

EURANDOM PREPRINT SERIES

2015-023

November 19, 2015

A queueing system with vacations after N services

O. Boxma, D. Claeys, L. Gulikers, O. Kella
ISSN 1389-2355

A queueing system with vacations after N services

Onno Boxma^{*}, Dieter Claeys[†], Lennart Gulikers[‡] and Offer Kella[§]

November 19, 2015

Abstract

This paper is devoted to the study of an M/G/1 queue with a particular vacation discipline. The server is due to take a vacation as soon as it has served exactly N customers since the end of the previous vacation. N may be either a constant or a random variable. If the system becomes empty before the server has served N customers, then it stays idle until the next customer arrival. Such a vacation discipline arises, e.g., in production systems and in order picking in warehouses.

We determine the joint transform of the length of a visit period and the number of customers in the system at the end of that period. We also derive the generating function of the number of customers at a random instant, and the Laplace-Stieltjes transform of the delay of a customer.

Note: accepted for publication in *Naval Research Logistics*.

1 Introduction

Queueing systems with vacations [6, 10, 11] have been studied extensively due to their wide area of applications including repair/maintenance models [13], ARQ systems [12], and resource-sharing systems such as priority [10] and polling systems [2]. The literature can roughly be divided into two categories depending on whether vacations are triggered by the system itself or are due to external factors. The category in which vacations are caused by external circumstances (e.g., server failures) can further be partitioned based on whether the vacation is nonpreemptive (i.e., the ongoing service is finished before the incoming vacation is initiated) or preemptive. The second subcase is often referred to as server interruptions or server breakdowns [7]. The category in which vacations are triggered by the system itself can be subdivided into exhaustive, gated, number-limited and time-limited vacations. In the exhaustive case, the server initiates a vacation when the system becomes

^{*}EURANDOM and Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands (o.j.boxma@tue.nl); research done in the framework of the IAP BESTCOM project, funded by the Belgian government

[†]Department of Telecommunications and Information Processing (TELIN), SMACS Research Group and Department of Industrial Systems Engineering and Product Design, Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium (dieter.claeys@telin.UGent.be); Postdoctoral Fellow with the Fund for Scientific Research, Flanders (FWO Vlaanderen), Belgium; research partially done in the framework of the Interuniversity Attraction Poles program initiated by the Belgian Science Policy office

[‡]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands (L.Gulikers@student.tue.nl); research done in the framework of an honours project of the department

[§]Department of Statistics, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel (offer.kella@huji.ac.il); supported in part by grant 1462/13 from the Israel Science Foundation and the Vigevani Chair in Statistics

empty, whereas in the gated case, it starts a new vacation when all customers that were present at the end of the previous vacation have been served. In number-limited systems, a vacation is initiated when the server has served a predetermined number of customers *or* when the system becomes empty. Similarly, in time-limited systems, a vacation starts when the server has served a predetermined amount of time or when it becomes empty. Another distinctive feature of vacation models is single versus multiple vacations, describing the action of the server upon finding an empty system at the end of a vacation. In the former scenario, the server remains idle and immediately resumes service when the next customer arrives, whereas another vacation is initiated in the latter case.

This paper is devoted to a continuous-time queueing system with server vacations of the second class, i.e., vacations are triggered by the system. We study a system where the server goes on vacation when *exactly* a predetermined number of customers, N , have been served; when the system becomes empty before having served N customers since the previous vacation, no vacation is initiated, yet the server is idle and starts serving the next incoming customer immediately. Hence, the system alternates between single vacations and visit periods during which exactly N customers receive service. As a consequence, our vacation mechanism can be considered as modified number-limited vacations (henceforth referred to as modified N -limited vacations to stress the number N).

The proposed queue with modified number-limited vacation policy can model various systems, such as production systems and order-picking systems. Let us look a bit closer at these two examples. Firstly, in production systems, one might program a machine to produce N products of a certain type, before it switches to another task. If not all material of this product type is available, the machine may be programmed to wait when the switching costs are high. The repetitive cyclic pattern of producing exactly N products of a particular type, followed by switching and doing other tasks, boils down to the above-described modified N -limited vacation model.

Secondly, in order-picking systems [3], an order picker at a remotely located order-picking workstation satisfies customer orders, each consisting of a certain number, N , of requested items. Totes containing a single Stock Keeping Unit (SKU) are automatically transported from the storage area to the workstation, where the order picker picks the requested items from these totes and puts them in the so-called order tote. When, occasionally, no more items for the current active order are available, the order picker waits until totes with these items arrive; he does not start handling another order to avoid picking errors. Once all items for an order are collected, the order picker moves on to another order. However, before starting to pick items requested by this new order, the order picker puts the finished order tote on another conveyor, takes another order tote, scans bar codes, etc. The repetitive cyclic pattern of collecting exactly N requested items followed by some finishing and preparatory work can be adequately modelled by the N -limited vacation policy.

In this paper, we determine the joint transform of the number of customers in the system at the beginning of a vacation and the length of the preceding visit period. From this joint transform, we obtain the marginal transforms, being respectively a probability generating function (PGF) and a Laplace-Stieltjes transform (LST). Moreover, from the PGF of the number of customers present at the beginning of a vacation, we extract the PGF of the system content at random time instants and the LST of the customer delay, by invoking Fuhrmann-Cooper's decomposition result [8].

In our analysis we consider generally distributed service times and vacation lengths and study several distributions for N . We focus in some detail on the geometric distribution, a finite mixture of geometric distributions and constant N . We determine explicit closed-form formulae in case of geometrically distributed N . When N is a finite mixture of geometrically distributed random

variables or when it is a constant, we obtain explicit formulae as well, albeit they contain a finite number of a priori unknown constants. We do, however, derive a set of linearly independent equations from which these constants can be computed numerically. Finally, we outline a method to treat the case of N having a PGF which is a rational function, i.e., a quotient of polynomials. Unlike the class of mixtures of geometric distributions (e.g., see [5]), the class of such distributions is dense in the class of all distributions of non-negative integer-valued random variables.

The paper is organized as follows. In Section 2 we provide a detailed model description. Section 3 covers the analysis of the joint and marginal transforms of the length of a visit period and of the number of customers in the system at the end of that visit period. System content at random time instants and customer delay are treated in Sections 4 and 5 respectively. Some numerical results are presented in Section 6 to elucidate the behaviour of the system. Section 7 contains conclusions and some suggestions for future research.

2 Model description: An M/G/1 queue with modified N -limited vacation policy

We study an M/G/1 queueing system with a vacation mechanism that, to the best of our knowledge, has not yet been studied analytically. The server is due to take a vacation as soon as it has served *exactly* N customers since the end of the previous vacation, with N a random variable. If the server observes an empty system before having served N customers, it does not take a vacation, yet it becomes idle, i.e., it waits for the arrival of the next customer, whose service is then immediately started. In addition, when the server finds an empty system upon returning from vacation, it becomes idle instead of commencing a new vacation. Hence, the system alternates between single vacations and visit periods during which exactly N customers receive service. The time duration of the n -th vacation is denoted by V_n and the consecutive vacation lengths are modelled by a sequence of independent and identically distributed (i.i.d.) random variables. Their common Laplace-Stieltjes Transform (LST) is denoted by ϕ .

The arrival rate is represented by λ , and the service time of the n -th customer by B_n . The consecutive service times are i.i.d. with common LST β . The service times are independent of the vacation lengths and the arrival process.

Related to ϕ and β , we define ϕ^* and β^* as the probability generating functions (PGFs) of the number of arrivals during a random vacation and the number of arrivals during a random service respectively:

$$\phi^*(p) := \phi(\lambda(1-p)) \quad , \quad \beta^*(p) := \beta(\lambda(1-p)) \quad , \quad |p| \leq 1.$$

The stability condition for this system reads

$$\lambda \left(\mathbb{E}[B] + \frac{\mathbb{E}[V]}{\mathbb{E}[N]} \right) < 1 \quad , \quad (1)$$

with B an arbitrary service time and V an arbitrary vacation duration. This can be understood by viewing the system as an M/G/1 queue without vacations and with possibly extended service times: a fraction $\frac{1}{\mathbb{E}[N]}$ of the items receives an extra service time of mean length $\mathbb{E}[V]$. We assume (1) to be fulfilled henceforth.

Before proceeding to the analysis of the system, we introduce short-hand notations for the open complex unit disk, the closed complex unit disk and its boundary:

$$B(0,1) := \{p \in \mathbb{C} : |p| < 1\} , \quad \overline{B(0,1)} := \{p \in \mathbb{C} : |p| \leq 1\} , \quad \partial B(0,1) := \{p \in \mathbb{C} : |p| = 1\} .$$

3 Analysis at embedded epochs

In this section, we examine the system at vacation initiation (embedded) epochs. We fix $n \in \{1, 2, \dots\}$, $\alpha \in \mathbb{C}$ with $\Re(\alpha) \geq 0$, $r \in B(0,1)$ and $p \in \overline{B(0,1)}$, and define C_n (E_n) as the number of customers at the beginning (end) of the n -th vacation and L_n as the duration of the n -th visit period, i.e., the period between the $(n-1)$ -th and n -th vacation. The notation $C := \lim_{n \rightarrow \infty} C_n$, $E := \lim_{n \rightarrow \infty} E_n$, and $L := \lim_{n \rightarrow \infty} L_n$ is used for their steady-state counterparts.

Note that in any case

$$\mathbb{E} [p^{E_n}] = \mathbb{E} [p^{C_n}] \phi^*(p) , \quad (2)$$

because the number of customers that arrive during a vacation is independent of the number of customers present at the start of that vacation. To calculate $\mathbb{E} [e^{-\alpha L_n} p^{C_n}]$ we condition on $E_{n-1} = i \in \mathbb{N}$ so that we end up in exactly the situation studied by Cohen in [4], Section II.4.3, albeit in a different context. There, the transient behaviour of the queue length at customer departure epochs in an ordinary M/G/1 queue is considered. To inherit Cohen's notation we let, for $k \in \mathbb{N}$, Z_k be the number of jobs left behind in the ordinary queue at the k -th departure epoch and r'_k the time between the 0-th and k -th departure epoch. The 0-th departure epoch is postulated to coincide with the origin of the time axis. Formula (4.23) on p.240 in [4] states that, for $|p| \leq 1$, $\Re(\alpha) \geq 0$, $|r| < 1$,

$$\begin{aligned} \pi_i(p, \alpha, r) &:= \sum_{k=1}^{\infty} r^k \mathbb{E} [p^{Z_k} e^{-\alpha r'_k} | Z_0 = i] \\ &= \frac{r\beta(\alpha + \lambda(1-p))}{p - r\beta(\alpha + \lambda(1-p))} \left\{ p^i - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1 - \mu(\alpha, r))} \mu^i(\alpha, r) \right\} , \end{aligned} \quad (3)$$

where $\mu(\alpha, r) := \mathbb{E} [e^{-\alpha P} r^K]$ is the transform of the steady-state joint distribution of busy period length P and number of customers K served during that busy period in the ordinary M/G/1 queue. It is the unique zero of the function $p \mapsto p - r\beta(\alpha + \lambda(1-p))$ in $B(0,1)$.

In the sequel we derive results for different distributions of N .

3.1 N geometrically distributed

In this subsection we assume that the probability distribution of N is geometric:

$$\mathbb{P}(N = k) = \begin{cases} (1-r)r^{k-1} & \text{if } k \in \{1, 2, \dots\} \\ 0 & \text{otherwise} \end{cases} ,$$

where $r \in (0, 1)$. Now, using (3),

$$\begin{aligned}
\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i] &= \sum_{k=1}^{\infty} \mathbb{P}(N = k | E_{n-1} = i) \mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i, N = k] \\
&= \frac{1-r}{r} \sum_{k=1}^{\infty} r^k \mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i, N = k] \\
&= \frac{1-r}{r} \pi_i(p, \alpha, r) \\
&= \gamma^{(\alpha)}(p) \left\{ p^i - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, r))} \mu^i(\alpha, r) \right\},
\end{aligned}$$

with

$$\gamma^{(\alpha)}(p) := \frac{(1-r)\beta(\alpha + \lambda(1-p))}{p - r\beta(\alpha + \lambda(1-p))}.$$

Summation over all possible values for E_{n-1} yields

$$\begin{aligned}
\mathbb{E} [e^{-\alpha L_n} p^{C_n}] &= \sum_{i=0}^{\infty} \mathbb{P}(E_{n-1} = i) \mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i] \\
&= \gamma^{(\alpha)}(p) \left\{ \mathbb{E} [p^{E_{n-1}}] - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, r))} \mathbb{E} [\mu^{E_{n-1}}(\alpha, r)] \right\}. \quad (4)
\end{aligned}$$

Using Relation (2) between E_{n-1} and C_{n-1} , we obtain

$$\mathbb{E} [e^{-\alpha L_n} p^{C_n}] = \gamma^{(\alpha)}(p) \left\{ \mathbb{E} [p^{C_{n-1}}] \phi^*(p) - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, r))} \mathbb{E} [\mu^{C_{n-1}}(\alpha, r)] \phi^*(\mu(\alpha, r)) \right\}. \quad (5)$$

It should be noticed that the factor in between curly brackets becomes zero for $p = \mu(\alpha, r)$. On the other hand, the function $p \mapsto p - r\beta(\alpha + \lambda(1-p))$, the denominator in $\gamma^{(\alpha)}(p)$, has exactly one zero in $B(0, 1)$ and that zero is $\mu(\alpha, r)$.

The steady-state behaviour of the system follows immediately from (5):

$$\begin{aligned}
\mathbb{E} [e^{-\alpha L} p^C] &:= \lim_{n \rightarrow \infty} \mathbb{E} [e^{-\alpha L_n} p^{C_n}] \\
&= \gamma^{(\alpha)}(p) \left\{ \mathbb{E} [p^C] \phi^*(p) - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, r))} \mathbb{E} [\mu^C(\alpha, r)] \phi^*(\mu(\alpha, r)) \right\}. \quad (6)
\end{aligned}$$

In particular (let $\alpha \rightarrow 0$),

$$\mathbb{E} [p^C] = \frac{(1-r)\beta^*(p)}{p - r\beta^*(p)} \left\{ \mathbb{E} [p^C] \phi^*(p) - \frac{1-p}{1-\mu(0, r)} \mathbb{E} [\mu^C(0, r)] \phi^*(\mu(0, r)) \right\},$$

or, equivalently, bringing the $\mathbb{E} [p^C]$ term in the right-hand side to the left-hand side:

$$\mathbb{E} [p^C] = \frac{\frac{1-r}{1-\mu(0, r)}(1-p)\beta^*(p)\phi^*(\mu(0, r))}{(1-r)\beta^*(p)\phi^*(p) - p + r\beta^*(p)} \mathbb{E} [\mu^C(0, r)]. \quad (7)$$

To determine $\mathbb{E} [\mu^C(0, r)]$, we let $p \rightarrow 1$, invoking the normalisation condition of PGFs and l'Hospital's rule, leading to

$$\mathbb{E} [\mu^C(0, r)] = \frac{1-\mu(0, r)}{\phi^*(\mu(0, r))} \mathbb{E} [N] \left\{ 1 - \lambda \left(\mathbb{E} [B] + \frac{\mathbb{E} [V]}{\mathbb{E} [N]} \right) \right\}, \quad (8)$$

where we have used that $\mathbb{E}[N] = \frac{1}{1-r}$. Substitution of (8) into (7) gives

$$\mathbb{E}[p^C] = \frac{\left\{1 - \lambda \left(\mathbb{E}[B] + \frac{\mathbb{E}[V]}{\mathbb{E}[N]} \right)\right\} (1-p)\beta^*(p)}{\beta^*(p) \left[1 - \frac{1}{\mathbb{E}[N]} + \frac{1}{\mathbb{E}[N]}\phi^*(p)\right] - p} . \quad (9)$$

Plugging $p = \mu(\alpha, r)$ into (9) yields $\mathbb{E}[\mu(\alpha, r)^C]$, and then $\mathbb{E}[e^{-\alpha L} p^C]$, given by (6), is in fact completely specified. From (6), we can also extract the LST of the length of a random visit period of the server in steady state:

$$\mathbb{E}[e^{-\alpha L}] = \frac{(1-r)\beta(\alpha)}{1-r\beta(\alpha)} \left\{1 - \frac{\alpha}{\alpha + \lambda(1-\mu(\alpha, r))} \mathbb{E}[\mu^C(\alpha, r)] \phi^*(\mu(\alpha, r))\right\} . \quad (10)$$

It easily follows that

$$\mathbb{E}[L] = \frac{\mathbb{E}[B]}{1-r} + \frac{1}{\lambda(1-\mu(0, r))} \mathbb{E}[\mu^C(0, r)] \phi^*(\mu(0, r)) . \quad (11)$$

Remark.

In this special case of geometrically distributed N , we could have obtained Expression (9) for $\mathbb{E}[p^C]$ also by invoking the classical result for the PGF of the distribution of the queue length in the M/G/1 queue, say, Z , immediately after a service completion. In an ordinary M/G/1 queue, it holds that $\mathbb{E}[p^Z] = \frac{\{1-\lambda\mathbb{E}[B]\}(1-p)\beta^*(p)}{\beta^*(p)-p}$. However, in the system under consideration, an ordinary service B is with probability $1-r$ extended by a vacation V , independently from customer to customer. Adapting the service time accordingly in the above expression for $\mathbb{E}[p^Z]$ yields:

$$\mathbb{E}[p^Z] = \frac{\left\{1 - \lambda \left(\mathbb{E}[B] + \frac{\mathbb{E}[V]}{\mathbb{E}[N]} \right)\right\} (1-p)\beta^*(p) \left[1 - \frac{1}{\mathbb{E}[N]} + \frac{1}{\mathbb{E}[N]}\phi^*(p)\right]}{\beta^*(p) \left[1 - \frac{1}{\mathbb{E}[N]} + \frac{1}{\mathbb{E}[N]}\phi^*(p)\right] - p} . \quad (12)$$

Owing to the memoryless property of the geometric distribution, the number of customers left behind at a non-extended service completion epoch is, in distribution, equal to C and the number of customers left behind at an extended service completion epoch is, in distribution, equal to $C+A_V$, with A_V the number of arrivals during the vacation. As a result, and since a fraction of $1/\mathbb{E}[N]$ services is extended, we obtain

$$\mathbb{E}[p^Z] = \mathbb{E}[p^C] \left[1 - \frac{1}{\mathbb{E}[N]} + \frac{1}{\mathbb{E}[N]}\phi^*(p)\right] .$$

Taking this into account, Formula (9) follows from (12).

3.2 N mixture of geometric distributions

Fix $m \in \mathbb{N}^+$, $r_1, \dots, r_m \in (0, 1)$ and $p_1, \dots, p_m \in (0, 1)$ in such a way that $p_1 + \dots + p_m = 1$. We let N have a probability distribution pointwise defined by

$$\mathbb{P}(N = k) = \begin{cases} \sum_{l=1}^m p_l (1-r_l) r_l^{k-1} & \text{if } k \in \{1, 2, \dots\}, \\ 0 & \text{otherwise.} \end{cases}$$

From a calculation that is almost identical to the one above, it emerges that (similar to (6)),

$$\mathbb{E}[e^{-\alpha L} p^C] = \sum_{l=1}^m p_l \gamma_l^{(\alpha)}(p) \left\{ \mathbb{E}[p^C] \phi^*(p) - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, r_l))} \mathbb{E}[\mu^C(\alpha, r_l)] \phi^*(\mu(\alpha, r_l)) \right\} , \quad (13)$$

where

$$\gamma_l^{(\alpha)}(p) := \frac{(1-r_l)\beta(\alpha + \lambda(1-p))}{p - r_l\beta(\alpha + \lambda(1-p))} .$$

We let $\alpha \rightarrow 0$ in (13) to obtain

$$\frac{1}{1-p} \left\{ \phi^*(p) \sum_{l=1}^m p_l \gamma_l^{(0)}(p) - 1 \right\} \mathbb{E}[p^C] = \sum_{l=1}^m p_l \gamma_l^{(0)}(p) \frac{\mathbb{E}[\mu^C(0, r_l)]}{1 - \mu(0, r_l)} \phi^*(\mu(0, r_l)) . \quad (14)$$

To determine $\mathbb{E}[\mu^C(0, r_1)], \dots, \mathbb{E}[\mu^C(0, r_m)]$, we examine the term in between curly brackets on the left-hand side of (14), denoted by $g(p)$ from now on, i.e.,

$$g(p) := \beta^*(p) \phi^*(p) \sum_{l=1}^m \frac{p_l(1-r_l)}{f_l(p)} - 1 ,$$

with

$$f_l(p) := p - r_l \beta^*(p) .$$

In Theorem 1 below we shall prove that $g(p)$ has exactly m zeroes in $\overline{B(0, 1)}$, one zero equaling 1 and the other zeroes lying within $[0, 1) \setminus \{\mu(0, r_1), \dots, \mu(0, r_m)\}$. See Figure 1 for an example in case of $m = 2$.

The following two lemmas provide information about $g(p)$, and will lead to the theorem.

Lemma 1. (i) $g(p)$ has exactly m poles in $B(0, 1)$, viz. $\mu(0, r_1), \dots, \mu(0, r_m)$.
(ii) $g(p)$ has at least $m - 1$ zeroes in $[0, 1) \setminus \{\mu(0, r_1), \dots, \mu(0, r_m)\}$.

Proof. As $\beta^*(p)$ and $\phi^*(p)$ are PGFs, they have no poles in $B(0, 1)$. Hence, poles from $g(p)$ in $B(0, 1)$ stem from zeroes of $f_l(p)$ that are no zeroes of $p_l(1-r_l)\beta^*(p)\phi^*(p)$. In addition, as $f_l(p)$ is the kernel equation of the PGF of the number of customers served during a busy period in an M/G/1 queue, $\mu(0, r_l)$ is the unique zero of $f_l(p)$ in $B(0, 1)$. Moreover, as $r_l \in (0, 1)$, $\mu(0, r_l)$ is a real number. Next, for $p \in [0, 1)$, it holds that

$$p_l(1-r_l)\beta^*(p)\phi^*(p) > 0 ,$$

and, on account of the stability condition (1),

$$f_l'(p) = 1 - r_l \lambda \mathbb{E} \left[B e^{-\lambda(1-p)B} \right] \geq 1 - r_l \lambda \mathbb{E}[B] > 1 - r_l > 0 .$$

As a result,

$$\begin{aligned} \lim_{p \uparrow \mu(0, r_l)} g(p) &= -\infty , \\ \lim_{p \downarrow \mu(0, r_l)} g(p) &= +\infty . \end{aligned}$$

As $g(p)$ is continuous in $[0, 1) \setminus \{\mu(0, r_1), \dots, \mu(0, r_m)\}$ (sums and/or products of PGFs), $g(p)$ has at least one zero between two consecutive poles (Bolzano's theorem), concluding the proof. \square

Lemma 2. $\tilde{g}(p) := -g(p) \prod_{l=1}^m (p - r_l \beta^*(p))$ has exactly $m - 1$ zeroes in $B(0, 1)$.

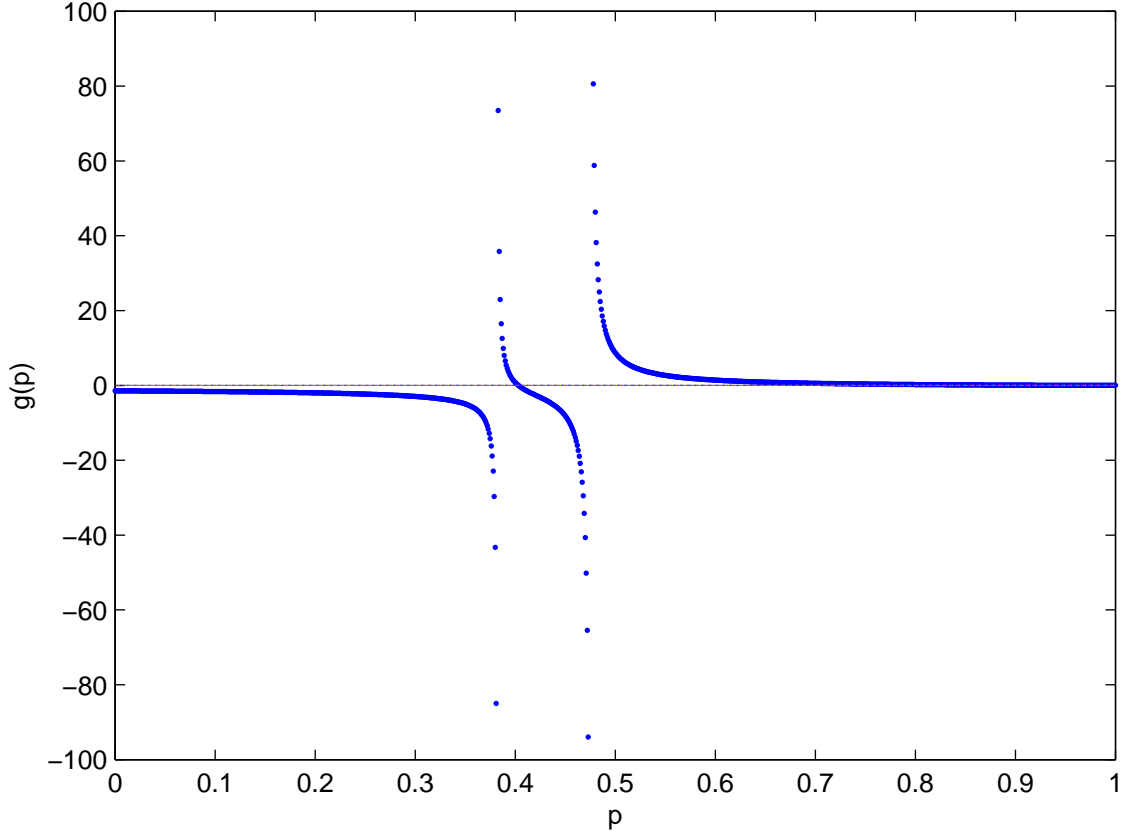


Figure 1: The function g has exactly 2 zeros $q_1 \simeq 0.4$ and $q_2 = 1$ on $[0, 1]$ in case $m = 2$.

Proof. We shall prove the lemma by using Klimenok's theorem [9]. We start with restructuring $\tilde{g}(p)$ as follows:

$$\tilde{g}(p) = f(p) + \psi(p) ,$$

with

$$f(p) := \prod_{l=1}^m (p - r_l \beta^*(p)) ,$$

$$\psi(p) := -\beta^*(p) \phi^*(p) \sum_{l=1}^m p_l (1 - r_l) \prod_{j=1, j \neq l}^m (p - r_j \beta^*(p)) .$$

Note that $f(p)$ has exactly m zeroes in $B(0, 1)$. First, it holds that $f(p)$ and $\psi(p)$ are analytic in $B(0, 1)$ and continuous on $\partial B(0, 1)$. Secondly, $f(1) = -\psi(1) \neq 0$. As a third step, we will prove that $|f(p)| > |\psi(p)|$ on $\partial B(0, 1) \setminus \{p = 1\}$. Since $\beta^*(p)$ and $\phi^*(p)$ are PGFs, it holds that $|\beta^*(p)| < 1$ and $|\phi^*(p)| < 1$ on $\partial B(0, 1) \setminus \{p = 1\}$ (note that equality cannot occur exactly because $p = 1$ is excluded). Furthermore,

$$|p - r_j \beta^*(p)| \geq ||p| - r_j \beta^*(p)|| = |1 - r_j| \beta^*(p)| > 1 - r_j ,$$

on $\partial B(0, 1) \setminus \{p = 1\}$. Hence,

$$\begin{aligned} \left| \frac{\psi(p)}{f(p)} \right| &\leq |\beta^*(p)\phi^*(p)| \sum_{l=1}^m p_l(1-r_l) \left| \frac{1}{p-r_l\beta^*(p)} \right| \\ &< \sum_{l=1}^m p_l(1-r_l) \frac{1}{|p-r_l\beta^*(p)|} \\ &< \sum_{l=1}^m p_l = 1 \quad , \end{aligned}$$

on $\partial B(0, 1) \setminus \{p = 1\}$. As a fourth and final step, we verify that $(f'(1) + \psi'(1))/f(1) > 0$. As $f(1) = \prod_{l=1}^m (1-r_l) > 0$, it boils down to verifying that $f'(1) + \psi'(1) > 0$. We find subsequently

$$\begin{aligned} f'(1) + \psi'(1) &= \frac{d}{dp} [f(p) + g(p)] \Big|_{p=1} \\ &= - \frac{d}{dp} g(p) \prod_{l=1}^m (p-r_l\beta^*(p)) \Big|_{p=1} \\ &= -g'(1) \prod_{l=1}^m (1-r_l) \quad . \end{aligned}$$

Hence, $f'(1) + \psi'(1) > 0 \Leftrightarrow g'(1) < 0$. Using that $\mathbb{E}[N] = \sum_{l=1}^m p_l/(1-r_l)$ and invoking the stability condition yields

$$g'(1) = \mathbb{E}[N] \left\{ \mathbb{E}[B] + \frac{\mathbb{E}[V]}{\mathbb{E}[N]} \lambda \right\} - \mathbb{E}[N] < 0 \quad .$$

Summarizing, we have verified all conditions of Klimenok's theorem [9]. As a consequence, the number of zeroes of $\tilde{g}(p)$ in $B(0, 1)$ equals the number of zeroes of $f(p)$ in $B(0, 1)$, which is $m-1$, concluding the proof. \square

Theorem 1. $g(\cdot)$ has exactly m zeroes in $\overline{B(0, 1)}$, one zero equals 1 and the other zeroes lie within $[0, 1) \setminus \{\mu(0, r_1), \dots, \mu(0, r_m)\}$.

Proof. This is an immediate consequence of $g(1) = 1$ and Lemmas 1-2. \square

Denote the zeroes different from 1 by q_i for $i \in \{1, \dots, m-1\}$. On account of the analyticity property of PGFs, the right-hand side of (14) vanishes at $p = q_i$:

$$\sum_{l=1}^m p_l \gamma_l^{(0)}(q_i) \frac{\mathbb{E}[\mu^C(0, r_l)]}{1-\mu(0, r_l)} \phi^*(\mu(0, r_l)) = 0, \quad i = 1, \dots, m-1. \quad (15)$$

Furthermore, application of the normalisation condition to (14) and an appeal to l'Hospital's rule yields

$$\begin{aligned} \sum_{l=1}^m p_l \frac{\mathbb{E}[\mu^C(0, r_l)]}{1-\mu(0, r_l)} \phi^*(\mu(0, r_l)) &= -\lambda(\mathbb{E}[B] + \mathbb{E}[V]) + \sum_{l=1}^m p_l \frac{1-r_l\lambda\mathbb{E}[B]}{1-r_l} \\ &= \mathbb{E}[N] \left\{ 1 - \lambda \left(\mathbb{E}[B] + \frac{\mathbb{E}[V]}{\mathbb{E}[N]} \right) \right\} \quad . \end{aligned} \quad (16)$$

All together, $\mathbb{E} [\mu^C(0, r_1)] , \dots, \mathbb{E} [\mu^C(0, r_m)]$ follow upon solving the set of m linearly independent Equations (15) - (16). In conclusion, $\mathbb{E} [p^C]$ is now completely specified for $p \in \overline{B(0, 1)}$, from (14):

$$\mathbb{E} [p^C] = \frac{(1-p) \sum_{l=1}^m p_l \gamma_l^{(0)}(p) \frac{\mathbb{E}[\mu^C(0, r_l)]}{1-\mu(0, r_l)} \phi^*(\mu(0, r_l))}{\phi^*(p) \sum_{l=1}^m p_l \gamma_l^{(0)}(p) - 1} . \quad (17)$$

Note that with this information, the joint transform $\mathbb{E} [e^{-\alpha L} p^C]$ is completely determined by Equation (13). The LST of L is found by letting $p \rightarrow 1$ in (13), leading to (cf. also (10)):

$$\mathbb{E} [e^{-\alpha L}] = \sum_{l=1}^m p_l \frac{(1-r_l)\beta(\alpha)}{1-r_l\beta(\alpha)} \left\{ 1 - \frac{\alpha}{\alpha + \lambda(1-\mu(\alpha, r_l))} \mathbb{E} [\mu^C(\alpha, r_l)] \phi^*(\mu(\alpha, r_l)) \right\} . \quad (18)$$

It readily follows from (18) that

$$\mathbb{E} [L] = \sum_{l=1}^m p_l \left[\frac{\mathbb{E} [B]}{1-r_l} + \frac{1}{\lambda(1-\mu(0, r_l))} \mathbb{E} [\mu^C(0, r_l)] \phi^*(\mu(0, r_l)) \right] . \quad (19)$$

3.3 N constant

In the case of constant N , $\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i]$ is the coefficient corresponding to r^N in $\pi_i(p, \alpha, r)$. In the sequel, we rewrite $\pi_i(p, \alpha, r)$ explicitly as a power series in r , so that the coefficient corresponding to r^N can be readily obtained. To this end, we need to express $\mu^m(0, r)$ as a power series in r . This expression is provided in [4] (p. 243):

$$\mu^m(\alpha, r) = \sum_{l=m}^{\infty} \frac{m}{l} r^l I_{l,m}^{(\alpha)} , \quad m \geq 1 , \quad (20)$$

with

$$I_{l,m}^{(\alpha)} := \int_0^{\infty} e^{-(\alpha+\lambda)t} \frac{(\lambda t)^{l-m}}{(l-m)!} dB^{l*}(t) , \quad m \geq 1 , \quad l \geq m . \quad (21)$$

For notational convenience, let us define

$$S_{l,i}^{(\alpha)} := \sum_{m=0}^{l-i} \left(\frac{\lambda}{\alpha + \lambda} \right)^m \frac{m+i}{l} I_{l,m+i}^{(\alpha)} , \quad l \geq 1 , \quad i \geq 0 . \quad (22)$$

Caution should be used in applying (20) as it only holds for $m \geq 1$, i.e. in calculating $\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i]$ we have to treat the cases $i = 0$ and $i \geq 1$ separately. We start with the case $i = 0$ and rewrite $\pi_0(p, \alpha, r)$ for fixed p and $|r| < |p/\beta(\alpha + \lambda(1-p))|$ as

$$\begin{aligned}
\pi_0(p, \alpha, r) &= \frac{r\beta(\alpha + \lambda(1-p))}{p - r\beta(\alpha + \lambda(1-p))} \left[1 - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1 - \mu(\alpha, r))} \right] \\
&= \sum_{k=1}^{\infty} r^k \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^k \left[1 - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \left\{ 1 + \sum_{m=1}^{\infty} \left(\frac{\lambda}{\alpha + \lambda} \right)^m \mu(\alpha, r)^m \right\} \right] \\
&= \sum_{k=1}^{\infty} r^k \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^k \left[\frac{\lambda p}{\alpha + \lambda} - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{l=1}^{\infty} r^l S_{l,0}^{(\alpha)} \right] \\
&= \frac{\lambda p}{\alpha + \lambda} \sum_{k=1}^{\infty} r^k \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^k \\
&\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{k=1}^{\infty} r^k \sum_{j=1}^{k-1} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{k-j,0}^{(\alpha)} . \tag{23}
\end{aligned}$$

Note that in the second equation we have employed that $|\mu(\alpha, r)| < 1$. Analogously, we obtain:

$$\begin{aligned}
\pi_i(p, \alpha, r) &= p^i \sum_{k=1}^{\infty} r^k \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^k \\
&\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{k=i+1}^{\infty} r^k \sum_{j=1}^{k-i} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{k-j,i}^{(\alpha)} , \quad i \geq 1 . \tag{24}
\end{aligned}$$

From (23) and (24), it is straightforward to pinpoint the coefficient corresponding to r^N . Hence, we obtain

$$\begin{aligned}
\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = 0] &= \frac{\lambda p}{\alpha + \lambda} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^N \\
&\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{j=1}^{N-1} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{N-j,0}^{(\alpha)} , \\
\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i] &= p^i \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^N \\
&\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{j=1}^{N-i} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{N-j,i}^{(\alpha)} , \quad i \geq 1 . \tag{25}
\end{aligned}$$

Note that when $i \geq N$, the second term in (25) vanishes. By introducing

$$S_{0,0}^{(\alpha)} := 1 ,$$

these cases can be combined as follows:

$$\begin{aligned}
\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i] &= p^i \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^N \\
&\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{j=1}^{N-i} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{N-j,i}^{(\alpha)} , \quad i \geq 0 . \tag{26}
\end{aligned}$$

In the appendix we give a probabilistic meaning to terms occurring in the above equation.

From now on, we assume that $N \geq 2$. Note that when $N = 1$, the system is equivalent to an M/G/1 queue without server vacations and with the LST of the service times given by $\beta(p)\phi(p)$.

Summing over all possible values of E_{n-1} leads to the unconditioned joint transform of L_n and C_n :

$$\begin{aligned} \mathbb{E} [e^{-\alpha L_n} p^{C_n}] &= \sum_{i=0}^{\infty} \mathbb{P}(E_{n-1} = i) \mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i] \\ &= \mathbb{E} [p^{E_{n-1}}] \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^N \\ &\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{i=1}^{N-1} \mathbb{P}(E_{n-1} = i) \sum_{j=1}^{N-i} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{N-j,i}^{(\alpha)}. \end{aligned}$$

Going to steady state and relying on Relation (2) between the E_n and C_n , we obtain:

$$\begin{aligned} \mathbb{E} [e^{-\alpha L} p^C] &= \mathbb{E} [p^C] \phi^*(p) \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^N \\ &\quad - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda} \sum_{i=1}^{N-1} \mathbb{P}(E = i) \sum_{j=1}^{N-i} \left(\frac{\beta(\alpha + \lambda(1-p))}{p} \right)^j S_{N-j,i}^{(\alpha)}. \end{aligned} \quad (27)$$

Letting $\alpha \rightarrow 0$ yields

$$\mathbb{E} [p^C] [p^N - \phi^*(p)\beta^*(p)^N] = -(1-p)p^N \sum_{i=1}^{N-1} \mathbb{P}(E = i) \sum_{j=1}^{N-i} \left(\frac{\beta^*(p)}{p} \right)^j S_{N-j,i}^{(0)}. \quad (28)$$

The boundary probabilities $\mathbb{P}(E = i)$ can be computed numerically by a combination of Rouché's theorem, the analyticity property and the normalisation condition of PGFs. First, the function $p \mapsto p^N - \phi^*(p)\beta^*(p)^N$ has N zeroes in $\overline{B(0,1)}$, one of which equals 1 [1]. Second, on account of the analyticity property, the right-hand side of (28) vanishes at those zeroes. This provides $N - 1$ useful equations in $\mathbb{P}(E = i)$ (the zero at 1 produces a trivial equation), which together with the normalisation condition provides N linear equations in the N boundary probabilities.

As a result, $\mathbb{E} [p^C]$ and $\mathbb{E} [e^{-\alpha L} p^C]$ are now completely determined. Letting $p \rightarrow 1$ in (27) yields

$$\mathbb{E} [e^{-\alpha L}] = \beta(\alpha)^N - \frac{\alpha}{\alpha + \lambda} \sum_{i=1}^{N-1} \mathbb{P}(E = i) \sum_{j=1}^{N-i} \beta(\alpha)^j S_{N-j,i}^{(\alpha)}. \quad (29)$$

It easily follows from (29) that

$$\mathbb{E} [L] = \mathbb{E} [B] \mathbb{E} [N] + \frac{1}{\lambda} \sum_{i=1}^{N-1} \mathbb{P}(E = i) \sum_{j=1}^{N-i} S_{N-j,i}^{(0)}. \quad (30)$$

3.4 N general

In this subsection we sketch an approach for generally distributed order size N . Starting-point is formed by the following two equations:

$$\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i] = \sum_{k=1}^{\infty} \mathbb{P}(N = k) \mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i, N = k],$$

and, as already observed above (3),

$$\sum_{k=1}^{\infty} r^k \mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i, N = k] = \pi_i(p, \alpha, r) , \quad (31)$$

with $\pi_i(p, \alpha, r)$ given in (3). The key observation of the approach is that the coefficient of r^k in (31) is obtained via contour integration (counterclockwise) of $z \mapsto \pi_i(p, \alpha, z)$ along a circle D with radius 1:

$$\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i, N = k] = \frac{1}{2\pi i} \int_D \frac{\pi_i(p, \alpha, z)}{z^{k+1}} dz .$$

We can write, using (3):

$$\begin{aligned} \mathbb{E} [e^{-\alpha L_n} p^{C_n}] &= \sum_{k=1}^{\infty} \mathbb{P}(N = k) \sum_{i=0}^{\infty} \mathbb{P}(E_{n-1} = i) \frac{1}{2\pi i} \int_D \frac{\pi_i(p, \alpha, z)}{z^{k+1}} dz \\ &= \sum_{k=1}^{\infty} \mathbb{P}(N = k) \frac{1}{2\pi i} \int_D \frac{z\beta(\alpha + \lambda(1-p))}{p - z\beta(\alpha + \lambda(1-p))} \\ &\quad \left[\mathbb{E} [p^{E_{n-1}}] - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1 - \mu(\alpha, z))} \mathbb{E} [\mu(\alpha, z)^{E_{n-1}}] \right] \frac{1}{z^{k+1}} dz \\ &= \frac{1}{2\pi i} \int_D \beta(\alpha + \lambda(1-p)) \mathbb{E} \left[\left(\frac{1}{z} \right)^N \right] \frac{\mathbb{E} [p^{E_{n-1}}] - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1 - \mu(\alpha, z))} \mathbb{E} [\mu(\alpha, z)^{E_{n-1}}]}{p - z\beta(\alpha + \lambda(1-p))} dz . \end{aligned} \quad (32)$$

According to the theory of Cauchy integration, the integral in the last line of (32) equals the sum of the residues for $z \in D^+$, the interior of D . To determine the sum of the residues, one first needs to determine which non-removable singularities lie in D^+ . First, observe that $\mu(\alpha, z)$ is analytic and bounded by one for $|z| \leq 1$. Hence, $\alpha + \lambda(1 - \mu(\alpha, z))$ is not equal to zero for $z \in D^+$, while $\mathbb{E} [\mu(\alpha, z)^{E_{n-1}}]$ is analytic in D^+ . At first sight it may seem that $z = \frac{p}{\beta(\alpha + \lambda(1-p))}$ is a pole, which is possibly located inside the unit circle D . However, remembering the definition of $\mu(\alpha, z)$ as the unique zero, inside the unit circle, of $p \mapsto p - z\beta(\alpha + \lambda(1-p))$ (see below (3)), and observing the numerator of the quotient in the last line of (32), it is seen that this z is a removable singularity. Hence, the only non-removable singularities in D^+ are the poles of $\mathbb{E} [(\frac{1}{z})^N]$ in D^+ ; otherwise, the integrand is analytic in z .

Let us consider the example of Subsection 3.1, viz., $N \sim \text{geom}(r)$. Then $\mathbb{E} [(\frac{1}{z})^N] = \frac{1-r}{z-r}$, and, as a result, it only has the pole $z = r$ in D^+ . The residue at this pole is readily seen to equal the righthand side of (4). One may similarly verify Formula (13) for the case of N being distributed according to a finite mixture of geometric terms, which was treated in Subsection 3.2; and from there, we can proceed as in those two subsections.

While the above two cases of N having an infinite support do not seem to give any serious difficulties (one should of course prove that interchange of the summation and integration is allowed), serious problems do arise when one considers the case of N having a Poisson(ζ) distribution. In this case, $\mathbb{E} [(\frac{1}{z})^N] = e^{-\zeta(1-\frac{1}{z})}$. Hence, $z = 0$ now is an essential singularity. Calculating the residue at this essential singularity seems impossible. Therefore, rather than trying to carry out the Cauchy integration by considering the non-removable singularities in D^+ , one could now focus on the non-removable singularities in the exterior of D . However, $\frac{1}{\alpha + \lambda(1 - \mu(\alpha, z))}$ and $\mathbb{E} [\mu(\alpha, z)^{E_{n-1}}]$ may have many non-removable singularities in the exterior of D , and determination of their residues might be a tedious task. We leave this as an open problem.

Let us now consider some examples of the easier case in which N has a finite support. Firstly, N is constant, viz., the case treated in Subsection 3.3. Again, we need to determine $\mathbb{E} [e^{-\alpha L_n} p^{C_n} | E_{n-1} = i]$. As observed in the introduction to Subsection 3.3, that is the coefficient corresponding to r^N in $\pi_i(p, \alpha, r)$. Formula (32) now simplifies to:

$$\mathbb{E} [e^{-\alpha L_n} p^{C_n}] = \frac{1}{2\pi i} \int_D \beta(\alpha + \lambda(1-p)) \left(\frac{1}{z}\right)^N \frac{\mathbb{E} [p^{E_{n-1}}] - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, z))} \mathbb{E} [\mu(\alpha, z)^{E_{n-1}}]}{p - z\beta(\alpha + \lambda(1-p))} dz . \quad (33)$$

According to the theory of Cauchy contour integration, the only nonzero contribution of this integral is the coefficient of z^{-1} in the integrand of (33). Hence, one needs to pinpoint the coefficient of z^{N-1} in the last part of that integrand. This exercise was done in Subsection 3.3 and won't be repeated here.

If $N - 1$ is binomially distributed, viz., $N \sim 1 + \text{bin}(K, q)$ (we assume N is always at least 1), then one should similarly determine the coefficients of z^K, z^{K-1}, \dots, z in the quotient featuring in the integrand of (33).

Note that the above examples are all special cases of distributions with rational PGFs, i.e.,

$$\mathbb{E} [z^N] = \frac{\sum_{n=1}^{D_1} c_n z^n}{\sum_{n=0}^{D_2} d_n z^n} ,$$

with D_1 and D_2 the degrees of respectively the numerator and denominator of $\mathbb{E} [z^N]$. By writing $\mathbb{E} [(1/z)^N]$ as

$$\mathbb{E} [(1/z)^N] = z^{D_2 - D_1} \frac{\sum_{n=1}^{D_1} c_n z^{D_1 - n}}{\sum_{n=0}^{D_2} d_n z^{D_2 - n}} ,$$

it is easy to determine its singularities in D^+ . Indeed, $z^{D_2 - D_1}$ produces a pole, of order $D_1 - D_2$, at $z = 0$ if and only if $D_1 > D_2$. In addition, as $\sum_{n=0}^{D_2} d_n z^{D_2 - n}$ is a polynomial of degree D_2 , it has D_2 zeroes in the complex plane (counted by their multiplicity). Some of these zeroes, say $\hat{z}_1, \dots, \hat{z}_\kappa$, may be located within D^+ , constituting poles of $\mathbb{E} [(1/z)^N]$. Hence, in combination with the residue theorem and (32), we obtain

$$\mathbb{E} [e^{-\alpha L_n} p^{C_n}] = \sum_{k=1}^{\kappa} \frac{1}{(m_k - 1)!} \frac{\partial^{m_k - 1}}{\partial z^{m_k - 1}} (z - \hat{z}_k)^{m_k} \mathbb{E} [(1/z)^N] g(\alpha, p, z) \Big|_{z=\hat{z}_k} + f(\alpha, p) ,$$

with m_k the multiplicity of \hat{z}_k and

$$g(\alpha, p, z) := \beta(\alpha + \lambda(1-p)) \frac{\mathbb{E} [p^{E_{n-1}}] - \frac{\alpha + \lambda(1-p)}{\alpha + \lambda(1-\mu(\alpha, z))} \mathbb{E} [\mu(\alpha, z)^{E_{n-1}}]}{p - z\beta(\alpha + \lambda(1-p))} ,$$

$$f(\alpha, p) := \begin{cases} \frac{1}{(D_1 - D_2 - 1)!} \frac{\partial^{D_1 - D_2 - 1}}{\partial z^{D_1 - D_2 - 1}} z^{D_1 - D_2} \mathbb{E} [(1/z)^N] g(\alpha, p, z) \Big|_{z=0} & \text{if } D_1 > D_2 , \\ 0 & \text{else .} \end{cases}$$

4 System content at arbitrary time instants

In the previous section, we have obtained the PGF of the number of customers present at the beginning of an arbitrary vacation. We now deduce a relation between this PGF and the PGF

$U(p)$ of the number of customers at random instants. As we examine a system with vacations, we invoke Fuhrmann-Cooper's decomposition result (Proposition 5 of [8]). It states that $U(p)$ equals the product of the PGF of the system content at a random instant in a regular M/G/1 queueing system (i.e., $\frac{(1-\lambda\mathbb{E}[B])(1-p)\beta^*(p)}{\beta^*(p)-p}$) and the PGF $\chi(p)$ of the number of customers χ present in the system under investigation at a random non-serving period. It is important to bear in mind that a non-serving period (called vacation period by Fuhrmann and Cooper) consists of (forced) vacations after having served N customers *and* periods during which the server is idle simply because the system is empty. Hence, we obtain

$$\chi(p) = 1 - q + q\mathbb{E}[p^\chi | \text{on vacation}] , \quad (34)$$

with q the probability that the server is on vacation if the server is in non-serving mode. The system content at a random vacation epoch equals the sum of the number of customers at the beginning of such a period and the number of customers that have arrived during the elapsed vacation time. The former's PGF equals $\mathbb{E}[p^C]$ and has been derived in Section 3 for various choices of the distribution of N . The latter's PGF is equal to

$$\frac{1 - \phi^*(p)}{\mathbb{E}[V] \lambda(1-p)} . \quad (35)$$

Let us now calculate q . By applying the relation between conditional and joint probabilities, we can write

$$q = \frac{\mathbb{P}(\text{on vacation})}{\mathbb{P}(\text{not serving})} .$$

In a system in steady state, the average number of customers arriving per time unit, λ , equals the average number of customers departing per time unit, which, in turn, equals $(1 - \mathbb{P}(\text{not serving}))/\mathbb{E}[B]$. As a result, $\mathbb{P}(\text{not serving}) = 1 - \lambda\mathbb{E}[B]$, leading to

$$q = \frac{\mathbb{P}(\text{on vacation})}{1 - \lambda\mathbb{E}[B]} . \quad (36)$$

We use a similar argument to deduce $\mathbb{P}(\text{on vacation})$. Consider the period between two successive starts of a vacation as one cycle. On average, $\mathbb{E}[N]$ customers are served in a cycle, so $\mathbb{E}[N]$ customers arrive on average per cycle. During a vacation, on average $\lambda\mathbb{E}[V]$ customers arrive. Hence

$$\mathbb{P}(\text{on vacation}) = \frac{\lambda\mathbb{E}[V]}{\mathbb{E}[N]} . \quad (37)$$

Substitution of (37) into (36) yields

$$q = \frac{\lambda\mathbb{E}[V]}{(1 - \lambda\mathbb{E}[B])\mathbb{E}[N]} . \quad (38)$$

The combination of Fuhrmann-Cooper's decomposition result with (34), (35) and (38) eventually yields:

$$U(p) = \left[1 - \frac{\lambda\mathbb{E}[V]}{(1 - \lambda\mathbb{E}[B])\mathbb{E}[N]} + \frac{\lambda\mathbb{E}[V]}{(1 - \lambda\mathbb{E}[B])\mathbb{E}[N]} \mathbb{E}[p^C] \frac{1 - \phi^*(p)}{\mathbb{E}[V] \lambda(1-p)} \right] \times \frac{(1 - \lambda\mathbb{E}[B])(1-p)\beta^*(p)}{\beta^*(p) - p} . \quad (39)$$

Hence, the mean number of customers in steady state is given by

$$\mathbb{E}[U] = \frac{\lambda^2 \mathbb{E}[B^2]}{2(1 - \lambda \mathbb{E}[B])} + \lambda \mathbb{E}[B] + \frac{\lambda \mathbb{E}[V]}{(1 - \lambda \mathbb{E}[B]) \mathbb{E}[N]} \left(\mathbb{E}[C] + \lambda \frac{\mathbb{E}[V^2]}{2 \mathbb{E}[V]} \right). \quad (40)$$

Remark.

An expression for $U(p)$ may also be obtained in the following way: First study the joint distribution of the number of customers Z immediately after a customer departure, and the index of that customer within an order (i.e., the departing customer is number $1, 2, \dots, N$ within that order). This may be done in a similar (but somewhat more complicated) way as the number of customers immediately after a departure has been studied in the classical M/G/1 queue, cf. Section II.4.2 of [4]. Then use that the marginal distribution of Z coincides with the distribution of the number of customers just before a customer arrival and, by PASTA, in steady state.

5 Sojourn time

Next to the above-used queue length decomposition result, Fuhrmann and Cooper also derive a similar decomposition result for sojourn times ([8], Proposition 4), viz., a relation between the sojourn time D (service time included) of a random customer in the vacation system and the sojourn time \hat{D} of a random customer in the M/G/1 system:

$$\mathbb{E}[e^{-sD}] = \chi(1 - s/\lambda) \mathbb{E}[e^{-s\hat{D}}].$$

Applying this result, we obtain

$$\begin{aligned} \mathbb{E}[e^{-sD}] &= \left[1 - \frac{\lambda \mathbb{E}[V]}{(1 - \lambda \mathbb{E}[B]) \mathbb{E}[N]} + \frac{\lambda \mathbb{E}[V]}{(1 - \lambda \mathbb{E}[B]) \mathbb{E}[N]} \mathbb{E}[(1 - s/\lambda)^C] \frac{1 - \phi(s)}{\mathbb{E}[V] s} \right] \\ &\quad \times \frac{(1 - \lambda \mathbb{E}[B]) \beta(s) s}{\lambda \beta(s) + s - \lambda}. \end{aligned} \quad (41)$$

As a result, the Laplace-Stieltjes transform of the waiting time (service time excluded) reads:

$$\begin{aligned} \mathbb{E}[e^{-sW}] &= \left[1 - \frac{\lambda \mathbb{E}[V]}{(1 - \lambda \mathbb{E}[B]) \mathbb{E}[N]} + \frac{\lambda \mathbb{E}[V]}{(1 - \lambda \mathbb{E}[B]) \mathbb{E}[N]} \mathbb{E}[(1 - s/\lambda)^C] \frac{1 - \phi(s)}{\mathbb{E}[V] s} \right] \\ &\quad \times \frac{(1 - \lambda \mathbb{E}[B]) s}{\lambda \beta(s) + s - \lambda}. \end{aligned} \quad (42)$$

6 Numerical results

In this section we present some compendious numerical examples. Their role is mainly to illustrate the influence of particular parameters on one key performance measure, viz., $\mathbb{E}[C]$; the influence on the other performance measures is similar and therefore not shown. We plot $\mathbb{E}[C]$ as a function of ρ and as a function of $\mathbb{E}[N]$. $\mathbb{E}[C]$ decreases as $\mathbb{E}[N]$ increases, because the server takes fewer vacations per unit of time.

In our first example, we fix $\mathbb{E}[N]$ at 3, λ at 1 and use $\exp(4)$ distributed service times. We investigate the dependence of $\mathbb{E}[C]$ on ρ by using for $j = 0, \dots, 22$, $V_j \sim E_j(10)$ as time duration of the vacations, so that $\rho_j = \frac{1}{4} + \frac{j}{30} \leq \frac{59}{60}$. We distinguish between three cases: (i) $N = 3$

constant, (ii) $N \sim geo(\frac{2}{3})$ and (iii) N is a mixture of geometrically distributed random variables ($p_1 = \frac{1}{4}, r_1 = \frac{1}{2}, p_2 = \frac{3}{4}, r_2 = \frac{7}{10}$). The result is shown in Fig. 2. Indeed, we see that $\mathbb{E}[C]$ is increasing and that for fixed ρ it is largest in case of N a mixture, while it is smallest exactly when N is taken to be constant. The last observation makes sense as the mixture of geometrically distributed variables has a larger variance than the singular geometric distribution which in its turn has a non-zero variance.

Our second example shows that $\mathbb{E}[C]$ decreases with increasing $\mathbb{E}[N]$. More specifically, in Fig. 3, $\mathbb{E}[C]$ has been plotted for different values of r , where the blue graph corresponds to a mixture of geometrically distributed random variables ($p_1 = \frac{1}{4}, r_1 = \frac{1}{2}, p_2 = \frac{3}{4}, r_2 = r$) and the green one to a geometric distribution with parameter $\frac{2}{3}$. In both cases we took $\exp(4)$ as distribution for the service times, λ fixed at 1 and the vacations $E_{10}(10)$ distributed. In Fig. 4 exactly the same setup is used, except that now the vacations are $E_{16}(10)$ distributed. Once more we see that for fixed r , $\mathbb{E}[C]$ is larger in the case of a mixture, which we can contribute to its larger variance.

7 Conclusions and suggestions for further research

We have provided a detailed exact analysis of an M/G/1 type queue with the special feature that the server serves exactly N customers (a deterministic or random number) before taking a vacation. We have in particular determined the joint transform of the length of a server visit period and the number of customers in the system at the end of that period. We have also derived the generating function and mean of the number of customers at a random instant, and the Laplace-Stieltjes transform of the delay of a customer.

As a next step, one might introduce a cost function in this system. For example, there could be customer holding costs C_h , switching costs C_s when the server leaves the queue, and idling costs C_i when the server is idle during a visit period, and our goal might be to choose N (when deterministic) or r (when $N \sim geo(r)$) such as to minimize the long-term average costs. This amounts to minimizing the cost function F , given by

$$F = C_h \mathbb{E}[U] + C_s \frac{1}{\mathbb{E}[L] + \mathbb{E}[V]} + C_i (\mathbb{E}[L] - \mathbb{E}[N] \mathbb{E}[B]) . \quad (43)$$

Notice that $\mathbb{E}[U]$ is given in (40) and $\mathbb{E}[L]$ in (11) for constant N and in (30) for $N \sim geo(r)$.

From a vacation theoretic point of view, it might be interesting to study variants of our model, like the case of multiple vacations (i.e., if the server returns from a vacation and finds the system empty, it immediately takes another vacation). It seems that our approach, which is based on studying the system at the time points immediately after N customers have been served, can be easily adapted to handle that case. From an order picking point of view, this model variant could also be of relevance, the order picker spending the extra vacations on some other tasks. Another model variant is the case in which the server also takes vacations anytime when the system becomes empty. This variant seems less natural from the order picking point of view, and might also require a different approach. Actually, this variant is called N -limited or E -limited in the vacation literature; see, e.g., Section 2.6 of Takagi [10].

Appendix: Interpretation of the probability generating function in (26)

In this appendix, we give a probabilistic meaning to terms occurring in Equation (26), taking $\alpha = 0$:

$$\begin{aligned} \mathbb{E} [p^{C_n} | E_{n-1} = i] &= p^i \left(\frac{\beta^*(p)}{p} \right)^N \\ &\quad - (1-p) \sum_{j=1}^{N-i} \left(\frac{\beta^*(p)}{p} \right)^j S_{N-j,i}^{(0)}, \quad i \geq 0. \end{aligned} \quad (44)$$

The main purpose is to shed some light on the occurrence of the various terms in the summation. First observe that, cf. (21), with $\Sigma_l := B_1 + \dots + B_l$ and with $A(\Sigma_l)$ the number of arrivals during Σ_l , one can write for $l \geq m$:

$$I_{l,m}^{(0)} = \mathbb{P}(A(\Sigma_l) = l - m). \quad (45)$$

In addition, we notice that for $j = 0, 1, \dots, N$,

$$\left(\frac{\beta^*(p)}{p} \right)^j = \left(\frac{\mathbb{E}[p^{A(B_1)}]}{p} \right)^j = \mathbb{E}[p^{A(\Sigma_j) - j}]. \quad (46)$$

With (45)-(46) at hand and invoking (22), we rewrite Expression (44):

$$\begin{aligned} \mathbb{E} [p^{C_n} | E_{n-1} = i] &= \mathbb{E}[p^{A(\Sigma_N) - N + i}] \\ &\quad - (1-p) \left\{ \sum_{j=1}^{N-i} \sum_{m=0}^{N-j-i} \frac{i+m}{N-j} \mathbb{P}(A(\Sigma_{N-j}) = N - j - (i+m)) \mathbb{E}[p^{A(\Sigma_N - \Sigma_{N-j}) - j}] \right\}, \end{aligned} \quad (47)$$

when $1 \leq i < N$ and,

$$\mathbb{E} [p^{C_n} | E_{n-1} = i] = \mathbb{E}[p^{A(\Sigma_N) - N + i}],$$

when $i \geq N$.

Now for $i \geq N$, the expression for $\mathbb{E} [p^{C_n} | E_{n-1} = i]$ is obvious, since then the server is not waiting for items to arrive and in this busy period precisely $A(\Sigma_N)$ items arrive during services $1, \dots, N$. Since the initial queue consisted of i items and N of those leave, $C_n = A(\Sigma_N) - N + i$.

For $1 \leq i < N$, $\mathbb{E} [p^{C_n} | E_{n-1} = i]$ seems to be obtained after conditioning on the different arrival patterns that can occur during the service of one full order. E.g., all other $N - i$ customers arrive timely, so that the server does not become idle during L_n (so that $C_n = A(\Sigma_N) - N + i$), or an order could consist of several busy periods. To illustrate this we will calculate $\mathbb{E} [p^{C_n} | E_{n-1} = i]$ for $i = N - 1$ and $i = N - 2$.

The case $i = N - 1$

For $i = N - 1$, (47) reduces to:

$$\mathbb{E} [p^{C_n} | E_{n-1} = N - 1] = \mathbb{E}[p^{A(\Sigma_N) - 1}] - (1-p) \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N) - 1}]. \quad (48)$$

We can interpret/rederive this formula as follows. Since the visit period starts with $N - 1$ items, there are two possibilities: (i) there is at least one arrival in Σ_{N-1} , so that the server is not idle after Σ_{N-1} and immediately continues to serve the N -th (last) item of the order; and (ii) there is no arrival in Σ_{N-1} , so that the server becomes idle after Σ_{N-1} , waiting for the N -th item of the

order to arrive. Hence,

$$\begin{aligned}
\mathbb{E} [p^{C_n} | E_{n-1} = N - 1] &= \mathbb{E}[p^{A(\Sigma_N)-1} I(A(\Sigma_{N-1}) > 0)] + \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N)}] \\
&= \mathbb{E}[p^{A(\Sigma_N)-1}] - \mathbb{E}[p^{A(\Sigma_N)-1} I(A(\Sigma_{N-1}) = 0)] \\
&\quad + \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N)}] \\
&= \mathbb{E}[p^{A(\Sigma_N)-1}] - (1 - p) \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N)-1}] ,
\end{aligned}$$

confirming (48).

The case $i = N - 2$

For $i = N - 2$, (47) reduces to:

$$\begin{aligned}
\mathbb{E} [p^{C_n} | E_{n-1} = N - 2] &= \mathbb{E}[p^{A(\Sigma_N)-2}] - (1 - p) \frac{N - 2}{N - 1} \mathbb{P}(A(\Sigma_{N-1}) = 1) \mathbb{E}[p^{A(B_N)-1}] \\
&\quad - (1 - p) \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N)-1}] \\
&\quad - (1 - p) \mathbb{P}(A(\Sigma_{N-2}) = 0) \mathbb{E}[p^{A(B_{N-1})+A(B_N)-2}] .
\end{aligned}$$

Let us call the four successive terms in the righthand side I,II,III and IV. We can interpret/rederive this formula as follows. Consider $\mathbb{E} [p^{C_n} | E_{n-1} = N - 2]$, distinguishing between the three cases $A(\Sigma_{N-2}) \geq 2$ (the first three terms in the righthand side below; in this case the visit period surely consists of only one busy period), $A(\Sigma_{N-2}) = 1$ (the fourth and fifth term; in this case the busy period possibly ends after the $(N - 1)$ -th service) and $A(\Sigma_{N-2}) = 0$ (the sixth and seventh term; in this case the busy period definitely ends after the $(N - 2)$ -th service):

$$\begin{aligned}
\mathbb{E} [p^{C_n} | E_{n-1} = N - 2] &= \mathbb{E}[p^{A(\Sigma_N)-2}] - \mathbb{E}[p^{A(\Sigma_N)-2} I(A(\Sigma_{N-2}) = 0)] \\
&\quad - \mathbb{E}[p^{A(\Sigma_N)-2} I(A(\Sigma_{N-2}) = 1)] \\
&\quad + \mathbb{E}[p^{A(B_{N-1})+A(B_N)-1} I(A(\Sigma_{N-2}) = 1) I(A(B_{N-1}) \geq 1)] \\
&\quad + \mathbb{E}[p^{A(B_N)} I(A(\Sigma_{N-2}) = 1) I(A(B_{N-1}) = 0)] \\
&\quad + \mathbb{E}[p^{A(B_{N-1})+A(B_N)-1} I(A(\Sigma_{N-2}) = 0) I(A(B_{N-1}) \geq 1)] \\
&\quad + \mathbb{E}[p^{A(B_N)} I(A(\Sigma_{N-2}) = 0) I(A(B_{N-1}) = 0)] .
\end{aligned}$$

We can rewrite this formula as follows, taking complements of some of the indicator functions:

$$\begin{aligned}
\mathbb{E} [p^{C_n} | E_{n-1} = N - 2] &= \mathbb{E}[p^{A(\Sigma_N)-2}] - \mathbb{P}(A(\Sigma_{N-2}) = 0) \mathbb{E}[p^{A(B_{N-1})+A(B_N)-2}] \\
&\quad - \mathbb{P}(A(\Sigma_{N-2}) = 1) \mathbb{E}[p^{A(B_{N-1})+A(B_N)-1}] \\
&\quad + \mathbb{P}(A(\Sigma_{N-2}) = 1) \mathbb{E}[p^{A(B_{N-1})+A(B_N)-1}] \\
&\quad - \mathbb{P}(A(\Sigma_{N-2}) = 1) \mathbb{P}(A(B_{N-1}) = 0) \mathbb{E}[p^{A(B_N)-1}] \\
&\quad + \mathbb{P}(A(\Sigma_{N-2}) = 1) \mathbb{P}(A(B_{N-1}) = 0) \mathbb{E}[p^{A(B_N)}] \\
&\quad + \mathbb{P}(A(\Sigma_{N-2}) = 0) \mathbb{E}[p^{A(B_{N-1})+A(B_N)-1}] \\
&\quad - \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N)-1}] \\
&\quad + \mathbb{P}(A(\Sigma_{N-1}) = 0) \mathbb{E}[p^{A(B_N)}] . \tag{49}
\end{aligned}$$

The righthand side of (49) contains nine terms, to be called T_1, \dots, T_9 . We now show that the sum of these nine terms equals $I + II + III + IV$. Firstly, $T_1 = I$. Secondly, $T_2 + T_7 = IV$. Thirdly, $T_8 + T_9 = III$. Fourthly, $T_3 + T_4 = 0$. Finally, it is seen that $T_5 + T_6 = II$, by observing that, by symmetry,

$$\mathbb{P}(A(\Sigma_{N-2}) = 1) \mathbb{P}(A(B_{N-1}) = 0) = \frac{N - 2}{N - 1} \mathbb{P}(A(\Sigma_{N-1}) = 1) .$$

The last formula gives an indication of the meaning of the terms $\frac{i+m}{N-j}$ in (47).

Acknowledgment. The authors gratefully acknowledge stimulating discussions with Ivo Adan.

References

- [1] I.J.B.F. Adan, J.S.H. van Leeuwen, E.M.M. Winands (2006), On the application of Rouché's theorem in queueing theory, *Operations Research Letters* 34, 355–360.
- [2] E. Altman, D. Fiems (2007), Expected waiting time in polling systems with correlated vacations, *Queueing Systems* 56(3-4), 241–253.
- [3] R. Andriansyah (2011), *Order-picking Workstations for Automated Warehouses*, PhD Thesis Eindhoven University of Technology.
- [4] J.W. Cohen (1982), *The Single Server Queue*, North-Holland Publ. Cy., Amsterdam.
- [5] H.E. Daniels (1961), Mixtures of geometric distributions, *Journal of the Royal Statistical Society, Series B (Methodological)* 23(2), 409–413.
- [6] B.T. Doshi (1986), Queueing systems with vacations – A survey, *Queueing Systems* 1(1), 29–66.
- [7] D. Fiems, H. Bruneel (2013), Discrete-time queueing systems with Markovian preemptive vacations, *Mathematical and Computer Modelling* 57, 782–792.
- [8] S.W. Fuhrmann, R.B. Cooper (1985), Stochastic decompositions in the M/G/1 queue with generalized vacations, *Operations Research* 33(5), 1117–1129.
- [9] V. Klimenok (2001), On the modification of Rouché's theorem for the queueing theory problems, *Queueing Systems* 38, 431–434.
- [10] H. Takagi (1991), *Queueing Analysis: A Foundation of Performance Evaluation, volume 1: Vacation and Priority Systems, Part 1*, North-Holland Publ. Cy., Amsterdam.
- [11] N. Tian and Z.G. Zhang (2006). *Vacation Queueing Models*, Springer, New York.
- [12] D. Towsley (1981), A statistical analysis of ARQ protocols operating in a nonindependent error environment, *IEEE Transactions on Communications* 29(7), 971–981.
- [13] F. Van der Duyn Schouten, S. Vanneste (1995), Maintenance optimization of a production system with buffer capacity, *European Journal of Operational Research* 82, 323–338.

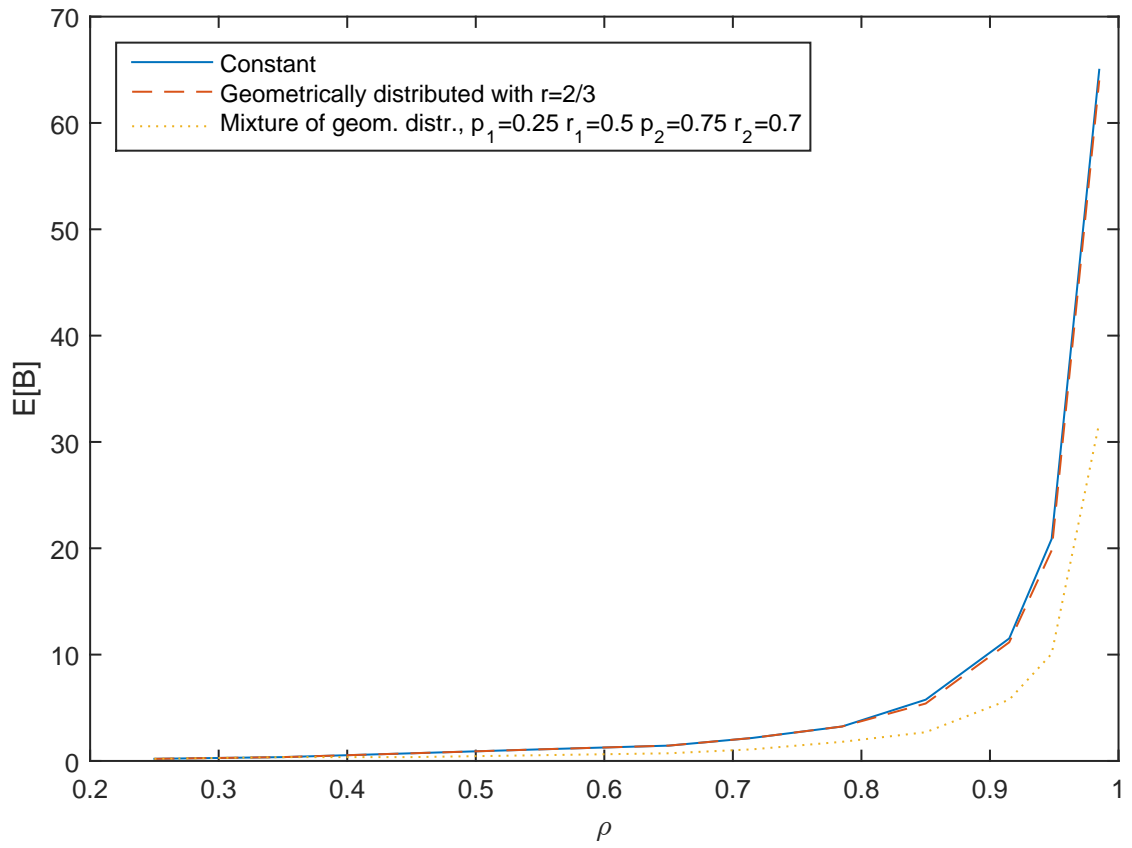


Figure 2: Relation between $\mathbb{E}[C]$ and ρ in case of $\mathbb{E}[N] = 3$.

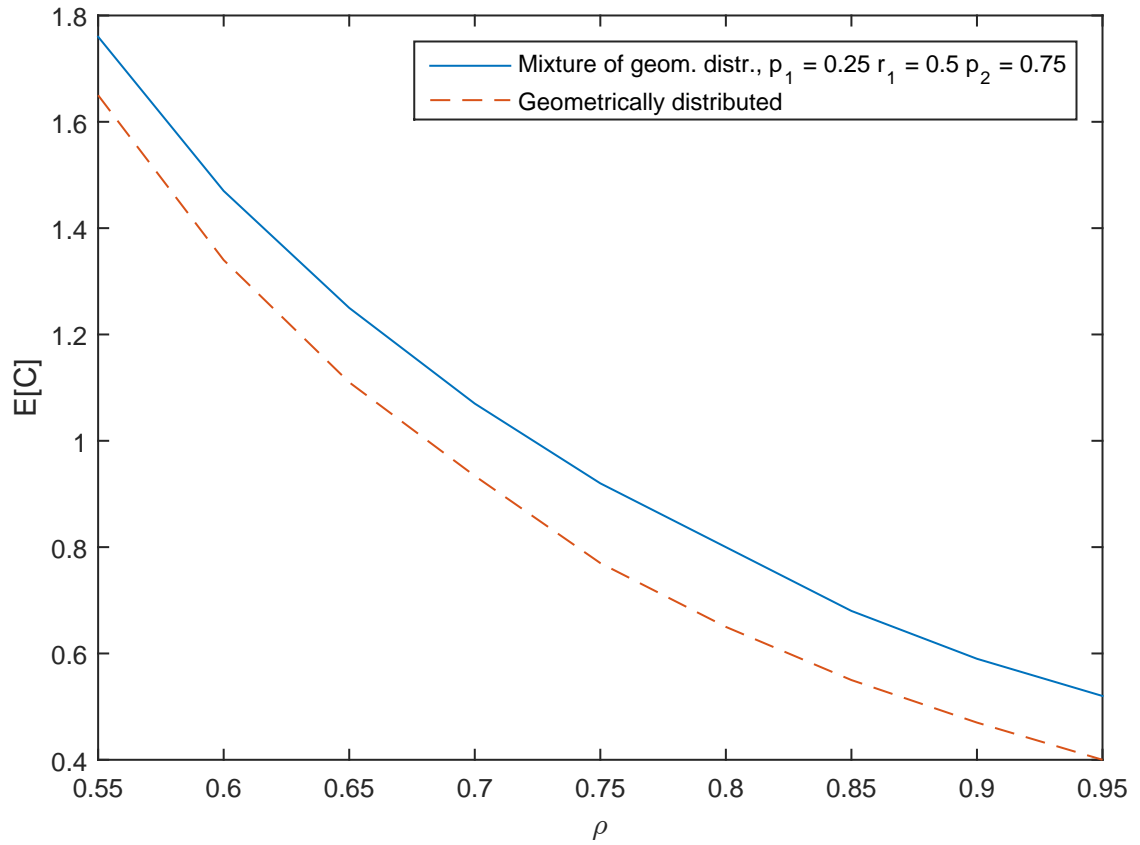


Figure 3: Relation between $\mathbb{E}[C]$ and $\mathbb{E}[N]$ in case of N having a geometric distribution ($r = 1 - 1/\mathbb{E}[N]$) and in case of a mixture of geometrically distributed random variables. The service time durations have an $\text{exp}(4)$ distribution and the vacations are $E_{10}(10)$ distributed.

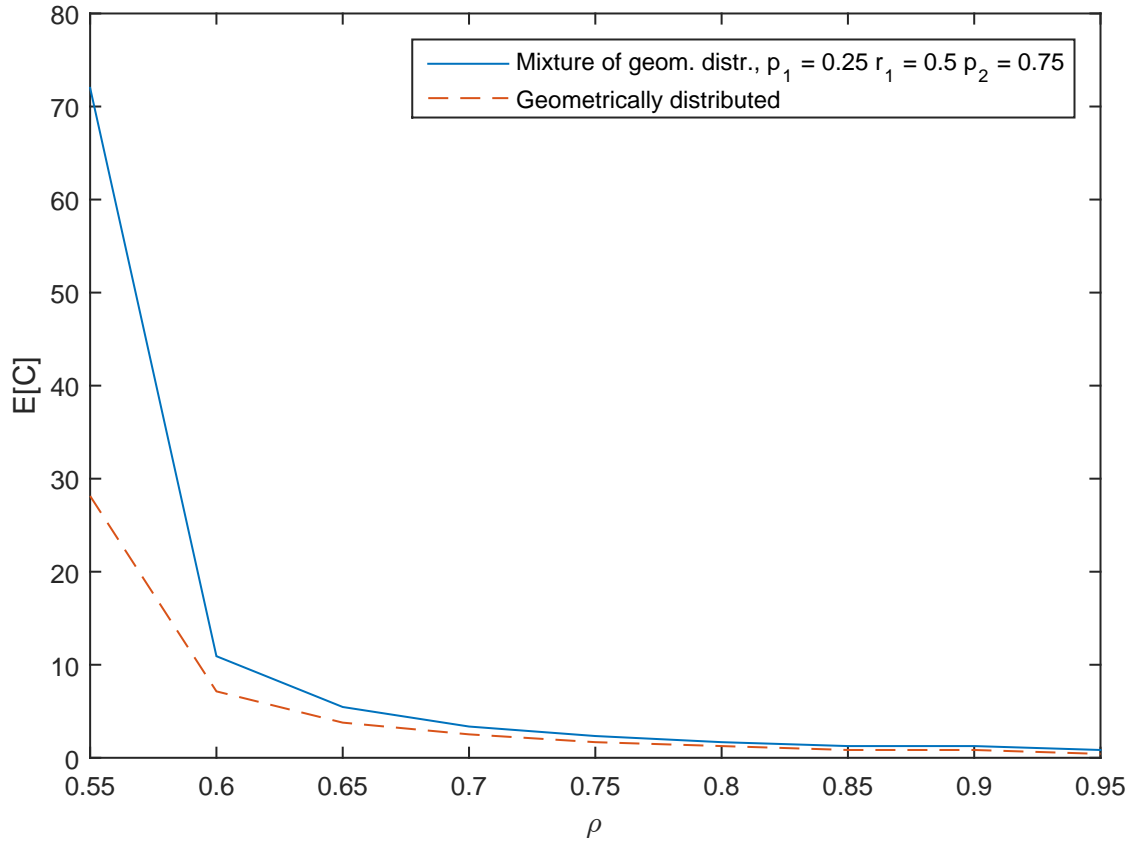


Figure 4: Relation between $\mathbb{E}[C]$ and $\mathbb{E}[N]$ in case of N having a geometric distribution ($r = 1 - 1/\mathbb{E}N$) and in case of a mixture of geometrically distributed random variables. The service time durations have an $\text{exp}(4)$ distribution and the vacations are $E_{16}(10)$ distributed.