2) \\
&\times \prod_{k=0}^{K-1} D\left(f_1(g^{(k)}(z_2)), g^{(k+1)}(z_2)\right) \times F_2\left(f_1(g^{(K)}(z_2)), g^{(K+1)}(z_2)\right) \\
&+ c_3 D(z_1, z_2) \sum_{k=0}^{K-1} E\left(f_1(g^{(k)}(z_2)), g^{(k+1)}(z_2)\right) \\
&\times \prod_{j=0}^{k-1} D\left(f_1(g^{(j)}(z_2)), g^{(j+1)}(z_2)\right) + c_3 E(z_1, z_2),
\end{aligned}$$

where

$$g^{(0)}(z) = z, \quad g^{(k)}(z) = g(g^{(k-1)}(z)), \quad k \geq 1.$$

Letting  $K \rightarrow \infty$  (for convergence see Boxma, Schlegel and Yechiali [10]) gives

$$\begin{aligned}
F_2(z_1, z_2) &= D(z_1, z_2) \\
&\times \prod_{k=0}^{\infty} D\left(f_1(g^{(k)}(z_2)), g^{(k+1)}(z_2)\right) \times F_2\left(f_1(g^{(\infty)}(z_2)), g^{(\infty)}(z_2)\right) \\
&+ c_3 D(z_1, z_2) \sum_{k=0}^{\infty} E\left(f_1(g^{(k)}(z_2)), g^{(k+1)}(z_2)\right) \\
&\times \prod_{j=0}^{k-1} D\left(f_1(g^{(j)}(z_2)), g^{(j+1)}(z_2)\right) + c_3 E(z_1, z_2), \tag{31}
\end{aligned}$$

and it remains to determine  $g^{(\infty)}$  and  $c_3$ . By definition

$$g^{(k)}(z_2) = \tilde{\theta}_2\left(\lambda_1(1 - f_1(g^{(k-1)}(z_2)))\right)$$

and thus

$$g^{(\infty)}(z_2) = \tilde{\theta}_2\left(\lambda_1(1 - f_1(g^{(\infty)}(z_2)))\right)$$

which is solved by  $g^{(\infty)} \equiv 1$ . Since  $\tilde{\theta}_2\left(\lambda_1(1 - f_1(x))\right)$  is a convex function of  $x$  with

$$\left. \frac{\partial}{\partial x} \tilde{\theta}_2\left(\lambda_1(1 - f_1(x))\right) \right|_{x=1} = \lambda_1 \mathbb{E}\theta_2 \lambda_2 \mathbb{E}\theta_1 = \lambda_1 \frac{\beta_2}{1 - \rho_2} \lambda_2 \frac{\beta_1}{1 - \rho_1} < 1$$

under the stability condition  $\rho_1 + \rho_2 < 1$ , there is no other solution. To get  $c_3$  we put  $z_1 = 0$  and  $z_2 = \tilde{\theta}_2(\lambda_1)$  in (31) (cf. (30)):

$$c_3 = D(0, z) \prod_{k=0}^{\infty} D\left(f_1(g^{(k)}(z)), g^{(k+1)}(z)\right) \times \left[1 - E(0, z) - D(0, z)\right]$$

$$\times \sum_{k=0}^{\infty} E\left(f_1(g^{(k)}(z)), g^{(k+1)}(z)\right) \prod_{j=0}^{k-1} D\left(f_1(g^{(j)}(z)), g^{(j+1)}(z)\right) \Big|_{z=\tilde{\theta}_2(\lambda_1)},$$

and  $F_2(z_1, z_2)$  is determined. Finally,  $F_1(z_1, z_2)$  is obtained by substituting (31) with  $z_2 = f_1(z_2)$  into (27).

#### 4. Total Workload and a Pseudoconservation Law

We now investigate the total workload in the system. We show that the amount of work in the system can be decomposed into the amount of work in the 'corresponding'  $M/G/1$  model (the model without timer and switchover times) and the amount of work in the system at an arbitrary epoch in a non-serving interval, i.e. in a timer period or a switchover period; cf. (34), (38). The fact that the total workload can also be expressed in terms of mean waiting times then allows us to obtain a pseudoconservation law for the mean waiting times, i.e., an exact expression for a particular weighted sum of the mean waiting times at  $Q_1$  and  $Q_2$ ; cf. (37) and Theorem 4.1. It should be noted that the results obtained in this section hold for general service disciplines and a timer at  $Q_1$ .

##### *Zero switchover times*

We first consider the case of zero switchover times but a nonzero timer. Then, whenever the system is empty, the server stays at  $Q_1$  waiting. Indeed, if there are no arrivals, neither to  $Q_1$  nor to  $Q_2$ , while the timer at  $Q_1$  is active, then, after expiration of the timer, the server switches to  $Q_2$  and immediately back to  $Q_1$ , where a new timer starts.

In order to derive the LST  $\varphi(s) = \int_{0-}^{\infty} e^{-sy} dV(y)$  of the (total) workload distribution  $V(\cdot)$  with density  $v(\cdot)$  we use the argument that in steady state the probability for downcrossing a level  $x$  is the same as that for upcrossing it. In our model a downcrossing is only possible if the timer is off. Denoting  $v_0(x) = \frac{d}{dx} \mathbb{P}(V \leq x, \text{timer on})$ ,  $[v(x) - v_0(x)]dt$  is the probability of downcrossing  $x$  in the next  $dt$ . Hence

$$v(x) - v_0(x) = \lambda \int_{0-}^x (1 - B(x - y)) dV(y). \quad (32)$$

Multiplying both sides of (32) by  $e^{-sx}$  and integrating over the positive real line yields, noticing that  $V(0) = \mathbb{P}(V = 0) = \mathbb{P}(V = 0, \text{timer on})$ ,

$$\begin{aligned} & \varphi(s) - \mathbb{E}\left[e^{-sV} \mathbb{I}(\text{timer on})\right] \\ &= \lambda \int_0^{\infty} e^{-sx} \int_{0-}^x (1 - B(x - y)) dV(y) dx = \rho \beta_e(s) \varphi(s), \end{aligned}$$

where  $\beta_e(s)$  is the LST of the residual service time distribution  $B_e(x) = \int_0^x (1 - B(y))dy/\beta$  with  $\beta = (\lambda_1\beta_1 + \lambda_2\beta_2)/\lambda$ . Thus,

$$\varphi(s) = \frac{1 - \rho}{1 - \rho\beta_e(s)} \mathbb{E}\left[e^{-sV} \mid \text{timer on}\right], \quad (33)$$

since, with zero switchover times,  $\mathbb{P}(\text{timer on}) = 1 - \rho$ . From (33) we immediately get the following work decomposition:

$$V \stackrel{d}{=} V_{M/G/1} + V|_{\text{timer on}}, \quad (34)$$

where the distribution of  $V_{M/G/1}$  is the stationary distribution of the workload in an  $M/G/1$  queue with arrival rate  $\lambda$  and service time distribution  $B(\cdot)$  having LST  $(1 - \rho)/(1 - \rho\beta_e(s))$ , see Cohen [11], and  $V|_{\text{timer on}}$  is independent of  $V_{M/G/1}$  and has the stationary workload distribution in our model given that the timer at  $Q_1$  is active. We now study this quantity more closely.

First we introduce the following quantities:  $M_i^{(*)}$  denotes the amount of unfinished work left by the server at  $Q_i$ , and  $M_2^{(1)}$  denotes the amount of unfinished work left by the server at  $Q_2$  on condition that  $Q_1$  is empty at that moment. When the timer starts,  $Q_1$  is empty whereas in  $Q_2$  the workload consists of the unfinished work left behind when the server leaves this queue. Moreover, while the timer is active, only arrivals to  $Q_2$  may occur. When the timer is on, as soon as there is an arrival at  $Q_1$  or the timer expires, whichever occurs first, the timer is switched off. Since  $\{V|_{\text{timer on}}(t), t \geq 0\}$  is a regenerative process, we get for its LST:

$$\begin{aligned} \mathbb{E}\left(e^{-sV|_{\text{timer on}}}\right) &= \frac{1}{\mathbb{E}M_1} \mathbb{E} \int_0^{M_1} e^{-sV|_{\text{timer on}}(t)} dt \\ &= \frac{1}{\mathbb{E}M_1} \mathbb{E} \int_0^\infty \mathbb{I}(M_1 > t) e^{-s(M_2^{(1)} + \sum_{k=1}^{A_2(t)} B_{2k})} dt \\ &= \frac{\mathbb{E}e^{-sM_2^{(1)}}}{\mathbb{E}M_1} \int_0^\infty \mathbb{P}(M_1 > t) e^{-\lambda_2(1 - \beta_2(s))t} dt \\ &= \mathbb{E}e^{-sM_2^{(1)}} \frac{1 - \tilde{M}_1(\lambda_2(1 - \beta_2(s)))}{\mathbb{E}M_1 \lambda_2(1 - \beta_2(s))}. \end{aligned} \quad (35)$$

From (34) and (35) we get

$$\mathbb{E}V = \mathbb{E}V_{M/G/1} + \mathbb{E}V|_{\text{timer on}} = \frac{\sum_{i=1}^2 \lambda_i \beta_i^{(2)}}{2(1 - \rho)} + \mathbb{E}M_2^{(1)} + \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1}.$$

In the special case of a timer that is exponentially distributed with parameter  $\xi$ , we get

$$\mathbb{E}\left(e^{-sV} | \text{timer on}\right) = \frac{\mathbb{E}e^{-sM_2^{(1)}} (\xi + \lambda_1)}{\xi + \lambda_1 + \lambda_2(1 - \beta_2(s))}$$

and

$$\mathbb{E}V = \frac{\sum_{i=1}^2 \lambda_i \beta_i^{(2)}}{2(1 - \rho)} + \mathbb{E}M_2^{(1)} + \frac{\rho_2}{\xi + \lambda_1}.$$

Also, for any service regime

$$\begin{aligned} \mathbb{E}V &= \sum_{i=1}^2 \beta_i \mathbb{E}[\text{number of customers waiting at } Q_i] + \sum_{i=1}^2 \rho_i \frac{\beta_i^{(2)}}{2\beta_i} \\ &= \sum_{i=1}^2 \rho_i \mathbb{E}W_i + \sum_{i=1}^2 \lambda_i \frac{\beta_i^{(2)}}{2}, \end{aligned} \quad (36)$$

where the last equality follows from Little's law and  $W_i$  denotes the stationary waiting time at  $Q_i$ . Therefore, for zero switchover times we get the following pseudoconservation law:

$$\sum_{i=1}^2 \rho_i \mathbb{E}W_i = \frac{\rho}{1 - \rho} \frac{\sum_{i=1}^2 \lambda_i \beta_i^{(2)}}{2} + \mathbb{E}M_2^{(1)} + \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1}. \quad (37)$$

#### *Nonzero switchover times*

We now turn to the case of nonzero switchover times, in which the (total) workload is not decreasing both when the timer is active as well as when the server is switching. Therefore, with  $v_s(x) = \frac{d}{dx} \mathbb{P}(V \leq x, \text{server switches})$ ,

$$v(x) - v_0(x) - v_s(x) = \lambda \int_{0-}^x (1 - B(x - y)) dV(y).$$

Analogous to the case of zero switchover times we get

$$\begin{aligned} \varphi(s) - V(0) - \left[ \mathbb{E}(e^{-sV} \mathbb{I}(\text{timer on})) - \mathbb{P}(V = 0, \text{timer on}) \right] \\ - \left[ \mathbb{E}(e^{-sV} \mathbb{I}(\text{switch})) - \mathbb{P}(V = 0, \text{switch}) \right] = \rho \beta_e(s) \varphi(s). \end{aligned}$$

Since  $V(0) = \mathbb{P}(V = 0, \text{timer on}) + \mathbb{P}(V = 0, \text{switch})$ , we have

$$\begin{aligned} \varphi(s) &= \frac{\mathbb{E}(e^{-sV} \mathbb{I}(\text{timer on})) + \mathbb{E}(e^{-sV} \mathbb{I}(\text{switch}))}{1 - \rho \beta_e(s)} \\ &= \frac{1 - \rho}{1 - \rho \beta_e(s)} \left\{ q \mathbb{E}(e^{-sV} | \text{timer on}) + (1 - q) \mathbb{E}(e^{-sV} | \text{switch}) \right\}, \end{aligned}$$



where  $q = \mathbb{P}(\text{timer on})/\mathbb{P}(\text{timer on or switch})$ . Thus, the following work decomposition holds:

$$V \stackrel{d}{=} V_{M/G/1} + Y, \quad (38)$$

with  $V_{M/G/1}$  and  $Y$  independent and  $Y$  is specified by  $\mathbb{P}(Y = V|_{\text{timer on}}) = q = 1 - \mathbb{P}(Y = V|_{\text{switch}})$ . We compute  $\mathbb{E}(V|_{\text{timer on}})$  and  $\mathbb{E}(V|_{\text{switch}})$ . When the timer starts,  $Q_1$  is empty. Now, other than before, in  $Q_2$  the workload consists of the unfinished work  $M_2^{(1)}$  left behind when the server leaves this queue ( $Q_1$  being empty) as well as the work that has arrived during the switchover from  $Q_2$  to  $Q_1$ . Note that this is not an ordinary switchover time  $D_2$ , but has to be conditioned on the event that there are no arrivals to  $Q_1$  during this period. Therefore,

$$\begin{aligned} \mathbb{E}(V|_{\text{timer on}}) &= \mathbb{E}M_2^{(1)} + \frac{\int_0^\infty \rho_2 t e^{-\lambda_1 t} dD_2(t)}{\int_0^\infty e^{-\lambda_1 t} dD_2(t)} + \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1} \\ &= \mathbb{E}M_2^{(1)} - \frac{\rho_2 \tilde{D}'_2(\lambda_1)}{\tilde{D}_2(\lambda_1)} + \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1}, \end{aligned} \quad (39)$$

where  $\tilde{D}'_2(\lambda_1) = \left. \frac{d}{ds} \tilde{D}_2(s) \right|_{s=\lambda_1}$ . The first two terms on the right hand side of (39) represent the workload at the instant when the timer starts, which consists of the unfinished work left behind at  $Q_2$  and of the arriving work to  $Q_2$  during the switchover  $D_2$ , given that there are no arrivals to  $Q_1$  during that time. The third term arises since work increases at  $Q_2$  at rate  $\rho_2$ . We now compute  $\mathbb{E}(V|_{\text{switch}}) = (p_1 \mathbb{E}(V|_{D_1}) + p_2 \mathbb{E}(V|_{D_2}))/\mathbb{P}(\text{switch})$ , where  $p_i = d_i/\mathbb{E}C$  is the probability that the server is switching from  $Q_i$  to the other queue,  $i = 1, 2$ , and  $V|_{D_i}$  denotes the workload at an arbitrary epoch in such a switchover. At the beginning of the switchover period from  $Q_1$  to  $Q_2$ , the total workload consists of the work left at  $Q_1$ , if there is any left, the work at  $Q_2$  that has been left there on the server's departure from  $Q_2$ , the work that has arrived to  $Q_2$  during the switchover time  $D_2$  from  $Q_2$  to  $Q_1$ , and the work that has arrived to  $Q_2$  during the server's stay at  $Q_1$ , where the mean of this stay is given by  $\rho_1 \mathbb{E}C + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)$ . Thus (remember that  $M_i^{(*)}$  is defined above (35) as the amount of unfinished work left by the server at  $Q_i$ ):

$$\begin{aligned} \mathbb{E}(V|_{D_1}) &= \mathbb{E}M_1^{(*)} + \mathbb{E}M_2^{(*)} + \rho_2 d_2 \\ &\quad + \rho_2 \rho_1 \frac{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{1 - \rho} + \rho_2 \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0) + \frac{\rho \mathbb{E}D_1^2}{2\mathbb{E}D_1}. \end{aligned}$$

Similarly,

$$\mathbb{E}(V|_{D_2}) = \mathbb{E}M_1^{(*)} + \mathbb{E}M_2^{(*)} + \rho_1 d_1 + \rho_1 \rho_2 \frac{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{1 - \rho} + \frac{\rho \mathbb{E}D_2^2}{2\mathbb{E}D_2}.$$

With

$$q = \frac{\mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}$$

and  $(1 - q)/\mathbb{P}(\text{switch}) = 1/\mathbb{P}(\text{timer on or switch}) = 1/(1 - \rho)$  we finally get

$$\begin{aligned} \mathbb{E}V &= \mathbb{E}V_{M/G/1} + q\mathbb{E}V|_{\text{timer on}} + (1 - q)\mathbb{E}V|_{\text{switch}} \\ &= \frac{\sum_{i=1}^2 \lambda_i \beta_i^{(2)}}{2(1 - \rho)} + \frac{\mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)} \left( \mathbb{E}M_2^{(1)} - \frac{\rho_2 \tilde{D}'_2(\lambda_1)}{\tilde{D}_2(\lambda_1)} + \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1} \right) \\ &\quad + \frac{d}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)} \left( \mathbb{E}M_1^{(*)} + \mathbb{E}M_2^{(*)} \right) + \frac{\rho d^{(2)}}{2(d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0))} \\ &\quad + \frac{\mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)} d_1 \rho_2 + \frac{d}{2(1 - \rho)} \left[ \rho^2 - \sum_{i=1}^2 \rho_i^2 \right]. \end{aligned} \quad (40)$$

For  $T = 0$ , i.e., when the server does not wait for an arrival at  $Q_1$ , the above result coincides with the well-known result, see e.g. [5]. Combining (40) and (36) we obtain the following result.

**Theorem 4.1** (Pseudoconservation Law) *In the 2-queue polling system under consideration with nonzero switchover times and a timer at  $Q_1$ ,*

$$\begin{aligned} \sum_{i=1}^2 \rho_i \mathbb{E}W_i &= \frac{\rho}{1 - \rho} \sum_{i=1}^2 \frac{\lambda_i \beta_i^{(2)}}{2} \\ &\quad + \frac{\mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)} \left( \mathbb{E}M_2^{(1)} - \frac{\rho_2 \tilde{D}'_2(\lambda_1)}{\tilde{D}_2(\lambda_1)} + \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1} \right) \\ &\quad + \frac{d}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)} \left( \mathbb{E}M_1^{(*)} + \mathbb{E}M_2^{(*)} \right) + \frac{\rho d^{(2)}}{2(d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0))} \\ &\quad + \frac{\mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)}{d + \mathbb{E}M_1 \mathbb{P}(X_1^1 = 0)} d_1 \rho_2 + \frac{d}{2(1 - \rho)} \left[ \rho^2 - \sum_{i=1}^2 \rho_i^2 \right]. \end{aligned} \quad (41)$$

Note that the terms  $\mathbb{E}M_i^{(*)}$  and  $\mathbb{E}M_2^{(1)}$  (the amount of unfinished work left by the server) depend on the service discipline at  $Q_i$  and can thus only be determined after specifying the service discipline at. For exhaustive service at  $Q_2$ , for example,  $\mathbb{E}M_2^{(1)} = \mathbb{E}M_2^{(*)} = 0$ .

**Corollary 4.1** *For 1-limited service at  $Q_1$  and exhaustive service at  $Q_2$ , we have*

$$\rho_1 \frac{1 - \rho - \lambda_1 d}{1 - \rho} \mathbb{E}W_1 + \rho_2 \mathbb{E}W_2$$

$$\begin{aligned}
&= \frac{\rho}{1-\rho} \sum_{i=1}^2 \frac{\lambda_i \beta_i^{(2)}}{2} + \frac{\mathbb{E}M_1(1-\rho-\lambda_1 d)}{(1-\rho)(d+\mathbb{E}M_1(1-\lambda_1 d))} \\
&\quad \times \left( \frac{\rho_2 \mathbb{E}M_1^2}{2\mathbb{E}M_1} - \frac{\rho_2 \tilde{D}'_2(\lambda_1)}{\tilde{D}_2(\lambda_1)} \right) + \frac{\rho_1^2 d}{1-\rho} + \frac{\rho(1-\rho+\lambda_1 \rho \mathbb{E}M_1)d^{(2)}}{2(1-\rho)(d+\mathbb{E}M_1(1-\lambda_1 d))} \\
&\quad + \frac{\mathbb{E}M_1(1-\rho-\lambda_1 d)}{(1-\rho)(d+\mathbb{E}M_1(1-\lambda_1 d))} d_1 \rho_2 + \frac{d}{2(1-\rho)} \left[ \rho^2 - \sum_{i=1}^2 \rho_i^2 \right].
\end{aligned}$$

*Proof.* The server can only leave unfinished work at  $Q_1$  if on departure he has just completed serving a customer, which happens with probability  $\lambda_1 \mathbb{E}C$ . The amount of work left behind then consists of the work that has arrived during this customer's sojourn time at  $Q_1$ . Thus,

$$\begin{aligned}
\mathbb{E}M_1^{(1)} &= \lambda_1 \mathbb{E}C \rho_1 (\mathbb{E}W_1 + \beta_1) \\
&= \lambda_1 \frac{d + \mathbb{E}M_1(1-\lambda_1 d)}{1-\rho + \lambda_1 \rho \mathbb{E}M_1} \rho_1 \mathbb{E}W_1 + \rho_1^2 \frac{d + \mathbb{E}M_1(1-\lambda_1 d)}{1-\rho + \lambda_1 \rho \mathbb{E}M_1},
\end{aligned}$$

where we have used  $\mathbb{P}(X_1^1 = 0) = \frac{1-\rho-\lambda_1 d}{1-\rho+\lambda_1 \rho \mathbb{E}M_1}$ , which is obtained from (6) with  $z_2 = 1$  and (11). Since  $\mathbb{E}M_2^{(1)} = \mathbb{E}M_2^{(*)} = 0$ , the assertion now immediately follows from Theorem 4.1.

## 5. Waiting Times

In general, it is difficult to calculate closed form expressions for the mean waiting times at isolated queues in a polling system. Then, the pseudoconservation law is often the only exact information on waiting times that can be obtained. However, if the scheduling discipline is not too complicated, explicit expressions for the mean waiting times are available; this is for example the case in the three models of Section 3.

We now specify the service disciplines at the two queues to be 1-limited at  $Q_1$  and exhaustive at  $Q_2$  as in Section 3.1. For this model we are able to derive the LST of the stationary waiting times  $W_1$  and  $W_2$ ; cf. (43), (44). Taking derivatives we also obtain explicit expressions for the mean waiting times  $\mathbb{E}W_1$  and  $\mathbb{E}W_2$ ; cf. (45), (46).

Let  $X_i$  denote the number of customers at  $Q_i$  at the beginning of a *servicing* interval at  $Q_i$  (exercising the timer at  $Q_1$  is considered as a *non-servicing* period), and let  $Y_i$  denote the number of customers at the end of a *visit* of the server at  $Q_i$ . Then the LST of  $W_i$  are given by, see e.g. [3],

$$\mathbb{E}e^{-sW_i} = \frac{(1-\rho_i)s}{s-\lambda_i(1-\tilde{B}_i(s))} \frac{\mathbb{E}[(1-s/\lambda_i)^{Y_i}] - \mathbb{E}[(1-s/\lambda_i)^{X_i}]}{(\mathbb{E}X_i - \mathbb{E}Y_i)s/\lambda_i}. \quad (42)$$

In (42), the first factor is the LST of the stationary waiting time  $W_i$  in the corresponding  $M/G/1$ -model. For arbitrary service discipline at  $Q_1$  we have

$$X_1 = X_1^1 \mathbb{I}(X_1^1 > 0) + \mathbb{I}(X_1^1 = 0) \mathbb{I}(IA_1 \leq T)$$

and therefore

$$\mathbb{E}[z^{X_1}] = F_1(z, 1) + a_1 \mathbb{P}(X_1^1 = 0)(z - 1)$$

and  $\mathbb{E}X_1 = \mathbb{E}X_1^1 + a_1 \mathbb{P}(X_1^1 = 0)$ . Further, since  $Y_1 = X_2^1 - A_1(D_1)$ ,

$$\begin{aligned} \mathbb{E}[z^{Y_1}] &= \frac{F_2(z, 1)}{D_1(z, 1)} = \tilde{B}_1(\lambda_1(1-z)) \frac{F_1(z, 1) - F_1(0, 1)}{z} \\ &\quad + F_1(0, 1)(a_1 \tilde{B}_1(\lambda_1(1-z)) + 1 - a_1), \end{aligned}$$

where in the last step we have used (4), and

$$\mathbb{E}Y_1 = \mathbb{E}X_1^1 - (1 - F_1(0, 1)) + \rho_1(1 - F_1(0, 1)) + \rho_1 a_1 F_1(0, 1).$$

Thus, with  $s = \lambda_1(1-z)$ ,

$$\begin{aligned} \mathbb{E}[e^{-\lambda_1(1-z)W_1}] &= \frac{F_1(z, 1) - F_1(0, 1)}{z(1 - F_1(0, 1) + a_1 F_1(0, 1))} + \frac{a_1 F_1(0, 1)}{1 - F_1(0, 1) + a_1 F_1(0, 1)}. \end{aligned} \quad (43)$$

Similarly, with  $X_2 = X_2^2$  and  $Y_2 = 0$ ,

$$\mathbb{E}[e^{-\lambda_2(1-z)W_2}] = \frac{1 - \rho_2}{\mathbb{E}X_2^2} \frac{1 - F_2(1, z)}{\tilde{B}_2(\lambda_2(1-z)) - z}. \quad (44)$$

Now the mean waiting times can be calculated from (43) and (44) by taking derivatives. Together with the expressions in (13) and (14) for the mean queue lengths  $\mathbb{E}X_1^1$  and  $\mathbb{E}X_2^2$  under the 1-limited discipline at  $Q_1$  and the exhaustive discipline at  $Q_2$  we obtain

$$\begin{aligned} \mathbb{E}W_1 &= \frac{\lambda_1 \beta_1^{(2)} + \lambda_2 \beta_2^{(2)}}{2(1 - \rho_2)(1 - \rho - \lambda_1 d)} \\ &\quad + \frac{1 - \rho + \lambda_1 \rho \mathbb{E}M_1}{1 - \rho_2} \frac{1}{1 - \rho - \lambda_1 d} \frac{d^{(2)} + 2\beta_1 d}{2(d + \mathbb{E}M_1(1 - \lambda_1 d))} \\ &\quad - \frac{\rho_2(1 - \rho + \lambda_1 \rho \mathbb{E}M_1)}{\lambda_1(d + \mathbb{E}M_1(1 - \lambda_1 d))(1 - \rho_2)} \left( \frac{\tilde{D}'_2(\lambda_1)}{\tilde{D}_2(\lambda_1)} - d_1 \right) - \frac{\rho}{\lambda_1} \\ &\quad + \frac{\rho \mathbb{E}M_1}{\lambda_1(d + \mathbb{E}M_1(1 - \lambda_1 d))} + \frac{\rho_2}{1 - \rho_2} \frac{\rho \mathbb{E}M_1^2}{2(d + \mathbb{E}M_1(1 - \lambda_1 d))}, \end{aligned} \quad (45)$$

and

$$\begin{aligned}
\mathbb{E}W_2 &= \frac{\lambda_1\beta_1^{(2)} + \lambda_2\beta_2^{(2)}}{2(1-\rho_2)} + \frac{1-\rho + \lambda_1\rho\mathbb{E}M_1}{1-\rho_2} \frac{d^{(2)} + 2\beta_1d}{2(d + \mathbb{E}M_1(1 - \lambda_1d))} \\
&+ \frac{1-\rho - \lambda_1d}{(d + \mathbb{E}M_1(1 - \lambda_1d))(1-\rho_2)} \left( \frac{\bar{D}'_2(\lambda_1)}{\bar{D}_2(\lambda_1)} - d_1 \right) (\beta_1 - \mathbb{E}M_1(1 + \rho_1)) \\
&+ \frac{(1-\rho - \lambda_1d)(1 + \rho_1)}{(d + \mathbb{E}M_1(1 - \lambda_1d))(1-\rho_2)} \frac{\mathbb{E}M_1^2}{2}, \tag{46}
\end{aligned}$$

and for  $T = 0$  this coincides with the known results for the model without timer, see e.g. [17] or [16], p.105.

## 6. Conclusions

The results of this paper may be generalized in several directions. Firstly, the analysis in Subsection 3.3 may be extended to the case of service disciplines at  $Q_1$  and  $Q_2$  that are of a *branching* type as studied by Resing [20]. This class includes not only the exhaustive and gated service disciplines, but also more general disciplines that allow the joint queue length process to be a multi-type branching process.

Secondly, the analysis in Subsection 3.3 may be further extended to the case of  $N > 2$  queues, but with a timer mechanism at only one queue. On the other hand, we see major problems in extending our results to cases where there are timers at more than one queue. It is also challenging to study the models of Subsections 3.1 and 3.2 when the exhaustive service discipline is replaced by the gated discipline.

## Acknowledgment

The authors are grateful to the referees for several useful suggestions.

## References

1. Avi-Itzhak, B., Maxwell, W.L. and Miller, L. (1965) Queuing with alternating priorities. *Oper. Res.* **13**, 306–318.
2. Borovkov, A.A. (1998) *Ergodicity and Stability of Stochastic Processes*. Wiley, Chichester.
3. Borst, S.C. and Boxma, O.J. (1997) Polling models with and without switchover times. *Oper. Res.* **45**, 536–543.
4. Boxma, O.J. and Down, D.G. (1997) Dynamic server assignment in a two-queue model. *Eur. J. Oper. Res.* **103**, 595–609.
5. Boxma, O.J. and Groenendijk, W.P. (1987) Pseudo-conservation laws in cyclic service systems. *J. Appl. Prob.* **24**, 949–964.

6. Boxma, O.J. and Groenendijk, W.P. (1988) Two queues with alternating service and switching times. In: *Queueing Theory and Its Applications*, eds. O.J. Boxma and R. Syski (North-Holland, Amsterdam), 261–282.
7. Boxma, O.J., Koole, G. and Mitrani, I. (1995) A two-queue polling model with a threshold service policy. In: *Proceeding MASCOTS '95*, ed. E. Gelenbe (IEEE Computer Society Press, Los Alamitos, CA), 84–89.
8. Boxma, O.J., Koole, G., and Mitrani, I. (1995) Polling models with threshold switching. In: *Quantitative Methods in Parallel Systems*, eds. F. Baccelli, A. Jean-Marie and I. Mitrani (Springer Verlag, Berlin), 129–140.
9. Boxma, O.J., Schlegel, S. and Yechiali, U. (2000) On the  $M/G/1$  queue with a waiting server, timer and vacations. Technical report 2000-038, Eurandom. Accepted for publication in *Advances in the Mathematical Sciences*, AMS Translation Series.
10. Boxma, O.J., Schlegel, S. and Yechiali, U. (2001) Two-queue polling models with a patient server. Technical report 2001-029, Eurandom.
11. Cohen, J.W. (1982) *The Single Server Queue*. North-Holland Publ. Cy., Amsterdam; 2nd ed.
12. Eisenberg, M. (1971) Two queues with changeover times. *Oper. Res.* **19**, 386–401.
13. Eliazar, I. and Yechiali, U. (1998) Randomly timed gated queueing systems. *SIAM J. Appl. Math.* **59**, 423–441.
14. Eliazar, I. and Yechiali, U. (2000) A communication multiplexer problem: Two alternating queues with dependent randomly-timed gated regime. Technical report, Dept. of Statistics and OR, Tel Aviv University.
15. Fricker, C. and Jaibi, M.R. (1994) Monotonicity and stability of periodic polling models. *Queueing Systems* **15**, 211–238.
16. Groenendijk, W.P. (1990) *Conservation Laws in Polling Systems*. Ph.D. Thesis, University of Utrecht, Utrecht.
17. Ibe, O.C. (1990) Analysis of polling systems with mixed service disciplines. *Stoch. Mod.* **6**, 667–689.
18. Lee, D.-S. (1996) A two-queue model with exhaustive and limited service disciplines. *Stoch. Mod.* **12**, 285–305.
19. Peköz, E.A. (1999) More on using forced idle time to improve performance in polling models. *Prob. Eng. and Inf. Sci.* **13**, 489–496.
20. Resing, J.A.C. (1993) Polling systems and multitype branching processes. *Queueing Systems* **13**, 409–426.
21. Takács, L. (1968) Two queues attended by a single server. *Oper. Res.* **16**, 639–650.
22. Takagi, H. (1986) *Analysis of Polling Systems* (MIT Press, Cambridge, MA).
23. Takagi, H. (1988) Queueing analysis of polling models. *ACM Comput. Surveys* **20**, 5–28.
24. Yechiali, U. (1993) Analysis and control of polling systems. In: *Performance Evaluation of Computer and Communication Systems*, eds. L. Donatiello and R. Nelson (Springer Verlag, Berlin), 630–650.