A two-station queue with dependent preparation and service times

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

A two-station queue with dependent preparation and service times

M. Vlasiou[*], I.J.B.F. Adan[**,***], and O.J. Boxma[**,***]

September 11, 2007

[*] Georgia Institute of Technology,
H. Milton Stewart School of Industrial & Systems Engineering,
765 Ferst Drive, Atlanta GA 30332-0205, USA.

[**] Eindhoven University of Technology,
Department of Mathematics & Computer Science,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

[***] EURANDOM,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

vlasiou@gatech.edu, iadan@win.tue.nl, boxma@win.tue.nl

Abstract

We discuss a single-server multi-station alternating queue where the preparation times and the service times are auto- and cross-correlated. We examine two cases. In the first case, preparation and service times depend on a common discrete time Markov chain. In the second case, we assume that the service times depend on the previous preparation time through their joint Laplace transform. The waiting time process is directly analysed by solving a Lindley-type equation via transform methods. Numerical examples are included to demonstrate the effect of the autocorrelation of and the cross-correlation between the preparation and service times.

1 Introduction

Consider a server alternating between two service points, serving one customer at a time. At each service point there is an infinite queue of customers waiting for service. Before being served by the server, a customer first must undergo a preparation phase, which starts immediately after the server has completed service at that particular service point and has moved to the next one. The server is not involved at all with the preparation phase. The server is obliged to alternate; therefore he serves all odd-numbered customers at one service point and all even-numbered customers at the other. Thus the server, after having finished serving a customer at one service point, may have to wait for the preparation phase of the customer at the other service point to be completed. Let W_n be the time the server has to wait before he can start serving the n-th customer. If B_n is the preparation time of the n-th customer and A_n is the service time of the n-th customer, then W_n can be defined recursively by

$$W_{n+1} = \max\{0, B_{n+1} - A_n - W_n\}, \quad n \geqslant 1. \tag{1.1}$$

1

This alternating service model occurs in many applications. For example, this strategy is followed by surgeons performing eye surgeries. Another example where (1.1) occurs comes from inventory theory, in particular from the analysis of two-carousel systems. This application is considered in [17, 20, 22]. Equation (1.1) is introduced in [22] and has been studied further in [29, 30, 31, 32, 33]; see these references for a further discussion on applications. A restriction of the previous work is that $\{A_n\}$ and $\{B_n\}$ are two mutually *independent* sequences of i.i.d. random variables. In the present study, we analyse the above queuing model under two different dependence structures.

In the first dependence structure we study, the distributions of the preparation and service times are regulated by an irreducible discrete-time Markov chain. Specifically, we assume that each transition of the Markov chain generates a new preparation time and its corresponding service time. Given the state of the Markov chain at times n and $n + 1$, the distributions of A_n and B_{n+1} are independent of one another for all n. However, the distributions of A_n and B_n depend on the state of the Markov chain. Such a dependence structure occurs naturally in many applications. For example, in the application involving two carousels that is described in [22, 30, 32], one can intuitively see that if an order consists of multiple items stored on one carousel, then there are strategies for the preparation of the carousel, where a long preparation time B_n implies that the service time A_n (i.e. the time necessary to pick an order on that carousel) will be relatively short, while being independent of all other past or future preparation and service times. Also in this example, the service and preparation times may depend on the carousel. That is, if one carousel is slower than the other, then the service and preparation times on the slower carousel will be longer than on the faster one. Hence, the service and preparation times are strictly periodic, which is a special case of Markov-dependence. Also queueing models with strictly periodic arrivals have been frequently studied; see, [11, 21, 23, 24]. These models arise, for example, in the modelling of inventory systems using periodic ordering policies; see [26] and [34].

The second dependence structure we study assumes that the random variables A_n and B_{n+1} have a joint distribution. In particular, given the length of the service time A_n, the following preparation time has a Laplace-Stieltjes transform of a specific form. The form we choose is rather general and allows for various specific dependence structures and preparation time distributions. Later on, we give a few specific examples. Dependencies between a service time and the *following* preparation time are also possible in applications. Again for the carousel model described in [30, 32] with orders consisting of multiple items, one can have the situation where a "smart" preparation strategy is followed, which anticipates the expected delay of the server for the previous order. Thus, knowing that the previous service time is relatively long, the other carousel rotates at a starting point that may be further away, but reduces the service time of the following order.

One should note here that (1.1) is, up to the minus sign in front of W_n, equal to Lindley's recursion, which is one of the most important and well-studied recursions in applied probability. Both dependence structures studied are also motivated by analogous cases studied for Lindley's recursion. In Lindley's recursion, A_n represents the interarrival time between customers n and $n+1$, B_n is the service time of the n-th customer, and W_n represents the time the n-th customer has to wait before starting his service; see, e.g., [2] and [10] for a comprehensive description. We shall compare Lindley's recursion to (1.1) for the two dependence structures mentioned above.

Queuing models with dependencies between interarrival and service time have been studied by several authors. A review of the early literature can be found in Bhat [4]. Such dependencies arise naturally in various applications. For example, the phenomenon of dependence among the interarrival times in the packet streams of voice and data traffic is well known; see, e.g., [18, 19, 27]. However, in [16] the authors argue that in packet communication networks one should

also expect two additional forms of dependence: between successive service times and among interarrival times and service times. These forms of dependence occur because of the presence of bursty arrivals and multiple sources with different mean service times (due to different packet lengths), and they may have a dominant effect on waiting times and queue lengths. In this paper, we study such forms of dependence for the model described by (1.1).

The presentation is organised as follows. In Section 2 we derive the steady-state waiting-time distribution F_W for the first case studied. We assume that the service times follow some general distribution that depends on the state of the Markov chain. For the preparation times, in Section 2.1 we assume that they are exponentially distributed (with a rate depending on the state of the Markov chain), while in Section 2.2 we extend the analysis to mixed-Erlang distributions. We complement the results with various numerical examples in Section 2.3. In Section 3 we derive F_W for the second case studied. We assume that the service times are exponentially distributed, although, as we remark later on, the analysis can be extended to mixed-Erlang distributions.

2 Markov-modulated dependencies

In this section we define the setting and derive the autocorrelation function of preparation or service times, as well as the crosscorrelation of these. We study the case where the preparation times and the service times depend on a common discrete-time Markov chain. This model allows dependencies between preparation and service times. The waiting time in this case is directly derived by using Laplace transforms. For Lindley's recursion, the analogous model has been analysed in Adan and Kulkarni [1].

We first introduce some notation. For a random variable Y we denote its distribution by F_Y and its density by f_Y. A second index is used when we need to distinguish between the different states of the underlying Markov chain. For example, we denote by $F_{W,j}(x)$ the steady-state probability that the waiting time $W \leqslant x$ and the Markov chain is in state j. Also, given an event E we use the convention that $\mathbb{P}[Y \leqslant x\,;E] = \mathbb{E}[\mathbb{1}_{[Y \leqslant x]} \cdot \mathbb{1}_{[E]}]$, and likewise for expectations.

As mentioned before, we assume that the sequences $\{A_n\}$ and $\{B_n\}$ are both autocorrelated and cross-correlated. The nature of the dependence we study in this first part is described below.

The distributions of the preparation and service times are regulated by an irreducible, aperiodic, discrete-time Markov chain $\{Z_n\}$, $n \geqslant 1$, with state space $\{1, 2, \ldots, M\}$ and transition probability matrix $\mathbf{P} = (p_{i,j})$. To be exact, we have that

$$
\begin{aligned}
\mathbb{P}[A_n \leqslant x\,; &B_{n+1} \leqslant y\,; Z_{n+1} = j \mid Z_n = i\,; B_n\,; (A_\ell, B_\ell, Z_\ell), 1 \leqslant \ell \leqslant n-1] \\
&= \mathbb{P}[A_1 \leqslant x\,; B_2 \leqslant y\,; Z_2 = j \mid Z_1 = i] \\
&= p_{i,j}\mathbb{P}[A_1 \leqslant x\,; B_2 \leqslant y \mid Z_1 = i\,; Z_2 = j] \\
&= p_{i,j}F_{A|i}(x)F_{B|j}(y),
\end{aligned}
\tag{2.1}
$$

where $x, y \geqslant 0$ and where $i, j = 1, 2, \ldots, M$. Thus, given Z_n and Z_{n+1}, the distributions of A_n and B_{n+1} are independent of one another for all n. The random variables A_n follow an arbitrary distribution that is independent of the past, given Z_n, while B_n follows in general a phase-type distribution that is depending on the state of Z_n.

We assume that there are an i and j such that $\mathbb{P}[X_n < 0, Z_{n+1} = j | Z_n = i] > 0$, where $X_n = B_{n+1} - A_n$. This implies that the Markov chain (W_n, Z_n) is stable. To see that, first notice that a regeneration point occurs if $W_{n+1} = 0$ and $Z_{n+1} = j$. Since the Markov chain $\{Z_n\}$ is finite and irreducible, the time between two occurrences of state j is finite in expectation. For each time state j is reached there is a positive probability that $X_n < 0$, which implies that

$W_{n+1} = 0$. Thus, we will have that in a geometric number of steps both events $W_{n+1} = 0$ and $Z_{n+1} = j$ will happen at the same step. In particular, the number of steps is finite in expectation. In other words, we will have a regeneration point within a finite expected time. Hence, since $\{Z_n\}$ is aperiodic, the limiting distribution of the Markov chain (W_n, Z_n) exists, and thus we can define for $\mathrm{Re}(s) \geqslant 0$, $n \geqslant 1$, and $j = 1, 2, \ldots, M$ the transforms

$$\omega_j(s) = \lim_{n \to \infty} \omega_j^n(s),$$

where

$$\omega_j^n(s) = \mathbb{E}[\mathrm{e}^{-sW_n} ; Z_n = j].$$

Denote by λ_i^{-1} the mean and by s_i the second moment of the service time distribution $F_{A|i}$. Analogously, define μ_i^{-1} as the mean of $F_{B|i}$ and σ_i as its second moment. Moreover, denote by $\boldsymbol{\varpi} = (\varpi_1, \varpi_2, \ldots, \varpi_M)$ the stationary distribution of the Markov chain $\{Z_n\}$. Then, in steady state, the autocorrelation between A_m and A_{m+n} is given by

$$\rho[A_m, A_{m+n}] = \rho[A_1, A_{n+1}] = \frac{\sum_{i=1}^{M} \sum_{j=1}^{M} \varpi_i \big(p_{i,j}^{(n)} - \varpi_j\big) \lambda_i^{-1} \lambda_j^{-1}}{\sum_{i=1}^{M} \varpi_i s_i - \left(\sum_{i=1}^{M} \varpi_i \lambda_i^{-1}\right)^2}, \tag{2.2}$$

where

$$p_{i,j}^{(n)} = \mathbb{P}[Z_{n+1} = j \mid Z_1 = i], \qquad n \geqslant 0, \quad 1 \leqslant i, j \leqslant M.$$

A similar expression can be derived for the autocorrelation between preparation times. Since $\mathbf{P}$ is aperiodic, we have that $p_{i,j}^{(n)}$ converges to ϖ_j geometrically as n tends to infinity. In other words, the autocorrelation function approaches zero geometrically fast as the lag goes to infinity. For the cross-correlation between A_n and B_n we have that

$$\rho[A_n, B_n] = \rho[A_1, B_1] = \frac{\sum_{i=1}^{M} \varpi_i \lambda_i^{-1} \mu_i^{-1} - \widehat{\mu}\,\widehat{\lambda}}{\left(\sum_{i=1}^{M} \varpi_i s_i - \widehat{\lambda}^2\right)^{1/2} \left(\sum_{j=1}^{M} \varpi_j \sigma_i - \widehat{\mu}^2\right)^{1/2}},$$

where $\widehat{\lambda} = \sum_{i=1}^{M} \varpi_i \lambda_i^{-1}$ and $\widehat{\mu} = \sum_{i=1}^{M} \varpi_i \mu_i^{-1}$.

When $\{Z_n\}$ is in state j, we denote by α_j the Laplace-Stieltjes transform of the service distribution $F_{A|j}$. We denote the derivative of order i of a function f by $f^{(i)}$ and we have by definition that $f^{(0)} = f$. For simplicity, in the following section we first derive the steady-state waiting-time distribution for this model in case the preparation times are exponentially distributed. Later on, in Section 2.2, we generalise this result to phase-type preparation times.

2.1 Exponential preparation times

In this section we assume that $F_{B|j}(x) = 1 - \mathrm{e}^{-\mu_j x}$. We are interested in the steady-state waiting-time distribution. The next theorem gives the equations satisfied by the waiting-time densities for every state of $\{Z_n\}$.

Theorem 1 (Exponential preparation times)**.** *Let* $F_{B|j}(x) = 1 - \mathrm{e}^{-\mu_j x}$. *Then, for all* $j = 1, \ldots, M$,

$$F_{W,j}(0) = \varpi_j - c_j,$$

where $c_j = \sum_{i=1}^{M} p_{i,j}\, \omega_i(\mu_j)\, \alpha_i(\mu_j)$, *and the waiting-time density is given by*

$$f_{W,j}(x) = \mu_j c_j\, \mathrm{e}^{-\mu_j x}.$$

The M^2 unknown constants $\omega_i(\mu_j)$ are the unique solution to the following system of linear equations

$$\omega_j(\mu_\ell) = \varpi_j - \frac{\mu_\ell}{\mu_j + \mu_\ell} \sum_{i=1}^{M} p_{i,j}\, \omega_i(\mu_j)\, \alpha_i(\mu_j), \qquad j, \ell = 1, \ldots, M.$$

Proof. From (1.1) we obtain the following equation for the transforms ω_j^{n+1}, $j = 1, \ldots, M$.

$$\omega_j^{n+1}(s) = \mathbb{E}[\mathrm{e}^{-sW_{n+1}}\,;\, Z_{n+1} = j] = \sum_{i=1}^{M} \mathbb{P}[Z_n = i]\mathbb{E}[\mathrm{e}^{-s\max\{0, B_{n+1} - A_n - W_n\}}\,;\, Z_{n+1} = j \mid Z_n = i]$$

$$= \sum_{i=1}^{M} \mathbb{P}[Z_n = i]p_{i,j}\Big(\mathbb{E}[\int_0^{A_n+W_n} f_{B_{n+1}}(x)\,\mathrm{d}x \mid Z_n = i\,;\, Z_{n+1} = j] +$$

$$+ \mathbb{E}[\int_{A_n+W_n}^{\infty} \mathrm{e}^{-s(x - A_n - W_n)} f_{B_{n+1}}(x)\,\mathrm{d}x \mid Z_n = i\,;\, Z_{n+1} = j] \Big). \tag{2.3}$$

Since $Z_{n+1} = j$, we have that B_{n+1} is now exponentially distributed with rate μ_j. Thus, the above equation becomes

$$\omega_j^{n+1}(s) = \sum_{i=1}^{M} \mathbb{P}[Z_n = i]p_{i,j}\Big(\mathbb{E}[\int_0^{A_n+W_n} \mu_j\mathrm{e}^{-\mu_j x}\,\mathrm{d}x \mid Z_n = i] +$$

$$+ \mathbb{E}[\int_{A_n+W_n}^{\infty} \mathrm{e}^{-s(x - A_n - W_n)} \mu_j\mathrm{e}^{-\mu_j x}\,\mathrm{d}x \mid Z_n = i] \Big)$$

$$= \sum_{i=1}^{M} \mathbb{P}[Z_n = i]p_{i,j}\mathbb{E}[1 - \mathrm{e}^{-\mu_j(A_n+W_n)} + \frac{\mu_j}{\mu_j + s}\mathrm{e}^{-\mu_j(A_n+W_n)} \mid Z_n = i]$$

$$= \sum_{i=1}^{M} \mathbb{P}[Z_n = i]p_{i,j}\Big(1 - \frac{s}{\mu_j + s}\mathbb{E}[\mathrm{e}^{-\mu_j(A_n+W_n)} \mid Z_n = i]\Big)$$

$$= \sum_{i=1}^{M} p_{i,j}\Big(\mathbb{P}[Z_n = i] - \frac{s}{\mu_j + s}\omega_i^n(\mu_j)\, \alpha_i(\mu_j)\Big).$$

So for $n \to \infty$ we have that $\omega_j(s)$ is given by

$$\omega_j(s) = \varpi_j - \sum_{i=1}^{M} p_{i,j}\, \omega_i(\mu_j)\, \alpha_i(\mu_j) + \frac{\mu_j}{\mu_j + s}\sum_{i=1}^{M} p_{i,j}\, \omega_i(\mu_j)\, \alpha_i(\mu_j). \tag{2.4}$$

Define the constants

$$c_j = \sum_{i=1}^{M} p_{i,j}\, \omega_i(\mu_j)\, \alpha_i(\mu_j).$$

Inverting the Laplace transform ω_j yields that the density of the waiting time is given by

$$f_{W,j}(x) = \mu_j c_j\, \mathrm{e}^{-\mu_j x},$$

and the corresponding distribution has mass $F_{W,j}(0) = \varpi_j - c_j$ at the origin. For $i, j = 1, \ldots, M$, the M^2 unknown constants $\omega_i(\mu_j)$ that are needed in order to determine the unknown constants c_j are the unique solution to the system of linear equations given by the expression

$$\omega_j(\mu_\ell) = \varpi_j - \frac{\mu_\ell}{\mu_j + \mu_\ell}\sum_{i=1}^{M} p_{i,j}\, \omega_i(\mu_j)\, \alpha_i(\mu_j), \qquad j, \ell = 1, \ldots, M; \tag{2.5}$$

see Equation (2.4). The uniqueness of the solution follows from the general theory of Markov chains, which states that there is a unique stationary distribution and thus also a unique solution to the system of equations formed by (2.5) for all $j, \ell = 1, \ldots, M$. $\qquad \square$

The result given in Theorem 1 is expected. Evidently, since the preparation time B_n is exponentially distributed (with a rate depending on the state of the Markov chain) and W_n is the residual preparation time, we have that for every state j of the Markov chain, the waiting-time distribution has mass at zero and the conditional waiting time is exponentially distributed with rate μ_j.

Observe that Theorem 1 generalises the statement of Theorem 1 in [30], which gives the steady-state waiting-time density in case $\{A_n\}$ and $\{B_n\}$ are mutually independent sequences of i.i.d. random variables and B follows an Erlang distribution. Specifically, if the Markov chain in Theorem 1 has only one state (and thus there is a unique service-time distribution and a unique rate μ for the exponentially distributed preparation times) and the Erlang distribution F_B in [30, Theorem 1] has only one phase, then the statements of these two theorems are identical. Observe, for example, that (2.5) reduces to [30, Eq. (3.2)] as now $\varpi_j = 1$, and $p_{i,j} = 1$.

2.2 Phase-type preparation times

Assume now that if the Markov chain is in state j, the preparation time is with probability κ_n equal to a random variable Y_n, $n = 1, \ldots, N$, that follows an Erlang distribution with parameter μ_j and n phases. In other words the distribution function of B is given by

$$F_{B|j}(x) = \sum_{n=1}^{N} \kappa_n \left(1 - e^{-\mu_j x} \sum_{\ell=0}^{n-1} \frac{(\mu_j x)^\ell}{\ell!} \right), \qquad x \geqslant 0. \tag{2.6}$$

This class of phase-type distributions may be used to approximate any given distribution on $[0, \infty)$ for the preparation times arbitrarily close; see Schassberger [25]. The waiting-time density for this case is given by the following theorem.

Theorem 2 (Mixed-Erlang preparation times)**.** *Let $F_{B|j}$ be given by* (2.6)*. Then, for $j = 1, \ldots, M$,*

$$F_{W,j}(0) = \varpi_j - \sum_{i=1}^{M} \sum_{n=1}^{N} \sum_{\ell=0}^{n-1} \sum_{m=0}^{\ell} \kappa_n p_{i,j} \frac{\mu_j^\ell}{\ell!} \binom{\ell}{m} (-1)^\ell \big(\alpha_i(\mu_j)\big)^{(\ell-m)} \big(\omega_i(\mu_j)\big)^{(m)} \tag{2.7}$$

and the waiting time density is given by

$$f_{W,j}(x) = \sum_{i=1}^{M} \sum_{n=1}^{N} \sum_{\ell=0}^{n-1} \sum_{m=0}^{\ell} \kappa_n p_{i,j} \frac{(-1)^\ell}{\ell!} \binom{\ell}{m} \big(\alpha_i(\mu_j)\big)^{(\ell-m)} \big(\omega_i(\mu_j)\big)^{(m)} \mu_j^n e^{-\mu_j x} \frac{x^{n-\ell-1}}{(n-\ell-1)!}.$$

Proof. For the proof, we shall refrain from presenting detailed computations, as the analysis is straightforward and similar to the one for the exponential case. We give, however, a few intermediate formulas. From (2.3) and for the present preparation time distributions we have that

$$\omega_j^{n+1}(s) = \sum_{i=1}^{M} \mathbb{P}[Z_n = i] p_{i,j} \left(1 - \sum_{n=1}^{N} \kappa_n \mathbb{E}[e^{-\mu_j(A_n + W_n)} \sum_{\ell=0}^{n-1} \frac{\mu_j^\ell (A_n + W_n)^\ell}{\ell!} \mid Z_n = i] + \right.$$

$$\left. + \sum_{n=1}^{N} \kappa_n \left(\frac{\mu_j}{\mu_j + s} \right)^n \mathbb{E}[e^{-\mu_j(A_n + W_n)} \sum_{\ell=0}^{n-1} \frac{(\mu_j + s)^\ell (A_n + W_n)^\ell}{\ell!} \mid Z_n = i] \right).$$

So for $n \to \infty$ we have that $\omega_j(s)$ is given by (cf. (2.4))

$$\omega_j(s) = u_j + \sum_{i=1}^{M} \sum_{n=1}^{N} \sum_{\ell=0}^{n-1} \sum_{m=0}^{\ell} \kappa_n p_{i,j} \frac{(-\mu_j)^\ell}{\ell!} \binom{\ell}{m} \left(\alpha_i(\mu_j)\right)^{(\ell-m)} \left(\omega_i(\mu_j)\right)^{(m)} \left(\frac{\mu_j}{\mu_j + s}\right)^{n-\ell}, \quad (2.8)$$

where u_j is the mass of the waiting time distribution at the origin and it is given by (2.7). Inverting the Laplace transform ω_j yields that the density of the waiting time is given by the expression presented in the theorem and the corresponding distribution has mass u_j at the origin. For $i, j = 1, \ldots, M$, the M^2 unknown constants $\omega_i(\mu_j)$ are the unique solution to the system of linear equations resulting from substituting μ_k, $k = 1, \ldots, M$, for s in (2.8). $\qquad \square$

Naturally, Theorem 2 reduces to Theorem 1 for $N = 1$, and to the i.i.d. case for $M = 1$; cf. [30, Theorem 2].

Dependence structures of the form of Equation (2.1), and several generalisations, have been studied extensively for Lindley's recursion. Asmussen and Kella [3] considered multidimensional martingales for a class of processes, and sketched the results for the Markov-modulated M/G/1 queue. The basic model was then described in detail in Adan and Kulkarni [1], where all results are given explicitly and the analysis extends to the study of the queue length distribution. Specifically, the authors study a single-server queue where the interarrival times and the service times depend on a common discrete-time Markov chain in a way similar to the one described by Equation (2.1). This model generalises the well-known MAP/G/1 queue by allowing dependencies between interarrival and service times. The MAP/G/1 queue provides a powerful framework to model dependencies between successive interarrival times [12], but typically the service times are assumed to be i.i.d. and independent of the arrival process.

Remark 1 (Periodic Markov chain). The results in Sections 2.1 and 2.2 are also valid in case the transition probability matrix $\mathbf{P}$ is periodic. Clearly, the transforms ω_j should then be defined as Cesaro limits, i.e., for $\text{Re}(s) \geqslant 0$, $n \geqslant 1$, and $j = 1, 2, \ldots, M$,

$$\omega_j(s) = \lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} \omega_j^m(s).$$

But contrary to the aperiodic case, the autocorrelation between service times and preparation times does not converge to zero as the lag tends to infinity; see (2.2).

2.3 Numerical examples

In this section we present some examples to demonstrate the effects of autocorrelation and crosscorrelation of the preparation and service times. In all of the examples, we have assumed that the underlying Markov chain has four states, i.e., $M = 4$. The transition matrices are chosen so that the limiting distributions for all cases are equal. Naturally, different transition matrices produce different dependence structures between and among the sequences $\{A_n\}$ and $\{B_n\}$. We also assume that the service times and the preparation times are exponentially distributed and that the service rates in each state of the Markov chain are given by $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (1, 100, 1, 100)$. We keep the mix of small and large preparation and service rates the same. Thus the examples differ only in the dependence structure among and between the sequences of preparation and service times.

First example. In this example we study the effect of autocorrelation in case $\{A_n\}$ and $\{B_n\}$ are positively crosscorrelated. In particular, we assume that the preparation rates are given by $(\mu_1, \mu_2, \mu_3, \mu_4) = (u/2, 10u, u/2, 10u)$, where we employ the parameter $u > 0$ in order to explore
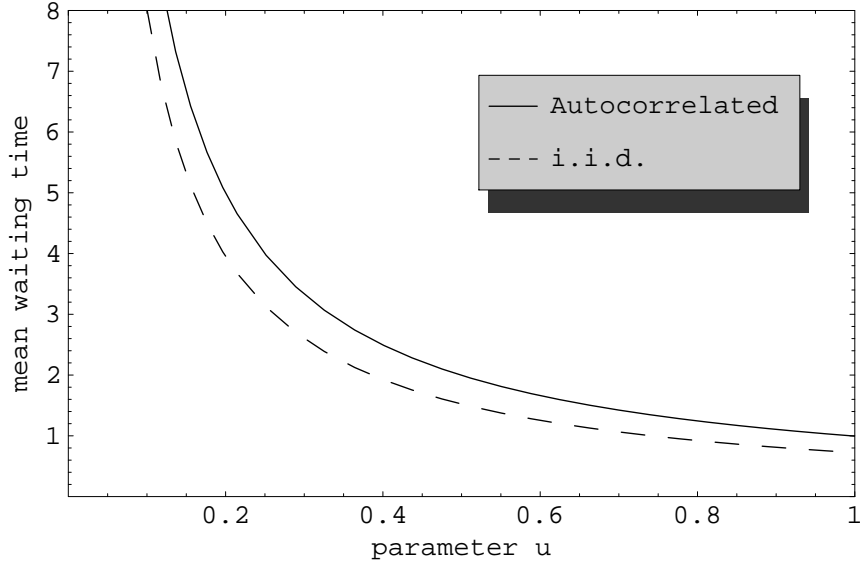
Figure 1: Positive crosscorrelation: the mean waiting time against the parameter controlling the preparation rate for autocorrelated and i.i.d. sequences.

the effect of the mean preparation time on the expected waiting time. For this setting, we have that the crosscorrelation between $\{A_n\}$ and $\{B_n\}$ is approximately equal to 0.3195.

We will compare the case where both the preparation and the service times are autocorrelated to the case that successive preparation and service times are i.i.d. In the first case, we take the transition probability matrix to be given by

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{2.9}$$

Thus, since we assumed that the service times are exponentially distributed (with the given rates), we have that their autocorrelation is given by

$$\rho[A_1, A_{n+1}] = (-1)^n \frac{9801}{29803}, \qquad n \geqslant 1, \tag{2.10}$$

and similarly, we can derive that the autocorrelation for the preparation times is equal to

$$\rho[B_1, B_{n+1}] = (-1)^n \frac{361}{1163}, \qquad n \geqslant 1.$$

In the second case, we take $\mathbf{P} = (0.25)$, i.e. we take all elements of the matrix $\mathbf{P}$ equal to 0.25, which implies that successive preparation and successive service times have zero autocorrelation. In Figure 1 we plot the mean waiting time of the server against the parameter u. Keep in mind that all means of the preparation times decrease as u increases. As is evident from Figure 1, autocorrelation leads to higher mean waiting times.

Second example. In Figure 2 we study the effect of autocorrelation in case $\{A_n\}$ and $\{B_n\}$ are *negatively* crosscorrelated. For this case, we take the preparation rates equal to
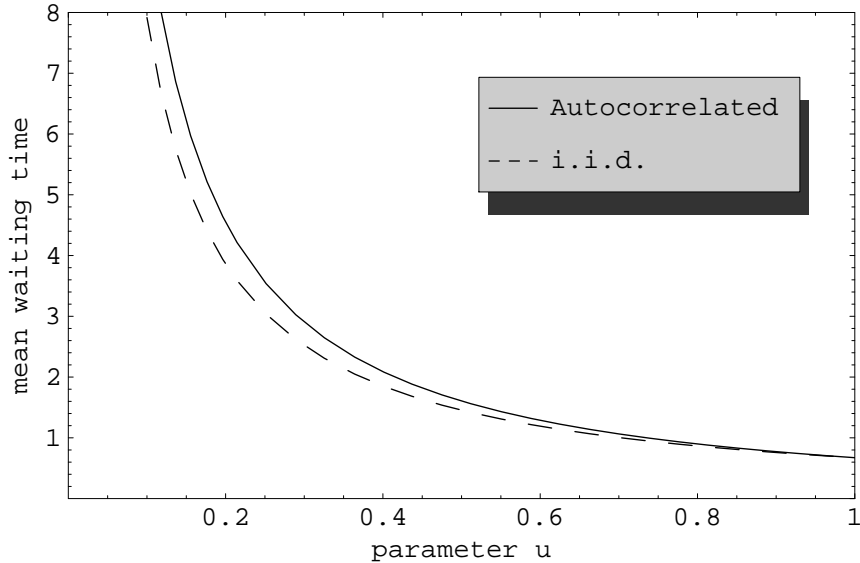
8

Figure 2: Negative crosscorrelation: the mean waiting time against the parameter controlling the preparation rate for autocorrelated and i.i.d. sequences.

$(\mu_1, \mu_2, \mu_3, \mu_4) = (10u, u/2, 10u, u/2)$, $u > 0$, which implies that the crosscorrelation in this example is approximately equal to -0.3195. We compare the previous two cases for the new preparation rates. Namely, we compare the autocorrelated case, where the transition matrix $\mathbf{P}$ is given by (2.9), to the i.i.d. case, where the transition matrix has all its entries equal to 0.25. The autocorrelation functions, both for the preparation times and for the service times remain the same. As before, autocorrelation leads to higher mean waiting times. However, in this case, the effect of the autocorrelation is not as big as in the first example.

Third example. Now the preparation rates are given by $(\mu_1, \mu_2, \mu_3, \mu_4) = (u/2, u/2, 10u, 10u)$, $u > 0$, which implies that the crosscorrelation in this example is equal to zero. As before, we compare two cases. In the first case, we take

$$\mathbf{P} = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix}$$

which leads to autocorrelated service times with the autocorrelation function given by Equation (2.10), and i.i.d. preparation times. We compare this case to the case where all entries of the transition matrix are equal to 0.25, where successive preparation and successive service times have zero autocorrelation. In Figure 3, the mean waiting times are shown for both cases. The difference in the mean waiting times is negligible in this case, although the two functions are not equal.

Comparing these numerical results to the analogous cases for Lindley's recursion, which are described in [1], illustrates the different effect of dependencies on these two recursions. In Lindley's recursion, autocorrelation leads to slightly lower mean waiting times. Moreover, negative crosscorrelation leads to a bigger difference between the autocorrelated and the i.i.d. case than the difference observed for the positively crosscorrelated case. The results for the
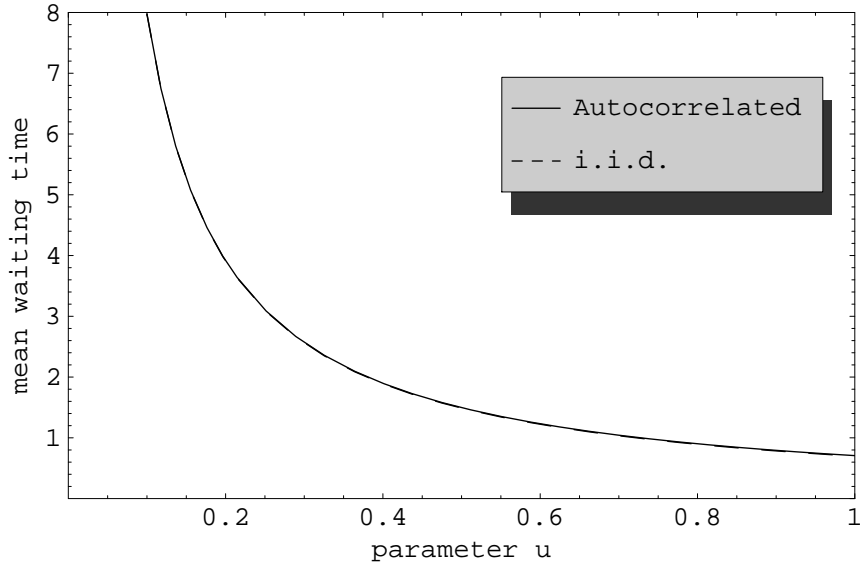
Figure 3: Zero crosscorrelation: the mean waiting time against the parameter controlling the preparation rate for autocorrelated and i.i.d. sequences.

Lindley-type recursion (1.1) are reversed. Autocorrelation leads to higher mean waiting times and the biggest differences are observed in case $\{A_n\}$ and $\{B_n\}$ are positively crosscorrelated (cf. Figures 1 and 2). Also, in the third example, where the crosscorrelation is eliminated, the outcome is again very different. For Lindley's recursion, even in the case of zero crosscorrelation, autocorrelation leads to significant differences for the mean waiting times. In our case, there is hardly any difference between the autocorrelated and the i.i.d. case (cf. Figure 3).

3   Services depending on the previous preparation time

In this section, we study the second dependence structure mentioned in the introduction. In particular, we assume that for all n, the service times A_n are distributed as A, which in turn is exponentially distributed with rate λ. Moreover, the Laplace-Stieltjes transform of the preparation time B_{n+1}, given that the previous service time A_n equals t, is of the form

$$\mathbb{E}[\mathrm{e}^{-sB_{n+1}} \mid A_n = t] = \mathbb{E}[\mathrm{e}^{-sB} \mid A = t] = \chi(s)\mathrm{e}^{-\psi(s)t}. \tag{3.1}$$

Observe that now the preparation time B_{n+1} depends only on the previous service time, while in the Markov-modulated case we have examined previously all preparation and service times are correlated between and among one another, since their distributions depend on a common Markov chain.

This dependence structure also occurs in simple queuing models. Consider the following situation. Work arrives at a single-server queue according to a process with stationary, non-negative independent increments. This work, however, does not immediately enter the queue of the server facility; instead it is accumulated behind a gate. At exponential interarrivals the gate is opened and – after the addition of an independent component – the work is collected and delivered as a single customer at the queue of the service facility. The additional component may be viewed as a set-up time.

Due to the exponentially distributed interarrival times of customers, we can view the service facility as an M/G/1 queue in which the interarrival and service time for each customer are positively correlated. Indeed, if the interval between two consecutive openings of the gate is relatively long (short), it is likely that a relatively large (small) amount of work has accumulated during that interval. This model is a unification and generalisation of the M/G/1 queue with a positive correlation between interarrival and service times [5, 6, 9, 14] and has been analysed in Boxma and Combé [7]. In Combé and Boxma [12] it is shown that the collect system can also be modelled by using the BMAP framework.

3.1 The waiting-time distribution

In order to derive the waiting-time distribution we shall further assume that the functions χ and ψ that appear in Equation (3.1) are rational functions; i.e.,

$$\chi(s) = \frac{P_1(s)}{Q_1(s)} \qquad \text{and} \qquad \psi(s) = \frac{P_2(s)}{Q_2(s)}, \tag{3.2}$$

where P_2, Q_1, and Q_2 are polynomials of degrees L, M, and N respectively. From the form of Equation (3.1) we see that a number of other assumptions have been implicitly made. For example, since for $s = 0$ the expectation $\mathbb{E}[e^{-sB} \mid A = t]$ should be equal to one, we have implicitly assumed that $\psi(0) = 0$ and $\chi(0) = 1$. We shall mention other implications of such type only when necessary.

The preparation time B_{n+1} consists of two parts: a component which depends on the previous service time, represented by $e^{-\psi(s)t}$, and an 'ordinary' preparation time with Laplace-Stieltjes transform $\chi(s)$, which does not depend on the service time. From (3.1) we have that the bivariate Laplace-Stieltjes transform of the generic preparation and service time is given by

$$\mathbb{E}[e^{-sB-zA}] = \int_0^\infty \lambda e^{-\lambda t} e^{-zt} \chi(s) e^{-\psi(s)t} \, \mathrm{d}t = \frac{\lambda \chi(s)}{\lambda + \psi(s) + z}, \tag{3.3}$$

for $\mathrm{Re}(\lambda + \psi(s) + z) > 0$. This expression leads to

$$\mathbb{E}[B] = \frac{\psi'(0)}{\lambda} - \chi'(0), \quad \text{and} \quad \mathbb{E}[AB] = \frac{2\psi'(0) - \lambda\chi'(0)}{\lambda^2},$$

from which we have that the covariance between a preparation time and a service time is given by

$$\mathrm{cov}[A, B] = \frac{\psi'(0)}{\lambda^2}.$$

The correlation between these two variables can be also computed, see Boxma and Combé [7]. Thus, one can construct a distribution function F_B that has any covariance (or correlation) between A and B that is desired.

In order to derive the steady-state waiting-time distribution, we shall first derive the Laplace-Stieltjes transform of F_W. We follow a method based on Wiener-Hopf decomposition. A straightforward calculation, starting from the steady-state version of (1.1),

$$W \overset{\mathcal{D}}{=} \max\{0, B - A - W\},$$

yields that

$$\omega(s) = \mathbb{E}[e^{-sW}] = \mathbb{P}[B \leqslant W + A] + \mathbb{E}[e^{-s(B-W-A)}] - \mathbb{E}[e^{-s(B-W-A)}; B \leqslant W + A]$$
$$= \mathbb{P}[B \leqslant W + A] + \mathbb{E}[e^{sW}]\,\mathbb{E}[e^{-s(B-A)}] - \mathbb{E}[e^{-s(B-W-A)}; B \leqslant W + A],$$

11

since $B - A$ and W are independent. Therefore, from (3.3) we have for $\text{Re}(s) = 0$ that

$$\omega(s) = \mathbb{P}[B \leqslant W + A] + \omega(-s)\frac{\lambda\chi(s)}{\lambda - s + \psi(s)} - \mathbb{E}[e^{-s(B-W-A)} \mid B \leqslant W + A]\mathbb{P}[B \leqslant W + A]$$

(3.4)

$$= \omega(-s)\frac{P_1(s)}{Q_1(s)}\frac{\lambda Q_2(s)}{(\lambda - s)Q_2(s) + P_2(s)} + \mathbb{P}[B \leqslant W + A]\big(1 - \mathbb{E}[e^{-s(B-W-A)} \mid B \leqslant W + A]\big),$$

which can be rewritten as

$$\omega(s)Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big) = \lambda\omega(-s)P_1(s)Q_2(s) + Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big)\times$$
$$\times \mathbb{P}[B \leqslant W + A]\big(1 - \mathbb{E}[e^{-s(B-W-A)} \mid B \leqslant W + A]\big). \quad (3.5)$$

We can observe that $Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big)$ is a polynomial of degree $K = \max\{M + N + 1, M + L\}$ and also that the left-hand side of (3.5) is analytic for $\text{Re}(s) > 0$ and continuous for $\text{Re}(s) \geqslant 0$, and the right-hand side of (3.5) is analytic for $\text{Re}(s) < 0$ and continuous for $\text{Re}(s) \leqslant 0$. Therefore, from Liouville's theorem [28], we conclude that both sides of (3.5) are the same K-th degree polynomial, say, $\sum_{i=0}^{K} q_i s^i$. Hence,

$$\omega(s) = \frac{\sum_{i=0}^{K} q_i s^i}{Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big)}. \quad (3.6)$$

In the expression above, the constants q_i are not determined so far. In order to obtain the transform, observe that ω is a fraction of two polynomials both of degree K. Let r_i, $i = 1, \ldots, K$, be the zeros of the denominator. Ignoring the special case of zeros with multiplicity greater than one, partial fraction decomposition yields that (3.6) can be rewritten as

$$\omega(s) = c_0 + \sum_{i=1}^{K} \frac{c_i}{s - r_i}, \quad (3.7)$$

which implies that the waiting-time distribution has a mass at the origin that is given by

$$\mathbb{P}[W = 0] = \lim_{s \to \infty} \mathbb{E}[e^{-sW}] = c_0$$

and has a density that is given by

$$f_W(x) = \sum_{i=1}^{K} c_i e^{r_i x}. \quad (3.8)$$

All that remains is to determine the $K + 1$ constants c_i. To do so, we work as follows. We express the terms $\mathbb{P}[B \leqslant W + A]$ and $\mathbb{E}[e^{-s(B-W-A)} \mid B \leqslant W + A]$ that appear at the right-hand side of (3.5) in terms of the constants c_i by using (3.7) and (3.8). Then, we substitute these expressions and (3.7) in the left-hand side of (3.5). Thus we obtain a new equation in terms of the constants c_i that we differentiate a total of K times. We evaluate each of these derivatives for $s = 0$ and thus we obtain a linear system of $K + 1$ equations, $i = 0, \ldots, K$, for the constants c_i (the original equation and the K derivatives). We summarise the above in the following theorem.

Theorem 3. *Let $F_A(x) = 1 - e^{-\lambda x}$. Under the assumptions* (3.1) *and* (3.2)*, we have that*

$$F_W(0) = c_0$$

12

and the waiting time density is given by

$$f_W(x) = \sum_{i=1}^{K} c_i \mathrm{e}^{r_i x}.$$

In the expression above, the constants r_i are the K zeros of the equation

$$Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big) = 0$$

and the coefficients c_i are derived as described above.

Remark 2. In case that for some i we have that the real part of r_i is greater than or equal to zero, then it follows that the corresponding coefficient c_i is equal to zero. This is a consequence of the fact that the limiting distribution exists and therefore the density f_W should be integrable. Note that the meaning of χ as the Laplace-Stieltjes transform of an ordinary preparation time already implies that the L zeroes of $Q_1(s)$ have a negative real part.

Remark 3. Although the roots r_i and coefficients c_i may be complex, the density and the mass c_0 at zero will be positive. This follows from the fact that there is a unique equilibrium distribution and thus a unique solution to the linear system for the coefficients c_i. Of course, it is also clear that each root r_i and coefficient c_i have a companion conjugate root and conjugate coefficient, which implies that the imaginary parts appearing in the density cancel.

Remark 4. When $Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big)$ has multiple zeros, the analysis proceeds in essentially the same way. For example, if $r_1 = r_2$, then the partial fraction decomposition of ω becomes

$$\omega(s) = c_0 + \frac{c_1}{(s - r_1)^2} + \sum_{i=2}^{K} \frac{c_i}{s - r_i},$$

yielding

$$f_W(x) = c_1 x \mathrm{e}^{r_1 x} + \sum_{i=2}^{K} c_i \mathrm{e}^{r_i x}.$$

Remark 5. For the service time A we have considered only the exponential distribution, mainly because we can illustrate the technique we use without complicating the analysis. However, we can extend this class by considering distributions with a mixed-Erlang distribution of the form of Equation (2.6) and the proof remains essentially the same. That is, let $F_A(x) = \sum_{i=1}^{n} \kappa_i E_i(x)$, where E_i is the Erlang distribution with i phases. Then, the resulting density of the waiting time is again of the form

$$f_W(x) = \sum_{i=1}^{K'} c_i \mathrm{e}^{r_i x},$$

where $K' = M + n \max\{N + 1, L\}$. The constants r_i are the zeros to the equation

$$Q_1(s)\big((\lambda - s)Q_2(s) + P_2(s)\big)^n = 0,$$

and the coefficients c_i are determined in a similar fashion as before. Naturally, if any of the zeros r_i has multiplicity greater than one, the form of the density changes analogously; see also the previous remark.

We now present a few examples where we show how some classic dependence structures fit into this class. In the examples below, we only discuss how to derive the function ψ. All of the examples presented here are examples of Lévy processes.

Independent case

The dependence structure described by Equation (3.1) contains a great variety of dependence structures, including the independent case. If for all s we have that $\psi(s) = 0$, then the Laplace-Stieltjes transform of B is independent of the length of the service time, and thus the function χ appearing in (3.1) is in fact the Laplace-Stieltjes transform of B. Observe that we have assumed that B has a rational Laplace-Stieltjes transform, which is necessary in order to decompose Equation (3.4) into functions that are analytic either in the left-half plane or in the right-half plane.

Linear Dependence

Assume that the service time A and the preparation time B are linearly dependent; that is, $B = cA$. Then,
$$\mathbb{E}[\mathrm{e}^{-sB} \mid A = t] = \mathbb{E}[\mathrm{e}^{-scA} \mid A = t] = \mathrm{e}^{-sct}.$$

Thus we have that $\chi(s) = 1$, and $\psi(s) = cs$, and both functions satisfy our assumptions. Queuing models with a linear dependence between the service time and the preceding interarrival time have been studied in [9, 13, 15]; see also [8].

Compound Poisson Process

In this case we assume that given that $A = t$, the preparation time B is equal to $\sum_{i=1}^{N(t)} C_i$, where $N(t)$ is a Poisson process with rate γ, and $\{C_i\}$ is a sequence of i.i.d. random variables, where each of them is distributed like C, and where C has a rational Laplace-Stieltjes transform. By convention, if $N(t) = 0$, we have that the preparation time is also zero. Under this assumption, we have

$$\mathbb{E}[\mathrm{e}^{-sB} \mid A = t] = \mathbb{E}[\mathrm{e}^{-s \sum_{i=1}^{N(t)} C_i} \mid A = t] = \sum_{k=0}^{\infty} \mathbb{E}[\mathrm{e}^{-s \sum_{i=1}^{k} C_i} \mid A = t] \mathrm{e}^{-\gamma t} \frac{(\gamma t)^k}{k!}$$

$$= \sum_{k=0}^{\infty} \big(\mathbb{E}[\mathrm{e}^{-sC}]\big)^k \mathrm{e}^{-\gamma t} \frac{(\gamma t)^k}{k!} = \mathrm{e}^{-\psi(s)t},$$

where $\psi(s) = \gamma\big(1 - \mathbb{E}[\mathrm{e}^{-sC}]\big)$. As before, in this case we have that $\chi(s) = 1$.

Brownian Motion

In this case we assume that given that $A = t$, the random variable B is normally distributed with mean μt and variance $\sigma^2 t$. Then we have that

$$\mathbb{E}[\mathrm{e}^{-sB} \mid A = t] = \int_{-\infty}^{\infty} \mathrm{e}^{-sx} \frac{\mathrm{e}^{-(x-\mu t)^2/(2\sigma^2 t)}}{\sigma\sqrt{2\pi t}} \, \mathrm{d}x = \mathrm{e}^{-\psi(s)t},$$

where $\psi(s) = \mu s - s^2\sigma^2/2$, and $\chi(s) = 1$. Of course, if B is interpreted as the preparation time of a customer, then assuming that B is normally distributed is not a natural assumption, since the preparation time of a customer is non-negative. However, in the analysis we do not need the condition of B being non-negative; therefore, it is mathematically possible to consider this case. For further examples of distributions satisfying the condition described by (3.1), see Boxma and Combé [7].

One additional assumption made in [7] is that the function ψ appearing in (3.1) has a completely monotone increasing derivative. This implies that the function $e^{-\psi(s)}$ is the Laplace-Stieltjes transform of an infinitely divisible probability distribution, which in its turn is a distribution of increments in processes with stationary independent increments, i.e. Lévy processes. However, the monotonicity of the derivative of ψ was not required for the analysis of our model.

3.2   Numerical examples

We shall demonstrate the effect of the dependence structure given by (3.1) to the mean waiting time of the server. In Figure 4 we plot the mean waiting time against the mean preparation time of a customer for the three first cases mentioned in the previous section. In all cases, we assume that the service times are exponentially distributed with mean 1.

As an example, we present the derivations for the linear dependence. In this case, we have that $K = 1$ and thus we have one root, given by $r_1 = -1/(c-1)$. Moreover, one can compute that

$$\mathbb{P}[B \leqslant W + A] = \begin{cases} 1, & c \leqslant 1 \\ 1 - c_0 + \frac{c_1}{2r_1}, & c > 1 \end{cases}$$

and that

$$\mathbb{E}[e^{-s(B-W-A)} \mid B \leqslant W + A] = \begin{cases} \frac{\omega(-s)}{1-(1-c)s}, & c \leqslant 1 \\ \frac{1}{1+(c-1)s} \frac{\omega(-s)-\omega(1/(c-1))}{1-c_0+\frac{c_1}{2r_1}}, & c > 1. \end{cases}$$

Thus, for $c \leqslant 1$, Equation (3.5) becomes

$$(c_0 + \frac{c_1}{s-r_1})(1-s+cs) = c_0 - \frac{c_1}{s+r_1} + (1-s+cs)\big(1 - \frac{c_0 - \frac{c_1}{s+r_1}}{1-(1-c)s}\big).$$

Evaluating this equation and its derivative at $s = 0$ we form a linear system of two equations, from which we conclude that for $c \leqslant 1$, $c_0 = 1$ and $c_1 = 0$. This is what was to be expected, as in this case there is no waiting time.

For $c > 1$, Equation (3.5) becomes

$$(c_0 + \frac{c_1}{s-r_1})(1-s+cs) = c_0 - \frac{c_1}{s+r_1} + (1-s+cs)(1-c_0+\frac{c_1}{2r_1})\big(1 - \frac{\frac{c_1}{2r_1} - \frac{c_1}{s+r_1}}{(1+(c-1)s)(1-c_0+\frac{c_1}{2r_1})}\big).$$

Evaluating this equation and its derivative at $s = 0$ we have that for $c > 1$, $c_0 = 1/3$ and $c_1 = 2/(3(c-1))$. Since the mean waiting time is equal to c_1/r_1^2, we have that $\mathbb{E}[W]$ is zero for $c \leqslant 1$ and is equal to $2(c-1)/3$ for $c > 1$. This case is depicted in Figure 4.

Furthermore, for the independent case we have assumed that the preparation times are also exponentially distributed, and for the compound Poisson process we have assumed that the jumps C_i are exponentially distributed with mean 1 (which implies that the mean preparation times are equal to the parameter γ of the underlying Poisson process).

As is evident from Figure 4, the precise nature of the dependence structure between the preparation and the service times has a significant impact on the mean waiting time (and on the waiting time distribution). In Section 2.3, where we have examined the effect of Markov modulated dependencies, correlation has led to longer waiting times. However, in this case, we observe that dependent preparation and service times do not necessarily lead to longer waiting times. In case B is linearly dependent on A, the resulting mean waiting time is always smaller than the mean waiting time for A and B being both exponentially distributed and independent of one another.
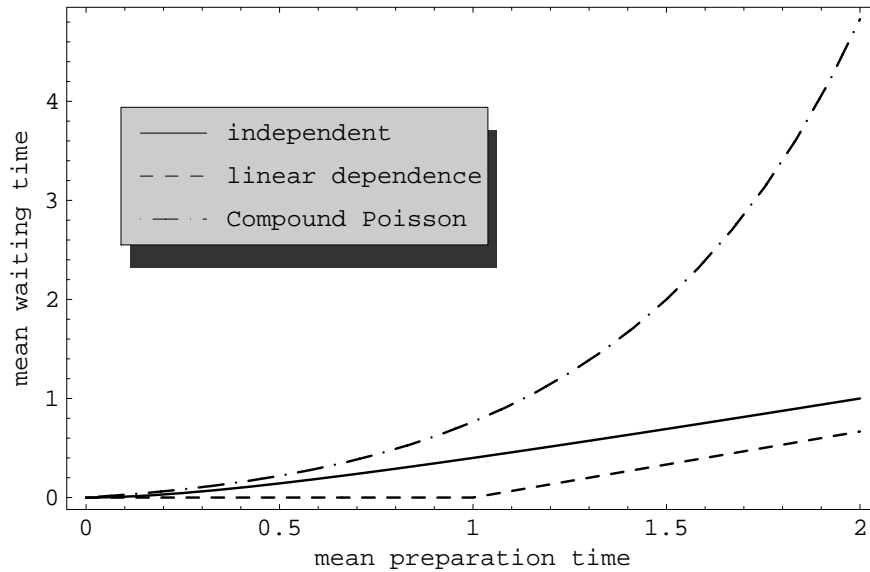
Figure 4: The mean waiting time against the mean preparation time

We further observe that the mean waiting time can vary significantly among various dependence structures. For example, we see that for the compound Poisson case, the mean waiting time grows faster than an exponential (which is the growth rate for the independent case) and diverges from the independent case faster than the mean waiting time for the linear case. One can infer from this example that the dependence among preparation and service times in Equation (1.1) cannot be ignored and should be modelled appropriately.

References

[1] I. J.-B. F. Adan and V. G. Kulkarni. Single-server queue with Markov-dependent inter-arrival and service times. *Queueing Systems. Theory and Applications*, 45(2):113–134, October 2003.

[2] S. Asmussen. *Applied Probability and Queues*. Springer-Verlag, New York, 2003.

[3] S. Asmussen and O. Kella. A multi-dimensional martingale for Markov additive processes and its applications. *Advances in Applied Probability*, 32(2):376–393, June 2000.

[4] U. N. Bhat. Queueing systems with first-order dependence. *Opsearch. The Journal of the Operational Research Society of India*, 6:1–24, 1969.

[5] S. C. Borst, O. J. Boxma, and M. B. Combé. Collection of customers: A correlated M/G/1 queue. *Performance Evaluation Review*, 20(1):47–59, June 1992.

[6] S. C. Borst, O. J. Boxma, and M. B. Combé. An M/G/1 queue with customer collection. *Communications in Statistics. Stochastic Models*, 9(3):341–371, 1993.

[7] O. J. Boxma and M. B. Combé. The correlated M/G/1 queue. *Archiv für Elektronik und Übertragungstechnik*, 47(5/6):330–335, 1993.

[8] O. J. Boxma and D. Perry. A queueing model with dependence between service and interarrival times. *European Journal of Operational Research*, 128(3):611–624, February 2001.

[9] I. Cidon, R. Guérin, A. Khamisy, and M. Sidi. On queues with interarrival times proportional to service times. *Probability in the Engineering and Informational Sciences*, 10(1):87–107, 1996.

[10] J. W. Cohen. *The Single Server Queue*. North-Holland Publishing Co., Amsterdam, 1982.

[11] J. W. Cohen. On periodic Pollaczek waiting time processes. In *Athens Conference on Applied Probability and Time Series Analysis, Vol. I (1995)*, volume 114 of *Lecture Notes in Statistics*, pages 361–378. Springer, New York, 1996.

[12] M. B. Combé and O. J. Boxma. BMAP modelling of a correlated queue. In J. Walrand, K. Bagchi, and G. W. Zobrist, editors, *Network Performance Modeling and Simulation*, pages 177–196. CRC, first edition, April 1998.

[13] B. W. Conolly. The waiting time process for a certain correlated queue. *Operations Research*, 16(5):1006–1015, September 1968.

[14] B. W. Conolly and Q. H. Choo. The waiting time process for a generalized correlated queue with exponential demand and service. *SIAM Journal on Applied Mathematics*, 37(2):263–275, October 1979.

[15] B. W. Conolly and N. Hadidi. A correlated queue. *Journal of Applied Probability*, 6(1):122–136, April 1969.

[16] K. W. Fendick, V. R. Saksena, and W. Whitt. Dependence in packet queues. *IEEE Transactions on Communications*, 37(11):1173–1183, 1989.

[17] E. Hassini and R. G. Vickson. A two-carousel storage location problem. *Computers & Operations Research*, 30(4):527–539, April 2003.

[18] H. Heffes. A class of data traffic processes – covariance function characterization and related queueing results. *The Bell System Technical Journal*, 59(6):897–929, July–August 1980.

[19] H. Heffes and D. M. Lucantoni. A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE Journal on Selected Areas in Communications*, 4(6):856–868, September 1986. Special Issue on Congestion Control in High-Speed Packet Switched Networks.

[20] E. Koenigsberg. Analysis of the efficiency of carousel and tote-stacker performance. In J. White, editor, *Proceedings of the 7th International Conference on Automation in Warehousing*, pages 173–183, San Francisco, California, October 1986. Springer.

[21] A. J. Lemoine. Waiting time and workload in queues with periodic Poisson input. *Journal of Applied Probability*, 26(2):390–397, June 1989.

[22] B. C. Park, J. Y. Park, and R. D. Foley. Carousel system performance. *Journal of Applied Probability*, 40(3):602–612, September 2003.

[23] T. Rolski. Approximation of periodic queues. *Advances in Applied Probability*, 19(3):691–707, September 1987.

[24] T. Rolski. Relationships between characteristics in periodic Poisson queues. *Queueing Systems. Theory and Applications*, 4(1):17–26, March 1989.

[25] R. Schassberger. *Warteschlangen.* Springer-Verlag, Wien, 1973.

[26] S. R. Smits, M. Wagner, and A. G. de Kok. Determination of an order-up-to policy in the stochastic economic lot scheduling model. *International Journal of Production Economics*, 90(3):377–389, August 2004.

[27] K. Sriram and W. Whitt. Characterizing superposition arrival processes in packet multiplexers for voice and data. *IEEE Journal on Selected Areas in Communications*, 4:833–846, 1986.

[28] E. C. Titchmarsh. *Theory of Functions.* Oxford University Press, London, second edition, 1968.

[29] M. Vlasiou. A non-increasing Lindley-type equation. *Queueing Systems. Theory and Applications*, 56(1):41–52, May 2007.

[30] M. Vlasiou and I. J.-B. F. Adan. An alternating service problem. *Probability in the Engineering and Informational Sciences*, 19(4):409–426, October 2005.

[31] M. Vlasiou and I. J.-B. F. Adan. Exact solution to a Lindley-type equation on a bounded support. *Operations Research Letters*, 35(1):105–113, January 2007.

[32] M. Vlasiou, I. J.-B. F. Adan, and J. Wessels. A Lindley-type equation arising from a carousel problem. *Journal of Applied Probability*, 41(4):1171–1181, December 2004.

[33] M. Vlasiou and B. Zwart. Time-dependent behaviour of an alternating service queue. *Stochastic Models*, 23(2):235–263, 2007.

[34] M. Wagner and S. R. Smits. A local search algorithm for the optimization of the stochastic economic lot scheduling problem. *International Journal of Production Economics*, 90(3):391–402, August 2004.