



On Time-to-Buffer Overflow Distribution in a Single-Machine Discrete-Time System with Finite Capacity

Wojciech M. Kempa

Silesian University of Technology, Faculty of Applied Mathematics

Kaszubska 23, 44-100 Gliwice, Poland

E-mail(*corresp.*): wojciech.kempa@polsl.pl

Received June 3, 2019; revised February 12, 2020; accepted February 13, 2020

Abstract. A model of a single-machine production system with finite magazine capacity is investigated. The input flow of jobs is organized according to geometric distribution of interarrival times, while processing times are assumed to be generally distributed. The closed-form formula for the generating function of the time to the first buffer overflow distribution conditioned by the initial buffer state is found. The analytical approach based on the idea of embedded Markov chain, the formula of total probability and linear algebra is applied. The corresponding result for next buffer overflows is also given. Numerical examples are attached as well.

Keywords: buffer overflow, geometric distribution, production line, queueing model, transient analysis.

AMS Subject Classification: 60K20; 60K25; 68M20; 90B22.

1 Introduction

The occurrence of a buffer overflow is a typical problem that arises in mass service systems where the accumulation of unprocessed elements is bounded. Of course, the increase in the number of requests that are waiting for service depends on the processing speed, the intensity of the input stream, as well as the number of service stations that can work (process tasks) in parallel. The phenomenon of a buffer overflow usually has a negative impact on the quality of service (QoS), because it involves the loss of a number of potential new customers (tasks, jobs, packets, etc.) that do not have a place where they could wait for service and, consequently, resign, or it is connected with the

necessity to propose an alternative service path (redirection to another service station or to another service system), which generates additional costs.

In particular, examples of practical operating systems in which a buffer storage overflow phenomenon can occur are given below:

- a production line, in which there are accumulating buffers of finite capacities between consecutive processing stations; a buffer overflow at one of the stations usually suspends or significantly slows down the service in the preceding stations;
- buffers accumulating data packets incoming to the packet network node (e.g. IP packets in the Internet routers) have finite capacities (counted in bytes); a buffer overflow results in sending feedback to the source emitting the packets, which makes the decision on their retransmission (in accordance with the TCP/IP protocol);
- on communication routes with a sequence of successive traffic lights as the “buffer” can be treated the section of the route between two successive traffic lights;
- the phenomenon of a buffer overflow also occurs in the case of the organization of parking spaces, e.g. near large shopping centers: full parking (accumulating “buffer”) results in the need for the driver to look for another parking space (there is a redirection of the customer or its loss).

The application of queueing theory in the modeling of practical systems in which a large number of customers is processed is well known. In recent years, stochastic characteristics of systems in which the service is limited in some way, which is connected with the use of, for example, a specific mechanism of the input traffic control, are particularly intensively studied. An example of such type of a mechanism affecting the system throughput is the use of a storage buffer with finite capacity. Moreover, an alternative to the classical approach of queueing systems analysis with a continuous time parameter is the use of discrete time, in which the time axis is divided into the so-called slots, indexed with consecutive natural numbers. This type of approach is sometimes much more desirable from the practical point of view. In production systems or logistic networks, the state of a magazine or a specific accumulating buffer is monitored once a day or at regular time intervals.

In books [3] and [21] one can find the overview of results for stochastic characteristics of different-type queueing models with discrete time parameter. The selection of methods and techniques, and some applications of discrete-time stochastic models can be found in [9]. The exact approach for buffer overflow calculations based on martingales is proposed in [2]. The representation for the distribution of the time to buffer overflow in a queueing model with autocorrelated input stream can be found e.g. in [5] and [6], where the cases of single and batch arrivals are considered, respectively. One can find some other analytical results for the problem of buffer overflow and loss of customers during this time period e.g. in [8] and [18]. In particular, in [8] the formula for the joint transform of the busy period and numbers of customers successfully

served and lost during the busy period is obtained in the model described by phase-type distributions. Similarly, in [18] the representation for the distribution of the number of buffer overflows during one busy period is found for the system with Markovian arrival stream. Analytical results for the cumulative distribution function of the time to the first buffer overflow can be found e.g. in [13] and [14], where the multiple vacation policy and the control mechanism based on setup and closedown times are considered, respectively. The model of a production line with unreliable machine is investigated in [15]. In [12] the distribution of the time to buffer overflow is analyzed in the case of a model with server working vacations. In [20] the representation for the mean time to reach the fixed level in the $M/G/1$ -type queue with infinite buffer is obtained. Similar characteristic for the corresponding finite-buffer queue is studied in [11], where the closed-form representation of the distribution of the time to start a crowded period (to hit the predetermined level) is derived. In [16] (see also [7]) the first overflow time in the $GI/G/1/N$ -type queue is studied via diffusion approximation. Obviously, the time to the first buffer overflow coincides with the first passage time from the fixed initial state to state N . In [4] the representation for the probability generating function of the time to the first buffer overflow is obtained for the $Geo/Geo/1/N$ -type model with arrival and service rates dependent on the number of jobs accumulated in the system. The result is given in terms of roots of the characteristic equation of a certain functional matrix. The result form [4] is generalized in [19] for the case of phase-type distributed interarrival and service times.

In the paper the production system with one processing station is modeled using the $Geo/G/1/B$ -type queue. Using the analytical approach based on the idea of embedded Markov chain and the formula of the total probability, a system of equations is constructed for the distribution of the time to the first buffer overflow, conditioned by the initial state of the accumulating buffer. The solution of the corresponding system written for generating function is found in a compact form using the algebraic potential method introduced by Korolyuk in [17]. So, in consequence, the obtained result generalizes the one derived in [4] for the case of generally-distributed service times. Moreover, in relation to [11] and [20], the discrete-time model is investigated, so instead of Poisson arrival process a binomial process is considered. Analytical results are given in terms of a certain additional sequence (called in [17] a potential). Successive terms of this sequence can be obtained recursively (however, such a procedure gives usually high computational complexity) or by using the Maclaurin expansion. The latter possibility is used in numerical examples.

The remaining part of the article is organized as follows. The next Section 2 presents the exact mathematical description of the considered model and states the basic facts about the binomial process (continuous analogue of the Poisson process) characterizing the input traffic. In Section 3, using the formula of the total probability and the idea of embedded Markov chain, a system of equations for the conditional distribution of the time to the first buffer overflow and a corresponding system written for generating functions are obtained. The solution of the latter system is found in Section 4, where the appropriate result for the subsequent buffer overflow periods is also given. Section 5 contains

numerical examples illustrating theoretical results, while the last Section 6 contains a brief summary and conclusions.

2 Description of the model

We deal with a discrete-time model of a production line with a single processing machine, modeled by a *Geo/G/1/B*-type queue. So, according to the well-known Kendall notation, it is assumed that the stream of incoming jobs is governed by a binomial process in that successive interarrival times are independent random variables with the common geometric distribution with parameter $a \in (0, 1)$, namely

$$a_k \stackrel{def}{=} a(1-a)^{k-1},$$

where $k \in \{1, 2, \dots\}$.

In the considered model service (processing) times are assumed to be generally distributed, where b_k denotes the probability that the service duration of a single customer equals k slots, where $\sum_{k=1}^{\infty} b_k = 1$. The capacity of accumulating buffer is equal to $B - 1$ places, so the maximal system state is B , that is reached in the case of the buffer being saturated and the service machine being busy with processing. In the case the arriving job finds the system in state B , it leaves the system without processing. The classical FIFO service discipline is assumed.

We also accept the situation in that at the same time slot both the arrival and service completion occur. In such a case it does not matter whether the arrival precedes the departure (Arrival First (AF) scheme) or the opposite (Departure First (DF) scheme), since the characteristic being analyzed is independent on it.

Introduce the following probability generating functions of the sequences (a_k) and (b_k) , $k \in \{1, 2, \dots\}$, as follows:

$$A(z) \stackrel{def}{=} \sum_{k=1}^{\infty} z^k a_k = \frac{az}{1 - (1-a)z}, \quad B(z) \stackrel{def}{=} \sum_{k=1}^{\infty} z^k b_k, \quad (2.1)$$

where $|z| < 1$.

3 Equations for conditional time-to-buffer overflow distribution

In this section, by using the paradigm of embedded Markov chain and the total probability law, we establish the system of equations for the conditional distribution of the time to the first buffer overflow and, next, we write the corresponding system for appropriate generating functions.

The binomial arrival process is a discrete-time analogue of the Poisson one. From the point of view of the further analysis, it is important to express the probability $p_i(j)$ that up to fixed time moment $i \in \{1, 2, \dots\}$ exactly $j \in \{0, \dots, i\}$ customers will arrive to the system.

Let us observe that, denoting by θ_k the k th interarrival time, we get

$$\begin{aligned}
 p_i(j) &= \mathbf{P}\left\{\sum_{k=1}^j \theta_k \leq i, \sum_{r=1}^{j+1} \theta_r \geq i + 1\right\} \\
 &= \sum_{l=1}^i \mathbf{P}\left\{\sum_{r=1}^{j+1} \theta_r \geq i + 1 \mid \sum_{k=1}^j \theta_k = l\right\} \mathbf{P}\left\{\sum_{k=1}^j \theta_k = l\right\} \\
 &= \sum_{l=1}^i \mathbf{P}\{\theta_{j+1} \geq i + 1 - l\} \mathbf{P}\left\{\sum_{k=1}^j \theta_k = l\right\} \\
 &= \sum_{l=1}^i (1 - a)^{i-l} \mathbf{P}\left\{\sum_{k=1}^j \theta_k = l\right\}. \tag{3.1}
 \end{aligned}$$

The sum of independent and identically distributed geometric random variables has negative binomial distribution, so we obtain

$$\mathbf{P}\left\{\sum_{k=1}^j \theta_k = l\right\} = \binom{l-1}{j-1} a^j (1-a)^{l-j}, \tag{3.2}$$

where $l \geq j$. Using (3.2) in (3.1) we obtain

$$p_i(j) = a^j (1-a)^{i-j} \sum_{l=j}^i \binom{l-1}{j-1},$$

where $i \in \{1, 2, \dots\}$ and $j \in \{0, \dots, i\}$.

Let us denote by $\beta_1 = \beta$ the time (counting from the initial epoch 0) to the start of the first buffer overflow period, so

$$\beta \stackrel{def}{=} \min\{n \in \mathbb{N} : X_n = B\},$$

where X_n stands for the number of customers (jobs, packets, etc.) present in the system at time n including the one in service (if any).

Introduce the following notation:

$$\beta_n(k) \stackrel{def}{=} \mathbf{P}\{\beta \geq k \mid X_0 = n\},$$

where $n \in \{0, \dots, N - 1\}$ and $k \in \mathbb{N}$.

Assume, firstly, that the accumulating buffer is empty at the starting moment, so $X_0 = 0$. Note that then the following equation is true:

$$\beta_0(k) = \sum_{i=1}^{k-1} a_i \beta_1(k-i) + \sum_{i=k}^{\infty} a_i. \tag{3.3}$$

Indeed, if the first arrival occurs at time $i \leq k - 1$, then at this time the operation of the system “renews” with one job present and the time for the buffer overflow initial moment, to exceed or to be equal to, shortens to $k - i$

(the first summand on the right side of (3.3)). If the first job enters the system at time k or later, then $\{\beta \geq k\}$ with probability one (the second summand).

Let us consider now the case at which $X_0 \geq 1$ and assume that, despite to the presence of a number of jobs in the accumulating buffer before the opening of the system, the processing (with the distribution described by the sequence (b_k)) begins at time 0. The memoryless property of geometric distribution of interarrival times implies that consecutive departure moments are Markov epochs in the evolution of the system. Namely, the subsequence $(X_{k_n}), k \in \mathbb{N}$, where k_j is the j th departure epoch in the evolution of the system, is a Markov chain, called an embedded one (see e.g. [3], [21]). In consequence, applying the formula of total probability with respect to the first departure moment after 0, we can write the following system of equations:

$$\begin{aligned} \beta_n(k) = & \sum_{i=1}^{k-1} b_i \sum_{j=0}^{B-n-1} p_{i-1}(j) [(1-a)\beta_{n+j-1}(k-i) + a\beta_{n+j}(k-i)] \\ & + b_k \sum_{j=0}^{B-n-1} p_{k-1}(j) + \sum_{i=k+1}^{\infty} b_i \sum_{j=0}^{B-n-1} p_k(j), \end{aligned} \tag{3.4}$$

where $1 \leq n \leq B - 1$.

Indeed, the first summand on the right side of (3.4) presents the situation in that the first service completes at the moment $i \leq k - 1$, so at the Markov moment i the system state is $n + j - 1$, where j denotes the number of arrivals up to time $i - 1$ if the epoch k is not the next arrival moment, and is $n + j$ otherwise. If the first processing completes exactly at time $i = k$, then, if only up to the time $k - 1$ the number of arrivals does not exceed $B - n - 1$, the random event $\{\beta \geq k\}$ occurs with probability one. Otherwise, the probability of $\{\beta \geq k\}$ equals zero (compare the second summand on the right side of (3.4)). The last summand on the right side of (3.4) relates to the case in that the first customer leaves the system after time k .

Let us transform the system (3.3)–(3.4) to the corresponding one written for generating functions. So, define

$$\tilde{\beta}_n(z) \stackrel{def}{=} \sum_{k=1}^{\infty} z^k \beta_n(k),$$

where $|z| < 1$. Observe that, since

$$\begin{aligned} & \sum_{k=2}^{\infty} z^k \sum_{i=1}^{k-1} a_i \beta_1(k-i) + \sum_{k=1}^{\infty} z^k \sum_{i=k}^{\infty} a_i \\ & = \sum_{i=1}^{\infty} a_i z^i \sum_{k=i+1}^{\infty} z^{k-i} \beta_1(k-i) + \frac{z}{1-z} \sum_{i=1}^{\infty} (1-z^i) a_i, \end{aligned}$$

then (3.3) gives

$$\tilde{\beta}_0(z) = A(z)\tilde{\beta}_1(z) + \frac{z}{1-z}[1 - A(z)], \tag{3.5}$$

where $A(z)$ is defined in (2.1).

Similarly, since

$$\begin{aligned} & \sum_{k=2}^{\infty} z^k \sum_{i=1}^{k-1} b_i \sum_{j=0}^{B-n-1} p_{i-1}(j) \beta_{n+j-1}(k-i) \\ &= \sum_{j=0}^{B-n-1} \sum_{i=1}^{\infty} z^i b_i p_{i-1}(j) \sum_{k=i+1}^{\infty} z^{k-i} \beta_{n+j-1}(k-i), \end{aligned}$$

then, introducing the following functional sequences:

$$\begin{aligned} \alpha_j(z) &\stackrel{def}{=} (1-a) \sum_{i=1}^{\infty} p_{i-1}(j) b_i z^i, \quad \bar{\alpha}_j(z) \stackrel{def}{=} a \sum_{i=1}^{\infty} p_{i-1}(j) b_i z^i, \\ T_n(z) &\stackrel{def}{=} \sum_{k=1}^{\infty} z^k \sum_{i=k}^{\infty} b_i \sum_{j=0}^{B-n-1} [p_k(j)(1-\delta_{i,k}) + \delta_{i,k} p_{k-1}(j)], \end{aligned}$$

where $|z| < 1$ and $\delta_{i,j}$ stands for the Kronecker delta function, we obtain from (3.4)

$$\tilde{\beta}_n(z) = \sum_{j=0}^{B-n-1} [\alpha_j(z) \tilde{\beta}_{n+j-1}(z) + \bar{\alpha}_j(z) \tilde{\beta}_{n+j}(z)] + T_n(z), \tag{3.6}$$

where $1 \leq n \leq B - 1$. Observe that, defining,

$$u_j(z) \stackrel{def}{=} \alpha_j(z) + (1 - \delta_{j,0}) \bar{\alpha}_{j-1}(z),$$

we can rewrite (3.6) in the following way:

$$\tilde{\beta}_n(z) = \sum_{j=0}^{B-n-1} u_j(z) \tilde{\beta}_{n+j-1}(z) + \bar{\alpha}_{B-n-1}(z) \tilde{\beta}_{B-1}(z) + T_n(z), \tag{3.7}$$

where $1 \leq n \leq B - 1$.

4 Compact-form solution

In this section, we obtain the solution of the system (3.5), (3.7) in terms of a certain functional sequence connected with the given sequence $(u_k(z))$.

Let us take into consideration the following infinite-size system of linear equations with known coefficients τ_n and σ_n and with unknowns x_2, x_3, \dots :

$$\sum_{j=-1}^{n-2} \tau_{j+1} x_{n-j} - x_n = \sigma_n, \quad n \geq 2. \tag{4.1}$$

In [17] one can find the in-depth analysis of such type systems according to the study of the so-called random walks continuous from the right. It is given

in [10] that each solution of (4.1) can be stated as

$$x_n = c \cdot \varkappa_{n-1} + \sum_{k=2}^n \varkappa_{n-k} \sigma_k, \quad n \geq 2, \tag{4.2}$$

where $c \in \mathbb{R}$ and successive terms of the sequence $(\varkappa_k)_{k=0}^\infty$ (called the potential in [17]) can be obtained by using two various approaches. The first one is a recursive way in that we have $\varkappa_0 = 0$, $\varkappa_1 = (\tau_0)^{-1}$ and for $k \geq 2$

$$\varkappa_k = \varkappa_1 \left(\varkappa_{k-1} - \sum_{j=0}^{k-1} \tau_{j+1} \varkappa_{k-1-j} \right).$$

The alternative approach is based on the formula (see [17])

$$\sum_{k=0}^\infty z^k \tau_k = (\tilde{\tau}(z) - 1)^{-1},$$

where for $|z| < 1$ we define

$$\tilde{\tau}(z) \stackrel{def}{=} \sum_{i=-1}^\infty z^i \tau_{i+1}.$$

Using next the Maclaurin power expansion we find

$$\varkappa_k = \frac{1}{k!} \frac{d^k}{dz^k} (\tilde{\tau}(z) - 1)^{-1} \Big|_{z=0},$$

where $k \geq 0$. Introducing the following substitution:

$$\tilde{v}_n(z) \stackrel{def}{=} \tilde{\beta}_{B-n}(z), \quad 1 \leq n \leq B, \tag{4.3}$$

we can rewrite the system of equations (3.5), (3.7) in the following form:

$$\sum_{j=-1}^{n-2} u_{j+1}(z) \tilde{v}_{n-j}(z) - \tilde{v}_n(z) = \phi_n(z), \tag{4.4}$$

where $1 \leq n \leq B - 1$,

$$\phi_n(z) \stackrel{def}{=} -\bar{\alpha}_{n-1}(z) \tilde{v}_1(z) - T_{B-n}(z), \tag{4.5}$$

$$\tilde{v}_B(z) = A(z) \tilde{v}_{B-1}(z) + \frac{z}{1-z} [1 - A(z)]. \tag{4.6}$$

Let us note that (4.4) has the same shape as (4.1) with the difference that unknowns $\tilde{v}_n(z)$ and coefficients $u_n(z)$ and $\phi_n(z)$ are now dependent on the variable z . However, referring to (4.2), we can represent $\tilde{v}_n(z)$ in the following way, having in mind that now c and \varkappa_n will also be functions of z :

$$\tilde{v}_n(z) = c(z) \cdot \varkappa_{n-1}(z) + \sum_{k=2}^n \varkappa_{n-k}(z) \phi_k(z), \quad n \geq 2. \tag{4.7}$$

Another observation is the fact that the number of equations in (4.4) is finite in comparing to (4.1). In such a case the equation (4.6) can be used to evaluate $c(z)$ in the explicit form.

Substituting $n = 1$ into (4.4) and using (4.5), we obtain

$$u_0(z)\tilde{v}_2(z) - \tilde{v}_1(z) = -\bar{\alpha}_0(z)\tilde{v}_1(z) - T_{B-1}(z).$$

Hence, due to the fact that (4.7) written for $n = 2$ has the form

$$\tilde{v}_2(z) = c(z)(u_0(z))^{-1},$$

we obtain

$$\tilde{v}_1(z) = [1 - \bar{\alpha}_0(z)]^{-1}[c(z) + T_{B-1}(z)]. \tag{4.8}$$

Now, applying (4.7) in (4.6) and referring to (4.5) and (4.8), we get

$$\begin{aligned} c(z)\varkappa_{B-1}(z) + \sum_{k=2}^B \varkappa_{B-k}(z) \left\{ -\bar{\alpha}_{k-1}(z)[1 - \bar{\alpha}_0(z)]^{-1}[c(z) + T_{B-1}(z)] \right. \\ \left. - T_{B-k}(z) \right\} = A(z) \left\{ c(z)\varkappa_{B-2}(z) + \sum_{k=2}^{B-1} \varkappa_{B-1-k}(z) \right. \\ \left. \times \left[-\bar{\alpha}_{k-1}(z)[1 - \bar{\alpha}_0(z)]^{-1}[c(z) + T_{B-1}(z)] - T_{B-k}(z) \right] \right\} + \frac{z}{1-z}[1 - A(z)]. \end{aligned}$$

As a consequence we can eliminate $c(z)$ as follows:

$$c(z) = D_1(z)D_2(z), \tag{4.9}$$

where

$$\begin{aligned} D_1(z) \stackrel{def}{=} [1 - \bar{\alpha}_0(z)]^{-1}T_{B-1}(z) \sum_{k=2}^{B-1} [\varkappa_{B-k}(z) - A(z)\varkappa_{B-1-k}(z)]\bar{\alpha}_{k-1}(z) \\ + \sum_{k=2}^{B-1} [\varkappa_{B-k}(z) - A(z)\varkappa_{B-1-k}(z)]T_{B-k}(z) + \frac{z}{1-z}[1 - A(z)] \end{aligned}$$

and

$$\begin{aligned} D_2(z) \stackrel{def}{=} \left\{ \varkappa_{B-1}(z) - A(z)\varkappa_{B-2}(z) \right. \\ \left. - [1 - \bar{\alpha}_0(z)]^{-1} \sum_{k=2}^{B-1} [\varkappa_{B-k}(z) - A(z)\varkappa_{B-1-k}(z)]\bar{\alpha}_{k-1}(z) \right\}^{-1}. \end{aligned}$$

Referring now to (4.7) and having in mind (4.5) and (4.9), we obtain

$$\begin{aligned} \tilde{v}_n(z) = D_1(z)D_2(z) \left[\varkappa_{n-1}(z) - (1 - \bar{\alpha}_0(z))^{-1} \sum_{k=2}^n \varkappa_{n-k}(z)\bar{\alpha}_{k-1}(z) \right] \\ - \sum_{k=2}^n \varkappa_{n-k}(z) \left[(1 - \bar{\alpha}_0(z))^{-1}T_{B-1}(z)\bar{\alpha}_{k-1}(z) + T_{B-k}(z) \right]. \end{aligned}$$

Taking into consideration (4.3) we can formulate the following main theorem:

Theorem 1. *The probability generating function $\tilde{\beta}_n(z)$ of the (tail) conditional distribution of the time β to the first buffer overflow in the Geo/G/1/B-type queueing model can be represented in the following way:*

$$\begin{aligned} \tilde{\beta}_n(z) = & D_1(z)D_2(z) \left[\varkappa_{B-n-1}(z) - (1 - \bar{\alpha}_0(z))^{-1} \sum_{k=2}^{B-n} \varkappa_{B-n-k}(z)\bar{\alpha}_{k-1}(z) \right] \\ & - \sum_{k=2}^{B-n} \varkappa_{B-n-k}(z) \left[(1 - \bar{\alpha}_0(z))^{-1} T_{B-1}(z)\bar{\alpha}_{k-1}(z) + T_{B-k}(z) \right], \end{aligned} \tag{4.10}$$

where $|z| < 1$, $n \in \{0, \dots, B - 2\}$, and (compare (4.8))

$$\tilde{\beta}_{B-1}(z) = [1 - \bar{\alpha}_0(z)]^{-1} [D_1(z)D_2(z) + T_{B-1}(z)]. \tag{4.11}$$

The following corollary is obvious:

Corollary 1. The mean time $\mathbf{E}_n\beta$ to the first buffer overflow conditioned by the initial buffer state $n \in \{0, \dots, B - 1\}$ can be computed as

$$\mathbf{E}_n\beta = \sum_{k=1}^{\infty} \mathbf{P}\{\beta \geq k \mid X_0 = n\} = \tilde{\beta}_n(1).$$

Let us consider now next buffer overflows. Denote by $\beta^{(r)}$, $r \geq 2$, the time to the initialization of the r th buffer overflow period, counting from the completion epoch of the $(r - 1)$ th buffer overflow. Since the number of jobs present in the system at the completion epoch of each buffer overflow period equals $B - 1$, then the following corollary is obvious:

Corollary 1. *For $r \geq 2$ and $|z| < 1$ the following representation is true:*

$$\sum_{k=1}^{\infty} z^k \mathbf{P}\{\beta^{(r)} \geq k\} = \sum_{k=1}^{\infty} z^k \mathbf{P}\{\beta \geq k \mid X_0 = B - 1\} = \tilde{\beta}_{B-1}(z).$$

In consequence, the mean time to the r th buffer overflow ($r \geq 2$) equals $\mathbf{E}\beta^{(r)} = \tilde{\beta}_{B-1}(1)$.

5 Numerical examples

In this section, we present numerical results illustrating theoretical formulae obtained above. The main goal of this section is to investigate and visualize the impact of key “input” system parameters (like the changing of the arrival rate, service speed and initial state of the accumulating buffer) on the probabilistic behaviour of the time to buffer overflow.

To find the values of probabilities $\beta_n(k)$ for fixed n and successive k 's we can use multiple differentiation of the right side of the formula (4.10) or (4.11). However, quite often we can meet some problems with symbolic or numerical differentiation. Thus, in the paper, we apply the algorithm of numerical inversion of the generating function proposed in [1], which is based on Cauchy integral formula and trapezoidal rule for numerical integration.

Namely, if $F(z) \stackrel{def}{=} \sum_{k=0}^{\infty} z^k f_k$, then we approximate the values f_k in the following way:

$$f_k \approx \frac{1}{2klr^k} \left[c_0(k, l, r) + (-1)^k c_k(k, l, r) + 2 \sum_{j=1}^{k-1} (-1)^j \operatorname{Re}(c_j(k, l, r)) \right],$$

where $c_j(k, l, r) \stackrel{def}{=} \sum_{n=0}^{l-1} e^{-\pi i n/l} F(re^{\pi i(n+l j)/lk})$ and typical values of parameters are $l = 1$ and $r = (0.0001)^{1/k}$.

Note that the numerical inversion of continuous and discrete transforms may result in poor accuracy for small values of the argument.

Let us consider a single-machine production system modeled via the *Geo/G/1/B*-type queueing system with finite buffer capacity and assume that $B = 6$. We study three different cases according to the service time: geometrically distributed service time, deterministic (constant) one and the bounded discrete distribution of the service time. To compare the impact of the shape of the processing time distribution on the probabilistic behaviour of the time to buffer overflow, we take into consideration the appropriate distributions with the same means.

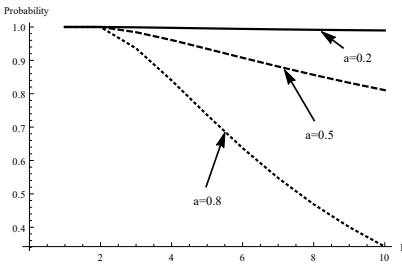


Figure 1. $\mathbf{P}\{\beta \geq k \mid X_0 = 3\}$ for geometrically distributed service times.

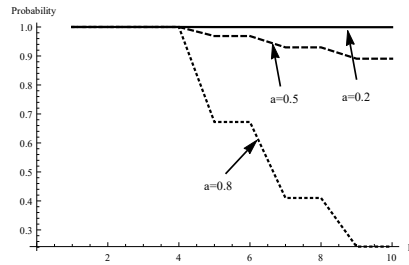


Figure 2. $\mathbf{P}\{\beta \geq k \mid X_0 = 3\}$ for deterministic service times.

Let us begin with investigation of the impact of arrival intensity. Assume firstly that the service time is geometrically distributed with parameter $b = 0.5$, so $b_k = 1/2^k$, where $k \geq 1$ (mean service time equals 2). In Figure 1 we visualize the probabilities $\mathbf{P}\{\beta \geq k \mid X_0 = 3\}$ for $k = 1, \dots, 10$, and three different values of the parameter a of geometric distribution of interarrival times: $a = 0.2, 0.5$ and 0.8 , corresponding to three different offered loads of the considered system: $\rho = 0.4, 1.0$ and 1.6 , respectively. For example, the probabilities $\mathbf{P}\{\beta \geq 5 \mid X_0 = 3\}$ take on the following values: $0.9958, 0.9345$ and 0.7363 for $\rho = 0.4, 1.0$ and 1.6 , respectively. So, for the case of $\rho = 1.6$ (overloaded system) in comparison to the case $\rho = 1.0$ the decrease of the value of probability is essential (about 21%).

For the case of constant (deterministic) service times equal to 2 the appropriate results are presented in Figure 2. The conclusion is similar. The “cyclic” behaviour of the curves is connected with constant service time equal to 2.

Finally, in Figure 3 we visualize results for the case of the service time with

the same mean 2, distributed as follows:

$$b_1 = 0.25, b_2 = 0.50, b_3 = 0.25, b_k = 0, k \geq 4.$$

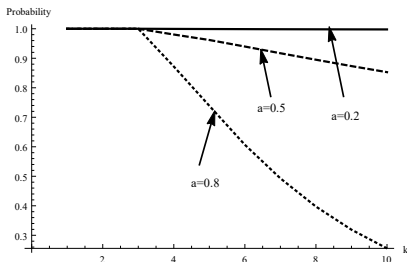


Figure 3. $P\{\beta \geq k | X_0 = 3\}$ for service times with 3 possible values.

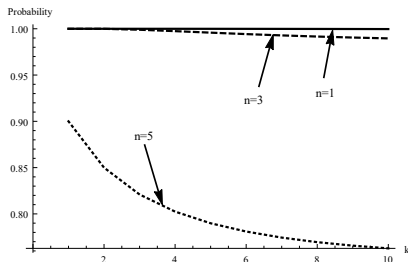


Figure 4. $P\{\beta \geq k | X_0 = n\}$ for $n = 1, 3$ and 5 , and $\rho = 0.4$.

Now let us analyze the impact of the initial buffer state on the distribution of the time to the first buffer overflow. In Figures 4–6 the probabilities $P\{\beta \geq k | X_0 = n\}$ for three initial levels of the buffer state, namely $n = 1, 3$ and 5 , and different values of the offered load ρ are visualized.

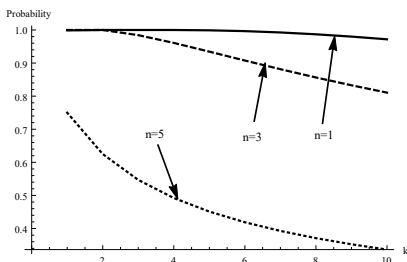


Figure 5. $P\{\beta \geq k | X_0 = n\}$ for $n = 1, 3$ and 5 , and $\rho = 1.0$.

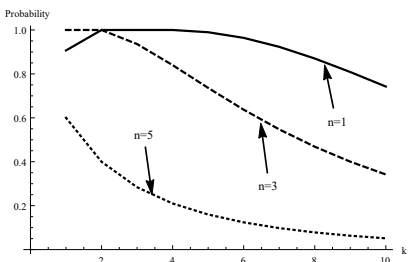


Figure 6. $P\{\beta \geq k | X_0 = n\}$ for $n = 1, 3$ and 5 , and $\rho = 1.6$.

The results are obtained for geometric service distribution with parameter $b = 0.5$ and three different values of the parameter of geometric distribution of interarrival times, namely $a = 0.2, 0.5$ and 0.8 (corresponding to Figure 4, Figure 5 and Figure 6, respectively). Evidently, as one can observe, the smaller the number of jobs accumulated in the buffer before the opening of the system, the greater the probability that the time to the first buffer overflow will exceed the concrete value k . Let us note that if the buffer is almost saturated ($n = 5$), the probability is essentially smaller than in two remaining cases ($n = 1$ and 3).

Lastly, in Table 1 the results for the mean time to the first buffer overflow are presented (for the same geometric distributions of interarrival and service times as in Figures 4–6). Obviously, corresponding values decrease as the number of jobs accumulated in the buffer initially increases, however the difference between values obtained for different ρ 's is very large.

Table 1. Mean times to the first buffer overflows in a function of n and ρ for geometric service distribution.

Initial buffer state n	$\rho = 0.4$	$\rho = 1.0$	$\rho = 1.6$
1	3766.0198	68.7803	14.9719
3	3709.2296	52.8547	8.9569
5	2828.4886	20.9511	2.3312

6 Conclusions

In the paper a model of a single-server production system with finite buffer capacity is investigated. Using the approach based on queueing theory, the probability distribution of the time to the first buffer overflow is studied. By using the method utilizing the idea of embedded Markov chain and the notion of the potential proposed by Korolyuk, the closed-form representation for the probability generating function of this distribution is derived. Moreover, corresponding results for next buffer overflow periods and for their means are obtained. Numerical analysis of sensitivity of the distribution of the time to the buffer overflow on key operating system parameters is attached.

References

- [1] J. Abate, G.L. Choudhury and W. Whitt. An introduction to numerical transform inversion and its application to probability models. In W. Grassmann(Ed.), *Computational Probability*, pp. 257–323, Boston, 2000. Kluwer Academic Publishers.
- [2] S. Asmussen, M. Jobmann and H.-P. Schwefel. Exact buffer overflow calculations for queues via martingales. *Queueing Syst.*, **42**(1):63–90, 2002. <https://doi.org/10.1023/A:1019994728099>.
- [3] H. Bruneel and B.G. Kim. *Discrete-time models for communication systems including ATM*. Kluwer Academic Publishers, Boston, 1993. <https://doi.org/10.1007/978-1-4615-3130-2>.
- [4] M.L. Chaudhry and Y.Q. Zhao. First-passage-time and busy-period distributions of discrete-time Markovian queues: Geom(n)/geom(n)/1/n. *Queueing Syst.*, **18**:5–26, 1994. <https://doi.org/10.1007/BF01158772>.
- [5] A. Chydziański. Time of buffer overflow in a MMPP queue. In *Proceedings of the International Conference on Research Networking: NETWORKING 2007*, volume 4479 of *Lecture Notes in Computer Science*, pp. 879–889, 2007.
- [6] A. Chydziański. Time to reach buffer capacity in a BMAP queue. *Stoch. Models*, **23**:95–209, 2007. <https://doi.org/10.1080/15326340701300746>.
- [7] A. Duda. Transient diffusion approximation for some queueing systems. In *Proceedings of the ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems*. ACM, 1983. <https://doi.org/10.1145/800040.801399>.
- [8] A. Al Hanbali. Busy period analysis of the level dependent PH/PH/1/K queue. *Queueing Syst.*, **67**:221–249, 2011. <https://doi.org/10.1007/s11134-011-9213-6>.
- [9] J.J. Hunter. *Mathematical techniques of applied probability, vol. II, Discrete-time models: techniques and applications*. Academic Press, New York, 1983.

- [10] W.M. Kempa. A comprehensive study on the queue-size distribution in a finite-buffer system with a general independent input flow. *Perform. Evaluation*, **108**:1–15, 2017. <https://doi.org/10.1016/j.peva.2016.11.002>.
- [11] W.M. Kempa. Time to start a crowded period in a finite-buffer queue with Poisson input flow and general processing times. In I. Dimov, I. Farago and L. Vulkov(Eds.), *Finite difference methods theory and applications. 7th International conference, FDM 2018, Lozenetz, Bulgaria, June 11-16, 2018*, volume 11386 of *Lecture Notes in Computer Science*, pp. 329–336, 2018.
- [12] W.M. Kempa and M. Kobielnik. Time to buffer overflow in a queueing model with working vacation policy. In P. Gaj, M. Sawicki, G. Suchacka and A. Kwiecień(Eds.), *Computer networks. 25th International conference (CN 2018), Gliwice, Poland, June 19-22, 2018*, volume 860 of *Communications in Computer and Information Science*, pp. 219–231, Cham, 2018. Springer.
- [13] W.M. Kempa and R. Marjasz. Distribution of the time to buffer overflow in the M/G/1/N-type queueing model with batch arrivals and multiple vacation policy. *J. Oper. Res. Soc.*, 2019. <https://doi.org/10.1080/01605682.2019.1567651>.
- [14] W.M. Kempa and I. Paprocka. Time to buffer overflow in a finite-capacity queueing model with setup and closedown times. In J. Świątek, Z. Wilimowska, L. Borzowski and A. Grzech(Eds.), *Information systems architecture and technology. Proceedings of 37th International Conference on Information Systems Architecture and Technology (ISAT 2016)*, volume 523 of *Advances in Intelligent Systems and Computing*, pp. 215–224, Cham, 2017. Springer.
- [15] W.M. Kempa, I. Paprocka, C. Grabowik and K. Kalinowski. Distribution of time to buffer overflow in a finite-buffer manufacturing model with unreliable machine. In L. Slatineanu et al.(Ed.), *21st Innovative Manufacturing Engineering & Energy International Conference (IManE&E 2017), Iasi, Romania, May 24-27, 2017*, volume 112 of *MATEC Web of Conferences*, pp. 1–6, Les Ulis, 2017. EDP Sciences. <https://doi.org/10.1051/mateconf/201711205005>.
- [16] T. Kimura, K. Ohno and H. Mine. Diffusion approximation for GI/G.G/1 queuing system with finite capacity: I - the first overflow time. *J. Oper. Res. Soc. Jpn.*, **22**:41–68, 1979. <https://doi.org/10.15807/jorsj.22.41>.
- [17] V.S. Korolyuk. *Boundary-Value Problems for Compound Poisson Processes*. Naukowa Dumka, Kiev, 1975.
- [18] E.Y. Lee and K.K.J. Kinader. The expected wet period of finite dam with exponential inputs. *Stoch. Proc. Appl.*, **90**:175–180, 2000. [https://doi.org/10.1016/S0304-4149\(00\)00034-X](https://doi.org/10.1016/S0304-4149(00)00034-X).
- [19] F. Machihara. First passage times of PH/PH/1/K and PH/PH/1 queues. *J. Oper. Res. Soc. Jpn.*, **30**, 1987. <https://doi.org/10.15807/jorsj.30.1>.
- [20] S.M. Ross and S. Seshadri. Hitting time in an M/G/1 queue. *J. Appl. Probab.*, **36**:934–940, 1999. <https://doi.org/10.1017/S0021900200017691>.
- [21] H. Takagi. *Queueing analysis - a foundation of performance evaluation, Vol. 3. Discrete-Time Systems*. North-Holland, Amsterdam, 1993.