



ANDREAS BRANDT<sup>1</sup>, MANFRED BRANDT

**A sample path relation for the sojourn  
times in  $G/G/1 - PS$  systems and its  
applications**

---

<sup>1</sup>Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin, Germany

# A sample path relation for the sojourn times in $G/G/1 - PS$ systems and its applications<sup>1</sup>

Andreas Brandt

*Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin,  
Spandauer Str. 1, D-10178 Berlin, Germany*

Manfred Brandt

*Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB),  
Takustr. 7, D-14195 Berlin, Germany*

## Abstract

For the general  $G/G/1$  processor sharing (PS) system a sample path result for the sojourn times in a busy period is proved, which yields a relation between the sojourn times under PS and FCFS discipline. In particular, the result provides a formula for the mean sojourn time in  $G/D/1 - PS$  in terms of the mean sojourn time in the corresponding  $G/D/1 - FCFS$ , generalizing known results for  $GI/M/1$  and  $M/GI/1$ . Extensions of the formula provide the basis for a two-moment approximation of the mean sojourn time in  $G/GI/1 - PS$  in terms of a related  $G/D/1 - FCFS$ .

**Mathematics Subject Classification (MSC 2000):** 60K25, 68M20, 60G10.

**Keywords:**  $G/G/1$ ;  $G/GI/1$ ;  $G/D/1$ ;  $G/M/1$ ; processor sharing; sojourn time; waiting time; busy period; sample path; first come first served; two-moment approximation.

## 1 Introduction

In this paper we consider the  $G/G/1-PS$  system, i.e., we assume that the input is given by a stationary ergodic marked point process  $\Phi = \{[T_\ell, S_\ell]\}_{\ell=-\infty}^{\infty}$  on  $\mathbb{R}$  with  $\dots \leq T_{-1} \leq T_0 \leq 0 < T_1 \leq T_2 \leq \dots$ , where  $T_\ell$  denotes the arrival instant of the  $\ell$ -th request and  $S_\ell$  its service requirement. The requests are served by a single server under processor sharing (PS) discipline, i.e., if  $n (> 0)$  requests are in the single server then each request receives  $1/n$  of the service capacity<sup>2</sup>. Besides the PS discipline, later we consider the corresponding single server system with the same input  $\Phi$  and infinite waiting

---

<sup>1</sup>This work was supported by a grant from the Siemens AG.

<sup>2</sup>This type of PS discipline is often called the egalitarian processor sharing discipline, cf. [Y] p.102.

room under the first come first served (FCFS) discipline, abbreviated by  $G/G/1 - FCFS$ .

Since the PS and FCFS discipline are work conserving disciplines, their corresponding work load processes and hence busy periods are identical, and under both disciplines the stability condition reads

$$\rho := \frac{m_B}{m_A} < 1, \tag{1.1}$$

where  $m_B$  is the mean service time and  $m_A$  the mean inter-arrival time, cf. e.g. [BB], [BFL], [FKAS]. Under (1.1) a uniquely determined stationary work load process can be constructed, having empty periods and thus busy periods of finite length. By means of this fact for the  $G/G/1 - PS$  system a uniquely determined stationary state process of the number of requests and their residual service times can be constructed along the lines of [BFL], but which will not be outlined here. Note, although the workload process and hence also the busy periods are identical under the PS and FCFS discipline this is clearly not the case for the sojourn times of requests, which are very sensitive with respect to the service discipline.

Processor sharing systems have been studied extensively in many papers, see e.g. [Y] and the references therein. Some recent papers are [ZB], [N], [UB], [BBJ]. However, since the purpose of these studies mainly was to determine sojourn time characteristics of a request, e.g. its mean and variance (given its service requirement), independence as well as distributional assumption are supposed. However, in the general case, few structural properties seem to be known. For general results see [BT], [BB2].

The aim of this paper is to prove in Section 2 (Lemma 2.1) a sample path result for the sojourn times of a busy period in  $G/G/1 - PS$ . As an application in Section 3.1 there is given a general relation (Theorem 3.1) between the sojourn times under PS and FCFS discipline, yielding in particular a formula (Corollary 3.1) for the mean stationary sojourn time in  $G/D/1 - PS$  in terms of the mean stationary sojourn time under FCFS, generalizing corresponding known results for  $GI/M/1$  and  $M/GI/1$ . Then, extensions of the formula (Theorem 3.2) are given for a large subset of  $G/GI/1$  systems. By means of these extensions in Section 3.2 we propose – based on a two-moment matching of the service time distribution – an approximation for the mean sojourn time in a  $G/GI/1 - PS$  in terms of the mean waiting time in a related  $G/D/1 - FCFS$ , thus reducing the complexity.

## 2 A sample path relation for the sojourn times in a busy period under PS discipline

For the  $G/G/1 - PS$  system let us consider a sample path of a busy period<sup>1</sup> during which  $n$  requests arrive and are served. Denote by  $v_k$  and  $s_k$ ,  $k = 1, \dots, n$ , the sojourn time and service requirement of the  $k$ -th arriving request ordered according to their arrivals, respectively. Further, let  $v_k^*$  be the sojourn time of the  $k$ -th request if the arrival process is stopped after the arrival of the  $k$ -th request.

**Lemma 2.1** *It holds*

$$\sum_{k=1}^n (v_k + s_k) = 2 \sum_{k=1}^n v_k^*. \quad (2.1)$$

**Proof.** The proof will be given by induction on  $n$ . Since for  $n = 1$  it holds  $v_1 = s_1 = v_1^*$ , obviously (2.1) is valid. Assume now that (2.1) is true for busy periods for which at most  $n$  requests arrive. Consider an arbitrary busy period during which  $n + 1$  requests are served and which arrive at the time instants

$$0 = t_1 \leq t_2 \leq \dots \leq t_n \leq t_{n+1}$$

and have service times  $s_k$ ,  $k = 1, \dots, n + 1$ . The departure instant  $\tau_k$  of the  $k$ -th request is given by

$$\tau_k = t_k + v_k, \quad k = 1, \dots, n + 1. \quad (2.2)$$

Note that the busy period has length  $\sum_{k=1}^{n+1} s_k$  in view of the work conservation law. The induction step is divided into several steps.

**1.** Consider the modification of the busy period where the  $(n + 1)$ -st arriving request is not considered. Hence the modified busy period consists of  $n$  request. Denote the corresponding sojourn times of the modified busy period by  $\bar{v}_k$ ,  $k = 1, \dots, n$ . Then the modified departure times  $\bar{\tau}_k$  are given by

$$\bar{\tau}_k = t_k + \bar{v}_k, \quad k = 1, \dots, n. \quad (2.3)$$

Applying the induction assumption to the modified busy period and taking into account that  $v_k^*$ ,  $k = 1, \dots, n$ , does not depend on the requests arriving

---

<sup>1</sup>See e.g. [GH] p. 11.

after the  $k$ -th request, we obtain

$$\begin{aligned}\sum_{k=1}^{n+1}(v_k + s_k) &= \sum_{k=1}^n(\bar{v}_k + s_k) + \sum_{k=1}^n(v_k - \bar{v}_k) + v_{n+1} + s_{n+1} \\ &= 2\sum_{k=1}^n v_k^* + \sum_{k=1}^n(v_k - \bar{v}_k) + v_{n+1} + s_{n+1}.\end{aligned}\quad (2.4)$$

**2.** Let  $\pi$  be a permutation of  $\{1, \dots, n\}$  such that

$$\bar{\tau}_{\pi(1)} \geq \bar{\tau}_{\pi(2)} \geq \dots \geq \bar{\tau}_{\pi(n)}.$$

In the following we will exploit several times the fact that under the PS discipline the service capacity is divided equally among the requests being in the system, which implies that the residual service times of the requests in the system are reduced with equal speed. Therefore it holds

$$\tau_{\pi(1)} \geq \tau_{\pi(2)} \geq \dots \geq \tau_{\pi(n)}, \quad (2.5)$$

too. In view of (2.5), there are indices  $i, j \in \{0, \dots, n\}$  such that

$$\tau_{\pi(1)} \geq \dots \geq \tau_{\pi(j)} \geq t_{n+1} > \tau_{\pi(j+1)} \geq \dots \geq \tau_{\pi(n)},$$

$$\tau_{\pi(1)} \geq \dots \geq \tau_{\pi(i)} \geq \tau_{n+1} > \tau_{\pi(i+1)} \geq \dots \geq \tau_{\pi(n)},$$

respectively. Obviously,  $1 \leq j \leq n$  and  $0 \leq i \leq j$ . Further, it holds

$$\bar{\tau}_{\pi(k)} = \tau_{\pi(k)}, \quad k > j, \quad (2.6)$$

$$\bar{\tau}_{\pi(k)} - \bar{\tau}_{\pi(k+1)} = \tau_{\pi(k)} - \tau_{\pi(k+1)}, \quad k < i, \quad (2.7)$$

in view of the PS discipline.

For the second summand on the r.h.s. of (2.4) from (2.2), (2.3), (2.6) and (2.7) we obtain

$$\begin{aligned}\sum_{k=1}^n(v_k - \bar{v}_k) &= \sum_{k=1}^n(v_{\pi(k)} - \bar{v}_{\pi(k)}) = \sum_{k=1}^n(\tau_{\pi(k)} - \bar{\tau}_{\pi(k)}) \\ &= \sum_{k=1}^{j-1}(k+1)\left((\tau_{\pi(k)} - \bar{\tau}_{\pi(k)}) - (\tau_{\pi(k+1)} - \bar{\tau}_{\pi(k+1)})\right) \\ &\quad + (j+1)(\tau_{\pi(j)} - \bar{\tau}_{\pi(j)}) - (\tau_{\pi(1)} - \bar{\tau}_{\pi(1)}) \\ &= \sum_{k=\max(i,1)}^{j-1}(k+1)\left((\tau_{\pi(k)} - \tau_{\pi(k+1)}) - (\bar{\tau}_{\pi(k)} - \bar{\tau}_{\pi(k+1)})\right) \\ &\quad + (j+1)(\tau_{\pi(j)} - \bar{\tau}_{\pi(j)}) - \tau_{\pi(1)} + \bar{\tau}_{\pi(1)}.\end{aligned}\quad (2.8)$$

**3.** Next, by expressing the summands on the r.h.s. of (2.8) in terms of departure times of the original busy period we will prove that

$$\sum_{k=1}^n (v_k - \bar{v}_k) = \tau_{n+1} - t_{n+1} - \sum_{k=1}^{n+1} s_k + \bar{\tau}_{\pi(1)}. \quad (2.9)$$

**3.1** Let  $0 < i = j \leq n$ . Then  $(\bar{\tau}_{\pi(j)} - t_{n+1}) - (\tau_{\pi(j)} - t_{n+1})$  is just the reduction of  $v_{\pi(j)}$  since  $t_{n+1}$  in the modified busy period compared to the original one as the portion of the service capacity assigned to each of the requests in the system is raised from  $1/(j+1)$  to  $1/j$  until  $\tau_{n+1}$ . Hence

$$\bar{\tau}_{\pi(j)} - \tau_{\pi(j)} = -\frac{1}{j+1} (\tau_{n+1} - t_{n+1}).$$

Thus (2.8) provides

$$\sum_{k=1}^n (v_k - \bar{v}_k) = \tau_{n+1} - t_{n+1} - \tau_{\pi(1)} + \bar{\tau}_{\pi(1)}. \quad (2.10)$$

Since  $i > 0$  implies  $\tau_{\pi(1)} \geq \tau_{n+1}$ , it follows that the original busy period, which has length  $\sum_{k=1}^{n+1} s_k$ , ends at  $\tau_{\pi(1)}$ , and thus (2.10) yields (2.9).

**3.2** Let  $0 \leq i < j \leq n$ . Then  $(\bar{\tau}_{\pi(j)} - t_{n+1}) - (\tau_{\pi(j)} - t_{n+1})$  is just the reduction of  $v_{\pi(j)}$  since  $t_{n+1}$  in the modified busy period compared to the original one as the portion of the service capacity assigned to each of the requests is raised from  $1/(j+1)$  to  $1/j$ . Hence

$$\bar{\tau}_{\pi(j)} - \tau_{\pi(j)} = -\frac{1}{j+1} (\tau_{\pi(j)} - t_{n+1}). \quad (2.11)$$

If  $i < k < j$ , then  $(\bar{\tau}_{\pi(k)} - \bar{\tau}_{\pi(k+1)}) - (\tau_{\pi(k)} - \tau_{\pi(k+1)})$  is the reduction of  $v_{\pi(k)}$  since  $\tau_{\pi(k+1)}$  as the portion of the service capacity assigned to each of the requests is raised from  $1/(k+1)$  to  $1/k$ . Hence

$$(\bar{\tau}_{\pi(k)} - \bar{\tau}_{\pi(k+1)}) - (\tau_{\pi(k)} - \tau_{\pi(k+1)}) = -\frac{1}{k+1} (\tau_{\pi(k)} - \tau_{\pi(k+1)}). \quad (2.12)$$

Finally, because  $(\bar{\tau}_{\pi(i)} - \bar{\tau}_{\pi(i+1)}) - (\tau_{\pi(i)} - \tau_{\pi(i+1)})$  is the reduction of  $v_{\pi(i)}$  since  $\tau_{\pi(i+1)}$  as the portion of the service capacity assigned to each of the requests is raised from  $1/(i+1)$  to  $1/i$  until  $\tau_{n+1}$ , it holds

$$(\bar{\tau}_{\pi(i)} - \bar{\tau}_{\pi(i+1)}) - (\tau_{\pi(i)} - \tau_{\pi(i+1)}) = -\frac{1}{i+1} (\tau_{n+1} - \tau_{\pi(i+1)}). \quad (2.13)$$

Thus from (2.8) and (2.11)–(2.13) it follows

$$\sum_{k=1}^n (v_k - \bar{v}_k) = \mathbb{1}\{i > 0\} (\tau_{n+1} - \tau_{\pi(1)}) - t_{n+1} + \bar{\tau}_{\pi(1)}. \quad (2.14)$$

If  $i = 0$  then  $\tau_{n+1} > \tau_{\pi(1)}$ , and therefore the original busy period ends at  $\tau_{n+1} = \sum_{k=1}^{n+1} s_k$ , and hence (2.14) implies (2.9). If  $i > 0$  then  $\tau_{n+1} \leq \tau_{\pi(1)}$ , and the original busy period ends at  $\tau_{\pi(1)} = \sum_{k=1}^{n+1} s_k$ , and hence (2.14) implies (2.9), too.

4. Since the modified busy period ends at  $\bar{\tau}_{\pi(1)}$ , from the work conservation law it follows  $\bar{\tau}_{\pi(1)} = \sum_{k=1}^n s_k$ , and hence from (2.4), (2.9) and  $v_{n+1} = \tau_{n+1} - t_{n+1}$  we obtain

$$\begin{aligned} \sum_{k=1}^{n+1} (v_k + s_k) &= 2 \sum_{k=1}^n v_k^* + \tau_{n+1} - t_{n+1} - \sum_{k=1}^{n+1} s_k + \sum_{k=1}^n s_k + v_{n+1} + s_{n+1} \\ &= 2 \sum_{k=1}^n v_k^* + 2v_{n+1} = 2 \sum_{k=1}^{n+1} v_k^*, \end{aligned}$$

where the last equality is valid as obviously  $v_{n+1}^* = v_{n+1}$ .

□

**Remark 2.1** *The difference of the sojourn and service time of a request in a PS system can be interpreted as the waiting time of the request, cf. [Y] p. 107, namely as the time which the request has to spend additionally to its service time in the system in view of the presence of other requests. In terms of waiting times, (2.1) is equivalent to*

$$\sum_{k=1}^n w_k = 2 \sum_{k=1}^n w_k^*, \quad (2.15)$$

where  $w_k := v_k - s_k$  and  $w_k^* := v_k^* - s_k$  are the waiting times of the  $k$ -th request in the busy period of the original and of the modified system, where the arrival process is stopped after the  $k$ -th request, respectively.

### 3 Applications

In this section we derive some relations between the stationary sojourn times for  $G/G/1$  under  $PS$  and  $FCFS$ . Moreover, we propose an approximation for the mean stationary sojourn time in  $G/GI/1 - PS$  in terms of the mean stationary waiting time in a related  $G/D/1 - FCFS$ .

### 3.1 Relations between sojourn times for $G/G/1$ under $PS$ and $FCFS$

Lemma 2.1 can be used for deriving an upper bound for the mean sojourn time in a  $G/G/1 - PS$  system in terms of the mean sojourn time in the corresponding  $G/G/1 - FCFS$  system. As in Lemma 2.1 consider a sample path of a busy period of  $n$  requests with service times  $s_k$ ,  $k = 1, \dots, n$ . Additionally to the sojourn time  $v_k^P (= v_k)$  of the  $k$ -th request under the PS discipline let  $v_k^F$ ,  $k = 1, \dots, n$ , be the corresponding sojourn time under the FCFS discipline.

**Theorem 3.1** *The sojourn times  $v_k^P$  of the  $G/G/1 - PS$  system and  $v_k^F$  of the  $G/G/1 - FCFS$  system satisfy*

$$\sum_{k=1}^n (v_k^P - s_k) \leq 2 \sum_{k=1}^n (v_k^F - s_k). \quad (3.1)$$

For the  $G/D/1$  system it holds equality in (3.1).

**Proof.** Consider a sample path of a busy period of  $n$  requests and let  $0 = t_1 \leq \dots \leq t_n$  be the arrival epochs of the  $n$  requests. Analogously to Lemma 2.1 let  $v_k^*$  be the sojourn time of the  $k$ -th request in the PS system if the arrival process is stopped after the arrival of the  $k$ -th request. Obviously, the work conservation law provides

$$v_k^* \leq \sum_{j=1}^k s_j - t_k, \quad k = 1, \dots, n, \quad (3.2)$$

as the busy period induced by the first  $k$  arrivals  $(t_1, s_1), \dots, (t_k, s_k)$  has length  $\sum_{j=1}^k s_j$ . If all service times are equal, i.e., in case of a  $G/D/1 - PS$  system, we have equality in (3.2) because then the requests leave the system in the order of their arrivals. In case of the  $G/G/1 - FCFS$  system it holds

$$v_k^F = \sum_{j=1}^k s_j - t_k, \quad k = 1, \dots, n. \quad (3.3)$$

Combining (2.1), where  $v_k = v_k^P$ , (3.2) and (3.3) we obtain (3.1), where in (3.1) we have equality in case of deterministic service times. □

By means of ergodicity arguments, cf. e.g. [BB], [BFL], [FKAS], the sample path result of Theorem 3.1 yields a corresponding result for the mean stationary sojourn times.



**Corollary 3.1** *Under the stability condition  $\rho < 1$  for the stationary sojourn times  $V^P$  of the  $G/G/1 - PS$  system and  $V^F$  of the corresponding  $G/G/1 - FCFS$  system it holds*

$$EV^P - ES \leq 2(EV^F - ES), \quad (3.4)$$

where  $S$  denotes a generic r.v. of the service time, with equality in case of deterministic service times, i.e., for the  $G/D/1$  system it holds

$$EV^P - ES = 2(EV^F - ES). \quad (3.5)$$

Now we will generalize (3.5) to a broader class of  $G/GI/1$  systems, covering in particular the corresponding known results for  $GI/M/1$  and  $M/GI/1$ . Consider the service time distributions  $B_i(x) := P(S \leq x)$ ,  $i \in \{1, 2\}$ , defined by

$$B_1(x) := 1 - p \mathbb{I}\{x < s\}, \quad x \in \mathbb{R}_+, \quad (3.6)$$

$$B_2(x) := 1 - p \exp(-x/s), \quad x \in \mathbb{R}_+, \quad (3.7)$$

where  $p \in (0, 1]$  and  $s \in \mathbb{R}_+ \setminus \{0\}$ , i.e.,  $B_i(x)$  is a mixture of a zero and a deterministic or exponential time with mean  $s$ , respectively.

**Theorem 3.2** *Let  $\rho < 1$ . Then for the systems  $G/GI/1$  with  $GI = B_i$ ,  $i \in \{1, 2\}$ , and  $M/GI/1$  it holds*

$$EV^P - ES = \frac{2(ES)^2}{ES^2} (EV^F - ES), \quad (3.8)$$

where  $V^P$  and  $V^F$  denote the stationary sojourn times in the corresponding systems under the PS and FCFS discipline, respectively.

**Proof. 1.** For the  $M/GI/1$  system (3.8) follows directly from the well-known formulae for  $EV^P$  and  $EV^F$ , cf. e.g. [Y] p.109 and [W] p.278.

**2.** Consider now the systems  $G/GI/1$  with  $GI = B_i$ ,  $i \in \{1, 2\}$ .

**2.1** First let  $p = 1$ , i.e., we deal with the  $G/D/1$  and  $G/M/1$  system, respectively. Since in case of a  $G/D/1$  system  $ES^2 = (ES)^2$ , (3.8) is an immediate consequence of Corollary 3.1. In case of a  $G/M/1$  system (3.8) is equivalent to  $EV^P = EV^F$  because of  $ES^2 = 2(ES)^2$ . In view of the exponential service times, under the PS as well as under the FCFS discipline the departure process is a Poisson process of intensity  $1/ES$  as long as there are requests in the system. Consequently, the numbers of requests in the

system are stochastically equivalent under both disciplines, and by Little's formula it follows  $EV^P = EV^F$ .

**2.2** Now let  $p \in (0, 1)$ , i.e., the service times are a proper mixture of a zero and a positive (deterministic or exponential) time. Besides the input  $\Phi = \{[T_\ell, S_\ell]\}_{\ell=-\infty}^\infty$  consider the modified input  $\tilde{\Phi} = \{[\tilde{T}_\ell, \tilde{S}_\ell]\}_{\ell=-\infty}^\infty$  with  $\dots \leq \tilde{T}_{-1} \leq \tilde{T}_0 \leq 0 < \tilde{T}_1 \leq \tilde{T}_2 \leq \dots$  consisting of all those requests of  $\Phi$  having a positive service time. In the following let us endow all quantities and variables which are related to  $\tilde{\Phi}$  with a tilde, thus  $\tilde{S}$  denotes the generic deterministic or exponential service time with mean  $s$ , and  $\tilde{V}^P, \tilde{V}^F$  are the stationary sojourn times in the  $G/G/1$  system with input  $\tilde{\Phi}$  under the PS and FCFS discipline, respectively. The following observation is crucial: the requests with a zero service time have under the PS discipline a sojourn time of length zero whereas under the FCFS discipline they do not have any impact on the workload process and hence on the waiting times. This, (3.6), (3.7), and since the service times  $\{S_\ell\}_{\ell=-\infty}^\infty$  are i.i.d. r.v.'s independent on the arrival process, yield in case of  $GI = B_i, i \in \{1, 2\}$ , the relations

$$EV^P = pE\tilde{V}^P, \quad EV^F - ES = E\tilde{V}^F - E\tilde{S}, \quad (3.9)$$

$$ES = pE\tilde{S}, \quad ES^2 = pE\tilde{S}^2. \quad (3.10)$$

Note that  $E\tilde{V}^P$  ( $E\tilde{V}^F - E\tilde{S}$ ) is the mean stationary sojourn time (waiting time) in the corresponding  $G/D/1 - PS$  (FCFS) or  $G/M/1 - PS$  (FCFS) system, respectively, where  $G$  stands for the process  $\{\tilde{T}_\ell\}_{\ell=-\infty}^\infty$  of arrival instants of  $\tilde{\Phi}$ ,  $D$  for the constant service times  $s$  in case of  $B_1(x)$  and  $M$  for the exponential service times with mean  $s$  in case of  $B_2(x)$ . (Note, a rigorous proof of (3.9) can be given by applying Palm's formula for marked point processes and using the independence assumptions of the service times.) From (3.9), (3.10) and step 2.1 above we obtain

$$\begin{aligned} EV^P - ES &= p(E\tilde{V}^P - E\tilde{S}) = p \frac{2(E\tilde{S})^2}{E\tilde{S}^2} (E\tilde{V}^F - E\tilde{S}) \\ &= \frac{2(ES)^2}{ES^2} (EV^F - ES). \end{aligned}$$

□

**Remark 3.1** (i) Note, for the  $GI/M/1$  system the fact  $EV^P = EV^F$  is well-known, cf. e.g. [R] p. 441.

(ii) Theorem 3.2 states that (3.8) holds for a much broader class than for  $GI/M/1$  and  $M/GI/1$  systems.

(iii) Using the squared coefficient of variation  $c_B^2 := ES^2/(ES)^2 - 1$ , relation (3.8) is equivalent to

$$EV^P - ES = \frac{2}{1+c_B^2} (EV^F - ES),$$

where  $c_{B_1}^2 = (1-p)/p \geq 0$  and  $c_{B_2}^2 = (2-p)/p \geq 1$ , cf. (3.6), (3.7).

Although Theorem 3.2 shows that (3.8) holds for a broad class of  $G/GI/1$  systems, this relation is not true for  $G/GI/1$  systems in general. Let us consider the  $M^{GI}/GI/1$  system. Note that the  $M^{GI}$  batch arrival process is a special  $G$  process, cf. e.g. [BB], [BFL]. There arrive batches of positive size  $X$  according to a Poisson process of intensity  $\lambda$ . The mean sojourn time  $EV^F$  of an arriving request under the FCFS discipline is given by, cf. e.g. [T] p. 277,

$$EV^F = ES + \frac{ES}{2(1-\varrho)} \left( \varrho \frac{ES^2}{(ES)^2} + b \right), \quad (3.11)$$

where  $\varrho = \lambda EX ES < 1$ , cf. (1.1), and  $b := EX^2/EX - 1$  is the mean number of requests that arrive together with the tagged arriving request other than this request. Since for the  $M^{GI}/GI/1-PS$  system an explicit formula for  $EV^P$  seems not to be available, let us restrict now to  $H_2$ -distributed service times with balanced means, i.e.,

$$H_2(x) := p_1(1 - \exp(-x/s_1)) + p_2(1 - \exp(-x/s_2)), \quad x \in \mathbb{R}_+, \quad (3.12)$$

where  $p_1 = 1 - p_2 \in (0, 1)$ ,  $s_1, s_2 \in \mathbb{R}_+ \setminus \{0\}$ ,  $s_1 \neq s_2$  and

$$p_1 s_1 = p_2 s_2. \quad (3.13)$$

Specializing the general results of [KMR], [B] for the  $M^{GI}/GI/1-PS$  and  $M^{GI}/GH/1-PS$  system, where GH stands for generalized hyperexponential distribution, respectively, to the  $H_2$  distribution with balanced means, after tedious algebra one obtains that for the  $M^{GI}/H_2/1-PS$  system with (3.13) it holds

$$EV^P = ES + \frac{ES}{2(1-\varrho)} \left( 2\varrho + b - \frac{ab}{2-\varrho} \right), \quad (3.14)$$

where

$$a := \frac{ES^2 - 2(ES)^2}{ES^2} \in (0, 1).$$

From (3.11), (3.14) it follows

$$EV^P - ES = \frac{2(ES)^2}{ES^2} \left( 1 + \frac{(1-\varrho)ab}{(2-\varrho)(2\varrho+(1-a)b)} \right) (EV^F - ES). \quad (3.15)$$

Therefore, for the  $M^{GI}/H_2/1$  system with (3.13) it holds

$$EV^P - ES \geq \frac{2(ES)^2}{ES^2} (EV^F - ES) \quad (3.16)$$

with equality iff  $b = 0$ , in which case we have a  $M/H_2/1$  system covered by Theorem 3.2.

### 3.2 Two-moment approximation for $EV^P$ in $G/GI/1 - PS$

Consider the general  $G/GI/1 - PS$  system with service time distribution  $B(x)$ . The mean  $m_B (> 0)$  and squared coefficient of variation  $c_B^2 (\geq 0)$  of the service times provide a two-moment characterization of  $B(x)$ . Choosing  $p$  and  $s$  as

$$p := \frac{1}{1 + c_B^2}, \quad s := (1 + c_B^2)m_B, \quad (3.17)$$

then  $B_1(x)$ , defined by (3.6), (3.17), is a two-moment approximation of  $B(x)$ , i.e., the first two moments of the service times coincide. Hence the mean sojourn time  $EV_1^P$  in the corresponding  $G/B_1/1 - PS$  system can be considered as an approximation for  $EV^P$ . From Theorem 3.2 and step 2.2 (3.9), (3.10) of its proof it follows

$$EV_1^P = ES + p \frac{2(E\tilde{S})^2}{E\tilde{S}^2} E\tilde{W}_1^F,$$

where  $E\tilde{W}_1^F := E\tilde{V}_1^F - E\tilde{S}$  is the mean stationary waiting time in the corresponding  $G/D/1 - FCFS$  system with  $D = s$  given by (3.17). In view of (3.17), therefore we obtain the two-moment approximation

$$EV^P \approx ES + \frac{2}{1 + c_B^2} E\tilde{W}_1^F \quad (3.18)$$

for  $EV^P$  in terms of  $E\tilde{W}_1^F$ , thus reducing the complexity. Note that  $E\tilde{W}_1^F$  can be computed efficiently in several cases. Numerical studies – not reported here – have shown that (3.18) provides a good approximation, e.g. in case of Poisson-distributed batch arrivals with deterministic inter-arrival times, where  $E\tilde{W}_1^F$  can be computed efficiently via Spitzer's identity for the mean waiting time in  $GI/GI/1 - FCFS$  systems by interpreting an arriving batch as one customer whose service time is proportional to the batch size.

**Remark 3.2** Note, for  $c_B^2 \geq 1$  one obtains an analogous approximation to (3.18) by using  $B_2(x)$  instead of  $B_1(x)$ . Choosing

$$p := \frac{2}{1 + c_B^2}, \quad s := \frac{1 + c_B^2}{2} m_B, \quad (3.19)$$

then  $B_2(x)$  is a two-moment approximation of  $B(x)$ , too, and analogously to (3.18) we obtain the approximation

$$EV^P \approx ES + \frac{2}{1 + c_B^2} E\tilde{W}_2^F, \quad (3.20)$$

where  $E\tilde{W}_2^F$  denotes the mean stationary waiting time in the corresponding  $G/M/1 - FCFS$  system with exponential service times with mean  $s$  given by (3.19). Note that (3.20) is of the same structure as (3.18) and that the r.h.s. of (3.20) is also well defined in the general case of  $c_B^2 \geq 0$ .

**Remark 3.3** Note that (3.18), (3.20) provide approximations for  $EV^P$  for a general  $G/GI$  input. For the  $GI/GI/1 - PS$  system Sengupta [S] suggested to approximate the stationary sojourn time by the product of two independent r.v.'s, where one of them is the service time and the other one has a Gamma distribution, where the parameters are determined from known analytical results for the  $M/GI/1 - PS$  and  $GI/M/1 - PS$  system. The proposed approximation is exact in some limiting cases and provides good approximations for  $GI/GI/1 - PS$  as reported.

## References

- [BB] Baccelli, F., Brémaud, P., *Elements of Queueing Theory*. Applications of Mathematics 26, Springer-Verlag, Berlin, 2003.
- [BT] Baccelli, F., Towsley, D., The customer response times in the processor sharing queue are associated. *Queueing Systems* 7 (1990) 269–282.
- [B] Bansal, N., Analysis of the  $M/G/1$  processor-sharing queue with bulk arrivals. To appear in *Operations Research Letters*.
- [BBJ] Borst, S., Boxma, O., Jelenković, P., Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows. *Queueing Systems* 43 (2003) 273–306.

- [BB1] Brandt, A., Brandt, M., On the sojourn times for many-queue head-of-the-line processor-sharing systems with permanent customers. *Math. Methods Oper. Res.* 47 (1998) 181–220.
- [BB2] Brandt, A., Brandt, M., A note on the stability of the many-queue head-of-the-line processor-sharing system with permanent customers. *Queueing Systems* 32 (1999) 363–381.
- [BFL] Brandt, A., Franken, P., Lisek, B., *Stationary Stochastic Models*. Akademie-Verlag, Berlin; Wiley, Chichester, 1990.
- [FKAS] Franken, P., König, D., Arndt, U., Schmidt, V., *Queues and Point Processes*. Akademie-Verlag, Berlin; Wiley, Chichester, 1982.
- [GH] Gross, D., Harris, C.M., *Fundamentals of Queueing Theory*. Wiley, New York, 1998.
- [KMR] Kleinrock, L., Muntz, R.R., Rodemich, E., The processor-sharing queueing model for time-shared systems with bulk arrivals. *Networks* 1 (1971) 1–13.
- [N] Núñez-Queija, R., Sojourn times in a processor sharing queue with service interruptions. *Queueing Systems* 34 (2000) 351–386.
- [R] Ramaswami, V., The sojourn time in the  $GI/M/1$  queue with processor sharing. *J. Appl. Probab.* 21 (1984) 437–442.
- [S] Sengupta, B., An approximation for the sojourn-time distribution for the  $GI/G/1$  processor-sharing queue. *Commun. Statist., Stochastic Models* 8 (1992) 35–57.
- [T] Tijms, H.C., *Stochastic Models: An Algorithmic Approach*. Wiley, Chichester, 1994.
- [UB] Van Uitert, M., Borst, S.C., A reduced-load equivalence for generalised processor sharing networks with long-tailed input flows. *Queueing Systems* 41 (2002) 123–163.
- [W] Wolff, R.W., *Stochastic Modeling and the Theory of Queues*. Prentice-Hall International, Englewood Cliffs, N.J., 1989.
- [Y] Yashkov, S.F., Mathematical problems in the theory of shared-processor systems. *J. Soviet Math.* 58 (1992) 101–147. (Original

Russian version in: Itogi Nauki i Tekhniki, Seriya Teoriya Veroyatnostei, Matematicheskaya Statistika, Teoreticheskaya Kibernetika 29 (1990) 3–82.)

[ZB] Zwart, A.P., Boxma, O.J., Sojourn time asymptotics in the M/G/1 processor sharing queue. Queueing Systems 35 (2000) 141–166.