

Batch sojourn time in polling systems on a circle

Tim Engels^{1,*}, Ivo Adan², Onno Boxma¹, and Jacques Resing¹

¹Department of Mathematics and Computer Science, Eindhoven University of Technology

²Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology

*Corresponding author: t.p.g.engels@tue.nl

July 28, 2023

Abstract

In this paper, we analyze a polling system on a circle. Random batches of customers arrive at a circle, where each customer, independently, obtains a uniform location. A single server cyclically travels over the circle to serve all customers. Using mean value analysis, we derive the expected number of waiting customers within a given distance of the server and a closed form expression for the mean batch sojourn time.

1 Introduction

Polling systems are systems of multiple queues, where a single server visits the queues in a cyclic order to complete the service requirements of customers. A vast amount of literature is devoted to the analysis of these systems; see for instance the literature reviews in Borst and Boxma [3], Takagi [15]. Polling models naturally arise in a wide range of application areas, for example in communication, traffic and transportation systems [2, 9, 12]. Throughout the years, numerous extensions of the model have been analyzed, in particular we focus on two extensions: extension to a continuum of queues, i.e. polling on a circle, and the case of batch arrivals.

Polling on a circle with batch arrivals naturally arises in warehouse logistics. Customer orders, often consisting of several different items, arrive at the warehouse, requesting to be picked in a timely manner. One way of picking these orders, is by using a so-called milk-run system [17]. In this system, a picker walks cyclically through the warehouse, picking any requested item that she encounters. This can be modeled as a polling system, where each requested item represents a customer, an order therefore can be seen as a customer batch. The server represents the picker walking through the warehouse. Warehouses often store a very large number of different items and therefore it is natural to approximate the warehouse by allowing for *continuous* pick locations, i.e. extending the model to polling on a circle. The batch sojourn time then represents the time it takes the picker to completely pick an order.

Polling on a circle, see e.g. Coffinan and Gilbert [5], Kroese and Schmidt [10, 11], assumes that customers arrive according to a Poisson process and obtain locations according to a continuous distribution on the circle. The former two papers assume that these locations are uniform, while the latter considers an extension where the customers arrive according to a certain distribution with respect to the server. Eliazar [7], considers an extended model: “the snowblower problem”, which assumes a more general arrival process of jobs, excluding bulk arrivals, on the circle and allows for a stochastic travel time over intervals. Using an analytical approach, the author derives

some results on the amount of work in intervals, in particular the expected amount of work in an interval is found. Eliazar [8] later proved that these snowblower models are limiting cases of discrete polling systems, where the number of queues tends to infinity. Hence, these models can prove to be useful for the analysis of large and complex polling models. For example, the mean waiting time has an explicit formula for many continuous polling models, whereas under discrete polling models no method has been found that is faster than solving a system of N^2 linear equations, where N is the number of queues [19]. On the other hand obtaining more intricate performance measures than the mean in polling on a circle is more challenging.

The extension to batch arrivals assumes that customers now arrive as a batch, where each such group can be considered as a single job. The locations of the customers in these batches are assumed to follow a certain distribution. In the current paper, we assume that these locations are uniformly distributed and independent from one another. The mean waiting time of customers in bulk arrival polling systems is previously analyzed in Levy and Sidi [13], providing a mean value analysis. Boxma [4] derives a conservation law for the weighted mean waiting time of customers in the system, making use of a stochastic decomposition. Other research on this topic includes that of Al Hanbali et al. [1], van der Mei [18]. The analysis of the batch sojourn time in polling systems is, as far as we know, restricted to discrete systems, see van der Gaast et al. [16]. In said paper the authors provide a mean value analysis for the mean sojourn time of a batch, again requiring one to solve a system of N^2 linear equations.

In the current paper we combine both extensions and present a mean value analysis for the sojourn time of a batch of customers under a continuous variant. For this, we analyze the number of waiting customers in the system within a certain distance of the server. Using Little's law and an explicit expression for the expected waiting time, we derive an integral equation for the number of customers at distance x of the server. Our main contributions are a novel technique used for the analysis of the mean batch sojourn time and an explicit expression for the mean batch sojourn time that lends itself for optimization purposes. The technique in this paper allows for the analysis of related systems, like systems with generally distributed arrival locations of customers. This is a topic for future research.

The paper is built up as follows: Section 2 is devoted to the description of the model and some preliminaries. We determine the steady-state mean number of waiting customers in Section 3. Section 4 is devoted to the determination of the mean of the total sojourn time of a batch of customers. Some numerical results are presented in Section 5.

2 Model description and preliminaries

Consider a circle with circumference 1. Assume that batches of customers arrive on this circle according to a Poisson process with intensity λ ; each batch is considered as a single job. Each arriving customer in the batch is assigned a location on this circle according to the uniform distribution, independent from all other customers in the same batch. The size of a customer batch, K , is assumed to be strictly positive and to follow a known distribution with probabilities p_k and probability generating function $\tilde{K}(\cdot)$ (without loss of generality we assume that $p_0 = 0$).

The service is provided by a single traveling server. We assume that the server travels at a fixed speed α^{-1} in a given direction and serves any customer she encounters, taking a randomly distributed time B . After finishing the service, the server is assumed to continue traveling in the same direction.

Extending the methodology of Kroese and Schmidt [11, Theorem 3.1] to batch arrivals shows that the system is stable when $\rho := \lambda\mathbb{E}[K]\mathbb{E}[B] < 1$, as is intuitively logical. Throughout this paper we

consider the limiting behavior of the system and in particular focus on (i) the steady-state number of waiting customers in the system, and (ii) the long-run sojourn time of a job, S^B , that is the time from the arrival of the batch of customers to the service completion of the last customer in that batch.

We define the distance between a point on the circle and the server as the distance that the server has to travel to reach that point, see Figure 1. Throughout this paper we rely on the symmetry of the model, and often instead use that *arriving customers obtain locations at a uniform distance from the server*. Additionally we focus on customers at a certain distance to the server, rather than at a location on the circle. We define $L(x)$ to be the number of waiting customers within distance x of the server, see Figure 1; we often write the shorthand L for all waiting customers in the system (i.e. $L(1)$).

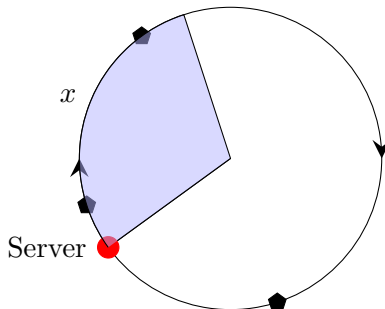


Figure 1: Illustration of the polling model and corresponding range within distance x of the server. In this example $L(x) = 2$.

3 The mean number of waiting customers

In this section we derive an expression for the expectation of the steady-state number of waiting customers in the system. We use an approach based on Mean Value Analysis. In two remarks, we briefly outline two other approaches.

Relating this model and one where customers are served in a FCFS order, we prove the following:

Lemma 1. *The expected number of waiting customers in the system, excluding the customer that is possibly in service, is given by:*

$$\mathbb{E}[L] = \frac{\lambda \mathbb{E}[K]}{2(1 - \rho)} \left(\alpha + \lambda \mathbb{E}[K] \mathbb{E}[B^2] + \mathbb{E}[B] \frac{\mathbb{E}[K(K - 1)]}{\mathbb{E}[K]} \right). \quad (1)$$

Proof. The number waiting in the system, L , can be expressed as $L_0 + L_1$, with $L_0 = L$ if the server is traveling and 0 otherwise, and $L_1 = L$ if the server is busy and 0 otherwise. According to Little's law, $E(L_0) = \lambda E(K)E(W_0)$ and $E(L_1) = \lambda E(K)E(W_1)$, where W_0 is the travel time of the server during the waiting time and W_1 is the busy time of the server during the waiting time. Clearly, $E(W_0) = \alpha/2$ and thus $E(L_0) = \lambda E(K)\alpha/2$. To determine $E(W_1)$, we interchange locations of customers to preserve service in order of arrival. So when a new batch arrives, these customers swap their locations with waiting customers that are overtaken. Interchanging locations of customers does not alter the number waiting in the system, and hence, by Little's law, it does not alter the mean waiting time, since all customers have the same service time distribution. In

case of service in order of arrival, we have

$$E(W_1) = \rho E(R) + (E(L_0) + E(L_1))E(B) + \frac{E(K(K-1))}{2E(K)}E(B).$$

Combined with $E(L_1) = \lambda E(K)E(W_1)$, both $E(W_1)$ and $E(L_1)$ follow. \square

Remark 1. *The expected number of waiting customers can also be analyzed by first considering a discrete cyclic system with N queues, and then letting N tend to ∞ ; [8] has proven that such a limiting operation, without batch arrivals, results in a continuous polling system. One can extend this result to the case of batch arrivals. Using the pseudo-conservation law for a discrete polling system with batch arrivals, cf. Equation (3.21) of [4], in combination with the symmetry of the model, then results once more in (1).*

Remark 2. *The lemma can also be proven by using the stochastic decomposition in Boxma [4, Theorem 2.1]: the expected queue length is given by the sum of the expected queue length in a corresponding system without travel time and the expected queue length at an epoch at which the server is traveling. The first term follows from a known result for the $M/G/1$ queue with batch arrivals (see, e.g., [6], Section 5.10) and equals the last two components of (1). The second term can be argued to equal the mean number of customers at the end of a cycle (one complete tour along the circle) of the server. A balancing argument shows that the mean number of customers arriving per cycle equals $\lambda \mathbb{E}[K]\alpha/(1-\rho)$, where $\alpha/(1-\rho)$ denotes the mean cycle time. A symmetry argument finally shows that, on average, half of those customers are still present at the end of the cycle.*

4 The mean sojourn time of a batch

In this section we derive the expected sojourn time $\mathbb{E}[S^B]$ of a job/batch, i.e., the time from the arrival of a batch until the service completion of its last customer. The service of jobs, now, is preempted by travel periods, as the server will have to travel between the customers. Moreover, it might happen that other jobs are (partially) served between the service of two customers in the same batch. Because of that, one cannot apply the stochastic decomposition and pseudo-conservation law in Boxma [4] to the number of batches in the system. Making use of a branching process construction, we analyze the total service time of customers served before the tagged customer, allowing us to derive an initial expression for the mean sojourn time of a batch.

Our approach consists of five steps, which are worked out in detail in the following five subsections. **Step 1** is the branching process construction. It results in an initial expression for $\mathbb{E}[S^B]$, that still involves one unknown function $f(\cdot)$. **Step 2** is the construction of an integral equation for that function $f(\cdot)$. **Step 3** is the solution of that integral equation; its uniqueness is discussed in **Step 4**. Finally, $\mathbb{E}[S^B]$ is determined in **Step 5**.

4.1 Step 1: A branching process construction

The waiting time of a tagged customer consists of the travel time of the server and the service time of all customers served before her. This latter term includes customers who might arrive later than the tagged customer, but with a location closer to the server. Inspired by Resing [14], we handle these future arrivals by linking each future arrival to either (i) a customer already present at the arrival moment of the tagged customer or (ii) the travel period, resulting in a multi-type branching process with immigration.

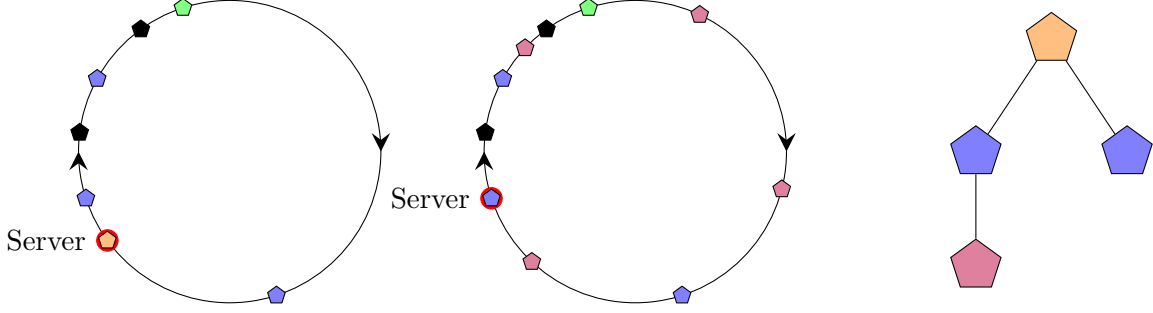


Figure 2: Illustration of the waiting time of a tagged customer (green) that is generated by a service (of the orange customer) and the corresponding branching process. During the service of the orange customer, blue customers arrive, of which only the first two are considered. During the service of the first blue customer, the red customers arrive, of which only one will be served before the tagged customer.

We construct the branching process by following two simple rules:

- A customer (a) is the offspring of a customer (b) when (a) arrives during the service of (b)
- A customer (a) is called an immigrant when she arrives during a travel period of the server.

From the perspective of a tagged customer, however, only customers who will be served before her are of interest. We therefore trim the branching process to only consider the customers who are served before the tagged customer, see Figure 2. As the offspring distribution of a customer now depends on the distance from this customer to the tagged customer, we take this distance as the type of a customer. Using this trimmed branching process we introduce the following.

Definition 1. $S(y)$ denotes the service time of a customer at distance y of the tagged customer plus the service times of all its descendants in the trimmed branching process. This is also referred to as the waiting time generated by a service at distance y .

Definition 2. $T(y)$ denotes the added service times of all customers in the branching processes of immigrants, arriving during the travel time of the server of distance y to the tagged customer. We call this the waiting time generated by the traveling of distance y .

The expectations of these random variables can be found by conditioning w.r.t. the size of the offspring, resulting in an integral equation for the first moment of $S(y)$, that allows us to obtain $\mathbb{E}[S(x)]$ and $\mathbb{E}[T(x)]$:

Lemma 2. The expected waiting time generated by a service at distance x of a customer, $\mathbb{E}[S(x)]$, is given by:

$$\mathbb{E}[S(x)] = \mathbb{E}[B] \exp(\rho x). \quad (2)$$

The mean total waiting time generated by the server traveling a distance x to the customer, $\mathbb{E}[T(x)]$, satisfies:

$$\mathbb{E}[T(x)] = \frac{\alpha}{\rho} \left[\exp(\rho x) - 1 \right]. \quad (3)$$

Proof. During a service, we know that in expectation a total of ρ customers arrive. Each of them initiates another additional service requirement of $S(y)$, where y is the location of the arriving

customer. Hence:

$$\mathbb{E}[S(x)] = \mathbb{E}[B] + \rho \int_{y=0}^x \mathbb{E}[S(y)]dy \implies \mathbb{E}[S(x)] = \mathbb{E}[B] \exp(\rho x),$$

where the last step follows from the fact that $\mathbb{E}[S(0)] = \mathbb{E}[B]$.

We can now use the above expression to prove (3). The server takes a time αdu to travel a small distance du . Therefore, in expectation, $\lambda \mathbb{E}[K] \alpha du$ customers arrive during the traveling of said distance. Each of these arrivals adds a total waiting time of $S(y)$ when arriving between the tagged customer and the server at a distance y to the tagged customer:

$$\begin{aligned} \mathbb{E}[T(x)] &= \alpha x + \int_{u=0}^x \lambda \mathbb{E}[K] \alpha \int_{y=0}^u \mathbb{E}[S(y)]dy du = \alpha x + \alpha \lambda \mathbb{E}[K] \cdot \int_{u=0}^x \frac{1}{\lambda \mathbb{E}[K]} \left[\exp(\rho u) - 1 \right] du \\ &= \frac{\alpha}{\rho} \left[\exp(\rho x) - 1 \right]. \end{aligned} \quad \square$$

Remark 3. *At the arrival instant of a customer she might encounter a residual service time first. The total expected waiting time generated by this residual service time, $\mathbb{E}[S^R(x)]$, satisfies relation (2) with $\mathbb{E}[B]$ replaced by $\mathbb{E}[B^R]$.*

For a direct analysis of the mean batch sojourn time we first condition on the distance of the furthest customer w.r.t. the server, denoted as X^B . This is the maximum of K uniform random variables, hence conditional on $K = k$ the c.d.f. of X^B is given by x^k , therefore:

$$\mathbb{E}[S^B] = \mathbb{E}[B] + \int_{x=0}^1 \sum_{k=1}^{\infty} p_k k x^{k-1} \mathbb{E}[W^B | X^B = x, K = k] dx, \quad (4)$$

where W^B denotes the waiting time of the last customer of a job. Remark that this waiting time of a customer can be split in four distinct elements: (i) the total waiting time generated by the traveling of the server over distance x : $T(x)$; (ii) the waiting time generated by the customer currently in service, if any: with probability ρ there is a contribution $S^R(x)$, where $S^R(x)$ denotes the total waiting time generated by the *residual* service of a customer; (iii) the waiting time generated by all customers in the same batch $S(Y_1) + \dots + S(Y_{k-1})$, where Y_1, \dots, Y_{k-1} denote the distance of these (unordered) customers to the tagged customer – remark that these are uniform on $[0, x]$; (iv) the waiting time generated by all customers present at the time of arrival and in front of the tagged customer. A customer at distance y of the server has a distance $x - y$ to the tagged customer, and hence generates a waiting time of $S(x - y)$. Remembering that $L(x)$ denotes the number of waiting customers within distance x of the server, let $f(x)$ denote the derivative of $\mathbb{E}[L(x)]$. Intuitively, $f(x)dx$ denotes the long-run expected number of customers with distances between x and $x + dx$ of the server. Then the mean of the maximal waiting time of a batch, conditional on furthest customer location and the batch size, is given by

$$\begin{aligned} \mathbb{E}[W^B | X^B = x, K = k] &= \mathbb{E}[T(x)] + \rho \mathbb{E}[S^R(x)] + (k - 1) \int_{y=0}^x \frac{1}{x} \mathbb{E}[S(y)] dy \\ &\quad + \int_{y=0}^x \mathbb{E}[S(x - y)] f(y) dy. \end{aligned} \quad (5)$$

Equation (5) reveals that, for a direct analysis of the waiting time of a job/batch, we require an expression for $f(x)$. Such an expression can be found by constructing an integral equation for $f(x)$, based on Little's law applied to intervals.

4.2 Step 2: An integral equation

In this subsection we construct an integral equation that $f(x)$ has to satisfy. The foundation of this subsection is an observation similar to Little's law. We relate the number of customers with distances further than $1 - x$ from the server to the waiting time of an arbitrary customer arriving in the range $[0, x]$.

Lemma 3 (Little's law). *Let $W(x)$ be the waiting time of an arbitrary customer arriving within distance x of the server. Then we have:*

$$\mathbb{E}[L] - \mathbb{E}[L(1 - x)] = \lambda \mathbb{E}[K] x \mathbb{E}[W(x)]. \quad (6)$$

Proof. Consider the case in which each customer obtains \$1 for each time unit it spends at a distance of at least $1 - x$ of the server. One could pay this out by giving each customer the dollar per time unit, i.e. each time unit you pay: $\mathbb{E}[L(1)] - \mathbb{E}[L(1 - x)]$ in expectation (observe that $L(1) = L$).

Equivalently one could pay out each customer leaving the interval $[1 - x, 1]$, where the customer thus receives a dollar for each time unit she spends in this interval: denoted as $\tau_{[1-x,1]}$. As the departure rate from this interval has to equal the arrival rate, due to the stability condition, this implies that each time unit you expect to pay out $\lambda \mathbb{E}[K] x$ customers.

Since both payout methods are equivalent we thus have:

$$\mathbb{E}[L] - \mathbb{E}[L(1 - x)] = \lambda \mathbb{E}[K] x \mathbb{E}[\tau_{[1-x,1]}].$$

Now consider a customer arriving at distance $1 - x + y$ and an imaginary customer arriving at y . Then the total time the first customer spends in the interval $[1 - x, 1]$ is equal to the total time it takes the server to reach point y , hence the waiting time of the imaginary customer. Therefore, due to the symmetry of the model, we have: $\mathbb{E}[\tau_{[1-x,1]}] = \mathbb{E}[W(x)]$. \square

Remark 4. *This observation is similar to that in Winands et al. [19, Equation (6)], regarding discrete polling models. The equivalence becomes especially apparent when one no longer focuses on customers at distance x of the server but rather on the customers at a certain fixed location.*

Lemmas 2 and 3 form the main ingredients for the integral equation that we construct in this subsection. We write the expected waiting time $\mathbb{E}[W(x)]$ as an integral by conditioning over the arrival location of the tagged customer, and by then taking the derivative w.r.t. x on both sides of (6) we find:

Proposition 4. *The function $f(x)$ satisfies the following integral equation:*

$$f(1 - x) = \lambda \mathbb{E}[K] \left\{ \frac{\alpha}{\rho} [\exp(\rho x) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho x) + \mathbb{E}[B] \frac{\mathbb{E}[K(K - 1)]}{\rho \mathbb{E}[K]} [\exp(\rho x) - 1] \right. \\ \left. + \int_{z=0}^x \mathbb{E}[B] \exp(\rho(x - z)) f(z) dz \right\}, \quad 0 \leq x < 1. \quad (7)$$

Proof. First we focus on $\mathbb{E}[W(x)]$ and use that the waiting time consists of the same four parts as before: (i) the waiting time generated by the traveling of the server to the customer; (ii) the waiting time generated by the possible residual service time; (iii) the total waiting time generated by the customers in the same batch who arrive in front of the customer; (iv) the waiting time generated

by customers already present on the circle at the moment of arrival. Conditioning on the exact location of the customer shows:

$$\begin{aligned} \mathbb{E}[W(x)] = \frac{1}{x} \int_{u=0}^x \left\{ \underbrace{\mathbb{E}[T(u)]}_{(i)} + \underbrace{\rho \mathbb{E}[S^R(u)]}_{(ii)} + \underbrace{\frac{\mathbb{E}[K(K-1)]}{\mathbb{E}[K]} \int_{z=0}^u \mathbb{E}[S(u-z)] dz}_{(iii)} \right. \\ \left. + \underbrace{\int_{z=0}^u \mathbb{E}[S(u-z)] f(z) dz}_{(iv)} \right\} du, \end{aligned}$$

and hence:

$$\begin{aligned} x \mathbb{E}[W(x)] = \int_{u=0}^x \left\{ \frac{\alpha}{\rho} [\exp(\rho u) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho u) + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho u) - 1] \right. \\ \left. + \int_{z=0}^u \mathbb{E}[B] \exp(\rho(u-z)) f(z) dz \right\} du. \end{aligned}$$

By writing the LHS of (6) as an integral over $f(u)$, i.e. $\mathbb{E}[L] - \mathbb{E}[L(1-x)] = \int_{u=1-x}^1 f(u) du$, we obtain the following equality:

$$\begin{aligned} \int_{u=1-x}^1 f(u) du = \lambda \mathbb{E}[K] \int_{u=0}^x \left\{ \frac{\alpha}{\rho} [\exp(\rho u) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho u) \right. \\ \left. + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho u) - 1] + \int_{z=0}^u \mathbb{E}[B] \exp(\rho(u-z)) f(z) dz \right\} du. \end{aligned} \quad (8)$$

Taking the derivative on both sides results in the integral equation:

$$\begin{aligned} f(1-x) = \lambda \mathbb{E}[K] \left\{ \frac{\alpha}{\rho} [\exp(\rho x) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho x) + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho x) - 1] \right. \\ \left. + \int_{z=0}^x \mathbb{E}[B] \exp(\rho(x-z)) f(z) dz \right\}. \quad \square \end{aligned} \quad (9)$$

Corollary 4.1. *Any twice-differentiable solution $f(x)$ to (7) satisfies $f''(x) = 0$.*

Proof. Taking the derivative on both sides of (7) results in:

$$\begin{aligned} -f'(1-x) = \rho \lambda \mathbb{E}[K] \left\{ \frac{\alpha}{\rho} [\exp(\rho x) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho x) + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho x) - 1] \right. \\ \left. + \int_{z=0}^x \mathbb{E}[B] \exp(\rho(x-z)) f(z) dz \right\} + \alpha \lambda \mathbb{E}[K] + \lambda \mathbb{E}[B] \mathbb{E}[K(K-1)] + \rho f(x). \end{aligned}$$

Here, the first part is $\rho f(1-x)$ by (9) and thus:

$$-f'(1-x) = \alpha \lambda \mathbb{E}[K] + \lambda \mathbb{E}[B] \mathbb{E}[K(K-1)] + \rho f(1-x) + \rho f(x).$$

Taking the derivative on both sides another time and using the above relation results in:

$$f''(1-x) = -\rho f'(1-x) + \rho f'(x) = \rho^2 f(1-x) + \rho^2 f(x) - \rho^2 f(x) - \rho^2 f(1-x) = 0. \quad \square$$

The corollary hints at a linear solution to the integral equation. This can also be argued intuitively. All customers at a distance x of the server need to have arrived since the last visit to said point. During this time, the server has traversed a total distance of $1 - x$. Due to the symmetry of the model, it makes sense that the time the server takes, on average, to traverse this distance is proportional to that distance.

4.3 Step 3: Solution of the integral equation

In this subsection we give a (linear) solution to the previously constructed integral equation with additional condition that the integral over $f(x)$ has to equal $\mathbb{E}[L]$, found in Lemma 1. We then prove that this is the unique solution and therefore that this solution gives the expected number of customers in a small interval around x . We then derive the expected batch sojourn time, see Theorem 7.

Proposition 5. *The function:*

$$f(x) = \rho \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2 \mathbb{E}[B]} + \frac{\alpha \lambda \mathbb{E}[K]}{1 - \rho} (1 - x) + \rho \frac{\lambda^2 \mathbb{E}[K]^2 \mathbb{E}[B^2]}{1 - \rho} (1 - x) + \frac{\rho \mathbb{E}[K(K - 1)]}{(1 - \rho) \mathbb{E}[K]} (1 - x), \quad (10)$$

is a solution to (7) with boundary condition:

$$\int_{x=0}^1 f(x) dx = \frac{\lambda \mathbb{E}[K]}{2(1 - \rho)} \left(\alpha + \lambda \mathbb{E}[K] \mathbb{E}[B^2] + \mathbb{E}[B] \frac{\mathbb{E}[K(K - 1)]}{\mathbb{E}[K]} \right). \quad (11)$$

Proof. It is readily verified that $f(\cdot)$ satisfies the boundary condition (11). It is thus left to prove that the expression in (10) also satisfies equation (7).

For the linearity of the solution to the integral equation, we substitute $f(1 - x) = ax + b$ in (7) and show that (10) gives a solution. By partial integration we have that:

$$\begin{aligned} \int_{z=0}^x \rho \exp(\rho(x - z)) f(z) dz &= \left[-\{a(1 - z) + b\} \cdot \exp(\rho(x - z)) \right]_{z=0}^x - a \int_{z=0}^x \exp(\rho(x - z)) dz \\ &= -a(1 - x) - b + (a + b) \exp(\rho x) + \frac{a}{\rho} [1 - \exp(\rho x)] \\ &= ax + \left(b - \frac{a(1 - \rho)}{\rho} \right) [\exp(\rho x) - 1]. \end{aligned}$$

Remark that ax cancels against the left-hand side in (7). Further, by taking $b = \lambda^2 \mathbb{E}[K]^2 \mathbb{E}[B^2]/2$, the only remaining terms in the right-hand side of (7) are $\exp(\rho x) - 1$ factors. These disappear when taking a as in (10), i.e.:

$$a = \frac{1}{1 - \rho} \left(\alpha \lambda \mathbb{E}[K] + \rho \lambda^2 \mathbb{E}[K]^2 \mathbb{E}[B^2] + \frac{\rho \mathbb{E}[K(K - 1)]}{\mathbb{E}[K]} \right). \quad \square$$

4.4 Step 4: uniqueness of the solution of the integral equation

In order to prove that the solution above indeed represents the expected number of customers in a small interval around x we prove that the integral equation with boundary conditions has a unique integrable solution.

Proposition 6. *The integral equation in (7) with boundary condition (11) has a unique integrable solution.*

Proof. We first use that the integral equation implies (8) by integrating from $1 - x$ to 1. Interchanging the order of integration now gives:

$$\begin{aligned}
\int_{u=1-x}^1 f(u)du &= \lambda \mathbb{E}[K] \int_{u=0}^x \left\{ \frac{\alpha}{\rho} [\exp(\rho u) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho u) \right. \\
&\quad \left. + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho u) - 1] \right\} du \\
&\quad + \int_{z=0}^x f(z) \int_{u=z}^x \rho \exp(\rho(u-z)) du dz \\
&= \lambda \mathbb{E}[K] \int_{u=0}^x \left\{ \frac{\alpha}{\rho} [\exp(\rho u) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho u) \right. \\
&\quad \left. + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho u) - 1] \right\} du \\
&\quad + \int_{z=0}^x f(z) [\exp(\rho(x-z)) - 1] dz.
\end{aligned}$$

Single out the integral $\int_{z=0}^x f(z) dz$ and take it to the left-hand side to conclude:

$$\begin{aligned}
\int_{u=1-x}^1 f(u)du + \int_{u=0}^x f(u)du &= \lambda \mathbb{E}[K] \int_{u=0}^x \left\{ \frac{\alpha}{\rho} [\exp(\rho u) - 1] + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \exp(\rho u) \right. \\
&\quad \left. + \mathbb{E}[B] \frac{\mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho u) - 1] \right\} du \\
&\quad + \int_{u=0}^x f(u) \exp(\rho(x-u)) du. \tag{12}
\end{aligned}$$

Now consider two functions f, g satisfying (7), and let $\|\cdot\|_\infty$ denote the infinity-norm (supremum norm). Since the first terms of (7) are not affected by the choice of f, g , we have:

$$\|f - g\|_\infty = \left\| \int_{u=0}^x \rho \exp(\rho(x-u)) (f(u) - g(u)) du \right\|_\infty.$$

It follows from (12) that we can rewrite this as follows:

$$\|f - g\|_\infty = \left\| \rho \int_{u=1-x}^1 (f(u) - g(u)) du + \rho \int_{u=0}^x (f(u) - g(u)) du \right\|_\infty.$$

If $x \geq 1/2$, then we can write the integral as the integral from 0 to 1 plus the part that is accounted for twice:

$$\begin{aligned}
\|f - g\|_\infty &= \left\| \rho \int_{u=0}^1 (f(u) - g(u)) du + \rho \int_{u=1-x}^x (f(u) - g(u)) du \right\|_\infty \\
&\leq \left\| \rho \cdot 0 \right\|_\infty + \rho \int_{u=1-x}^x \|f - g\|_\infty du \\
&\leq \rho(2x - 1) \cdot \|f - g\|_\infty \leq \rho \|f - g\|_\infty,
\end{aligned}$$

where the first integral disappeared due to the boundary condition in (11). When $x < 1/2$ we have:

$$\begin{aligned}
\|f - g\|_\infty &\leq \rho \left\| \int_{u=1-x}^1 (f(u) - g(u)) du \right\|_\infty + \rho \left\| \int_{u=0}^x (f(u) - g(u)) du \right\|_\infty \\
&\leq \rho \cdot 2x \cdot \|f - g\|_\infty \leq \rho \|f - g\|_\infty.
\end{aligned}$$

Since $\rho < 1$ this implies that $\|f - g\|_\infty = 0$, and hence $f \equiv g$. \square

4.5 Step 5: Determination of $\mathbb{E}[S^B]$

For the maximal waiting time of a batch we use a construction of the waiting time similar to the proof of Proposition 4. Conditioning on the size of the batch and location of the furthest customer we find:

Theorem 7. *The expected batch sojourn time satisfies:*

$$\begin{aligned}
\mathbb{E}[S^B] &= \mathbb{E}[B] + \frac{1}{1-\rho} \cdot \left[\alpha + \rho\lambda\mathbb{E}[K]\mathbb{E}[B^2] + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\mathbb{E}[K]} \right] \mathbb{E}\left[\frac{K}{K+1}\right] \\
&\quad + \frac{\lambda\mathbb{E}[K]\mathbb{E}[B^2]}{2} + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\rho\mathbb{E}[K]} + \frac{1}{\lambda} \left(\exp(\rho) - 1 \right) \\
&\quad - \left(\mathbb{E}[B] + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\rho\mathbb{E}[K]} \right) \exp(\rho) \\
&\quad + \left(\mathbb{E}[B]\rho + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\mathbb{E}[K]} \right) \int_{x=0}^1 \exp(\rho x) \tilde{K}(x) dx.
\end{aligned} \tag{13}$$

Proof. We use, again, that the batch sojourn time is given by the waiting time of the last customer in the batch, plus the service time of said customer. This latter term immediately translates to the first term in (7). For the waiting time of the last customer we recall (5):

$$\begin{aligned}
\mathbb{E}[W^B|X^B = x, K = k] &= \mathbb{E}[T(x)] + \rho\mathbb{E}[S^R(x)] + (k-1) \int_{y=0}^x \frac{1}{x} \mathbb{E}[S(y)] dy \\
&\quad + \int_{y=0}^x \mathbb{E}[S(x-y)] f(y) dy \\
&= \underbrace{\frac{\alpha}{\rho} [\exp(\rho x) - 1]}_{(i)} + \underbrace{\frac{\lambda\mathbb{E}[K]\mathbb{E}[B^2]}{2} \exp(\rho x)}_{(ii)} + \underbrace{\frac{(k-1)\mathbb{E}[B]}{\rho x} [\exp(\rho x) - 1]}_{(iii)} \\
&\quad + \underbrace{\int_{y=0}^x \mathbb{E}[B] \exp(\rho(x-y)) f(y) dy}_{(iv)}.
\end{aligned}$$

Applying the same partial integration as in the proof of Proposition 5, we find:

$$\begin{aligned}
(iv) &= -\frac{\alpha}{\rho} (\exp(\rho x) - 1) - \frac{\lambda\mathbb{E}[K]\mathbb{E}[B^2]}{2} \exp(\rho x) - \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\rho\mathbb{E}[K]} (\exp(\rho x) - 1) \\
&\quad + \frac{\alpha x}{1-\rho} + \frac{\lambda\mathbb{E}[K]\mathbb{E}[B^2]}{2} + \frac{\rho\lambda\mathbb{E}[K]\mathbb{E}[B^2]x}{1-\rho} + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]x}{\mathbb{E}[K](1-\rho)}.
\end{aligned}$$

Now (i) and (ii) cancel against the first elements and thus we are left with:

$$\begin{aligned}
\mathbb{E}[W^B|X^B = x, K = k] &= \frac{\alpha x}{1-\rho} + \frac{\lambda\mathbb{E}[K]\mathbb{E}[B^2]}{2} + \frac{\rho\lambda\mathbb{E}[K]\mathbb{E}[B^2]x}{1-\rho} + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]x}{\mathbb{E}[K](1-\rho)} \\
&\quad + \frac{(k-1)\mathbb{E}[B]}{\rho x} [\exp(\rho x) - 1] - \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\rho\mathbb{E}[K]} [\exp(\rho x) - 1].
\end{aligned}$$

We return to Equation (4) and decondition the expression above with respect to the batch size.

Rewriting everything in terms of the probability generating function of K gives:

$$\begin{aligned}
\mathbb{E}[W^B] &= \int_{x=0}^1 \sum_{k=1}^{\infty} kp_k x^{k-1} \cdot \left[\frac{\alpha x}{1-\rho} + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} + \frac{\rho \lambda \mathbb{E}[K] \mathbb{E}[B^2] x}{1-\rho} + \frac{\mathbb{E}[B] \mathbb{E}[K(K-1)] x}{\mathbb{E}[K](1-\rho)} \right. \\
&\quad \left. + \frac{(k-1) \mathbb{E}[B]}{\rho x} [\exp(\rho x) - 1] - \frac{\mathbb{E}[B] \mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho x) - 1] \right] dx \\
&= \int_{x=0}^1 \left[\frac{\alpha}{1-\rho} + \frac{\rho \lambda \mathbb{E}[K] \mathbb{E}[B^2]}{1-\rho} + \frac{\mathbb{E}[B] \mathbb{E}[K(K-1)]}{\mathbb{E}[K](1-\rho)} \right] \tilde{K}'(x) x dx \\
&\quad + \int_{x=0}^1 \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} \tilde{K}'(x) dx + \int_{x=0}^1 \frac{\mathbb{E}[B]}{\rho} [\exp(\rho x) - 1] \tilde{K}''(x) dx \\
&\quad - \int_{x=0}^1 \frac{\mathbb{E}[B] \mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} [\exp(\rho x) - 1] \tilde{K}'(x) dx.
\end{aligned}$$

We now apply partial integration to the first, third and last terms:

$$\begin{aligned}
\int_{x=0}^1 \tilde{K}'(x) x dx &= \left[x \tilde{K}(x) \right]_{x=0}^1 - \int_{x=0}^1 \tilde{K}(x) dx = \mathbb{E} \left[\frac{K}{K+1} \right]; \\
\int_{x=0}^1 \tilde{K}''(x) [\exp(\rho x) - 1] dx &= \left[[\exp(\rho x) - 1] \tilde{K}'(x) \right]_{x=0}^1 - \int_{x=0}^1 \rho \exp(\rho x) \tilde{K}'(x) dx \\
&= [\exp(\rho) - 1] \mathbb{E}[K] - \int_{x=0}^1 \rho \exp(\rho x) \tilde{K}'(x) dx; \\
\int_{x=0}^1 \tilde{K}'(x) \exp(\rho x) dx &= \exp(\rho) - \int_{x=0}^1 \rho \exp(\rho x) \tilde{K}(x) dx.
\end{aligned}$$

Substituting this in the expression for $\mathbb{E}[W^B]$ results in:

$$\begin{aligned}
\mathbb{E}[W^B] &= \left[\frac{\alpha}{1-\rho} + \frac{\rho \lambda \mathbb{E}[K] \mathbb{E}[B^2]}{1-\rho} + \frac{\mathbb{E}[B] \mathbb{E}[K(K-1)]}{\mathbb{E}[K](1-\rho)} \right] \mathbb{E} \left[\frac{K}{K+1} \right] \\
&\quad + \frac{\lambda \mathbb{E}[K] \mathbb{E}[B^2]}{2} + \frac{\mathbb{E}[B]}{\rho} \left(\exp(\rho) \mathbb{E}[K] - \mathbb{E}[K] - \rho \exp(\rho) + \rho \int_{x=0}^1 \rho \exp(\rho x) \tilde{K}(x) dx \right) \\
&\quad - \frac{\mathbb{E}[B] \mathbb{E}[K(K-1)]}{\rho \mathbb{E}[K]} \left(\exp(\rho) - 1 - \int_{x=0}^1 \rho \exp(\rho x) \tilde{K}(x) dx \right). \quad \square
\end{aligned}$$

Remark 5. Taking $\mathbb{E}[B] \rightarrow 0$, gives a total mean batch sojourn time of $\alpha \cdot \mathbb{E}[K/(K+1)]$. Note that this is exactly equal to the expected travel time of the server to the furthest customer.

Remark 6. The case $K \equiv 1$, i.e. unit batch sizes, results in:

$$\mathbb{E}[S^B] = \mathbb{E}[B] + \frac{\lambda \mathbb{E}[B^2]}{2(1-\rho)} + \frac{\alpha}{2(1-\rho)}.$$

This is in line with expectations as the system now simplifies to the system discussed in Kroese and Schmidt [10].

Remark 7. The batch sojourn time for light traffic, i.e. $\lambda \rightarrow 0$, can be found by reordering the

terms:

$$\begin{aligned} \mathbb{E}[S^B] &= \frac{1}{1-\rho} \cdot \left[\alpha + \rho\lambda\mathbb{E}[K]\mathbb{E}[B^2] + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\mathbb{E}[K]} \right] \mathbb{E}\left[\frac{K}{K+1}\right] \\ &\quad + \frac{\lambda\mathbb{E}[K]\mathbb{E}[B^2]}{2} + \frac{1}{\lambda}(\exp(\rho) - 1) \\ &\quad + \left(\rho\mathbb{E}[B] + \frac{\mathbb{E}[B]\mathbb{E}[K(K-1)]}{\mathbb{E}[K]} \right) \int_{x=0}^1 \exp(\rho x) [\tilde{K}(x) - 1] dx. \end{aligned}$$

Note that the last integral simplifies to $-\mathbb{E}[K/(K+1)]$ as $\rho \rightarrow 0$. Further realizing that $[\exp(\lambda\mathbb{E}[K]\mathbb{E}[B]) - 1]/\lambda \rightarrow \mathbb{E}[K]\mathbb{E}[B]$ as $\lambda \rightarrow 0$, then shows that:

$$\mathbb{E}[S^B] \rightarrow \alpha\mathbb{E}\left[\frac{K}{K+1}\right] + \mathbb{E}[K]\mathbb{E}[B],$$

i.e. in light traffic the mean batch sojourn time reduces to the sum of the mean travel time to the furthest customer plus the mean service time of a batch.

5 Numerical results

The results in this paper allow for a direct performance analysis and can be used to obtain direct insights in the performance effect of, for instance, the distribution of the batch size. Additionally, the given expression lends itself nicely for optimization. In previous work, van der Gaast et al. [16] present a mean value analysis for the discrete polling system with batch arrivals. They derive the expected waiting time of a batch by first solving a system of N^2 equations, in which N is the total number of queues. For symmetric systems, this is reduced to a system of N linear equations. For systems with large N this becomes computationally expensive, and in particular insights into the effect of e.g. the distribution of the batches, can no longer be directly obtained. In this section, we approximate a discrete (symmetric) polling system by its continuous variant and investigate the rate of convergence of the mean batch sojourn time as well as the effect of the batch size distribution.

We consider a symmetric discrete polling systems, consisting of N queues and fixed switch-over times. Additionally, we assume that each customer in a batch independently is assigned a queue with equal probability ($1/N$). For the approximation with the continuous variant we take α to be the total switch-over time between queues and B the generic service time in each separate queue. Moreover, the total size of each customer batch is taken for K .

First, we study the convergence of the mean batch sojourn time for different values of ρ and K . Consider the model with deterministic batch sizes and exponential service times (with mean 1). We consider four values of the workload, $\rho = 0.2, 0.45, 0.7, 0.95$. Figure 3 shows the effect of the batch size on the batch sojourn time and illustrates the rate of convergence of the discretized model to the continuous model. Logically, larger batch sizes result in bigger mean sojourn times due to the higher service requirement of a batch. Moreover, the continuous model already proves to be a good approximation for a relatively small number of queues: for $N = 10$ the difference is $< 2\%$ and for $N = 20$ the difference is $< 1\%$. The rate of the convergence seems to be much higher for lightly loaded systems.

Secondly, we consider different batch size distributions, see Figure 4. In each example, the service requirement of a job is on average 1 minute, where each service of a customer is assumed to take a deterministic time. Again, it is apparent that, for already small values of N , the continuous model provides a reasonable approximation for the batch sojourn time in discrete models, here the

difference is $< 6\%$ for $N = 10$ and $< 3\%$ for $N = 20$. The effect of the batch size distribution on the sojourn time appears to be closely linked to the variance of this distribution; distributions with higher variance result in bigger mean sojourn times. This is in line with results on $M^K/G/1$ queues, where the same relation to the variance of the bulk sizes holds.

Remark that the continuous system always results in larger mean batch sojourn times. This is caused by the fact that the distance from the server to the furthest customer in a batch is larger in expectation. In a light-traffic situation, i.e. $\lambda \rightarrow 0$, this becomes even more apparent as the batch sojourn time is then solely built up from the total service time of a batch and the travel time of the server, cf. Remark 7.

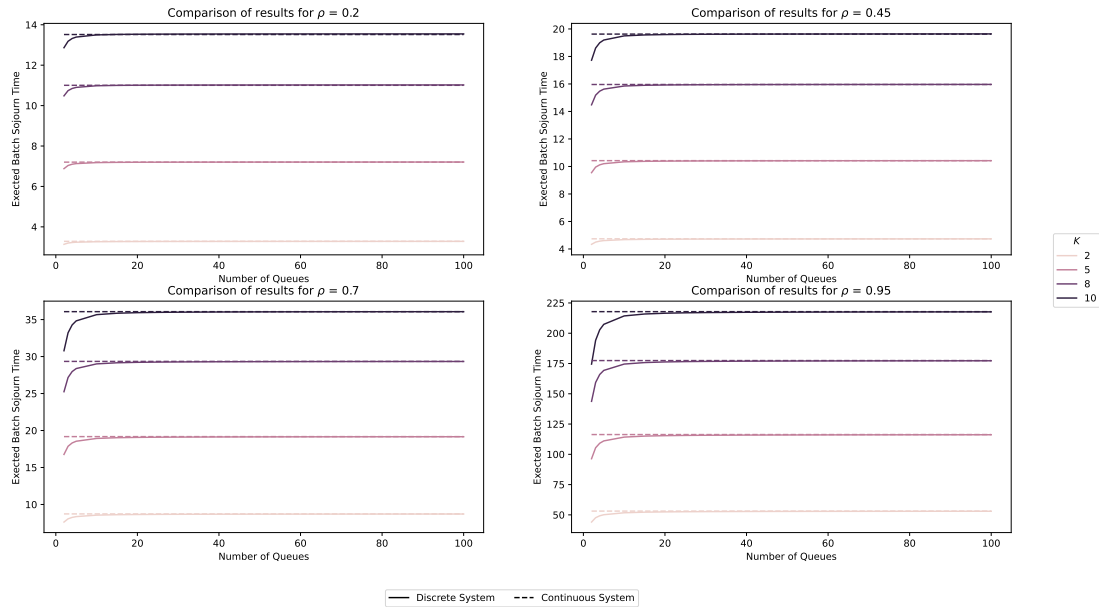


Figure 3: The expected batch sojourn time for deterministic batch sizes and exponential service requirements with unit mean, comparing the discrete and continuous polling model.

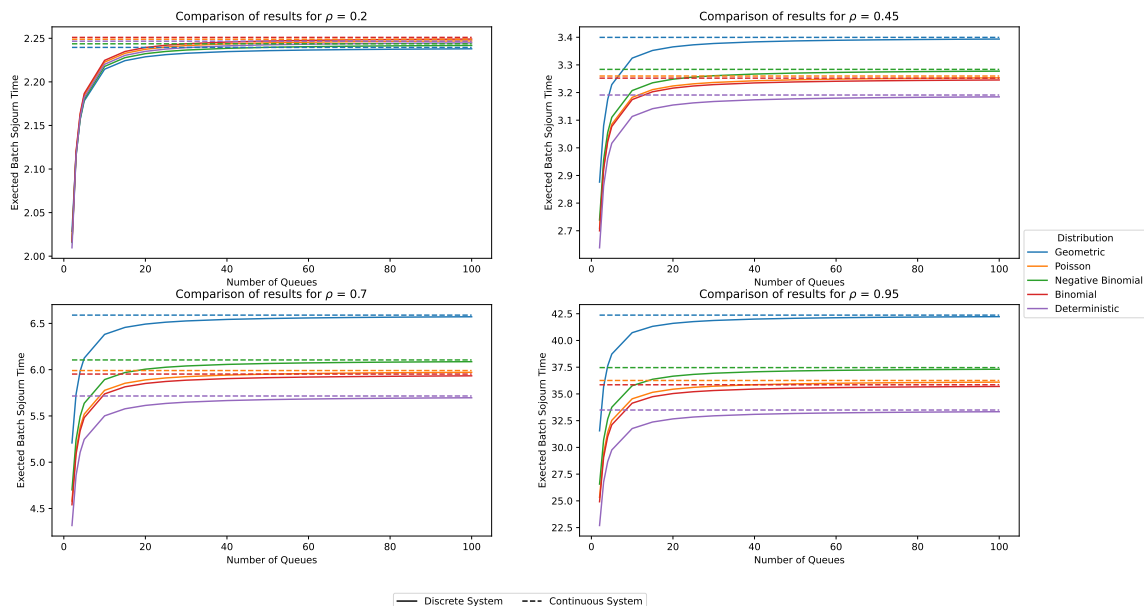


Figure 4: Comparison of the expected batch sojourn time for different batch size distributions. Services take $1/5$ time units and the average batch size was taken 5.

Acknowledgement

The research of Tim Engels and Onno Boxma is partly funded by the NWO Gravitation project NETWORKS, grant number 024.002.003.

References

- [1] A. Al Hanbali, R. de Haan, R. J. Boucherie, and J. K. van Ommeren. Time-limited polling systems with batch arrivals and phase-type service times. *Annals of Operations Research*, 198: 57–82, 2012.
- [2] M. Boon, R. van der Mei, and E. Winands. Applications of polling systems. *Surveys in Operations Research and Management Science*, 16:67–82, 2011.
- [3] S. Borst and O. Boxma. Polling: past, present, and perspective. *TOP*, 26:335–369, 2018. ISSN 18638279.
- [4] O. Boxma. Workloads and waiting times in single-server systems with multiple customer classes. *Queueing Systems*, 5:185–214, 1989.
- [5] E. Coffman and E. Gilbert. A continuous polling system with constant service times. *IEEE Transactions on Information Theory*, 32:584–591, 1986.
- [6] R. Cooper. *Introduction to Queueing Theory*. Macmillan, 1972.
- [7] I. Eliazar. The snowblower problem. *Queueing Systems*, 45:357–380, 2003.

- [8] I. Eliazar. From polling to snowplowing. *Queueing Systems*, 51:115–133, 2005.
- [9] D. Grillo. Polling mechanism models in communication systems-some application examples. In H. Takagi, editor, *Stochastic Analysis of Computer and Communication Systems*, pages 659–698. Elsevier Science Publishers BV (North-Holland), 1990.
- [10] D. Kroese and V. Schmidt. A continuous polling system with general service times. *The Annals of Applied Probability*, 2:906–927, 1992.
- [11] D. Kroese and V. Schmidt. Queueing systems on a circle. *Mathematical Methods of Operations Research*, 37:303–331, 1993.
- [12] H. Levy and M. Sidi. Polling systems : Applications, modeling, and optimization. *IEEE Transactions on Communications*, 38:1750–1760, 1990.
- [13] H. Levy and M. Sidi. Polling systems with simultaneous arrivals. *IEEE Transactions on Communications*, 39:823–827, 1991.
- [14] J. Resing. Polling systems and multitype branching processes. *Queueing Systems*, 13:409–426, 1993.
- [15] H. Takagi. Analysis and application of polling models. In *Performance Evaluation: Origins and Directions*, pages 423–442. Springer, 2001.
- [16] J. van der Gaast, I. Adan, and R. de Koster. The analysis of batch sojourn-times in polling systems. *Queueing Systems*, 85:313–335, 2017.
- [17] J. P. van der Gaast, R. B. de Koster, and I. J. Adan. Optimizing product allocation in a polling-based milkrun picking system. *IIE Transactions*, 51:486–500, 2019. ISSN 24725862. doi: 10.1080/24725854.2018.1493758.
- [18] R. van der Mei. Waiting-time distributions in polling systems with simultaneous batch arrivals. *Annals of Operations Research*, 113:155–173, 2002.
- [19] E. Winands, I. Adan, and G. van Houtum. Mean value analysis for polling systems. *Queueing Systems*, 54:35–44, 2006.