



Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

J.L. van den Berg

Simple approximations for second moment characteristics
of the sojourn time in the M/G/1 processor sharing queue

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Simple Approximations for Second Moment Characteristics of the Sojourn Time in the M/G/1 Processor Sharing Queue

J.L. van den Berg

*Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands*

In this paper simple approximations are derived for the second moment of the conditional and unconditional sojourn time in the M/G/1 processor sharing queue. The approximations are mainly based on general asymptotic results (e.g. heavy traffic) and on simple exact expressions for specific service time distributions. They depend on the service time distribution only through its first and second moment. Numerous numerical examples show that these simple approximations are accurate enough for many practical purposes. A refinement of the approximations is obtained by taking the third moment of the service time into account.

1980 Mathematics Subject Classification: 60K25, 68M20.

Key Words & Phrases: processor sharing queue, sojourn time, approximations.

1. INTRODUCTION

The processor sharing (PS) service discipline is widely used to model time sharing in computer systems. During the last ten years considerable attention has been paid to the analysis of the M/G/1 PS queue, see e.g. the surveys by Jaiswal [4] and Yashkov [13]. The major problem in processor sharing queues is the problem of characterizing the sojourn time distribution. Only recently exact expressions for the Laplace-Stieltjes transform of the sojourn time distribution have been obtained by Yashkov [11], Ott [6] and Schassberger [7]. Due to their complexity these formulas are not attractive for practical applications. Only for the mean sojourn time a simple explicit expression exists; this expression is insensitive to the service time distribution apart from its first moment (Kleinrock [5]). Formulas for the second moment of the sojourn time (Ott [6], Yashkov [13]) require perfect information about the service time distribution, which is almost never available in practice. Moreover, these formulas contain a (double) integral which, in general, can only be evaluated numerically. As far as we know no attention has been paid to the derivation of approximations or asymptotic formulas which are useful for practical evaluation, apart from [1] and a paper by Yashkov [12]. Yashkov derives some asymptotic estimates for the conditional sojourn time variance for customers with small or large service times. An extension of these results is presented in [1].

The aim of the present study is to derive approximations for the second moment of the sojourn time distribution, which are quite simple and yet accurate enough for most practical purposes. We first show that a lower and upper bound for the second moment of the sojourn time can be expressed in terms of the first and second moment of the service time. Next some very simple approximation formulas based on the first and second moment of the service time are presented. The accuracy of the approximations is tested for a large number of different service time distributions and a wide range of traffic intensities. A refinement of the approximation is obtained by taking the third moment of the service time into account. This refinement yields remarkably accurate results with relative errors less than 1.5 percent in most cases.

The organization of the rest of the paper is as follows. In Section 2 we introduce the notations and give a summary of those known sojourn time results which are relevant for our study. We also present some extensions and new results. In particular an upper bound for the second moment of the

sojourn time is derived. Section 3 is concerned with the second moment of the *conditional* sojourn time of a customer with service demand x . We propose an approximation formula which is based on an asymptotic result for $x \rightarrow 0$ derived in [1]. In Section 4 approximations are developed for the second moment of the *unconditional* sojourn time. We first propose an approximation which uses only information about the first and second moment of the service time distribution. A refinement of the approximation is derived for the case that the squared coefficient of variation of the service time distribution is smaller than one. Finally, for service time distributions with a squared coefficient of variation larger than one, we construct a very accurate approximation formula which is based on the first three moments of the service time distribution.

2. NOTATIONS AND PRELIMINARY RESULTS

Customers arrive at a single server queue according to a Poisson process with rate λ . Their service requirements are i.i.d. non-negative random variables with a general distribution $B(\cdot)$ with first and second moment β and β_2 . All customers are served simultaneously according to the processor sharing (PS) service discipline, i.e. whenever i customers are present, each customer receives service at a rate of $1/i$. We assume that $\rho := \lambda\beta < 1$ and that the system is in steady state. Let $S(x)$ be the sojourn time of a tagged customer who requires an amount x of service at his arrival. It is well-known that $E S(x)$ is linear in x (see Kleinrock [5]):

$$E S(x) = \frac{x}{1-\rho}. \quad (2.1)$$

The second moment of the distribution of $S(x)$ has been obtained by Yashkov [11] and Ott [6]:

$$E S^2(x) = \frac{x^2}{(1-\rho)^2} + \frac{2}{(1-\rho)^2} \int_0^x (x-t)(1-R(t)) dt, \quad (2.2)$$

where $R(t)$ represents the waiting time distribution for the M/G/1 first come first served (FCFS) queue with service time distribution $B(\cdot)$,

$$R(t) = (1-\rho) \sum_{n=0}^{\infty} \rho^n F^{n*}(t), \quad (2.3)$$

$$F(t) = \frac{1}{\beta} \int_0^t (1-B(u)) du.$$

Note, for the waiting time distribution $R(t)$ in (2.2), that $1-R(t) \leq 1-R(0) = \rho$, $t \geq 0$. Hence,

$$E S^2(x) \leq \frac{1+\rho}{(1-\rho)^2} x^2. \quad (2.4)$$

A lower bound for $E S^2(x)$ follows immediately from (2.1) and Schwartz' inequality (see also (2.2)),

$$E S^2(x) \geq \frac{x^2}{(1-\rho)^2}. \quad (2.5)$$

So,

$$\frac{x^2}{(1-\rho)^2} \leq E S^2(x) \leq \frac{1+\rho}{(1-\rho)^2} x^2. \quad (2.6)$$

Note that the upper bound is 100% higher than the lower bound and that these bounds depend only on the first moment of the service time distribution; this supports a certain robustness of $E S^2(x)$ for the service time distribution.

The heavy traffic behaviour of $E S^2(x)$ can be derived from (2.2) by noting that the heavy traffic behaviour of the waiting time distribution for the M/G/1 FCFS queue is, for $\rho \rightarrow 1$, negative exponential, i.e. (see Cohen [3])

$$R\left(\frac{t}{1-\rho}\right) \sim 1 - e^{-t/d}, \quad \text{for } \rho \rightarrow 1, \quad (2.7)$$

where $d = \frac{1}{2} \rho \beta_2 / \beta$.

Substituting (2.7) into (2.2) yields

$$E S^2(x) \sim \frac{1+\rho}{(1-\rho)^2} x^2, \quad \text{for } \rho \rightarrow 1. \quad (2.8)$$

Other asymptotic results apply to the case that the required service time x of the tagged customer becomes either very small or very large. In [1] it is shown that

$$\text{var}(S(x)) = E S^2(x) - (E S(x))^2 \sim \frac{\rho}{(1-\rho)^2} x^2 - \frac{\rho}{3\beta(1-\rho)} x^3, \quad \text{for } x \rightarrow 0. \quad (2.9)$$

The x^2 -term in this formula has already been obtained by Yashkov [12].

The quadratic behaviour of the sojourn time variance, $\text{var}(S(x))$, for $x \rightarrow 0$ contrasts with the linear behaviour for large x (see Yashkov [12], and [1]):

$$\text{var}(S(x)) \sim \frac{\rho \beta_2}{\beta(1-\rho)^2} x - \frac{1}{(1-\rho)^4} \left[\frac{1}{2} \rho^2 \frac{\beta_2^2}{\beta^2} + \frac{1}{3} \rho \frac{\beta_3}{\beta} (1-\rho) \right], \quad \text{for } x \rightarrow \infty. \quad (2.10)$$

For exponential and deterministic service times, simple, explicit expressions for $E S^2(x)$ are known, see e.g. Ott [6]. For future use we state these expressions:

$$E S^2(x)_{EXP} = \frac{2\rho\beta}{(1-\rho)^3} x - \frac{2\rho\beta^2}{(1-\rho)^4} (1 - e^{-x(1-\rho)/\beta}), \quad x \geq 0, \quad (2.11)$$

$$E S^2(x)_{DET} = \frac{2}{(1-\rho)^2} x^2 - \frac{2\beta^2}{\rho^2(1-\rho)} (e^{\rho x/\beta} - 1 - \rho x/\beta), \quad 0 \leq x \leq \beta. \quad (2.12)$$

Let S be the unconditional sojourn time, i.e.

$$\Pr\{S \leq t\} = \int_{x=0}^{\infty} \Pr\{S(x) \leq t\} dB(x), \quad t \geq 0. \quad (2.13)$$

From the above results for $S(x)$ it follows immediately that

$$E S = \frac{\beta}{1-\rho}, \quad (2.14)$$

$$\frac{\beta_2}{(1-\rho)^2} \leq E S^2 \leq \frac{1+\rho}{(1-\rho)^2} \beta_2, \quad (2.15)$$

$$E S^2 \sim \frac{1+\rho}{(1-\rho)^2} \beta_2, \quad \text{for } \rho \rightarrow 1. \quad (2.16)$$

For exponential and deterministic service times,

$$E S^2_{EXP} = \left(1 + \frac{2+\rho}{2-\rho}\right) \frac{\beta^2}{(1-\rho)^2}, \quad (2.17)$$

$$E S^2_{DET} = \frac{2\beta^2}{(1-\rho)^2} - \frac{2\beta^2}{\rho^2(1-\rho)}(e^\rho - 1 - \rho). \quad (2.18)$$

REMARK. (2.15) implies that, for the M/G/1 PS queue, the dependence of $E S^2$ on the third moment of the service time distribution is limited. This should be contrasted with the behaviour of the second moment of the sojourn time distribution for the M/G/1 FCFS queue. For the FCFS discipline it depends linearly on the third moment of the service time distribution.

In the following sections the above results are exploited to develop simple approximations for $E S^2(x)$ and $E S^2$. We present extensive tables comparing the approximations with exact values. The service time distributions which we have chosen to test the approximations are:

- exponential distribution
- deterministic distribution
- k -stage Erlang distribution (E_k)
- two-stage hyper exponential distribution (H_2), in particular
 - H_2 with balanced means (H_2^{BM}), and
 - H_2 with gamma normalization (H_2^{GN})
- two-stage Coxian distribution (C_2)
- three-stage hyper exponential distribution (H_3)

These types of service time distributions are often used for practical applications in queueing theory, see Tijms [8] and Whitt [9,10].

In practice service times are often characterized by the mean, β , and the squared coefficient of variation, cv , defined by

$$cv = \frac{\sigma^2}{\beta^2},$$

where σ^2 denotes the service time variance, see Tijms [8]. In the rest of this paper we shall use cv rather than σ^2 to characterize the variability of the service times.

The H_2^{BM} and H_2^{GN} distributions have been introduced to reduce the number of parameters of the H_2 distribution, see Tijms [8]. The H_2^{BM} and H_2^{GN} distributions are uniquely determined by their first two moments. In particular, the H_2^{GN} distribution with mean β and $cv \geq 1$ has the same third moment as the gamma distribution with mean β and squared coefficient of variation cv . In Section 4 the class of H_2 distributions will be considered in more detail.

The tables presented at the end of the paper contain relative errors of the approximations for various service time distributions. The relative error is defined as

$$100\% \frac{\text{approximation result} - \text{exact result}}{\text{exact result}}$$

The exact values of $E S^2(x)$ and $E S^2$ have been obtained from the formulas derived in [1]. For H_k and C_k (and E_k) service time distributions these formulas require the roots of a polynomial of degree k and the solution of a set of k linear equations. Even for the case $k=2$, the resulting expressions are very large and complicated and do not give much insight into the influence of the parameters.

3. APPROXIMATION OF $E S^2(x)$

In this section we show that the asymptotic result (2.9) yields a good approximation for $E S^2(x)$ for an important range of x -values.

We define (cf. (2.9)),

$$E S^2(x)_{APPX} = \frac{1+\rho}{(1-\rho)^2} x^2 - \frac{\rho}{3\beta(1-\rho)} x^3. \quad (3.1)$$

Note that $E S^2(x)_{APPX}$ satisfies the heavy traffic behaviour of $E S^2(x)$ (see (2.8)) and that $E S^2(x)_{APPX}$ is smaller than the upper bound of $E S^2(x)$ given by (2.4). Approximation $E S^2(x)_{APPX}$ is independent of the service time distribution apart from its first moment. Obviously it can not be applied for too large values of x because it becomes negative for $x > 3\beta(1+\rho)/(\rho(1-\rho))$. Moreover, assuming that the variance of $S(x)$ is a convex function of x (cf. (2.9) and (2.10)), we may not expect that $E S^2(x)_{APPX}$ is a good approximation for $x > x_1$, where $x_1 = \beta/(1-\rho) = E S$ is the point of inflection of (cf. (2.9))

$$f(x) = \frac{\rho}{(1-\rho)^2} x^2 - \frac{\rho}{3\beta(1-\rho)} x^3. \quad (3.2)$$

For $x < x_1$, $E S^2(x)_{APPX}$ is within the bounds of $E S^2(x)$ given by (2.6).

In Table 1 approximation results are compared with exact results for a number of different values of x ($x = \frac{1}{2}\beta, \beta, \frac{3}{2}\beta, 2\beta, \frac{\beta/2}{1-\rho}, \frac{\beta}{1-\rho}$) and different service time distributions. For each of these cases ρ varies from 0.1 to 0.9. For the sake of clarity only the relative approximation errors are given. It appears that for most cases the relative approximation errors are negative. As expected, the approximation becomes less accurate when x grows. For $0 \leq x \leq \beta$ the relative errors are less than 2.34% in absolute value. For $0 \leq x \leq 2\beta$ the maximum relative error is 6.56%. When x remains constant the maximum errors occur for $\rho \approx 0.3$. For $x = \beta/(2(1-\rho))$ and $x = \beta/(1-\rho)$ the relative errors tend to increase when ρ grows. For $x = \beta/(1-\rho)$ the maximum error is 11.29%.

It is seen from the results for different service time distributions that the accuracy of the approximation tends to decrease when cv becomes larger.

REMARK. In [2] we have derived an approximation for $E S^2(x)$ for the whole range of possible x -values ($x \geq 0$) by appropriately combining the two asymptotic formulas (2.9) and (2.10). The idea is as follows. Two values x_1 and x_2 are determined, such that for $x \leq x_1$ (2.9) yields a good approximation and for $x \geq x_2$ (2.10) yields a good approximation. For $x_1 \leq x \leq x_2$, $E S^2(x)$ is approximated by the term $x^2/(1-\rho)^2$ plus a linear function of x , cf. (2.10). We took $x_1 = \beta/(1-\rho)$. Details about the determination of x_2 are given in [2]. The approximation yields reasonably good results for service

time distributions with cv not too large ($cv \leq 4$). For these cases we found relative errors which are typically less than 10%.

4. APPROXIMATION OF $E S^2$

In this section we propose some approximations for the second moment of the unconditional sojourn time S . First we derive a very simple approximation which is based on the exact formula of $E S^2$ for exponentially distributed service times. This approximation uses only the first two moments of the service time distribution. Next it is shown how this simple approximation can be improved. For that purpose we distinguish between models with a service time squared coefficient of variation cv between zero and one, and models with cv larger than one. In the latter case also the third moment of the service time distribution is taken into account.

4.1. A simple approximation

It follows from (2.15) that an approximation $E S^2_{APP}$ of $E S^2$, which satisfies

$$\frac{\beta_2}{(1-\rho)^2} \leq E S^2_{APP} \leq \frac{1+\rho}{(1-\rho)^2} \beta_2, \quad (4.1)$$

yields relative errors which are bounded by 100% in absolute value. This observation and the relations for $E S^2$ given in Section 2 support the idea to derive an approximation for $E S^2$ which is based only on the first two moments of the service time distribution. We start with the exact formulas (2.17) and (2.18) for the case of exponential and deterministic service times. These formulas can be rewritten as follows:

$$E S^2_{EXP} = \left(1 + \frac{\rho}{2-\rho}\right) \frac{\beta_2}{(1-\rho)^2} = \left(1 + \frac{1}{2}\rho + \frac{1}{4}\rho^2 + \frac{1}{8}\rho^3 + \dots\right) \frac{\beta_2}{(1-\rho)^2}, \quad (4.2)$$

$$\begin{aligned} E S^2_{DET} &= 2\left(1 - \frac{(1-\rho)}{\rho^2}(e^\rho - 1 - \rho)\right) \frac{\beta_2}{(1-\rho)^2} \\ &= \left(1 + \frac{2}{3}\rho + \frac{1}{4}\rho^2 + \frac{1}{15}\rho^3 + \dots\right) \frac{\beta_2}{(1-\rho)^2}. \end{aligned} \quad (4.3)$$

(4.2) and (4.3) suggest that $E S^2$ is almost linear in β_2 . This observation and relation (2.15) lead to an approximation, $E S^2_{APP}$, for $E S^2$ which reads as follows:

$$E S^2_{APP} = \alpha \frac{1+\rho}{(1-\rho)^2} \beta_2 + (1-\alpha) \frac{\beta_2}{(1-\rho)^2},$$

with $0 \leq \alpha \leq 1$.

To determine a suitable choice of the weight factor α we shall require that the approximation is exact for exponential service times. It is easily seen from (4.2) that this requirement yields $\alpha = 1/(2-\rho)$. So, we propose

$$E S^2_{APP} = \left(1 + \frac{\rho}{2-\rho}\right) \frac{\beta_2}{(1-\rho)^2}. \quad (4.4)$$

Note that this approximation has the following appealing properties:

APPROXIMATION PROPERTIES

- (1) The approximation is exact for exponentially distributed service times.
- (2) The approximation yields values between the lower and upper bound of $E S^2$ given by (2.15).
- (3) The approximation satisfies the heavy traffic behaviour of $E S^2$ (see (2.16)).
- (4) The approximation yields the exact value of $E S^2$ for $\rho=0$:

$$E S^2_{APP} = E S^2 = \beta_2, \quad \text{for } \rho=0.$$

The approximation results for the test set of service time distributions and traffic intensities are presented in Table 2. It appears that the approximation yields reasonably good results. In all tested cases the relative approximation error is smaller than 9%. In particular for service time distributions with cv close to one ($0 \leq cv \leq 2$) the relative errors are less than 5.17%. Obviously this is due to the fact that the approximation is exact for exponential service times. For larger values of cv ($cv > 2$) the approximation becomes worse. It is noticeable that the approximation is significantly better for the H_2 distribution with gamma normalization (H_2^{GN}) than for the H_2 distribution with balanced means (H_2^{BM}). In the next subsection we shall show that this is due to the influence of the third moment of the service time distribution on $E S^2$.

4.2 Detailed approximations

The simple approximation (4.4) tends to be less accurate if the squared coefficient of variation of the service time distribution becomes larger. In this subsection we shall develop two new approximations, one for the case that $0 \leq cv \leq 1$ (*APP1*) and one for the case that $cv \geq 1$ (*APP2*). *APP1* is obtained by appropriately weighing the exact values of $E S^2$ for exponential and deterministic service times. *APP2* is based on simple exact expressions for $E S^2$ for two classes of H_2 distributions.

The case $0 \leq cv \leq 1$

For service time distributions with $0 \leq cv \leq 1$ we propose a refinement of $E S^2_{APP}$, $E S^2_{APP1}$, which is based on the exact formulas (2.17) and (2.18) for exponential and deterministic service times. For $0 \leq cv \leq 1$ it is natural to approximate $E S^2$ by a linear interpolation between $E S^2_{EXP}$ and $E S^2_{DET}$:

$$E S^2_{APP1} = cv E S^2_{EXP} + (1-cv) E S^2_{DET}. \quad (4.5)$$

So,

$$E S^2_{APP1} = cv \left(1 + \frac{2+\rho}{2-\rho}\right) \frac{\beta^2}{(1-\rho)^2} + (1-cv) \left(\frac{2\beta^2}{(1-\rho)^2} - \frac{2\beta^2}{\rho^2(1-\rho)}(e^\rho - 1 - \rho)\right). \quad (4.6)$$

Additional to possessing the approximation properties of $E S^2_{APP}$ given in Subsection 4.1, $E S^2_{APP1}$ yields exact results for deterministic service times.

We tested *APP1* for a number of service time distributions: E_4 ($cv=0.25$), E_3 ($cv=0.33$), E_2 ($cv=0.50$), C_2 ($cv=0.75, 0.92$). The results are given in Table 3. It appears, as expected, that *APP1* is much more accurate than the simple approximation *APP* proposed in the previous subsection. The relative error of *APP1* is less than 0.50% in all (test) cases.

The case $cv \geq 1$

For service time distributions with $cv \geq 1$ we shall develop an approximation, *APP2*, for $E S^2$ which is based on simple exact formulas for two classes of extreme H_2 distributions. This approximation

contains the first three moments of the service time distribution.

We start with recalling some characteristics of the class of H_2 distributions. The H_2 distribution function is given by

$$B_{H_2}(t) = \alpha(1 - e^{-t/\tilde{\beta}^{(1)}}) + (1 - \alpha)(1 - e^{-t/\tilde{\beta}^{(2)}}), \quad (4.7)$$

where $0 \leq \alpha \leq 1$, $0 \leq \tilde{\beta}^{(1)} \leq \tilde{\beta}^{(2)}$.

So, there are three parameters. Given the mean $\beta = \alpha\tilde{\beta}^{(1)} + (1 - \alpha)\tilde{\beta}^{(2)}$ and $cv \geq 1$ there is thus one remaining degree of freedom, r , defined by

$$r = \frac{\alpha\tilde{\beta}^{(1)}}{\alpha\tilde{\beta}^{(1)} + (1 - \alpha)\tilde{\beta}^{(2)}}.$$

$r = 1/2$ yields the class of H_2 distributions with balanced means (H_2^{BM}).

Obviously, if β and cv are given, r determines the third moment, β_3 , of the H_2 distribution. For fixed β and β_2 (cv), the smallest possible value of β_3 is obtained for $r = 0$. In that case $\beta_3 = \frac{3}{2} \beta_2^2 / \beta$. For $r \rightarrow 1$, $\beta_3 \rightarrow \infty$, (see Whitt [9,10]).

Our numerical experience with respect to H_2 distributions indicates that $E S^2$ becomes smaller when β_3 grows (β and β_2 constant). So (cf. (2.15)), we expect that $E S^2_{H_2}$ has a limit for $\beta_3 \rightarrow \infty$, β and β_2 fixed. From the formulas for $E S^2$ given in Section 3 of [1] it is found that, for β and β_2 fixed,

$$E S^2_{H_2^{-1}} = \lim_{\beta_3 \rightarrow \infty (r \rightarrow 1)} E S^2_{H_2} = \frac{\beta_2}{(1 - \rho)^2} + \frac{2\rho}{2 - \rho} \frac{\beta^2}{(1 - \rho)^2}, \quad (4.8)$$

and, for $\beta_3 = \frac{3}{2} \beta_2^2 / \beta$ ($r = 0$),

$$E S^2_{H_2^{-0}} = \left(1 + \frac{\rho}{2 - \rho}\right) \frac{\beta_2}{(1 - \rho)^2}. \quad (4.9)$$

It is easily seen that, for $cv \geq 1$,

$$E S^2_{H_2^{-1}} \leq E S^2_{H_2^{-0}}, \quad 0 \leq \rho \leq 1. \quad (4.10)$$

In (4.10), equality holds if $cv = 1$ ($\beta_2 = 2\beta^2$), i.e. if the service times are exponentially distributed. Note that (cf. (4.4)),

$$E S^2_{H_2^{-0}} = E S^2_{APP}.$$

This explains why approximation APP yields better results for H_2^{GN} service time distributions (with a relatively small third moment) than for H_2^{BM} service time distributions.

Now we introduce two approximation assumptions to extend the above results with respect to H_2 distributions to general service time distributions.

Assumption 1: $E S^2$ depends only on the first three moments (β , β_2 , β_3) of the service time distribution.

Assumption 2: $E S^2$ decreases if β_3 grows (β and β_2 fixed).

Under these assumptions it follows from (4.8) and (4.9) that, for $cv \geq 1$, $\beta_3 \geq \frac{3}{2}\beta_2/\beta$,

$$\frac{\beta_2}{(1-\rho)^2} + \frac{2\rho}{2-\rho} \frac{\beta^2}{(1-\rho)^2} \leq E S^2 \leq \left(1 + \frac{\rho}{2-\rho}\right) \frac{\beta_2}{(1-\rho)^2}. \quad (4.11)$$

(4.8), (4.9) and (4.11) suggest an approximation, *APP2*, for $E S^2$ which reads as follows:

$$E S^2_{APP2} = \gamma \left(1 + \frac{\rho}{2-\rho}\right) \frac{\beta_2}{(1-\rho)^2} + (1-\gamma) \left(\frac{\beta_2}{(1-\rho)^2} + \frac{2\rho}{2-\rho} \frac{\beta^2}{(1-\rho)^2}\right), \quad (4.12)$$

where $\gamma := \gamma(\rho, \beta, \beta_2, \beta_3)$, $0 \leq \gamma \leq 1$.

The choice of the weight factor γ will be partially determined by the approximation properties listed below (4.4). Besides these four properties we require that

(5) for β, β_2 fixed,

$$\lim_{\beta_3 \rightarrow \infty} E S^2_{APP2} = \frac{\beta_2}{(1-\rho)^2} + \frac{2\rho}{2-\rho} \frac{\beta^2}{(1-\rho)^2},$$

(6) for $\beta_3 = \frac{3}{2}\beta_2/\beta$,

$$E S^2_{APP2} = \left(1 + \frac{\rho}{2-\rho}\right) \frac{\beta_2}{(1-\rho)^2}.$$

Note, that, without any further specification of γ , *APP2* satisfies the approximation properties (1) and (4). Considering the other required properties ((2), (3), (5) and (6)) it is natural to choose γ as follows,

$$\gamma = \frac{1}{1 + \gamma_1 \left(\beta_3 - \frac{3}{2}\beta_2/\beta\right)(1-\rho)}, \quad (4.13)$$

where γ_1 represents the relative influence of β_3 on $E S^2$. γ_1 remains to be specified. We assume that γ_1 depends only on β and β_2 . Note that γ_1 has to be chosen such that γ is dimensionless. The most obvious choices are $\gamma_1 = 1/\beta^3$ or $\gamma_1 = 1/(\beta\beta_2)$. For both cases we compared approximation results with exact results. Our test set consisted of H_2 service time distributions with cv ranging from 1 to 20. For each value of cv a large number of β_3 values was considered. It appeared that the choice $\gamma_1 = 1/(\beta\beta_2)$ yields much better results than $\gamma_1 = 1/\beta^3$. However, in most cases the choice $\gamma_1 = 1/(\beta\beta_2)$ underestimated $E S^2$. In particular for larger values of cv the approximation results became worse. Extensive tests of the approximation for some variants of $\gamma_1 = 1/(\beta\beta_2)$ led to a modification which yields remarkably accurate results:

$$\gamma_1 = \frac{1}{(cv-1)} \frac{1}{\beta\beta_2}.$$

So, the ultimate approximation formula is given by (4.12), with

$$\gamma = \frac{1}{1 + (1-\rho)\left(\frac{\beta_3}{\beta\beta_2} - \frac{3}{2}\frac{\beta_2}{\beta^2}\right) / \left(\frac{\beta_2}{\beta^2} - 2\right)}. \quad (4.14)$$

It is seen from Table 4 that for H_2^{BM} and H_2^{GN} service time distributions (with $cv=2, 4, 6$) *APP2* yields very accurate results with relative errors less than 1%.

Table 5 illustrates the influence of β_3 on $E S^2$. This table shows exact values of $E S^2$ for a number of H_2 distributions with the same first and second moment but with a different third moment. The traffic intensity varies from 0.1 to 0.95. The relative approximation errors of *APP2* are indicated below the exact values of $E S^2$. As we stated before $E S^2$ decreases when β_3 grows. Note that even for large cv ($cv=10$) the relative approximation errors are less than 1.5%. It may be concluded from Table 5 that the influence of the third moment of the service time distribution on $E S^2$ increases when cv grows, cf. (2.15).

In Table 6 *APP2* is tested for some arbitrarily chosen H_3 and C_2 service time distributions. The relative errors are in all cases less than 1.5%.

Originally, *APP2* has been developed for service time distributions with $cv \geq 1$, $\beta_3 \geq \frac{3}{2} \beta_2^2 / \beta$. For these cases $E S^2_{APP2}$ can be interpreted as an interpolation formula, see (4.12). Nevertheless, approximation formula (4.12) (together with (4.14)) can be applied to service time distributions with $cv < 1$ or $\beta_3 < \frac{3}{2} \beta_2^2 / \beta$ as well. In Table 7 some results are shown for deterministic, C_2 and E_k service time distributions with $cv < 1$. It appears that the accuracy of *APP2* for these cases is about the same as the accuracy of *APP1*.

5. CONCLUSIONS

In this paper we have studied the second moment of the conditional and unconditional sojourn time, $E S^2(x)$ and $E S^2$, for the M/G/1 processor sharing queue. An upper bound and some asymptotic properties (like the heavy traffic behaviour) have been derived. Based on these properties and on exact expressions for specific service time distributions we developed some simple approximations. The approximations have been compared with exact results for a large number of different service time distributions and a wide range of traffic intensities. We conclude as follows.

- The influence of the third and higher moments of the service time distribution on $E S^2(x)$ and $E S^2$ is limited. An upper and a lower bound for $E S^2(x)$ can be expressed in terms of x (the service demand of a tagged customer) and the traffic intensity ρ , see (2.6). The corresponding upper and lower bound for $E S^2$ contain only the second moment of the sojourn time distribution and ρ , see (2.15).

- Approximation *APPX* for $E S^2(x)$, given by (3.1), is based on an asymptotic result for $x \rightarrow 0$ derived in [1]. It depends on the service time distribution only through its first moment. *APPX* yields reasonably good results for not too large values of x , see Table 1. For $0 \leq x \leq \beta$ the relative error of the approximation is a few percent. The approximation becomes less accurate when x increases. For $x = 2\beta$ the relative errors are typically less than 7%. *APPX* satisfies the heavy traffic behaviour of $E S^2(x)$, see (2.8).

- The approximations for $E S^2$, *APP*, *APP1* and *APP2*, given by (4.4), (4.6) and (4.12), have been constructed in such a way that they have the following appealing properties:

- * they are exact for exponential service times
- * they yield values between the lower and upper bound of $E S^2$
- * they satisfy the heavy traffic behaviour of $E S^2$
- * they yield the exact value of $E S^2$ for $\rho=0$.

In addition, *APP1* yields exact results for deterministic service times; *APP2* is exact for two classes

of extreme H_2 distributions.

- Approximation *APP* is the most simple approximation. It depends on the first two moments of the service time distribution. For not too large values of cv ($cv \leq 6$) it yields fairly accurate results, see Table 2. In practical situations *APP* may be applied as a first order approximation for $E S^2$.

- *APP1*, a refinement of *APP*, has been constructed for service time distributions with $0 \leq cv \leq 1$. It depends also on the first two moments of the service time distribution. *APP1* is very accurate. The relative approximation error is less than 0.4% in all of our examples, see Table 3.

- *APP2* depends on the first three moments of the service time distribution. It is based on exact formulas of $E S^2$ for two classes of extreme H_2 distributions. The details of the construction of *APP2* are rather heuristic. Nevertheless, it yields remarkably accurate results. *APP2* has been tested for a large number of different service time distributions with cv ranging from 0 to 10, see Tables 4 through 7. In all of these cases the relative error is less than 1.5%.

ACKNOWLEDGEMENT

The author is indebted to Prof. O.J. Boxma and Prof. J.W. Cohen for some useful comments.

REFERENCES

- 1 VAN DEN BERG, J.L., BOXMA, O.J. (1989). Sojourn times in feedback and processor sharing queues. In: *Teletraffic Science for new Cost-Effective Systems, Networks and Services*, ITC 12, ed. M. Bonatti (North-Holland Publ. Cy., Amsterdam).
- 2 VAN DEN BERG, J.L. (1988). Unpublished work.
- 3 COHEN, J.W. (1982). *The Single Server Queue*, 2nd ed. (North-Holland Publ. Cy., Amsterdam).
- 4 JAISWAL, N.K. (1982). Performance evaluation studies for time sharing computer systems, *Perf. Eval.* **2**, 223-236.
- 5 KLEINROCK, L. (1976). *Queueing Systems*, Vol. II (Wiley, New York).
- 6 OTT, T.J. (1984). The sojourn-time distribution in the M/G/1 queue with processor sharing, *J. Appl. Prob.* **21**, 360-378.
- 7 SCHAASBERGER, R. (1984). A new approach to the M/G/1 processor-sharing queue, *Adv. Appl. Prob.* **16**, 202-213.
- 8 TIJMS, H.C. (1986). *Stochastic Modeling and Analysis* (Wiley, New York).
- 9 WHITT, W. (1982). Approximating a point process by a renewal process, I: Two basic methods, *Oper. Res.* **30**, 125-147.
- 10 WHITT, W. (1984). On approximations for queues, III: Mixtures of exponential distributions, *AT&T Bell Lab. Tech. J.* **63**, 163-175.
- 11 YASHKOV, S.F. (1983). A derivation of response time distribution for an M/G/1 processor-sharing queue, *Problems Contr. & Info. Theory* **12**, 133-148.
- 12 YASHKOV, S.F. (1986). A note on asymptotic estimates of the sojourn time variance in the M/G/1 queue with processor-sharing, *Syst. Anal. Model. Simul.* **3**, 267-269.
- 13 YASHKOV, S.F. (1987). Processor-sharing queues: some progress in analysis, *Queueing Systems* **2**, 1-17.

TABLE 1

Approximation of $E S^2(x)$. The table contains relative errors (%) of approximation APPX (given by (3.1)) for various service time distributions.

Service time distribution: H_2^{BM} , $cv = 2$.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.19	-0.66	-1.35	-2.18	-0.23	-0.80
0.3	-0.32	-1.16	-2.40	-3.94	-0.62	-2.20
0.5	-0.27	-0.99	-2.08	-3.46	-0.99	-3.46
0.7	-0.15	-0.56	-1.19	-1.99	-1.44	-4.80
0.9	-0.04	-0.15	-0.30	-0.51	-2.31	-6.61

Service time distribution: $\overline{H_2^{BM}}$, $cv = 4$.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.22	-0.77	-1.55	-2.46	-0.27	-0.93
0.3	-0.39	-1.39	-2.83	-4.58	-0.76	-2.61
0.5	-0.35	-1.26	-2.58	-4.23	-1.26	-4.23
0.7	-0.21	-0.77	-1.60	-2.64	-1.93	-6.11
0.9	-0.06	-0.23	-0.48	-0.79	-3.39	-8.98

Service time distribution: H_2^{BM} , $cv = 6$.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.24	-0.82	-1.63	-2.59	-0.29	-0.99
0.3	-0.42	-1.50	-3.02	-4.86	-0.81	-2.78
0.5	-0.38	-1.37	-2.80	-4.56	-1.37	-4.56
0.7	-0.24	-0.86	-1.78	-2.92	-2.14	-6.67
0.9	-0.07	-0.27	-0.55	-0.91	-3.85	-10.01

TABLE 1 (Cont'd)

Service time distribution: H_2^{GN} , $cv = 2$.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.25	-0.83	-1.58	-2.45	-0.30	-0.98
0.3	-0.45	-1.51	-2.91	-4.56	-0.85	-2.69
0.5	-0.41	-1.38	-2.68	-4.21	-1.38	-4.21
0.7	-0.26	-0.87	-1.69	-2.64	-1.99	-5.66
0.9	-0.08	-0.27	-0.52	-0.79	-2.82	-7.15

Service time distribution: H_2^{GN} , $cv = 4$.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.34	-1.07	-1.98	-2.98	-0.41	-1.26
0.3	-0.64	-2.03	-3.79	-5.75	-1.17	-3.52
0.5	-0.61	-1.96	-3.68	-5.61	-1.96	-5.61
0.7	-0.41	-1.33	-2.50	-3.81	-2.93	-7.72
0.9	-0.14	-0.46	-0.86	-1.30	-4.33	-10.02

Service time distribution: H_2^{GN} , $cv = 6$.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.38	-1.18	-2.16	-3.23	-0.45	-1.39
0.3	-0.72	-2.26	-4.17	-6.29	-1.31	-3.88
0.5	-0.69	-2.21	-4.11	-6.22	-2.21	-6.21
0.7	-0.48	-1.53	-2.85	-4.32	-3.33	-8.63
0.9	-0.17	-0.54	-1.00	-1.52	-4.98	-11.29

TABLE 1 (Cont'd)

Service time distribution: exponential.

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.14	-0.53	-1.11	-1.85	-0.17	-0.64
0.3	-0.23	-0.86	-1.86	-3.17	-0.45	-1.70
0.5	-0.17	-0.66	-1.46	-2.53	-0.66	-2.53
0.7	-0.08	-0.30	-0.67	-1.19	-0.83	-3.19
0.9	-0.01	-0.04	-0.09	-0.16	-0.96	-3.74

Service time distribution: E_2 .

ρ	$x = \frac{1}{2}\beta$	$x = \beta$	$x = \frac{3}{2}\beta$	$x = 2\beta$	$x = \frac{1}{2} \frac{\beta}{1-\rho}$	$x = \frac{\beta}{1-\rho}$
0.1	-0.03	-0.24	-0.68	-1.32	-0.05	-0.32
0.3	-0.00	-0.24	-0.86	-1.88	-0.06	-0.75
0.5	0.07	0.05	-0.26	-0.91	0.05	-0.91
0.7	0.11	0.27	0.33	0.24	0.32	-0.84
0.9	0.06	0.19	0.34	0.48	0.85	-0.53

TABLE 2

Approximation of $E S^2$. The table contains relative errors (%) of approximation APP (given by (4.4)) for various service time distributions.

Service time distribution: H_2^{BM} .

ρ	$cv=2$	$cv=4$	$cv=6$
0.10	0.82	1.48	1.77
0.30	2.30	4.22	5.07
0.50	3.38	6.31	7.63
0.70	3.63	6.87	8.37
0.90	2.08	4.00	4.91
0.95	1.19	2.30	2.83

Service time distribution: H_2^{GN} .

ρ	$cv=2$	$cv=4$	$cv=6$
0.10	0.40	0.73	0.86
0.30	1.08	1.97	2.36
0.50	1.52	2.78	3.34
0.70	1.53	2.83	3.40
0.90	0.81	1.51	1.82
0.95	0.45	0.84	1.02

Service time distributions with $cv < 1$.

ρ	DET	E_2	C_2^*
0.10	-1.55	-0.40	-0.22
0.30	-3.92	-1.06	-0.60
0.50	-5.11	-1.46	-0.83
0.70	-4.79	-1.45	-0.84
0.90	-2.34	-0.75	-0.45
0.95	-1.29	-0.42	-0.25

* $cv=0.75$

TABLE 3

Approximation of $E S^2$. The table contains relative errors (%) of approximation APP1 (given by (4.6)) for various service time distributions with $cv < 1$.

ρ	E_4	E_3	E_2	$C_2^{(1)}$	$C_2^{(2)}$
0.10	0.18	0.14	0.13	0.00	0.00
0.30	0.35	0.35	0.28	0.00	-0.02
0.50	0.38	0.38	0.32	-0.07	-0.09
0.70	0.25	0.24	0.21	-0.13	-0.11
0.90	0.05	-0.01	0.04	-0.11	-0.07
0.95	0.02	-0.01	0.02	-0.08	-0.04

$C_2^{(1)}$: $cv = 0.75$, $C_2^{(2)}$: $cv = 0.92$.

TABLE 4

Approximation of $E S^2$. The table contains relative errors (%) of approximation APP2 (given by (4.12) and (4.14)) for various service time distributions with $cv \geq 1$, $\beta_3 \geq \frac{3}{2}\beta_2^2/\beta$.

Service time distribution: H_2^{BM} .

ρ	$cv = 2$	$cv = 4$	$cv = 6$
0.10	-0.15	-0.27	-0.32
0.30	-0.32	-0.58	-0.69
0.50	-0.31	-0.53	-0.61
0.70	-0.12	-0.09	-0.04
0.90	0.08	0.34	0.51
0.95	0.07	0.26	0.39

Service time distribution: H_2^{GN} .

ρ	$cv = 2$	$cv = 4$	$cv = 6$
0.10	-0.12	-0.21	-0.25
0.30	-0.23	-0.41	-0.48
0.50	-0.17	-0.30	-0.35
0.70	-0.01	0.01	0.03
0.90	0.09	0.20	0.26
0.95	0.07	0.14	0.18

TABLE 5

The influence of the third moment of the service time distribution (β_3) on $E S^2$.
 In the table the exact values of $E S^2$ are given. The relative approximation errors (%) of APP2 are indicated in parentheses below the exact values of $E S^2$.

H_2 service time distributions with $\beta=1$, $cv=4$.

β_3	$\rho=0.10$	$\rho=0.30$	$\rho=0.50$	$\rho=0.70$	$\rho=0.90$	$\rho=0.95$
37.500*	6.498 (0.00)	12.00 (0.00)	26.67 (0.00)	85.47 (0.00)	909.1 (0.00)	3810 (0.00)
49.084	6.434 (-0.25)	11.68 (-0.50)	25.65 (-0.40)	82.10 (-0.01)	889.2 (0.26)	3762 (0.19)
68.329	6.388 (-0.26)	11.43 (-0.58)	24.78 (-0.56)	78.74 (-0.15)	864.3 (0.34)	3698 (0.29)
105.42	6.353 (-0.19)	11.24 (-0.47)	24.02 (-0.54)	75.35 (-0.28)	830.6 (0.24)	3609 (0.26)
190.02	6.329 (-0.11)	11.09 (-0.29)	23.41 (-0.38)	72.17 (-0.31)	785.3 (0.01)	3441 (0.08)
310.86	6.318 (-0.07)	11.02 (-0.19)	23.12 (-0.26)	70.47 (-0.25)	751.6 (-0.13)	3295 (-0.09)
716.53	6.309 (-0.03)	10.97 (-0.08)	22.86 (-0.12)	68.85 (-0.14)	709.3 (-0.17)	3063 (-0.22)
1391.8	6.306 (-0.02)	10.95 (-0.04)	22.77 (-0.06)	68.21 (-0.08)	689.0 (-0.13)	2927 (-0.20)
2291.9	6.305 (-0.01)	10.94 (-0.03)	22.73 (-0.04)	67.94 (-0.05)	679.6 (-0.09)	2856 (-0.16)
4541.9	6.304 (-0.00)	10.93 (-0.01)	22.70 (-0.02)	67.74 (-0.03)	671.9 (-0.05)	2794 (-0.10)
∞	6.303 (0.00)	10.92 (0.00)	22.67 (0.00)	67.52 (0.00)	663.6 (0.00)	2724 (0.00)

$$* \beta_3 = \frac{3}{2} \beta_2^2 / \beta$$

TABLE 5 (Cont'd)

 H_2 service time distributions with $\beta=1$, $cv=10$.

β_3	$\rho=0.10$	$\rho=0.30$	$\rho=0.50$	$\rho=0.70$	$\rho=0.90$	$\rho=0.95$
181.50*	14.29 (0.00)	26.41 (0.00)	58.67 (0.00)	188.0 (0.00)	2000 (0.00)	8381 (0.00)
223.21	14.17 (-0.29)	25.80 (-0.50)	56.78 (-0.35)	181.9 (0.05)	1965 (0.28)	8298 (0.19)
330.00	14.01 (-0.40)	24.95 (-0.80)	53.89 (-0.68)	171.3 (0.03)	1891 (0.70)	8110 (0.53)
502.34	13.90 (-0.31)	24.34 (-0.73)	51.62 (-0.72)	161.6 (-0.04)	1802 (1.00)	7859 (0.86)
710.16	13.84 (-0.24)	24.00 (-0.60)	50.26 (-0.65)	155.0 (-0.09)	1724 (1.13)	7612 (1.09)
1323.2	13.78 (-0.14)	23.61 (-0.37)	48.65 (-0.44)	146.4 (-0.11)	1588 (1.13)	7093 (1.37)
1845.6	13.76 (-0.10)	23.49 (-0.27)	48.11 (-0.34)	143.2 (-0.10)	1523 (1.04)	6797 (1.42)
4162.2	13.73 (-0.05)	23.31 (-0.13)	47.31 (-0.17)	138.4 (-0.06)	1401 (0.69)	6129 (1.26)
12263	13.72 (-0.02)	23.22 (-0.04)	46.89 (-0.06)	135.7 (-0.03)	1315 (0.30)	5543 (0.71)
30488	13.71 (-0.01)	23.19 (-0.02)	46.76 (-0.02)	134.8 (-0.01)	1285 (0.13)	5305 (0.34)
∞	13.71 (0.00)	23.17 (0.00)	46.67 (0.00)	134.2 (0.00)	1264 (0.00)	5124 (0.00)

$$* \beta_3 = \frac{3}{2} \beta_2^2 / \beta$$

TABLE 6

Relative approximation errors (%) of APP2 for H_3 and C_2 service time distributions.

ρ	$H_3^{(1)}$	$H_3^{(2)}$	$C_2^{(1)}$	$C_2^{(2)}$	$C_2^{(3)}$
0.10	-0.20	-0.29	-0.15	-0.30	-0.32
0.30	-0.54	-0.71	-0.31	-0.63	-0.63
0.50	-0.65	-0.82	-0.25	-0.54	-0.48
0.70	-0.44	-0.48	-0.04	-0.04	0.05
0.90	-0.00	0.20	0.12	0.40	0.39
0.95	0.00	0.10	0.09	0.31	0.28

$H_3^{(1)}$: $cv = 2.778$, $\beta_3 = 40.963$.

$H_3^{(2)}$: $cv = 4.130$, $\beta_3 = 85.622$.

$C_2^{(1)}$: $cv = 2.200$, $\beta_3 = 18.240$.

$C_2^{(2)}$: $cv = 5.000$, $\beta_3 = 84.000$.

$C_2^{(3)}$: $cv = 8.556$, $\beta_3 = 187.33$.

In all cases: $\beta = 1$.

TABLE 7

Relative approximation errors (%) of APP2 for various service time distributions with $cv \leq 1$ and $\beta_3 \leq \frac{3}{2} \beta_2^2 / \beta$.

ρ	DET	E_4	E_2	C_2^*
0.10	-0.02	0.16	0.12	0.06
0.30	-0.18	0.23	0.22	0.11
0.50	-0.36	0.16	0.18	0.10
0.70	-0.44	-0.02	0.05	0.04
0.90	-0.25	-0.10	-0.04	-0.02
0.95	-0.13	-0.08	-0.03	-0.02

* $cv = 0.75$

