The two-queue E/1-L polling model with regularly varying service and/or switchover times

The two-queue $E/1$-L polling model with regularly varying service and/or switchover times

Q. Deng

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

November 12, 2001

Abstract

We consider the cyclic polling system with two queues. One queue is served according to the exhaustive discipline, and the other queue is served according to the 1-limited discipline. At least one of the service and/or switchover times has a regularly varying tail. We obtain the tail behavior of the waiting time distributions. When one of the service and/or switchover times has an infinite second moment, we derive the heavy-traffic behavior of the waiting time distribution at the 1-limited queue.

1991 Mathematical Subject Classification: 60K25, 90B22.
Keywords and Phrases: polling system, exhaustive service, 1-limited service, waiting time distribution, regularly varying distribution, heavy traffic.

1 Introduction

Cyclic polling systems are queueing systems in which a single server visits several queues in cyclic order. They have a wide range of applications in, e.g., computer communications, manufacturing and road traffic. The abundant literature on polling systems (see [18, 19]) contains the exact analysis of polling systems for a large number of service disciplines, like the 1-limited and exhaustive disciplines.

In this paper, we study the following polling system. The system consists of 2 queues, Q_1 and Q_2, attended by a single server S. Customers arrive at Q_k, $k = 1, 2$, according to a Poisson process with rate λ_k and require a generally distributed service time B_k having distribution function $B_k(\cdot)$, finite first moment β_k and LST (Laplace-Stieltjes Transform) $\beta_k(\cdot)$. In the sequel, customers arriving at Q_k are also referred to as type-k customers. The server visits the queues in a strictly cyclic order, i.e., $Q_1, Q_2, Q_1, \ldots$. The service policy is exhaustive service at Q_1 and 1-limited service at Q_2. In exhaustive service, the server continues to work at a queue until it becomes empty. In 1-limited service, the server serves at most one customer at a queue before switching to the other queue. When moving from Q_k to $Q_{(k \bmod 2)+1}$, the server incurs a generally distributed switchover time S_k, having distribution function $S_k(\cdot)$ with finite first moment σ_k and LST $\sigma_k(\cdot)$. The server continues switching even when the whole system is empty. The various interarrival, service and switchover times are independent.

Let us denote by $\lambda := \lambda_1 + \lambda_2$ the total arrival intensity of customers, by $\rho_k := \lambda_k \beta_k$ the traffic load at Q_k, by $\rho := \rho_1 + \rho_2$ the total traffic load, and by $\sigma := \sigma_1 + \sigma_2$ the mean of the total switchover time in one cycle.

The stability condition of this system is: $\rho + \lambda_2 \sigma < 1$ [13]. We assume this condition to hold in the remainder of this paper.

In the above-described model, we assume that at least one of the service and/or switchover time distributions has a regularly varying tail of index $-\nu$ $(1 < \nu < 2)$, and thus at least one of the service and/or switchover times does not have a finite second moment. A typical example of the class of regularly varying distributions is a power-tailed distribution like the Pareto distribution: $F(t) = 1 - (\frac{\theta}{\theta+t})^{\nu}$ $(\theta > 0)$, if $1 < \nu < 2$. The motivation of this study comes from the measurements in high-speed telecommunication networks (cf. [2, 20]), which exhibit the property of self-similarity and long-range dependence. As has been pointed out in e.g. [4, 8], fluid or ordinary queues with heavy-tailed input distributions are useful and tractable models for analyzing the effect of self-similar traffic on system performance.

In view of the central role of polling in computer-communication networks and the often observed occurrence of heavy-tailed traffic in computer communications, it is of importance to study the effect of heavy-tailed service and/or switchover time distributions on the waiting time tail behavior in polling systems. In this paper, we have two goals: (i) we want to determine the waiting time tail behavior, and (ii) we want to derive the heavy-traffic behavior of the waiting time distribution for the above-described model with nonzero switchover times.

A reason for studying this particular polling model is that it is a natural extension of the classical 2-class priority model: if there are no switchover times, then this particular model reduces to the 2-class priority model. Another reason is the tractability of this particular polling model which allows us to obtain detailed insight into the effect of heavy-tailed input characteristics on system performance.

We now describe some related work. The heavy-traffic behavior of the $M/G/1$ queue with priorities has been studied by Boxma et al. [6]. A cyclic polling model with gated or exhaustive service is studied by Boxma et al. [7]. They derived the tail behavior of the waiting time distribution when at least one of the service and/or switchover time distributions is regularly varying or index $-\nu$ $(\nu > 1)$.

Next we describe two classical heavy-traffic limit theorems, since we shall consider the waiting time distribution in a heavy-traffic situation. For the $G/G/1$ queue, we introduce some notation first. Denote by λ the arrival rate, by β and β_2 the first and second moment of the service time, and by $\rho := \lambda\beta$ the traffic load. When both service time and interarrival time distributions have a finite second moment, a standard heavy-traffic limit theorem for the stationary waiting time W in the $G/G/1$ queue holds (cf. [15]):

$$\lim_{\rho \uparrow 1} \mathbf{P}(\delta(\rho) W \leq t) = 1 - e^{-t},$$

where $\delta(\rho) := 2\lambda(1-\rho)/[1+\lambda^2(\beta_2 - \beta^2)]$. Boxma and Cohen [5] generalize the above results to the $G/G/1$ queue for which the service times may have *infinite* second moments. Their result is as follows: Assume that the tail of the service time distribution is regularly varying of index $-\nu$ $(1 < \nu < 2)$ and the tail of the interarrival time distribution is less heavy than that of the service time distribution. Then the distribution of the contracted waiting time $\Delta(\rho)W$ converges for $\rho \uparrow 1$ to the Mittag-Leffler distribution $R_{\nu-1}(t)$, which is specified by:

$$\int_0^{\infty} e^{-st} \mathrm{d}R_{\nu-1}(t) = \frac{1}{1+s^{\nu-1}}.$$

The coefficient of contraction $\Delta(\rho)$ is the unique solution to a contraction equation with the property that $\Delta(\rho) \downarrow 0$ for $\rho \uparrow 1$. In the case that the service time distribution is Pareto, the coefficient of contraction is given by $\Delta(\rho) = C(1-\rho)^{\frac{1}{\nu-1}}$ where C is a constant.

The model with nonzero switchover times has been studied by Groenendijk [13] and Ibe [14]. They derived the explicit LSTs of the waiting time distributions, which is the basis of the results in the present paper.

This paper is organized in the following way. In Section 2, we give some preliminary results. Based on these results, we investigate in Section 3 the tail asymptotics of the waiting times at both queues when at least one of the service and/or switchover times has a regularly varying tail. Section 4 proves a heavy-traffic limit theorem for the waiting time at Q_2 when at least one of the service and/or switchover times has a regularly varying distribution with an *infinite* second moment. Finally we show some numerical results in Section 5 to test the accuracy of the approximation for the waiting time distribution at Q_2 suggested by the heavy-traffic limit theorem.

2 Preliminaries

In this section we first present three lemmas, which simplify the proofs for the main results later on. Next we introduce the expressions for the LSTs of the waiting time distributions, cf. [13]. Finally we make some assumptions on the service and switchover times.

Let us start with the lemmas. Suppose $f(s)$ is the LST of some non-negative random variable X with finite moment ϕ_k of order k for $k = 0, ..., n$. We define

$$f_n(s) := (-1)^{n+1}\left(f(s) - \sum_{j=0}^{n} \phi_j \frac{(-s)^j}{j!}\right) = \mathrm{o}(s^n),\ \ s \downarrow 0. \tag{2.1}$$

The following lemma (cf. Lemma 2.2 in [8]), which is an extension of Theorem 8.1.6 (it is a special case of Karamata's Tauberian Theorem) in [3], links the regularly varying tail behavior of $\mathbf{P}(X > t)$ for $t \to \infty$ to the behavior of its LST $f(s)$.

Lemma 2.1 *Let X be a random variable with LST $f(s)$, $L(\cdot)$ a slowly varying function, $\nu \in (n, n+1)$ $(n \in \mathbb{N})$ and $C \geq 0$. Then the following two statements are equivalent:*
(i) $\mathbf{P}(X > t) = [C + \mathrm{o}(1)]L(t)/t^{\nu},\ t \to \infty.$
(ii) $\mathrm{E}[X^n] < \infty$ *and* $f_n(s) = (-1)^n\Gamma(1-\nu)[C + \mathrm{o}(1)]L(1/s)s^{\nu},\ s \downarrow 0$*, where $f_n(s)$ is defined by (2.1).*

Here $\Gamma(\cdot)$ denotes the Gamma function. The following lemma (cf. Lemma 7.7 in [8]), which can be derived easily from Karamata's Theorem and the Monotone Density Theorem (cf. [3] Sections 1.5.6 and 1.7.3), shows the equivalence between the tail behavior of X and the tail behavior of X^{res}. In this paper we follow the convention that X^{res} stands for the residual lifetime of X which has density function $(1 - X(t))/\mathrm{E}X$.

Lemma 2.2 *For all $\nu > 0$, X^{res} has a regularly varying tail of index $1 - \nu$ if and only if X has a regularly varying tail of index $-\nu$, and if either is the case then:*

$$\mathbf{P}(X^{res} > t) \sim \frac{t}{(\nu - 1)\mathrm{E}X}\mathbf{P}(X > t), \quad \text{as } t \to \infty.$$

Because some key formulas of this paper involve iterated functions, the following lemma (cf. [7]) is useful in this respect.

Lemma 2.3 *Suppose $\phi(\cdot)$, $\psi(\cdot)$ can be written as*

$$\phi(x) = \sum_{i=1}^{n} \phi_i x^i + \phi_\nu x^\nu L(1/x) + \mathrm{o}(x^\nu L(1/x)), \quad \textit{for } x \downarrow 0, \tag{2.2}$$

$$\psi(x) = \sum_{i=1}^{n} \psi_i x^i + \psi_\nu x^\nu L(1/x) + \mathrm{o}(x^\nu L(1/x)), \quad \textit{for } x \downarrow 0, \tag{2.3}$$

where $\phi_1, \psi_1 > 0$, $n < \nu < n+1$ ($n \in \mathbb{N}$), $\phi_i, \psi_i < \infty$ for $i = 1, ..., n$ and $L(\cdot)$ is a slowly varying function. Then the asymptotic expansion of the function $\phi(\psi(x))$ at point 0 is given by

$$\phi(\psi((x)) = \sum_{i=1}^{n} \theta_i x^i + (\phi_1\psi_\nu + \phi_\nu \psi_1^\nu) x^\nu L(1/x) + \mathrm{o}(x^\nu L(1/x)), \quad \textit{for } x \downarrow 0,$$

where $\theta_i < \infty$ for $i = 1, ..., n$.

Next we introduce the expressions for the LSTs of the waiting time distributions, cf. [13]. Let us give some notation first. Denote by $\eta_1(\cdot)$ the LST of the length of the busy period at Q_1 starting with one customer. Let W_k be the steady-state waiting time at Q_k with distribution function $W_k(\cdot)$ and LST $\omega_k(\cdot)$ for $k = 1, 2$. From (6.70) in [13], $\omega_k(s)$ ($k = 1, 2$) are, for Re $s \geq 0$, given by

$$\begin{aligned}
\omega_1(s) &= \frac{\sigma_1(s)\sigma_2(s)\beta_2(s) - 1}{\lambda_1 - s - \lambda_1\beta_1(s)} \frac{1-\rho}{\sigma} \\
&\quad + \frac{1-\rho-\lambda_2\sigma}{\sigma} \frac{\sigma_1(\lambda_2+s)\sigma_2(s)}{\sigma_1(\lambda_2)} \frac{1-\beta_2(s)}{\lambda_1 - s - \lambda_1\beta_1(s)}, \qquad (2.4)\\
\omega_2(s) &= \frac{1-\rho-\lambda_2\sigma}{\sigma} \sigma_1(s) \frac{\sigma_1(\lambda_2 + \lambda_1(1-\eta_1(s)))}{\sigma_1(\lambda_2)} \sigma_2(f(s)) \\
&\quad \frac{\lambda_2 - s - \lambda_2\beta_2(f(s))}{\lambda_2 - s - \lambda_2\sigma_1(f(s))\sigma_2(f(s))\beta_2(f(s))} \frac{1}{\lambda_2 - s} \\
&\quad - \frac{1}{\lambda_2 - s} \frac{1-\rho-\lambda_2\sigma}{\sigma}, \qquad (2.5)
\end{aligned}$$

where

$$f(s) := s + \lambda_1(1 - \eta_1(s)).$$

Concerning the service and switchover times, we are only interested in the properties of their tail behavior, i.e., the behavior of $1 - B_k(t)$ and $1 - S_k(t)$ for $t \to \infty$. We assume the following holds:

Assumption 2.1 *For the service and switchover time distributions, we have:*

$$1 - B_k(t) = [b_k + \mathrm{o}(1)]t^{-\nu}L(t), \quad t \to \infty, \qquad (2.6)$$

$$1 - S_k(t) = [s_k + \mathrm{o}(1)]t^{-\nu}L(t), \quad t \to \infty, \qquad (2.7)$$

where $b_k, s_k \geq 0$, $L(\cdot)$ is a slowly varying function and $k = 1, 2$. We assume $\sum_{k=1}^{2}(s_k + b_k) > 0$, i.e., at least one of the service and/or switchover times has a regularly varying tail of index $-\nu$.

Here a slowly varying function $L(\cdot)$ is a function which satisfies that $\lim_{t\to\infty} L(at)/L(t) = 1$ for any $a > 0$. For ease of presentation, we take the same function $L(\cdot)$ for all distributions, but one can easily change this into different slowly varying functions for different distributions. Note that the possibility that $b_k = 0$ or $s_k = 0$ implies that we do allow the possibility that some of the service and switchover time distributions have an exponential tail, or regularly varying of index strictly smaller than $-\nu$. According to Lemma 2.1, the tail behavior of the service and switchover time distributions as given in (2.6) and (2.7) is equivalent with the following behavior of their LSTs $\beta_k(s)$ (of the service time distributions) and $\sigma_k(s)$ (of the switchover time distributions):

$$1-\beta_k(s) = \sum_{j=1}^{m}(-1)^{j+1}\beta_{k,j}s^j + (-1)^m\beta_{k,\nu}s^\nu L(1/s) + \mathrm{o}(s^\nu L(1/s)), \tag{2.8}$$

$$1-\sigma_k(s) = \sum_{j=1}^{m}(-1)^{j+1}\sigma_{k,j}s^j + (-1)^m\sigma_{k,\nu}s^\nu L(1/s) + \mathrm{o}(s^\nu L(1/s)), \tag{2.9}$$

where $m < \nu < m+1$ ($m \in \mathbb{N}$), $\beta_{k,j} > 0$ and $\sigma_{k,j} > 0$ for $j = 1,...,m$, $k = 1,2$. Note that $\beta_{k,1} = \beta_k$, $\beta_{k,\nu} = (-1)^m\Gamma(1-\nu)b_k$, $\sigma_{k,1} = \sigma_k$, and $\sigma_{k,\nu} = (-1)^m\Gamma(1-\nu)s_k$ for $k = 1,2$.

It follows from the main result of [16] that the asymptotic behavior of the LST $\eta_k(s)$ of the length of the busy period in the 'corresponding' isolated $M/G/1$ queue of Q_k is given by

$$1-\eta_k(s) = \sum_{j=1}^{m}(-1)^{j+1}\eta_{k,j}s^j + (-1)^m\eta_{k,\nu}s^\nu L(1/s) + \mathrm{o}(s^\nu L(1/s)), \tag{2.10}$$

where $\eta_{k,1} = \beta_k/(1-\rho_k)$ and $\eta_{k,\nu} = \beta_{k,\nu}/(1-\rho_k)^{\nu+1}$ and $\eta_{k,j} > 0$ for $j = 1,...,m$, $k = 1,2$. Here the 'corresponding' isolated $M/G/1$ queue of Q_k stands for the single-server queue with the same arrival rate and service time distributions as Q_k.

3 The tail behavior of the waiting time distributions

In this section we derive the asymptotic behavior of the waiting times when at least one of the service and/or switchover times has a regularly varying tail.

Let us first consider the asymptotic expansions of the functions $\sigma_1(f(s))$, $\sigma_2(f(s))$ and $\beta_2(f(s))$ in the neighborhood of the origin, because these functions appear in (2.5). By using Lemma 2.3, we immediately get, for $k = 1,2$,

$$\begin{aligned}\sigma_k(f(s)) &= 1 + \sum_{j=1}^{m} g_{k,j}s^j + (-1)^{m+1}\left(\frac{\lambda_1\beta_{1,\nu}\sigma_k}{(1-\rho_1)^{\nu+1}} + \frac{\sigma_{k,\nu}}{(1-\rho_1)^{\nu}}\right)s^\nu L(1/s) \\ &\quad + \mathrm{o}(s^\nu L(1/s)), \quad s \downarrow 0, \end{aligned} \tag{3.1}$$

$$\begin{aligned}\beta_2(f(s)) &= 1 + \sum_{j=1}^{m} g_{3,j}s^j + (-1)^{m+1}\left(\frac{\lambda_1\beta_{1,\nu}\beta_2}{(1-\rho_1)^{\nu+1}} + \frac{\beta_{2,\nu}}{(1-\rho_1)^{\nu}}\right)s^\nu L(1/s) \\ &\quad + \mathrm{o}(s^\nu L(1/s)), \quad s \downarrow 0, \end{aligned} \tag{3.2}$$

where $g_{k,j}$ ($k = 1,2,3,\ j = 1,...,m$) are some constants. Moreover, for $k = 1,2$, $g_{k,1} = -\dfrac{\sigma_k}{1-\rho_1}$ and $g_{3,1} = -\dfrac{\beta_2}{1-\rho_1}$.

Again, applying Theorem 2.3 to (2.4) and (2.5), straightforward calculations lead to

$$\begin{aligned}
\omega_1(s) &= 1+\sum_{j=1}^{m-1}(-1)^j\omega_{1,j}s^j+(-1)^m\left(\frac{\lambda_1\beta_{1,\nu}+\lambda_2\beta_{2,\nu}}{1-\rho_1}+\frac{(1-\rho)(\sigma_{1,\nu}+\sigma_{2,\nu})}{\sigma(1-\rho_1)}\right)s^{\nu-1}L(1/s)\\
&\quad +\mathrm{o}(s^{\nu-1}L(1/s)), \quad s\downarrow 0, \qquad (3.3)\\
\omega_2(s) &= 1+\sum_{j=1}^{m-1}(-1)^j\omega_{2,j}s^j\\
&\quad +(-1)^m\left[\frac{1-\rho}{1-\rho-\lambda_2\sigma}\left(\frac{\lambda_1\beta_{1,\nu}(\rho_2+\lambda_2\sigma)}{(1-\rho_1)^{\nu+1}}+\frac{\lambda_2(\sigma_{1,\nu}+\sigma_{2,\nu}+\beta_{2,\nu})}{(1-\rho_1)^{\nu}}\right)\right.\\
&\quad \left.+\frac{\lambda_1\beta_{1,\nu}\rho_2}{(1-\rho_1)^{\nu+1}}+\frac{\lambda_2\beta_{2,\nu}}{(1-\rho_1)^{\nu}}\right]s^{\nu-1}L(1/s)+\mathrm{o}(s^{\nu-1}L(1/s)), \quad s\downarrow 0, \qquad (3.4)
\end{aligned}$$

where $j!\omega_{k,j}$ $(k=1,2,\ j=1,...,m-1)$ equals the jth moment of the waiting time W_k. Applying Lemma 2.1 to (3.3) and (3.4), we then get the following theorem.

Theorem 3.1 *If Assumption 2.1 holds, then the waiting times at both queues have a regularly varying tail of index which is one higher than the heaviest of the service and switchover times. In particular, we have*

$$\begin{aligned}
1-W_1(t) &\sim \frac{1}{\nu-1}\left(\frac{\lambda_1 b_1+\lambda_2 b_2}{1-\rho_1}+\frac{(1-\rho)(s_1+s_2)}{(1-\rho_1)\sigma}\right)t^{1-\nu}L(t), \quad t\to\infty,\\
1-W_2(t) &\sim \frac{1}{\nu-1}\left(\frac{\lambda_1 b_1+\lambda_2(s_1+s_2+b_2)}{(1-\rho_1)^{\nu-1}(1-\rho-\lambda_2\sigma)}+\frac{s_1+s_2}{(1-\rho_1)^{\nu-1}\sigma}\right)t^{1-\nu}L(t),\\
&\qquad t\to\infty.
\end{aligned}$$

We now relate the waiting time distribution to the residual service and switchover time distributions. Applying Lemma 2.2 to (2.6) and (2.7), we obtain the asymptotic behavior of the residual service times B_k^{res} and the residual switchover times S_k^{res}. For $k=1,2$, if $b_k,\ s_k>0$, then we have

$$\begin{aligned}
\mathbf{P}(B_k^{res}>t) &\sim \frac{b_k}{(\nu-1)\beta_k}t^{1-\nu}L(t), \quad t\to\infty,\\
\mathbf{P}(S_k^{res}>t) &\sim \frac{s_k}{(\nu-1)\sigma_k}t^{1-\nu}L(t), \quad t\to\infty,
\end{aligned}$$

which in combination with Theorem 3.1 implies the following corollary.

Corollary 3.1 *If Assumption 2.1 holds, then for $t \to \infty$,*

$$\begin{aligned}
1 - W_1(t) &\sim \frac{\rho_1 I_{\{b_1>0\}}}{1-\rho_1}\mathbf{P}(B_1^{res} > t) + \frac{\rho_2 I_{\{b_2>0\}}}{1-\rho_1}\mathbf{P}(B_2^{res} > t) \\
&\quad + \frac{(1-\rho)\sigma_1 I_{\{s_1>0\}}}{(1-\rho_1)\sigma}\mathbf{P}(S_1^{res} > t) \\
&\quad + \frac{(1-\rho)\sigma_2 I_{\{s_2>0\}}}{(1-\rho_1)\sigma}\mathbf{P}(S_2^{res} > t), \\
1 - W_2(t) &\sim \frac{\rho_1 I_{\{b_1>0\}}}{1-\rho-\lambda_2\sigma}\mathbf{P}(B_1^{res} > (1-\rho_1)t) \\
&\quad + \frac{\rho_2 I_{\{b_2>0\}}}{1-\rho-\lambda_2\sigma}\mathbf{P}(B_2^{res} > (1-\rho_1)t) \\
&\quad + \frac{(1-\rho)\sigma_1 I_{\{s_1>0\}}}{(1-\rho-\lambda_2\sigma)\sigma}\mathbf{P}(S_1^{res} > (1-\rho_1)t) \\
&\quad + \frac{(1-\rho)\sigma_2 I_{\{s_2>0\}}}{(1-\rho-\lambda_2\sigma)\sigma}\mathbf{P}(S_2^{res} > (1-\rho_1)t),
\end{aligned}$$

where $I_{\{A\}}$ is the indicator function of event $\{A\}$.

In the following we give a heuristic explanation of the above corollary. These heuristic arguments are similar to those in [17] for a fluid queue with $M/G/\infty$ input. We should point out that the heuristic arguments below are not rigorous in a mathematical sense, and do not really give a strict proof, but only identify a possible way for the desired event to occur, and thus provide a lower bound for the corresponding probability. However, the fact that the lower bound coincides with the formula we found analytically, implies that the probability of any other scenario is negligible. Hence, the scenario that we identify must actually represent the only plausible way in which the event occurs. For more complicated models, like the $M/G/k$ queue, this technique may be a starting point to find the exact waiting time tail behavior.

The heuristic arguments below are based on the following two preliminary observations:

1. At the scale of large t, one may think of the evolution of the workload as approximately linear.

2. Due to the PASTA property, the waiting time has the same distribution as the virtual waiting time (which is equal to the workload in some cases) at any time.

We consider a special case, $b_1 > 0$, $b_2 = s_1 = s_2 = 0$, i.e., the service time B_1 at Q_1 has the heaviest tail. The general case allows a similar intuitive explanation. We use heuristic arguments to verify

$$\mathbf{P}(W_1 > t) \sim \frac{\rho_1}{1-\rho_1}\mathbf{P}(B_1^{res} > t), \quad t \to \infty, \tag{3.5}$$

$$\mathbf{P}(W_2 > t) \sim \frac{\rho_1}{1-\rho-\lambda_2\sigma}\mathbf{P}(B_1^{res} > (1-\rho_1)t), \quad t \to \infty. \tag{3.6}$$

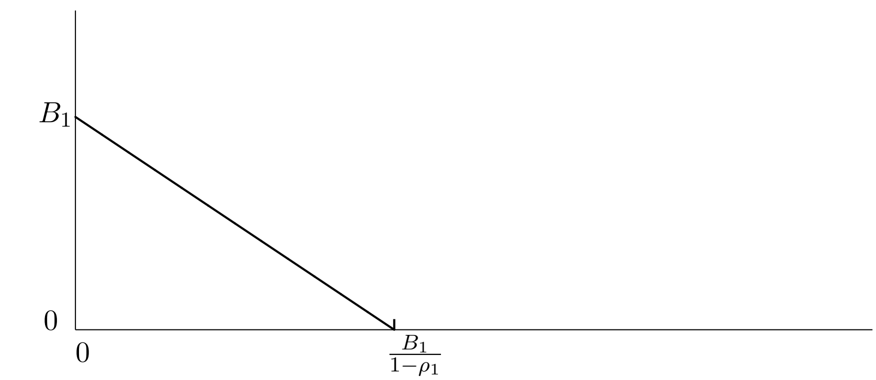

Figure 1: Evolution of the workload at Q_1.

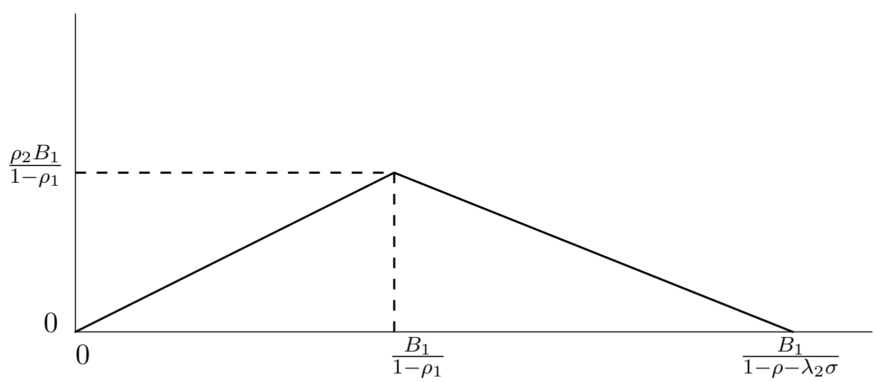

Figure 2: Evolution of the workload at Q_2.

Suppose a customer with a large service time B_1 enters Q_1 in steady state at time 0. Assume that the total workloads at both queues are very small compared to B_1. So at time 0 the workload at Q_1 is roughly B_1 and the workload at Q_2 is roughly 0. The workload at Q_1 decreases at rate $1-\rho_1 > 0$ until it becomes 0 at time $\frac{B_1}{1-\rho_1}$, see Figure 1.

Now we consider the workload at Q_2. During the time interval $(0, \frac{B_1}{1-\rho_1})$, the server stays at Q_1. Therefore, the workload at Q_2 increases at rate ρ_2. Notice that: When Q_2 is not empty, the long term fraction that the server stays at Q_2 is $\frac{\rho_2}{\rho_2+\lambda_2\sigma}(1-\rho_1)$ because the server incurs a vacation time (which is the sum of the switchover time and the period that the server stays at Q_1) for every service provided at Q_2. Hence, after time $\frac{B_1}{1-\rho_1}$, the service speed at Q_2 is $\frac{\rho_2}{\lambda_2\sigma+\rho_2}(1-\rho_1)$. The workload decreases at rate $\frac{\rho_2}{\lambda_2\sigma+\rho_2}(1-\rho-\lambda_2\sigma) > 0$ until time $\frac{B_1}{1-\rho-\lambda_2\sigma}$. After time $\frac{B_1}{1-\rho-\lambda_2\sigma}$, the effect of the customer with the large service time B_1 has disappeared, see Figure 2.

Suppose we observe the system at time y ($y \geq 0$). The virtual waiting time at Q_1 is large, i.e., $W_1 > t$ (t large), because at time 0 a customer with a large service time B_1 entered Q_1. The arrival rate of customers at Q_1 is λ_1. Consider Figure 1. In order to make $W_1 > t$, it is necessary to require $0 < (1-\rho_1)y < B_1 - t$. So,

$$\begin{aligned}\mathbf{P}(W_1 > t) &\approx \int_{y=0}^{\infty} \mathbf{P}(B_1 > (1-\rho_1)y + t)\lambda_1 \mathrm{d}y \\ &= \frac{\lambda_1}{1-\rho_1}\int_{y=t}^{\infty} \mathbf{P}(B_1 > y)\mathrm{d}y \\ &= \frac{\rho_1}{1-\rho_1}\mathbf{P}(B_1^{res} > t),\end{aligned}$$

which coincides with (3.5).

Now consider Figure 2. For the waiting time at Q_2 to become large, there are two possibilities:

1. $0 < y < \frac{B_1}{1-\rho_1}$. Note that the service speed is $\frac{\rho_2}{\rho_2+\lambda_2\sigma}(1-\rho_1)$ instead of 1. In this case, the waiting time can be represented in terms of y as

$$W_2 = \frac{B_1}{1-\rho_1} - y + \frac{(\rho_2+\lambda_2\sigma)y}{1-\rho_1} = \frac{B_1}{1-\rho_1} - \frac{1-\rho-\lambda_2\sigma}{1-\rho_1}y.$$

2. $\frac{B_1}{1-\rho_1} < y < \frac{B_1}{1-\rho-\lambda_2\sigma}$. In this case, the waiting time W_2 is related to y as

$$W_2 = \frac{B_1}{1-\rho_1} - \frac{1-\rho-\lambda_2\sigma}{1-\rho_1}y.$$

In both scenarios, note that when the customer arrives $\mathrm{d}y$ time units later (here $\mathrm{d}y$ stands for a small positive number), the waiting time is reduced by

$$\mathrm{d}y - \frac{\rho_2\mathrm{d}y}{\frac{\rho_2}{\rho_2+\lambda_2\sigma}(1-\rho_1)} = \frac{1-\rho-\lambda_2\sigma}{1-\rho_1}\mathrm{d}y,$$

which explains why the waiting time behaves the same way in both cases. So in order to get $W_2 > t$, we need $B_1 > (1-\rho_1)t + (1-\rho-\lambda_2\sigma)y$. The tail behavior of the waiting time distribution is thus given by

$$\begin{aligned}&\mathbf{P}(W_2 > t) \\ \approx& \int_{y=0}^{\infty} \mathbf{P}(B_1 > (1-\rho-\lambda_2\sigma)y + (1-\rho_1)t)\lambda_1 \mathrm{d}y \\ =& \frac{\lambda_1}{1-\rho-\lambda_2\sigma}\int_{y=(1-\rho_1)t}^{\infty} \mathbf{P}(B_1 > y)\mathrm{d}y \\ =& \frac{\rho_1}{1-\rho-\lambda_2\sigma}\mathbf{P}(B_1^{res} > (1-\rho_1)t),\end{aligned}$$

which coincides with (3.6).

Remark 3.1 Ibe [14] studies a more general model, a K-station mixed polling system in which station 1 is served exhaustively and stations $2, ..., K$ are served according to the 1-limited policy. Suppose at least one of the service and/or switchover times is regularly varying. By using the above heuristic arguments, it is not difficult to obtain the waiting time asymptotics for this K-station polling model.

4 A heavy-traffic limit theorem

Because the service discipline at Q_1 is exhaustive, type-1 customers do not suffer heavy traffic for $\rho+\lambda_2\sigma \uparrow 1$ unless $\rho_1 \uparrow 1$. In that case, the standard heavy-traffic limit theorems (cf. [5, 15] or Section 1) hold for the waiting time distribution at Q_1. This section is devoted to a heavy-traffic limit theorem for the waiting time distribution at Q_2 when at least one of the service and/or switchover times does not have a finite second moment. A similar technique as in [5] is applied.

In the sequel, we assume $1 < \nu < 2$, i.e., indeed at least one of the service and/or switchover times does not have a finite second moment. Then (3.1) and (3.2) reduce to, for $k = 1, 2$,

$$\begin{aligned}\sigma_k(f(s)) &= 1 - \frac{\sigma_k}{1-\rho_1}s + \left(\frac{\lambda_1\beta_{1,\nu}\sigma_k}{(1-\rho_1)^{\nu+1}} + \frac{\sigma_{k,\nu}}{(1-\rho_1)^{\nu}}\right)s^{\nu}L(1/s) \\ &\quad + \mathrm{o}(s^{\nu}L(1/s)), \qquad (4.1)\\ \beta_2(f(s)) &= 1 - \frac{\beta_2}{1-\rho_1}s + \left(\frac{\lambda_1\beta_{1,\nu}\beta_2}{(1-\rho_1)^{\nu+1}} + \frac{\beta_{2,\nu}}{(1-\rho_1)^{\nu}}\right)s^{\nu}L(1/s) \\ &\quad + \mathrm{o}(s^{\nu}L(1/s)). \qquad (4.2)\end{aligned}$$

Just like the priority queue in [6], it is easy to see that the waiting time W_1 at Q_1 is not subject to heavy traffic when $\rho + \lambda_2\sigma \uparrow 1$ unless $\rho_1 \uparrow 1$. In the following we derive a heavy-traffic limit theorem for the waiting time W_2 at Q_2.

Consider the contraction equation

$$\frac{F(\rho)x^{\nu-1}L(1/x)}{1-\rho} = 1, \quad x > 0, \qquad (4.3)$$

where $F(\cdot)$ is a function of ρ such that $F(\rho) > c$ for some positive constant c and $L(x)$ is a slowly varying function. We say that the solution $\Delta(\rho)$ to Equation (4.3) is a unique solution with the property that $\Delta(\rho) \downarrow 0$ for $\rho \uparrow 1$, if for two solutions $\Delta_j(\rho)(j = 1, 2)$ to Equation (4.3) such that $\Delta_j(\rho) \downarrow 0$ for $\rho \uparrow 1$, the limit of their ratio for $\rho \uparrow 1$ is equal to 1, i.e., $\lim_{\rho\uparrow 1} \Delta_1(\rho)/\Delta_2(\rho) = 1$. In the following we provide a lemma (cf. Lemma 3.5.1 in [10]) which characterizes the property of the solution to Equation (4.3).

Lemma 4.1 *If $L(t)$ is continuous, then there exists a unique solution $\Delta(\rho)$ to Equation (4.3) with the property that $\Delta(\rho) \downarrow 0$ for $\rho \uparrow 1$.*

Now we are ready to prove the following theorem.

Theorem 4.1 *If Assumption 2.1 holds with $1 < \nu < 2$, then the contracted waiting time $\Delta(\lambda_1, \lambda_2)W_2$ at Q_2 converges in distribution for $\rho + \lambda_2\sigma \uparrow 1$. The limit distribution function $R_{\nu-1}(t)$ has the following LST:*

$$\int_0^{\infty} e^{-st}\mathrm{d}R_{\nu-1}(t) = \frac{1}{1+s^{\nu-1}}, \quad s > 0, \qquad (4.4)$$

and the coefficient of contraction $\Delta(\lambda_1, \lambda_2)$ is the unique root of the following equation

$$x^{\nu-1}L(1/x) = -\frac{(1-\rho_1)^{\nu-1}}{\Gamma(1-\nu)(\lambda_1 b_1 + \lambda_2 s_1 + \lambda_2 s_2 + \lambda_2 b_2)}(1-\rho-\lambda_2\sigma), \qquad (4.5)$$

with the property that $\Delta(\lambda_1, \lambda_2) \downarrow 0$ for $\rho + \lambda_2\sigma \uparrow 1$.

Proof. We have

$$\begin{aligned}
&\frac{1-\rho-\lambda_2\sigma}{(\lambda_2-s)\sigma}\left(\frac{\lambda_2-s-\lambda_2\beta_2(f(s))}{\lambda_2-s-\lambda_2\sigma_1(f(s))\sigma_2(f(s))\beta_2(f(s))}-1\right)\\
=\;&\frac{\lambda_2\beta_2(f(s))}{(\lambda_2-s)\sigma}\frac{1-\sigma_1(f(s))\sigma_2(f(s))}{s}\frac{1-\rho-\lambda_2\sigma}{1-\lambda_2\frac{1-\sigma_1(f(s))\sigma_2(f(s))\beta_2(f(s))}{s}}. \qquad (4.6)
\end{aligned}$$

For ease of notation, we define

$$H_1(s) := \sigma_1(s)\frac{\sigma_1(\lambda_2+\lambda_1(1-\eta_1(s)))}{\sigma_1(\lambda_2)}\sigma_2(s+\lambda_1(1-\eta_1(s))), \qquad (4.7)$$

$$H_2(s) := \frac{\lambda_2\beta_2(f(s))}{(\lambda_2-s)\sigma}\frac{1-\sigma_1(f(s))\sigma_2(f(s))}{s}, \qquad (4.8)$$

$$H_3(s) := \frac{1-\rho-\lambda_2\sigma}{1-\lambda_2\frac{1-\sigma_1(f(s))\sigma_2(f(s))\beta_2(f(s))}{s}}. \qquad (4.9)$$

Inserting (4.6), ..., (4.9) into (2.5), we may rewrite $\omega_2(s)$ as

$$\omega_2(s) = H_1(s)H_2(s)H_3(s) + \frac{1-\rho-\lambda_2\sigma}{(\lambda_2-s)\sigma}(H_1(s)-1). \qquad (4.10)$$

Let $\Delta(\lambda_1,\lambda_2)$ be the solution of Equation (4.5) with the property that $\Delta(\lambda_1,\lambda_2)\downarrow 0$ for $\rho+\lambda_2\sigma\uparrow 1$. As has been proved in Lemma 4.1, the solution $\Delta(\lambda_1,\lambda_2)$ with that property exists and is unique. To simplify the notation, we make the convention that Δ stands for $\Delta(\lambda_1,\lambda_2)$. It is observed that, for a small number $\delta>0$, there exists a large number $N>0$, such that for any $0<s<\delta$,

$$|H_1(s)-1| < Ns \quad\text{and}\quad \left|H_2(s)-\frac{1}{1-\rho_1}\right| < Ns$$

hold uniformly for $0<\rho+\lambda_2\sigma<1$. Therefore, we have

$$\lim_{\rho+\lambda_2\sigma\uparrow 1} H_1(\Delta s) = 1, \qquad (4.11)$$

$$\lim_{\rho+\lambda_2\sigma\uparrow 1} H_2(\Delta s) = \frac{1}{1-\rho_1}. \qquad (4.12)$$

By (4.10), (4.11) and (4.12), in order to show that

$$\lim_{\rho+\lambda_2\sigma\uparrow 1}\omega_2(\Delta s) = \frac{1}{1+s^{\nu-1}}, \qquad (4.13)$$

it remains to prove that $\lim_{\rho+\lambda_2\sigma\uparrow 1} H_3(\Delta s) = \frac{1-\rho_1}{1+s^{\nu-1}}$. From (4.1) and (4.2), we may write

$$\begin{aligned}
&\frac{1-\sigma_1(f(s))\sigma_2(f(s))\beta_2(f(s))}{s} = \frac{\sigma+\beta_2}{1-\rho_1}\\
&-\left(\frac{\lambda_1\beta_{1,\nu}(\sigma+\beta_2)}{(1-\rho_1)^{\nu+1}}+\frac{\sigma_{1,\nu}+\sigma_{2,\nu}+\beta_{2,\nu}}{(1-\rho_1)^{\nu}}\right)s^{\nu-1}L(1/s)+G(s)s, \qquad (4.14)
\end{aligned}$$

where $G(s)$ is a function of s. For simplicity, we omit the expression for $G(s)$ here. One can easily prove that there exist a large number $N>0$ and a small number $\epsilon>0$ such that for any $0<s<\epsilon$,

$$|G(s)| < N.$$

Note that N is independent of $\rho + \lambda_2\sigma$. Thus, inserting (4.14) into (4.9) gives

$$H_3(s) = (1-\rho_1)\left[1 + \frac{1-\rho_1}{1-\rho-\lambda_2\sigma}\left(\frac{\lambda_1\beta_{1,\nu}(\lambda_2\sigma+\rho_2)}{(1-\rho_1)^{\nu+1}} + \frac{\lambda_2\sigma_{1,\nu}+\lambda_2\sigma_{2,\nu}+\lambda_2\beta_{2,\nu}}{(1-\rho_1)^{\nu}}\right) s^{\nu-1}L(1/s) + \frac{\lambda_2(1-\rho_1)}{1-\rho-\lambda_2\sigma}G(s)s\right]^{-1}. \quad (4.15)$$

Since $G(s)$ is in the neighborhood of the origin uniformly bounded for $0 < \rho + \lambda_2\sigma < 1$, it is easy to see that

$$\lim_{\rho+\lambda_2\sigma\uparrow 1} \frac{1-\rho_1}{1-\rho-\lambda_2\sigma} G(\Delta s)\Delta s = 0.$$

Therefore, replacing s in (4.15) by Δs, we get

$$\lim_{\rho+\lambda_2\sigma\uparrow 1} H_3(\Delta s) = \frac{1-\rho_1}{1+s^{\nu-1}}. \quad (4.16)$$

Let $R_{\nu-1}(t)$ denote the distribution which has LST $\frac{1}{1+s^{\nu-1}}$. Using the convergence theorem of Feller for Laplace-Stieltjes transforms, cf. [12], it follows from (4.13) that ΔW_2 converges in distribution and the limit distribution $R_{\nu-1}(t)$ satisfies (4.4). □

Remark 4.1 Take $\sigma = s_1 = s_2 = 0$ in (4.5). We then obtain a heavy-traffic limit theorem for the low-priority waiting time distribution in the 2-class priority queue. It coincides with the one in [6].

5 Application of the heavy-traffic limit theorem

Theorem 4.1 suggests the following heavy-traffic approximation for the waiting time distribution $W_2(t)$ at Q_2: For $\rho + \lambda_2\sigma < 1$,

$$1 - W_2(t) = \mathbf{P}(W_2 > t) \approx 1 - R_{\nu-1}(\Delta(\lambda_1, \lambda_2)t), \quad t > 0,$$

where $\Delta(\lambda_1, \lambda_2)$ is specified by Equation (4.5). According to the heavy-traffic limit theorem, this approximation should perform very well when $\rho + \lambda_2\sigma$ is sufficiently close to 1. In this section, we numerically test the accuracy of the approximation suggested by the heavy-traffic limit theorem. We conclude that this approximation is useful in some cases. In the following, we follow the same procedure as in [5, 6] to numerically investigate the accuracy of this heavy-traffic approximation for different values of ν and $\rho + \lambda_2\sigma$. Theorem 3.1 suggests the following asymptotic approximation for $1 - W_2(t)$:

$$1 - W_{RV}(t) := \min\{1, C_{RV}t^{1-\nu}L(t)\}$$

with

$$C_{RV} := \frac{1}{\nu-1}\left(\frac{\lambda_1 b_1 + \lambda_2(s_1+s_2+b_2)}{(1-\rho_1)^{\nu-1}(1-\rho-\lambda_2\sigma)} + \frac{s_1+s_2}{(1-\rho_1)^{\nu-1}\sigma}\right),$$

and Theorem 4.1 suggests the following heavy-traffic approximation:

$$1 - W_{HT}(t) := 1 - R_{\nu-1}(\Delta(\lambda_1, \lambda_2)t), \quad t > 0.$$

Suppose the switchover times are exponentially distributed and the service time distributions are of the following form

$$B_j(t) = 1 - \frac{1}{\Gamma(2-\nu_j)} \int_0^\infty e^{-\theta} \frac{\theta}{(\theta+t)^{\nu_j}} \mathrm{d}\theta, \quad j = 1, 2, \tag{5.1}$$

with $1 < \nu_j < 2$. Hence, we have

$$1 - W_2(t) \sim 1 - W_{HT}(t) \sim 1 - W_{RV}(t), \quad t \to \infty.$$

We have tested the approximations $1-W_{HT}(t)$ and $1-W_{RV}(t)$ for a large number of parameter combinations. Tables 1-6 below display numerical results for the following 6 cases: (i) $\nu = 1.25$, $\rho+\lambda_2\sigma = 0.9$, (ii) $\nu = 1.25$, $\rho+\lambda_2\sigma = 0.5$, (iii) $\nu = 1.5$, $\rho+\lambda_2\sigma = 0.9$, (iv) $\nu = 1.5$, $\rho+\lambda_2\sigma = 0.5$, (v) $\nu = 1.75$, $\rho+\lambda_2\sigma = 0.9$, and (vi) $\nu = 1.75$, $\rho+\lambda_2\sigma = 0.5$. We use the Fourier-series method for inverting transforms of probability distributions (cf. [1]) to compute $1 - W_2(t)$. Similar conclusions as in [6] can be made:

(i) When t is large, e.g., $t \geq 50000$, the heavy-traffic approximation $1-W_{HT}(t)$ is very accurate in all cases; while the asymptotic approximation $1 - W_{RV}(t)$ performs much worse when ν is small.

(ii) The larger the value of $\rho + \lambda_2\sigma$, the better the heavy-traffic approximation $1 - W_{HT}(t)$ performs.

(iii) When ν is small and $\rho + \lambda_2\sigma$ is large, e.g., $\nu \leq 1.5$ and $\rho + \lambda_2\sigma \geq 0.9$, the heavy-traffic approximation $1 - W_{HT}(t)$ is very good even for small t.

(iv) When ν is large, e.g., $\nu \geq 1.75$, the heavy-traffic approximation $1 - W_{HT}(t)$ performs poorly for small t; it is not better than the asymptotic approximation $1 - W_{RV}(t)$.

Acknowledgements
The author is indebted to Onno Boxma and Sem Borst for interesting discussions and valuable comments.

References

[1] J. Abate and W. Whitt (1992). *The Fourier-series method for inverting transforms of probability distributions.* Queueing Systems **10**, 5-88.

[2] J. Beran, R. Sherman, M.S. Taqqu and W. Willinger (1995). *Long-range dependence in variable-bit-rate video traffic.* IEEE Transactions on Communications **43**, 1566-1579.

[3] N.H. Bingham, C.M. Goldie and J.L. Teugels (1987). Regular Variation. Cambridge University Press, Cambridge.

[4] O.J. Boxma and J.W. Cohen (2000). *The single server queue: heavy tails and heavy traffic.* In: Self-Similar Network Traffic and Performance Evaluation. Eds. K. Park and W. Willinger. John Wiley & Sons, New York, 143-170.

[5] O.J. Boxma and J.W. Cohen (1999). *Heavy-traffic analysis for the GI/G/1 queue with heavy-tailed distributions.* Queueing Systems **33**, 177-204.

[6] O.J. Boxma, J.W. Cohen and Q. Deng (1999). *Heavy-traffic analysis of the M/G/1 queue with priority classes.* In: Teletraffic Engineering in a Competitive World, Proceedings of ITC-16. Eds. P. Key and D. Smith. North-Holland, Amsterdam, 1157-1167.

[7] O.J. Boxma, Q. Deng and J.A.C. Resing (2000). *Polling systems with regularly varying service and/or switchover times.* Advances in Performance Analysis **3**, 71-107.

[8] O.J. Boxma and V. Dumas (1998). *Fluid queues with heavy-tailed activity period distributions.* Computer Communications **21**, 1509-1529.

[9] G.L. Choudhury and W. Whitt (1996). *Computing distributions and moments in polling models by numerical transform inversion.* Performance Evaluation **25**, 267-292.

[10] Q. Deng (2001). Queues with Regular Variation. Ph.D. Thesis, Eindhoven University of Technology, The Netherlands.

[11] N.G. Duffield (1997). *Exponents for the tail of distributions in some polling models.* Queueing Systems **26**, 105-119.

[12] W. Feller (1970). An Introduction to Probability Theory and its Applications, Vol. II. John Wiley & Sons, New York.

[13] W.P. Groenendijk (1990). Conservation Laws in Polling Systems. Ph.D. Thesis, University of Utrecht.

[14] O.C. Ibe (1990). *Analysis of polling systems with mixed service disciplines.* Communications in Statistics, Stochastic Models **6**, 667-689.

[15] J.F.C. Kingman (1965). *The heavy traffic approximation in the theory of queues.* In: Proceedings of the Symposium on Congestion Theory. Eds. W.L. Smith and W.E. Wilkinson. The University of North Carolina Press, Chapel Hill, 137-159.

[16] A. De Meyer and J.L. Teugels (1980). *On the asymptotic behavior of the distributions of the busy period and service time in M/G/1.* Journal of Applied Probability **17**, 802-813.

[17] S. Resnick and G. Samorodnitsky (2001). *Steady state distribution of the buffer content for $M/G/\infty$ input fluid queues.* Bernoulli **7**, 191-210.

[18] H. Takagi (1990). *Queueing analysis of polling models: an update.* In: Stochastic Analysis of Computer and Communication Systems. Ed. H. Takagi. Elsevier, Amsterdam, 267-318.

[19] H. Takagi (1997). *Queueing analysis of polling models: Progress in 1990-1994.* In: Frontiers in Queueing. Ed. J.H. Dshalalow. CRC Press, Boca Raton, 119-146.

[20] W. Willinger, M.S. Taqqu, W.E. Leland and D.V. Wilson (1995). *Self-similarity in high-speed packet traffic: analysis and modeling of Ethernet traffic measurements.* Statistical Science **10**, 67-85.

Table 1: Approximations for the waiting time tails at Q_2.

$\nu = 1.25;\ \rho + \lambda_2\sigma = 0.9$ $\nu_1 = \nu_2 = 1.25;\ \sigma_1 = \sigma_2 = 0.05;\ \lambda_1 = \lambda_2 = 0.1475$					
t	$1 - W_2(t)$	$1 - W_{HT}(t)$	$\%\text{error}_{HT}$	$1 - W_{RV}(t)$	$\%\text{error}_{RV}$
1	0.886	0.925	4.40	1	12.88
2	0.878	0.912	3.89	1	13.94
5	0.863	0.891	3.30	1	15.88
10	0.849	0.873	2.93	1	17.85
20	0.831	0.853	2.60	1	20.33
50	0.803	0.821	2.26	1	24.52
100	0.778	0.794	2.04	1	28.52
200	0.750	0.764	1.86	1	33.38
500	0.707	0.719	1.67	1	41.37
1000	0.672	0.682	1.54	1	48.85
2000	0.634	0.643	1.42	1	57.82
5000	0.580	0.587	1.27	1	72.46
10000	0.537	0.544	1.17	1	57.82
20000	0.494	0.499	1.07	0.937	89.72
50000	0.437	0.441	0.94	0.745	70.68
100000	0.394	0.398	0.84	0.627	59.00
200000	0.353	0.356	0.75	0.527	49.26
500000	0.302	0.304	0.64	0.419	38.81

Table 2: Approximations for the waiting time tails at Q_2.

$\nu = 1.25;\ \rho + \lambda_2\sigma = 0.5$ $\nu_1 = \nu_2 = 1.25;\ \sigma_1 = \sigma_2 = 0.05;\ \lambda_1 = \lambda_2 = 0.082$					
t	$1 - W_2(t)$	$1 - W_{HT}(t)$	$\%\text{error}_{HT}$	$1 - W_{RV}(t)$	$\%\text{error}_{RV}$
1	0.462	0.551	19.34	1	116.50
2	0.441	0.507	14.95	0.966	118.97
5	0.406	0.448	10.34	0.768	88.99
10	0.376	0.405	7.71	0.646	71.80
20	0.343	0.363	5.74	0.543	58.25
50	0.299	0.310	3.92	0.432	44.57
100	0.266	0.274	2.99	0.363	36.64
200	0.234	0.240	2.33	0.305	30.26
500	0.196	0.200	1.73	0.243	23.63
1000	0.171	0.173	1.41	0.204	19.67
2000	0.148	0.149	1.17	0.172	16.41
5000	0.121	0.122	0.92	0.137	12.94
10000	0.104	0.104	0.78	0.115	10.83
20000	0.089	0.089	0.65	0.097	9.07
50000	0.072	0.072	0.52	0.077	7.18
100000	0.061	0.061	0.44	0.065	6.02
200000	0.052	0.052	0.37	0.054	5.05
500000	0.041	0.042	0.37	0.043	4.08

Table 3: Approximations for the waiting time tails at Q_2.

$\nu = 1.5;\ \rho + \lambda_2\sigma = 0.9$					
$\nu_1 = \nu_2 = 1.5;\ \sigma_1 = \sigma_2 = 0.15;\ \lambda_1 = \lambda_2 = 0.3913$					
t	$1 - W_2(t)$	$1 - W_{HT}(t)$	$\%\text{error}_{HT}$	$1 - W_{RV}(t)$	$\%\text{error}_{RV}$
1	0.869	0.946	8.90	1	15.10
2	0.846	0.925	9.40	1	18.25
5	0.808	0.886	9.61	1	23.76
10	0.771	0.844	9.53	1	29.73
20	0.724	0.791	9.22	1	38.09
50	0.646	0.700	8.37	1	54.83
100	0.574	0.616	7.38	1	74.19
200	0.494	0.524	6.15	0.800	62.13
500	0.382	0.398	4.32	0.506	32.62
1000	0.300	0.310	3.01	0.358	19.11
2000	0.229	0.233	1.92	0.253	10.74
5000	0.153	0.154	0.95	0.160	4.73
10000	0.110	0.111	0.52	0.113	2.47
20000	0.079	0.079	0.29	0.080	1.28
50000	0.050	0.050	0.12	0.051	0.51
100000	0.036	0.036	0.08	0.036	0.28
200000	0.025	0.025	0.20	0.025	0.30
500000	0.016	0.016	-0.17	0.016	-0.13

Table 4: Approximations for the waiting time tails at Q_2.

$\nu = 1.5;\ \rho + \lambda_2\sigma = 0.5$ $\nu_1 = \nu_2 = 1.5;\ \sigma_1 = \sigma_2 = 0.15;\ \lambda_1 = \lambda_2 = 0.2174$					
t	$1 - W_2(t)$	$1 - W_{HT}(t)$	$\%\text{error}_{HT}$	$1 - W_{RV}(t)$	$\%\text{error}_{RV}$
1	0.420	0.611	45.37	1	138.10
2	0.368	0.518	40.73	0.784	112.90
5	0.301	0.393	30.43	0.496	64.72
10	0.249	0.305	22.35	0.351	40.85
20	0.199	0.229	15.18	0.248	24.79
50	0.140	0.151	8.09	0.157	11.99
100	0.104	0.109	4.64	0.111	6.59
200	0.076	0.078	2.53	0.078	3.50
500	0.049	0.049	1.08	0.050	1.46
1000	0.035	0.035	0.55	0.035	0.74
2000	0.025	0.025	0.28	0.025	0.38
5000	0.016	0.016	0.12	0.016	0.15
10000	0.011	0.011	0.06	0.011	0.08
20000	0.008	0.008	0.03	0.008	0.04
50000	0.005	0.005	0.03	0.005	0.03
100000	0.003	0.004	0.35	0.004	0.36
200000	0.002	0.002	-0.15	0.002	-0.15
500000	0.002	0.002	0.14	0.002	0.14

Table 5: Approximations for the waiting time tails at Q_2.

$\nu = 1.75;\ \rho + \lambda_2\sigma = 0.9$ $\nu_1 = \nu_2 = 1.75;\ \sigma_1 = \sigma_2 = 0.45;\ \lambda_1 = \lambda_2 = 0.5745$					
t	$1 - W_2(t)$	$1 - W_{HT}(t)$	$\%\text{error}_{HT}$	$1 - W_{RV}(t)$	$\%\text{error}_{RV}$
1	0.893	0.942	5.49	1	11.98
2	0.818	0.904	10.50	1	22.25
5	0.691	0.821	18.90	1	44.72
10	0.577	0.723	25.33	0.881	52.85
20	0.452	0.589	30.40	0.524	15.97
50	0.288	0.379	31.56	0.264	-8.53
100	0.184	0.232	25.92	0.157	-15.05
200	0.110	0.128	16.48	0.093	-15.15
500	0.053	0.056	6.47	0.047	-10.93
1000	0.030	0.031	2.89	0.028	-7.49
2000	0.017	0.018	1.29	0.017	-4.86
5000	0.009	0.009	0.47	0.008	-2.62
10000	0.005	0.005	0.55	0.005	-1.29
20000	0.003	0.003	0.01	0.003	-1.09
50000	0.001	0.001	0.00	0.001	-0.34
100000	0.001	0.001	0.00	0.001	-0.03
200000	0.001	0.001	0.00	0.001	0.00
500000	0.000	0.000	0.00	0.000	0.00

Table 6: Approximations for the waiting time tails at Q_2.

$\nu = 1.75$; $\rho + \lambda_2\sigma = 0.5$ $\nu_1 = \nu_2 = 1.75$; $\sigma_1 = \sigma_2 = 0.45$; $\lambda_1 = \lambda_2 = 0.3191$					
t	$1 - W_2(t)$	$1 - W_{HT}(t)$	$\%\text{error}_{HT}$	$1 - W_{RV}(t)$	$\%\text{error}_{RV}$
1	0.538	0.582	8.16	0.511	-5.05
2	0.327	0.423	29.41	0.304	-7.10
5	0.154	0.226	47.01	0.153	-0.64
10	0.092	0.124	34.36	0.091	-1.61
20	0.056	0.066	18.56	0.054	-3.33
50	0.028	0.030	6.86	0.027	-3.67
100	0.017	0.017	3.08	0.016	-3.03
200	0.010	0.010	1.36	0.010	-2.23
500	0.005	0.005	0.44	0.005	-1.35
1000	0.003	0.003	0.18	0.003	-0.88
2000	0.002	0.002	0.07	0.002	-0.56
5000	0.001	0.001	0.02	0.001	-0.29
10000	0.001	0.001	0.00	0.001	-0.15
20000	0.000	0.000	0.00	0.000	0.03
50000	0.000	0.000	0.00	0.000	0.00
100000	0.000	0.000	0.00	0.000	0.00
200000	0.000	0.000	0.00	0.000	0.00
500000	0.000	0.000	0.00	0.000	0.00