



UvA-DARE (Digital Academic Repository)

Infinite-server systems with Coxian arrivals

Boxma, O.; Kella, O.; Mandjes, M.

DOI

[10.1007/s11134-019-09613-2](https://doi.org/10.1007/s11134-019-09613-2)

Publication date

2019

Document Version

Final published version

Published in

Queueing Systems

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Boxma, O., Kella, O., & Mandjes, M. (2019). Infinite-server systems with Coxian arrivals. *Queueing Systems*, 92(3-4), 233-255. <https://doi.org/10.1007/s11134-019-09613-2>

General rights


It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Infinite-server systems with Coxian arrivals

Onno Boxma^{1,2}  · Offer Kella³ · Michel Mandjes⁴

Received: 22 May 2018 / Revised: 3 April 2019 / Published online: 13 May 2019
© The Author(s) 2019

Abstract

We consider a network of infinite-server queues where the input process is a Cox process of the following form: The arrival rate is a vector-valued linear transform of a multivariate generalized (i.e., being driven by a subordinator rather than a compound Poisson process) shot-noise process. We first derive some distributional properties of the multivariate generalized shot-noise process. Then these are exploited to obtain the joint transform of the numbers of customers, at various time epochs, in a single infinite-server queue fed by the above-mentioned Cox process. We also obtain transforms pertaining to the joint stationary arrival rate and queue length processes (thus facilitating the analysis of the corresponding departure process), as well as their means and covariance structure. Finally, we extend to the setting of a network of infinite-server queues.

Keywords Coxian process · $M/G/\infty$ · Multivariate shot-noise process · Subordinator

Mathematics Subject Classification Primary: 60K25; Secondary: 90B15

✉ Onno Boxma
o.j.boxma@tue.nl

Offer Kella
offer.kella@gmail.com

Michel Mandjes
m.r.h.mandjes@uva.nl

¹ Eurandom, Eindhoven, The Netherlands

² Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

³ Department of Statistics, The Hebrew University of Jerusalem, 9190501 Jerusalem, Israel

⁴ Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

1 Introduction

In queueing theory it is commonly assumed that the customer arrival process is a Poisson process. Some recent empirical studies (see, for example, [4] for references) suggest, however, that arrival processes may exhibit *overdispersion*, i.e., the variance of the number of arrivals in an interval is larger than the corresponding mean. This has triggered research on queueing systems with overdispersed input.

The focus of the present paper is on infinite-server queues with, as input process, a doubly stochastic process, also known as a *Cox process*. Such a process can be seen as a Poisson process in which the rate is not a constant; instead, the rate process $\{\Lambda(t), t \in \mathbb{R}\}$ itself is a (nonnegative) stochastic process. As an immediate consequence of the law of total variance, Cox processes indeed are overdispersed. Before describing the main contributions of the present paper, we first provide a brief account of the existing literature in this area.

Literature In [7], the arrival process of the infinite-server queues is a Cox process in which the arrival rate is a *shot-noise process*. More specifically, the jumps of the shot-noise process occur according to a homogeneous Poisson process and are i.i.d. (independent and identically distributed) with a general distribution, and between jumps, the shot-noise process decays exponentially fast. The main object of study in [7] is a feedforward network of infinite-server queues, and the main result is the joint transform of the shot-noise-driven arrival rates and the numbers of customers in the queues.

Infinite-server queues are also studied in [2] and [8], but the arrival process there is a *Hawkes process*, the so-called self-exciting process—jumps in the shot-noise process in turn increase the intensity of the occurrence of the jumps. Daw and Pender [2] present several interesting motivating examples. They consider deterministic jump sizes in the shot-noise process and study in particular the Hawkes/Ph/ ∞ and Hawkes/D/ ∞ queues, obtaining detailed expressions for moments and autocovariances. Koops *et al.* [8] allow generally distributed jump sizes. For the case of exponentially distributed service times, the joint Laplace and z-transform of the Hawkes intensity and the number of customers is characterized via a partial differential equation, and that PDE is exploited to obtain recursive expressions for (joint) moments of the Hawkes intensity and the number of customers. For the case of generally distributed service times, the Hawkes process is viewed as a branching process, which allows expressing the z-transform of the number of customers in terms of the solution of a fixed-point equation.

Main goals and results In this paper we aim to develop a general framework for the study of infinite-server queues with, as input, a quite general Cox process and to provide an exact analysis of networks of such queues. The modeling framework of [7] is extended in several ways; in particular, the shot-noise process is multivariate and not driven by a compound Poisson process, but by a Lévy subordinator. The class of Lévy subordinators not only contains compound Poisson processes, but also, for example, linear drifts and Gamma processes (where the latter process has the notable property that it exhibits infinitely many jumps in any time interval of positive length).

The main results of the paper are compact, elegant, transform expressions for joint distributions of arrival rates and numbers of customers, and properties of the departure processes, which reveal an interesting generalization of the classical results for the

number-of-customers process and the departure process in the $M/G/\infty$ queue. While large parts of the paper are quite technical, we also try to make clear why a detailed analysis of this complicated network is possible. In a nutshell, crucial for the analysis are the following convenient properties: (i) Lévy processes have stationary and independent increments, (ii) in a shot-noise process with exponential decay, all shots simultaneously decay at the same exponential rate, independently of each other, and (iii) in (a network of) infinite-server queues, all customers move through the queues independently, without interfering with each other.

Motivation Our motivation for this study is threefold. (i) We started studying queues with Coxian input as a framework naturally allowing the incorporation of overdispersion. (ii) When we realized that a detailed analysis of an isolated infinite-server queue with Coxian input is possible, we were motivated to develop a very general framework for studying networks of infinite-server queues with a multivariate Coxian input process that is of great generality but still allows a detailed exact analysis. Together with product-form networks and linear stochastic fluid networks (cf. [6]) these are some of the rare examples of queueing networks for which one can obtain the joint distribution of key performance measures. (iii) Finally, while infinite-server queues with overdispersed input have various applications, we were in particular inspired by Web server applications. Consider the number of simultaneous visitors to a popular Web site or online video [11,12]. While the arrival process of visitors to the Web site may well be a Poisson process, its rate may jump up due to some (external) event, then decay gradually, only to jump up again because of another event. Such an example formed one of the motivations for [2,7,8], which all study infinite-server queues with an overdispersed arrival process.

Cox input processes driven by a *multivariate* Lévy subordinator are quite relevant for the above-mentioned example of a Web site with a visit rate that jumps up because of a certain event and subsequently decays again. They allow one to take into account *multiple*, possibly related, events that affect the visit rate of the Web site. Those events could occur instantaneously (the Lévy subordinator being a compound Poisson process), but there might also, for example, be a linear (fluid) component.

Since we also consider *networks* of infinite-server queues, we can also model the situation in which a visitor to a Web site, after a while, clicks in order to move to another Web site.

Organization of the paper In Sect. 2 we introduce the multivariate shot-noise process and we describe distributional properties of this process. Section 3 studies a single infinite-server queue with, as input process, the above-described Cox process. The general case of a network of infinite-server queues is dealt with in Sect. 4. Each of these three sections starts with a brief overview of related known results: the one-dimensional shot-noise process (Sect. 2.1), the classical $M/G/\infty$ queue (Sect. 3.1), and a tandem network of $M/G/\infty$ queues (Sect. 4.1). Section 5 contains some conclusions and suggestions for further research.

2 Multivariate shot-noise

In this preliminary section, we first briefly introduce shot-noise (Sect. 2.1) and define multivariate shot-noise and its stationary version (Sect. 2.2), then we recall some basic facts on Poisson random measures (Sect. 2.3), and we conclude in Sect. 2.4 with describing distributional properties of the multivariate shot-noise process.

2.1 Shot-noise processes

We start our exposition by describing the classical shot-noise process $\{X(t), t \geq 0\}$, as follows: This process increases according to a compound Poisson process $\{J(t), t \geq 0\}$ with jump rate c of the Poisson jump process $\{N(t), t \geq 0\}$ and independent, identically distributed upward jumps B_1, B_2, \dots with Laplace–Stieltjes transform $\beta(\cdot)$; see Fig. 1.

In between jumps, the process decreases with a state-dependent rate $rX(t)$. Possible representations are hence, with T_1, T_2, \dots denoting the jump epochs of the compound Poisson process,

$$X(t) = e^{-rt} X(0) + \int_0^t e^{-r(t-s)} dJ(s) = e^{-rt} X(0) + \sum_{j=1}^{N(t)} B_j e^{-r(t-T_j)}, \quad t \geq 0; \quad (1)$$

alternatively, $X(\cdot)$ is the unique solution of the stochastic integral equation

$$X(t) = X(0) + \sum_{j=1}^{N(t)} B_j - r \int_0^t X(s) ds, \quad t \geq 0. \quad (2)$$

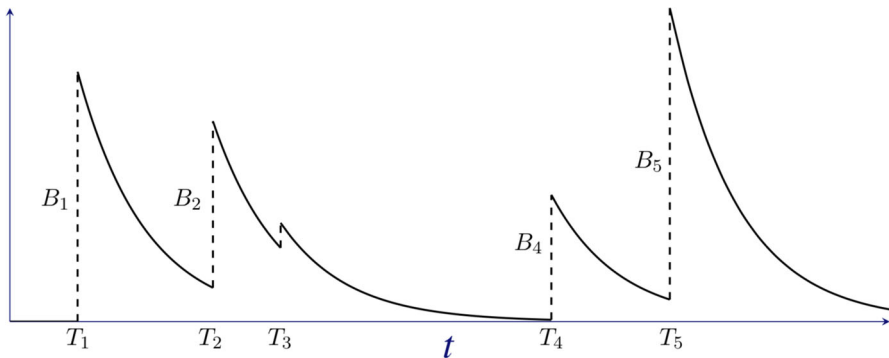


Fig. 1 The rate $X(t)$ as a function of time t

2.2 Definition of multivariate shot-noise

After having defined the classical shot-noise process, we now describe the generalized version of the shot-noise process that we work with in this paper.

Let $X(\cdot)$ be a (generalized) d -dimensional multivariate shot-noise process, which is defined as follows: Let $J(\cdot)$ be a d -dimensional subordinator, i.e., a d -dimensional Lévy process which is non-decreasing in all components. Examples of such non-decreasing Lévy processes include linear drifts, compound Poisson processes and Gamma processes (but, obviously, not Brownian motion). We define the exponent of $J(\cdot)$ by $-\eta(\cdot)$; it satisfies, for $\alpha \in \mathbb{R}_+^d$ and with $'$ denoting transposition,

$$\eta(\alpha) = -\log \left(\mathbb{E} e^{-\alpha' J(1)} \right) = \alpha' c + \int_0^\infty (1 - e^{-\alpha' x}) \nu(dx),$$

where $c \in \mathbb{R}_+^d$ and ν is an associated Lévy measure satisfying

$$\nu \left(\left(\mathbb{R}_+^d \right)^c \cup \{0\} \right) = 0 \quad \text{and} \quad \int_{\mathbb{R}_+^d} (\|x\| \wedge 1) \nu(dx) < \infty;$$

cf. [9, 13].

Also, let $Q = (q_{ij})$ be a $(d \times d)$ -matrix with nonnegative diagonal and non-positive off-diagonal entries, and with all eigenvalues having strictly positive real parts. An example of such a matrix is $Q = (I - P')D$, where D is a positive diagonal matrix and P is a substochastic matrix satisfying $P^n \rightarrow 0$ as $n \rightarrow \infty$. A detailed motivation for studying this setting is provided in [6, Section 4].

We now introduce, for $X(0)$ componentwise strictly positive and independent of $J(\cdot)$, the multivariate shot-noise process through

$$X(t) = e^{-Qt} X(0) + \int_{(0,t]} e^{-Q(t-s)} dJ(s); \tag{3}$$

cf. (1) and (2) above. Alternatively, $X(\cdot)$ is the unique (strong) solution to the stochastic integral equation

$$X(t) = X(0) + J(t) - Q \int_0^t X(s) ds. \tag{4}$$

The cumulative external input to station i is $X_i(0) + J_i(t)$, the internal input rate from station i to station j is $\sum_{j \neq i} q_{ij} X_j(t)$, and the output rate (some of which goes to other stations, while the rest leaves the system) is $q_i X_i(t)$, where $q_i = -q_{ii} \geq 0$. The reason for the above assumption that the eigenvalues of Q have a strictly positive real part is that it ensures that all the coordinates of the matrix e^{-Qt} vanish as $t \rightarrow \infty$ and thus, from (3), it follows that the effect of each arriving particle vanishes as well.

The process $X(\cdot)$ is Markovian and strictly positive (i.e., never hits or crosses any of the axes). Moreover, it has a unique stationary/ergodic/limiting distribution. This limiting distribution is identical to that of

$$X(\infty) = \int_{(0, \infty)} e^{-Qs} dJ(s). \quad (5)$$

So far we have considered $J(t)$ for positive values of t , but we can extend $J(\cdot)$ to the whole real line (with $J(0) = \mathbf{0}$). It follows directly that

$$X^*(t) := \int_{(-\infty, t]} e^{-Q(t-s)} dJ(s) \quad (6)$$

is a stationary process satisfying the same conditions that X does, that is, relations (3) and (4). From here on we assume that $X = X^*$.

In Sect. 3 we shall consider an infinite-server queue with as arrival process a non-homogeneous Poisson process with arrival rate function $\Lambda(t) = \mathbf{a}'X(t)$ at time t . But first, in the next subsection, we review some well-known properties and results for Poisson random measures. We need those properties and results in order to provide an elegant treatment of the transforms of the arrival process and numbers of customers in networks of infinite-server queues with Coxian arrivals. In Sect. 2.4 we already demonstrate that by deriving an expression for the d -dimensional Laplace–Stieltjes transform (LST) of \mathbf{X} .

2.3 Properties of Poisson random measures

Readers who are not familiar with Poisson random measures are referred to, for example, Chapter 6 of Çinlar [1]. To proceed, we first recall that a Poisson random measure N on some measurable space $(\mathbb{X}, \mathcal{G})$ is a random measure having the property that, for every pairwise disjoint $A_1, \dots, A_n \in \mathcal{G}$, the random variables $N(A_1), \dots, N(A_n)$ are independent with $N(A_i) \sim \text{Poisson}(\mu(A_i))$, where $\mu(A) = \mathbb{E}N(A)$ is assumed to be σ -finite. When $\mu(A)$ is 0 or ∞ then $N(A)$ is defined to be almost surely 0 or ∞ , respectively. We note that it is well known that N is a Poisson random measure with a (σ -finite) mean measure μ if and only if for any \mathcal{G} -measurable $f : \mathbb{X} \rightarrow \mathbb{R}_+$ we have that (cf. [1, Thm. 2.9 on p. 252]),

$$\mathbb{E} \exp \left(- \int f(s) dN(s) \right) = \exp \left(\int (1 - e^{-f(s)}) d\mu(s) \right).$$

Observe that $\mathbb{E} \int f(s) N(ds) = \int f(s) d\mu(s)$ (which holds even if N is not Poisson). Finally, it is also known, and actually easy to check (first for indicators, then simple functions, etc.), that if $f_1(\cdot)$ and $f_2(\cdot)$ are nonnegative \mathcal{G} -measurable, then

$$\mathbb{E} \int f_1(s) dN(s) \int f_2(s) dN(s) = \int f_1(s) f_2(s) d\mu(s) + \int f_1(s) d\mu(s) \int f_2(s) d\mu(s).$$

Clearly, when $\int f_1(s) d\mu(s)$, $\int f_2(s) d\mu(s)$, and $\int f_1(s) f_2(s) d\mu(s)$ are all finite, then this is equivalent to

$$\text{Cov} \left(\int f_1(s) dN(s), \int f_2(s) dN(s) \right) = \int f_1(s) f_2(s) d\mu(s).$$

Let $\varrho = (\varrho_i)$ be the mean rate of growth of the subordinator $\mathbf{J}(\cdot)$, that is, $\varrho_i := c_i + \int_{\mathbb{R}_+^d} x_i \nu(d\mathbf{x})$. We can also write $\varrho_i = c_i + \int_{\mathbb{R}_+^d} x_i \nu_i(d\mathbf{x}_i)$, using the notation

$$\nu_i(A) = \nu \left(\mathbb{R}_+^{i-1} \times A \times \mathbb{R}_+^{d-i} \right).$$

It is also well known (see, for example, Chapter 4 of [13]) that there exists a Poisson random measure $N(\cdot, \cdot)$ on $\mathbb{R} \times \mathbb{R}_+^d$ (equipped with its Borel σ -field) with mean measure $\ell \otimes \nu$, where ℓ is Lebesgue measure, such that

$$J_i(t) = \begin{cases} c_i t + \int_{(0,t] \times \mathbb{R}_+^d} x_i dN(s, \mathbf{x}), & t \geq 0, \\ c_i t - \int_{[t,0) \times \mathbb{R}_+^d} x_i dN(s, \mathbf{x}), & t < 0, \end{cases}$$

where $\mathbf{x} = (x_1, \dots, x_d)'$. This property entails that if we take $f_1(s, \mathbf{x}) = \sum_{i=1}^d g_i(s) x_i$ and $f_2(s, \mathbf{x}) = \sum_{i=1}^d h_i(s) x_i$, then we can immediately conclude that for each Borel $\mathbf{g}, \mathbf{h} : \mathbb{R} \rightarrow \mathbb{R}_+^d$ the following three identities hold:

$$\mathbb{E} \int_{\mathbb{R}} \mathbf{g}(s)' d\mathbf{J}(s) = \int_{\mathbb{R}} \mathbf{g}(s)' \varrho ds, \tag{7}$$

$$\mathbb{E} \exp \left(- \int_{\mathbb{R}} \mathbf{g}(s)' d\mathbf{J}(s) \right) = \exp \left(- \int_{\mathbb{R}} \eta(\mathbf{g}(s)) ds \right), \tag{8}$$

and

$$\text{Cov} \left(\int_{\mathbb{R}} \mathbf{g}(s)' d\mathbf{J}(s), \int_{\mathbb{R}} \mathbf{h}(s)' d\mathbf{J}(s) \right) = \int_{\mathbb{R}} \mathbf{g}(s)' \Sigma \mathbf{h}(s) ds,$$

with

$$\Sigma := -\nabla^2 \eta(\mathbf{0}) = \int_{\mathbb{R}_+^d} \mathbf{x} \mathbf{x}' \nu(d\mathbf{x}),$$

provided that the corresponding variances $\int_{\mathbb{R}} \mathbf{g}(s)' \Sigma \mathbf{g}(s) ds$ and $\int_{\mathbb{R}} \mathbf{h}(s)' \Sigma \mathbf{h}(s) ds$ are both finite. With σ_{ij} denoting the (i, j) th coordinate of Σ , it also holds that

$$\sigma_{ij} = \int_{\mathbb{R}_+^d} x_i x_j \nu(d\mathbf{x}) = \int_{\mathbb{R}_+^2} x_i x_j \nu_{ij}(dx_i, dx_j), \tag{9}$$

where ν_{ij} is the (i, j) th marginal measure associated with ν (i.e., the Lévy measure associated with the (i, j) th coordinate of $\mathbf{J}(\cdot)$).

2.4 Distributional properties of multivariate shot-noise

Combining (5) with the powerful identity (8), we immediately obtain the following

Proposition 2.1 *The d -dimensional Laplace–Stieltjes transform of \mathbf{X} is given by*

$$\mathbb{E} e^{-\boldsymbol{\alpha}' \mathbf{X}} = \exp \left(- \int_0^\infty \eta \left(e^{-Q's} \boldsymbol{\alpha} \right) ds \right). \tag{10}$$

In the usual manner, moments can be found from the LST. The first moment takes the following form: With $\nabla \eta(\mathbf{0})$ the vector of first derivatives,

$$\mathbb{E} \mathbf{X} = Q^{-1} \nabla \eta(\mathbf{0}) = Q^{-1} \boldsymbol{\rho}.$$

We now identify the covariance matrix of \mathbf{X} . From (9) it follows, as was also shown in [6, Thm. 5.2], that the covariance matrix of \mathbf{X} is given by

$$\Sigma_0 = \int_0^\infty e^{-Qs} \Sigma e^{-Q's} ds, \tag{11}$$

and is the unique solution of

$$Q \Sigma_0 + \Sigma_0 Q' = \Sigma. \tag{12}$$

The next objective is to find an expression for the covariance between $\mathbf{X}(t)$ and $\mathbf{X}(t + h)$, again bearing in mind that we started the process at $-\infty$. Now, clearly

$$\mathbf{Y}(t, h) := \int_{(t, t+h]} e^{-Q(t+h-s)} d\mathbf{J}(s)$$

is independent of $\mathbf{X}(t)$, due to the independent increments property of $\mathbf{J}(\cdot)$. As a consequence, with

$$\mathbf{Z}(t, h) := \int_{(-\infty, t]} e^{-Q(t+h-s)} d\mathbf{J}(s),$$

we obtain, by splitting $\mathbf{X}(t + h)$ into the sum of $\mathbf{Y}(t, h)$ and $\mathbf{Z}(t, h)$ (cf. (6)), that

$$\text{Cov}(\mathbf{X}(t), \mathbf{X}(t + h)) = \text{Cov}(\mathbf{X}(t), \mathbf{Y}(t, h) + \mathbf{Z}(t, h)) = \text{Cov}(\mathbf{X}(t), \mathbf{Z}(t, h)).$$

Denote (for $h \geq 0$) by Σ_h a matrix of which the (i, j) th entry is $\text{Cov}(X_i(t), X_j(t + h))$. It now follows from (9) that, by taking $\mathbf{g}(s) := e^{-Q'(t-s)}\mathbf{x}$ and $\mathbf{h}(s) := e^{-Q'(t+h-s)}\mathbf{y}$,

$$\mathbf{x}'\Sigma_h\mathbf{y} = \int_{-\infty}^t \mathbf{x}'e^{-Q(t-s)}\Sigma e^{-Q'(t+h-s)}\mathbf{y} ds = \int_0^\infty \mathbf{x}'e^{-Qs}\Sigma e^{-Q'(s+h)}\mathbf{y} ds,$$

for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. It thus follows, using (11), that

$$\Sigma_h = \int_0^\infty e^{-Qs}\Sigma e^{-Q'(s+h)} ds = \Sigma_0 e^{-Q'h}. \tag{13}$$

Combining the above, we find the following result.

Proposition 2.2 *The $(\mathbb{R}^{2d}$ -valued) covariance matrix of $(\mathbf{X}(t), \mathbf{X}(t + h))$ is given by*

$$\begin{pmatrix} \Sigma_0 & \Sigma_0 e^{-Q'h} \\ e^{-Qh}\Sigma_0 & \Sigma_0 \end{pmatrix}.$$

Since Q' and $e^{-Q'h}$ commute, it also follows from (12) that

$$Q\Sigma_h + \Sigma_h Q' = \Sigma e^{-Q'h}.$$

In fact, employing the method of proof of [6, Thm. 5.2], Σ_h is the unique solution of this equation.

Remark 2.3 In the one-dimensional case, letting σ^2 and q denote the one-dimensional versions of the matrices Σ and Q , respectively, we conclude that the covariance between $X(t)$ and $X(t + h)$ is given by $\frac{1}{2}e^{-qh}\sigma^2/q$.

3 Single infinite-server queue with a Cox input process

We consider an infinite-server queue in which, conditioned on $\mathbf{J}(\cdot)$, the arrival process is a non-homogeneous Poisson process with rate function $\Lambda(t) = \mathbf{a}'\mathbf{X}(t)$ at time t , where $\mathbf{a} \in \mathbb{R}_+^d$. It is throughout assumed that service times are i.i.d. and are independent of $\mathbf{J}(\cdot)$ (hence also of $\mathbf{X}(\cdot)$) and the arrival process. The service times have distribution function $F(\cdot)$ and complementary distribution function $\bar{F}(\cdot)$, and in addition we define $F(s, t) := F(t) - F(s)$ for $s < t$.

In this section, we start by reminding the reader of some well-known results concerning the classical M/G/∞ queue (Sect. 3.1), subsequently we derive, for the infinite-server queue with a Cox input process, the transform of queue lengths at various time epochs jointly with the numbers of arrivals in various time intervals (Sect. 3.2),

and we conclude, in Sect. 3.3, with several calculations yielding moments as well as the joint steady-state transform of the number of customers in the infinite-server queue and the arrival rate.

3.1 The M/G/∞ queue

In this subsection we consider the classical M/G/∞ queue with fixed arrival rate λ and generic service time B . It is well known [3] that the number of customers $L(t)$ at time t , starting with $L(0) = 0$, is Poisson distributed with mean

$$\lambda t \int_0^t \mathbb{P}(B > t - u) \frac{du}{t} = \lambda \int_0^t \mathbb{P}(B > u) du.$$

This is easily seen by observing that the number of arrivals in $[0, t]$ is $\text{Poisson}(\lambda t)$, that different customers do not interact and that, given that there were n arrivals in $[0, t]$, the arrival epochs were independent and uniformly distributed on $[0, t]$. So the Poisson arrival process is thinned, and the number of customers still present at t follows a Poisson distribution. Thus, the generating function of $L(t)$ immediately follows:

$$\mathbb{E}[w^{L(t)} | L(0) = 0] = \exp \left((w - 1) \lambda \int_0^t \mathbb{P}(B > u) du \right). \tag{14}$$

The pleasing properties of non-interfering customer behavior and of the uniformly distributed arrival epochs also quickly lead to elegant results for joint distributions of customer numbers at various epochs. Another well-known result about the M/G/∞ queue, going back to Mirasol [10], is that in stationarity the departure process of the queue also is Poisson. We now study the question of how these results change in the setting of the Cox input process.

3.2 Transform of queue lengths and numbers of arrivals

Our first objective is to derive the transform of queue lengths at various points in time, jointly with the number of arrivals in corresponding intervals.

To this end, we consider $n \in \mathbb{N}$ time intervals, say $(t_0, t_1]$ up to $(t_{n-1}, t_n]$ (where it is assumed that $t_{i-1} < t_i$ for $i = 1, \dots, n$). Our goal is to establish the joint transform of the queue lengths L_i at each of the t_i and the numbers of arrivals A_i in each of the intervals $(t_{i-1}, t_i]$, i.e., we shall compute the transform, for $\mathbf{w} \in \mathbb{R}^{n+1}$ and $\mathbf{z} \in \mathbb{R}^n$,

$$\Psi(\mathbf{w}, \mathbf{z}) := \mathbb{E} \left(\prod_{i=0}^n w_i^{L_i} \cdot \prod_{i=1}^n z_i^{A_i} \right).$$

Notice that a job arriving in the interval $(t_{i-1}, t_i]$ contributes to A_i (and does not contribute to any of the other A_j) and potentially contributes to L_i up to L_n . More precisely, supposing that the job arrives at $s \in (t_{i-1}, t_i]$, with probability $F(t_j -$

$s, t_{j+1} - s$) it contributes to L_i up to L_j , for $j \in \{i, \dots, n\}$ and defining $t_{n+1} := \infty$. Conditional on $\Lambda(\cdot) = \lambda(\cdot)$ this concerns a Poisson contribution with parameter

$$\int_{t_{i-1}}^{t_i} \lambda(s)F(t_j - s, t_{j+1} - s) ds;$$

conditional on $\Lambda(\cdot) = \lambda(\cdot)$ all these contributions are independent. In addition, we recall that the probability-generating function (evaluated in z) of a Poisson random variable with parameter λ equals $\exp(\lambda(z - 1))$. Combining the above elements, we obtain

$$\Psi(\mathbf{w}, \mathbf{z}) = \mathbb{E} \exp \left(\int_{-\infty}^{t_n} \Lambda(s)\psi(s | \mathbf{w}, \mathbf{z}) ds \right), \tag{15}$$

where, with $t_{-1} := -\infty$,

$$\psi(s | \mathbf{w}, \mathbf{z}) = \sum_{i=0}^n \psi_i(s | \mathbf{w}, \mathbf{z}) \mathbf{1}_{\{s \in (t_{i-1}, t_i]\}};$$

the individual $\psi_i(s | \mathbf{w}, \mathbf{z})$ are defined by

$$\begin{aligned} \psi_0(s | \mathbf{w}, \mathbf{z}) &:= \sum_{j=0}^n F(t_j - s, t_{j+1} - s) \left(\prod_{i=0}^j w_i - 1 \right) \\ &= \sum_{j=0}^n F(t_j - s, t_{j+1} - s) \prod_{i=0}^j w_i + F(t_0 - s) - 1 \end{aligned}$$

for $i = 0$, and

$$\begin{aligned} \psi_i(s | \mathbf{w}, \mathbf{z}) &:= \sum_{j=i}^n F(t_j - s, t_{j+1} - s) \left(z_i \prod_{k=i}^j w_k - 1 \right) \\ &= \sum_{j=i}^n F(t_j - s, t_{j+1} - s) z_i \prod_{k=i}^j w_k + F(t_i - s) - 1 \end{aligned}$$

for $i \in \{1, \dots, n\}$.

Representation (15) generally holds, i.e., for any nonnegative arrival rate process $\Lambda(\cdot)$. For the case of multivariate shot-noise, however, the expression can be made more explicit. To this end, we substitute (cf. (6))

$$\Lambda(s) = \mathbf{a}' \mathbf{X}(s) = \mathbf{a}' \int_{-\infty}^s e^{-\mathcal{Q}(s-r)} d\mathbf{J}(r).$$

By applying Eq. (10), we obtain that (15) equals

$$\begin{aligned} \Psi(\mathbf{w}, \mathbf{z}) &= \mathbb{E} \exp \left(\mathbf{a}' \int_{-\infty}^{t_n} \int_{-\infty}^s e^{-Q(s-r)} d\mathbf{J}(r) \psi(s | \mathbf{w}, \mathbf{z}) ds \right) \\ &= \mathbb{E} \exp \left(\mathbf{a}' \int_{-\infty}^{t_n} \int_r^{t_n} e^{-Q(s-r)} \psi(s | \mathbf{w}, \mathbf{z}) ds d\mathbf{J}(r) \right) \\ &= \exp \left(- \int_{-\infty}^{t_n} \eta \left(- \int_r^{t_n} e^{-Q'(s-r)} \psi(s | \mathbf{w}, \mathbf{z}) ds \mathbf{a} \right) dr \right). \end{aligned}$$

We have thus established the following result.

Theorem 3.1 For any $\mathbf{w} \in \mathbb{R}^{n+1}$ and $\mathbf{z} \in \mathbb{R}^n$,

$$\begin{aligned} \Psi(\mathbf{w}, \mathbf{z}) &= \mathbb{E} \left(\prod_{i=0}^n w_i^{L_i} \cdot \prod_{i=1}^n z_i^{A_i} \right) \\ &= \exp \left(- \int_{-\infty}^{t_n} \eta \left(- \int_r^{t_n} e^{-Q'(s-r)} \psi(s | \mathbf{w}, \mathbf{z}) ds \mathbf{a} \right) dr \right). \end{aligned}$$

Interestingly, the above result directly enables us to describe the distribution of the number of departures in all intervals. Let D_i be the number of departures in $(t_{i-1}, t_i]$. Then we would like to evaluate

$$\Phi(\mathbf{v}) := \mathbb{E} \left(\prod_{i=1}^n v_i^{D_i} \right).$$

Now observe that $L_i = L_{i-1} + A_i - D_i$, and hence $D_i = L_{i-1} - L_i + A_i$. As a consequence,

$$\Phi(\mathbf{v}) = \mathbb{E} \left(\prod_{i=1}^n v_i^{L_{i-1} - L_i + A_i} \right) = \mathbb{E} \left(v_1^{L_0} \cdot \prod_{i=1}^{n-1} \left(\frac{v_{i+1}}{v_i} \right)^{L_i} \cdot v_n^{-L_n} \cdot \prod_{i=1}^n v_i^{A_i} \right).$$

With $w_1(\mathbf{v}) := v_1$, $w_i(\mathbf{v}) := v_{i+1}/v_i$ for $i \in \{1, \dots, n - 1\}$ and $w_n(\mathbf{v}) := v_n^{-1}$, we thus find the following result.

Proposition 3.2 For any $\mathbf{v} \in \mathbb{R}^n$,

$$\Phi(\mathbf{v}) = \Psi(\mathbf{w}(\mathbf{v}), \mathbf{v}).$$

3.3 Explicit calculations

Let $L(t)$ be the number of customers present at time t . In this subsection we compute (i) the mean and variance of $L(0)$, (ii) the joint transform of $\Lambda(0)$ and $L(0)$, and (iii) the covariance between $L(0)$ and $L(t)$ for some $t > 0$. As before, we assume that the

process started at $-\infty$, so that it displays stationary behavior at time 0 (and consequently at t as well). In principle, (factorial) moments can be derived by differentiating $\Psi(\mathbf{w}, z)$ suitably often and plugging in $\mathbf{w} = \mathbf{1}$ and $z = \mathbf{1}$, but elegant direct arguments can be given, as we show now.

We first introduce some notation. Define β as the mean service time:

$$\beta = \int_0^\infty \bar{F}(s) ds.$$

The density of the residual service time B_e is given by $f_e(s) := \bar{F}(s)/\beta$.

o **Mean and variance of $\Lambda(0)$ and $L(0)$**

It is easily checked that $\mathbb{E} \Lambda(0) = \lambda := \mathbf{a}' \mathcal{Q}^{-1} \mathbf{q}$ and $\mathbb{V}ar \Lambda(0) = \mathbf{a}' \Sigma_0 \mathbf{a}$. We therefore now concentrate on the mean and variance of $L(0)$.

Let $\mathcal{P}(\mu)$ denote a Poisson random variable with parameter μ . It is straightforward that

$$\mathbb{E} L(0) = \mathbb{E} \left(\mathcal{P} \left(\int_{-\infty}^0 \Lambda(s) \bar{F}(-s) ds \right) \right) = \mathbb{E} \left(\int_{-\infty}^0 \Lambda(s) \bar{F}(-s) ds \right).$$

We thus find, with \mathbf{J} shorthand notation for the entire path of the process $\mathbf{J}(\cdot)$,

$$\begin{aligned} \mathbb{E}[L(0) | \mathbf{J}] &= \int_{-\infty}^0 \Lambda(s) \bar{F}(-s) ds = \int_0^\infty \Lambda(-s) \bar{F}(s) ds \\ &= \beta \int_0^\infty \Lambda(-s) f_e(s) ds = \beta \mathbb{E}[\Lambda(-B_e) | \mathbf{J}]. \end{aligned} \tag{16}$$

For later use, we rewrite this expression as

$$\begin{aligned} \beta \mathbb{E}[\Lambda(-B_e) | \mathbf{J}] &= \beta \mathbb{E}[\mathbf{a}' \mathbf{X}(-B_e) | \mathbf{J}] = \beta \mathbf{a}' \mathbb{E}[\mathbf{X}(-B_e) | \mathbf{J}] \\ &= \beta \int_{-\infty}^0 \left(\mathbb{E} \mathbf{a}' e^{-\mathcal{Q}(-B_e-s)} \mathbf{1}_{\{s \leq -B_e\}} \right) d\mathbf{J}(s) \\ &\stackrel{d}{=} \beta \int_0^\infty \left(\mathbb{E} e^{-\mathcal{Q}'(s-B_e)} \mathbf{a} \mathbf{1}_{\{B_e \leq s\}} \right)' d\mathbf{J}(s), \end{aligned} \tag{17}$$

where the last step is due to time-reversibility of $\mathbf{J}(\cdot)$.

Applying (7), we obtain

$$\mathbb{E} L(0) = \beta \mathbb{E} \Lambda(0) = \lambda \beta. \tag{18}$$

Now we move to computing $\mathbb{V}ar L(0)$. Due to the fact that $L(0)$ is mixed Poisson, we have that $\mathbb{V}ar(L(0) | \mathbf{J}) = \mathbb{E}(L(0) | \mathbf{J})$ and thus $\mathbb{E} \mathbb{V}ar(L(0) | \mathbf{J}) = \mathbb{E} L(0) = \lambda \beta$.

Using the law of total variance, (16), (18) and (9),

$$\begin{aligned} \text{Var } L(0) &= \mathbb{E} \text{Var} (L(0) | \mathbf{J}) + \text{Var} (\mathbb{E}[L(0) | \mathbf{J}]) \\ &= \lambda\beta + \text{Var} (\mathbb{E}(\Lambda(-B_e) | \mathbf{J})) \beta^2 \\ &= \lambda\beta + \beta^2 \int_0^\infty \mathbf{a}' \left(\mathbb{E} e^{-Q(s-B_e)} 1_{\{B_e \leq s\}} \right) \Sigma \left(\mathbb{E} e^{-Q'(s-B_e)} 1_{\{B_e \leq s\}} \right) \mathbf{a} \, ds. \end{aligned} \tag{19}$$

Expression (19) can be substantially simplified, which we do later in this section. Observe that $\text{Var } L(0) \geq \mathbb{E} L(0)$, reflecting the fact that Coxian arrival rates lead to overdispersed arrival processes; see, for example, [4,7].

o **Joint transform of $\Lambda(0)$ and $L(0)$**

Here the goal is to determine $\mathbb{E} e^{-v\Lambda(0)} w^{L(0)}$. By arguments similar to those used in Sect. 3,

$$\begin{aligned} \mathbb{E} e^{-v\Lambda(0)} w^{L(0)} &= \mathbb{E} \exp \left((w-1) \int_{-\infty}^0 \Lambda(s) \bar{F}(-s) ds - v \Lambda(0) \right) \\ &= \mathbb{E} \exp \left((w-1) \mathbf{a}' \int_{-\infty}^0 X(s) \bar{F}(-s) ds - v \mathbf{a}' X(0) \right). \end{aligned}$$

Combining this relation with (cf. (6))

$$\begin{aligned} \int_{-\infty}^0 X(s) \bar{F}(-s) ds &= \int_{-\infty}^0 \int_{-\infty}^s e^{-Q(r-s)} d\mathbf{J}(r) \bar{F}(-s) ds \\ &= \int_{-\infty}^0 \int_r^0 e^{-Q(r-s)} \bar{F}(-s) ds d\mathbf{J}(r) \\ &= \int_0^\infty \int_0^s e^{-Q(s-r)} \bar{F}(r) dr d\mathbf{J}(s), \end{aligned}$$

yields

$$\mathbb{E} e^{-v\Lambda(0)} w^{L(0)} = \mathbb{E} \exp \left(\mathbf{a}' \int_0^\infty \Omega(v, w, s) d\mathbf{J}(s) \right),$$

where

$$\begin{aligned} \Omega(v, w, s) &:= (w-1) \int_0^s e^{-Q(s-r)} \bar{F}(r) dr - v e^{-Qs} \\ &= (w-1)\beta \mathbb{E} e^{-Q(s-B_e)} 1_{\{B_e \leq s\}} - v e^{-Qs}. \end{aligned}$$

Applying (8), we thus obtain the following result.

Proposition 3.3

$$\mathbb{E} e^{-v\Lambda(0)} w^{L(0)} = \exp \left(- \int_0^\infty \eta(-\Omega(v, w, s)' \mathbf{a}) ds \right).$$

It is possible to apply this joint transform to the computation of moments of $L(0)$ and $\Lambda(0)$, as well as their mixed moments, but lower moments can typically be computed by direct arguments. The mean and variance of $\Lambda(0)$ and $L(0)$ have been identified above. We therefore now focus on the covariance between $\Lambda(0)$ and $L(0)$. Since $\mathbb{E}[\Lambda(0) | \mathbf{J}] = \Lambda(0)$ (as $\mathbf{X}(\cdot)$ is a functional of $\mathbf{J}(\cdot)$), and since B_e is independent of \mathbf{J} , we find

$$\begin{aligned} \text{Cov}(\Lambda(0), L(0)) &= \text{Cov}(\mathbb{E}[L(0) | \mathbf{J}], \Lambda(0)) = \text{Cov}(\mathbb{E}[\Lambda(-B_e) | \mathbf{J}], \Lambda(0)) \\ &= \text{Cov}(\Lambda(-B_e), \Lambda(0)) = \mathbb{E} \text{Cov}(\Lambda(-B_e), \Lambda(0) | B_e) \\ &= \mathbf{a}' \mathbb{E} \Sigma_{B_e} \mathbf{a} = \mathbf{a}' \Sigma_0 \mathbb{E} e^{-Q' B_e} \mathbf{a}, \end{aligned}$$

where we employ the fact that the covariance matrix between $\mathbf{X}(t)$ and $\mathbf{X}(t + h)$, which is the same as that between $\mathbf{X}(t - h)$ and $\mathbf{X}(t)$, is given by $\Sigma_h = \Sigma_0 e^{-Q'h}$.

◦ **Covariance between $L(0)$ and $L(t)$**

The starting point is the law of total covariance:

$$\text{Cov}(L(0), L(t)) = \mathbb{E} \text{Cov}(L(0), L(t) | \mathbf{J}) + \text{Cov}(\mathbb{E}[L(0) | \mathbf{J}], \mathbb{E}[L(t) | \mathbf{J}]).$$

We evaluate the two terms separately. The first term can be rewritten as follows: Let $C_1(t)$ denote the number of customers that arrive before time 0 and depart in $(0, t]$; $C_2(t)$ is the number of customers that arrive before time 0 and depart after t ; finally, $C_3(t)$ is the number of customers that arrive in $(0, t]$ and depart after t . Evidently, due to the conditional independence of these three quantities,

$$\begin{aligned} \text{Cov}(L(0), L(t) | \mathbf{J}) &= \text{Cov}(C_1(t) + C_2(t), C_2(t) + C_3(t) | \mathbf{J}) \\ &= \text{Var}(C_2(t) | \mathbf{J}) = \mathbb{E}(C_2(t) | \mathbf{J}), \end{aligned}$$

with, mimicking the above arguments,

$$\mathbb{E}(C_2(t) | \mathbf{J}) \stackrel{d}{=} \beta \int_0^\infty \mathbb{E} \left(e^{-Q'(s+t-B_e)} \mathbf{a} 1_{\{t < B_e \leq t+s\}} \right) d\mathbf{J}(s),$$

where the last equality is due to the fact that the conditional distribution of $C_2(t)$ given \mathbf{J} is Poisson. Thus, with $F_e(\cdot)$ denoting the distribution function of B_e and $\bar{F}_e(\cdot)$ the corresponding complementary distribution function,

$$\begin{aligned}
 \mathbb{E} \text{Cov} (L(0), L(t) | \mathbf{J}) &= \beta \int_0^\infty \mathbb{E} \left(\mathbf{a}' e^{-Q(s+t-B_e)} \boldsymbol{\varrho} 1_{\{t < B_e \leq t+s\}} \right) ds \\
 &= \beta \mathbb{E} \left(\int_{B_e-t}^\infty \left(\mathbf{a}' e^{-Q(s-(B_e-t))} \boldsymbol{\varrho} \right) ds 1_{\{B_e > t\}} \right) \\
 &= \beta \mathbb{E} \left(\int_0^\infty \mathbf{a}' e^{-Qs} ds 1_{\{B_e > t\}} \right) \boldsymbol{\varrho} \\
 &= \beta \mathbf{a}' Q^{-1} \boldsymbol{\varrho} \mathbb{P}(B_e > t) = \lambda \beta \bar{F}_e(t).
 \end{aligned}$$

Next, we move to the second term. To this end, we first recall (cf. (16) and (17)) that

$$\begin{aligned}
 \mathbb{E}[L(0) | \mathbf{J}] &= \beta \int_0^\infty \mathbb{E} \left(e^{-Q'(s-B_e)} \mathbf{a} 1_{\{B_e \leq s\}} \right) d\mathbf{J}(s), \\
 \mathbb{E}[L(t) | \mathbf{J}] &= \beta \int_{-t}^\infty \mathbb{E} \left(e^{-Q'(s+t-B_e)} \mathbf{a} 1_{\{B_e \leq s+t\}} \right) d\mathbf{J}(s).
 \end{aligned}$$

As \mathbf{J} has independent increments, when considering the covariance between $\mathbb{E}[L(0) | \mathbf{J}]$ and $\mathbb{E}[L(t) | \mathbf{J}]$, we can restrict ourselves to integrating over $s > 0$ only; more concretely,

$$\begin{aligned}
 &\text{Cov}(\mathbb{E}[L(0) | \mathbf{J}], \mathbb{E}[L(t) | \mathbf{J}]) \\
 &= \beta^2 \text{Cov} \left(\int_0^\infty \mathbb{E} \left(e^{-Q'(s-B_e)} \mathbf{a} 1_{\{B_e \leq s\}} \right) d\mathbf{J}(s), \right. \\
 &\quad \left. \int_0^\infty \mathbb{E} \left(e^{-Q'(s+t-B_e)} \mathbf{a} 1_{\{B_e \leq s+t\}} \right) d\mathbf{J}(s) \right) \\
 &= \beta^2 \int_0^\infty \mathbb{E} \left(\mathbf{a}' e^{-Q(s-B_e)} 1_{\{B_e \leq s\}} \right) \Sigma \mathbb{E} \left(e^{-Q'(s+t-B_e)} \mathbf{a} 1_{\{B_e \leq s+t\}} \right) ds, \quad (20)
 \end{aligned}$$

where the last step follows by using (9).

As it turns out, the last formula (as well as (19)) may be simplified, as follows: Let $B_{e,1}$ and $B_{e,2}$ be two i.i.d. copies of B_e . Recalling (12) and denoting $x^+ := x \vee 0$ and $x^- := -x \wedge 0 = (-x)^+$, (20) can be rewritten as

$$\begin{aligned}
 &\beta^2 \mathbf{a}' \int_0^\infty \mathbb{E} \left(e^{-Q(s-B_e)} 1_{\{B_e \leq s\}} \right) \Sigma \mathbb{E} \left(e^{-Q'(s+t-B_e)} 1_{\{B_e \leq s+t\}} \right) ds \mathbf{a} \\
 &= \beta^2 \mathbf{a}' \mathbb{E} \left(\int_{B_{e,1} \vee (B_{e,2}-t)}^\infty e^{-Q(s-B_{e,1})} \Sigma e^{-Q'(s+t-B_{e,2})} ds \right) \mathbf{a} \\
 &= \beta^2 \mathbf{a}' \mathbb{E} \left(e^{-Q(B_{e,1}-B_{e,2}+t)^-} \left(\int_0^\infty e^{-Qs} \Sigma e^{-Q's} ds \right) e^{-Q'(B_{e,1}-B_{e,2}+t)^+} \right) \mathbf{a} \\
 &= \beta^2 \mathbf{a}' \mathbb{E} \left(e^{-Q(B_{e,1}-B_{e,2}+t)^-} \Sigma_0 e^{-Q'(B_{e,1}-B_{e,2}+t)^+} \right) \mathbf{a},
 \end{aligned}$$

where, in the last equality, (11) has been used. For any $s \in \mathbb{R}$ we have that, recalling Equation (13),

$$\mathbf{a}' e^{-Qs^-} \Sigma_0 e^{-Q's^+} \mathbf{a} = \mathbf{a}' \Sigma_0 e^{-Q'|s|} \mathbf{a} = \mathbf{a}' \Sigma_{|s|} \mathbf{a}.$$

Thus, denoting by $R(t)$ the autocorrelation function (with $R(0) = 1$), then adding the two terms yields

$$R(t) \cdot \text{Var } L(0) = \text{Cov}(L(0), L(t)) = \lambda\beta \bar{F}_e(t) + \beta^2 \mathbf{a}' \mathbb{E} (\Sigma_{|B_{e,1} - B_{e,2} + t|}) \mathbf{a}. \tag{21}$$

In particular, when $t = 0$, (21) provides us with a simplified expression for $\text{Var } L(0)$. The density of $B_{e,1} - B_{e,2}$ is clearly symmetric around zero and is given by

$$g(x) = \int_0^\infty f_e(y) f_e(y + |x|) dy.$$

Since $f_e(\cdot) = \beta^{-1} \bar{F}(\cdot)$ is non-increasing on $[0, \infty)$, this implies that $g(\cdot)$ is unimodal.

We end this subsection by considering the one-dimensional case. Since both $h(x) := e^{-q|x|}$ ($q > 0$) and $g(x)$ are symmetric and unimodal (around zero), it follows by [15] that so is their convolution. Alternatively, this follows also from [5] as $h(\cdot)$ is also log-concave. This implies that

$$\begin{aligned} \mathbb{E} e^{-q|B_{e,1} - B_{e,2} + t|} &= \int_{-\infty}^\infty e^{-q|x+t|} g(x) dx = \int_{-\infty}^\infty e^{-q|-x+t|} g(-x) dx \\ &= \int_{-\infty}^\infty e^{-q|t-x|} g(x) dx \end{aligned}$$

is unimodal with a mode at zero, hence non-increasing on $[0, \infty)$, and thus so is $R(\cdot)$. A similar result holds if \mathbf{a} is an eigenvector associated with a real-valued (positive) eigenvalue q of Q' since then $\mathbf{a}' \Sigma_0 e^{-Q'|x|} \mathbf{a} = \mathbf{a}' \Sigma_0 \mathbf{a} e^{-q|x|}$. However, in general, even though $\mathbf{a}' \Sigma_0 e^{-Q'|x|} \mathbf{a}$ is a symmetric function, it is not clear if it is decreasing on $[0, \infty)$ or if it is log-concave. However, since it vanishes as $|x| \rightarrow \infty$, it is clear that $R(t)$ vanishes as $t \rightarrow \infty$.

4 The network case

In this section we consider *networks* of infinite-server queues with a Cox input process. There are n queues, with the arrival rate of queue k being a non-homogeneous Poisson process with rate function $\Lambda_k(t) = \mathbf{a}'_k \mathbf{X}(t)$. This is the same as defining the vector of arrival processes to be $\mathbf{\Lambda}(t) = A\mathbf{X}(t)$ for some matrix A with rows \mathbf{a}'_k . The $\mathbf{X}(t)$ process is the same multivariate shot-noise process as before. Notice that in this construction the arrival processes at the various queues are potentially dependent.

Define by $p_{km}(t)$ the probability that a job entering at queue k at time 0 is at queue m at time t ; likewise, $p_{k0}(t)$ is the probability that a job entering at queue k at time 0

has left the network by time t . In the case of exponentially distributed service times and probabilistic routing, these $p_{km}(t)$'s can be computed more explicitly (relying on the machinery developed for phase-type distributions).

Before studying the joint queue length distribution in the network (Sect. 4.2) and presenting an example (Sect. 4.3), we briefly remind the reader of the structure of the joint queue length distribution in the case of two M/G/∞ queues in tandem (Sect. 4.1). This helps sharpen our intuition and will make it easier to digest and understand the results of Sect. 4.2.

4.1 Two M/G/∞ queues in tandem

Consider a model of two M/G/∞ queues 1, 2 in series, with arrival rate λ at queue 1, generic service time B_i at queue i , $i = 1, 2$, and all customers moving from queue 1 to 2 before leaving the system. Queue 2 has no external arrivals. The same concepts that were underlying the results mentioned in Sect. 3.1 (see, in particular, (14)) quickly lead to the following expression for the generating function of the joint distribution of the two queue lengths $L_1(t), L_2(t)$:

$$\begin{aligned} &\mathbb{E}[w_1^{L_1(t)} w_2^{L_2(t)} | L_1(0) = L_2(0) = 0] \\ &= \exp\left((w_1 - 1)\lambda \int_0^t \mathbb{P}(B_1 > u) du + (w_2 - 1)\lambda \int_0^t \mathbb{P}(B_1 < u, B_1 + B_2 > u) du\right). \end{aligned}$$

Furthermore, the generating function of the steady-state joint queue length distribution is given by

$$\mathbb{E}[w_1^{L_1} w_2^{L_2}] = \exp((w_1 - 1)\lambda \mathbb{E}B_1 + (w_2 - 1)\lambda \mathbb{E}B_2).$$

4.2 Joint queue length

We focus on analyzing the stationary joint queue length distribution. In principle, virtually all quantities studied in the previous section can be derived again, at the expense of introducing rather heavy notation.

In this section, we let K_m denote the stationary queue length at node $m \in \{1, \dots, n\}$. Our objective is to compute the joint probability-generating function

$$\Pi(\mathbf{w}) = \mathbb{E}\left(\prod_{m=1}^n w_m^{K_m}\right).$$

Using the same arguments as in the previous section, we conclude that K_m has a mixed Poisson distribution. More precisely, K_m has a Poisson distribution with (random) parameter

$$\int_{-\infty}^0 \sum_{k=1}^n \Lambda_k(s) p_{km}(-s) ds. \tag{22}$$

It follows that

$$\begin{aligned} \Pi(\mathbf{w}) &= \mathbb{E} \left(\prod_{m=1}^n w_m^{K_m} \right) = \mathbb{E} \exp \left(\sum_{m=1}^n (w_m - 1) \int_{-\infty}^0 \sum_{k=1}^n \Lambda_k(s) p_{km}(-s) ds \right) \\ &= \mathbb{E} \exp \left(\sum_{m=1}^n (w_m - 1) \int_{-\infty}^0 \sum_{k=1}^n \mathbf{a}'_k \mathbf{X}(s) p_{km}(-s) ds \right). \end{aligned} \tag{23}$$

Substituting (cf. (6))

$$\mathbf{X}(s) = \int_{-\infty}^s e^{-Q(s-r)} d\mathbf{J}(r)$$

in (23) gives

$$\Pi(\mathbf{w}) = \mathbb{E} \exp \left(\sum_{k=1}^n \mathbf{a}'_k \int_{-\infty}^0 \int_r^0 e^{-Q(s-r)} \pi_k(s | \mathbf{w}) ds d\mathbf{J}(r) \right),$$

with

$$\pi_k(s | \mathbf{w}) := \sum_{m=1}^n p_{km}(-s) (w_m - 1) = \sum_{m=1}^n p_{km}(-s) w_m + p_{k0}(-s) - 1.$$

This leads to the following result.

Theorem 4.1 For any $\mathbf{w} \in \mathbb{R}^n$,

$$\begin{aligned} \Pi(\mathbf{w}) &= \mathbb{E} \left(\prod_{m=1}^n w_m^{K_m} \right) \\ &= \exp \left(- \int_{-\infty}^0 \eta \left(- \sum_{k=1}^n \int_r^0 e^{-Q'(s-r)} \pi_k(s | \mathbf{w}) ds \mathbf{a}_k \right) dr \right). \end{aligned}$$

4.3 Example

In this illustrative example we consider a two-node tandem system in which the service times at both nodes are exponential with parameter $\kappa > 0$. We assume for ease that there is only input at the first queue and that all customers move from the first queue to the second queue. It is immediate that, for $t \geq 0$,

$$p_{11}(t) = e^{-\kappa t}, \quad p_{12}(t) = \kappa t e^{-\kappa t}, \quad p_{10}(t) = 1 - e^{-\kappa t} - \kappa t e^{-\kappa t};$$

use that the sojourn time in the system is Erlang(2). We find that

$$\pi_1(t | w_1, w_2) = (w_1 - 1) e^{\kappa t} - (w_2 - 1) \kappa t e^{\kappa t}.$$

Assume for ease that $d = 1$ (i.e., one-dimensional shot-noise). We choose $a_1 = 1$ (whereas $a_2 = 0$, as we assumed no external input to the second queue). As a consequence, we have that $\Lambda_1(t) = X(t)$ and $\Lambda_2(t) = 0$, where $X(t)$ can be represented as

$$X(t) = \int_{(-\infty, t]} e^{-q(t-s)} dJ(s),$$

for some $q > 0$ and a (scalar) Lévy subordinator $J(\cdot)$ (with exponent $-\eta(\cdot)$).

Appealing to Thm. 4.1,

$$\Pi(w_1, w_2) = \mathbb{E}[w_1^{K_1} w_2^{K_2}] = \exp\left(-\int_{-\infty}^0 \eta\left(-\int_r^0 e^{-q(s-r)} \pi_1(s | w_1, w_2) ds\right) dr\right). \tag{24}$$

To evaluate this expression, observe that

$$\begin{aligned} \int_r^0 e^{-q(s-r)} \pi_1(s | w_1, w_2) ds &= \int_r^0 e^{-q(s-r)} ((w_1 - 1) e^{\kappa s} - (w_2 - 1) \kappa s e^{\kappa s}) ds \\ &= (w_1 - 1) \frac{e^{qr} - e^{\kappa r}}{\kappa - q} + (w_2 - 1) \left(\frac{\kappa r e^{\kappa r}}{\kappa - q} + \frac{\kappa(e^{qr} - e^{\kappa r})}{(\kappa - q)^2} \right). \end{aligned} \tag{25}$$

It is easy to obtain moments of K_1 and K_2 from (24) and (25) by differentiation with respect to w_1 and w_2 , respectively. In particular, straightforward algebra yields

$$\mathbb{E}[K_1] = \mathbb{E}[K_2] = \frac{\eta'(0)}{\kappa q}.$$

Remark 4.2 Similar calculations can be performed for considerably more general networks. Application of Thm. 4.1 requires (i) expressions for the matrix exponential of Q' and (ii) expressions for $\pi_k(s | \mathbf{w})$ (and therefore for $p_{km}(s)$) for $s \geq 0$. To this end, evaluate, using common techniques from linear algebra, the matrix exponential of Q' ; this can be done relying on standard computer algebra. Assuming the eigenvalues of Q to have multiplicity 1 (where we note that higher multiplicities can be dealt with too, but require some additional care), we thus obtain a matrix whose entries are weighted sums of exponentials. Likewise, if the service times are of phase type, then the probabilities $p_{km}(s)$ can be written in terms of (specific entries of) a matrix exponential, and therefore (again under mild conditions) also as a weighted sum of exponentials. It thus follows that the integrand $e^{-Q'(s-r)} \pi_k(s | \mathbf{w})$ is a mixture of exponentials (in s) and can be integrated by elementary calculus. This allows the computation of the (mixed) moments pertaining to the stationary queue lengths.

We now return to our tandem example. Suppose that $J(\cdot)$ corresponds to a Gamma process [9, Section 1.2.4] with (without loss of generality) rate and shape parameters both equal to 1, i.e.,

$$\eta(\alpha) = -\log\left(\frac{1}{1+\alpha}\right) = \log(1+\alpha).$$

The probability-generating function of the queue length in the first queue is therefore

$$\Pi(w, 1) = \exp\left(-\int_{-\infty}^0 \log\left(1 + (1-w)\frac{e^{qr} - e^{\kappa r}}{\kappa - q}\right) dr\right).$$

Using the Taylor expansion of the logarithm, the exponent of this expression can be rewritten as

$$\int_{-\infty}^0 \sum_{n=1}^{\infty} \frac{(-1)^n}{n} (v(w)(e^{qr} - e^{\kappa r}))^n dr,$$

with $v(w) := (1-w)/(\kappa - q)$. Relying on the binomium, this equals

$$\int_{-\infty}^0 \sum_{n=1}^{\infty} \frac{(-1)^n}{n} (v(w))^n \left(\sum_{k=0}^n \binom{n}{k} (-1)^{n-k} e^{qrk + \kappa r(n-k)}\right) dr.$$

Swapping the order of the summations and the integral, we obtain

$$\sum_{n=1}^{\infty} \frac{1}{n} (v(w))^n \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{1}{qk + \kappa(n-k)}.$$

We thus end up with

$$\mathbb{E}[w^{K_1}] = \exp\left(-\sum_{n=1}^{\infty} \frac{1}{n} (v(w))^n \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{1}{qk + \kappa(n-k)}\right).$$

A similar procedure yields the probability-generating function of the second queue, but the resulting expressions are slightly more complicated.

5 Conclusions and suggestions for further research

In this paper we have considered a network of infinite-server queues where the input process is a Cox process that allows much modeling flexibility; the arrival rate is represented as a linear combination of the components of a multivariate generalized shot-noise process. We have derived various distributional properties of the multivariate shot-noise process, subsequently exploiting them to obtain the joint transform

of the numbers of customers, at consecutive time epochs, in an infinite-server queue with, as input process, such a Cox process. We have also derived the joint steady-state transform of the vectors of arrival rate and queue length, as well as their means and covariance structure, and we have studied the departure process from the queue. Finally, we extended our analysis to the setting of a network of infinite-server queues, allowing the arrival processes at the various queues to be dependent Cox processes.

In a follow-up study [14] we investigate several related aspects. Specifically, (i) we develop a recursive scheme that allows us to obtain higher-order moments of $(\Lambda(t), L(t))$, (ii) we derive asymptotics of the queue length process, under assumptions regarding the tail behavior of the shot-noise process, (iii) we study the heavy-traffic behavior of the queue length process, and (iv) we obtain several numerical results for means, variances and correlations; this includes a numerical exploration of the sensitivity of various performance metrics with respect to the service time distribution. It could also be investigated to what extent the assumption of exponential decay in the shot-noise process can be relaxed.

Acknowledgements OK's work has been supported by Israel Science Foundation (Grant Number 1647/17) and the Vigevani Chair in Statistics. The research of OB and MM is partly funded by the NWO Gravitation Programme NETWORKS (Grant Number 024.002.003) and an NWO Top Grant (Grant Number 613.001.352).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Çinlar, E.: Probability and Stochastics. Graduate Texts in Mathematics, vol. 261. Springer, New York (2011)
2. Daw, A., Pender, J.: Queues driven by Hawkes processes. *Stoch. Syst.* **8**, 192–229 (2018)
3. Eick, S., Massey, W., Whitt, W.: The physics of the $M_t/G/\infty$ queue. *Oper. Res.* **41**, 731–742 (1993)
4. Heemskerck, M., van Leeuwen, J., Mandjes, M.: Scaling limits for infinite-server systems in a random environment. *Stoch. Syst.* **7**, 1–31 (2017)
5. Ibragimov, I.A.: On the composition of unimodal distributions. *Theory Probab. Appl.* **1**, 255–260 (1956)
6. Kella, O., Whitt, W.: Linear stochastic fluid networks. *J. Appl. Probab.* **36**, 244–260 (1999)
7. Koops, D., Boxma, O., Mandjes, M.: Networks of $M_t/G/\infty$ queues with shot-noise-driven arrival intensities. *Queueing Syst.* **86**, 301–325 (2017)
8. Koops, D., Saxena, M., Boxma, O., Mandjes, M.: Infinite-server queues with Hawkes input. *J. Appl. Probab.* **55**, 920–943 (2018)
9. Kyprianou, A.: Fluctuations of Lévy Processes with Applications. Introductory Lectures. Springer, Heidelberg (2014)
10. Mirasol, N.M.: The output of an $M/G/\infty$ queueing system is Poisson. *Oper. Res.* **11**, 282–284 (1963)
11. Riziou, M.-A., Xie, L., Sanner, S., Cebrian, M., Yu, H., Van Henteryck, P.: Expecting to be HIP: Hawkes intensity processes for social media popularity. In: Proceedings of the 26th International Conference on World Wide Web (Republic and Canton of Geneva, Switzerland, 2017), WWW '17, pp. 735–744 (2017)
12. Riziou, M.-A., Xie, L.: Online popularity under promotion: Viral potential, forecasting, and the economics of time. [arXiv:1703.01012v3](https://arxiv.org/abs/1703.01012v3) (2017)
13. Sato, K.-I.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge (2005)

14. Saxena, M., Boxma, O., Mandjes, M.: An infinite-server system with Lévy shot-noise modulation: moments and asymptotics. Eurandom Report 2019-003 (accepted for publication in Markov Processes and Related Fields). <https://www.eurandom.nl/reports/2019/003-report.pdf> (2018)
15. Wintner, A.: Asymptotic Distributions and Infinite Convolutions. Edwards Brothers, Ann Arbor (1938)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.