An Alternating-Priority Server with Correlated Switchover Times

Eitan Altman, Robin Groenevelt

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

An Alternating-Priority Server with Correlated Switchover Times

Robin Groenevelt — Eitan Altman

N° 5368

Novembre 2004

Thème COM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5368--FR+ENG

An Alternating-Priority Server with Correlated Switchover Times

Robin Groenevelt* , Eitan Altman*

Thème COM — Systèmes communicants
Projet MAESTRO

Rapport de recherche n° 5368 — Novembre 2004 — 55 pages

Abstract: This document analyzes a single server queueing system in which service is alternated between two queues and the server requires a (finite) switchover time to switch from one queue to the other. The distinction from classical results is that the sequence of switchover times from each of the queues need not be i.i.d. nor independent from each other; each sequence is merely required to form a stationary ergodic sequence. With the help of stochastic recursive equations explicit expressions are derived for a number of performance measures, most notably for the average delay of a customer and the average queue lengths under different service disciplines. With these expressions a comparison is made between the service disciplines and the influence of correlation is studied. Finally, through a number of examples it is shown that the correlation can significantly increase the mean delay and the average queue lengths indicating that the correlation between switchover times should not be ignored. This has important implications for communication systems in which a common communication channel is shared amongst various users and where the time between consecutive data transfers is correlated (for example in ad-hoc networks).

Key-words: Polling systems, queuing theory, alternating server, correlated switchover times, stochastic recursive equations, stochastic processes

* Email: {robin.groenevelt}{eitan.altman}@sophia.inria.fr

This work was partially sponsored by the EURONGI Network of Excellence.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Un Système à Polling avec Deux Files d'Attente et des Temps de Passage Corrélés

Résumé : Ce document résout un système à Polling qui consiste en deux files d'attente servies par un serveur. Après avoir servi une file d'attente, le serveur a besoin d'un temps de passage après lequel il commence à servir l'autre file d'attente. Les suites des temps de passages peuvent être corrélées. Nous obtenons l'expression de plusieurs quantités, dont notamment le temps d'attente moyen et la taille moyenne de la file d'attente. Grâce à ces expressions, nous comparons les différentes disciplines de service. Finalement, par des exemples nous montrons que la corrélation des temps de passage peut augmenter significativement le temps d'attente moyen et la taille des files d'attente, indiquant que cette corrélation ne peut pas être ignorée. Cela a des implications importantes pour des systèmes de communication dans lesquels un canal de communication commun est partagé par plusieurs utilisateurs et où le temps entre des transferts de données consécutifs est corrélé (par exemple dans les réseaux ad-hoc).

Mots-clés : Système à *Polling*, files d'attente, équations récursives stochastiques, processus stochastiques

Contents

1 Introduction and Motivation

So far only few explicit results have been known in queueing theory for systems whose evolution is described by general stationary ergodic processes. One line of research that allows one to handle stationary ergodic sequences is based on identifying measures that are insensitive to correlations. For example, the probability of finding a G/G/1 queue non-empty is just the ratio between the expected service time and the expected interarrival time of customers (which follows directly from Little's Law). The expected cycle duration in a polling system (under fairly general condition) too, depends on the interarrival, service and vacation times only through their expectations under general stationary ergodic assumptions (see e.g. [4]). An example of performance measures that depend on the whole distribution of service times but is insensitive to correlations is the growth rate of the number of customers or of the sojourn time in a (discriminatory) processor sharing queue in overload [3, 14]. Other insensitivity results on bandwidth sharing in a network can be found in [7, 6].

The polling models we study do not exhibit insensitivity. Approximating correlated vacations by independent ones can result in large errors in the performance, see e.g. [21] in the context of bluetooth. To study these systems we make use of stochastic recursions equations (SRE) introduced in [1] which extend branching processes with migration on one hand, and linear stochastic recursive equations[1] on the other. It has already been shown in [1] that vector valued SRE can be used to describe some embedded processes appearing in polling models.[2]

In this paper we identify one dimensional SRE which we use in order to compute the expected waiting times and queue lengths in a system with two queues where a single server alternates between two queues and requires switch-over times (modeled as vacations) to move from one queue to the other. We consider the exhaustive service discipline where the server serves a queue until it empties before switching to the next queue as well as the gated discipline where only customers present upon the arrival of the server are served. Two systems are studied: one in which both queues are served exhaustively and one in which one queue is served exhaustively and the other according to the gated discipline. Our analytical results are then used to study numerically the impact of correlated switchover times on the performance, as well as the difference in performance due to the service discipline used.

The polling system studied in this paper, but without the correlation, has been used in the past [17] to model communication systems in which transmission between two stations can take place only in one direction at a time. The position of the server then corresponds to the direction data is travelling in. A similar situation arrises in ad-hoc network; there is a common channel which needs to be shared amongst various users. The more users there are, the longer one has to wait before being able to capture the channel necessary to (re)transmit data. In particular, if one has to wait a long time before being able to transmit data, then it is very likely that there are many users around and that the next time one has to wait once again for a long period of time. For this reason the correlation of the

[1] Linear SRE have already been used to study the impact of correlation of the loss process on TCP throughput [5].

[2] SRE have also been used recently to study the infinite server queue with correlated arrivals [2].

number of users over time in an ad-hoc network inherently introduces correlation between the waiting (switchover) times, and this in turn leads to an increase in the mean delay and queue lengths.

The remainder of this document is structured as follows. In section 2 the polling system is described in more detail and the notation and some formulas are established. Section 3 (respectively section 4) is entirely devoted the analysis of the exhaustive/exhaustive (respectively exhaustive/gated) queueing system. This is done by first deriving a SRE in section 3.1 (4.1) which leads to the derivation of a number of performance measures in section 3.2 (4.2), of which most notably the expected waiting times and queue lengths in section 3.2.6 (4.2.7). The performance measures are used then used in the examples of section 5 to show the influence of correlated switchover times. Finally, conclusions are given in section 6. To aid the reader, a list of notations is given appendix E.

2 Model Description

We examine the polling of two queues, i.e. one queue is served after which the other queue is served. No limit is specified for the length of either queue. After serving queue i ($i = 1, 2$) for the n-th time, the server requires a switchover time of duration $V_{n,i}$. Assume all $V_{n,i}$ have the same distribution as V_i ($V_{n,i} \sim V_i$), where V_i is assumed to form a general distribution with first and second moment v_i and $v_i^{(2)}$, and with variance $\delta_i^2 := v_i^{(2)} - v_i^2$, $i = 1, 2$. Let $R := v_1 + v_2$ and $\Delta^2 := \delta_1^2 + \delta_2^2$. The sequences of switchover times are assumed to be stationary ergodic instead of the usual i.i.d., and possibly dependent on each other. This implies that there can be a correlation between the switchover times of the two queues and/or within the sequence of switchover times for each queue. The arrival of customers at queue i is Poisson with rate λ_i and the service times are nonnegative, i.i.d. random variables with (finite) first and second moments for queue i given by, respectively, b_i and $b_i^{(2)}$. The *load* at queue i is $\rho_i := \lambda_i b_i$ and the system is stable [19, page 280] if and only if the overall *load* $\rho := \rho_1 + \rho_2 < 1$, which we assume throughout. Furthermore, we will continuously assume that the queues are operating under stationary regime.

Introduce the covariance functions ($i = 1, 2$)

$$c_i(n) = \mathbb{E}[V_{0,i} V_{n,i}] - \mathbb{E}[V_{0,i}]\mathbb{E}[V_{n,i}], \qquad n \in \mathbb{N}, \tag{1}$$

$$c_{12}(n) = \mathbb{E}[V_{0,1} V_{n,2}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{n,2}], \qquad n \in \mathbb{Z}. \tag{2}$$

Note that $c_{12}(n)$ is defined for $n \in \mathbb{Z}$. With this convention it is not necessary to work with $c_{21}(n) := \mathbb{E}[V_{0,2} V_{n,1}] - \mathbb{E}[V_{0,2}]\mathbb{E}[V_{n,1}]$, since under stationary regime it follows that $c_{21}(n) = \mathbb{E}[V_{0,2} V_{n,1}] - \mathbb{E}[V_{0,2}]\mathbb{E}[V_{n,1}] = \mathbb{E}[V_{-n,2} V_{0,1}] - \mathbb{E}[V_{-n,2}]\mathbb{E}[V_{0,1}] = c_{12}(-n)$. In particular, if for each queue the sequence of switchover times is uncorrelated, then $c_i(0) = \delta_i^2$ and $c_i(n) = 0$, for $n \in \mathbb{N}$. If there is no correlation between the switchover times of the two queues, then $c_{12}(n) = 0$, for $n \in \mathbb{Z}$.

Because of the assumption of the queues operating under stationary regime (1) and (2) can be rewritten as

$$\mathbb{E}[V_{0,i}V_{n,i}] = v_i^2 + c_i(n), \quad i=1,2, \quad n \in \mathbb{N}, \tag{3a}$$
$$\mathbb{E}[V_{0,1}V_{n,2}] = v_1 v_2 + c_{12}(n), \qquad n \in \mathbb{Z}. \tag{3b}$$

In order to establish the SRE, let $\mathcal{D}_{n,i}(N)$ be the duration of the busy period in the i^{th} queue, initiated by N customers waiting in that queue when the server arrives at that queue for the n^{th} time. Similarly, let $\mathcal{N}_{n,i}(T)$ be the number of customers arriving at queue i during a period of time T during the server's n^{th} visit to queue i.

Let us now establish a number of formulas which will be used throughout. First recall that if D_i is a random sequence with $\mathbb{E}[D_i] = d$ and $\mathbb{E}[D_i^2] = d^{(2)}$, independent of a random variable N, and

$$\tau(N) := \sum_{i=1}^{N} D_i, \tag{4}$$

then

$$\begin{aligned}\mathbb{E}[\tau^2(N)] &= \sum_{n=1}^{\infty} n\mathbb{E}\left[\sum_{i=1}^{n} D_i \cdot \sum_{i=1}^{n} D_i | N=n\right] P(N=n)\\ &= \sum_{n=1}^{\infty} n\Big(\mathbb{E}[D_1^2] + (n-1)(\mathbb{E}[D_1])^2\Big) P(N=n)\\ &= d^2\mathbb{E}[N^2] + (d^{(2)} - d^2)\mathbb{E}[N].\end{aligned} \tag{5}$$

Similarly,

$$\begin{aligned}\mathbb{E}[\mathcal{N}_{n,i}^2(T)] &= \int_0^{\infty} \mathbb{E}[\mathcal{N}_{n,i}^2(t)|T=t]dT(t) = \int_0^{\infty} \Big((\lambda_i t)^2 + \lambda_i t\Big) dT(t)\\ &= \lambda_i^2\mathbb{E}[T^2] + \lambda_i\mathbb{E}[T].\end{aligned} \tag{6}$$

Next we proceed in a similar manner to obtain the second moment of the busy period generated by Y customers initially in the system. First recall that $\mathcal{D}_{n,i}(1)$ is a single busy period initiated by a single customer in an M/G/1 queue with Poisson arrivals with rate λ_i and general service time with first and second moments b_i and $b_i^{(2)}$ respectively. The first two moments of a single busy period initiated by a single customer are given by [15, equations 5.141 and 5.142]

$$d_i \quad := \quad \mathbb{E}[\mathcal{D}_{n,i}(1)] = \frac{b_i}{1-\rho_i}, \tag{7a}$$
$$d_i^{(2)} \quad := \quad \mathbb{E}[\mathcal{D}_{n,i}^2(1)] = \frac{b_i^{(2)}}{(1-\rho_i)^3}, \quad i=1,2. \tag{7b}$$

The busy period, $D_{n,i}(Y)$, generated by Y customers is the sum of Y independent single busy periods, each with distribution $D_{n,i,k} \sim D_{n,i}(1)$. Hence the second moment is

$$\mathbb{E}[\mathcal{D}_{n,i}^2(Y)] = \mathbb{E}\left[\sum_{k=1}^{Y} \mathcal{D}_{n,i,k}\right]^2 = \mathbb{E}\left[\sum_{k=1}^{Y} \mathcal{D}_{n,i}(1)\right]^2 = d_i^2 \mathbb{E}[Y^2] + \left(d_i^{(2)} - d_i^2\right)\mathbb{E}[Y]. \qquad (8)$$

By taking $Y = \mathcal{N}_{n,i}(T)$ we obtain $(i = 1, 2)$

$$\begin{aligned} \mathbb{E}[\mathcal{D}_{n,i}^2(\mathcal{N}_{n,i}(T))] =& d_i^2 \mathbb{E}[\mathcal{N}_{n,i}^2(T)] + (d_i^{(2)} - d_i^2)\mathbb{E}[\mathcal{N}_{n,i}(T)] \\ =& \lambda_i^2 d_i^2 \mathbb{E}[T^2] + d_i^{(2)} \lambda_i \mathbb{E}[T]. \end{aligned} \qquad (9)$$

3 Exhaustive/Exhaustive Service System

3.1 Introduction

We start by examining the exhaustive polling of two queues, i.e. one queue is served until it is empty after which the other queue is served until emptied. Consider the system at the moment the server starts serving the first queue for the n^{th} time with $L^*_{n,1}$ customers waiting in the queue. From here on the following steps take place (see Figure 1 for a visual representation of this decomposition):

- *Exhausting the first queue.* The $L^*_{n,1}$ customers in the first queue require a busy period duration of $D_{n,1} := \mathcal{D}_{n,1}(L^*_{n,1})$ to exhaust.
- *Switching to the second queue.* After serving the first queue the server requires a switchover time of $V_{n,1}$ units of time.
- *Exhausting the second queue.* In the time needed to switch from the second to the first queue ($V_{n-1,2}$), to exhaust the first queue ($D_{n,1}$), and to switch back to the second queue ($V_{n,1}$), there have been $L^*_{n,2} := \mathcal{N}_{n-1,2}(V_{n-1,2})+\mathcal{N}_{n,2}(\mathcal{D}_{n,1}(L^*_{n,1})+V_{n,1})$ customers arriving at the second queue. It requires $D_{n,2} := \mathcal{D}_{n,2}(L^*_{n,2})$ units of time to empty this queue.
- *Switching back to the first queue.* After serving the second queue the server requires a switchover time of $V_{n,2}$ units of time.

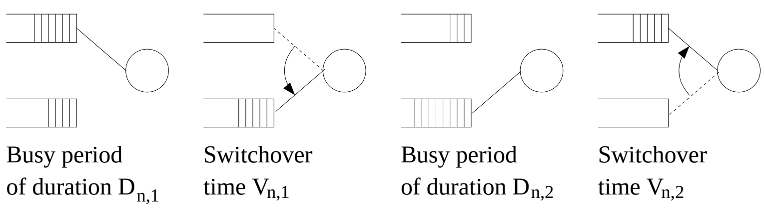

Figure 1: Decomposition of the n^{th} cycle into busy periods ($D_{n,1}$ and $D_{n,2}$) and switchover times ($V_{n,1}$ and $V_{n,2}$).

After this the process starts over again and a new *cycle* begins. Hence the n^{th} cycle is made up of

$$C_n = D_{n,1} + V_{n,1} + D_{n,2} + V_{n,2}.$$

The time between the server finishing work at queue i and returning to queue i in the next cycle is the *intervisit* time $I_{n,i}$ and is given by

$$I_{n,1} = V_{n,1} + D_{n,2} + V_{n,2}, \tag{10a}$$
$$I_{n,2} = V_{n,2} + D_{n+1,1} + V_{n+1,1}. \tag{10b}$$

A SRE will be established for this quantity and we will see that it plays a central role for the derivation of the expected waiting times and queue lengths. The time $D_{n+1,i}$ spent at queue i in the $(n+1)^{th}$ cycle is related to the intervisit time according to

$$D_{n+1,1} = \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(I_{n,1})), \tag{11a}$$
$$D_{n+1,2} = \mathcal{D}_{n+1,2}\left[\mathcal{N}_{n,2}(V_{n,2})+\mathcal{N}_{n+1,2}\left(V_{n+1,1} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(I_{n,1}))\right)\right] \tag{11b}$$

The expectation is the sum of expected busy periods [15, p.217] and thus

$$\mathbb{E}[D_{n,1}] = \frac{b_1\mathbb{E}[\mathcal{N}_{n,1}(I_{n,1})]}{1-\rho_1} = \frac{\rho_1\mathbb{E}[I_{n,1}]}{1-\rho_1}. \tag{12a}$$

Using the stationarity and the divisibility ($\mathcal{N}_{n,2}(a+b) = \mathcal{N}_{n,2}(a)+\mathcal{N}_{n,2}(b)$) of the arrival process it can be shown that

$$\mathbb{E}[D_{n,2}] = \frac{\rho_2\mathbb{E}[I_{n,2}]}{1-\rho_2}. \tag{12b}$$

Since the busy periods are sums of service times, the divisibility property also holds for $\mathcal{D}_{n,i}$. This means that from (substitute 11b into 10a)

$$I_{n+1,1} = V_{n+1,1} + V_{n+1,2} + \mathcal{D}_{n+1,2}\Big[\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}\Big(V_{n+1,1} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(I_{n,1}))\Big)\Big],$$

$n \in \mathbb{N}$, we see a SRE (as presented and solved for stationary ergodic sequences in [1]) arising. Although the system is two dimensional (as there are two queues), the reduction to a one dimensional SRE is a key element in obtaining explicit formulas for the performance measures.

Theorem 3.1. (SRE for exhaustive/exhaustive system). *The intervisit time of the first queue allows itself to be written as a one-dimensional SRE,*

$$I_{n+1,1} = \mathcal{A}_n(I_{n,1}) + \mathcal{B}_n \tag{13}$$

with

$$\mathcal{A}_n(\cdot) := \mathcal{D}_{n+1,2}\Big(\mathcal{N}_{n+1,2}\Big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(\cdot))\Big)\Big), \tag{14}$$

$$\mathcal{B}_n := V_{n+1,1} + V_{n+1,2} + \mathcal{D}_{n+1,2}\Big(\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}(V_{n+1,1})\Big). \tag{15}$$

Note that from (12a) and (12b) we have $\mathbb{E}[\mathcal{A}_n(I_{n,1})] = \alpha\mathbb{E}[I_{n,1}]$ *where* $\alpha := \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}$. ◇

The existence of a stationary ergodic $I_{n,1}$ which satisfies (13) is given in the following theorem which is proven in Appendix A.

Theorem 3.2. *If $\rho < 1$ then there exists a stationary ergodic regime $I^*_{n,1}$ which satisfies (13).* $\diamond$

3.2 Performance Measures

In the next couple of sections the first two moments of various performance measures will be derived. Starting with the first two moments of the intervisit time of the first queue, we see that all of the other quantities can easily be expressed in terms of the intervisit time. Most importantly, closed form expressions for the average waiting time and the mean queue length will be derived in the presence of correlated switchover times.

3.2.1 Intervisit Time

Taking the expectation of the square of (13) gives

$$\mathbb{E}[I^2_{n+1,1}] = \mathbb{E}[\mathcal{A}_n(I_{n,1})]^2 + \mathbb{E}[\mathcal{B}^2_n] + 2\mathbb{E}[\mathcal{A}_n(I_{n,1})\mathcal{B}_n]. \tag{16}$$

This expression is central to the proof of the following theorem.

Theorem 3.3. (Intervisit time in exhaustive/exhaustive system). *Under the stationary regime the expected intervisit time of queue i is given by*

$$E[I_{n,i}] = \frac{R(1-\rho_i)}{1-\rho}, \qquad \rho := \rho_1 + \rho_2, \qquad i = 1, 2. \tag{17}$$

The second moment is given by

$$\beta\mathbb{E}[I^2_{n,1}] = \frac{R}{1-\rho}\left(\frac{\lambda_1\rho_2^2 b_1^{(2)}}{(1-\rho_1)^2} + \lambda_2 b_2^{(2)}\right) + \delta_1^2 + \left(1 - \frac{2\rho_2(1-\rho)}{1-\rho_1}\right)\delta_2^2 + \left(\frac{1-\rho+2\rho_1\rho_2}{1-\rho}\right)R^2 + 2C \tag{18}$$

where

$$\alpha := \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}, \qquad \beta := \frac{(1-\rho)(1-\rho+2\rho_1\rho_2)}{(1-\rho_1)^2}.$$

and

$$C := \sum_{j=1}^{\infty}\left[c_1(j) + c_2(j) + \frac{(1-\rho)^2}{\rho_1(1-\rho_1)}c_2(j) + \frac{1-\rho_2}{\rho_1}c_{12}(-j) + \frac{1-\rho_2(1-\alpha)}{\alpha}c_{12}(j-1)\right]\alpha^j \tag{19}$$

is the addition to the intervisit time due to the correlation between the switchover times. $\diamond$

Remark: By taking $\lambda_2 = 0$, one obtains an ordinary M/G/1 queue with correlated vacations where $\mathbb{E}[I_{n,1}^2]$ $(= v_1^{(2)} + v_2^{(2)} + 2v_1v_2)$ no longer depends on the correlation (since $\alpha = 0$).

The proof is given in Appendix B. On the basis of this theorem a number of other results quickly follow.

Corollary 3.1. (Intervisit Time With Uncorrelated Switchover Times).
If the succesive switchover times in each queue are uncorrelated then $c_1(j) = c_2(j) = 0$ *for* $j \geq 1$. *Furthermore, if the switchover times between the two queues are uncorrelated, then* $c_{12}(j) = 0$ *for* $j \in \mathbb{Z}$. *This leads to* $C = 0$. *The expression then agrees with the classical result mentioned in [11, equation (55)] or in [18, equation*[3] *(4.36a)].* ◇

3.2.2 Number of Customers Waiting When Server Arrives

The exhaustive nature of the server implies that the number of customers building up at queue i is exactly the number of customers that arrived at that queue during its intervisit time. Thus,

$$L_{n+1,1}^* = \mathcal{N}_{n,1}(I_{n,1}).$$

From this we immediately obtain

$$\mathbb{E}[L_{n+1,i}^*] = \lambda_i \mathbb{E}[I_{n,i}] = \frac{R\lambda_i(1-\rho_i)}{1-\rho} \tag{20}$$

as the expected length of the queue, under stationary regime, at the moment the server arrives at queue i $(i = 1, 2)$. The second moment follows through squaring,

$$\mathbb{E}[(L_{n+1,1}^*)^2] = \mathbb{E}[\mathcal{N}_{n,1}^2(I_{n,1})] = \lambda_1^2 \mathbb{E}[I_{n,1}^2] + \lambda_1 \mathbb{E}[I_{n,1}],$$

which leads to

$$\mathbb{E}[(L_{n,i}^*)^2] := \lambda_i^2 \mathbb{E}[I_{n,i}^2] + \frac{R\lambda_i(1-\rho_i)}{1-\rho}. \tag{21}$$

3.2.3 Duration of Busy Periods

The expected time per cycle, in steady state, for the server to work on queue i $(i = 1, 2)$ is given by

$$\mathbb{E}[D_{n,i}] = \frac{R\rho_i}{1-\rho}. \tag{22}$$

[3]Instead of switchover times this work refers to reply intervals. Unfortunately, the definition or the indices in (4.36a) are not coherent for the case $N = 2$ as the indices of ones and two have to be interchanged (for δ_1 and δ_2, as well as for r_1 and r_2) to obtain the correct result. The result is mentioned correctly in [11, equation (55)].

This follows directly from $\mathbb{E}[D_{n,i}] = \mathbb{E}[\mathcal{D}_{n,i}(L^*_{n,i})] = \frac{b_i \mathbb{E}[L^*_{n,i}]}{1-\rho_i}$.

Since $\mathbb{E}[D^2_{n,1}] = \mathbb{E}[\mathcal{D}^2_{n,1}(\mathcal{N}_{n,1}(I_{n,1}))]$, the second moment follows with (9) and is given by

$$\mathbb{E}[D^2_{n,1}] = \frac{\rho_1^2 \mathbb{E}[I^2_{n,1}]}{(1-\rho_1)^2} + \frac{R\lambda_1 b_1^{(2)}}{(1-\rho_1)^2(1-\rho)}. \tag{23}$$

3.2.4 Cycle time

By definition $C_{n+1,i} = D_{n+1,i} + I_{n+1,i}$. From this the expected duration of a cycle under stationary regime immediately follows:

$$\mathbb{E}[C_{n+1,i}] = \frac{R}{1-\rho}, \qquad i = 1, 2. \tag{24}$$

Taking the expectation over the square produces

$$\mathbb{E}[C^2_{n+1,1}] = \mathbb{E}[D^2_{n+1,1}] + \mathbb{E}[I^2_{n+1,1}] + 2\mathbb{E}[D_{n+1,1} I_{n+1,1}].$$

The last term on the right hand side can be derived using (13) to give

$$\begin{aligned}
\mathbb{E}[D_{n+1,1} I_{n+1,1}] &= \mathbb{E}[D_{n+1,1} \cdot (\mathcal{A}_n(I_{n,1}) + \mathcal{B}_n)] = \mathbb{E}[D_{n+1,1} \cdot \mathcal{A}_n(I_{n,1})] + \mathbb{E}[D_{n+1,1} \cdot \mathcal{B}_n] \\
&= \mathbb{E}\big[D_{n+1,1} \cdot \mathcal{D}_{n+1,2}\big(\mathcal{N}_{n+1,2}(D_{n+1,1})\big)\big] + \mathbb{E}\big[\mathcal{D}_{n+1,1}\big(\mathcal{N}_{n,1}(I_{n,1}) \cdot \mathcal{B}_n\big)\big] \\
&= \frac{\rho_2}{1-\rho_2} \mathbb{E}[D^2_{n+1,1}] + \frac{\rho_1}{1-\rho_1} \mathbb{E}[I_{n,1} \mathcal{B}_n].
\end{aligned}$$

Plugging in the expressions for $\mathbb{E}[I_{n,1}\mathcal{B}_n]$ (equations (72) and (75)) reveals that

$$\begin{aligned}
\mathbb{E}[C^2_{n+1,1}] &= \mathbb{E}[D^2_{n+1,1}] + \mathbb{E}[I^2_{n+1,1}] + \frac{2\rho_2}{1-\rho_2} \mathbb{E}[D^2_{n+1,1}] + \frac{2\rho_1}{1-\rho_1} \mathbb{E}[I_{n,1}\mathcal{B}_n] \\
&= \left(\frac{1+\rho_2}{1-\rho_2}\right) \mathbb{E}[D^2_{n+1,1}] + \mathbb{E}[I^2_{n,1}] + \frac{2\rho_1}{1-\rho_2} \left(\frac{R^2}{1-\rho} + \frac{\rho_2 \delta_2^2}{1-\rho_1}\right) \\
&\quad + \frac{2}{\rho_2} \sum_{j=1}^{\infty} \alpha^j \left[\frac{c_1(j) + c_2(j)}{1-\rho_2} + \frac{(1-\rho)^2 c_2(j)}{\rho_1(1-\rho_1)(1-\rho_2)} + \frac{c_{12}(-j)}{\rho_1} + c_{12}(j)\Big(1 + \frac{\alpha\rho_2}{1-\rho_2}\Big)\right] \\
&\quad - 2c_2(1) + \frac{2(-c_{12}(-1) + \alpha c_{12}(0))}{1-\rho_2}.
\end{aligned}$$

The expression for $\mathbb{E}[D^2_{n+1,1}] = \mathbb{E}[D^2_{n,1}]$ is given in equation (23) and leads to

$$\begin{aligned}
\mathbb{E}[C^2_{n+1,1}] &= \left(\left(\frac{1+\rho_2}{1-\rho_2}\right)\left(\frac{\rho_1^2}{(1-\rho_1)^2}\right) + 1\right) \mathbb{E}[I^2_{n+1,1}] + \frac{2\rho_1}{1-\rho_2} \left(\frac{R^2}{1-\rho} + \frac{\rho_2 \delta_2^2}{1-\rho_1}\right) \\
&\quad + \frac{2}{\rho_2} \sum_{j=1}^{\infty} \alpha^j \left[\frac{c_1(j) + c_2(j)}{1-\rho_2} + \frac{(1-\rho)^2 c_2(j)}{\rho_1(1-\rho_1)(1-\rho_2)} + \frac{c_{12}(-j)}{\rho_1} + c_{12}(j)\Big(1 + \frac{\alpha\rho_2}{1-\rho_2}\Big)\right] \\
&\quad - 2c_2(1) + \frac{2(-c_{12}(-1) + \alpha c_{12}(0))}{1-\rho_2}.
\end{aligned}$$

Remark: A cycle can also be defined as $C_{n,1} = I_{n,1} + D_{n+1,1}$. The first moment of the cycle time is then the same. However, the second moment is then different since

$$\begin{aligned}\mathbb{E}[C_{n,1}^2] =& \mathbb{E}[I_{n,1}^2] + \mathbb{E}[D_{n+1,1}^2] + 2\mathbb{E}[I_{n,1} \cdot \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(I_{n,1}))] \\ =& \mathbb{E}[I_{n,1}^2] + \lambda_1^2 d_1^2 \mathbb{E}[I_{n,1}^2] + \lambda_1 d_1^{(2)} \mathbb{E}[I_{n,1}] + 2\lambda_1 d_1 \mathbb{E}[I_{n,1}^2] \\ =& \frac{\mathbb{E}[I_{n,1}^2]}{(1-\rho_1)^2} + \frac{\lambda_1 b_1^{(2)} \mathbb{E}[I_{n,1}]}{(1-\rho_1)^3}.\end{aligned}$$

This last expression corresponds to [18, equation[4] (4.23b)]. ∎

3.2.5 Number Served per Cycle

To derive the first and second moments of the number of customers served per cycle, consider an $M/G/1$ queue with arrival rate λ_i, average service time b_i, and the second moment of the service time $b_i^{(2)}$. Then the expectation [15, equation (5.153)] and the variance [15, equation (5.154)] of the number of customers served in a single busy period are

$$\mathbb{E}[\Gamma_i] = \frac{1}{1-\rho_i}, \qquad \mathrm{Var}[\Gamma_i] = \frac{\rho_i(1-\rho_i) + \lambda_i^2 b_i^{(2)}}{(1-\rho_i)^3}.$$

Let $\mathcal{T}_{n,i}(N)$ be the number of customers served at queue i during the n^{th} cycle if there are N customers in the queue at the moment of polling. Since the number served is the sum of the number served during N busy periods, (4) tells us that the expected number of customers served, per cycle, at queue i is

$$\mathbb{E}[T_{n,i}] = \mathbb{E}[\mathcal{T}_{n,i}(L_{n,i}^*)] = \frac{\mathbb{E}[L_{n,i}^*]}{1-\rho_i} = \frac{\lambda_i R}{1-\rho}, \tag{25}$$

for $i = 1, 2$. To derive the second moment note that

$$\begin{aligned}\mathbb{E}[T_{n,i}^2] =& \mathbb{E}[\mathcal{T}_{n,i}^2(L_{n,i}^*)] \\ =& (\mathbb{E}[\Gamma_i])^2 \mathbb{E}[(L_{n,i}^*)^2] + \mathrm{Var}[\Gamma_{n,i}] \mathbb{E}[L_{n,i}^*] \\ =& \frac{\mathbb{E}[(L_{n,i}^*)^2]}{(1-\rho_i)^2} + \frac{R\lambda_i\left(\rho_i(1-\rho_i) + \lambda_i^2 b_i^{(2)}\right)}{(1-\rho_i)^2(1-\rho)}.\end{aligned}$$

Using equation (21) gives

$$\mathbb{E}[T_{n,i}^2] = \frac{1}{(1-\rho_i)^2}\left(\lambda_i^2 \mathbb{E}[I_{n,i}^2] + \frac{R\lambda_i\left(1-\rho_i^2+\lambda_i^2 b_i^{(2)}\right)}{1-\rho}\right) \qquad i = 1, 2. \tag{26}$$

[4] There is in fact a factor λ_i missing in [18, formula (4.23b)]. It should read $\ldots + \frac{\lambda_i[-(1-\rho_i)+\lambda_i^2 b_i^{(2)}]\sum_{k=1}^N r_k}{1-\sum_{k=1}^N \rho_k}$.

3.2.6 Expected Waiting Time and Average Queue Length

Two of the most important performance measures, the expected waiting time and the average queue length, still remain to be given. The following theorem provides us with this result.

Theorem 3.4. (Expected waiting time and queue length for the exhaustive/exhaustive service discipline). *The expected waiting time (total time in system minus service time) of a customer going through queue i is decomposed of two parts, namely*

$$\mathbb{E}[W_{q,i}] = \frac{\lambda_i b_i^{(2)}}{2(1-\rho_i)} + \frac{(1-\rho)\mathbb{E}[I_{n,i}^2]}{2R(1-\rho_i)}, \tag{27}$$

$i = 1, 2$. *This gives*

$$\begin{aligned}\mathbb{E}[W_{q,1}] = {} & \frac{\lambda_1 b_1^{(2)}}{2(1-\rho_1)} + \frac{\lambda_1 \rho_2^2 b_1^{(2)} + \lambda_2 (1-\rho_1)^2 b_2^{(2)}}{2(1-\rho_1)(1-\rho)(1-\rho+2\rho_1\rho_2)} \\ & + \frac{(1-\rho_1)R}{2(1-\rho)} + \left(\frac{\Delta^2}{2} - \frac{\rho_2(1-\rho)}{1-\rho_1}\delta_2^2 + C\right)\psi,\end{aligned} \tag{28}$$

where $\psi := \frac{1-\rho_1}{R(1-\rho+2\rho_1\rho_2)}$ *and* C *(defined in (19)) is the increase in the expected waiting time due to correlated switchover times.*

The average number of customers at queue i (in service and in the queue) follows directly from Little and is

$$\mathbb{E}[L_{s,i}] = \lambda_i \mathbb{E}[W_{q,i}] + \rho_i, \qquad i = 1, 2. \qquad \diamond$$

In the uncorrelated case ($C = 0$) this is in correspondence with [19, formula[5] (3.12)].

Proof: For a customer arriving at queue i ($i = 1, 2$) the system behaves as an M/G/1 queue where the server goes on vacation as soon as the queue is empty. The random variable for the n^{th} "vacation" from the first queue is exactly $I_{n,i}$. Conditioning the waiting time in the queue on whether or not a customer arrives when the server is busy or on vacation produces

$$\mathbb{E}[W_{q,i}] = \frac{\mathbb{E}[I_{n,i}]}{\mathbb{E}[C_{n,i}]}\mathbb{E}[W_{q,i}|vac] + \frac{\mathbb{E}[D_{n,i}]}{\mathbb{E}[C_{n,i}]}\mathbb{E}[W_{q,i}|busy]. \tag{29}$$

A tagged customer that arrives during a vacation has to wait for the vacation to finish plus the time needed to serve the customers that arrived before him/her in the vacation. The expected remaining vacation time is $\mathbb{E}[I_{n,i}^2]/2\mathbb{E}[I_{n,i}]$ and the expected number of customer that arrived before the tagged customer is $\lambda_i \mathbb{E}[I_{n,i}^2]/2\mathbb{E}[I_{n,i}]$. This means that

$$\mathbb{E}[W_{q,i}|vac] = \frac{\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}(1 + \lambda_i b_i). \tag{30}$$

[5] The last term in [19, formula (3.12)] is missing a factor two and there is a mix up between δ_1 and δ_2. The expression should read $\cdots + \frac{[1-\rho_1-2\rho_2(1-\rho)]\delta_2^2 + (1-\rho_1)^2\delta_1^2}{2R(1-\rho+2\rho_1\rho_2)}$.

A tagged customer that arrives when the server is busy has to wait for the current customer in service to finish plus the expected time needed to serve the $L_{q,i}$ customers that arrived at (and still are in) the queue before the tagged customer did. This gives

$$\mathbb{E}[W_{q,i}|busy] = \frac{b_i^{(2)}}{2b_i} + b_i\mathbb{E}[L_{q,i}|busy]. \tag{31}$$

To obtain the number of customers in the queue, first realize that the expected waiting time of a customer in the system is $\mathbb{E}[W_{s,i}] = b_i + \mathbb{E}[W_{q,i}]$. Little [13] tells us that the expected number of customers, $L_{s,i}$, at the queue i (in service and in the queue) is

$$\mathbb{E}[L_{s,i}] := \lambda_i\mathbb{E}[W_{s,i}] = \rho_i + \lambda_i\mathbb{E}[W_{q,i}].$$

On the other hand,

$$\begin{aligned}\mathbb{E}[L_{s,i}] =& \frac{\mathbb{E}[D_{n,i}]}{\mathbb{E}[C_{n,i}]}\mathbb{E}[L_{s,i}|busy] + \frac{\mathbb{E}[I_{n,i}]}{\mathbb{E}[C_{n,i}]}\mathbb{E}[L_{s,i}|vac]\\ =&\rho_i(1+\mathbb{E}[L_{q,i}|busy]) + (1-\rho_i)\frac{\lambda_i\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}.\end{aligned}$$

Combining these last two equations gives

$$\mathbb{E}[L_{q,i}|busy] =\frac{1}{b_i}\left(\mathbb{E}[W_{q,i}] - \frac{(1-\rho_i)\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}\right). \tag{32}$$

Putting together equations (29)-(32) gives

$$\begin{aligned}\mathbb{E}[W_{q,i}] =&(1-\rho_i)\frac{\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}(1+\lambda_i b_i) + \rho_i\left(\frac{b_i^{(2)}}{2b_i} + \mathbb{E}[W_{q,i}] - \frac{(1-\rho_i)\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}\right)\\ =&\frac{\lambda_i b_i^{(2)}}{2(1-\rho_i)} + \frac{\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}. \end{aligned} \tag{33}$$

The theorem follows by plugging in the values of $\mathbb{E}[I_{n,i}]$ and $\mathbb{E}[I_{n,i}^2]$. Notice that not once have we assumed the switchover times and the busy periods to be uncorrelated!

An alternative proof is given in Appendix C. ∎

3.3 Identical queues

Let both queues be identically distributed, i.e. $\lambda := \lambda_i$, $c(j) := c_i(j)$, $b^{(2)} := b_i^{(2)}$, $v := v_i$, $\delta^2 := \delta_i^2$, and $\hat{\rho} := \rho_i$, for $i = 1, 2$. The stability condition is then $\hat{\rho} = \lambda d < 1/2$. In stationary regime,

$$\mathbb{E}[I_{n,1}] = \frac{2v(1-\hat{\rho})}{1-2\hat{\rho}} \qquad \text{(Intervisit time)}$$

$$\mathbb{E}[L^*_{n,1}] = \frac{2v\lambda(1-\hat{\rho})}{1-2\hat{\rho}} \qquad \text{(Queue length when server arrives)}$$

$$\mathbb{E}[D_{n,1}] = \frac{2v\hat{\rho}}{1-2\hat{\rho}} \qquad \text{(Busy period duration)}$$

$$\mathbb{E}[C_{n,1}] = \frac{2v}{1-2\hat{\rho}} \qquad \text{(Cycle time)}$$

$$\mathbb{E}[T_{n,1}] = \frac{2v\lambda}{1-2\hat{\rho}} \qquad \text{(Number served per cycle)}$$

The second moments are

$$\begin{aligned}
E[I^2_{n,1}] &= 2v\left(\frac{\lambda b^{(2)} + 2v(1-\hat{\rho})^2}{(1-2\hat{\rho})^2}\right) + \frac{2(1-\hat{\rho})\delta^2}{1-2\hat{\rho}} \\
&\quad + \frac{2(1-3\hat{\rho}+3\hat{\rho}^2)}{\beta(1-\hat{\rho})^2}\sum_{j=0}^{\infty}\alpha^j c_{12}(j) + \frac{2(1-\hat{\rho})}{\hat{\rho}}\sum_{j=1}^{\infty}\left[\frac{c(j)}{1-2\hat{\rho}} + \frac{1}{\beta}c_{12}(-j)\right]\alpha^j \\
\mathbb{E}[(L^*_{n,1})^2] &= \lambda^2\mathbb{E}[I^2_{n,1}] + \frac{2v\lambda(1-\hat{\rho})}{1-2\hat{\rho}} \\
\mathbb{E}[D^2_{n,1}] &= \frac{1}{(1-\hat{\rho})^2}\left(\hat{\rho}^2\mathbb{E}[I^2_{n,1}] + \frac{2v\lambda b^{(2)}}{1-2\hat{\rho}}\right) \\
\mathbb{E}[C^2_{n,1}] &= \frac{\mathbb{E}[I^2_{n,1}]}{(1-\hat{\rho})^2} + \frac{2v\lambda b^{(2)}}{(1-\hat{\rho})^2(1-2\hat{\rho})} \\
\mathbb{E}[T^2_{n,1}] &= \frac{1}{(1-\hat{\rho})^2}\left(\lambda^2\mathbb{E}[I^2_{n,1}] + \frac{2v\lambda(1-\hat{\rho}^2+\lambda^2 b^{(2)})}{1-2\hat{\rho}}\right)
\end{aligned}$$

with

$$\alpha = \frac{\hat{\rho}^2}{(1-\hat{\rho})^2}, \qquad \beta = \frac{(1-\hat{\rho})^4-\hat{\rho}^4}{(1-\hat{\rho})^2} = \frac{(1-2\hat{\rho})(1-2\hat{\rho}+2\hat{\rho}^2)}{(1-\hat{\rho})^2}.$$

The expected waiting time and average queue length are

$$\begin{aligned}
\mathbb{E}[W_{q,1}] &= \frac{\lambda b^{(2)} + (1-\hat{\rho})v}{1-2\hat{\rho}} + \frac{\delta^2}{2v} \qquad \text{(Waiting time)} \\
&\quad + \frac{1-3\hat{\rho}+3\hat{\rho}^2}{v(1-\hat{\rho})(1-2\hat{\rho}+2\hat{\rho}^2)}\sum_{j=0}^{\infty}\alpha^j c_{12}(j) + \frac{1-2\hat{\rho}}{2v\hat{\rho}}\sum_{j=1}^{\infty}\left[\frac{c(j)}{1-2\hat{\rho}} + \frac{1}{\beta}c_{12}(-j)\right]\alpha^j \\
\mathbb{E}[L_{s,1}] &= \lambda\mathbb{E}[W_{q,1}] + \hat{\rho} \qquad \text{(Average queue length)}
\end{aligned}$$

In the case of uncorrelated switchover times these expressions coincide with the known results for the intervisit time (equation (4.21) in [18]), number of customers at polling instant (equation (4.12b) in [18]), duration of a busy period (4.15b), cycle time (4.24), number of customers served per cycle (4.18b), waiting time (4.33b), and the average queue length (4.34).

3.4 Heavily Unbalanced Traffic

What happens to the waiting times if we let the traffic at the second queue approach zero? By letting $\lambda_2 \to 0$ in (28) we see that the expected waiting times,

$$\mathbb{E}[W_{q,1}] = \frac{\lambda_1 b_1^{(2)}}{2(1-\rho_1)} + \frac{R}{2} + \frac{\Delta^2}{2R},$$

$$\mathbb{E}[W_{q,2}] = \frac{\lambda_1 b_1^{(2)}}{2(1-\rho_1)^2} + \frac{R}{2(1-\rho_1)} + \left(\frac{\Delta^2}{2} - \rho_1(1-\rho_1)\delta_2^2\right)\frac{1}{R(1-\rho_1)},$$

no longer depends on the correlation. Because of this the conclusions of [18, page 83] hold which state that if $\rho_1 \geq 0.5$, or if all switching times of the second queue are constant ($\delta_2 = 0$), we have $\mathbb{E}[W_{q,2}] \geq \mathbb{E}[W_{q,1}]$.

3.5 Switchover Times Equal to Zero

Let us examine what happens to the expected waiting time when the switchover times are constant and equal to zero. By looking at the expressions for the expected waiting times (equation (28)), we see that the variances and (cross) correlations of the switchover times must go faster to zero than the mean switchover times does since otherwise the expected waiting times could explode (for example, $\lim_{R\to 0} \delta_2^2/R$ must go to zero). Sending the appropriate variables to zero in the right order produces

$$\mathbb{E}[W_{q,1}] = \frac{\lambda_1 b_1^{(2)}}{2(1-\rho_1)} + \frac{\lambda_1 \rho_2^2 b_1^{(2)} + \lambda_2(1-\rho_1)^2 b_2^{(2)}}{2(1-\rho_1)(1-\rho)(1-\rho+2\rho_1\rho_2)}.$$

This is naturally no longer dependent on the correlation and is identical to (4.36b) in [18].

4 Exhaustive/Gated Service System

4.1 Introduction

Now let the first queue be served exhaustively and the second be served in a gated manner. The time needed to serve N customers in the second queue in the n^{th} cycle is denoted by

$\mathcal{S}_{n,2}(N)$. Naturally, $\mathbb{E}[\mathcal{S}_{n,2}(N)] = b_i\mathbb{E}[N]$. The service time of the second queue, $S_{n,2}$, satisfies the following recursive relationship

$$S_{n+1,2} = \mathcal{S}_{n+1,2}\big(\mathcal{N}_{n,2}(S_{n,2} + V_{n,2}) + \mathcal{N}_{n+1,2}(D_{n+1,1} + V_{n+1,1})\big).$$

At the same time, the time the server works per cycle at the first queue, $D_{n,1}$, satisfies

$$D_{n+1,1} = \mathcal{D}_{n+1,1}\Big(\mathcal{N}_{n,1}(V_{n,1} + S_{n,2} + V_{n,2})\Big).$$

By combining these two expression we obtain the following theorems.

Theorem 4.1. (SRE for the exhaustive/gated system). *The SRE for the service time at the gated queue is given by*

$$S_{n+1,2} = \mathcal{X}_n(S_{n,2}) + \mathcal{Y}_n, \tag{34}$$

with

$$\mathcal{X}_n(\cdot) := \mathcal{S}_{n+1,2}\Big(\mathcal{N}_{n,2}(\,\cdot\,) + \mathcal{N}_{n+1,2}\big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(\cdot))\big)\Big), \tag{35}$$

$$\mathcal{Y}_n := \mathcal{S}_{n+1,2}\Big(\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}\big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1} + V_{n,2})) + V_{n+1,1}\big)\Big). \tag{36}$$

Note that $\mathbb{E}[\mathcal{X}_n(S_{n,2})] = \gamma\mathbb{E}[S_{n,2}]$ *where* $\gamma := \frac{\rho_2}{1-\rho_1}$. ◇

Theorem 4.2. *If* $\rho < 1$ *then there exists a stationary ergodic regime* $S^*_{n,2}$ *which satisfies (34).* ◇

The proof of the last theorem is omitted due to the strong similarity with the proof of Theorem 3.2.

4.2 Performance Measures

Starting with the service time at the second queue, the next couple of sections will present a number of performance measures of the mixed exhaustive/gated service system.

4.2.1 Service Time Second Queue

Theorem 4.3 (Service Time Second Queue). *Under the stationary regime the expected time per cycle spend on service time at the second queue is given by*

$$\mathbb{E}[S_{n+1,2}] = \frac{\rho_2 R}{1-\rho}. \tag{37}$$

The second moment of the same variable is

$$\omega \mathbb{E}[S_{n+1,2}^2] = \frac{R}{1-\rho}\left(\lambda_1 b_1^{(2)} + \frac{\lambda_2(1-\rho_1)^2 b_2^{(2)}}{\rho_2^2}\right) + (1-2\rho_1(1-\rho))\,\delta_1^2 + \delta_2^2 + \left(\frac{1-\rho_1+\rho_2}{1-\rho}\right) R^2$$
$$+ 2\rho \sum_{j=0}^{\infty} c_{12}(j)\gamma^j + 2\sum_{j=1}^{\infty}\left(c_2(j) + \frac{1-\rho_1(2-\rho)}{\rho_2}\Big(c_{12}(-j)+\rho c_1(j)\Big)\right)\gamma^j. \tag{38}$$

with

$$\gamma := \frac{\rho_2}{1-\rho_1}, \qquad \omega := \frac{(1-\rho_1)^2 - \rho_2^2}{\rho_2^2}. \qquad \diamond$$

The proof forwarded to Appendix D

4.2.2 Cycle Time Starting from the Second Queue

The cycle time, $C_{n,2}$, starting from the server arriving at the second queue can be derived by using the relationship

$$S_{n+1,2} = \mathcal{S}_{n+1,2}\big(\dot{\mathcal{N}}_{n,2}(C_{n,2})\big),$$

where $\dot{\mathcal{N}}_{n,2}(\cdot)$ expresses[6] the number of arrivals at the second queue during the n^{th} cycle, with the cycle starting from the arrival of the server at the second queue. The time the server works, per cycle, on the second queue is equal to the time needed to serve the customers that arrived at the second queue during the previous cycle. Taking the expectation ($\mathbb{E}[S_{n+1,2}] = \rho_2 \mathbb{E}[C_{n,2}]$) gives the first moment,

$$\mathbb{E}[C_{n,2}] = \frac{R}{1-\rho}. \tag{39}$$

By taking the expectation of the square one obtains

$$\mathbb{E}[S_{n+1,2}^2] = \mathbb{E}[\mathcal{S}_{n+1,2}^2(\dot{\mathcal{N}}_2(C_{n,2}))] = \rho_2^2 \mathbb{E}[C_{n,2}^2] + \lambda_2 b_2^{(2)} \mathbb{E}[C_{n,2}],$$

where (76) was used. Substituting (38) into this gives the second moment of the cycle time starting from the polling instant of the second queue:

$$\eta \mathbb{E}[C_{n,2}^2] = \frac{R}{1-\rho}\left(\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}\right) + \big(1-2\rho_1(1-\rho)\big)\delta_1^2 + \delta_2^2 + \left(\frac{1-\rho_1+\rho_2}{1-\rho}\right) R^2$$
$$+ 2\rho \sum_{j=0}^{\infty} c_{12}(j)\gamma^j + 2\sum_{j=1}^{\infty}\left(c_2(j) + \frac{1-\rho_1(2-\rho)}{\rho_2}\Big(c_{12}(-j) + \rho c_1(j)\Big)\right)\gamma^j. \tag{40}$$

where $\gamma := \frac{\rho_2}{1-\rho_1}$ and $\eta := (1-\rho_1)^2 - \rho_2^2$.

[6]This represents a shift from the original arrival process $\mathcal{N}_{n,2}(\cdot)$. By doing so the calculations become significantly less involved. Note that this process is correlated to $\mathcal{N}_{n,2}$ and $\mathcal{N}_{n+1,2}$, but it is independent of $\mathcal{D}_{n,1}(\cdot)$ and $\mathcal{S}_{n,2}(\cdot)$.

4.2.3 Intervisit Time First Queue

The time between the server leaving the first queue and arriving back at the first queue is the intervisit time and it is given by

$$I_{n,1} = V_{n,1} + S_{n,2} + V_{n,2}.$$

It's expectation is straightforward,

$$\mathbb{E}[I_{n,1}] = \frac{(1-\rho_1)R}{1-\rho}. \tag{41}$$

For the second moment a little more work is needed as

$$\mathbb{E}[I_{n,1}^2] = \mathbb{E}[(V_{n,1} + S_{n,2} + V_{n,2})^2] = \delta_1^2 + \delta_2^2 + R^2 + 2\mathbb{E}[S_{n,2}(V_{n,1} + V_{n,2})] + \mathbb{E}[S_{n,2}^2]. \tag{42}$$

The second last term remains to be assessed. Similar to the step taking (80) to (81), Theorem 2 of [1] provides us with the answer. Applying the theorem just as in (81) gives

$$\begin{aligned}
\mathbb{E}[S_{n,2}(V_{n,1}+V_{n,2})] &= \sum_{j=0}^{\infty} \gamma^j \mathbb{E}[\mathcal{Y}_0(V_{j+1,1} + V_{j+1,2})] \\
&= \sum_{j=0}^{\infty} \gamma^j \mathbb{E}\Big[\mathcal{S}_{1,2}\Big(\mathcal{N}_{0,2}(V_{0,2}) + \mathcal{N}_{1,2}\big(\mathcal{D}_{1,1}\big(\mathcal{N}_{0,1}(V_{0,1}+V_{0,2})\big) + V_{1,1}\big)\Big)(V_{j+1,1} + V_{j+1,2})\Big] \\
&= \rho_2 \sum_{j=0}^{\infty} \gamma^j \mathbb{E}\Big[\Big(V_{0,2} + \frac{\rho_1}{1-\rho_1}(V_{0,1} + V_{0,2}) + V_{1,1}\Big)(V_{j+1,1} + V_{j+1,2})\Big].
\end{aligned}$$

Using the covariance functions and collecting terms gives

$$\begin{aligned}
\mathbb{E}[S_{n,2}(V_{n,1}+V_{n,2})] = \rho_2 \sum_{j=0}^{\infty} \gamma^j \Big(& \frac{R^2 + c_{12}(-j-1) + c_2(j+1)}{1-\rho_1} \\
& + \frac{\rho_1\big(c_1(j+1) + c_{12}(j+1)\big)}{1-\rho_1} + c_1(j) + c_{12}(j)\Big).
\end{aligned}$$

Here the term R^2 can be taken out of the summation and the covariance terms can be re-indexed (for example, $\gamma \sum_{j=0}^{\infty} \gamma^j c_1(j+1) = \sum_{j=1}^{\infty} \gamma^j c_1(j)$) to give

$$\begin{aligned}
\mathbb{E}[S_{n,2}(V_{n,1} + V_{n,2})] =& \frac{\rho_2}{1-\rho} R^2 + \rho_2\big(c_1(0) + c_{12}(0)\big) \\
&+ \sum_{j=1}^{\infty} \Big[c_{12}(-j) + c_2(j) + \rho_1\big(c_1(j) + c_{12}(j)\big) + \rho_2\big(c_1(j) + c_{12}(j)\big)\Big]\gamma^j \\
=& \frac{\rho_2 R^2}{1-\rho} + \rho_2\big(\delta_1^2 + c_{12}(0)\big) + \sum_{j=1}^{\infty} \Big[c_{12}(-j) + c_2(j) + \rho\big(c_1(j) + c_{12}(j)\big)\Big]\gamma^j.
\end{aligned}$$

This expression can be put back into (42) to disclose that

$$
\begin{aligned}
\mathbb{E}[I_{n,1}^2] =& \delta_1^2 + \delta_2^2 + \left(1 + \frac{2\rho_2}{1-\rho}\right) R^2 + 2\rho_2\big(\delta_1^2 + c_{12}(0)\big) \\
&+ 2\sum_{j=1}^{\infty} \Big[c_{12}(-j) + c_2(j) + \rho\big(c_1(j) + c_{12}(j)\big)\Big]\gamma^j \\
&+ \frac{1}{\omega}\left[\frac{R}{1-\rho}\left(\lambda_1 b_1^{(2)} + \frac{\lambda_2(1-\rho_1)^2 b_2^{(2)}}{\rho_2^2}\right) + (1-2\rho_1(1-\rho))\delta_1^2 + \delta_2^2 + \left(1+\frac{2\rho_2}{1-\rho}\right) R^2 \right. \\
&\qquad \left. + 2\rho\sum_{j=0}^{\infty} c_{12}(j)\gamma^j + 2\sum_{j=1}^{\infty}\left(c_2(j) + \frac{1-\rho_1(2-\rho)}{\rho_2}\Big(c_{12}(-j) + \rho c_1(j)\Big)\right)\gamma^j\right],
\end{aligned}
$$

with $\gamma := \frac{\rho_2}{1-\rho_1}$ and $\omega := \frac{(1-\rho_1)^2-\rho_2^2}{\rho_2^2}$. Collecting terms gives the second moment of the intervisit time of the first queue,

$$
\begin{aligned}
\mathbb{E}[I_{n,1}^2] =& \frac{\lambda_1\rho_2^2 b_1^{(2)} + \lambda_2(1-\rho_1)^2 b_2^{(2)}}{(1-\rho)^2(1-\rho_1+\rho_2)} R + \frac{(1-\rho_1)^2\Delta^2}{(1-\rho)(1-\rho_1+\rho_2)} + \frac{2\rho_2(1-\rho_1)(1+\rho_2)}{1-\rho_1+\rho_2}\delta_1^2 \\
&+ \frac{(1-\rho_1)^2}{(1-\rho)^2} R^2 + \frac{2\rho_2}{1-\rho}\left(\frac{1-\rho_1(2-\rho)}{1-\rho_1+\rho_2}\right) c_{12}(0) \qquad (43) \\
&+ \frac{2(1-\rho_1)^2}{(1-\rho)(1-\rho_1+\rho_2)} \sum_{j=1}^{\infty}\left[\left(1+\frac{\rho_2(1-\rho)}{1-\rho_1}\right)\Big(\rho c_1(j) + c_{12}(-j)\Big) + c_2(j) + \rho c_{12}(j)\right]\gamma^j.
\end{aligned}
$$

4.2.4 Duration of the Busy Period at the First Queue

The time the server spends, per cycle, at the first queue is made up of

$$D_{n+1,1} = \mathcal{D}_{n+1,1}\big(\mathcal{N}_{n,1}(I_{n,1})\big),$$

the time needed to empty a queue (exhaustively), starting with the number of customers that have arrived at and accumulated at the first queue since the last time the server served the first queue, i.e. since the intervisit time $I_{n,1}$. The expected value is

$$\mathbb{E}[D_{n+1,1}] = \frac{\rho_1 R}{1-\rho_1}. \qquad (44)$$

Taking once again the expectation over the square gives $\mathbb{E}[D_{n+1,1}^2] = \mathbb{E}[\mathcal{D}_{n+1,1}^2\big(\mathcal{N}_{n,1}(I_{n,1})\big)]$, which with (9) gives

$$\mathbb{E}[D_{n+1,1}^2] = \frac{\rho_1^2\mathbb{E}[I_{n,1}^2]}{(1-\rho_1)^2} + \frac{\lambda_1 R b_1^{(2)}}{(1-\rho_1)^2(1-\rho)}, \qquad (45)$$

as the second moment of the duration of the busy period at the first queue.

4.2.5 Number of Customers Waiting When Server Arrives

The number of customers, $L^*_{n+1,i}$, waiting in a queue at the moment the queue is polled is equal to either ($i = 1$) the number of customers that arrived during the last intervisit time or ($i = 2$) the number of customers that have arrived since the gate closed. This gives

$$L^*_{n+1,1} = \mathcal{N}_{n,1}(I_{n,1})$$
$$L^*_{n+1,2} = \dot{\mathcal{N}}_{n,2}(C_{n,2}).$$

Thus the expected number of customers waiting at the polling instances are

$$\mathbb{E}[L^*_{n+1,1}] = \frac{\lambda_1(1-\rho_1)R}{1-\rho} \tag{46a}$$
$$\mathbb{E}[L^*_{n+1,2}] = \frac{\lambda_2 R}{1-\rho}. \tag{46b}$$

The second moments follows through squaring to give (see (6))

$$\mathbb{E}[(L^*_{n+1,1})^2] = \lambda_1^2 \mathbb{E}[I^2_{n,1}] + \frac{\lambda_1(1-\rho_1)R}{1-\rho} \tag{47a}$$
$$\mathbb{E}[(L^*_{n+1,2})^2] = \lambda_2^2 \mathbb{E}[C^2_{n,2}] + \frac{\lambda_2 R}{1-\rho}. \tag{47b}$$

4.2.6 Number of Customers Served per Cycle

The derivation of the expected number of customers served, per cycle, at the first queue is completely identical to that of the exhaustive/exhaustive service discipline given in section 3.2.5 (with $i = 1$). Equations (25) and (26) give

$$\mathbb{E}[T_{n,1}] = \frac{\mathbb{E}[L^*_{n,1}]}{1-\rho_1} = \frac{\lambda_1 R}{1-\rho}, \tag{48}$$

$$\begin{aligned} \mathbb{E}[T^2_{n,1}] &= \frac{1}{(1-\rho_1)^2}\left(\mathbb{E}[(L^*_{n,1})^2] + \frac{\lambda_1 R\Big(\rho_1(1-\rho_1) + \lambda_1^2 b_1^{(2)}\Big)}{1-\rho}\right) \\ &= \frac{1}{(1-\rho_1)^2}\left(\lambda_1^2 \mathbb{E}[I^2_{n,1}] + \frac{\lambda_1 R\Big(1-\rho_1^2 + \lambda_1^2 b_1^{(2)}\Big)}{1-\rho}\right). \end{aligned} \tag{49}$$

The number of customers served, per cycle, at the second queue is easier since, because of its gated nature, it is equal to the number of customers waiting in the queue at the moment the queue gets polled. Thus $T_{n,2} = L^*_{n,2}$ and, in particular,

$$\mathbb{E}[T_{n,2}] = \mathbb{E}[L^*_{n,2}] = \frac{\lambda_2 R}{1-\rho}, \tag{50}$$
$$\mathbb{E}[T^2_{n,2}] = \mathbb{E}[(L^*_{n,2})^2] = \lambda_2^2 \mathbb{E}[C^2_{n,2}] + \frac{\lambda_2 R}{1-\rho}. \tag{51}$$

4.2.7 Expected Waiting Time and Average Queue Length

Theorem 4.4 (Expected Waiting Times for Exhaustive/Gated Served Queues). *If queue 1 is serviced exhaustively and queue 2 has gated service, then the expected time a customer waits in queue i $(i = 1, 2)$ until being served is given by*

$$\mathbb{E}[W_{q,1}] = \frac{(1-\rho_1)R}{2(1-\rho)} + \frac{\rho_2(1-\rho)(1+\rho_2)\delta_1^2}{R(1-\rho_1+\rho_2)} + \frac{1-\rho_1}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho} + \frac{\Delta^2+2C_1}{R}\right] \quad (52a)$$

$$\mathbb{E}[W_{q,2}] = \frac{(1+\rho_2)R}{2(1-\rho)} - \frac{\rho_1(1-\rho)(1+\rho_2)\delta_1^2}{R(1-\rho_1+\rho_2)} + \frac{1+\rho_2}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho} + \frac{\Delta^2+2C_2}{R}\right] \quad (52b)$$

where

$$C_1 = \sum_{j=1}^{\infty}\left[c_2(j)+\rho c_{12}(j) + \left(1+\frac{\rho_2(1-\rho)}{1-\rho_1}\right)\left(c_{12}(-j)+\rho c_1(j)\right)\right]\gamma^j + \frac{\rho_2(1-\rho_1(2-\rho))}{(1-\rho_1)^2}c_{12}(0)$$

$$C_2 = \sum_{j=1}^{\infty}\left[c_2(j) + \rho c_{12}(j) + \frac{1-\rho_1(2-\rho)}{\rho_2}\left(c_{12}(-j) + \rho c_1(j)\right)\right]\gamma^j + 2\rho c_{12}(0)$$

is the increase due to correlated switchover times. Here $\gamma := \frac{\rho_2}{1-\rho_1}$.

The average queue lengths,

$$\mathbb{E}[L_{s,i}] = \lambda_i \mathbb{E}[W_{q,i}] + \rho_i, \qquad i = 1, 2. \quad (53)$$

follow immediately because of Little. ◇

If the switchover times for each of the two queues are indepedent, as well as the switchover times between the two queues, then $C_1 = 0 = C_2$ and we obtain the results given in [16, formulas (25) and (28)] or [19, formula[7] 4.1].

Proof: For a customer arriving at the first (exhaustive) queue, the derivation of the waiting time is identical to that leading to (33) in Theorem 3.4. Therefore we have

$$\mathbb{E}[W_{q,1}] = \frac{\lambda_1 b_1^{(2)}}{2(1-\rho_1)} + \frac{\mathbb{E}[I_{n,1}^2]}{2\mathbb{E}[I_{n,1}]}.$$

Filling in the values of $\mathbb{E}[I_{n,1}]$ and $\mathbb{E}[I_{n,1}^2]$ gives the expected waiting time of a customer in the first queue.

Now imagine a tagged customer arriving at the second queue. The tagged customer has to wait for the cycle to finish (the expected remaining cycle time is $\mathbb{E}[C_{n,2}^2]/2\mathbb{E}[C_{n,2}]$) plus the time needed to serve the customers which arrived in the same cycle, but the before the tagged customer did (the expected number of customers arriving before the tagged customer

[7] The formula presented in this reference is taken incorrectly from [16]. The first term for the waiting time for customers arriving at the second queue should contain $(1+\rho_2)$ instead of $(1+\rho_1)$.

is $\lambda_2 \mathbb{E}[C_{n,2}^2]/2\mathbb{E}[C_{n,2}]$ and they require an expected total time of $\lambda_2 b_2 \mathbb{E}[C_{n,2}^2]/2\mathbb{E}[C_{n,2}]$ to serve). Therefore, the expected waiting time for a customer arriving at the second queue is

$$\mathbb{E}[W_{q,2}] = \frac{\mathbb{E}[C_{n,2}^2]}{2\mathbb{E}[C_{n,2}]}(1 + \rho_2).$$

Filling in the value for the second moment of the cycle time qives $\mathbb{E}[W_{q,2}]$. ∎

4.3 Identical Queues

In the case of identical queues ($\lambda := \lambda_i$, $\hat{\rho} = \rho_i$, $v := v_i$, $\delta := \delta_i$, $b^{(2)} := b_i^{(2)}$, and $c(j) := c_i(j)$, for $i = 1, 2$) we have as the stability condition $\hat{\rho} = \lambda d < 1/2$ and

$$\mathbb{E}[D_{n,1}] = \frac{2\hat{\rho}v}{1-\hat{\rho}} \qquad \text{(Service Time/Busy Period Q1)}$$

$$\mathbb{E}[S_{n,2}] = \frac{2\hat{\rho}v}{1-2\hat{\rho}} \qquad \text{(Service time/Busy Period Q2)}$$

$$\mathbb{E}[I_{n,1}] = \frac{2(1-\hat{\rho})v}{1-2\hat{\rho}} \qquad \text{(Intervisit time Q1)}$$

$$\mathbb{E}[C_{n,2}] = \frac{2v}{1-2\hat{\rho}} \qquad \text{(Cycle time Q2)}$$

$$\mathbb{E}[L_{n,1}^*] = \frac{2\lambda(1-\hat{\rho})v}{1-2\hat{\rho}} \qquad \text{(Number Waiting Q1)}$$

$$\mathbb{E}[L_{n,2}^*] = \frac{2\lambda v}{1-2\hat{\rho}} \qquad \text{(Number Waiting Q2)}$$

$$\mathbb{E}[T_{n,i}] = \frac{2\lambda v}{1-2\hat{\rho}} \qquad \text{(Number Served per Cycle)}$$

Writing out the second moments of these quantities does not provide much insight. The expected waiting times are

$$\mathbb{E}[W_{q,1}] = \frac{(1-\hat{\rho})v}{1-2\hat{\rho}} + \frac{\hat{\rho}(1-2\hat{\rho})(1+\hat{\rho})\delta^2}{2v} + \frac{1-\hat{\rho}}{2}\left[\frac{2\lambda b^{(2)}}{1-2\hat{\rho}} + \frac{\delta^2 + C_1}{v}\right] \tag{54a}$$

$$\mathbb{E}[W_{q,2}] = \frac{(1+\hat{\rho})v}{1-2\hat{\rho}} - \frac{\hat{\rho}(1-2\hat{\rho})(1+\hat{\rho})\delta^2}{2v} + \frac{1+\hat{\rho}}{2}\left[\frac{2\lambda b^{(2)}}{1-2\hat{\rho}} + \frac{\delta^2 + C_2}{v}\right] \tag{54b}$$

with

$$C_1 = \sum_{j=1}^{\infty} \left[2\hat{\rho}c_{12}(j) + \frac{1-2\hat{\rho}^3}{1-\hat{\rho}} c(j) + \frac{1-2\hat{\rho}^2}{1-\hat{\rho}} c_{12}(-j)) \right] \gamma^j + \frac{\hat{\rho}(1-2\hat{\rho}(1-\hat{\rho}))}{(1-\hat{\rho})^2} c_{12}(0) \quad (55a)$$

$$C_2 = \sum_{j=1}^{\infty} \left[2\hat{\rho}c_{12}(j) + \left(3 - 4\hat{\rho} + 4\hat{\rho}^2\right) c(j) + \frac{1-2\hat{\rho}(1-\hat{\rho})}{\hat{\rho}} c_{12}(-j) \right] \gamma^j + 4\hat{\rho}c_{12}(0) \quad (55b)$$

being the increase due to correlated switchover times. In this case $\gamma := \frac{\hat{\rho}}{1-\hat{\rho}}$.

The average queue lengths are given by

$$\mathbb{E}[L_{s,i}] = \lambda \mathbb{E}[W_{q,i}] + \hat{\rho}, \quad i = 1, 2. \quad (56)$$

4.4 Comparison of Waiting Times in Exhaustive/Gated System

In the subsequent sections the effect of the service discipline at the two queues will be studied. We start by examining the expected waiting times in the exhaustive/gated system. Although the discussions are about the expected waiting times, the same arguments hold for the average queue length.

4.4.1 Single Server Queue with Correlated Vacations

By turning off one of the queues one obtains an M/G/1 queue with multiple correlated vacations. Let us start by turning off the second queue (by setting $\lambda_2 = 0$, $\rho_2 = 0$, $v_2 = 0$, and $v_2^{(2)} = 0$, which leads to $c_2(j) = 0$, $c_{12}(j) = 0$, and $\gamma = 0$) to end up with an exhaustively served M/G/1 queue where the expected waiting time

$$\mathbb{E}[W_{q,1}] = \frac{\lambda_1 b_1^{(2)}}{2(1-\rho_1)} + \frac{v_1^{(2)}}{2v_1} \quad \text{(Exhaustive M/G/1)}$$

is independent of the correlation between the vacations! This result was previously pointed out in [1, paragraph 3.6] which causes it to correspond to the expression for the expected waiting time but with i.i.d. vacation times [20, page 123].

On the other hand, by turning off the first queue in the exhaustive/gated system, we are left with an M/G/1 queue with a gated service discipline. After setting the appropriate parameters to zero we obtain

$$\mathbb{E}[W_{q,2}] = \frac{\lambda_2 b_2^{(2)}}{2(1-\rho_2)} + \frac{v_2^{(2)}}{2v_2} + \frac{\rho_2 v_2}{1-\rho_2} + \frac{1}{v_2} \sum_{j=1}^{\infty} c_2(j) \rho_2^j \quad \text{(Gated M/G/1)}$$

as the waiting time of a customer arriving at a gated M/G/1 queue with correlated vacations. If there is no correlation then this expression is in agreement with the result previously obtained in [1, Theorem 5] and [20, equation (5.24a)].

It is interesting to compare the difference between these two waiting times due to the server behaving differently. Assuming queues with identical parameters (by dropping the indices of ones and twos and setting $\hat{\rho} := \rho_1 = \rho_2$) we see that

$$\mathbb{E}[W]_{gated\ M/G/1} - \mathbb{E}[W]_{exhaustive\ M/G/1} = \frac{\hat{\rho} v}{1-\hat{\rho}} + \frac{1}{v}\sum_{j=1}^{\infty} c(j)\hat{\rho}^j, \tag{57}$$

where the first term on the right hand side is the mean length of a service period (which is the same for the exhaustive and the gated service systems). If there is no correlation, then it is well known that the expected waiting time in an exhaustively served queue is less than that in a gated serviced queue. In the presence of correlated vacation times this difference is larger but remains a surprisingly simple expression.

4.4.2 Nonidentical Queues

The difference between the waiting times at the two queues is found by subtracting (52b) from (52a) to give $\mathbb{E}[W_{q,2}] - \mathbb{E}[W_{q,1}] =$

$$\frac{\rho R}{2(1-\rho)} - \frac{\rho(1-\rho)(1+\rho_2)\delta_1^2}{R(1-\rho_1+\rho_2)} + \frac{\rho}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho} + \frac{\Delta^2 + 2(C_2 - C_1)}{R}\right]$$

$$= \frac{\rho R}{2(1-\rho)} - \frac{\rho(1-2(\rho_1+\rho_2\rho))\delta_1^2}{2R(1-\rho_1+\rho_2)} + \frac{\rho}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho} + \frac{\delta_2^2 + 2(C_2 - C_1)}{R}\right],$$

with C_1 and C_2 defined in Theorem 4.4. Although not shown here, it can be verified that for any choice of parameters $C_2 \geq C_1$. From this expresion we see that if $1 - 2(\rho_1 + \rho_2\rho) > 0$, and $C_1 = 0 = C_2$, and if δ_1^2 is sufficiently large, that it may very well be possible that the expected waiting time at the gated queue is smaller than the expected waiting time at the exhaustive queue! However, the range of parameter settings for which this is the case is fairly small. In particular, in the next sections it will be shown that this does not happen when all parameters are equal, if there are no switchover times, or if the system is heavily loaded.

4.4.3 Identical Queues

Assume that, except for the service discipline, the two queues in the exhaustive/gated system have equal parameter settings. The difference between the waiting time of a customer arriving at the gated or at the exhaustive queue is

$$\mathbb{E}[W_{q,2}] - \mathbb{E}[W_{q,1}] = \hat{\rho}\left(\frac{2\lambda b^{(2)} + 2v}{1-2\hat{\rho}} + \frac{\hat{\rho}(1+2\hat{\rho})\delta^2 + C_2 - C_1}{v}\right),$$

with C_1 and C_2 defined in (55). For all parameter settings $C_2 \geq C_1$. From this we can deduct that, under equal parameter settings, *the average waiting time at the gated queue is always larger than the average waiting time at the exhaustive queue.* As pointed out in the previous section, this is not always the case if the parameter values are not identical.

4.4.4 Switchover Times Equal to Zero

Let us now examine the two service disciplines in the interesting situation when the switchover times are equal to zero. By looking at the expressions for the expected waiting times (equation (52)), we see that the variances and (cross) correlations of the switchover times must go faster to zero than the mean switchover times does since otherwise the expected waiting times could explode (for example, $\lim_{R\to 0} \delta_1^2/R$ must go to zero).

By first sending the second moments and the (cross) correlations of the switchover times to zero, and then the mean switchover times to zero, we see that the expected waiting times for the exhaustive/gated service discipline are

$$\mathbb{E}[W_{q,1}] = \frac{1-\rho_1}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho}\right],$$
$$\mathbb{E}[W_{q,2}] = \frac{1+\rho_2}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho}\right].$$

From this it is clear that a customer arriving at the second queue is expected to wait

$$\mathbb{E}[W_{q,2}] - \mathbb{E}[W_{q,1}] = \frac{\rho}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho}\right]$$

longer than a customer that arrives at the first queue.

4.4.5 Heavily Loaded System

Let us now consider what happens if the system is heavily loaded (ρ close to 1). In this case (with the exception of the case where $\rho_1 \to 1$ and $\rho_2 \to 0$, in which case we refer to section 4.4.1) the expected waiting times (taken from 52),

$$\mathbb{E}[W_{q,1}] \to \frac{(1-\rho_1)R}{2(1-\rho)} + \frac{1-\rho_1}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho}\right]$$
$$\mathbb{E}[W_{q,2}] \to \frac{(1+\rho_2)R}{2(1-\rho)} + \frac{1+\rho_2}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho}\right]$$

explode due to the factor $1-\rho$ in the denominators. As we can see, the correlation no longer plays a role. The difference between the expected waiting times tends towards

$$\mathbb{E}[W_{q,2}] - \mathbb{E}[W_{q,1}] \to \frac{\rho R}{2(1-\rho)} + \frac{\rho}{2(1-\rho_1+\rho_2)}\left[\frac{\lambda_1 b_1^{(2)} + \lambda_2 b_2^{(2)}}{1-\rho}\right]$$

which is always positive.

More discussions on the differences arising due to the server behaviour can be found in the examples of the next section.

5 Examples

In the following sections we will consider a number of examples in which the sequences of switchover times are correlated. The covariance functions will be calculated explicitely after which the effect of the correlation on the waiting times will be studied. In all of the examples the expected waiting time of a customer arriving at the first queue of the exhaustive/exhaustive system is given by (28), whereas in the exhaustive/gated system the expected waiting times are given by (52). The difference between each of the examples is that C, C_1, and C_2 take on different values.

5.1 Example 1: Correlated Switchover Times

Consider a sequence of switchover times where there is no correlation between the switchover times of the two queues (this gives $c_{12}(j) = 0$, for $j \in \mathbb{Z}$). Let the individual sequence of switchover times per queue satisfy

$$V_{n+1,i} = x_i V_{n,i} + (1 - x_i)\varepsilon_{n,i}, \qquad i = 1, 2, \tag{58}$$

where $x_i \in [0, 1)$ is a constant and $\varepsilon_{n,i}$ are positive i.i.d. variables with finite expectation $\mathbb{E}[\varepsilon_{n,i}] =: \bar{\varepsilon}_i$ and second moment $\mathbb{E}[\varepsilon^2_{n,i}] =: \varepsilon_i^{(2)}$. The parameter x_i determines the amount of correlation in the sequence; with $x_i = 0$ the sequence is i.i.d., whereas when x_i tends to one the correlation is maximal. Notice that there exists a stationary ergodic sequence of switchover times which satisfies (58). By taking the expectation it follows that $\mathbb{E}[V_{n+1,i}] = x_i\mathbb{E}[V_{n,i}] + (1 - x_i^n)\bar{\varepsilon}_i$. Due to the stationarity of the process $\mathbb{E}[V_{0,i}] = \mathbb{E}[V_{n,i}] = v_i$ is independent of x_i, and therefore $v_i = \bar{\varepsilon}_i$. A similar relationship can be derived for the second moments by taking the expectation over the square of (58) to give

$$\begin{aligned}\mathbb{E}[V^2_{n+1,i}] =& x_i^2\mathbb{E}[V^2_{n,1}] + (1 - x_i)^2\mathbb{E}[\varepsilon^2_{n,i}] \\ &+ 2x_i(1 - x_i)\mathbb{E}[\varepsilon_{n,i}]\mathbb{E}[V_{n,i}].\end{aligned}$$

Due to the stationarity ($\mathbb{E}[V^2_{n+1,i}] = \mathbb{E}[V^2_{n,i}] = v_i^{(2)}$) this implies that

$$v_i^{(2)} = \frac{(1 - x_i)\varepsilon_i^{(2)} + 2x_i\bar{\varepsilon}_i v_i}{1 + x_i},$$

which gives a second relationship (since $\bar{\varepsilon}_i = v_i$),

$$\delta_i^{(2)} = \frac{1 - x_i}{1 + x_i}\, Var(\varepsilon_{n,i}).$$

Thus we see that for $x_i \in [0, 1)$ there exists a $\varepsilon_{n,i}$ such that any desired values of v_i and δ_i^2 can be obtained. Now we will derive the covariance functions and the expected waiting time.

By iterating (58) a number of times it is quickly seen that

$$V_{n,i} = x_i^n V_{0,i} + (1 - x_i) \sum_{k=0}^{n-1} \varepsilon_{n-1-k,i} x_i^k. \tag{59}$$

From this we obtain

$$\mathbb{E}[V_{0,i} V_{j,i}] = x_i^j \mathbb{E}[V_{0,i}^2] + (1 - x_i^j)\bar{\varepsilon}_i \mathbb{E}[V_{0,i}] = x_i^j v_i^{(2)} + (1 - x_i^j)\bar{\varepsilon}_i v_i.$$

This means that the covariance functions, $c_i(j) = \mathbb{E}[V_{0,i}V_{j,i}] - \mathbb{E}[V_{0,i}]\mathbb{E}[V_{j,i}]$, are given by

$$c_i(j) = x_i^j v_i^{(2)} + (1 - x_i^j)\bar{\varepsilon}_i v_i - v_i \left(x_i^j v_i + (1 - x_i^j)\bar{\varepsilon}_i \right) = x_i^j \left(v_i^{(2)} - v_i^2 \right) = x_i^j \delta_i^2. \tag{60}$$

Since

$$\sum_{j=1}^{\infty} c_i(j)\alpha^j = \delta_i^2 \sum_{j=1}^{\infty} (\alpha x_i)^j = \frac{\alpha x_i \delta_i^2}{1 - \alpha x_i}, \tag{61}$$

we have from Theorem 3.4 that the expected waiting time in the exhaustive/exhaustive system is given by (28) with

$$C := \frac{\alpha x_1 \delta_1^2}{1 - \alpha x_1} + \frac{\alpha x_2 \delta_2^2}{1 - \alpha x_2} \left(1 + \frac{(1-\rho)^2}{\rho_1(1-\rho_1)} \right)$$

and $\alpha = \frac{\rho_1 \rho_2}{(1-\rho_1)(1-\rho_2)}$.

Equivalently, in the exhaustive/gated system we have from equation (61) and from Theorem 4.4 that the expected waiting times are given by (52) where $\gamma = \frac{\rho_2}{1-\rho_1}$ and

$$\begin{aligned} C_1 &= \frac{\gamma x_1 \delta_1^2}{1 - \gamma x_1} \left(1 + \frac{\rho_2(1-\rho)}{1-\rho_1} \right) \rho + \frac{\gamma x_2 \delta_2^2}{1 - \gamma x_2} \\ C_2 &= \frac{\gamma x_1 \delta_1^2}{1 - \gamma x_1} \left(1 + \frac{\rho_1}{\rho_2}(1-\rho)^2 \right) + \frac{\gamma x_2 \delta_2^2}{1 - \gamma x_2}. \end{aligned}$$

Numerical examples of the influence of the correlation on the expected waiting times can be found in Figure 2. Shown in each of the figures is the expected waiting time divided by the expected waiting time for uncorrelated sequences of switchover times. There are Poisson arrivals with $\lambda_i = 0.4$. The first two moments of the switchover times are always kept fixed (first moment for each of the switchover time distributions is fixed at $v_i = 3$) and the service times are taken to be exponential with $b_i = 0.4$ or $b_i = 1.2$.

Based on the figures and equations above the following important conclusions can be made:

- If $x_1 = x_2 = 0$ then there is no correlation between the sequences of switchover times and $C = 0$, $C_1 = 0$, and $C_2 = 0$;

- The increase in expected waiting times due to correlated switchover times can be up to several times (3.5 times in the example) the expected waiting times if there would be uncorrelated switchover times.
- The increase in the expected waiting times due to correlation grows linearly with the variance δ_i^2 of the switchover times;
- Under light traffic (α and γ are small and so) the increase in the expected waiting time is (approximately) linear in x_i.
- Under heavy traffic α and γ are close to one and, due to the factor $1-\alpha x_i$ or $1-\gamma x_i$ in the denominators, the increase in waiting time due to correlated switchover times can be significant. Hence the presence of correlation has the biggest impact on the waiting time if the system has a heavy load (and the switching times have a high variance). This can be seen clearly in Figure 2.
- In can be shown that, under identical parameter setting, in the exhaustive/gated system the expected waiting time at the exhaustive queue is always larger than at the gated queue. In addition to this, we see from Figure 2 that in lightly loaded systems the gated queue (Q2) suffers most from correlated switchover times whereas in heavily loaded traffic both queues are effected (relatively) equally by the correlated switchover times.

5.2 Example 2: Stochastic Recursive Sequence of Switchover Times

Consider a sequence of switchover times which satisfy the following stochastic recursive relationship

$$V_{n+1,i} = \mathcal{F}_{n,i}(V_{n,i}) + \mathcal{E}_{n,i}, \tag{62}$$

where $\mathcal{F}_{n,i}(\cdot)$ are independent, infinitely divisible stochastic processes with $\mathbb{E}[\mathcal{F}_{n,i}(T)] = x_i\mathbb{E}[T]$ and $\mathbb{E}[\mathcal{F}_{n,i}^2(T)] = x_i^{(2)}\mathbb{E}[T^2] + y_i\mathbb{E}[T]$. Here $x_i \in [0,1)$, $y_i \geq 0$ and $x_i^{(2)} \geq x_i^2$. The sequence $\mathcal{E}_{n,i}$ is a sequence of independent variables with $\mathbb{E}[\mathcal{E}_{n,i}] = \bar{\varepsilon}_i$ and $\mathbb{E}[\mathcal{E}_{n,i}^2] = \varepsilon_i^{(2)}$. Iterating gives

$$V_{n,i} = \left(\prod_{k=0}^{n-1} \mathcal{F}_{k,i}\right) V_{0,i} + \sum_{k=0}^{n-1} \left(\prod_{l=k+1}^{n-1} \mathcal{F}_{l,i}\right) \mathcal{E}_{n-k,i},$$

and so

$$\mathbb{E}[V_{n,i}] = x_i^n v_i + \bar{\varepsilon}_i \sum_{k=0}^{n-1} x_i^k = x_i^n v_i + \frac{1-x_i^n}{1-x_i}\bar{\varepsilon}_i.$$

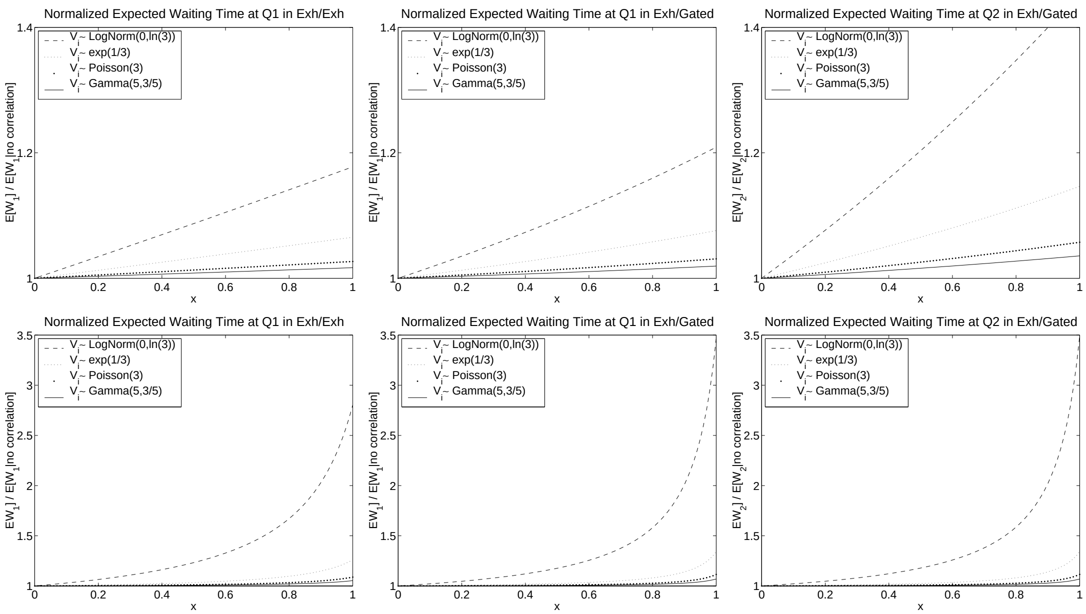

Figure 2: *Example: Correlated Switchover Times.* The expected waiting time divided by the expected waiting time with uncorrelated switchover times. The different lines correspond to different switchover time distributions, all with mean $v_i = 3$. Here $x := x_1 = x_2$ determines the level of correlation, there is no cross correlation, the service times are exponential, and $\lambda_i = 0.4$. The top figures are with mean service times $b_i = 0.4$ ($\rho = 0.32$) whereas the second row of figures are under heavy traffic with $b_i = 1.2$ ($\rho = 0.96$). The first column of figures correspond to the exhaustive/exhaustive system, whereas the second and third column of figures show the normalized waiting times for, respectively, the exhaustive queue (Q1) and the gated queue (Q2) in the exhaustive/gated system.

This leads to the condition $v_i = \frac{\bar{\varepsilon}_i}{1-x_i}$. The second moment of the switchover times is found by taking the expectation over the square of (62). Doing this produces

$$\mathbb{E}[V_{n+1,i}^2] = \mathbb{E}[\mathcal{F}_{n,i}^2(V_{n,i})] + \mathbb{E}[\mathcal{E}_{n,i}^2] + 2\mathbb{E}[\mathcal{F}_{n,i}(V_{n,i}) \cdot \mathcal{E}_{n,i}] = x_i^{(2)} v_i^{(2)} + y_i v_i + \varepsilon_i^{(2)} + 2x_i v_i \bar{\varepsilon}_i.$$

Since, by definition, $\mathbb{E}[V_{n+1,i}^2]$ also equals $v_i^{(2)}$, we have

$$v_i^{(2)} = \frac{\varepsilon_i^{(2)} + (y_i + 2x_i\bar{\varepsilon}_i)v_i}{1 - x_i^{(2)}}.$$

Thus we see that there exists a $\mathcal{E}_{n,i}$ which generates a sequence $V_{n,i}$ with arbitrary correlation and any desired values of v_i and $v_i^{(2)}$. Similarly, it is easy to show that

$$\mathbb{E}[V_{0,i}V_{n,i}] = x_i^n v_i^{(2)} + \frac{1 - x_i^n}{1 - x_i}\bar{\varepsilon}_i v_i.$$

This gives the covariance functions

$$c_i(j) = \mathbb{E}[V_{0,i}V_{j,i}] - \mathbb{E}[V_{0,i}]\mathbb{E}[V_{j,i}] = x_i^j v_i^{(2)} + \frac{1 - x_i^j}{1 - x_i}\bar{\varepsilon}_i v_i - v_i \left(x_i^j v_i + \frac{1 - x_i^j}{1 - x_i}\bar{\varepsilon}_i \right) = x_i^j \delta_i^2$$

which are completely identical to (60)! This means that if the first two moments of the switchover times are kept fixed while varying $x \in [0, 1)$, that the expected waiting times are once again given by (28) and (52) with C, C_1, and C_2 as given in the first example. Furthermore, the conclusions of the first example also hold here.

As a special case of (62) we can take $V_{n+1,i} = x_i V_n + \varepsilon_{n,i}$ where $x_i \in [0, 1)$ is a constant and $\varepsilon_{n,i}$ is a positive sequence of i.i.d. variables.

5.3 Example 3: Identical Switchover Times

Set $V_{n,2} = V_{n,1}$. This introduces cross-correlation between the two sequences of switchover times and it gives $v_2 = v_1$ and $\delta_2^2 = \delta_1^2$. In addition to this let $V_{n+1,1} = xV_{n,1} + (1 - x)\varepsilon_{n,1}$ just as in the first example. From (60) we have $c_1(j) = x^j \delta_1^2$ after which

$$\begin{aligned}
c_2(j) &= \mathbb{E}[V_{0,2}V_{n,2}] - \mathbb{E}[V_{0,2}]\mathbb{E}[V_{n,2}] = \mathbb{E}[V_{0,1}V_{n,1}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{n,1}] = c_1(j) = x^j \delta_1^2 \\
c_{12}(j) &= \mathbb{E}[V_{0,1}V_{n,2}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{n,2}] = \mathbb{E}[V_{0,1}V_{n,1}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{n,1}] = c_1(j) = x^j \delta_1^2 \\
c_{12}(-j) &= \mathbb{E}[V_{0,2}V_{n,1}] - \mathbb{E}[V_{0,2}]\mathbb{E}[V_{n,1}] = \mathbb{E}[V_{0,1}V_{n,1}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{n,1}] = c_1(j) = x^j \delta_1^2
\end{aligned}$$

immediately follow. This means that $(i = 1, 2)$

$$\sum_{j=1}^{\infty} c_i(j)\alpha^j = \sum_{j=1}^{\infty} c_{12}(j)\alpha^j = \sum_{j=1}^{\infty} c_{12}(-j)\alpha^j = \frac{\alpha x \delta_1^2}{1 - \alpha x}$$

can all be plugged into Theorem 3.4 so that the expected waiting time in the exhaustive/exhaustive system is given by (28) with

$$C = \frac{\alpha x \delta_1^2}{1-\alpha x}\left(\frac{1-\rho_2(1-\alpha)}{\alpha x} + 2 + \frac{(1-\rho)^2}{\rho_1(1-\rho_1)} + \frac{1-\rho_2}{\rho_1}\right),$$

$\psi := \frac{1-\rho_1}{2v_1(1-\rho+2\rho_1\rho_2)}$, and $\alpha = \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}$.

Equivalently, the expected waiting times in the exhaustive/gated system are given by (52) with

$$C_1 = \frac{\gamma x \delta_1^2}{1-\gamma x}\left(2+\frac{\rho_2(1-\rho)}{1-\rho_1}\right)(1-\rho)+\frac{\rho_2(1-\rho_1(2-\rho))}{(1-\rho_1)^2}\delta_1^2,$$
$$C_2 = \frac{\gamma x \delta_1^2}{1-\gamma x}\left(1+\frac{1-\rho_1(2-\rho)}{\rho_2}\right)(1-\rho) + 2\rho\delta_1^2.$$

To get a feeling of the impact of the cross correlation, the expected waiting times are plotted in Figure 3 for various switchover time distributions and traffic loads. Shown in each of the figures is the expected waiting time divided by the expected waiting time for uncorrelated sequences of switchover times. There are Poisson arrivals with $\lambda_i = 0.4$. The first two moments of the switchover times are always kept fixed (first moment for each of the switchover time distributions is fixed at $v_i = 3$) and the service times are taken to be exponential with $b_i = 0.4$ or $b_i = 1.2$.

Striking is the impact of the cross correlation on the waiting times. For example, if there is no correlation within each sequence of switchover times ($x = 0$), then there is still an increase in the expected waiting time due to the cross-correlation. For the exhaustive/exhaustive system this increase is $(1-\rho_2(1-\alpha))\psi\delta_1^2$ and for the exhaustive/gated system this increase is given by $\frac{\rho_2(1-\rho_1(2-\rho))}{(1-\rho_1)^2}\delta_1^2$ and $2\rho\delta_1^2$ for, respectively, the exhaustive and the gated queue. For exponentially distributed switchover times this can mean an increase of tens of percents in the expected waiting time. Besides this, all of the conclusions made in the first example also hold here, with the exception that the increase in expected waiting time can up to a factor 5.

5.4 Example 4: Switchover Times Coming from the Same Sequence

As a last example consider the case where the two sequences of switchover times come from a single sequence:

$$Y_{n+1} = xY_n + (1-x)\varepsilon_n, \qquad \begin{matrix} V_{n,1} := Y_{2n} \\ V_{n,2} := Y_{2n+1}. \end{matrix} \tag{63}$$

Here Y_n is a stationary ergodic sequence, $\mathbb{E}[\varepsilon_n] = \varepsilon$, and $\mathbb{E}[\varepsilon_n^2] = \varepsilon^{(2)}$. In this case there is a large correlation between the two sequences, but it is not as strong as in example 3. Once

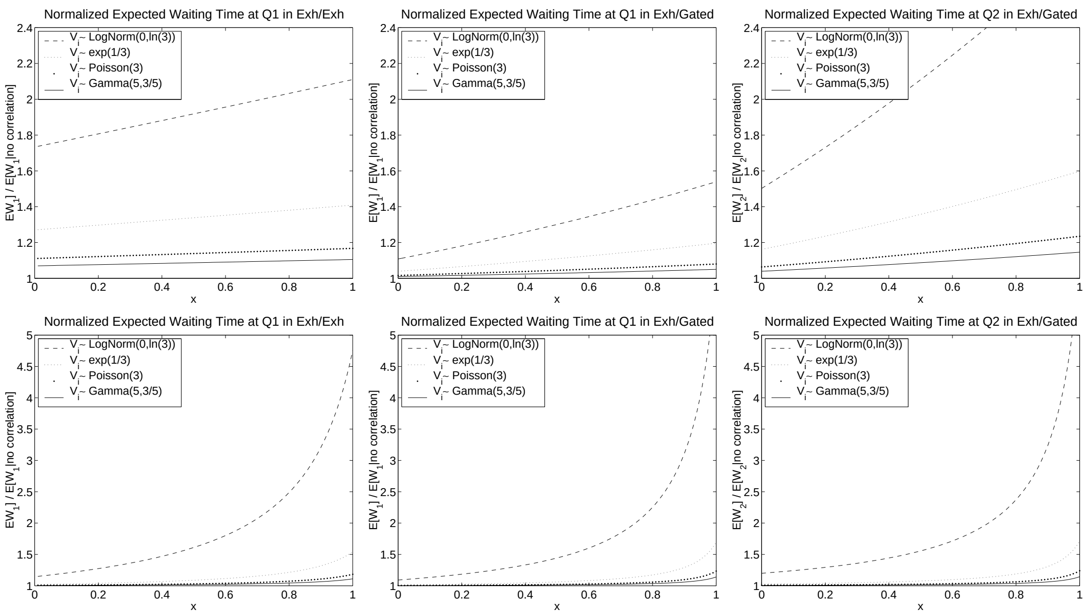

Figure 3: *Example: Identical Switchover Times.* The expected waiting time divided by the expected waiting time if there would be no correlation between the switchover times. The different lines correspond to different switchover time distributions. Here x determines the level of correlation, cross correlation is introduced by setting $V_{n,2} = V_{n,1}$, the service times are exponential, and $\lambda_i = 0.4$. The figures in the first row are with mean service time $b_i = 0.4$ ($\rho = 0.32$) and the figures on the bottom row are under heavy traffic with $b_i = 1.2$ ($\rho = 0.96$). The first column of figures correspond to the exhaustive/exhaustive system, whereas the second and third column of figures show the normalized waiting times for, respectively, the exhaustive queue (Q1) and the gated queue (Q2) in the exhaustive/gated system.

again we will derive an explicit expression for the expected waiting times. Equation (63) can be rewritten as

$$\begin{aligned} V_{n,1} &= xV_{n-1,2} + (1-x)\varepsilon_{2n} \\ V_{n,2} &= xV_{n,1} + (1-x)\varepsilon_{2n+1}. \end{aligned}$$

Iterating twice gives

$$\begin{aligned} V_{n,1} &= x^2 V_{n-1,1} + (1-x)\big(x\varepsilon_{2n-1} + \varepsilon_{2n}\big) \\ V_{n,2} &= x^2 V_{n-1,2} + (1-x)\big(x\varepsilon_{2n} + \varepsilon_{2n+1}\big), \end{aligned}$$

and iterating a few more times shows that

$$\begin{aligned} V_{n,1} &= x^{2n} V_{0,1} + \sum_{k=0}^{n-1} \left(x\varepsilon_{2(n-k)-1} + \varepsilon_{2(n-k)}\right) x^{2k} \\ V_{n,2} &= x^{2n} V_{0,2} + \sum_{k=0}^{n-1} \left(x\varepsilon_{2(n-k)} + \varepsilon_{2(n-k)+1}\right) x^{2k}. \end{aligned}$$

Taking the expectation over this gives

$$\mathbb{E}[V_{n,i}] = x^{2n} v_i + \frac{1-x^{2n}}{1-x}\varepsilon, \quad i = 1, 2,$$

which with the stationarity gives $v_i = \varepsilon$. This last result can be generalized to saying that all of the moments of $V_{n,1}$ and $V_{n,2}$ are equal to each other, in particular, $\delta_1^2 = \delta_2^2$. This follows from the stationarity of Y_n, i.e. $\mathbb{E}[V_{n,2}^k] = \mathbb{E}[Y_{2n+1}^k] = \mathbb{E}[Y_{2n}^k] = \mathbb{E}[V_{n,1}^k]$.
The covariance functions can be derived in the same way as in the previous examples,

$$c_1(j) = \mathbb{E}[V_{0,1}V_{j,1}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{j,1}] = x^{2j}\left(\mathbb{E}[V_{0,1}^2] - \mathbb{E}[V_{0,1}]^2\right) = x^{2j}\delta_1^2. \tag{64}$$

In a similar manner it can be derived that $c_2(j) = x^{2j}\delta_2^2 = x^{2j}\delta_1^2$. The cross covariance function for $j \in \mathbb{N}$ is

$$\begin{aligned} c_{12}(j) &= \mathbb{E}[V_{0,1}V_{j,2}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{j,2}] \\ &= \mathbb{E}\big[V_{0,1}\big(xV_{j,1} + (1-x)\varepsilon_{2j+1}\big)\big] + \mathbb{E}\big[V_{0,1}\big]\mathbb{E}\big[xV_{j,1} + (1-x)\varepsilon_{2j+1}\big] \\ &= x\mathbb{E}\big[V_{0,1}^2\big] + x\mathbb{E}\big[V_{0,1}\big]\mathbb{E}[V_{j,1}] = xc_1(j) = x^{2j+1}\delta_1^2. \end{aligned} \tag{65}$$

Similarly,

$$\begin{aligned} c_{12}(-j) &= \mathbb{E}[V_{-j,2}V_{0,1}] - \mathbb{E}[V_{-j,2}]\mathbb{E}[V_{0,1}] = \mathbb{E}[V_{0,2}V_{j,1}] - \mathbb{E}[V_{0,2}]\mathbb{E}[V_{j,1}] \\ &= \mathbb{E}[V_{0,2} \cdot \big(xV_{j-1,2} + (1-x)\varepsilon_{2n}\big)] - \mathbb{E}[V_{0,2}]\mathbb{E}\big[xV_{j-1,2} + (1-x)\varepsilon_{2n}\big] \\ &= xc_2(j-1) = x^{2j-1}\delta_1^2 \end{aligned} \tag{66}$$

Putting these expressions for the covariance functions into (19) tells us that the expected waiting time for a customer arriving at the first queue of an exhaustive/exhaustive system is given by (28) where $\psi := \frac{1-\rho_1}{R(1-\rho+2\rho_1\rho_2)}$, $\alpha = \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}$, and

$$C = \left(\frac{1-\rho_2(1-\alpha)}{\alpha x} + 2 + \frac{(1-\rho)^2}{\rho_1(1-\rho_1)} + \frac{1-\rho_2}{x\rho_1} \right) \frac{\alpha x^2 \delta_1^2}{1-\alpha x^2}. \tag{67}$$

Similarly, in the exhaustive/gated system the expected waiting times follow by putting (64)-(66) into Theorem 4.4. This leads to the expected waiting time being given by (52) with $\gamma = \frac{\rho_2}{1-\rho_1}$ and

$$C_1 = \frac{\gamma x^2 \delta_1^2}{1-\gamma x^2} \left(\rho + \frac{1}{x} \right) \left(x + 1 + \frac{\rho_2(1-\rho)}{1-\rho_1} \right) + \frac{\rho_2(1-\rho_1(2-\rho))}{(1-\rho_1)^2} x \delta_1^2 \tag{68a}$$

$$C_2 = \frac{\gamma x^2 \delta_1^2}{1-\gamma x^2} \left(p + \frac{1}{x} \right) \left(x + \frac{1-\rho_1(2-\rho)}{\rho_2} \right) + 2\rho x \delta_1^2 \tag{68b}$$

The waiting times as a function of x and for different switchover time distributions can be seen in Figure 4. Due to the strong similarities the same conclusions can be made as for the first examples, with the exception that in this case the increase is not entirely linear in x for systems with a low traffic load.

6 Conclusions

We have studied the performance of alternating-priority queues with very weak assumptions on the switchover time sequences; all we assume is that these sequences are stationary ergodic. In spite of this generality we were able to derive explicit expressions for the expected waiting times and number of customers in each queue. The expressions obtained involve the weighted sum of all correlations where the weights decrease exponentially fast to zero. With the help of our explicit expressions, we studied numerically the role of correlation and gave examples where they add up to 400% to the expected waiting times. This has important implications for (ad-hoc) networks where a common communication channel is shared amongst a number of users and the number of users between consecutive data transfers are correlated.

Acknowledgements

This work was partly funded by the EURONGI Network of Excellence. The authors would like to thank Uri Yechiali for proofreading the original manuscript as well as Daniele Miorandi for his discussions on polling systems.

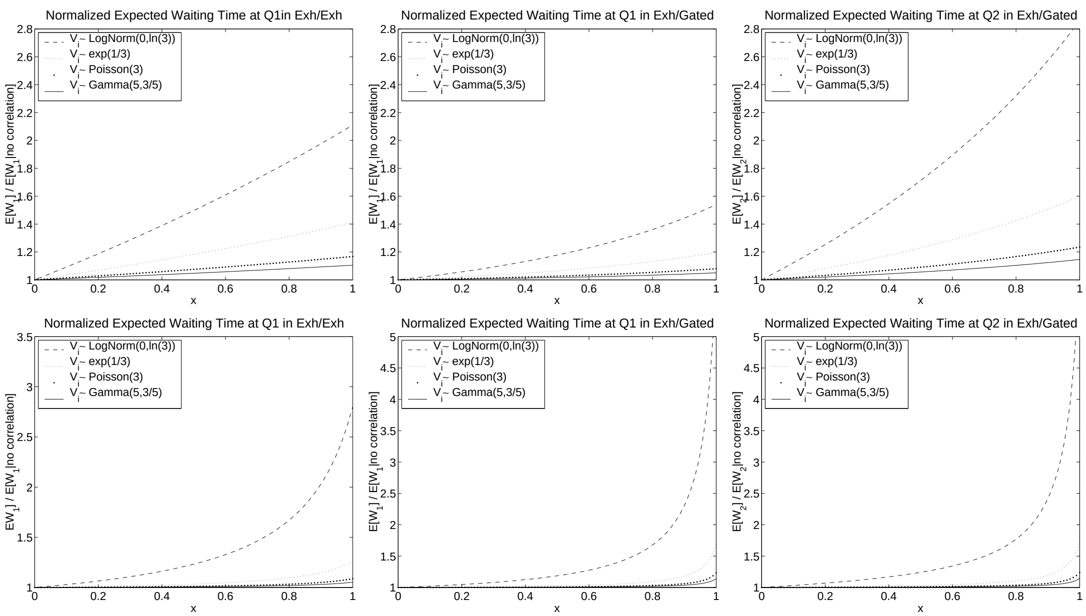

Figure 4: *Switchover Times from the Same Sequence.* The expected waiting time divided by the expected waiting time if there would be no correlation between the switchover times. The different lines correspond to different switchover time distributions. Here x determines the level of correlation, cross correlation is introducted having the switchover times come from the same sequence, the service times are exponential, and $\lambda_i = 0.4$. The figures in the first row are with mean service time $b_i = 0.4$ ($\rho = 0.32$) and the figures on the bottom row are under heavy traffic with $b_i = 1.2$ ($\rho = 0.96$). The first column of figures correspond to the exhaustive/exhaustive system, whereas the second and third column of figures show the normalized waiting times for, respectively, the exhaustive queue (Q1) and the gated queue (Q2) in the exhaustive/gated system.

A Proof of Theorem 3.2.

The proof of Theorem 3.2 will make use the following lemma [8, Propositions 6.6 and 6.31][10, p.292 and p.295]

Lemma A.1. *Let $X_1, X_2, \ldots$ be a stationary and ergodic process, and $\phi(\mathbf{x})$ be Borel measurable, then the process $B_1, B_2, \ldots$ defined by $B_n = \phi(X_n, X_{n+1}, \ldots)$ is stationary ergodic.*

First of all note that $\mathcal{A}_n(y)$ and $\mathcal{B}_n$ are nonnegative (componentwise) for all n and y. Moreover, $\mathcal{A}_n(y)$ are monotone increasing in y for all n and for each sample path. Theorem 1 in [1] then tells us that

(i) if the sequence $\{(\mathcal{A}_n(\cdot), \mathcal{B}_n), -\infty < n < \infty\}$ is stationary ergodic[8] and

(ii) if $P(\overline{\lim}_{n\to\infty} I_{n,1} \text{ is finite}) > 0$ for $I_{0,1} = 0$,

then there exists a stationary ergodic regime $I^*_{n,1}$ defined on the probability space and $I^*_{n,1}$ satisfies (13).

We start by verifying that $\mathcal{B}_n$ is stationary ergodic by analyzing it term by term. First of all the processes $V_{n,1}$ and $V_{n,2}$ are by assumption stationary ergodic. Furthermore it can be shown that $\mathcal{N}_{n,2}(V_{n,i})$, $i = 1, 2$, is also stationary ergodic. To do this introduce the sequence of i.i.d. uniform variables U_n on $(0, 1)$ and let $\mathcal{N}_{n,2}(\omega)$, $\omega \in \mathbb{R}$, have cummulative distribution function $F_{\mathcal{N}_{n,2}(\omega)}(x) := P(\mathcal{N}_{n,2}(\omega) \le x)$. With the help of the inverse-transformation method for drawing random numbers it follows that $\mathcal{N}_{n,2}(\omega) = F^{-1}_{\mathcal{N}_{n,2}(\omega)}(U_n)$ which leads to $\mathcal{N}_{n,2}(V_{n,i}) = F^{-1}_{\mathcal{N}_{n,2}(V_{n,i})}(U_n)$. In other words, $\mathcal{N}_{n,2}(V_{n,i})$ is a function of two parameters, U_n and $V_{n,i}$. Call this function $f(U, V) := F^{-1}_{\mathcal{N}_{n,2}(V)}(U)$. This function is Borel measurable, and since $X_n := (U_n, V_{n,i})$ is stationary ergodic it follows with Lemma A.1 that $\mathcal{N}_{n,2}(V_{n,2}) = f(U_n, V_{n,i})$ is stationary ergodic. An identical argumentation with the sequence of i.i.d. uniform variables W_n on $(0, 1)$ leads to the function $g(W_n, N) := F^{-1}_{\mathcal{D}_{n,2}(N)}(W_n) = \mathcal{D}_{n,2}(N)$ with which it can be shown that $\mathcal{D}_{n,2}(N_n) = g(W_n, N_n)$ is also stationary ergodic. Getting back to $\mathcal{B}_n$, with the vector $X_n := (V_{n,1}, V_{n,2}, U_n, W_n)$ and the projection function $P_i(X_n)$ which gives the i-th coordinate of X_n, we can rewrite $\mathcal{B}_n$ as

$$\begin{aligned}
\mathcal{B}_n &= V_{n+1,1} + V_{n+1,2} + g(W_{n+1}, f(U_n, V_{n,2}) + f(U_{n+1}, V_{n+1,1})) \\
&= P_1(X_{n+1}) + P_2(X_{n+1}) + g(P_4(X_{n+1}), f(P_3(X_n), P_2(X_n)) + f(P_3(X_{n+1}), P_1(X_{n+1})))
\end{aligned}$$

Since projections, additions and compositons of measurable functions are once again measurable it follows with Lemma A.1 that $\mathcal{B}_n$ is stationary ergodic.

Next we show that the sequence $\mathcal{A}_n(\cdot)$ is stationary ergodic by starting with the innermost term of (14) and working outwards. Since the number of arrivals in cycle n is independent of the number of arrivals in cycle k, for $k \ne n$, it follows that $\mathcal{N}_{n,1}(\cdot)$ is an i.i.d. sequence in n. With a similar argumentation it follows that $\mathcal{D}_{n,1}(\cdot)$ is an i.i.d. sequence in n. Furthermore,

[8] We mean by that notation that the sequence $\{((\mathcal{A}_n(y))_{y\in\mathcal{Y}}, \mathcal{B}_n), -\infty < n < \infty\}$ is stationary ergodic rather than $\{(\mathcal{A}_n(I_{n,1}), \mathcal{B}_n), -\infty < n < \infty\}$. Here $\mathcal{Y}$ is a subset of $\mathbb{R}$.

a composition of independent processes results once again in an i.i.d. sequence. To prove this let $\mathcal{D}_k(\cdot)$ and $\mathcal{N}_m(\cdot)$ be i.i.d. sequences and independent of each other. Conditioning on the inner process gives ($k \neq l$ and $m \neq n$)

$$\begin{aligned} P(\mathcal{D}_k(N_m) \leq x; \mathcal{D}_l(N_n) \leq y) = & \int_0^\infty \int_0^\infty P(\mathcal{D}_k(r) \leq x; \mathcal{D}_l(s) \leq y) dP(N_m \leq r) dP(N_n \leq s) \\ = & \int_0^\infty \int_0^\infty P(\mathcal{D}_k(r) \leq x) P(\mathcal{D}_l(s) \leq y) dP(N_m \leq r) dP(N_n \leq s) \\ = & P(\mathcal{D}_k(N_m) \leq x) P(\mathcal{D}_l(N_n) \leq y). \end{aligned}$$

In particular, this means that the sequence $\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(\cdot))$ is an i.i.d. sequence in n. Performing two more compositions like this proves that $\mathcal{A}_n(\cdot)$ is an i.i.d. sequence (and therefore also stationary ergodic). This proves (i).

To show that (ii) holds it is sufficient to show that certain conditions of Lemma 1 in [1] hold as that lemma leads to (ii). The first condition is that $\mathcal{A}_n(\cdot)$ forms an i.i.d. sequence independent of the stationary ergodic sequence $\mathcal{B}_n$. The second condition is that $\mathbb{E}[|\mathcal{A}_0(y)|] \leq \alpha|y|$ for some $\alpha < 1$ and all $y \geq D$ where $D > 0$ is some constant.

It has already been shown that the sequence $\mathcal{A}_n(\cdot)$ forms an i.i.d. sequence and that $\mathcal{B}_n$ is a stationary ergodic sequence. The sequence $\mathcal{A}_n(\cdot)$ is also independent of $\mathcal{B}_n$ since (compare (15) to (14)) $V_{n+1,1}$ and $V_{n+1,2}$ are independent of $\mathcal{A}_n(\cdot)$, $\mathcal{N}_{n,2}$ is independent of $\mathcal{N}_{n+1,2}$, and $V_{n+1,1}$ is independent of $\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(\cdot))$.

Finally, since $\mathbb{E}[|\mathcal{A}_0(y)|] = \mathbb{E}[\mathcal{A}_n(y)] = \alpha y$, with $\alpha = \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}$, it follows that the second condition holds since $\alpha < 1$ whenever $\rho < 1$, which was assumed throughout. Thus (ii) holds if $\rho < 1$.

Since (i) and (ii) hold there exists (Theorem 1 in [1]) a stationary regime $I^*_{n,1}$ which satisfies (13). ∎

B Proof of Theorem 3.3.

Proof: Taking the expectation on both sides of equation (13) gives

$$\mathbb{E}[I_{n+1,1}] = R + \frac{\rho_2}{1-\rho_2}\Big(R + \frac{\rho_1 \mathbb{E}[I_{n,1}]}{1-\rho_1}\Big).$$

Under the stationary regime $\mathbb{E}[I_{n+1,1}] = \mathbb{E}[I_{n,1}]$ which immediately leads to the first moment

$$E[I_{n,1}] = \frac{R(1-\rho_1)}{1-\rho}. \tag{69}$$

To obtain the second moment we need to derive the terms on the right hand side of (16). First of all,

$$\begin{aligned}\mathbb{E}[\mathcal{A}_n^2(I_{n,1})] =& \mathbb{E}\left[\mathcal{D}_{n+1,2}^2\Big(\mathcal{N}_{n+1,2}\Big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(I_{n,1}))\Big)\Big)\right] \\ =& \lambda_2^2 d_2^2 \mathbb{E}[\mathcal{D}_{n+1,1}^2(\mathcal{N}_{n,1}(I_{n,1}))] + \lambda_2 d_2^{(2)} \mathbb{E}[\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(I_{n,1}))] \\ =& \lambda_2^2 d_2^2 \Big(\lambda_1^2 d_1^2 \mathbb{E}[I_{n,1}^2] + \lambda_1 d_1^{(2)} \mathbb{E}[I_{n,1}]\Big) + \lambda_1 d_1 \lambda_2 d_2^{(2)} \mathbb{E}[I_{n,1}].\end{aligned}$$

Plugging equations (7) and (17) into this results in

$$\mathbb{E}[\mathcal{A}_n^2(I_{n,1})] = \frac{\rho_1^2\rho_2^2 \mathbb{E}[I_{n,1}^2]}{(1-\rho_1)^2(1-\rho_2)^2} + \frac{R}{(1-\rho_2)^2(1-\rho)}\left(\frac{\lambda_1\rho_2^2 b_1^{(2)}}{(1-\rho_1)^2} + \lambda_2\rho_1 b_2^{(2)}\right). \tag{70}$$

Next we proceed with the second unknown of expression (16), $\mathbb{E}[\mathcal{B}_n^2]$, where we recall that $\mathcal{B}_n$ is defined in equation (15). Making use of (3) gives

$$\begin{aligned}\mathbb{E}[\mathcal{B}_n^2] =& \mathbb{E}\Big[V_{n+1,1} + V_{n+1,2} + \mathcal{D}_{n+1,2}\Big(\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}(V_{n+1,1})\Big)\Big]^2 \\ =& v_1^{(2)} + v_2^{(2)} + 2v_1v_2 + 2c_{12}(0) + 2\lambda_2 d_2\left(v_1^{(2)} + v_2^2 + 2v_1v_2 + c_2(1) + c_{12}(-1) + c_{12}(0)\right) \\ &+ \lambda_2^2 d_2^2\Big(v_1^{(2)} + v_2^{(2)} + 2v_1v_2 + 2c_{12}(-1)\Big) + \lambda_2 d_2^{(2)} R \\ =& \frac{\Delta^2 + R^2}{(1-\rho_2)^2} - \frac{2\rho_2\delta_2^2}{1-\rho_2} + \frac{R\lambda_2 b_2^{(2)}}{(1-\rho_2)^3} + \frac{2\rho_2 c_2(1) + 2c_{12}(0)}{1-\rho_2} + \frac{2\rho_2 c_{12}(-1)}{(1-\rho_2)^2}.\end{aligned} \tag{71}$$

To solve the last part first notice that the processes $\mathcal{N}_{n,1}(\cdot)$, $\mathcal{N}_{n,2}(\cdot)$, $\mathcal{N}_{n+1,1}(\cdot)$, $\mathcal{N}_{n+1,2}(\cdot)$, $\mathcal{D}_{n+1,1}(\cdot)$, and $\mathcal{D}_{n+1,1}(\cdot)$ are all independent of each other, and each of them is independent of $I_{n,1}$, $V_{n,2}$, $V_{n+1,1}$, and $V_{n+1,2}$, This means that

$$\mathbb{E}[\mathcal{A}_n(I_{n,1})\mathcal{B}_n] = \alpha\mathbb{E}[I_{n,1}\mathcal{B}_n], \tag{72}$$

with $\alpha := \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}$. The last piece of the puzzle can be derived with the help of Theorem 2 in [1] which states that

$$I_{n,1} = \sum_{j=0}^{\infty}\left(\prod_{i=n-j}^{n-1}\mathcal{A}_i^{(n-j)}\right)(\mathcal{B}_{n-j-1}), \quad n \in \mathbb{Z},$$

where for each integer i, $\{\mathcal{A}_i^{(-j)}\}_j$ are independent of each other and have the same distribution as $\mathcal{A}_i(\cdot)$. To apply the theorem it is sufficient to have $\alpha < 1$, which turns out to be equivalent to $\rho < 1$, and that $\mathbb{E}[\mathcal{B}_n] < \infty$ (see Lemma 1 in [1]). The latter indeed holds as $\mathbb{E}[\mathcal{B}_n] = \frac{R}{1-\rho_2}$.

Applying the theorem gives

$$\begin{aligned}\mathbb{E}[I_{n,1}\mathcal{B}_n] &= \sum_{j=0}^{\infty} \mathbb{E}\Big[\prod_{i=n-j}^{n-1} \Big(\mathcal{D}_{i+1,2}\Big(\mathcal{N}_{i+1,2}\big(\mathcal{D}_{i+1,1}(\mathcal{N}_{i,1}(\mathcal{B}_{n-j-1}))\big)\Big)\Big) \cdot \mathcal{B}_n\Big] \\ &= \sum_{j=0}^{\infty} \alpha^j \mathbb{E}\left[\mathcal{B}_{n-j-1}\mathcal{B}_n\right] = \sum_{j=0}^{\infty} \alpha^j \mathbb{E}\left[\mathcal{B}_0\mathcal{B}_{j+1}\right]. \end{aligned} \tag{73}$$

because of the independence of the processes $\mathcal{D}_{i,1}(\cdot)$, $\mathcal{D}_{n,2}(\cdot)$, $\mathcal{N}_{i,1}(\cdot)$, and $\mathcal{N}_{i,2}(\cdot)$, for all $i \in \mathbb{Z}$. Writing out the last term yields

$$\begin{aligned}\mathbb{E}\left[\mathcal{B}_0\mathcal{B}_{j+1}\right] =& \mathbb{E}\Big[\Big(V_{1,1} + V_{1,2} + \mathcal{D}_{1,2}\Big(\mathcal{N}_{0,2}(V_{0,2}) + \mathcal{N}_{1,2}(V_{1,1})\Big)\Big) \\ & \cdot \Big(V_{j+2,1} + V_{j+2,2} + \mathcal{D}_{j+2,2}\Big(\mathcal{N}_{j+1,2}(V_{j+1,2}) + \mathcal{N}_{j+2,2}(V_{j+2,1})\Big)\Big)\Big] \\ =& v_1^2 + c_1(j+1) + v_1v_2 + c_{12}(j+1) + \lambda_2 d_2\Big(v_1v_2 + c_{12}(j) + v_1^2 + c_1(j+1)\Big) \\ & + v_1v_2 + c_{12}(-j-1) + v_2^2 + c_2(j+1) + \lambda_2 d_2\Big(v_2^2 + c_2(j) + v_1v_2 + c_{12}(-j-1)\Big) \\ & + \lambda_2 d_2\Big(v_1v_2 + c_{12}(-j-2) + v_2^2 + c_2(j+2) + \lambda_2 d_2\Big(v_2^2 + c_2(j+1) + v_1v_2 + c_{12}(-j-2)\Big)\Big) \\ & + \lambda_2 d_2\Big(v_1^2 + c_1(j+1) + v_1v_2 + c_{12}(j+1) + \lambda_2 d_2\Big(v_1v_2 + c_{12}(j) + v_1^2 + c_1(j+1)\Big)\Big) \\ =& \frac{R^2}{(1-\rho_2)^2} + \frac{c_1(j+1) + c_2(j+1)}{(1-\rho_2)^2} + \frac{\rho_2(c_2(j) - 2c_2(j+1) + c_2(j+2))}{1-\rho_2} \\ & + \frac{c_{12}(-j-1) + c_{12}(j+1)}{1-\rho_2} + \frac{\rho_2(c_{12}(-j-2) + c_{12}(j))}{(1-\rho_2)^2}. \end{aligned} \tag{74}$$

Putting (72)-(74) together and re-indexing the summation (for example, $\sum_{j=0}^{\infty} \alpha^{j+1} c_2(j) = \sum_{j=1}^{\infty} \alpha^j c_2(j-1) = \alpha c_2(0) + \sum_{j=1}^{\infty} \alpha^j \alpha c_2(j)$) produces

$$\begin{aligned}\mathbb{E}[\mathcal{A}_n(I_{n,1})\mathcal{B}_n] = \sum_{j=1}^{\infty} \Bigg[& \frac{R^2 + c_1(j) + c_2(j)}{(1-\rho_2)^2} + \frac{\rho_2\Big(\alpha - 2 + \frac{1}{\alpha}\Big)c_2(j)}{1-\rho_2} \\ & + \frac{c_{12}(-j)}{1-\rho_2}\Big(1 + \frac{\rho_2}{\alpha(1-\rho_2)}\Big) + \frac{c_{12}(j)}{1-\rho_2}\Big(1 + \frac{\alpha\rho_2}{1-\rho_2}\Big)\Bigg]\alpha^j \\ & + \frac{\rho_2(\alpha c_2(0) - c_2(1))}{1-\rho_2} + \frac{\rho_2(-c_{12}(-1) + \alpha c_{12}(0))}{(1-\rho_2)^2}. \end{aligned}$$

All of the terms with v_1 and v_2 can be pulled out of the summation and under stationary regime $c_2(0) = \delta_2^2$. This gives

$$\begin{aligned}\mathbb{E}[\mathcal{A}_n(I_{n,1})\mathcal{B}_n] = &\frac{\rho_1\rho_2 R^2}{(1-\rho_2)^2(1-\rho)} + \frac{\rho_1\rho_2^2\delta_2^2}{(1-\rho_1)(1-\rho_2)^2} - \frac{\rho_2 c_2(1)}{1-\rho_2} + \frac{\rho_2(-c_{12}(-1)+\alpha c_{12}(0))}{(1-\rho_2)^2}\\ &+ \frac{1}{(1-\rho_2)^2}\sum_{j=1}^{\infty}\left[c_1(j)+c_2(j)+\frac{(1-\rho)^2 c_2(j)}{\rho_1(1-\rho_1)}+\frac{(1-\rho_2)c_{12}(-j)}{\rho_1}\right]\alpha^j\\ &+ \frac{1-\rho_2(1-\alpha)}{(1-\rho_2)^2}\sum_{j=1}^{\infty} c_{12}(j)\alpha^j. \qquad (75)\end{aligned}$$

Putting equations (70),(71), and (75) into (16) and collecting terms gives

$$\begin{aligned}\mathbb{E}[I_{n+1,1}^2] = &\frac{\rho_1^2\rho_2^2\mathbb{E}[I_{n,1}^2]}{(1-\rho_1)^2(1-\rho_2)^2} + \frac{R}{(1-\rho_2)^2(1-\rho)}\left(\frac{\lambda_1\rho_2^2 b_1^{(2)}}{(1-\rho_1)^2} + \lambda_2 b_2^{(2)}\right)\\ &+ \frac{1}{(1-\rho_2)^2}\left(\Delta^2 - \frac{2\rho_2(1-\rho)\delta_2^2}{1-\rho_1} + \left(\frac{1-\rho+2\rho_1\rho_2}{1-\rho}\right)R^2\right)\\ &+ \frac{2(1-\rho_2(1-\alpha))}{(1-\rho_2)^2}\sum_{j=0}^{\infty} c_{12}(j)\alpha^j\\ &+ \frac{2}{(1-\rho_2)^2}\sum_{j=1}^{\infty}\left[c_1(j)+c_2(j)+\frac{(1-\rho)^2 c_2(j)}{\rho_1(1-\rho_1)}+\frac{(1-\rho_2)c_{12}(-j)}{\rho_1}\right]\alpha^j.\end{aligned}$$

Under stationary regime $\mathbb{E}[I_{n+1,1}^2] = \mathbb{E}[I_{n,1}^2]$. The Theorem follows by putting these terms on the same side and by making use of the identity

$$1 - \left(\frac{\rho_1}{1-\rho_1}\right)^2\left(\frac{\rho_2}{1-\rho_2}\right)^2 = \frac{(1-\rho)(1-\rho+2\rho_1\rho_2)}{(1-\rho_1)^2(1-\rho_2)^2}.$$ ∎

C Alternative proof of Theorem 3.4.

Proof: Under a number of assumptions [12] presents a decomposition theory which states that the expected waiting time in an M/G/1 queue with vacations can be decomposed into two parts, namely

$$\mathbb{E}[W_{q,i}] = \frac{\lambda_i b_i^{(2)}}{2(1-\rho_i)} + \mathbb{E}[\tilde{V}_i].$$

The first part is the Pollaczek-Khinchin formula for the expected waiting time in an ordinary M/G/1 queue without vacations. The second part is the forward-recurrence time defined as

$$\mathbb{E}[\tilde{V}_i] = \frac{\mathbb{E}[I_{n,i}^2]}{2\mathbb{E}[I_{n,i}]}.$$

See [9] for an application of this theory. Since a customer in a polling system arriving at queue i sees the system as an M/G/1 with vacations, and the theorem does not assume independence between vacations and service times (or busy periods), we can apply the theorem to our polling system. ∎

D Proof of Theorem 4.3.

Proof: First of all note that the expected service time of N customers at the second queue is the sum of their individual service times and hence $\mathbb{E}[S_{n,2}] = b_2\mathbb{E}[N]$. Taking the expectation over (34) then gives

$$\begin{aligned}\mathbb{E}[S_{n+1,2}] =& \rho_2\Big(\mathbb{E}[S_{n,2}] + R + \mathbb{E}\big[\mathcal{D}_{n+1,1}\big(\mathcal{N}_{n,1}(S_{n,2} + V_{n,1} + V_{n,2})\big)\big]\Big)\\ =& \frac{\rho_2}{1-\rho_1}\Big(\mathbb{E}[S_{n,2}] + R\Big),\end{aligned}$$

where the last line follows since $\mathbb{E}[D_{n,1}] = \rho_1\mathbb{E}[I_{n,1}])/(1-\rho_1)$ (equation (12a)) . Because of the stationarity ($\mathbb{E}[S_{n+1,2}] = \mathbb{E}[S_{n,2}]$), this gives the first moment, $\mathbb{E}[S_{n+1,2}] = \rho_2 R/(1-\rho)$.

Before we derive the second moment, we first need to point out (with the help of (4) and (6)) that

$$\begin{aligned}\mathbb{E}[\mathcal{S}_{n,2}^2(\mathcal{N}_{n,2}(T))] =& b_2^2\mathbb{E}[\mathcal{N}_{n,2}^2(T)] + (b_2^{(2)} - b_2^2)\mathbb{E}[\mathcal{N}_{n,2}(T)]\\ =& \rho_2^2\mathbb{E}[T^2] + \lambda_2 b_2^{(2)}\mathbb{E}[T].\end{aligned} \tag{76}$$

Squaring (34) gives the second moment of the service time of the second queue as

$$\mathbb{E}[S_{n+1,2}^2] = \mathbb{E}[\mathcal{X}_n^2(S_{n,2})] + \mathbb{E}[\mathcal{Y}_n^2] + 2\mathbb{E}[\mathcal{X}_n(S_{n,2})\mathcal{Y}_n]. \tag{77}$$

The right hand side will be solved piece by piece. First of all, using (76), and then (6) along with the independence of $\mathcal{N}_{n,2}(\cdot)$ and $\mathcal{N}_{n+1,2}(\cdot)$ gives

$$\begin{aligned}
\mathbb{E}[\mathcal{X}_n^2(S_{n,2})] =& \mathbb{E}\left[\mathcal{S}^2_{n+1,2}\Big(\mathcal{N}_{n,2}(S_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2})))\Big)\right] \\
=& b_2^2\mathbb{E}\left[\Big(\mathcal{N}_{n,2}(S_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2})))\Big)^2\right] \\
&+ (b_2^{(2)} - b_2^2)\mathbb{E}\left[\mathcal{N}_{n,2}(S_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2})))\right] \\
=& \lambda_2 b_2^2\Big(\lambda_2\mathbb{E}[S_{n,2}^2] + \mathbb{E}[S_{n,2}] + 2\lambda_2\mathbb{E}\left[S_{n,2}\cdot\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right] \\
&\quad + \lambda_2\mathbb{E}\left[\mathcal{D}^2_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right] + \mathbb{E}\left[\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right]\Big) \\
&+ \lambda_2(b_2^{(2)} - b_2^2)\mathbb{E}\left[S_{n,2} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right].
\end{aligned}$$

A couple of these terms can be crossed out to give

$$\begin{aligned}
\mathbb{E}[\mathcal{X}_n^2(S_{n,2})] =& \rho_2^2\Big(\mathbb{E}[S_{n,2}^2] + 2\mathbb{E}\left[S_{n,2}\cdot\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right] + \mathbb{E}\left[\mathcal{D}^2_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right]\Big) \\
&+ \lambda_2 b_2^{(2)}\mathbb{E}\left[S_{n,2} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))\right].
\end{aligned}$$

Applying formula (8) and then substituting $d_1 = b_1/(1-\rho_1)$ and $d_1^{(2)} = b_1^{(2)}/(1-\rho_1)^3$ (see (7)) leads to

$$\begin{aligned}
\mathbb{E}[\mathcal{X}_n^2(S_{n,2})] =& \rho_2^2\Big(\mathbb{E}[S_{n,2}^2]\Big(1 + \frac{2\rho_1}{1-\rho_1} + \lambda_1^2 d_1^2\Big) + \lambda_1 d_1^{(2)}\mathbb{E}[S_{n,2}]\Big) + \frac{\lambda_2 b_2^{(2)}}{1-\rho_1}\mathbb{E}[S_{n,2}] \\
=& \frac{\rho_2^2\mathbb{E}[S_{n,2}^2]}{(1-\rho_1)^2} + \frac{\mathbb{E}[S_{n,2}]}{1-\rho_1}\left(\frac{\lambda_1 b_1^{(2)}\rho_2^2}{(1-\rho_1)^2} + \lambda_2 b_2^{(2)}\right) \\
=& \frac{\rho_2^2\mathbb{E}[S_{n,2}^2]}{(1-\rho_1)^2} + \frac{\rho_2 R}{(1-\rho_1)(1-\rho)}\left(\frac{\lambda_1 b_1^{(2)}\rho_2^2}{(1-\rho_1)^2} + \lambda_2 b_2^{(2)}\right) \qquad (78)
\end{aligned}$$

The second part of (77) can be derived along similar lines. Taking the expectation of the square of (36),

$$
\begin{aligned}
\mathbb{E}[\mathcal{Y}_n^2] =&\mathbb{E}\Big[\mathcal{S}_{n+1,2}^2\Big(\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1})\Big)\Big] \\
=&b_2^2\mathbb{E}\Big[\Big(\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1})\Big)^2\Big] \\
&+ \Big(b_2^{(2)} - b_2^2\Big)\mathbb{E}\Big[\mathcal{N}_{n,2}(V_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1})\Big] \\
=&\lambda_2 b_2^2\Big(\lambda_2\mathbb{E}[V_{n,2}^2] + \mathbb{E}[V_{n,2}] + 2\lambda_2\mathbb{E}\Big[V_{n,2}\cdot\Big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\Big)\Big] \\
&\quad + \lambda_2\mathbb{E}\Big[\Big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\Big)^2\Big] \\
&\quad + \mathbb{E}\big[\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\big]\Big) \\
&+ \lambda_2\Big(b_2^{(2)} - b_2^2\Big)\mathbb{E}\Big[V_{n,2} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\Big].
\end{aligned}
$$

Also here terms can be crossed out[9]. Doing this gives

$$
\begin{aligned}
\mathbb{E}[\mathcal{Y}_n^2] =&\rho_2^2\Big(\mathbb{E}[V_{n,2}^2] + 2\mathbb{E}\Big[V_{n,2}\cdot\Big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\Big)\Big] \\
&+ \mathbb{E}\Big[\Big(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\Big)^2\Big]\Big) \\
&+ \lambda_2 b_2^{(2)}\mathbb{E}\Big[V_{n,2} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(V_{n,1}+V_{n,2})) + V_{n+1,1}\Big],
\end{aligned}
$$

after which the expectations can be moved farther back in the equation to give

$$
\begin{aligned}
\mathbb{E}[\mathcal{Y}_n^2] =&\rho_2^2\Big(\mathbb{E}[V_{n,2}^2] + \frac{2\rho_1}{1-\rho_1}\mathbb{E}\big[V_{n,1}V_{n,2} + V_{n,2}^2\big] + 2\mathbb{E}[V_{n+1,1}V_{n,2}] \\
&+ \lambda_1^2 d_1^2\mathbb{E}\big[(V_{n,1}+V_{n,2})^2\big] + \lambda_1 d_1^{(2)}\mathbb{E}[V_{n,1}+V_{n,2}] + \mathbb{E}[V_{n+1,1}^2] \\
&+ \frac{2\rho_1}{1-\rho_1}\mathbb{E}[V_{n,1}V_{n+1,1} + V_{n+1,1}V_{n,2}]\Big) + \frac{\lambda_2 b_2^{(2)} R}{1-\rho_1}.
\end{aligned}
$$

[9]The slightly more elaborate calculations and then the crossing out of terms is due to $N_{n,2}(\cdot)$ and $N_{n+1,2}(\cdot)$ being treated seperately. Although it is tempting to combine them into one term, it is better not to do so since the independence and correlations between the different processes $\mathcal{N}_{n,1}(T)$, $\mathcal{N}_{n,2}(T)$, $\mathcal{D}_{n+1,2}(T)$, $\mathcal{S}_{n,1}(T)$, and $\mathcal{S}_{n+1,1}(T)$ are not always obvious and have to be treated with care when squaring them.

In terms of the covariance functions $c_1(j)$, $c_2(j)$, and $c_{12}(\pm j)$, $j \in \mathbb{N}$ (see (3)),

$$\begin{aligned}
\mathbb{E}[\mathcal{Y}_n^2] =& \rho_2^2 \Big(v_2^{(2)} + \frac{2\rho_1(v_1 v_2 + c_{12}(0) + v_2^{(2)})}{1-\rho_1} + 2v_1v_2 + 2c_{12}(-1) \\
&+ \frac{\rho_1^2(v_1^{(2)} + v_2^{(2)} + 2v_1v_2 + c_{12}(0))}{(1-\rho_1)^2} + \frac{\lambda_1 b_1^{(2)} R}{(1-\rho_1)^3} + v_1^{(2)} \\
&+ \frac{2\rho_1(v_1^2 + c_1(1) + v_1v_2 + c_{12}(-1))}{1-\rho_1} \Big) + \frac{\lambda_2 b_2^{(2)} R}{1-\rho_1} \\
=& \frac{\rho_2^2}{1-\rho_1} \Big(\frac{\delta_1^2 + \delta_2^2 + R^2}{1-\rho_1} + 2\rho_1\delta_1^2 + \frac{\lambda_1 b_1^{(2)} R}{(1-\rho_1)^2} \Big) + \frac{\lambda_2 b_2^{(2)} R}{1-\rho_1} \\
&+ \frac{2\rho_2^2}{1-\rho_1} \Big(\rho_1 c_1(1) + c_{12}(-1) + \frac{\rho_1 c_{12}(0)}{1-\rho_1} \Big).
\end{aligned} \tag{79}$$

The third part of (77) is

$$\begin{aligned}
\mathbb{E}[\mathcal{X}_n(S_{n,2})\mathcal{Y}_n] =& \mathbb{E}\Big[\mathcal{S}_{n+1,2}\Big(\mathcal{N}_{n,2}(S_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))) \Big) \cdot \mathcal{Y}_n \Big] \\
=& \lambda_2 \mathbb{E}\Big[\Big(\mathcal{N}_{n,2}(S_{n,2}) + \mathcal{N}_{n+1,2}(\mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2}))) \Big) \cdot \mathcal{Y}_n \Big] \\
=& \rho_2 \mathbb{E}\Big[\Big(S_{n,2} + \mathcal{D}_{n+1,1}(\mathcal{N}_{n,1}(S_{n,2})) \Big) \cdot \mathcal{Y}_n \Big] = \gamma \mathbb{E}[S_{n,2}\mathcal{Y}_n]
\end{aligned} \tag{80}$$

where in each step the indepence and non-overlapping of $S_{n,2}$ and $V_{n,i}$ $(i = 1, 2)$ was used. Theorem 2 in [1] provides us with a mean to work out (80). The theorem states that

$$S_{n,2} = \sum_{j=0}^{\infty} \Big(\prod_{i=n-j}^{n-1} \mathcal{X}_i^{(n-j)} \Big) (\mathcal{Y}_{n-j-1}), \quad n \in \mathbb{Z},$$

where for each integer i, $\{\mathcal{X}_i^{(-j)}\}_j$ are independent of each other and have the same distribution as $\mathcal{X}_i(\cdot)$. To apply the theorem it is sufficient to have $\gamma < 1$, which turns out to be equivalent to $\rho < 1$, and that $\mathbb{E}[\mathcal{Y}_n] < \infty$ (see Lemma 1 in [1]). The latter indeed holds as $\mathbb{E}[\mathcal{Y}_n] = \frac{\rho_2 R}{1-\rho_1}$.

Using the theorem and because of the independence of the processes $\mathcal{D}_{n,1}(\cdot)$, $\mathcal{S}_{n,2}(\cdot)$, $\mathcal{N}_{n,1}(\cdot)$, and $\mathcal{N}_{n,2}(\cdot)$, for all $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
\mathbb{E}[S_{n,2}\mathcal{Y}_n] &= \sum_{j=0}^{\infty} \mathbb{E}\left[\prod_{i=n-j}^{n-1} \Big(\mathcal{S}_{i+1,2}\Big(\mathcal{N}_{i,2}\big(\mathcal{Y}_{n-j-1}\big) + \mathcal{N}_{i+1,2}\big(\mathcal{D}_{i+1,1}\big(\mathcal{N}_{i,1}(\mathcal{Y}_{n-j-1})\big)\big)\Big)\Big) \cdot \mathcal{Y}_n\right] \\
&= \sum_{j=0}^{\infty} b_2^j \mathbb{E}\left[\prod_{i=n-j}^{n-1} \Big(\mathcal{N}_{i,2}\big(\mathcal{Y}_{n-j-1}\big) + \mathcal{N}_{i+1,2}\big(\mathcal{D}_{i+1,1}\big(\mathcal{N}_{i,1}(\mathcal{Y}_{n-j-1})\big)\big)\Big) \cdot \mathcal{Y}_n\right] \\
&= \sum_{j=0}^{\infty} \rho_2^j \mathbb{E}\left[\prod_{i=n-j}^{n-1} \Big(\mathcal{Y}_{n-j-1} + \mathcal{D}_{i+1,1}\big(\mathcal{N}_{i,1}(\mathcal{Y}_{n-j-1})\big)\Big) \cdot \mathcal{Y}_n\right] \\
&= \sum_{j=0}^{\infty} \left(\frac{\rho_2}{1-\rho_1}\right)^j \mathbb{E}\left[\mathcal{Y}_{n-j-1}\mathcal{Y}_n\right] = \sum_{j=0}^{\infty} \gamma^j \mathbb{E}\left[\mathcal{Y}_0\mathcal{Y}_{j+1}\right]. \qquad (81)
\end{aligned}$$

with $\gamma := \frac{\rho_2}{1-\rho_1}$. Finally, writing out the last term of this expression yields

$$\begin{aligned}
\mathbb{E}\left[\mathcal{Y}_0\mathcal{Y}_{j+1}\right] =& \mathbb{E}\Big[\mathcal{S}_{1,2}\Big(\mathcal{N}_{0,2}\big(V_{0,2}\big) + \mathcal{N}_{1,2}\big(\mathcal{D}_{1,1}\big(\mathcal{N}_{0,1}(V_{0,1}+V_{0,2})\big) + V_{1,1}\big)\Big) \\
&\cdot \mathcal{S}_{j+2,2}\Big(\mathcal{N}_{j+1,2}\big(V_{j+1,2}\big) + \mathcal{N}_{j+2,2}\big(\mathcal{D}_{j+2,1}\big(\mathcal{N}_{j+1,1}(V_{j+1,1}+V_{j+1,2})\big) + V_{j+2,1}\big)\Big)\Big] \\
=& \rho_2^2 \mathbb{E}\Big[\Big(V_{0,2} + \mathcal{D}_{1,1}\big(\mathcal{N}_{0,1}(V_{0,1}+V_{0,2})\big) + V_{1,1}\Big) \\
&\cdot \Big(V_{j+1,2} + \mathcal{D}_{j+2,1}\big(\mathcal{N}_{j+1,1}(V_{j+1,1}+V_{j+1,2})\big) + V_{j+2,1}\Big)\Big] \\
=& \rho_2^2 \mathbb{E}\Big[V_{1,1}V_{j+2,1} + V_{0,2}V_{j+1,2} + V_{1,1}V_{j+1,2} + V_{j+2,1}V_{0,2} \\
&+ \frac{\rho_1}{1-\rho_1}\Big(V_{1,1}V_{j+1,1} + V_{0,1}V_{j+2,1} + 2V_{0,2}V_{j+1,2} \\
&\quad + V_{j+1,1}V_{0,2} + V_{0,1}V_{j+1,2} + V_{j+2,1}V_{0,2} + V_{1,1}V_{j+1,2}\Big) \\
&+ \frac{\rho_1^2}{(1-\rho_1)^2}\Big(V_{0,1}V_{j+1,1} + V_{0,2}V_{j+1,2} + V_{0,1}V_{j+1,2} + V_{j+1,1}V_{0,2}\Big)\Big].
\end{aligned}$$

Using the covariance functions (3)

$$\begin{aligned}
\mathbb{E}\,[\mathcal{Y}_0\mathcal{Y}_{j+1}] =& \rho_2^2\Big[v_1^2 + c_1(j{+}1) + v_2^2 + c_2(j{+}1) + 2v_1v_2 + c_{12}(j) + c_{12}(-j{-}2) \\
&+ \frac{\rho_1}{1-\rho_1}\Big(2v_1^2 + c_1(j) + c_1(j{+}2) + 2v_2^2 + 2c_2(j{+}1) \\
&\qquad + 4v_1v_2 + c_{12}(-j{-}1) + c_{12}(j{+}1) + c_{12}(-j{-}2) + c_{12}(j)\Big) \\
&+ \frac{\rho_1^2}{(1-\rho_1)^2}\Big(v_1^2 + c_1(j{+}1) + v_2^2 + c_2(j{+}1) + 2v_1v_2 + c_{12}(j{+}1) + c_{12}(-j{-}1)\Big)\Big] \\
=& \frac{\rho_2^2}{1-\rho_1}\Big[\frac{R^2}{1-\rho_1} + \frac{c_1(j{+}1) + c_2(j{+}1)}{1-\rho_1} + \rho_1(c_1(j) - 2c_1(j{+}1) + c_1(j{+}2)) \\
&+ c_{12}(-j{-}2) + c_{12}(j) + \frac{\rho_1(c_{12}(-j{-}1) + c_{12}(j{+}1))}{1-\rho_1}\Big].
\end{aligned} \tag{82}$$

Putting equations (81) and (82) into (80) and re-indexing the summation gives

$$\begin{aligned}
\mathbb{E}[\mathcal{X}_n(S_{n,2})\mathcal{Y}_n] =& \frac{\rho_2^2}{1-\rho_1}\sum_{j=1}^{\infty}\Big(\frac{R^2 + c_1(j) + c_2(j)}{1-\rho_1} + \rho_1(c_1(j{-}1) - 2c_1(j) + c_1(j{+}1)) \\
&+ c_{12}(-j{-}1) + c_{12}(j{-}1) + \frac{\rho_1(c_{12}(-j) + c_{12}(j))}{1-\rho_1}\Big)\gamma^j.
\end{aligned}$$

In this expression the term R^2 can be pulled out of the summation and the various covariance functions can be collected together (for example, $\sum_{j=1}^{\infty}\gamma^j c_2(j{-}1) = \gamma c_1(0) + \sum_{j=1}^{\infty}\gamma^j\gamma c_2(j)$) to reveal

$$\begin{aligned}
\mathbb{E}[\mathcal{X}_n(S_{n,2})\mathcal{Y}_n] =& \frac{\rho_2^2}{1-\rho_1}\Big[\frac{\rho_2 R^2}{(1-\rho_1)(1-\rho)} + \rho_1\big(\gamma c_1(0) - c_1(1)\big) - c_{12}(-1) + \gamma c_{12}(0) \\
&+ \sum_{j=1}^{\infty}\gamma^j\Big(\frac{c_1(j) + c_2(j)}{1-\rho_1} + \rho_1 c_1(j)\big(\gamma - 2 + \frac{1}{\gamma}\big) + c_{12}(-j)\big(\frac{1}{\gamma} + \frac{\rho_1}{1-\rho_1}\big) + c_{12}(j)\big(\gamma + \frac{\rho_1}{1-\rho_1}\big)\Big)\Big] \\
=& \frac{\rho_2^2}{1-\rho_1}\Big[\frac{\rho_2 R^2}{(1-\rho_1)(1-\rho)} + \rho_1\big(\gamma c_1(0) - c_1(1)\big) - c_{12}(-1) + \gamma c_{12}(0)\Big] \\
&+ \frac{\rho_2^2}{(1-\rho_1)^2}\sum_{j=1}^{\infty}\Big(c_1(j) + c_2(j) + \frac{\rho_1}{\rho_2}(1-\rho)^2 c_1(j) + \frac{1-\rho_1(2-\rho)}{\rho_2}c_{12}(-j) + \rho c_{12}(j)\Big)\gamma^j.
\end{aligned} \tag{83}$$

Putting equations (78), (79), and (83) into (77), using $c_1(0) = v_1^{(2)} - v_1^2$, and collecting terms gives

$$
\begin{aligned}
\mathbb{E}[S_{n+1,2}^2] = {} & \frac{\rho_2^2 \mathbb{E}[S_{n,2}^2]}{(1-\rho_1)^2} + \frac{\rho_2 R}{(1-\rho_1)(1-\rho)} \left(\frac{\lambda_1 \rho_2^2 b_1^{(2)}}{(1-\rho_1)^2} + \lambda_2 b_2^{(2)} \right) \\
& + \frac{\rho_2^2}{1-\rho_1} \left(\frac{\delta_1^2 + \delta_2^2 + R^2}{1-\rho_1} + 2\rho_1 \delta_1^2 + \frac{\lambda_1 b_1^{(2)} R}{(1-\rho_1)^2} \right) + \frac{\lambda_2 b_2^{(2)} R}{1-\rho_1} \\
& + \frac{2\rho_2^2}{1-\rho_1} \left(\rho_1 c_1(1) + c_{12}(-1) + \frac{\rho_1 c_{12}(0)}{1-\rho_1} \right) \\
& + \frac{2\rho_2^2}{1-\rho_1} \left[\frac{\rho_2 R^2}{(1-\rho_1)(1-\rho)} + \rho_1 \big(\gamma \delta_1^2 - c_1(1) \big) - c_{12}(-1) + \gamma c_{12}(0) \right] \\
& + \frac{2\rho_2^2}{(1-\rho_1)^2} \sum_{j=1}^{\infty} \gamma^j \left(c_2(j) + \frac{1-\rho_1(2-\rho)}{\rho_2} \Big(c_{12}(-j) + \rho c_1(j) \Big) + \rho c_{12}(j) \right) \\
= {} & \frac{\rho_2^2 \mathbb{E}[S_{n,2}^2]}{(1-\rho_1)^2} + \frac{R}{1-\rho} \left(\frac{\lambda_1 \rho_2^2 b_1^{(2)}}{(1-\rho_1)^2} + \lambda_2 b_2^{(2)} \right) \\
& + \frac{\rho_2^2}{(1-\rho_1)^2} \left(\big(1 - 2\rho_1(1-\rho)\big) \delta_1^2 + \delta_2^2 + \left(\frac{1-\rho_1+\rho_2}{1-\rho} \right) R^2 + 2\rho c_{12}(0) \right) \\
& + \frac{2\rho_2^2}{(1-\rho_1)^2} \sum_{j=1}^{\infty} \left(c_2(j) + \frac{1-\rho_1(2-\rho)}{\rho_2} \Big(c_{12}(-j) + \rho c_1(j) \Big) + \rho c_{12}(j) \right) \gamma^j .
\end{aligned}
$$

The final expression is obtained by assuming stationary, putting the terms $\mathbb{E}[S_{n+1,2}^2] = \mathbb{E}[S_{n,2}^2]$ on the same side, and multiplying both sides by $(1-\rho_1)^2/\rho_2^2$. ∎

References

[1] Eitan Altman. Stochastic recursive equations with applications to queues with dependent vacations. *Annals of Operations Research*, 112(1):43–61, 2002.

[2] Eitan Altman. On stochastic recursive equations and infinite server queues. Technical Report RR-5211, INRIA, June 2004.

[3] Eitan Altman, Tanja Jimenez, and Daniel Kofman. DPS queues with stationary ergodic service times and the performance of TCP in overload. In *Proceedings of IEEE Infocom*, Hong-Kong, March 2004.

[4] Eitan Altman, Panagiotis Konstantopoulos, and Zhen Liu. Stability, monoticity and invariant quantities in general polling systems. *Queuing Systems*, 11:35–57, 1992. Special issue on 'Polling Systems'.

[5] Eiten Altman, Chadi Barakat, and Konstantin Avratchenkov. A stochastic model of TCP/IP with stationary ergodic random losses. In *ACM-Sigcomm*, Aug. 28 - Sept. 1 2000. See also INRIA Research Report RR-3824.

[6] Thomas Bonald, M. Jonckheere, and A. Proutière. Insensitive load balancing. In *ACM Sigmetrics*, 2004.

[7] Thomas Bonald, A. Proutière, G. Régnié, and J.W. Roberts. Insensitivity results in statistical bandwidth sharing. In *ITC*, number 17, Salvador, Brazil, 2001.

[8] Leo Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. SIAM, Philadelphia, 1992.

[9] Robert B. Cooper, Shun-Chen Niu, and Mandyam M Srinivasan. A decomposition theorem for polling models: The switchover times are effectively additive. *Operations Researchs*, 44(4):629–633, Jul-Aug 1996.

[10] Richard Durrett. *Probability: Theory and Examples*. Wadsworth & Brooks/Cole Statistics/Probability Series. Brooks/Cole, 1991.

[11] Martin Eisenberg. Queues with periodic service and changeover time. *Operations Research*, 20(2):440–451, March-April 1972.

[12] S. W. Fuhrmann and Robert B. Cooper. Stochastic decompositions in the M/G/1 queue with generalized vacations. *Operations Research*, 33(5):1117–1129, 1985.

[13] J.D.C.Little. A proof for the queueing formula: $L = \lambda W$. *Operations Research*, 9:383–387, May-June 1961.

[14] Alain Jean-Marie and Philippe Robert. On the transient behaviour of the processor sharing queue. *QUESTA*, 17(1-2):129–136, 1994.

[15] Leonard Kleinrock. *Queueing Systems, Volume I: Theory*, volume I. John Wiley and Sons, 1976.

[16] T. Ozawa. An analysis for multi-queueing systems with cyclic-service discipline - models with exhaustive and gated services. Technical Report 4, The Institute of Electronics, Information and Communication Engineers (IEICE) (in Japanese), 1987.

[17] Jack S. Sykes. Simplified analysis of an alternating-priority queueing model with setup times. *Operations Research*, 18(6):1182–1192, November-December 1970.

[18] Hideaki Takagi. *Analysis of Polling Systems*. MIT Press, Cambridge, Massachusetts, 1986.

[19] Hideaki Takagi. Queueing analysis of polling models: An update. *Stochastic Analysis of Computer and Communication Systems*, pages 267–318, 1990.

[20] Hideaki Takagi. *Queueing Analysis, Vacations and Priority Systems, Part 1*, volume 1. Elsevier Science Publishers B.V., The Netherlands, 1991.

[21] Gil Zussman, Uri Yechiali, and Adrian Segall. Exact probabilistic analysis of the limited scheduling algorithm for symmetrical bluetooth piconets. In *Personal Wireless Communications (PWC)*, Venice, Italy, Sept. 23-25 2003.

E List of Notations

$\mathcal{A}_n(\cdot)$ = A nested combination of stochastic processes defined in (14).
$\mathcal{B}_n$ = A nested combination of stochastic processes defined in (15).
$\mathcal{D}_{n,i}(N)$ = Total busy period generated by N customers in queue i with arrival rate λ_i and first and second moment of the service time b_i and $b_i^{(2)}$, respectively.
$\mathcal{D}_{n,i,k}(N)$ = Single busy period generated by the k-th customer in queue i with arrival rate λ_i and first and second moment of the service time b_i and $b_i^{(2)}$, respectively.
$\mathcal{S}_{n,i}(N)$ = Service time of N customers at queue i with first and second moment of the service time b_i and $b_i^{(2)}$, respectively.
$\mathcal{N}_{n,i}(T)$ = Number of arrivals at queue i in time T in the n^{th} cycle (cycle starting from the polling instant of the first queue).
$\dot{\mathcal{N}}_{n,2}(\cdot)$ = The number of arrivals at the second queue during the n^{th} cycle, with the cycle starting from the arrival of the server at the second queue.
$\mathcal{T}_{n,i}(N)$ = The number of customers served at queue i during the n^{th} cycle if there are N customers in the queue at the moment of polling.
$\mathcal{X}_n(\cdot)$ = A nested combination of stochastic processes defined in (35).
$\mathcal{Y}_n$ = A nested combination of stochastic processes defined in (36).

$C_{n,i}$ = Duration of the n^{th} cycle starting from the polling instant of the i^{th} queue.
$D_{n,i}$ = Duration of the busy period at queue i in the n^{th} cycle.
$I_{n,i}$ = Intervisit of the i^{th} queue in the n^{th} cycle. This is the time between the server switching away from queue i until the time that the server comes back to queue i. ($I_{n,1} = V_{n,1} + D_{n,2} + V_{n,2}$ for exhaustive/exhaustive queues and $I_{n,1} = V_{n,1} + S_{n,2} + V_{n,2}$ for exhaustive/gated queues).
$L^*_{n,i}$ = Number of customers in queue i in the n^{th} cycle at the moment the queue is polled.
$L_{q,i}$ = Average number of customers in queue i. This is also the number of customers that arrived at queue i during a vacation, and are still in the queue, before a tag customer arrived in that same vacation.
$L_{s,i}$ = Average number of customers at queue i (including the customer in service). This is also the number of customers that arrived at queue i during a vacation (and are in the queue or in service) before a tag customer arrived in that same vacation.
$S_{n,i}$ = Service time at queue i in the n^{th} cycle (=similar to the duration $D_{n,i}$ of the busy period but then with no arrivals).
$T_{n,i}$ = The number of customers served, per cycle, at queue i.
$V_{n,i}$ = Switching time from queue i to the other queue in the n^{th} cycle.
$W_{q,i}$ = Random variable for the waiting time of a customer in queue i (not including service).
$W_{s,i}$ = Random variable for the total sojourn time of a customer at queue i (waiting time plus service time).
Γ_i = The number of customers served during a busy period, where the arrival rate is λ_i, average service time is b_i, and the second moment of the service time is $b_i^{(2)}$.

$\alpha = \frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)}$ = Central quantity in the exhaustive/exhaustive queueing system. Comes forth from $\mathbb{E}[\mathcal{A}_n(I)] = \alpha\mathbb{E}[I]$, with $\mathcal{A}_n(\cdot)$ defined in (14).

$b_i = \mathbb{E}[B_i]$ = Expected service time at queue i.

$b_i^{(2)} = \mathbb{E}[B_i]^2$ = Second moment of the service time at queue i.

$d_i = \rho_i/(1-\rho_i)$ = The expected duration of a single busy period, where the arrival rate is λ_i and the average service time is b_i.

$d_i^{(2)} = b_i^{(2)}/(1-\rho_i)^3$ = Second moment of the duration of a single busy period, where the arrival rate is λ_i, the average service time is b_i, and the second moment of the service time is $b_i^{(2)}$.

$\delta_i^2 = v_i^{(2)} - v_i^2$ = Variance of the switching time from queue i to the other queue.

$\gamma = \frac{\rho_2}{1-\rho_1}$ = Central quantity in the exhaustive/gated queueing system. Comes forth from $\mathbb{E}[\mathcal{X}_n(I)] = \gamma\mathbb{E}[I]$, with $\mathcal{X}_n(\cdot)$ defined in (36).

λ_i = (Poisson) arrival rate at queue i.

$\rho_i = \lambda_i d_i$ = load at queue i.

$\rho = \rho_1 + \rho_2$ = load of the system. The assumption is made throughout that $\rho < 1/2$.

$\hat{\rho} = \rho_1 = \rho_2$ if the parameter settings for the two queues are equal.

$R = v_1 + v_2$.

$v_i = E[V_{n,i}]$ = Expected switching time from queue i to the other queue.

$v_i^{(2)} = E[V_{n,i}^2]$ = Expected second moment of the switching time from queue i to the other queue.

$c_i(n) = \mathbb{E}[V_{0,i}V_{n,i}] - \mathbb{E}[V_{0,i}]\mathbb{E}[V_{n,i}]$ = Covariance function for the vacation sequences at queue i.

$c_{12}(n) = \mathbb{E}[V_{0,1}V_{n,2}] - \mathbb{E}[V_{0,1}]\mathbb{E}[V_{n,2}]$ = Covariance function for the vacation sequences between the two queues.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399