



HAL
open science

Analysis of the infinite server queues with semi-Markovian multivariate discounted inputs

Landy Rabehasaina, Jae-Kyung Woo

► **To cite this version:**

Landy Rabehasaina, Jae-Kyung Woo. Analysis of the infinite server queues with semi-Markovian multivariate discounted inputs. *Queueing Systems*, 2020, 94 (3), pp.393-420. hal-01894005v2

HAL Id: hal-01894005

<https://hal.science/hal-01894005v2>

Submitted on 19 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analysis of the infinite server queues with semi-Markovian multivariate discounted inputs

Landy Rabehasaina and Jae-Kyung Woo

*Laboratory of Mathematics, University Bourgogne Franche Comté,
16 route de Gray, 25030 Besançon cedex, France.*

e-mail: lrabehas@univ-fcomte.fr

*School of Risk and Actuarial Studies, Australian School of Business,
University of New South Wales, Australia.*

e-mail: j.k.woo@unsw.edu.au

Abstract: We consider a general k dimensional discounted infinite server queues process (alternatively, an Incurred But Not Reported (IBNR) claim process) where the multivariate inputs (claims) are given by a k dimensional finite state Markov Chain and the arrivals follow a renewal process. After deriving a multidimensional integral equation for the moment generating function jointly to the state of the input at time t given the initial state of the input at time 0, asymptotic results for the first and second (matrix) moments of the process are provided. In particular, when the interarrival or service times are exponentially distributed, transient expressions for the first two moments are obtained. Also, the moment generating function for the process with deterministic interarrival times is considered to provide more explicit expressions. Finally, we demonstrate the potential of the present model by showing how it allows us to study a semi-Markov modulated infinite server queues where the customers (claims) arrival and service (reporting delay) times depend on the state of the process immediately before and on the switching times.

Keywords and phrases: Semi-Markovian multivariate discounted inputs, Infinite server queues, IBNR process, Markov modulation.

1. Introduction

In the context of an infinite server queue with correlated batch arrivals, the total number of customers still in the system is related to an aggregation of correlated risks (multivariate risks) where the arrival times of those risks are adjusted by adding a random delay. Without this time delay, probability modeling of aggregate risk processes has been studied in various areas such as applied probability, reliability theory, and actuarial science. In particular, the research on an aggregation of correlated risks is striving to develop techniques to estimate the combined effect of different types of risks on the infrastructure or system.

In an infinite server queue with correlated batch arrivals, the random batch size is multivariate and the service time distribution is dependent on the type of input (queue). A similar idea but with multiple Markovian batch arrival streams can be found in [10] where a time-dependent matrix joint generating function of the number of customers in the system was derived. In a renewal process with correlated batch arrivals, [16] provides the transient expressions for the joint moments of the number of customers in an infinite server queue which are recursively obtainable. Then, [12] develop asymptotic approximation methods to study these

joint moments and also provided some queueing theoretic applications, including the workload of the queue and infinite server queues in tandem.

It is worthwhile to mention that research on the actuarial application to the Incurred But Not Reported (IBNR) claims exists in much the same way as an analysis of the infinite server queues with batch arrivals, see e.g. [14, 15, 16]. As discussed in [8], the service times in an infinite server queue can be interpreted as the time lag between the occurrence of a claim and the report of that claim in the IBNR model. In this respect, the proposed model in the present paper can be utilized to analyze different quantities of interest in (at least) two areas including queueing theory and actuarial science.

The present paper considers the model which is an extension of the one given in [12]. In each batch arrival, the model consists of multivariate queues (claims) which are modeled by some finite Markov Chain, and a renewal process is assumed for batch arrivals. The Markovian assumption for a vector of queues (claims) enable us to study the infinite server queues (IBNR process) in more realistic situations such that arrival times and service times are dependent on the states of the external semi-Markovian process immediately before and on the switching times, as will be illustrated later in Section 6. It is natural to model that the arrival process and service time are modulated by some external process, in particular, when this process impacts on the intensity of claim arrival processes and in turn, the type of service times. For example, the number of multiple types of claims in catastrophe insurance varies depending on the environmental condition and also it could lead to different types of reporting/settlement time delays. In addition, this model provides the capability to incorporate the information of successive states of the external background process before and on the switching times, which accommodates a link between consecutive states around the jump times. Moreover, when the batch sizes are regarded as claim amounts a discount factor introduced in the model is certainly important to study a present value of total IBNR claim amounts. Also, as shown in [12] it allows us to analyze the workload and covariance of the workload and queue size, and reinterpret Little's law in the presence of a positive discount factor.

To the best of our knowledge, there is no study of the current setting of the model (especially in the presence of a discount factor) in the literature of queueing theory or actuarial science. Instead, similar settings of the model such as Markov modulated infinite server systems are found. For example, in [4] the particle arrives according to a Poisson process and the Poisson arrival rate and the distribution of service times are dependent on the state of an external Markov process (background process). When the interarrival times in our model is exponential, the one-dimensional case in Section 6 is similar to the one studied in [4]. In a system with multiple infinite server queues, [9] consider that both the arrival rates and the parameter of the exponentially distributed service times are modulated by a common background process. In [3], a similar model but a single queue with a Poisson arrival is revisited to study the asymptotic behavior of the number of customers in the system in the large-deviations regime. In some papers, arrival and service rates in an infinite server queue are governed by an external semi-Markov process. See [5] and [7] for instance. [11] study $M/M/\infty$ queue model modulated by an external continuous-time Markov Chain.

The remainder of the paper is structured as follows: In Section 2, we provide a description of the mathematical model. After deriving some general results on the (joint) moment generating function (mgf)/Laplace Transform (LT) and the first two moments in Section 3, we show in Section 4 that the limiting second-order joint moments are explicitly available when the service times are (potentially degenerated) exponentially distributed and the transient moments are obtainable when the interarrival times are exponentially distributed. Some numerical illus-

trations for the limiting behaviour of the first and second joint moments are provided at the end of Section 4.1. Section 5 is concerned with the particular case of deterministic interarrival times, where we show that the mgf has a simple expression as a matrix product (see Theorem 17). Application to a model related to the infinite server queues is provided in Section 6. It is assumed that the infinite server queues are modulated by an external semi-Markovian process, such that arrival times and service times depend on the state of the modulating process immediately before and after it switches states.

2. Model description

Let $\{N_t, t \geq 0\}$ be a renewal process associated with a non decreasing sequence $(T_i)_{i \in \mathbb{N}}$ with $T_0 = 0$, such that $(T_i - T_{i-1})_{i \geq 1}$ is independent and identically distributed (iid). Also let $\tau = T_1$ with a cumulative distribution function (cdf) $F(x) = 1 - \bar{F}(x)$ and the LT $\mathcal{L}^\tau(u) = \mathbb{E}(e^{-u\tau})$ for $u \geq 0$. We introduce a stationary ergodic finite Markov Chain $(X_i)_{i \in \mathbb{N}}$ with a state space $\mathcal{S} = \{0, \dots, K\}^k$ for some $K \in \mathbb{N}$ and $k \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$, so that X_i is for all i of the form $X_i = (X_{i1}, \dots, X_{ik})$ with $X_{ij} \in \{0, \dots, K\}$ for $j = 1, \dots, k$. Then for $\alpha \geq 0$, the discounted process $\{Z(t) = Z(t; \alpha) \in \mathbb{R}^k, t \geq 0\}$ is a vector of k processes $Z(t) = (Z_1(t), \dots, Z_k(t))$ with each process defined as

$$Z_j(t) := \sum_{i=1}^{N_t} X_{ij} e^{-\alpha(L_{ij}+T_i)} \mathbb{1}_{[t < L_{ij}+T_i]} = \sum_{i=1}^{\infty} X_{ij} e^{-\alpha(L_{ij}+T_i)} \mathbb{1}_{[T_i \leq t < L_{ij}+T_i]}, \tag{1}$$

where $(L_{ij})_{i \in \mathbb{N}, j=1, \dots, k}$ is a sequence of independent random variables (rvs) such that $(L_{i1}, \dots, L_{ik})_{i \in \mathbb{N}}$ is iid (although L_{i1}, \dots, L_{ik} may have different distributions). We set (L_1, \dots, L_k) to be a generic random vector distributed as the (L_{i1}, \dots, L_{ik}) 's, with each L_j having the LT denoted by $\mathcal{L}_j(u) = \mathbb{E}(e^{-uL_j})$ for $u \geq 0$. As in [12], we let $\tilde{Z}(t) = \tilde{Z}(t; \alpha) := e^{\alpha t} Z(t; \alpha)$. The processes described in (1) are viewed as different quantities of interest in the following two areas. In queueing theory, and especially when $\alpha = 0$, X_{ij} represents the number of customers arriving in queue $j \in \{1, \dots, k\}$ at time T_i , each of those customers with same service time L_{ij} . In actuarial science, when the severities of claims of different types occurring due to a common accident or catastrophe event and there are some time delays for insurers to hear (or settle) these claims, X_{ij} represents amounts of j -type of claim arriving at the time T_i and X_i is a vector of multivariate claims caused by this i th event. Since the present value of each claim amount is calculated by discount claim amounts X_{ij} at the actual realization time $T_i + L_{ij}$, $Z_j(t)$ in (1) is regarded as discounted IBNR amounts of j -type of claim by time t and $Z(t)$ is a vector of multivariate discounted IBNR claim processes involving k types of claims. Hence, X_{ij} and L_{ij} will in what follows be invariably referred to the claim/batch sizes and delay/service times respectively. In particular, when $\alpha = 0$, this model is a generalization of the Model II in [9] which considers the case where τ is exponentially distributed, i.e. when the set of queues is modulated by a common continuous-time Markov Chain, as will be discussed in Section 4.

A few words describing the technical difficulties that stand out against those in [12]. Although the structure of the latter paper is similar to the present one, the correlation structure of the X_i 's (namely, forming a Markov Chain) considerably increases the technical challenges leading to the results proved here. For example, the iid assumption on X_i 's enables us to determine all joint moments of $Z_1(t), \dots, Z_k(t)$ recursively in [12], whereas some matrix issues result in the expressions only available for the first two first joint moments in this paper, see

Theorems and 11 and 15. Besides, contrarily to [12], we prove in Theorem 17 that in a particular Markovian setting with constant interarrival times the distribution of the joint vector $(\tilde{Z}(t), X_{N_t})$ is converging and an explicit expression for the limiting distribution is available. Finally, the application of the model described in Section 6 is novel in a sense that, to the best of our knowledge, infinite server queues (IBNR process) where service time (reporting delay) for incoming customers (claims) depend on both states of the modulating exterior process at the switching time and prior to this time were not analytically studied in the literature.

Notation. Let $P = (p(x, x'))_{(x, x') \in \mathcal{S}^2}$ and $\pi = (\pi(x))_{x \in \mathcal{S}}$ (written as a row vector) be respectively the transition matrix and stationary distribution of the Markov Chain. We next define for all $r \geq 0$ and $s = (s_1, \dots, s_k) \in \mathbb{R}^k$,

$$\tilde{\pi}(s, r) := \text{diag} \left[\mathbb{E} \left(\exp \left\{ \sum_{j=1}^k s_j x_j e^{-\alpha(L_j - r)} \mathbb{1}_{[L_j > r]} \right\} \right), x = (x_1, \dots, x_k) \in \mathcal{S} \right], \quad (2)$$

$$\tilde{Q}(s, r) := \tilde{\pi}(s, r) P', \quad (3)$$

where P' denotes the transpose of matrix P . We also introduce some notations in the following. I is the identity matrix, $\mathbf{0}$ is a column vector with zeroes, and $\mathbf{1}$ is a column vector with 1's, of appropriate dimensions. When a random variable (rv) X is exponentially distributed with mean $1/\beta$, it is denoted as $X \sim \mathcal{E}(\beta)$. Also, we let the $\mathcal{S} \times \mathcal{S}$ diagonal matrices

$$\begin{aligned} \Delta_j &:= \text{diag} [x_j, x = (x_1, \dots, x_k) \in \mathcal{S}], \quad j = 1, \dots, k, \\ \Delta_\pi &:= \text{diag}(\pi(x), x \in \mathcal{S}), \end{aligned} \quad (4)$$

and δ is used to denote the Kronecker symbol, e.g. $\delta_{x,y}$ equals to 1 iff $x = y$ and 0 else. The mgf of the process $\tilde{Z}(t) = \tilde{Z}(t; \alpha)$ jointly to the state of X_{N_t} given the initial state of X_0 is denoted by

$$\tilde{\psi}(s, t) = \tilde{\psi}(s, t; \alpha) := \left[\mathbb{E} \left(e^{\langle s, \tilde{Z}(t) \rangle} \mathbb{1}_{[X_{N_t} = y]} \middle| X_0 = x \right) \right]_{(x, y) \in \mathcal{S}^2}, \quad t \geq 0, \quad (5)$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product on \mathbb{R}^k . Note that $s = (s_1, \dots, s_k)$ is assumed to be such that $s_j \in \mathbb{R}$ for all $j = 1, \dots, k$ so that (5) is well defined, i.e. the expectation is finite. Definition (5) may include the case where the s_j 's are complex and purely imaginary, in which case $\tilde{\psi}(s, t)$ is the characteristic function of $\tilde{Z}(t)$ jointly to X_{N_t} ; this will particularly be the case in the proof of Theorem 17. Note also that X_0 in (5) has no direct physical interpretation here, as the batch sizes/claims sizes are given by X_i , $i \geq 1$, and is rather introduced for technical purpose. We define the first and second (matrix) moments of $\tilde{Z}(t)$ jointly to the state of the Markov Chain X_{N_t} given the initial state X_0 as

$$\begin{aligned} M_j(t) &:= \left[\mathbb{E} \left(\tilde{Z}_j(t) \mathbb{1}_{[X_{N_t} = y]} \middle| X_0 = x \right) \right]_{(x, y) \in \mathcal{S}^2}, \quad j = 1, \dots, k, \\ M_{jj'}(t) &:= \left[\mathbb{E} \left(\tilde{Z}_j(t) \tilde{Z}_{j'}(t) \mathbb{1}_{[X_{N_t} = y]} \middle| X_0 = x \right) \right]_{(x, y) \in \mathcal{S}^2}, \quad j, j' = 1, \dots, k, \end{aligned} \quad (6)$$

respectively. We remark that the mgf defined in (5) is different from the one studied in [12] which does not consider Markovian assumption for a vector X_i and joint structure with the state X_{N_t} conditioning on the initial state X_0 .

This introductory section is completed by giving some results of independent interest that will be used in the rest of the paper. The following lemma is important for some computations on Markov Chains, which may be found in [6, Lemma 1]:

Lemma 1. Let $(S_n)_{n \in \mathbb{N}}$ be a stationary Markov Chain with a state space E , the transition matrix P and (stationary) distribution $\pi = (\pi(x))_{x \in E}$. For all functions f_1, \dots, f_{l+1} we have

$$\mathbb{E}(f_1(S_1) \cdots f_l(S_l)) = \mathbf{1}' \prod_{i=0}^{l-2} Q_{f_{l-i}} \pi_{f_1}, \quad (7)$$

where $Q_{f_i} := \text{diag}(f_i(z), z \in E)P'$ for $i = 1, \dots, l$, and $\pi_{f_1} := \text{diag}(f_1(z), z \in E)\pi'$.

The following is a direct consequence of (7). Let e_x (resp. e_y) be the column vector of which z th entry is $\delta_{x,z}$ (resp. $\delta_{y,z}$). One has then for all x, y in E that

$$\mathbb{E}(f_1(S_1) \cdots f_l(S_l) \mathbb{1}_{[S_l=y]} | S_1 = x) = e_y' \prod_{i=0}^{l-2} Q_{f_{l-i}} \text{diag}(f_1(z), z \in E) e_x,$$

which, because it is a scalar, is equal to its transpose, i.e.

$$\mathbb{E}(f_1(S_1) \cdots f_l(S_l) \mathbb{1}_{[S_l=y]} | S_1 = x) = e_x' \text{diag}(f_1(z), z \in E) \prod_{i=2}^l Q_{f_i}' e_y,$$

which immediately implies the following corollary.

Corollary 2. Under the same notation as in Lemma 1, one has the matrix equality

$$\left[\mathbb{E}(f_1(S_1) \cdots f_l(S_l) \mathbb{1}_{[S_l=y]} | S_1 = x) \right]_{(x,y) \in E^2} = \text{diag}(f_1(z), z \in E) \prod_{i=2}^l Q_{f_i}'.$$

3. General results

3.1. The Laplace transform

The aim of this subsection is to establish some properties verified by the mgf $\tilde{\psi}(s, t)$ in (5).

Proposition 3. The mgf of $\tilde{Z}(t)$ defined by (5) satisfies

$$\tilde{\psi}(s, t) = \mathbb{E} \left(\prod_{i=1}^{N_t} \tilde{Q}(s, t - T_i)' \right) = \bar{F}(t)I + \mathbb{E} \left(\mathbb{1}_{[N_t > 0]} \prod_{i=1}^{N_t} \tilde{Q}(s, t - T_i)' \right) \quad (8)$$

for all $s \in \mathbb{R}^k$, $t \geq 0$, with the usual convention $\prod_{i=1}^{N_t} \tilde{Q}(s, t - T_i)' = I$ if $N_t = 0$. Besides, it satisfies the following multidimensional integral equation:

$$\tilde{\psi}(s, t) = \bar{F}(t)I + \int_0^t \tilde{Q}(s, t - y)' \tilde{\psi}(s, t - y) dF(y), \quad \forall s \in \mathbb{R}^k, t \geq 0. \quad (9)$$

Proof. Decomposing according to $N_t = 0$ and $N_t > 0$ yields that

$$\tilde{\psi}(s, t) = [\mathbb{P}(X_0 = y | X_0 = x) \mathbb{P}(N_t = 0)]_{(x,y) \in \mathcal{S}^2} + \left[\mathbb{E} \left(e^{<s, \tilde{Z}(t)>} \mathbb{1}_{[X_{N_t} = y]} \mathbb{1}_{[N_t > 0]} \middle| X_0 = x \right) \right]_{(x,y) \in \mathcal{S}^2}. \quad (10)$$

Note that $\mathbb{P}(X_0 = y | X_0 = x) \mathbb{P}(N_t = 0) = \delta_{x,y} \bar{F}(t)$, so that the first term on the right-hand side of (10) is given by the the first term on the right-hand side of (8). We turn to the second

term on the right-hand side of (10). Let us define $\mathcal{F} = \sigma(T_i, i \in \mathbb{N})$ the sigma field generated by T_i 's for $i \in \mathbb{N}$ as well as the set of rvs

$$s_t(x, y) := \mathbb{E} \left(e^{\langle s, \tilde{Z}(t) \rangle} \mathbb{1}_{[X_{N_t}=y]} \mathbb{1}_{[N_t>0]} \middle| X_0 = x, \mathcal{F} \right), \quad x, y \in \mathcal{S},$$

where $s = (s_1, \dots, s_k) \in \mathbb{R}^k$ is fixed throughout the proof. Using $\tilde{Z}(t) = e^{\alpha t} Z(t; \alpha)$ and (1), one obtains

$$s_t(x, y) = \mathbb{E} \left(\mathbb{1}_{[N_t>0]} \prod_{i=1}^{N_t} \exp \left\{ \sum_{j=1}^k s_j X_{ij} e^{-\alpha(L_{ij} - (t-T_i))} \mathbb{1}_{[L_{ij}>t-T_i]} \right\} \mathbb{1}_{[X_{N_t}=y]} \middle| X_0 = x, \mathcal{F} \right). \tag{11}$$

In order to compute $s_t(x, y)$, we use the fact that the Markov Chain $(X_i)_{i \in \mathbb{N}}$ is independent from \mathcal{F} and $(L_{ij})_{i \in \mathbb{N}, j=1, \dots, k}$. Using the result in Corollary 2 with replacement of $S_i := X_{i-1}$, $f(S_1) = f(X_0) = 1$ and $f_i(S_i) = f_i(X_{i-1}) := \exp \left\{ \sum_{j=1}^k s_j X_{i-1,j} e^{-\alpha(L_{(i-1),j} - (t-T_{i-1}))} \mathbb{1}_{[L_{(i-1),j}>t-T_{i-1}]} \right\}$ for $i = 2, \dots, l$ when $l = N_t + 1$, (11) may be expressed as

$$[s_t(x, y)]_{(x,y) \in \mathcal{S}^2} = \mathbb{1}_{[N_t>0]} I \cdot \prod_{i=2}^{N_t+1} \tilde{Q}(s, t - T_{i-1})' = \mathbb{1}_{[N_t>0]} \prod_{i=1}^{N_t} \tilde{Q}(s, t - T_i)',$$

where we recall that I is the identity matrix and $\tilde{Q}(\cdot, \cdot)$ is defined in (3). Since $\mathbb{E}([s_t(x, y)]_{(x,y) \in \mathcal{S}^2})$ is the second term in the right-hand side of (10), one thus obtains (8).

Finally, (9) is obtained by considering again $[N_t = 0] \iff [T_1 > t]$ and $[N_t > 0] \iff [T_1 \leq t]$ and conditioning with respect to T_1 . \square

It is known that a multidimensional integral equation such as (9) cannot be solved in general. One particular case is when arrivals occur according to a Poisson process, in which case one has the following result.

Proposition 4. *If $\tau \sim \mathcal{E}(\lambda)$ for $\lambda > 0$, then $\tilde{\psi}(s, t)$ is the unique solution to the first-order linear (matrix) differential equation*

$$\partial_t \tilde{\psi}(s, t) = [-\lambda I + \lambda \tilde{Q}(s, t)'] \tilde{\psi}(s, t) = [\lambda(P - I) + \lambda P(\tilde{\pi}(s, t) - I)] \tilde{\psi}(s, t) \tag{12}$$

with the initial condition $\tilde{\psi}(s, 0) = I$ from (8).

Note that, even though (12) admits a unique solution, $\tilde{\psi}(s, t)$ is not explicit except for particular cases e.g. when $P = I$. That is, when there is no Markov interference in the model, the differential equation (12) can be solved componentwise. Also note that a similar differential equation was obtained when $\alpha = 0$ in [10, Theorem 3.1] for the joint generating function, when interarrival times are matrix exponentially distributed.

Proof. Setting $dF(y) = \lambda e^{-\lambda y} dy$ and $\bar{F}(t) = e^{-\lambda t}$ in (9) and differentiating with respect to t yields (12). \square

3.2. The first and second moments

We are now interested in the first two moments defined in (6). It is standard that $M_j(t)$ and $M_{jj'}(t)$ are linked to $\tilde{\psi}(s, t)$ by

$$M_j(t) = \partial_{s_j} \tilde{\psi}(s, t) \Big|_{s=0}, \quad M_{jj'}(t) = \partial_{s_j} \partial_{s_{j'}} \tilde{\psi}(s, t) \Big|_{s=0}, \quad j, j' = 1, \dots, k.$$

It requires to differentiate $\tilde{\pi}(s, r)$ in (2) with respect to s_j or s_j and $s_{j'}$ followed by putting $s = \mathbf{0}$. One obtains

$$\partial_{s_j} \tilde{\pi}(s, r) \Big|_{s=\mathbf{0}} = \mathbb{E} \left(e^{-\alpha(L_j-r)} \mathbb{1}_{[L_j>r]} \right) \Delta_j, \tag{13}$$

$$\partial_{s_{j'}} \partial_{s_j} \tilde{\pi}(s, r) \Big|_{s=\mathbf{0}} = \mathbb{E} \left(e^{-\alpha(L_j-r)} \mathbb{1}_{[L_j>r]} e^{-\alpha(L_{j'}-r)} \mathbb{1}_{[L_{j'}>r]} \right) \Delta_j \Delta_{j'}, \tag{14}$$

where Δ_j is given by (4). Moreover, one also needs to compute $\tilde{\psi}(\mathbf{0}, r)$, namely from (8)

$$\tilde{\psi}(\mathbf{0}, r) = \mathbb{E} \left(\prod_{i=1}^{N_r} \tilde{Q}(\mathbf{0}, r - T_i)' \right) = \mathbb{E} (P^{N_r}), \quad r \geq 0, \tag{15}$$

as indeed $\tilde{Q}(\mathbf{0}, r) = P'$, again with the convention $P^0 = I$. It is convenient to introduce the following notations for all $j, j' = 1, \dots, k$:

$$b_j(t) := \int_0^t \left[\partial_{s_j} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}} \right] P \tilde{\psi}(\mathbf{0}, t-y) dF(y), \tag{16}$$

$$\begin{aligned} b_{jj'}(t) &:= \int_0^t \left[\partial_{s_{j'}} \partial_{s_j} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}} \right] P \tilde{\psi}(\mathbf{0}, t-y) dF(y) \\ &+ \int_0^t \left[\partial_{s_j} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}} \right] P M_{j'}(t-y) dF(y) + \int_0^t \left[\partial_{s_{j'}} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}} \right] P M_j(t-y) dF(y), \end{aligned} \tag{17}$$

where $\partial_{s_j} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}}$, $\partial_{s_{j'}} \partial_{s_j} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}}$, $\partial_{s_{j'}} \partial_{s_j} \tilde{\pi}(s, t-y) \Big|_{s=\mathbf{0}}$ and $\tilde{\psi}(\mathbf{0}, t-y)$ are given by (13), (14) and (15). Note that $b_j(t)$ is not explicit since (15) does not in general have a closed-form expression, and so is $b_{jj'}(t)$. We may then obtain general results concerning $M_j(t)$ and $M_{jj'}(t)$. Following the notation in [2, Section 2], we define, for a $N \times N$ dimensional matrix of non decreasing right continuous functions $(t \mapsto F_{ij}(t))_{i,j=1,\dots,N}$ and a $N \times N$ dimensional matrix of bounded measurable functions $(t \mapsto H_{ij}(t))_{i,j=1,\dots,N}$, the convolution $t \mapsto F \star H(t) = (F \star H)_{i,j=1,\dots,N}(t)$ by

$$(F \star H)_{i,j}(t) := \sum_{h=1}^N \int_0^t H_{hj}(t-u) dF_{ih}(u) \quad i, j = 1, \dots, N, \quad t \geq 0.$$

Then, differentiating (9) with respect to s_j or s_j and $s_{j'}$ followed by putting $s = \mathbf{0}$ yields the following proposition.

Proposition 5. *For $j, j' = 1, \dots, k$, $M_j(t)$ and $M_{jj'}(t)$ satisfy the following multidimensional renewal equations*

$$M_j(t) = b_j(t) + (PF) \star M_j(t), \quad t \geq 0, \tag{18}$$

$$M_{jj'}(t) = b_{jj'}(t) + (PF) \star M_{jj'}(t), \quad t \geq 0, \tag{19}$$

where $(PF)(t) := (p(x, y)F(t))_{(x,y) \in \mathcal{S}^2}$, and $b_j(t)$ and $b_{jj'}(t)$ are given by (16) and (17).

Although the solution for (9) does not have a closed-form expression, it turns out that (16) and (17) have solutions which can be expressed in terms of a multidimensional renewal

function. Indeed, since interarrival times satisfy $\tau > 0$ a.s., one has that $(PF)(0)$ is the zero matrix, of which largest eigenvalue is thus 0. [2, Lemma 2.1] entails that

$$M_j(t) = U \star b_j(t), \quad M_{jj'}(t) = U \star b_{jj'}(t), \quad t \geq 0, \quad (20)$$

where $U(t)$ is the renewal function defined by $U(t) := \sum_{n=0}^{\infty} (PF)^{\star(n)}(t)$, an $\mathcal{S} \times \mathcal{S}$ matrix, see [2, Definition (2.3)]. At this point, solutions in (20) are still not satisfactory because $U(t)$, $b_j(t)$, and $b_{jj'}(t)$ are not explicit. Hence, limiting behaviours of $M_j(t)$ and $M_{jj'}(t)$ are studied instead given as below.

Lemma 6. *Let us suppose that τ is non lattice, then one has the following*

$$M_j(t) \longrightarrow \frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \int_0^{\infty} b_j(t) dt, \quad t \rightarrow \infty, \quad j = 1, \dots, k, \quad (21)$$

$$M_{jj'}(t) \longrightarrow \frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \int_0^{\infty} b_{jj'}(t) dt, \quad t \rightarrow \infty, \quad j, j' = 1, \dots, k. \quad (22)$$

Proof. Since $(PF)(\infty) = P$ has a spectral radius equal to 1 and has (row vector) π and $\mathbf{1}$ (column vector) as left and right eigenvectors associated to the eigenvalue 1, the renewal equation (18) satisfied by $M_j(t)$ and [2, Theorem 2.2 (iii)] yield (21). The same method applied to the renewal equation (19) satisfied by $M_{jj'}(t)$ yields (22). \square

Note that the limit in (21) is not clearly available because an explicit expression for $\mathbb{E}(P^{N_t})$ is required to integrate $b_j(t)$ (see (16) with $\tilde{\psi}(\mathbf{0}, t-y)$ given by (15)). Likewise, the limit in (22) is not explicit either as it requires analytic expressions for $M_j(t)$ and $M_{j'}(t)$ in (17). However, the limiting first moment is explicitly available as below:

Proposition 7. *Assuming that τ is non lattice, the expectation $M_j(t)\mathbf{1} = [\mathbb{E}(\tilde{Z}_j(t) | X_1 = x)]'_{x \in \mathcal{S}}$ asymptotically behaves as*

$$M_j(t)\mathbf{1} \longrightarrow \frac{\mathbb{E}(X_j)}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}_j(\alpha)}{\alpha} \right] \mathbf{1}, \quad t \rightarrow \infty, \quad j = 1, \dots, k. \quad (23)$$

Proof. Let us prove that one obtains (23) by post multiplying (21) by $\mathbf{1}$. First note that, as $P\mathbf{1} = \mathbf{1}$, $b_j(t)\mathbf{1}$ reduces thanks to (16), (13) and (15) to

$$\begin{aligned} b_j(t)\mathbf{1} &= \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - (t-y))} \mathbb{1}_{[L_j > t-y]} \right) \Delta_j \cdot P\mathbb{E} \left(P^{N_{t-y}} \mathbf{1} dF(y) \right) \\ &= \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - (t-y))} \mathbb{1}_{[L_j > t-y]} \right) dF(y) \cdot \Delta_j \mathbf{1} \\ &= \mathbb{E} \left(e^{-\alpha(L_j - (t-\tau))} \mathbb{1}_{[0 \leq t-\tau < L_j]} \right) \Delta_j \mathbf{1}. \end{aligned} \quad (24)$$

In order to compute $\int_0^{\infty} b_j(t)\mathbf{1} dt$, one calculates

$$\begin{aligned} \int_0^{\infty} \mathbb{E} \left(e^{-\alpha(L_j - (t-\tau))} \mathbb{1}_{[L_j > t-\tau \geq 0]} \right) dt &= \int_0^{\infty} \mathbb{E} \left(e^{-\alpha(L_j - (t-\tau))} \mathbb{1}_{[\tau \leq t < L_j + \tau]} \right) dt \\ &= \mathbb{E} \left(\int_{\tau}^{L_j + \tau} e^{-\alpha(L_j - (t-\tau))} dt \right) = \mathbb{E} \left(\frac{1}{\alpha} [1 - e^{-\alpha L_j}] \right) = \frac{1 - \mathcal{L}_j(\alpha)}{\alpha}, \end{aligned}$$

so that one obtains from (24) and (21) that, as $t \rightarrow \infty$,

$$M_j(t)\mathbf{1} \longrightarrow \frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \int_0^{\infty} b_j(t)\mathbf{1} dt = \frac{1}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}_j(\alpha)}{\alpha} \right] \mathbf{1}\pi \Delta_j \mathbf{1}.$$

One checks easily that $\pi \Delta_j \mathbf{1} = \mathbb{E}(X_j)$ where Δ_j is given in (4), yielding (23). \square

3.3. The workload

In this section, in the context of infinite server queues, following [12, Section 5.2], we then consider the workload of the queues as the vector of k processes $D(t) = (D_1(t), \dots, D_k(t))$ with each process defined as

$$D_j(t) := \sum_{i=1}^{N_t} X_{ij}(T_i + L_{ij} - t) \mathbb{1}_{[t < T_i + L_{ij}]}, \quad t \geq 0, \quad j = 1, \dots, k.$$

and that one has for all $t \geq 0$ and $j = 1, \dots, k$

$$D_j(t) = -\frac{\partial}{\partial \alpha} \tilde{Z}_j(t; \alpha) \Big|_{\alpha=0}.$$

Here we define the joint expectation of the workload and the state of X_{N_t} given the initial state of X_0 as

$$\begin{aligned} W_j(t) &:= \left[\mathbb{E} \left(D_j(t) \mathbb{1}_{[X_{N_t}=y]} \mid X_0 = x \right) \right]_{(x,y) \in \mathcal{S}^2} \\ &= -\frac{\partial}{\partial \alpha} M_j(t) \Big|_{\alpha=0} = \left[-\frac{\partial}{\partial \alpha} \left[\partial_{s_j} \tilde{\psi}(s, t; \alpha) \right]_{s=0} \right]_{\alpha=0}. \end{aligned} \quad (25)$$

The following results the analogs of Proposition 5, Lemma 6 and Proposition 7. First, let us define and compute

$$\ell_j(t) := \int_0^t \left[-\frac{\partial}{\partial \alpha} \left[\partial_{s_j} \tilde{\pi}(s, t-y) \right]_{s=0} \right]_{\alpha=0} P \tilde{\psi}(\mathbf{0}, t-y) dF(y), \quad (26)$$

where it follows from (13) that

$$\left[-\frac{\partial}{\partial \alpha} \left[\partial_{s_j} \tilde{\pi}(s, r) \right]_{s=0} \right]_{\alpha=0} = \mathbb{E} \left((L_j - r) \mathbb{1}_{[L_j > r]} \right) \Delta_j. \quad (27)$$

Consequently, the following proposition is provided.

Proposition 8. *The joint expectation of the workload and the state of X_{N_t} satisfies*

$$W_j(t) = \ell_j(t) + (PF) \star W_j(t), \quad t \geq 0, \quad j = 1, \dots, k, \quad (28)$$

of which asymptotic expression is given by

$$W_j(t) \longrightarrow \frac{1}{\mathbb{E}(\tau)} \mathbf{1} \pi \int_0^\infty \ell_j(t) dt, \quad t \rightarrow \infty, \quad j = 1, \dots, k. \quad (29)$$

Moreover, the asymptotic expected workload $W_j(t) \mathbf{1} = [\mathbb{E}(D_j(t) \mid X_0 = x)]'_{x \in \mathcal{S}}$ is given by

$$W_j(t) \mathbf{1} \longrightarrow \left[\frac{\mathbb{E}(L_j^2)}{2\mathbb{E}(\tau)} + \frac{\mathbb{E}(\tau^2)}{2\mathbb{E}(\tau)} + \mathbb{E}(L_j) \right] \mathbb{E}(X_j) \mathbf{1}, \quad t \rightarrow \infty, \quad j = 1, \dots, k. \quad (30)$$

Proof. Since the proof of (28) and (29) is analogous to the one of (18) and (21), our focus is on (30). As in the proof of Proposition 7 similar to (24), using $P\mathbf{1} = \mathbf{1}$ one finds

$$\ell_j(t) \mathbf{1} = \mathbb{E} \left((L_j + \tau - t) \mathbb{1}_{[L_j + \tau > t]} \right) \Delta_j \mathbf{1},$$

with $\int_0^\infty \mathbb{E} \left((L_j + \tau - t) \mathbb{1}_{[L_j + \tau > t]} \right) dt = \mathbb{E} \left(\int_0^{L_j + \tau} (L_j + \tau - t) dt \right) = \mathbb{E} \left((L_j + \tau)^2 / 2 \right)$. Using independence of L_j and τ in the last expectation as well as $\pi \Delta_j \mathbf{1} = \mathbb{E}(X_j)$ yields (30) by post multiplying (29) by $\mathbf{1}$. \square

4. Special cases

The results given in Proposition 3, Lemma 6 and Proposition 8 hold under general assumptions on the service times (L_j) and interarrival times (τ). We present here some particular cases where those results are more explicitly obtainable with a specific distributional assumption for L_j or τ . Namely, as is customary when studying infinite server queues, one expects reasonably to obtain more information when one of those two rvs are exponentially distributed, see [13, Chapter 3].

4.1. Exponentially distributed service times

First, it is assumed that service time L_j for j -type customer is exponentially distributed with parameter $\mu_j > 0$. To provide explicit expressions for the limits of $M_j(t)$, $M_{jj'}(t)$ and $W_j(t)$ as $t \rightarrow \infty$ in (21), (22) and (29), we define the Laplace transforms for $\tilde{\psi}(s, t)$, $M_j(t)$, and $b_j(t)$ by

$$\begin{aligned} \hat{\psi}(s, h) &= \int_0^\infty \tilde{\psi}(s, t)e^{-ht} dt, \quad s \in \mathbb{R}^k, \\ \hat{M}_j(h) &= \int_0^\infty M_j(t)e^{-ht} dt, \\ \hat{b}_j(h) &= \int_0^\infty b_j(t)e^{-ht} dt, \end{aligned} \tag{31}$$

respectively for all $h > 0$. Next, some relations between the above quantities are first given.

Proposition 9. *The Laplace transforms verify for all $h > 0$*

$$\hat{\psi}(\mathbf{0}, h) = \frac{1 - \mathcal{L}^\tau(h)}{h} (I - \mathcal{L}^\tau(h)P)^{-1}, \tag{32}$$

$$\hat{M}_j(h) = (I - \mathcal{L}^\tau(h)P)^{-1} \hat{b}_j(h), \quad j = 1, \dots, k. \tag{33}$$

Proof. Recalling that $\tilde{\pi}(\mathbf{0}, r) = I$ from (2), (9) with $s = \mathbf{0}$ becomes the renewal equation

$$\tilde{\psi}(\mathbf{0}, t) = \bar{F}(t)I + (PF) \star \tilde{\psi}(\mathbf{0}, \cdot)(t),$$

which, upon taking Laplace transforms on both sides, yields

$$\hat{\psi}(\mathbf{0}, h) = \frac{1 - \mathcal{L}^\tau(h)}{h} I + P\mathcal{L}^\tau(h)\hat{\psi}(\mathbf{0}, h).$$

Then (32) is obtained by noting that, since $\mathcal{L}^\tau(h) < 1$ and P is a stochastic matrix, the matrix $\mathcal{L}^\tau(h)P$ has spectral radius less than 1 hence $I - \mathcal{L}^\tau(h)P$ is invertible. Similarly, (33) is obtained by taking Laplace transforms in the renewal equation (18). \square

Theorem 10. *The asymptotic result for the first moment jointly to the state of X_{N_t} in (21) can be precisely expressed as*

$$M_j(t) \longrightarrow \frac{1}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}^\tau(\mu_j)}{\mu_j + \alpha} \right] \mathbf{1}\pi\Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1}, \quad t \rightarrow \infty, \quad j = 1, \dots, k. \tag{34}$$

Proof. When $L_j \sim \mathcal{E}(\mu_j)$, one computes

$$\mathbb{E} \left(e^{-\alpha(L_j-r)} \mathbb{1}_{[L_j>r]} \right) = \int_r^\infty e^{-\alpha(t-r)} \mu_j e^{-\mu_j t} dt = \mu_j e^{\alpha r} \int_r^\infty e^{-(\mu_j+\alpha)t} dt = \frac{\mu_j}{\mu_j + \alpha} e^{-\mu_j r}, \quad r \geq 0, \quad (35)$$

so that one has (16) from (13) and (35) that

$$\begin{aligned} b_j(t) &= \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \int_0^t e^{-\mu_j(t-y)} \tilde{\psi}(\mathbf{0}, t-y) dF(y) \\ &= \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \mathbb{E} \left[e^{-\mu_j(t-\tau)} \tilde{\psi}(\mathbf{0}, t-\tau) \mathbb{1}_{[t \geq \tau]} \right]. \end{aligned} \quad (36)$$

The right-hand side of (21) is thus computed as

$$\begin{aligned} &\frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \int_0^\infty b_j(t) dt \\ &= \frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \int_0^\infty \mathbb{E} \left[e^{-\mu_j(t-\tau)} \tilde{\psi}(\mathbf{0}, t-\tau) \mathbb{1}_{[t \geq \tau]} \right] dt \\ &= \frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \mathbb{E} \left[\int_\tau^\infty e^{-\mu_j(t-\tau)} \tilde{\psi}(\mathbf{0}, t-\tau) dt \right] \\ &= \frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \hat{\psi}(\mathbf{0}, \mu_j), \end{aligned}$$

and in turn, (34) is obtained thanks to (32). \square

Let us note that the previous proof enables us to similarly obtain the expression of $\hat{b}_j(h)$ defined in (31) thanks to (36) as follows

$$\begin{aligned} \hat{b}_j(h) &= \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \int_0^\infty e^{-ht} \mathbb{E} \left[e^{-\mu_j(t-\tau)} \tilde{\psi}(\mathbf{0}, t-\tau) \mathbb{1}_{[t \geq \tau]} \right] dt \\ &= \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \mathbb{E} \left[\int_\tau^\infty e^{-h(t-\tau)} e^{-\mu_j(t-\tau)} \tilde{\psi}(\mathbf{0}, t-\tau) dt \cdot e^{-h\tau} \right] \\ &= \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \hat{\psi}(\mathbf{0}, \mu_j + h) \cdot \mathcal{L}^\tau(h), \quad h > 0. \end{aligned} \quad (37)$$

Theorem 11. *The asymptotic result for the second moment jointly to the state of X_{N_t} in (22) can be precisely expressed as*

$$\begin{aligned} M_{jj'}(t) &\longrightarrow \frac{1}{\mathbb{E}(\tau)} \left(\frac{\mu_j}{\mu_j + \alpha} \right) \left(\frac{\mu_{j'}}{\mu_{j'} + \alpha} \right) \left[\frac{1 - \mathcal{L}^\tau(\mu_j + \mu_{j'})}{\mu_j + \mu_{j'}} \right] \mathbf{1}\pi \left\{ \Delta_j \Delta_{j'} + \mathcal{L}^\tau(\mu_j) \Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1} \Delta_{j'} \right. \\ &\quad \left. + \mathcal{L}^\tau(\mu_{j'}) \Delta_{j'} P (I - \mathcal{L}^\tau(\mu_{j'})P)^{-1} \Delta_j \right\} P (I - \mathcal{L}^\tau(\mu_{j'} + \mu_j)P)^{-1} \end{aligned} \quad (38)$$

as $t \rightarrow \infty$, when $j, j' = 1, \dots, k$, $j \neq j'$, and

$$\begin{aligned} M_{jj'}(t) &\longrightarrow \frac{1}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}^\tau(\mu_j)}{\mu_j + 2\alpha} \right] \mathbf{1}\pi \Delta_j^2 P (I - \mathcal{L}^\tau(\mu_j)P)^{-1} \\ &\quad + \frac{2\mathcal{L}^\tau(\mu_j)}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}^\tau(2\mu_j)}{2\mu_j} \right] \left(\frac{\mu_j}{\mu_j + \alpha} \right)^2 \mathbf{1}\pi \Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1} \Delta_j P (I - \mathcal{L}^\tau(2\mu_j)P)^{-1} \end{aligned} \quad (39)$$

as $t \rightarrow \infty$, when $j, j' = 1, \dots, k$, $j = j'$. We remark that (38) and (39) still hold when μ_j or $\mu_{j'}$ is infinite, i.e. when the corresponding delays L_j or $L_{j'}$ are 0.

Proof. One first computes that

$$\mathbb{E} \left(e^{-\alpha(L_j - r)} \mathbb{1}_{[L_j > r]} e^{-\alpha(L_{j'} - r)} \mathbb{1}_{[L_{j'} > r]} \right) = \begin{cases} \frac{\mu_j}{\mu_j + \alpha} \frac{\mu_{j'}}{\mu_{j'} + \alpha} e^{-\mu_j r} e^{-\mu_{j'} r} & \text{if } j \neq j', \\ \frac{\mu_j}{\mu_j + 2\alpha} e^{-\mu_j r} & \text{if } j = j' \end{cases} \quad (40)$$

for $r \geq 0$. To evaluate the integral in (22) with (17), we thus need to compute the following integrals:

$$\int_0^\infty \int_0^t \partial_{s_{j'}} \partial_{s_j} \tilde{\pi}(s, t - y) \Big|_{s=0} P \tilde{\psi}(\mathbf{0}, t - y) dF(y), \quad (41)$$

$$\int_0^\infty \int_0^t \partial_{s_j} \tilde{\pi}(s, t - y) \Big|_{s=0} P M_{j'}(t - y) dF(y), \quad (42)$$

for $j, j' = 1, \dots, k$. When $j \neq j'$, using (14) with (40) followed by applying (32), (41) may be expressed as

$$\begin{aligned} & \left(\frac{\mu_j}{\mu_j + \alpha} \right) \left(\frac{\mu_{j'}}{\mu_{j'} + \alpha} \right) \Delta_j \Delta_{j'} P \int_0^\infty \mathbb{E} \left[e^{-(\mu_j + \mu_{j'})(t - \tau)} \tilde{\psi}(\mathbf{0}, t - \tau) \mathbb{1}_{[t \geq \tau]} \right] dt \\ &= \left(\frac{\mu_j}{\mu_j + \alpha} \right) \left(\frac{\mu_{j'}}{\mu_{j'} + \alpha} \right) \Delta_j \Delta_{j'} P \hat{\psi}(\mathbf{0}, \mu_j + \mu_{j'}) \\ &= \left(\frac{\mu_j}{\mu_j + \alpha} \right) \left(\frac{\mu_{j'}}{\mu_{j'} + \alpha} \right) \left[\frac{1 - \mathcal{L}^\tau(\mu_j + \mu_{j'})}{\mu_j + \mu_{j'}} \right] \Delta_j \Delta_{j'} P (I - \mathcal{L}^\tau(\mu_j + \mu_{j'})P)^{-1}. \end{aligned} \quad (43)$$

When $j = j'$, similar computation yields that (41) is expressed as

$$\left[\frac{1 - \mathcal{L}^\tau(\mu_j)}{\mu_j + 2\alpha} \right] \Delta_j^2 P (I - \mathcal{L}^\tau(\mu_j)P)^{-1}, \quad (44)$$

where $\Delta_j^2 = \text{diag}[x_j^2, x = (x_1, \dots, x_k) \in \mathcal{S}]$ for $j = 1, \dots, k$. Turning to (42), replacing (13) with (35) followed by using (33) and (37) with (32) yields

$$\begin{aligned} & \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \int_0^\infty \mathbb{E} \left[e^{-\mu_j(t - \tau)} M_{j'}(t - \tau) \mathbb{1}_{[t \geq \tau]} \right] dt = \frac{\mu_j}{\mu_j + \alpha} \Delta_j P \hat{M}_{j'}(\mu_j) \\ &= \frac{\mu_j}{\mu_j + \alpha} \Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1} \hat{b}_{j'}(\mu_j) \\ &= \left(\frac{\mu_j}{\mu_j + \alpha} \right) \left(\frac{\mu_{j'}}{\mu_{j'} + \alpha} \right) \mathcal{L}^\tau(\mu_j) \Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1} \Delta_{j'} P \hat{\psi}(\mathbf{0}, \mu_{j'} + \mu_j) \\ &= \left(\frac{\mu_j}{\mu_j + \alpha} \right) \left(\frac{\mu_{j'}}{\mu_{j'} + \alpha} \right) \left[\frac{1 - \mathcal{L}^\tau(\mu_{j'} + \mu_j)}{\mu_{j'} + \mu_j} \right] \mathcal{L}^\tau(\mu_j) \Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1} \Delta_{j'} P (I - \mathcal{L}^\tau(\mu_{j'} + \mu_j)P)^{-1}. \end{aligned} \quad (45)$$

Then, gathering expressions (43) and (45) for (41) and (42) respectively yields (38). Also, (39) is obtained with the help of (44) and (45). \square

Theorem 12. *The asymptotic result for the expectation of the workload jointly to the state of X_{N_t} in (29) can be precisely expressed as*

$$W_j(t) \longrightarrow \frac{1}{\mu_j^2} \left[\frac{1 - \mathcal{L}^\tau(\mu_j)}{\mathbb{E}(\tau)} \right] \mathbf{1}\pi \Delta_j P (I - \mathcal{L}^\tau(\mu_j)P)^{-1}, \quad t \rightarrow \infty, \quad j = 1, \dots, k. \quad (46)$$

Proof. When $L_j \sim \mathcal{E}(\mu_j)$, one straightforward verifies that $\mathbb{E}((L_j - r)\mathbb{1}_{[L_j > r]}) = e^{-\mu_j r} / \mu_j$. Then, one has from (26) and (27) that

$$\ell_j(t) = \frac{1}{\mu_j} \Delta_j P \mathbb{E} \left(e^{-\mu_j(t-\tau)} \tilde{\psi}(\mathbf{0}, t - \tau) \mathbb{1}_{[t \geq \tau]} \right),$$

from which the computation of $\frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \int_0^\infty \ell_j(t) dt$ in (29) is led similarly to that of $\frac{1}{\mathbb{E}(\tau)} \mathbf{1}\pi \int_0^\infty b_j(t) dt$ in Theorem 10. Hence, the result (46) follows by using (32). \square

Example 13. This example illustrates numerically convergences of (34) and (39) for the first and second joint moments when $\alpha = 0.1$. This was done by simulating $(Z(t), X_{N_t})$ in (1) for (small) $t = 30$ and (large) $t = 100$ and estimating the left-hand side of (34) and (39) thanks to the Law of Large Numbers (Monte Carlo) through $n = 500$ iterations. The right-hand side of (34) and (39) were computed explicitly by considering for the interarrival τ a Gamma distribution with shape a and rate b with the LT $\mathcal{L}^\tau(u) = \frac{1}{(1+u/b)^a}$ and $\mathbb{E}(\tau) = \frac{a}{b}$. We consider two cases of (a, b) choosing $(1, 10)$ (i.e. $\tau \sim \mathcal{E}(10)$), so that $\mathbb{E}(\tau) = 0.1$, and $(0.75, 15)$ with $\mathbb{E}(\tau) = 0.05$. Suppose that $k = 1$, i.e. a one dimensional process $\{Z(t), t \geq 0\}$ and a Markov Chain $(X_i)_{i \in \mathbb{N}}$ with state space $\{0, 1\}$ (i.e. $K = 1$). Assume that the transition matrix is given by $P = \begin{pmatrix} 0.25 & 0.75 \\ 0.5 & 0.5 \end{pmatrix}$, with stationary distribution $\pi = (0.4, 0.6)$. All simulations and computations were done with `Scilab`. We finally suppose that all delays have same distribution $L \sim \mathcal{E}(1)$.

	$(a, b) = (1, 10)$	$(a, b) = (0.75, 15)$
Exact	$\begin{pmatrix} 2.222 & 3.232 \\ 2.222 & 3.232 \end{pmatrix}$	$\begin{pmatrix} 4.405 & 6.504 \\ 4.405 & 6.504 \end{pmatrix}$
Monte Carlo for $t = 30$	$\begin{pmatrix} 1.758 & 3.152 \\ 2.218 & 3.566 \end{pmatrix}$	$\begin{pmatrix} 4.185 & 6.645 \\ 3.925 & 6.503 \end{pmatrix}$
Monte Carlo for $t = 100$	$\begin{pmatrix} 2.242 & 3.153 \\ 1.892 & 3.574 \end{pmatrix}$	$\begin{pmatrix} 4.476 & 6.344 \\ 4.411 & 6.509 \end{pmatrix}$

TABLE 1
First-order joint moments $M_1(t)$

	$(a, b) = (1, 10)$	$(a, b) = (0.75, 15)$
Exact	$\begin{pmatrix} 14.196 & 20.188 \\ 14.196 & 20.188 \end{pmatrix}$	$\begin{pmatrix} 52.596 & 76.623 \\ 52.596 & 76.623 \end{pmatrix}$
Monte Carlo for $t = 30$	$\begin{pmatrix} 11.017 & 23.043 \\ 12.322 & 22.652 \end{pmatrix}$	$\begin{pmatrix} 48.302 & 78.761 \\ 47.395 & 82.891 \end{pmatrix}$
Monte Carlo for $t = 100$	$\begin{pmatrix} 12.543 & 21.094 \\ 14.192 & 22.089 \end{pmatrix}$	$\begin{pmatrix} 53.635 & 77.217 \\ 45.289 & 83.454 \end{pmatrix}$

TABLE 2
Second-order joint moments $M_{11}(t)$

Both tables above illustrate the convergence of the first-order and the second-order joint moments to the values calculated in (34) and (39) respectively. Also note that this simple

example gives us some idea of potential applications of the model. First, from a queuing point of view, the numerical model described here explains some infinite server queue system where, if an arriving customer is not admitted in the queue at time T_i (i.e. $X_i = 0$) then the next one arriving at time T_{i+1} is accepted with high probability 0.75; this is especially interesting in a congestion regulation context, where one may choose to accept incoming customers more easily when the previous ones were rejected with high probability. Furthermore, it would be interesting to utilize the model for an insurance company facing a situation where claims occurring at time T_i are either immediately reported when $X_i = 0$ or reported with delay L_i when $X_i = 1$. From P , some interesting feature appears as if a claim is not reported (resp. reported) at time T_i then the next one is reported (resp. not reported) at time T_{i+1} with probability 0.5 (resp. with probability 0.75). This model could reflect the policyholder's certain type of behaviour, e.g. after immediately reporting a claim at time T_i , the policyholder prefers to delay reporting of the next claim at time $L_{i+1} + T_{i+1}$ with probability 0.75 to avoid the increase of premium when the policyholder renews the insurance. On the other hand, it can also explain the opposite situation with a different transition matrix P . For example, a worker understands that if the reporting delay of workplace injury is longer, then it is harder to prove the injury is work related under the workers compensation claim. In this case, the transition probability from $X_i = 0$ to $X_{i+1} = 1$ is much lower, that is, the policyholder prefers to report the claim immediately.

4.2. Exponentially distributed interarrival times

We now suppose in this subsection that $\tau \sim \mathcal{E}(\lambda)$, i.e. that arrivals occur according to a Poisson process with intensity $\lambda > 0$. From Proposition 4, it has been shown that the transient mgf $\tilde{\psi}(s, t)$ is the unique solution to an ordinary differential equation. Under this Poisson arrival setting, we shall derive closed-form expressions for the the transient behavior of the first and second moments as well as the expectation of the workload. To begin, a closed-form expression for $b_j(t)$ in (16) is obtained. In this case, one finds that (15) becomes

$$\tilde{\psi}(\mathbf{0}, r) = \mathbb{E}(P^{Nr}) = e^{\lambda r(P-I)}, \quad r > 0, \tag{47}$$

whence $b_j(t)$ in (16) with (13) may be expressed as

$$\begin{aligned} b_j(t) &= \Delta_j \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - (t-y))} \mathbb{1}_{[L_j > t-y]} \right) P e^{\lambda(t-y)(P-I)} \lambda e^{-\lambda y} dy \\ &= \lambda \Delta_j e^{-\lambda t} \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - y)} \mathbb{1}_{[L_j > y]} \right) P e^{\lambda y(P-I)} dy. \end{aligned}$$

Furthermore, one checks easily that for all $t \geq 0$

$$b'_j(t) + \lambda b_j(t) = \lambda \mathbb{E} \left(e^{-\alpha(L_j - t)} \mathbb{1}_{[L_j > t]} \right) \Delta_j P e^{\lambda t(P-I)}. \tag{48}$$

Theorem 14. *One has the exact expression for the first joint moment given by*

$$M_j(t) = \lambda e^{\lambda t(P-I)} \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - v)} \mathbb{1}_{[L_j > v]} \right) e^{-\lambda v(P-I)} \Delta_j P e^{\lambda v(P-I)} dv, \quad t \geq 0, \quad j = 1, \dots, k. \tag{49}$$

Proof. We aim at obtaining a differential equation satisfied by $M_j(t)$. Remember from (18) that it satisfies the renewal matrix equation, with $dF(y) = \lambda e^{-\lambda y} dy$ as well as a change of variable $t - y := y$,

$$M_j(t) = b_j(t) + P \int_0^t M_j(t - y) \lambda e^{-\lambda y} dy = b_j(t) + e^{-\lambda t} P \int_0^t M_j(y) \lambda e^{\lambda y} dy,$$

which, upon differentiation, and thanks to (48), leads to the first-order matrix differential equation

$$\begin{aligned} M_j'(t) &= b_j'(t) + \lambda b_j(t) - \lambda M_j(t) + \lambda P M_j(t) \\ &= \lambda \mathbb{E} \left(e^{-\alpha(L_j - t)} \mathbb{1}_{[L_j > t]} \right) \Delta_j P e^{\lambda t(P - I)} + \lambda(P - I) M_j(t) \end{aligned} \quad (50)$$

with initial condition $M_j(0) = 0$. The solution to (50) is given by (49). \square

Next, using (47), (17) with (13) and (14) in this case is given by

$$\begin{aligned} b_{jj'}(t) &= \Delta_j \Delta_{j'} \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - (t - y))} \mathbb{1}_{[L_j > t - y]} e^{-\alpha(L_{j'} - (t - y))} \mathbb{1}_{[L_{j'} > t - y]} \right) P e^{\lambda(t - y)(P - I)} \lambda e^{-\lambda y} dy \\ &\quad + \Delta_j \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - (t - y))} \mathbb{1}_{[L_j > t - y]} \right) P M_{j'}(t - y) \lambda e^{-\lambda y} dy \\ &\quad + \Delta_{j'} \int_0^t \mathbb{E} \left(e^{-\alpha(L_{j'} - (t - y))} \mathbb{1}_{[L_{j'} > t - y]} \right) P M_j(t - y) \lambda e^{-\lambda y} dy, \end{aligned}$$

and thus one finds the following relation

$$\begin{aligned} b_{jj'}'(t) + \lambda b_{jj'}(t) &= \lambda \mathbb{E} \left(e^{-\alpha(L_j - t)} \mathbb{1}_{[L_j > t]} e^{-\alpha(L_{j'} - t)} \mathbb{1}_{[L_{j'} > t]} \right) \Delta_j \Delta_{j'} P e^{\lambda t(P - I)} \\ &\quad + \lambda \mathbb{E} \left(e^{-\alpha(L_j - t)} \mathbb{1}_{[L_j > t]} \right) \Delta_j P M_{j'}(t) + \lambda \mathbb{E} \left(e^{-\alpha(L_{j'} - t)} \mathbb{1}_{[L_{j'} > t]} \right) \Delta_{j'} P M_j(t), \end{aligned} \quad (51)$$

where $M_j(t)$ is given in (49).

Theorem 15. *One has the exact expression for the second moment given by*

$$\begin{aligned} M_{jj'}(t) &= \lambda e^{\lambda t(P - I)} \int_0^t e^{-\lambda v(P - I)} \left\{ \mathbb{E} \left(e^{-\alpha(L_j - v)} \mathbb{1}_{[L_j > v]} e^{-\alpha(L_{j'} - v)} \mathbb{1}_{[L_{j'} > v]} \right) \Delta_j \Delta_{j'} P e^{\lambda v(P - I)} \right. \\ &\quad \left. + \mathbb{E} \left(e^{-\alpha(L_j - v)} \mathbb{1}_{[L_j > v]} \right) \Delta_j P M_{j'}(v) + \mathbb{E} \left(e^{-\alpha(L_{j'} - v)} \mathbb{1}_{[L_{j'} > v]} \right) \Delta_{j'} P M_j(v) \right\} dv, \quad t \geq 0, \end{aligned} \quad (52)$$

for $j, j' = 1, \dots, k$, where $M_j(v)$ and $M_{j'}(v)$ are given by (49) in Theorem 14.

Let us note that the structure of the expression of $M_{jj'}(t)$ is different according to whether $j = j'$ or $j \neq j'$, as $\mathbb{E}(e^{-\alpha(L_j - v)} \mathbb{1}_{[L_j > v]} e^{-\alpha(L_{j'} - v)} \mathbb{1}_{[L_{j'} > v]})$ is equal to $\mathbb{E}(e^{-2\alpha(L_j - v)} \mathbb{1}_{[L_j > v]})$ if $j = j'$, or $\mathbb{E}(e^{-\alpha(L_j - v)} \mathbb{1}_{[L_j > v]}) \mathbb{E}(e^{-\alpha(L_{j'} - v)} \mathbb{1}_{[L_{j'} > v]})$ when $j \neq j'$, by independence.

Proof. Similar to the proof of Theorem 14, we write the renewal equation (19) satisfied by $M_{jj'}(t)$ as $M_{jj'}(t) = b_{jj'}(t) + P \int_0^t M_{jj'}(t - y) \lambda e^{-\lambda y} dy$, $t \geq 0$. The same first-order differential equation analysis with the expression of $b_{jj'}'(t) + \lambda b_{jj'}(t)$ given in (51) yields thus the explicit expression (52). \square

Finally, a transient expression for the expectation of the workload has the same structure as the first moment, and the following result is given without proof:

Theorem 16. *One has the exact expression for the expectation of the workload given by*

$$W_j(t) = \lambda e^{\lambda t(P-I)} \int_0^t \mathbb{E} \left((L_j - v) \mathbb{1}_{[L_j > v]} \right) e^{-\lambda v(P-I)} \Delta_j P e^{\lambda v(P-I)} dv, \quad t \geq 0, \quad j = 1, \dots, k.$$

5. Moment generating function for deterministic interarrival times

So far, it has been shown that the (transient or limiting) distribution of process $\tilde{Z}(t)$ is hard to study explicitly in general, except for the Poisson arrivals. Hence we shall consider a specific distribution for the interarrival times being deterministic, and equal to 1 without loss of generality to obtain some results on the mgf $\tilde{\psi}(s, t)$.

Theorem 17. *Suppose that $\tau = 1$ a.s., then $(\tilde{Z}(t) = \tilde{Z}(t; \alpha), X_{N_t})$ has a closed-form expression for the mgf given by*

$$\tilde{\psi}(s, t) = \tilde{\psi}(s, t; \alpha) = \left[\prod_{m=0}^{t-1} \tilde{Q}(s, m) \right]' = \prod_{m=1}^t \tilde{Q}(s, t - m)', \quad t \in \mathbb{N}, \quad (53)$$

where $\tilde{\pi}(s, t)$ and $\tilde{Q}(s, t)$ are given in (2) and (3) respectively. Besides, when $\mathbb{E}(L_j)$ is finite for all $j = 1, \dots, k$, $\lim_{t \rightarrow \infty} \prod_{m=0}^t \tilde{Q}(s, m) = \prod_{m=0}^{\infty} \tilde{Q}(s, m)$ exists, and $(\tilde{Z}(t), X_{N_t})$ converges in distribution as $t \rightarrow \infty$ given $X_0 = x$ towards $(\mathcal{Z}_\infty, \mathcal{X}_\infty) \in \mathbb{R}^k \times \mathcal{S}$ with the joint mgf given by

$$\tilde{\psi}_\infty(s) = \tilde{\psi}_\infty(s; \alpha) = [\mathbb{E} (e^{\langle s, \mathcal{Z}_\infty \rangle} \mathbb{1}_{[\mathcal{X}_\infty^x = y]})]_{(x,y) \in \mathcal{S}^2} = \left[\prod_{m=0}^{\infty} \tilde{Q}(s, m) \right]', \quad s \in \mathbb{R}^k. \quad (54)$$

Proof. Since $T_m = m \in \mathbb{N}$ and $N_t = t \in \mathbb{N}$, (53) is a straightforward application of (8) in Proposition 3.

We recall that $s = (s_1, \dots, s_k)$ may belong to the set $\mathbb{S} := \{(s_1, \dots, s_k) \in \mathbb{C}^k \mid s_j \in i\mathbb{R}, j = 1, \dots, k\}$ mentioned shortly after Definition (5), in such as $\tilde{\psi}(s, t)$ is the characteristic function of $\tilde{Z}(t)$ jointly to X_{N_t} . Then, in order to prove the convergence in distribution of $(\tilde{Z}(t), X_{N_t})$ given $X_0 = x$, it suffices by Lévy's convergence theorem to show that $\tilde{\psi}(s, t)$ in (53) converges towards $\tilde{\psi}_\infty(s)$ given in (54) for all $s \in \mathbb{S}$ and $\tilde{\psi}_\infty(s)$ is continuous at $s = \mathbf{0}$. This part constitutes the main bulk of the proof of the theorem. Let $\|\cdot\|$ be a submultiplicative norm on $\mathcal{S} \times \mathcal{S}$ matrices, i.e. such that $\|MN\| \leq \|M\| \cdot \|N\|$ for all matrices M and N . We write from (3) that

$$\tilde{Q}(s, m) = P' + (\tilde{\pi}(s, m) - I)P'.$$

To apply the result given in [1], we first introduce the two following norms defined respectively on complex valued and matrices valued sequences (defined similarly as in [1])

$$\|(u_m)_{m \in \mathbb{N}}\|_E := \sum_{m=0}^{\infty} |u_m|, \quad \|(M_m)_{m \in \mathbb{N}}\|_E = \sum_{m=0}^{\infty} \|M_m\|,$$

where $u_m \in \mathbb{R}$ and M_m is an $\mathcal{S} \times \mathcal{S}$ matrix for all $m \in \mathbb{N}$. Let us set $A_m = A_m(s) := (\tilde{\pi}(s, m) - I)P'$ (so as to comply with the notation of the latter paper). Since $(P')^m$ converges

towards $(1\pi)'$ as $m \rightarrow \infty$, from [1, Theorem 2.1] it is sufficient to prove that $\|(A_m)_{m \in \mathbb{N}}\|_E = \sum_{m=0}^{\infty} \|A_m\| < +\infty$ for the existence of $\prod_{m=0}^{\infty} \tilde{Q}(s, m)$. One has

$$\|A_m\| \leq \|\tilde{\pi}(s, m) - I\| \|P'\| \quad (55)$$

with $\tilde{\pi}(s, m) - I$ a diagonal matrix of which the (x, x) th component, $x \in \mathcal{S}$, is given from (2) by $\mathbb{E}(\exp(\sum_{j=1}^k s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]})) - 1$. Using the inequality $|e^u - 1| \leq e^{|u|} - 1$ for all $u \in \mathbb{C}$, and remembering that $x_j \in \{0, \dots, K\}$ is non negative for all $j = 1, \dots, k$, one finds for all $m \in \mathbb{N}$ that

$$\begin{aligned} & \left| \mathbb{E} \left(\exp \left\{ \sum_{j=1}^k s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right\} \right) - 1 \right| \leq \mathbb{E} \left(\left| \exp \left\{ \sum_{j=1}^k s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right\} - 1 \right| \right) \\ & \leq \mathbb{E} \left(\exp \left\{ \left| \sum_{j=1}^k s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right| \right\} - 1 \right) \leq \mathbb{E} \left(\exp \left\{ \sum_{j=1}^k |s_j| x_j \mathbb{1}_{[L_j > m]} \right\} \right) - 1 \\ & = \prod_{j=1}^k \left[1 + (e^{|s_j| x_j} - 1) \mathbb{P}(L_j > m) \right] - 1 = \sum_{I \subset \{1, \dots, k\}} \prod_{\ell \in I} \left[(e^{|s_\ell| x_\ell} - 1) \mathbb{P}(L_\ell > m) \right], \end{aligned} \quad (56)$$

where the independence of L_1, \dots, L_k was used. Note now that for all $I \subset \{1, \dots, k\}$,

$$\begin{aligned} \sum_{m=1}^{\infty} \prod_{\ell \in I} \left[(e^{|s_\ell| x_\ell} - 1) \mathbb{P}(L_\ell > m) \right] & \leq \left(e^{\max(|s_1|, \dots, |s_k|) \cdot K} - 1 \right)^k \sum_{m=1}^{\infty} \prod_{\ell \in I} \mathbb{P}(L_\ell > m) \\ & = \left(e^{\max(|s_1|, \dots, |s_k|) \cdot K} - 1 \right)^k \mathbb{E} \left(\max_{\ell \in I} L_\ell \right), \end{aligned}$$

which is finite thanks to the assumption that $\mathbb{E}(L_j) < +\infty$ for all $j = 1, \dots, k$. We thus deduce from (56) that $\sum_{m=0}^{\infty} \|\tilde{\pi}(s, m) - I\| < +\infty$ and in turn, from (55) $\|(A_m)_{m \in \mathbb{N}}\|_E = \sum_{m=0}^{\infty} \|A_m\| < +\infty$.

Now it remains to prove that $\psi_\infty(s)$ in (54) is continuous at $s = \mathbf{0}$. Let us first recall the inequality $|e^u - 1| \leq e|u|$ for all $u \in \mathbb{C}$ such that $|u| \leq 1$. If $x = (x_1, \dots, x_k)$ is in \mathcal{S} , this entails that, for all $j = 1, \dots, k$ and $m \in \mathbb{N}$,

$$\left| \exp \left(s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right) - 1 \right| \leq e |s_j| x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \quad (57)$$

for all s_j such that $|s_j| x_j \leq 1$, which is satisfied if $s = (s_1, \dots, s_k) \in [-i/K, i/K]^k \subset \mathbb{S}$. Letting $u_m^j(s_j) := \mathbb{E} \left(\exp \left(s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right) \right)$, we deduce from (57) that

$$\begin{aligned} |u_m^j(s_j) - 1| & \leq \mathbb{E}(e |s_j| x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]}) \leq e |s_j| x_j \mathbb{P}(L_j > m), \\ |u_m^j(s_j)| & \leq e |s_j| x_j \mathbb{P}(L_j > m) + 1 \leq e + 1 := M, \end{aligned} \quad (58)$$

for all $m \in \mathbb{N}$, $s_j \in [-i/K, i/K]$ and $j = 1, \dots, k$. Then it follows from (58) that for all $m \in \mathbb{N}$ and $s = (s_1, \dots, s_k) \in [-i/K, i/K]^k$, again by independence of L_1, \dots, L_k ,

$$\left| \mathbb{E} \left(\exp \left\{ \sum_{j=1}^k s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right\} \right) - 1 \right| = \left| \prod_{j=1}^k u_m^j(s_j) - 1 \right| = \left| \sum_{r=1}^k \left[\prod_{j=1}^{r-1} u_m^j(s_j) \right] [u_m^r(s_r) - 1] \right|$$

$$\leq \sum_{r=1}^k \prod_{j=1}^{r-1} |u_m^j(s_j)| |u_m^r(s_r) - 1| \leq \sum_{r=1}^k M^{r-1} e^{|s_r| x_r} \mathbb{P}(L_r > m),$$

which, summing from $m = 0$ to $+\infty$, yields the following bound for the $|\cdot|_E$ norm for all $x = (x_1, \dots, x_k) \in \mathcal{S}$

$$\left| \left(\mathbb{E} \left(\exp \left\{ \sum_{j=1}^k s_j x_j e^{-\alpha(L_j - m)} \mathbb{1}_{[L_j > m]} \right\} \right) - 1 \right) \right|_{m \in \mathbb{N}} \Big|_E \leq \sum_{r=1}^k M^{r-1} e^{|s_r| x_r} (\mathbb{E}(L_r) + 1)$$

and $s = (s_1, \dots, s_k) \in [-i/K, i/K]^k$. The right-hand side of the above inequality tends to 0 as $s \rightarrow \mathbf{0}$, $s \in [-i/K, i/K]^k \subset \mathbb{S}$, and is valid for all $x = (x_1, \dots, x_k) \in \mathcal{S}$. By the definition of matrices $\tilde{\pi}(s, m)$, $m \in \mathbb{N}$, this immediately implies that

$$\|(\tilde{\pi}(s, m) - I)_{m \in \mathbb{N}}\|_E \rightarrow 0, \quad s \rightarrow \mathbf{0}, \quad s \in \mathbb{S}.$$

One then deduces from [1, (2.20) in Corollary 2.1] that the infinite product $s \mapsto \tilde{\psi}_\infty(s) = \left[\prod_{m=0}^\infty \tilde{Q}(s, m) \right]' = \left[\prod_{m=0}^\infty (P' + A_m(s)) \right]'$ is continuous at $s = \mathbf{0}$, $s \in \mathbb{S}$. This completes the proof. \square

One interesting consequence of Theorem 17 is that the limiting mgf is expressed conveniently when the L_j 's are bounded by some constant M . In that case, one has from (2) and (3) that $\tilde{\pi}(s, r) = I$ and $\tilde{Q}(s, r) = P'$ when $r \geq M$, and we thus obtain the following result for this particular case:

Corollary 18. *Suppose that $\tau = 1$ a.s. and rvs L_j , $j = 1, \dots, k$, are all upper bounded such that $L_j \leq M$ a.s. for some deterministic $M \in \mathbb{N}^*$. Then the transient mgf in (53) simplifies as*

$$\tilde{\psi}(s, t) = \left[\prod_{m=0}^{M-1} \tilde{Q}(s, m) (P')^{t-M} \right]', \quad t \geq M,$$

and the limiting mgf is given by

$$\tilde{\psi}_\infty(s) = \mathbf{1}\pi \left[\prod_{m=0}^{M-1} \tilde{Q}(s, m) \right]'$$

6. Application: Infinite server queues modulated by an external semi-Markovian process

The model described in Section 2 is flexible enough to study the following process in queueing theory and actuarial science. We consider here a semi-Markov process $\{Y(t), t \geq 0\}$ with finite state space $\{1, \dots, \kappa\}$, jump times $(T_i)_{i \in \mathbb{N}}$ such that $(T_i - T_{i-1})_{i \in \mathbb{N}^*}$ is iid distributed as τ with cdf F , and the embedded Markov Chain $\{Y(T_n), n \in \mathbb{N}\}$ having transition matrix and stationary distribution denoted by $P_Y = (p_Y(\ell, m))_{\ell, m=1, \dots, \kappa}$ and $\pi_Y = (\pi_Y(\ell))_{\ell=1, \dots, \kappa}$ respectively. Let us suppose that $\{Y(t), t \geq 0\}$ models the arrival of customers or claims, such that the n th arriving customer has service time/delay $L_{n, (Y(T_{n-1}), Y(T_n))}^*$ where $(L_{n, (\ell, m)}^*)_{n \in \mathbb{N}, \ell, m=1, \dots, \kappa}$ is an iid sequence of rvs $L_{n, (\ell, m)}^*$'s, $n \in \mathbb{N}$, $\ell, m = 1, \dots, \kappa$. We also let $(L_{(\ell, m)}^*)_{\ell, m=1, \dots, \kappa}$ a generic corresponding rv with LT denoted as $\mathcal{L}_{\ell, m}(u) = \mathbb{E}(e^{-uL_{(\ell, m)}^*})$ for $u \geq 0$. In other words, if N_t

denotes the number of clients arrived by time t , the N_t th customer has service time which depends on both states of the semi-Markov process at the switching time T_{N_t} and prior to this switching time (i.e. depending on both states $Y(T_{N_t})$ and $Y(T_{N_t}-)$), as illustrated in Figure 1. This model has potential applications in queueing theory where an incoming customer may inspect the state of the environment $Y(T_{N_t}-)$ before deciding to join the queue; in an actuarial setting, there are different reasons for the determination of reporting delay for the IBNR claims. In particular, in this case, this model allows random fluctuations in the underlying delay distribution influenced by external process. For example, policyholder may decide to delay the submission of claims under special circumstances such as the external environment process is in a particular state. Also note that some flexibility for this rv is available, for example one may have $\mathbb{P}(L_{(\ell,m)}^* = 0) > 0$, implying that a customer finding the environment in state ℓ before it switches to state m decides not to join the queue with positive probability.

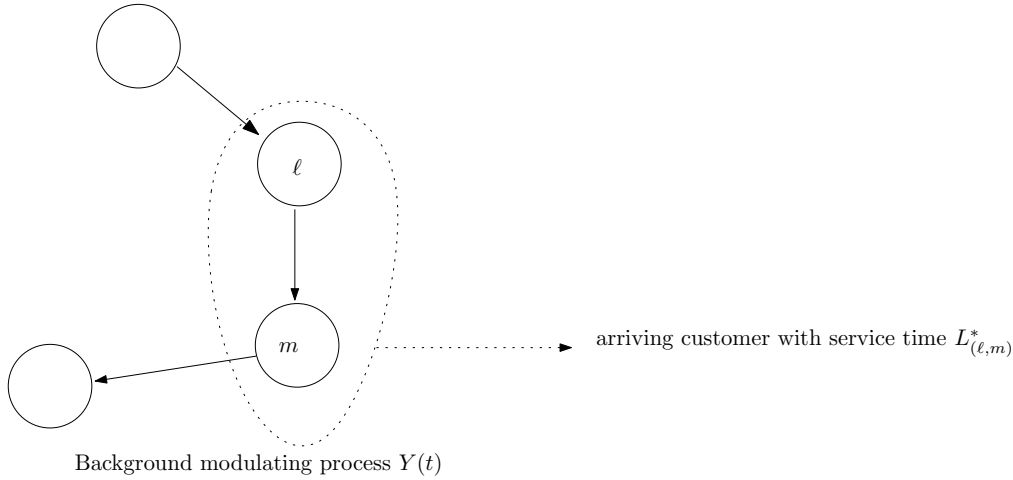


FIG 1. Modulating semi-Markov process and service time.

We then define the corresponding discounted processes $\{\mathcal{Z}(t), t \geq 0\}$ and $\{\tilde{\mathcal{Z}}(t), t \geq 0\}$ by

$$\mathcal{Z}(t) := \sum_{i=1}^{\infty} e^{-\alpha(T_i + L_{i,(Y(T_{i-1}), Y(T_i))}^*)} \mathbb{1}_{[T_i \leq t < T_i + L_{i,(Y(T_{i-1}), Y(T_i))}^*]}, \quad \tilde{\mathcal{Z}}(t) = e^{\alpha t} \mathcal{Z}(t). \quad (59)$$

As such, the process defined in (59) is different from the one introduced in (1) because the arrival times and the service times are now modulated by some semi-Markov process. However in the following, we shall show that (59) is actually embedded into (1), and this embedding procedure is essentially the central point of this section. Note in particular that this includes the particular case where $\{Y(t), t \geq 0\}$ is a continuous-time Markov Chain by considering $\tau \sim \mathcal{E}(\lambda)$ for some $\lambda > 0$, of which infinitesimal generating matrix is given by $Q_Y = (q_Y(\ell, m))_{\ell, m=1, \dots, \kappa}$ with $q_Y(\ell, m) = \lambda p_Y(\ell, m)$ if $\ell \neq m$ and $q_Y(\ell, \ell) = -\lambda \sum_{m \neq \ell} p_Y(\ell, m)$. In that case, inter-arrivals may not be identically distributed by considering the generic rv $(L_{(\ell, m)}^*)_{\ell, m=1, \dots, \kappa}$ to satisfy $L_{(\ell, \ell)}^* = 0$ for all $\ell = 1, \dots, \kappa$, so that a new customer actually arrives exactly at each time when $Y(t)$ switches to a new state, with service time distributed as $L_{(\ell, m)}^*$ when switching from state ℓ to state m .

So, we need to define the corresponding Markov Chain $(X_i)_{i \in \mathbb{N}}$. Let us introduce for all ℓ and m in $\{1, \dots, \kappa\}$ the $\kappa \times \kappa$ matrix $e(\ell, m)$ of which the (j, j') th entry is $\delta_{(j, j'), (\ell, m)}$. We then

define the state space of $(X_i)_{i \in \mathbb{N}}$ as

$$\mathcal{S} = \{e(\ell, m), (\ell, m) \in \{1, \dots, \kappa\}^2\} \subset \{0, 1\}^{\kappa \times \kappa}$$

so that one sets $k = \kappa^2$ and $K = 1$, sticking with the notation in Section 2. Then for all $i \in \mathbb{N}$, $X_i = (X_{i,(j,j')})_{(j,j') \in \{1, \dots, \kappa\}^2}$ is of the form $X_i = e(\ell, m)$ for some ℓ and m in $\{1, \dots, \kappa\}$, in which case one has

$$X_{i,(j,j')} = \delta_{(j,j'),(\ell,m)}, \quad \forall (j, j') \in \{1, \dots, \kappa\}^2.$$

The only difference here is in \mathcal{S} which is a strict subset of $\{0, 1\}^{\kappa \times \kappa}$, however this will not raise any additional technical difficulty in the following analysis. We then define the corresponding transition matrix as $P = (p(x, x'))_{(x, x') \in \mathcal{S}^2}$ with

$$p(e(\ell, m), e(\ell', m')) = \begin{cases} p_Y(\ell', m') & \text{if } m = \ell', \\ 0 & \text{otherwise,} \end{cases} \quad \ell, \ell', m, m' = 1, \dots, \kappa, \quad (60)$$

so that a transition from state $e(\ell, m)$ to state $e(\ell', m')$ of the Markov Chain $(X_i)_{i \in \mathbb{N}}$ is only possible if $m = \ell' \iff e(\ell, m)e(\ell', m') = e(\ell, m')$. One checks that (60) defines a proper transition matrix, i.e. the sum over each row is equal to 1, and $(X_i)_{i \in \mathbb{N}}$ is stationary and ergodic iff $\{Y(T_n), n \in \mathbb{N}\}$ is, with corresponding stationary distribution $(\pi(x))_{x \in \mathcal{S}}$ given by

$$\pi(x) = \pi(e(\ell, m)) = p_Y(\ell, m)\pi_Y(\ell), \quad \forall x = e(\ell, m) \in \mathcal{S}, \quad \ell, m = 1, \dots, \kappa. \quad (61)$$

Finally, we let $(L_{i,(j,j')})_{i \in \mathbb{N}, (j,j') \in \{1, \dots, \kappa\}^2}$ a sequence of independent rvs with corresponding distribution given by

$$L_{i,(j,j')} \sim L_{i,(j,j')}^*, \quad \forall (j, j') \in \{1, \dots, \kappa\}^2.$$

We now arrive at the embedding result. We let $\{Z(t) = Z(t; \alpha) = (Z_{(j,j')}(t))_{(j,j') \in \{1, \dots, \kappa\}^2} \in \mathbb{R}^{\{1, \dots, \kappa\}^2}, t \geq 0\}$ and $\tilde{Z}(t) = e^{\alpha t} Z(t)$ defined by (1), i.e.

$$Z_{(j,j')}(t) = \sum_{i=1}^{\infty} X_{i,(j,j')} e^{-\alpha(L_{i,(j,j')} + T_i)} \mathbb{1}_{[T_i \leq t < T_i + L_{i,(j,j')}]}, \quad (j, j') \in \{1, \dots, \kappa\}^2.$$

Then, one checks immediately that the following relation between $\tilde{Z}(t)$ and $\tilde{\mathcal{Z}}(t)$ defined in (59) holds

$$\left\{ \tilde{Z}_{(j,j')}(t), t \geq 0 \right\} \stackrel{D}{=} \left\{ \tilde{\mathcal{Z}}(t) \mathbb{1}_{[Y(T_{N_t-})=j, Y(T_{N_t})=j']}, t \geq 0 \right\}, \quad \forall (j, j') \in \{1, \dots, \kappa\}^2. \quad (62)$$

We remark that the above relation is interesting as it enables us to study $(\tilde{\mathcal{Z}}(t), Y(T_{N_t-}), Y(T_{N_t}))$ through $\tilde{Z}(t)$ i.e. through the analysis developed in Section 3. More precisely, one checks that

$$\begin{aligned} & \left(\mathcal{M}_{(j_0, j_1), (j_2, j_3)}^{(1)}(t) \right)_{(j_0, j_1) \in \{1, \dots, \kappa\}^2, (j_2, j_3) \in \{1, \dots, \kappa\}^2} \\ & := \left(\mathbb{E} \left(\tilde{\mathcal{Z}}(t) \mathbb{1}_{[Y(T_{N_t-})=j_2, Y(T_{N_t})=j_3]} \middle| Y(T_{-1}) = j_0, Y(T_0) = j_1 \right) \right)_{(j_0, j_1) \in \{1, \dots, \kappa\}^2, (j_2, j_3) \in \{1, \dots, \kappa\}^2} \\ & = \left(\mathbb{E} \left(\tilde{Z}_{(j_2, j_3)}(t) \middle| X_0 = e(j_0, j_1) \right) \right)_{(j_0, j_1) \in \{1, \dots, \kappa\}^2, (j_2, j_3) \in \{1, \dots, \kappa\}^2} \\ & = \left(M_{(j_2, j_3)}(t) \mathbf{1} \right)_{(j_2, j_3) \in \{1, \dots, \kappa\}^2} \end{aligned} \quad (63)$$

where $M_j(t)$ is defined in (6). Similarly, one can consider for all j_0, j_1, j_2 and j_3 the second moment $\mathcal{M}_{(j_0,j_1),(j_2,j_3)}^{(2)}(t) := \mathbb{E} \left(\tilde{\mathcal{Z}}(t)^2 \mathbb{1}_{[Y(T_{N_t}-)=j_2, Y(T_{N_t})=j_3]} \middle| Y(T_{-1}) = j_0, Y(T_0) = j_1 \right)$, which verifies

$$\left(\mathcal{M}_{(j_0,j_1),(j_2,j_3)}^{(2)}(t) \right)_{(j_0,j_1) \in \{1, \dots, \kappa\}^2, (j_2,j_3) \in \{1, \dots, \kappa\}^2} = \left(M_{(j_2,j_3),(j_2,j_3)}(t) \mathbf{1} \right)_{(j_2,j_3) \in \{1, \dots, \kappa\}^2}, \quad (64)$$

where $M_{j,j'}(t)$ is also defined in (6). Also, for the workload $\mathcal{D}(t)$ defined as

$$\mathcal{D}(t) := \sum_{i=1}^{\infty} \left(T_i + L_{i,(Y(T_{i-1}), Y(T_i))}^* - t \right) \mathbb{1}_{[T_i \leq t < T_i + L_{i,(Y(T_{i-1}), Y(T_i))}^*]}$$

(which is related with $D(t)$ in Section 3.3 as similar to (62), i.e. its expectation jointly to $(Y(T_{N_t}-), Y(T_{N_t}))$ is given by

$$\begin{aligned} & \left(\mathcal{W}_{(j_0,j_1),(j_2,j_3)}(t) \right)_{(j_0,j_1) \in \{1, \dots, \kappa\}^2, (j_2,j_3) \in \{1, \dots, \kappa\}^2} \\ := & \left(\mathbb{E} \left(\mathcal{D}(t) \mathbb{1}_{[Y(T_{N_t}-)=j_2, Y(T_{N_t})=j_3]} \middle| Y(T_{-1}) = j_0, Y(T_0) = j_1 \right) \right)_{(j_0,j_1) \in \{1, \dots, \kappa\}^2, (j_2,j_3) \in \{1, \dots, \kappa\}^2} \\ = & \left(W_{(j_2,j_3)}(t) \mathbf{1} \right)_{(j_2,j_3) \in \{1, \dots, \kappa\}^2}, \end{aligned}$$

where $W_j(t)$ is defined in (25).

Furthermore, one could find the following relation between the joint mgf of $\tilde{\mathcal{Z}}(t)$ and distribution of $(Y(T_{N_t}-), Y(T_{N_t}))$ defined for all $z \in \mathbb{R}$ and $t \geq 0$ by

$$\left[\tilde{\Psi}(z, t) \right]_{(j_0,j_1),(j_2,j_3)} = \mathbb{E} \left(e^{z \tilde{\mathcal{Z}}(t)} \mathbb{1}_{[Y(T_{N_t}-)=j_2, Y(T_{N_t})=j_3]} \middle| Y(T_{-1}) = j_0, Y(T_0) = j_1 \right),$$

where j_0, j_1, j_2 and j_3 are in $\{1, \dots, \kappa\}$. Then one notices that the above mgf is linked to the joint mgf of $\tilde{\psi}(s, t)$ of $\left\{ (\tilde{Z}_{(j,j')}(t))_{j,j'=1, \dots, \kappa}, t \geq 0 \right\}$ thanks to (62) by the relation

$$\left[\tilde{\Psi}(z, t) \right]_{(j_0,j_1),(j_2,j_3)} = \left[\tilde{\psi}(z.e(j_2, j_3), t) \right]_{(j_0,j_1),(j_2,j_3)}, \quad (65)$$

where we recall that $e(j_2, j_3)$ is the $\{1, \dots, \kappa\} \times \{1, \dots, \kappa\}$ matrix of which the (j, j') th entry is $\delta_{(j,j'),(j_2,j_3)}$.

The embedding relations (63), (64) and (65) thus enable us to derive results on the present model with semi-Markovian input, in particular thanks to Sections 3.2 and 4. Let us present some of them in the following. For example, (63) together with Proposition 7 yields the asymptotic first moment given by

$$\begin{aligned} & \left(\mathcal{M}_{(j_0,j_1),(j_2,j_3)}^{(1)}(t) \right)_{(j_0,j_1) \in \{1, \dots, \kappa\}^2, (j_2,j_3) \in \{1, \dots, \kappa\}^2} \\ & \rightarrow \left(\frac{1}{\mathbb{E}(\tau)} \frac{1 - \mathcal{L}_{(j_2,j_3)}(\alpha)}{\alpha} \mathbb{E}(X_{(j_2,j_3)}) \mathbf{1} \right)_{(j_2,j_3) \in \{1, \dots, \kappa\}^2}, \end{aligned}$$

as $t \rightarrow \infty$, where $\mathbb{E}(X_{(j_2,j_3)}) = \pi(e(j_2, j_3)) = p_Y(j_2, j_3) \pi_Y(j_2)$ by (61). When service times are exponentially distributed with $L_{(j_2,j_3)}^* \sim \mathcal{E}(\mu_{(j_2,j_3)})$, then (64) and Theorem 11 results in

$$\begin{aligned} & \left(\mathcal{M}_{(j_0, j_1), (j_2, j_3)}^{(2)}(t) \right)_{(j_0, j_1) \in \{1, \dots, \kappa\}^2, (j_2, j_3) \in \{1, \dots, \kappa\}^2} \longrightarrow \\ & \frac{1}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}^\tau(\mu_{(j_2, j_3)})}{\mu_{(j_2, j_3)} + 2\alpha} \right] \mathbf{1} \pi \Delta_{(j_2, j_3)}^2 P (I - \mathcal{L}^\tau(\mu_{(j_2, j_3)})P)^{-1} \\ & + \frac{2\mathcal{L}^\tau(\mu_{(j_2, j_3)})}{\mathbb{E}(\tau)} \left[\frac{1 - \mathcal{L}^\tau(2\mu_{(j_2, j_3)})}{2\mu_{(j_2, j_3)}} \right] \left(\frac{\mu_{(j_2, j_3)}}{\mu_{(j_2, j_3)} + \alpha} \right)^2 \mathbf{1} \pi \Delta_{(j_2, j_3)} P (I - \mathcal{L}^\tau(\mu_{(j_2, j_3)})P)^{-1} \\ & \quad \cdot \Delta_{(j_2, j_3)} P (I - \mathcal{L}^\tau(2\mu_{(j_2, j_3)})P)^{-1}, \end{aligned}$$

$t \rightarrow \infty$. When $\tau \sim \mathcal{E}(\lambda)$, i.e. when arrivals occur according to a Poisson process and the model is Markov modulated, the transient moment is explicit thanks to (49) in Theorem 14 and one computes easily for all j_2 and j_3 , using $PP'\Delta_\pi \mathbf{1} = \mathbf{1}$, $e^{\lambda v P} \mathbf{1} = e^{\lambda v} \mathbf{1}$, that

$$\begin{aligned} & \left(\mathcal{M}_{(j_0, j_1), (j_2, j_3)}^{(1)}(t) \right)_{(j_0, j_1) \in \{1, \dots, \kappa\}^2} = M_{(j_2, j_3)}(t) \mathbf{1} \\ & = \lambda e^{\lambda t(P-I)} \int_0^t \mathbb{E} \left(e^{-\alpha(L_j - v)} \mathbb{1}_{[L_j > v]} \right) e^{-\lambda v(P-I)} \Delta_j P e^{\lambda v(P-I)} dv \mathbf{1}, \quad t \geq 0. \end{aligned}$$

Further, when interarrival are deterministic equal to 1, then (65) and Theorem 17 entail that the limiting joint mgf is given for all $(j_0, j_1), (j_2, j_3)$ in $\{1, \dots, \kappa\}^2$, by

$$\lim_{t \rightarrow \infty} [\tilde{\Psi}(z, t)]_{(j_0, j_1), (j_2, j_3)} = \left[\prod_{i=0}^{\infty} \tilde{Q}(z, e(j_2, j_3), i) \right]_{(j_0, j_1), (j_2, j_3)}'.$$

Acknowledgements

The authors would like to thank two anonymous reviewers for their helpful comments and suggestions. This work was supported by Joint Research Scheme France/Hong Kong Procore Hubert Curien grant No 35296, F-HKU710/15T, and the UNSW Business School 2018 International Research Collaboration Travel Funds. Also, Jae-Kyung Woo gratefully acknowledges the support from and 2019 Business School Research Grant.

References

- [1] Artzrouni, M. (1986) On the convergence of infinite products of matrices. *Linear Algebra and its Applications*, 74: 11–21.
- [2] Athreya, A.B. & Ramamurthy, K. (1976) Feller’s renewal theorem for systems of renewal equations. *Journal of the Indian Institute of Science*, 58: 437–459.
- [3] Blom, J., De Turck K., & Mandjes, M. (2017) Refined large deviations asymptotics for Markov-modulated infinite-server systems. *European Journal of Operation Research*, 259: 1036–1044.
- [4] Blom, J., Kella, O., Mandjes, M., & Thorsdottir, H. (2014) Markov-modulated infinite-server queues with general service times. *Queueing Systems*, 76: 403–424.
- [5] D’Auria, B. (2008) $M/M/\infty$ queues in semi-Markovian random environment. *Queueing Systems*, 58: 221–237.
- [6] Francq, C. & Gautier, A. (2004) Estimation of time-varying ARMA models with Markovian changes in regime. *Statistics and Probability Letters*, 70: 243–251.

- [7] Fralix, B.H. & Adan, I.J.B.F. (2009) An infinite-server queue influenced by a semi-Markovian environment. *Queueing Systems*, 61: 65–84.
- [8] Karlsson, J.-E. (1974) A stochastic model for time lag in reporting of claims *Journal of Applied Probability*, 11(2): 382–387.
- [9] Mandjes, M. & De Turck, K. (2016) Markov-modulated infinite-server queues driven by a common background process. *Stochastic Models*, 32(2): 206–232.
- [10] Masuyama, H. & Takine, T. (2002) Analysis of an infinite-server queue with batch markovian arrival streams. *Queueing Systems*, 42(3): 269–296.
- [11] O’Cinneide, C., & Purdue, P. (1986) The $M/M/\infty$ in a random environment. *Journal of Applied Probability*, 23: 175–184.
- [12] Rabehasaina, L. & Woo, J.-K. (2018) On a multivariate renewal-reward process involving time delays and discounting: Applications to IBNR process and infinite server queues. *Queueing Systems*, 90(3): 307–350.
- [13] Takács, L. (1962) *Introduction to the theory of queues*. Oxford University press.
- [14] Willmot, G.E. (1990) A queueing theoretic approach to the analysis of the claims payment process. *Transaction of Society of Actuaries*, 42: 447-497.
- [15] Willmot, G.E. & Drekić, S. (2001) On the transient analysis of the $M^X/M/\infty$ queue. *Operations Research Letters*, 28: 137-142.
- [16] Woo, J.-K. (2016) On multivariate discounted compound renewal sums with time-dependent claims in the presence of reporting/payment delays. *Insurance: Mathematics and Economic*, 70: 354-363.