



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

PNA

Probability, Networks and Algorithms



Probability, Networks and Algorithms

Transient analysis of Markov-fluid-driven queues

A. Es-Saghouani, M.R.H. Mandjes

REPORT PNA-R0808 MAY 2008

Centrum voor Wiskunde en Informatica (CWI) is the national research institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organisation for Scientific Research (NWO). CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

Copyright © 2008, Stichting Centrum voor Wiskunde en Informatica
P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-3711

Transient analysis of Markov-fluid-driven queues

ABSTRACT

In this paper we study two transient characteristics of a Markov-fluid-driven queue, viz., the busy period and the covariance function of the workload process. Both metrics are captured in terms of their Laplace transforms. Relying on sample-path large deviations we also identify the logarithmic asymptotics of the probability that the busy period lasts longer than t , as $t \rightarrow \infty$. Examples are included that illustrate the theory.

2000 Mathematics Subject Classification: 60K25

Keywords and Phrases: queueing, Markov fluid, transient analysis

Transient analysis of Markov-fluid-driven queues

Abdelghafour Es-Saghouani & Michel Mandjes *

April 28, 2008

Abstract

In this paper we study two transient characteristics of a Markov-fluid-driven queue, viz., the busy period and the covariance function of the workload process. Both metrics are captured in terms of their Laplace transforms. Relying on sample-path large deviations we also identify the logarithmic asymptotics of the probability that the busy period lasts longer than t , as $t \rightarrow \infty$. Examples are included that illustrate the theory.

1 Introduction

Markov fluid models have been widely studied in a variety of application domains, with significant contributions made in the areas of queueing theory, storage processes, communication networking, insurance risk, etc., see for instance [2, 4, 12, 13, 17, 19]. A *Markov-fluid-driven queue* is a storage system which is fed by a source whose transmission rate modulates between multiple values in a Markovian manner, and which is emptied at constant speed. Traditionally in the literature emphasis was laid on computing *steady-state* characteristics of this class of queueing systems — in particular the distribution of the stationary workload; see for a nice (recent) literature overview for instance the introduction of [7] — whereas considerably less attention has been paid to *transient analysis*. The main goal of the present paper is the analysis of two such transient characteristics: (i) the distribution of the busy period, and (ii) the correlation function of the workload process.

Let us first give a brief (non-exhaustive) account of the literature on transient analysis of Markov-fluid-driven queues. Restricting themselves to the special case in which the Markov fluid source is actually a superposition of on-off sources, Ren and Kobayashi [18] were able to convert the partial differential equations [2, 12, 13] that govern the queue's transient behavior, into a matrix

*Both authors are with Korteweg-de Vries Institute for Mathematics, Plantage Muidergracht 24, 1018 TV Amsterdam, the Netherlands; {aessagho|mmandjes}@science.uva.nl. M. Mandjes is also with CWI, Amsterdam, the Netherlands, and EURANDOM, Eindhoven, the Netherlands; part of this work was done while he was at Stanford University, Stanford, CA 94305, US.

equation in the Laplace domain. Roughly simultaneously a paper by Asmussen [3] appeared, where the Laplace transform of the busy period was computed, mainly relying on martingale techniques; the resulting transform is in terms of a matrix of probabilities, which are not given explicitly, but are fixed points of a specific integral equation (which are proven to be unique). Narayanan and Kulkarni [15] showed that the probability distribution of the first time the buffer becomes empty (starting at an arbitrary positive level x), satisfies a system of partial differential equations; furthermore, they show that its Laplace transform is a solution of a specific differential equation. Barbot *et al.* [5] mainly focused on numerical issues: using the fact that the probability distribution of the busy period obeys a certain backward differential equation, they proposed an efficient numerical procedure. Finally, recent work by Ahn and Ramaswami [1] provides an efficient (quadratically convergent) procedure for computing the Laplace transform of the busy period, exploiting relations with so-called quasi-birth-death processes. To our best knowledge, no results on the correlation function have been reported so far.

As said above, this paper focuses on the distribution of the busy period, as well as the workload's correlation function. More specifically, the main contributions are the following.

- *Busy period.* In the first place we adopt a new approach for computing the Laplace transform of the busy period. This approach, some steps of which resonate elements of [1], first uses elementary calculations to express the Laplace transform in terms of a number of auxiliary transforms (as many as there are states with net buffer increase). Then an important role is played by a lemma that provides us with a sufficient number of additional constraints in order to uniquely determine these auxiliary transforms. The proof of this lemma is based on a powerful result by Sonneveld [20], in conjunction with Geršgorin's circle theorem, see [14].

It is stressed that this new methodology has a number of attractive properties. Most of the analysis is based on first principles; the above-mentioned lemma is the only technical element. In addition, the analysis essentially carries over to the correlation function, see below. Also, in special cases, such as the case of a two-state modulating Markov chain, the analysis can be done explicitly, and the resulting transform can be inverted.

Then we focus on the logarithmic asymptotics of the tail distribution of the busy period. It is shown that these can be written in terms of the minimum of the so-called *cumulant function* (i.e., the asymptotic log-moment generating function) of the input process. The upper bound is elementary, viz. a direct application of the Gärtner-Ellis theorem. The lower bound, on the contrary, is considerably more technical and relies on sample-path large deviations [6].

- *Correlation function.* With $Q(t)$ denoting the buffer content at time t , the covariance function $\mathbb{Cov}(Q(0), Q(t))$ is a measure of dependence between the workload at time 0, and the

workload at time t (and the correlation function is defined as the covariance function divided by $\sqrt{\text{Var}Q(0)\text{Var}Q(t)}$). Using the methodology that we developed for analyzing the busy period, we uniquely characterize the covariance function by its Laplace transform. Assuming that the workload was in stationarity at time 0, this has been done before in the cases of compound Poisson input [16] and spectrally-positive Lévy input [9], but the case of Markov-fluid input was not addressed yet. We do not restrict ourselves to the case that $Q(0)$ is distributed according to the workload's equilibrium distribution; in fact, for any initial distribution of phase-type the Laplace transform of $\text{Cov}(Q(0), Q(t))$ has a relatively manageable form.

Again for the case of on-off Markov-fluid input, the correlation function can be determined explicitly. Interestingly, its asymptotics are equal (up to a multiplicative constant) to those of the tail distribution of the busy period — a property that was observed before in the case of queues with spectrally-positive Lévy input [9].

The remainder of this paper is organized as follows. In Section 2 we describe our model and recapitulate a number of known results on the steady-state workload distribution. In Section 3 we identify the Laplace transform of the busy period. Its logarithmic asymptotics are derived in Section 4, invoking sample-path large deviations. In Section 5 we concentrate on the covariance function of the workload process. Section 6 presents an example that illustrates the results obtained in this paper. We draw conclusions and identify a number of open problems in Section 7.

2 Model and preliminaries

Let $\{X(t), t \geq 0\}$ be an irreducible continuous-time Markov process with finite state space $\mathcal{E} = \{1, 2, \dots, N\}$. This modulating Markov process drives a buffer in the following way: if it is in state i , the buffer content changes at rate r_i (which can be both positive and negative); there is *reflection at zero*, meaning that if the buffer is empty, and the Markov process is in a state i with $r_i < 0$, then the buffer remains empty. We denote by $\{Q(t), t \geq 0\}$ the buffer content process (or: workload process). The buffer size is assumed to be infinite, and hence $Q(t)$ can attain any value in $[0, \infty)$.

In order to avoid confusions in the notation, in the sequel the bold small letters will denote vectors, and bold large letters will denote matrices. Let $\mathbf{\Lambda} = (\lambda_{ij})_{1 \leq i, j \leq N}$ be the intensity matrix (or: rate matrix) of the Markov process $X(t)$, with $\lambda_i = -\lambda_{ii}$. Also, denote by $\boldsymbol{\pi} \equiv (\pi_1, \pi_2, \dots, \pi_N)^T$ its invariant distribution (where the superscript T denotes the transpose); then π_i is the stationary probability that $X(t)$ is in state i . Because of the above assumptions, this distribution exists and is unique. Furthermore, let \mathcal{E}^+ be the states i in \mathcal{E} such that $r_i > 0$ ('up-states'), and \mathcal{E}^- the states such that $r_i < 0$ ('down-states'); we assume for ease that $r_i \neq 0$ for all $i \in \mathcal{E}$. Define the traffic rate matrix $\mathbf{R} = \text{diag}\{r_1, \dots, r_N\}$. Let N^+ be the number of up-states, and N^- the number of

down-states. For ease, we let the state-space of the modulating process $X(t)$ be labeled such that the first N^+ states correspond to the up-states, whereas states $N^+ + 1$ up to N correspond to the down-states. In self-evident notation, we write

$$\mathbf{r} \equiv \begin{pmatrix} \mathbf{r}^+ \\ \mathbf{r}^- \end{pmatrix}.$$

This straightforward partition is used frequently in the paper.

The transient probabilities $q_i(t, x) := \mathbb{P}(Q(t) \leq x, X(t) = i)$ of the bivariate Markov process $(Q(t), X(t))$, defined on $[0, \infty) \times \mathcal{E}$, are known to satisfy a system of partial differential equations [12, 13], namely

$$\frac{\partial}{\partial t} q_i(t, x) + r_i \frac{\partial}{\partial x} q_i(t, x) = \sum_{j \in \mathcal{E}} \lambda_{ji} q_j(t, x), \quad \forall i \in \mathcal{E}, x > 0.$$

It is also well-known that under the condition

$$\sum_{i=1}^N r_i \pi_i < 0, \tag{1}$$

the workload process is stable, that is, ergodic. Moreover, there exists a stochastic vector (Q, X) to which the process $(Q(t), X(t))$ converges in distribution as $t \rightarrow \infty$, and the stationary distribution of $(Q(t), X(t))$, say $\mathbf{q}(x) \equiv (q_1(x), \dots, q_N(x))^T$, exists and satisfies

$$\mathbf{R} \frac{d}{dx} \mathbf{q}(x) = \mathbf{\Lambda}^T \mathbf{q}(x).$$

As a solution to the above system one could try $\mathbf{q}(x) = e^{\xi x} \mathbf{v}$, where $\xi \in \mathbb{C}$ and \mathbf{v} is a N -dimensional vector. Inserting this into the differential equation yields $(\xi \mathbf{R} - \mathbf{\Lambda}^T) \mathbf{v} = 0$. A non-trivial solution \mathbf{v} exists if

$$\det(\xi \mathbf{R} - \mathbf{\Lambda}^T) = 0. \tag{2}$$

Sonneveld [20] showed that there are N eigenvalues ξ_j (counting multiplicities) satisfying Equation (2), of which N^+ have negative real parts, one is zero, and $N^- - 1$ have positive real parts.

If the eigenvalues ξ_j are simple then

$$\mathbf{q}(x) = \boldsymbol{\pi} + \sum_{j=1}^{N^+} c_j e^{\xi_j x} \mathbf{v}^j; \tag{3}$$

where (ξ_j, \mathbf{v}^j) satisfy $(\xi_j \mathbf{R} - \mathbf{\Lambda}^T) \mathbf{v}^j = 0$, and the constants c_j , $j = 1, \dots, N^+$ are determined by the boundary conditions $\pi_i + \sum_{j=1}^{N^+} c_j v_i^j = 0$ if $r_i > 0$; if the eigenvalues are non-simple, elementary results from the theory of linear differential equations entail that there are terms in the above spectral expansion of the form $c_j x^\ell e^{\xi_j x} \mathbf{v}^j$, with $\ell = 0, \dots, k-1$, in case of an eigenvalue of multiplicity of order k .

Finally, we mention that we let $p_i(x)$ denote the density corresponding to the distribution function $q_i(x)$ and $\mathbf{p}(x) \equiv (p_1(x), \dots, p_N(x))^T$.

3 Analysis of the busy period

In this section we determine the Laplace transform of the busy period. Denoting by P the remaining time till the buffer becomes empty, $P := \inf\{t \geq 0 : Q(t) = 0\}$, the Laplace transform of the busy period is defined by

$$f_i(s) := \mathbb{E}(e^{-sP} \mid Q(0) = 0, X(0) = i). \quad (4)$$

Realize that any busy period starts in a state $i \in \mathcal{E}^+$.

In our analysis, the following transforms play an important role: for $i \in \mathcal{E}$, $x \geq 0$, $s \geq 0$, and $t > 0$, define

$$\begin{aligned} \zeta(s \mid x, i) &:= \mathbb{E}(e^{-sP} \mid Q(0) = x, X(0) = i); \\ f_i(s, t) &:= \int_0^\infty e^{-tx} \zeta(s \mid x, i) dx. \end{aligned}$$

Furthermore, let $t_i \equiv t_i(s) := (\lambda_i + s)/r_i$ for $i \in \mathcal{E}^+$.

For now we will focus on the double transforms $f_i(s, t)$; later we explain how these relate to the Laplace transforms of the busy period, i.e., the $f_i(s)$. Notice that when analyzing $f_i(s, t)$, we have to consider both $i \in \mathcal{E}^+$ and $i \in \mathcal{E}^-$; we deal with these cases separately.

Let us first consider $i \in \mathcal{E}^+$. It is evident that the busy period cannot end before a transition of the modulating Markov process $X(\cdot)$. By virtue of the memoryless property of the exponential distribution, one immediately obtains

$$f_i(s, t) = \sum_{k \neq i} \frac{\lambda_{ik}}{\lambda_i} \int_0^\infty e^{-tx} \int_0^\infty \lambda_i e^{-\lambda_i u} e^{-su} \zeta(s \mid x + r_i u, k) du dx.$$

After some algebra (change-of-variables and interchanging order of integrals) this reduces to

$$f_i(s, t) = \sum_{k \neq i} \frac{\lambda_{ik}}{\lambda_i + s - tr_i} (f_k(s, t) - f_k(s, t_i)).$$

Now proceed with $i \in \mathcal{E}^-$. Then the busy period may end before the first transition of the modulating process. We obtain

$$\begin{aligned} f_i(s, t) &= \sum_{k \neq i} \frac{\lambda_{ik}}{\lambda_i} \int_0^\infty e^{-tx} \int_0^{-x/r_i} \lambda_i e^{-\lambda_i u} e^{-su} \zeta(s \mid x + r_i u, k) du dx \\ &\quad + \sum_{k \neq i} \frac{\lambda_{ik}}{\lambda_i} \int_0^\infty e^{-tx} \int_{-x/r_i}^\infty \lambda_i e^{-\lambda_i u} e^{sx/r_i} du dx. \end{aligned}$$

It is readily verified (again by a change-of-variable and in addition interchanging the order of the integrals) that this reduces to

$$f_i(s, t) = \sum_{k \neq i} \frac{\lambda_{ik}}{\lambda_i + s - tr_i} f_k(s, t) - \frac{r_i}{\lambda_i + s - tr_i}.$$

Summarizing, we have found that $f_i(s, t)$, $i \in \mathcal{E}$ is a solution of the following system

$$(-\lambda_i - s + tr_i)f_i(s, t) + \sum_{k \neq i} \lambda_{ik} f_k(s, t) = \sum_{k \neq i} \lambda_{ik} f_k(s, t_i), \quad i \in \mathcal{E}^+; \quad (5)$$

$$(-\lambda_i - s + tr_i)f_i(s, t) + \sum_{k \neq i} \lambda_{ik} f_k(s, t) = r_i, \quad i \in \mathcal{E}^-. \quad (6)$$

Now consider equations (5) and (6). It is clear that if the auxiliary transforms $f_k(s, t_i)$ would be known, for $(k, i) \in \mathcal{E} \times \mathcal{E}^+$, then $\mathbf{f}(s, t)$ follow from Cramer's rule:

$$f_i(s, t) = \frac{\det(\mathbf{\Lambda} + t\mathbf{R} - s\mathbf{I} \mid \mathbf{g}(s), i)}{\det(\mathbf{\Lambda} + t\mathbf{R} - s\mathbf{I})}, \quad i \in \mathcal{E}; \quad (7)$$

where for $i \in \mathcal{E}$, $(\mathbf{\Lambda} + t\mathbf{R} - s\mathbf{I} \mid \mathbf{g}(s), i)$ is equal to the matrix $(\mathbf{\Lambda} + t\mathbf{R} - s\mathbf{I})$ with its i^{th} column replaced by a vector $\mathbf{g}(s)$ defined by

$$g_i(s) := \begin{cases} \sum_{k \neq i} \lambda_{ik} f_k(s, t_i), & i \in \mathcal{E}^+ \\ r_i, & i \in \mathcal{E}^-. \end{cases} \quad (8)$$

Therefore, it remains to identify the $N \cdot N^+$ auxiliary transforms $f_k(s, t_i)$, for $(k, i) \in \mathcal{E} \times \mathcal{E}^+$, and $s > 0$ given.

Now first observe that inserting $t = t_j$ (for $j \in \mathcal{E}^+$) into the system (5)–(6), yields $N^+ \cdot (N - 1)$ linear equations for the $f_k(s, t_i)$; realize that for $t = t_i$ we obtain a meaningless relation (that is, $0 = 0$). In other words, we can express all $f_k(s, t_i)$, for $i \in \mathcal{E}^+$ and $k \in \mathcal{E} \setminus \{i\}$, in terms of the N^+ unknowns $f_j(s, t_j)$, where $j \in \mathcal{E}^+$. Put differently, we have identified functions $\gamma_{kij}(s)$ and $\sigma_{ki}(s)$ such that

$$f_k(s, t_i) = \sum_{j \in \mathcal{E}^+} \gamma_{kij}(s) f_j(s, t_j) + \sigma_{ki}(s), \quad i \in \mathcal{E}^+, \quad k \in \mathcal{E} \setminus \{j\}. \quad (9)$$

Hence it remains to identify the $f_j(s, t_j)$, for $j \in \mathcal{E}^+$; if we would know them we could rewrite the vector $\mathbf{g}^+(s)$ as

$$g_i(s) = \sum_{j \in \mathcal{E}^+} \sum_{k \neq i} \lambda_{ik} \gamma_{kij}(s) f_j(s, t_j) + \sum_{k \neq i} \lambda_{ik} \sigma_{ki}(s). \quad (10)$$

With $\mathbf{f}^+(s, t_+) \equiv (f_1(s, t_1), \dots, f_{N^+}(s, t_{N^+}))^{\text{T}}$, the above means that we have, in self-evident notation, constructed a matrix $\mathbf{M}(s)$ (square; of dimension N^+) and a vector $\boldsymbol{\omega}^+(s)$ such that

$$\mathbf{g}^+(s) = \mathbf{M}(s) \mathbf{f}^+(s, t_+) + \boldsymbol{\omega}^+(s); \quad (11)$$

where the matrix \mathbf{M} and the vector $\boldsymbol{\omega}^+$ are given by

$$m_{ij}(s) := \sum_{k \neq i} \lambda_{ik} \gamma_{kij}(s), \quad i, j \in \mathcal{E}^+; \quad (12)$$

$$\omega_i(s) := \sum_{k \neq i} \lambda_{ik} \sigma_{ki}(s), \quad i \in \mathcal{E}^+. \quad (13)$$

As mentioned above, it remains to determine $f_j(s, t_j)$, $j \in \mathcal{E}^+$ for a given $s > 0$. This can be done by using the following powerful lemma.

Lemma 3.1. *For fixed $s > 0$, consider the equation $\det(\mathbf{\Lambda} - s\mathbf{I} + t\mathbf{R}) = 0$ for $t \in \mathbb{C}$. There are N^+ values of t with $\operatorname{Re} t > 0$ satisfying this equation.*

For a special choice of $\mathbf{\Lambda}$ and \mathbf{R} the above lemma was proven in [14], but is readily checked that the result carries over to general $\mathbf{\Lambda}$ and \mathbf{R} (as long as $\mathbf{\Lambda}$ corresponds to an irreducible Markov chain, and as long as the stability condition (1) is fulfilled). The proof is identical, and is based on [20], and intensively uses Geršgorin's circle theorem [10].

Fixing $s > 0$, realize that the transform (7) should have a finite norm for any t in the right half-plane. Hence, for any t in the right half plane for which the denominator in (7) equals 0 (that is, $\det(\mathbf{\Lambda} + t\mathbf{R} - s\mathbf{I}) = 0$), also the numerator should equal 0. From Lemma 3.1 we know that there are, for any $s > 0$, exactly N^+ such zeros in the right half-plane. Inserting these zeros into the numerator of (7) and equating it 0, we obtain exactly N^+ linear equations that determine $\mathbf{f}^+(s, t_+)$. Using (11) we can now determine $\mathbf{g}(s)$, and using (7) also $\mathbf{f}(s, t)$.

We conclude this section by arguing that, knowing the $g_i(s)$, we have also identified the Laplace transform of the busy period $f_i(s)$. This is seen as follows. Considering $f_i(s)$, with $i \in \mathcal{E}^+$, straightforward arguments yield

$$\begin{aligned} f_i(s) &= \sum_{k \neq i} \lambda_{ik} \int_0^\infty e^{-\lambda_i u} e^{-su} \mathbb{E}(e^{-sP} \mid Q(0) = r_i u, X(0) = k) du \\ &= \sum_{k \neq i} \frac{\lambda_{ik}}{r_i} \int_0^\infty \exp\left(-\frac{\lambda_i + s}{r_i} v\right) \mathbb{E}(e^{-sP} \mid Q(0) = v, X(0) = k) dv \\ &= \sum_{k \neq i} \frac{\lambda_{ik}}{r_i} f_k(s, t_i) = \frac{g_i(s)}{r_i}. \end{aligned}$$

The above findings are summarized in the following theorem.

Theorem 3.2. *For $s, t > 0$,*

$$\mathbf{f}(s, t) = (\mathbf{\Lambda} - s\mathbf{I} + t\mathbf{R})^{-1} \mathbf{g}(s),$$

with $\mathbf{g}^-(s) \equiv \mathbf{r}^-$ and $\mathbf{g}^+(s)$ given by (11) and obtained by solving

$$\det(\mathbf{\Lambda} - s\mathbf{I} + \tau_i(s)\mathbf{R} \mid \mathbf{g}(s), k) = 0,$$

for $i = 1, \dots, N^+$; here $\tau_i(s)$, $i = 1, \dots, N^+$, are, for $s > 0$ given, the N^+ values of τ in the right half-plane that satisfy $\det(\mathbf{\Lambda} - s\mathbf{I} + \tau\mathbf{R}) = 0$.

Furthermore, the Laplace transform of the busy period, starting in state $i \in \mathcal{E}^+$, is given by

$$f_i(s) = \frac{g_i(s)}{r_i}, \quad i = 1, \dots, N^+. \tag{14}$$

4 Busy period asymptotics

In this section we derive the logarithmic asymptotics of the probability that, starting in an up-state i , the busy period lasts longer than t , for $t \rightarrow \infty$. To this end, we define the *cumulant function*

$$\Gamma(\vartheta) := \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}(\exp(\vartheta A(t))),$$

here $A(t) := \int_0^t r_{X(s)} ds$; notice that the cumulant function can be regarded as an asymptotic log-moment generating function, and is, as a consequence, convex. $\Gamma'(0)$ equals the drift $\sum_{i=1}^N r_i \pi_i$, which we assumed to be negative. Define by ϑ^* the minimizer of $\Gamma(\vartheta)$; observe that necessarily $\vartheta^* > 0$ and $\Gamma(\vartheta^*) < 0$.

We introduce the short notation $\mathbb{P}_i(\cdot) := \mathbb{P}(\cdot \mid X(0) = i)$, and $\varrho_i(t) := \mathbb{P}_i(P > t \mid Q(0) = 0)$. We can now state the main result of this section.

Theorem 4.1. *For $i \in \mathcal{E}^+$*

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \varrho_i(t) = \Gamma(\vartheta^*);$$

here ϑ^* is the minimizing point of $\Gamma(\vartheta)$.

Proof. Let $i \in \mathcal{E}^+$. First we prove the upper bound. Evidently, we have

$$\varrho_i(t) \leq \mathbb{P}(A(t) > 0 \mid X(0) = i).$$

Now the Gärtner-Ellis theorem [8] immediately yields, with $A[k] := A(k+1) - A(k)$, that

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \varrho_i(t) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_i \left(\sum_{k=0}^{n-1} A[k] > 0 \right) \\ &= - \sup_{\vartheta \geq 0} (\vartheta \cdot 0 - \Gamma(\vartheta)) = \inf_{\vartheta \geq 0} \Gamma(\vartheta). \end{aligned}$$

The infimum of $\Gamma(\vartheta)$ is attained at ϑ^* as introduced above, which proves the upper bound.

We now proceed by proving the lower bound. For any $\delta \in (0, 1)$,

$$\varrho_i(t) \geq \mathbb{P}_i(\forall u \in [0, \delta t] : X(u) = i; \forall s \in (\delta t, t] : A(s) > 0).$$

But using the conditional independence, the expression in the right-hand side of the previous display equals

$$\mathbb{P}_i(\forall u \in [0, \delta t] : X(u) = i) \mathbb{P}_i(A(s) > -r_i \delta t, \forall s \in (0, (1 - \delta)t]).$$

The first factor of the above display is equal to $e^{-\lambda_i \delta t}$. Now concentrate on the second factor, with $r_{\max} := \max_{i \in \mathcal{E}} r_i$:

$$\begin{aligned}
& \frac{1}{t} \log \mathbb{P}_i (A(s) > -r_i \delta t, \forall s \in (0, (1-\delta)t]) = \\
& = \frac{1}{t} \log \mathbb{P}_i (A(0, \gamma(1-\delta)t) > -r_i \delta t, \forall \gamma \in [0, 1]) \\
& = \frac{1}{t} \log \mathbb{P}_i \left(\frac{A(0, \gamma(1-\delta)t)}{(1-\delta)t} > -r_i \frac{\delta}{1-\delta}, \forall \gamma \in [0, 1] \right) \\
& \geq \frac{1}{t} \log \mathbb{P}_i \left(\frac{A(0, \lceil \gamma(1-\delta)t \rceil)}{\lceil (1-\delta)t \rceil} > -r_i \frac{\delta}{1-\delta} + r_{\max} \frac{1}{(1-\delta)t}, \forall \gamma \in [0, 1] \right) \\
& \geq \frac{1}{t} \log \mathbb{P}_i \left(\frac{A(0, \lceil \gamma(1-\delta)t \rceil)}{\lceil (1-\delta)t \rceil} > -\frac{r_i}{2} \frac{\delta}{(1-\delta)}, \forall \gamma \in [0, 1] \right);
\end{aligned}$$

in the first inequality we use the fact that

$$A(0, \gamma(1-\delta)t) = A(0, \lceil \gamma(1-\delta)t \rceil) - A(\gamma(1-\delta)t, \lceil \gamma(1-\delta)t \rceil) \geq A(0, \lceil \gamma(1-\delta)t \rceil) - r_{\max},$$

and the last inequality holds for all $t > t^* := 2r_{\max}/(r_i \delta)$.

The process $A_n(\gamma) := n^{-1} \cdot A(0, \lceil n\gamma \rceil)$, with $\gamma \in [0, 1]$ and $n := \lceil (1-\delta)t \rceil$ fits in the framework of the sample-path large deviations principle in Example 2.5 of Chang [6]. As a consequence,

$$\begin{aligned}
& (1-\delta) \liminf_{t \rightarrow \infty} \frac{1}{(1-\delta)t} \log \mathbb{P}_i \left(\frac{A(0, \lceil \gamma(1-\delta)t \rceil)}{\lceil (1-\delta)t \rceil} > -\frac{r_i}{2} \frac{\delta}{(1-\delta)}, \forall \gamma \in [0, 1] \right) \\
& \geq -(1-\delta) \inf_{f \in \mathcal{A}^\circ} \mathbb{I}(f),
\end{aligned}$$

where $\mathbb{I}(\cdot)$ is the ‘rate functional’:

$$\mathbb{I}(f) := \int_0^1 \sup_{\vartheta} (\vartheta f'(t) - \Gamma(\vartheta)) dt,$$

and \mathcal{A}° is the interior of \mathcal{A} , which is the set of paths of interest:

$$\mathcal{A} := \left\{ f \in \text{AC}([0, 1], (\mathbb{R}, \|\cdot\|_\infty)) : f(\gamma) > -\frac{r_i}{2} \frac{\delta}{(1-\delta)}, \forall \gamma \in [0, 1] \right\}.$$

Here $\text{AC}([0, 1], (\mathbb{R}, \|\cdot\|_\infty))$ is the space of absolutely continuous function f such that $f(0) = 0$, equipped with the supremum norm topology, i.e.,

$$\|f\|_\infty = \sup_{t \in [0, 1]} |f(t)|.$$

It is seen that the set \mathcal{A} is open (and consequently $\mathcal{A} = \mathcal{A}^\circ$), as follows. Since the functions considered are absolutely continuous, thus continuous, over the *closed* interval $[0, 1]$, any function f attains a minimum at some point $\gamma_f \in [0, 1]$; as $f \in \mathcal{A}$, we have that

$$f(\gamma_f) > -\frac{r_i}{2} \frac{\delta}{(1-\delta)}.$$

Then consider the ball $B(f, \epsilon)$ around f with radius

$$\epsilon := \frac{1}{2} \left(f(\gamma_f) + \frac{r_i}{2} \frac{\delta}{(1-\delta)} \right) > 0;$$

this ball is evidently contained in \mathcal{A} , and hence \mathcal{A} is open.

Then observe that the path $f_0 \equiv 0$ is in $\mathcal{A}^o = \mathcal{A}$. Hence

$$-\inf_{f \in \mathcal{A}} \mathbb{I}(f) \geq -\mathbb{I}(f_0) = -\sup_{\vartheta} (0 - \Gamma(\vartheta)) = \Gamma(\vartheta^*).$$

Summarizing, we have

$$\begin{aligned} \liminf_{t \rightarrow \infty} \frac{1}{t} \log \varrho_i(t) &\geq \liminf_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}_i(X(\delta t) = i) \\ &\quad + \liminf_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}_i(A(s) > -r_i \delta t, \forall s \in (0, (1-\delta)t]) \\ &\geq -\lambda_i \delta + (1-\delta) \Gamma(\vartheta^*). \end{aligned}$$

Now letting $\delta \downarrow 0$ yields the lower bound. □

5 The covariance function of the workload process

In this section we analyze the Laplace transform of the covariance function of the workload process $Q(u)$. Let $R(u) := \text{Cov}(Q(0), Q(u))$ and $\gamma(\vartheta)$ be its Laplace transform, i.e.,

$$\gamma(\vartheta) := \int_0^\infty e^{-\vartheta u} R(u) du = \int_0^\infty e^{-\vartheta u} [\mathbb{E}(Q(0)Q(u)) - \mathbb{E}Q(0)\mathbb{E}Q(u)] du.$$

As an important special case, we later consider the situation that $Q(0)$ is distributed according to the stationary distribution (3). Then the correlation coefficient between $Q(0)$ and $Q(u)$ reads

$$\text{Corr}(Q(0), Q(u)) := \frac{\text{Cov}(Q(0), Q(u))}{\sqrt{\text{Var}Q(0) \cdot \text{Var}Q(u)}} = \frac{\mathbb{E}(Q(0)Q(u)) - (\mathbb{E}Q(0))^2}{\text{Var}Q(0)},$$

using that the queue is still in equilibrium at time u .

In our derivation of the Laplace transform $\gamma(\vartheta)$, we condition on the state of the system at time zero. More specifically, we throughout assume that the probability distribution of $(Q(0), X(0))$ is given, and denote its density by $\mathbf{p}^0(x) \equiv (p_1^0(x), \dots, p_N^0(x))^T$, for $x \geq 0$ and $i \in \mathcal{E}$; as indicated above we will later specialize to the special case of $\mathbf{p}^0(x) = \mathbf{p}(x)$, with $\mathbf{p}(x)$ distributed according to (3). In the sequel let T be an exponentially distributed random variable with parameter ϑ , independently of the modulating process $X(t)$. For $s \geq 0$, $t > 0$, we also introduce

$$\begin{aligned} \eta_i(\vartheta, s | x) &:= \mathbb{E} \left(e^{-sQ(T)} \mid Q(0) = x, X(0) = i \right); \\ \ell_i(\vartheta, s, t) &:= \int_0^\infty e^{-tx} \eta_i(\vartheta, s | x) dx. \end{aligned} \tag{15}$$

For later use, define $\vartheta_i := \vartheta + \lambda_i$, and in addition the ‘derivatives’ of $\boldsymbol{\eta}(\vartheta, s | x)$ and $\boldsymbol{\ell}(\vartheta, s, t)$, for $i \in \mathcal{E}$:

$$\eta_i^{(s)}(\vartheta, s | x) := \frac{\partial}{\partial s} \eta_i(\vartheta, s | x), \quad \ell_i^{(s)}(\vartheta, s, t) := \frac{\partial}{\partial s} \ell_i(\vartheta, s, t), \quad \ell_i^{(s,t)}(\vartheta, s, t) := \frac{\partial^2}{\partial s \partial t} \ell_i(\vartheta, s, t).$$

These functions will turn out to play a pivotal role in determining the Laplace transform of the covariance function $R(u)$. The following lemma relates $\boldsymbol{\eta}(\vartheta, s | 0)$ and $\boldsymbol{\ell}(\vartheta, s, t)$.

Lemma 5.1. *The vectors $\boldsymbol{\eta}(\vartheta, s | 0)$ and $\boldsymbol{\ell}(\vartheta, s, t)$ satisfy the following relation:*

$$(tr_i - \vartheta) \ell_i(\vartheta, s, t) + \sum_{k \in \mathcal{E}} \lambda_{ik} \ell_k(\vartheta, s, t) = r_i \eta_i(\vartheta, s | 0) - \frac{\vartheta}{s + t}, \quad (16)$$

for $i \in \mathcal{E}$. In addition,

$$\eta_i(\vartheta, s | 0) = \sum_{k \neq i} \frac{\lambda_{ik}}{r_i} \cdot \ell_k \left(\vartheta, s, \frac{\vartheta_i}{r_i} \right) + \frac{\vartheta}{sr_i + \vartheta_i}, \quad i \in \mathcal{E}^+; \quad (17)$$

$$\eta_i(\vartheta, s | 0) = \sum_{k \neq i} \frac{\lambda_{ik}}{\vartheta_i} \cdot \eta_k(\vartheta, s | 0) + \frac{\vartheta}{\vartheta_i}, \quad i \in \mathcal{E}^-. \quad (18)$$

Proof. First we stress that there are strong similarities between this proof and the steps used in Section 3, when we determined the Laplace transform of the busy period.

Notice that for $i \in \mathcal{E}^+$ the buffer cannot become empty before the first jump of the modulating Markov process, whereas for $i \in \mathcal{E}^-$ this is possible; we deal with the two cases differently. First consider the case $i \in \mathcal{E}^+$. Then, conditioning on the jump epoch of the modulating Markov process,

$$\begin{aligned} \ell_i(\vartheta, s, t) &= \int_0^\infty e^{-tx} \int_0^\infty \vartheta_i e^{-\vartheta_i u} \left\{ \sum_{k \neq i} \frac{\lambda_{ik}}{\vartheta_i} \eta_k(\vartheta, s | x + r_i u) + \frac{\vartheta}{\vartheta_i} e^{-s(x+r_i u)} \right\} dx du \\ &= \sum_{k \neq i} \frac{\lambda_{ik}}{tr_i - \vartheta_i} \left(\ell_k \left(\vartheta, s, \frac{\vartheta_i}{r_i} \right) - \ell_k(\vartheta, s, t) \right) + \frac{\vartheta}{(s+t)(sr_i + \vartheta_i)}; \end{aligned}$$

the last equality followed after elementary calculus. The above relation can then be rewritten to

$$(tr_i - \vartheta_i) \ell_i(\vartheta, s, t) + \sum_{k \neq i} \lambda_{ik} \ell_k(\vartheta, s, t) = \sum_{k \neq i} \lambda_{ik} \ell_k \left(\vartheta, s, \frac{\vartheta_i}{r_i} \right) + \frac{\vartheta(tr_i - \vartheta_i)}{(s+t)(sr_i + \vartheta_i)}. \quad (19)$$

Now consider the case $i \in \mathcal{E}^-$. Taking into account that the buffer can become empty before the first jump of the modulating Markov process,

$$\begin{aligned} \ell_i(\vartheta, s, t) &= \int_0^\infty e^{-tx} \int_0^\infty \vartheta_i e^{-\vartheta_i u} \left\{ \sum_{k \neq i} \frac{\lambda_{ik}}{\vartheta_i} \eta_k(\vartheta, s | (x + r_i u)^+) + \frac{\vartheta}{\vartheta_i} e^{-s(x+r_i u)^+} \right\} dx du \\ &= - \sum_{k \neq i} \frac{\lambda_{ik}}{tr_i - \vartheta_i} \left\{ \ell_k(\vartheta, s, t) - \frac{r_i}{\vartheta_i} \eta_k(\vartheta, s | 0) \right\} + \frac{\vartheta((s+t)r_i - \vartheta_i)}{(s+t)\vartheta_i(tr_i - \vartheta_i)}, \end{aligned}$$

which can be written as

$$(tr_i - \vartheta_i)\ell_i(\vartheta, s, t) + \sum_{k \neq i} \lambda_{ik} \ell_k(\vartheta, s, t) = \sum_{k \neq i} \frac{\lambda_{ik} r_i}{\vartheta_i} \eta_k(\vartheta, s | 0) + \frac{\vartheta((s+t)r_i - \vartheta_i)}{\vartheta_i(s+t)}. \quad (20)$$

Now let us evaluate $\eta_i(\vartheta, s | 0)$ further. First consider the case $i \in \mathcal{E}^+$; then the buffer immediately becomes non-empty, and hence

$$\begin{aligned} \eta_i(\vartheta, s | 0) &= \int_0^\infty \vartheta_i e^{-\vartheta_i u} \left\{ \sum_{k \neq i} \frac{\lambda_{ik}}{\vartheta_i} \eta(\vartheta, s | r_i u, k) + \frac{\vartheta}{\vartheta_i} e^{-sr_i u} \right\} du \\ &= \sum_{k \neq i} \frac{\lambda_{ik}}{r_i} \cdot \ell_k \left(\vartheta, s, \frac{\vartheta_i}{r_i} \right) + \frac{\vartheta}{sr_i + \vartheta_i}, \end{aligned}$$

which proves (17). For $i \in \mathcal{E}^-$ the buffer remains empty until the first jump, and hence

$$\begin{aligned} \eta_i(\vartheta, s | 0) &= \int_0^\infty \vartheta_i e^{-\vartheta_i u} \left\{ \sum_{k \neq i} \frac{\lambda_{ik}}{\vartheta_i} \cdot \eta_k(\vartheta, s | 0) + \frac{\vartheta}{\vartheta_i} \right\} du \\ &= \sum_{k \neq i} \frac{\lambda_{ik}}{\vartheta_i} \eta_k(\vartheta, s | 0) + \frac{\vartheta}{\vartheta_i}, \end{aligned}$$

which proves (18). Equation (16) is obtained by inserting (17)–(18) into (19)–(20). \square

We now explain how the transform $\ell(\vartheta, s, t)$ can be identified. Equation (16) can be rewritten in matrix form:

$$(\mathbf{\Lambda} - \vartheta \mathbf{I} + t \mathbf{R}) \boldsymbol{\ell}(\vartheta, s, t) = \mathbf{w}(\vartheta, s, t); \quad (21)$$

here $w_i(\vartheta, s, t) := r_i \eta_i(\vartheta, s | 0) - \vartheta/(s+t)$, for $i \in \mathcal{E}$. In other words, assuming for the moment that $\boldsymbol{\eta}(\vartheta, s | 0)$ is known, application of Cramer's rule leads to

$$\ell_i(\vartheta, s, t) = \frac{\det(\mathbf{\Lambda} - \vartheta \mathbf{I} + t \mathbf{R} | \mathbf{w}(\vartheta, s, t), i)}{\det(\mathbf{\Lambda} - \vartheta \mathbf{I} + t \mathbf{R})}, \quad i \in \mathcal{E}. \quad (22)$$

It is observed that, if we are able to determine $\boldsymbol{\eta}(\vartheta, s | 0)$, then we have found $\ell(\vartheta, s, t)$. We now identify N linear equations that enable us to compute $\boldsymbol{\eta}(\vartheta, s | 0)$. Equation (18) already gives N^- equations, so that it remains to determine the other N^+ linear equations. Since ϑ is fixed, using Lemma 3.1 with ϑ instead of s , there are N^+ values $\tau_i \equiv \tau_i(\vartheta)$ ($i = 1, \dots, N^+$) in the right half-plane satisfying $\det(\mathbf{\Lambda} - \vartheta \mathbf{I} + \tau_i \mathbf{R}) = 0$. Since Equation (22) should give a finite norm for any $\vartheta > 0$, these N^+ values $\tau_i \equiv \tau_i(\vartheta)$ should also satisfy

$$\det(\mathbf{\Lambda} - \vartheta \mathbf{I} + \tau_i \mathbf{R} | \mathbf{w}(\vartheta, s, \tau_i), k) = 0, \quad i \in \mathcal{E}^+.$$

In other words: we have now obtained N linear equations; solving these yields $\boldsymbol{\eta}(\vartheta, s | 0)$.

Above we developed a procedure for determining $\ell(\vartheta, s, t)$. Clearly $\ell(\vartheta, s, t)$ uniquely defines $\boldsymbol{\eta}(\vartheta, s | x)$, see (15). We now focus on how this procedure can be used to obtain an expression for the Laplace transform $\gamma(\cdot)$ of the covariance function.

Theorem 5.2. For $\vartheta > 0$,

$$\gamma(\vartheta) = \frac{1}{\vartheta} \cdot \sum_{i \in \mathcal{E}} \int_0^\infty [\mathbb{E}Q(0) - x] \eta_i^{(s)}(\vartheta, 0 | x) p_i^0(x) dx, \quad (23)$$

where

$$\mathbb{E}Q(0) = \sum_{i \in \mathcal{E}} \int_0^\infty x p_i^0(x) dx.$$

Proof. With $\mathbb{E}_{x,i}(\cdot) := \mathbb{E}(\cdot | Q(0) = x, X(0) = i)$, conditioning on the state of the system at time 0 yields

$$\begin{aligned} \vartheta \gamma(\vartheta) &= \int_0^\infty \vartheta e^{-\vartheta u} R(u) du \\ &= \int_0^\infty \vartheta e^{-\vartheta u} \left(\sum_{i \in \mathcal{E}} \int_0^\infty [x \mathbb{E}_{x,i}(Q(u)) - \mathbb{E}Q(0) \mathbb{E}_{x,i}(Q(u))] p_i^0(x) dx \right) du \\ &= \sum_{i \in \mathcal{E}} \int_0^\infty [x - \mathbb{E}Q(0)] \left(\int_0^\infty \vartheta e^{-\vartheta u} \mathbb{E}_{x,i}(Q(u)) du \right) p_i^0(x) dx \\ &= \sum_{i \in \mathcal{E}} \int_0^\infty [\mathbb{E}Q(0) - x] \eta_i^{(s)}(\vartheta, 0 | x) p_i^0(x) dx, \end{aligned}$$

where in the last equality we used the fact that $\eta_i^{(s)}(\vartheta, 0 | x) = - \int_0^\infty \vartheta e^{-\vartheta u} \mathbb{E}_{x,i} Q(u) du$. \square

Now we consider a number of special cases. If the density $\mathbf{p}^0(x)$ is given by

$$p_i^0(x) = \sum_{j=1}^k \sigma_{ij} e^{-\zeta_j x}, \quad x \geq 0, \quad i \in \mathcal{E}; \quad (24)$$

for constants σ_{ij} and $\zeta_j > 0$, then the Laplace transform $\gamma(\vartheta)$ is given by

$$\gamma(\vartheta) = \frac{1}{\vartheta} \cdot \sum_{i \in \mathcal{E}} \sum_{j=1}^k \sigma_{ij} \left[\ell_i^{(s,t)}(\vartheta, 0, \zeta_j) + \mathbb{E}Q(0) \ell_i^{(s)}(\vartheta, 0, \zeta_j) \right]. \quad (25)$$

Formula (25) extends in a straightforward way to the case in which among the ζ_j there are pairs of complex conjugates (with necessarily positive real parts). Importantly, this observation entails that we have now identified the Laplace transform of the covariance function in case $Q(0)$ obeys the stationary workload distribution (3). Also the case of eigenvalues with multiplicity k larger than 1 can be solved; then the density of the stationary workload has terms proportional to $x^j e^{-\zeta_j x}$, with $j = 0, \dots, k-1$, which is reflected in the appearance of higher order derivatives of $\ell_i(\vartheta, s, t)$ (in which $s = 0$ and $t = \zeta_j$ should be inserted) in the expression for $\gamma(\vartheta)$.

6 Example

The following example illustrates the results of this paper. We concentrate on the two-state case, and compute the busy-period distribution, as well as the covariance function. For α, β, λ and μ positive, we denote

$$\mathbf{\Lambda} := \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}, \quad \mathbf{r} := \begin{pmatrix} \alpha \\ -\beta \end{pmatrix}, \quad \boldsymbol{\pi} := \psi \begin{pmatrix} \mu \\ \lambda \end{pmatrix}, \quad \text{with } \psi := \frac{1}{\lambda + \mu};$$

it is easily verified that $\boldsymbol{\pi}$ is the invariant distribution of $\mathbf{\Lambda}$. We call the state space $\mathcal{E} = \{+, -\}$. The stability condition is satisfied if $\beta\lambda > \alpha\mu$, which in the sequel is assumed to hold true. The pairs of eigenvalues-vectors of $\mathbf{R}^{-1}\mathbf{\Lambda}^T$ are given by

$$\xi_0 = 0, \quad \text{with } v^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad \xi_+ = \frac{\alpha\mu - \beta\lambda}{\alpha\beta}, \quad \text{with } v^{(+)} = \begin{pmatrix} \lambda/\alpha \\ \mu/\beta \end{pmatrix};$$

note that $\xi_+ < 0$. The stationary distribution of $(Q \leq x, X = \pm)$ and its density are

$$\begin{aligned} \mathbb{P}(Q \leq x, X = +) &= \mu\psi (1 - \exp(\xi_+x)), & p_+(x) &= -\mu\xi_+\psi \exp(\xi_+x); \\ \mathbb{P}(Q \leq x, X = -) &= \lambda\psi \left(1 - \frac{\alpha\mu}{\beta\lambda} \exp(\xi_+x)\right), & p_-(x) &= -\frac{\alpha}{\beta}\mu\xi_+\psi \exp(\xi_+x). \end{aligned} \quad (26)$$

The mean and variance of Q are finite and given by

$$\mathbb{E}Q = \frac{\alpha\mu(\alpha + \beta)}{(\lambda + \mu)(\beta\lambda - \alpha\mu)}, \quad \text{Var } Q = \alpha^2 \left(\frac{\beta^2}{(\beta\lambda - \alpha\mu)^2} - \frac{1}{(\lambda + \mu)^2} \right).$$

Busy period. We first determine the distribution of the busy period P , as well as its tail asymptotics. The system (5)–(6) can be rewritten as, with $f_{ij}(s) := f_i(s, t_j(s))$, for $i, j \in \{+, -\}$,

$$\begin{cases} -(\lambda + s - \alpha t)f_+(s, t) + \lambda f_-(s, t) &= \lambda f_{-+}(s); \\ \mu f_+(s, t) - (\mu + s + \beta t)f_-(s, t) &= -\beta. \end{cases}$$

From the second equation we have, by instering $t = (\lambda + s)/\alpha$,

$$f_{-+}(s) = \frac{\alpha}{\alpha(\mu + s) + \beta(\lambda + s)} (\mu f_{++}(s) + \beta),$$

and hence the vector $\mathbf{g}(s)$ is given by

$$\mathbf{g}(s) = \left(\frac{\lambda\alpha(\mu f_{++}(s) + \beta)}{\alpha(\mu + s) + \beta(\lambda + s)}, -\beta \right)^T.$$

To determine $f_{++}(s)$, we first compute the zeros of the determinant of $(\mathbf{\Lambda} + t\mathbf{R} - s\mathbf{I})$ for given $s > 0$:

$$\tau_{\pm}(s) = \frac{1}{2\alpha\beta} \cdot \left(\beta(\lambda + s) - \alpha(\mu + s) \pm \sqrt{[\beta(\lambda + s) + \alpha(\mu + s)]^2 - 4\alpha\beta\lambda\mu} \right),$$

and focus on the positive root $\tau_+(s)$. It is clear that $\tau_+(s)$ must also be a zero of the determinant of $(\mathbf{A} + t\mathbf{R} - s\mathbf{I} \mid \mathbf{g}(s), +)$. It can now be verified that

$$f_{++}(s) = \frac{\beta^2}{\mu\alpha} \cdot \frac{\lambda + s - \alpha\tau_+(s)}{\mu + s + \beta\tau_+(s)}, \quad f_{-+}(s) = \frac{\beta}{\mu + s + \beta\tau_+(s)}.$$

The Laplace transform of P , starting off at buffer level 0 (so that the busy period necessarily starts in +), is then given by

$$f_+(s) = \frac{1}{2\alpha\mu} \left((\alpha + \beta)s + (\beta\lambda + \alpha\mu) - \sqrt{[(\alpha + \beta)s + (\beta\lambda + \alpha\mu)]^2 - 4\alpha\beta\lambda\mu} \right).$$

This transform can be explicitly inverted, yielding the density of the busy-period; with, as introduced in Section 4, $\varrho_i(t) := \mathbb{P}_i(P > t \mid Q(0) = 0)$, we have that the density of the busy period equals

$$\frac{d}{dt}\varrho_+(t) = \sqrt{\frac{\beta\lambda}{\alpha\mu}} \cdot \mathbb{I}_1 \left(2\frac{\sqrt{\alpha\beta\lambda\mu}}{\alpha + \beta} t \right) \cdot \frac{1}{t} \exp \left(-\frac{\alpha\mu + \beta\lambda}{\alpha + \beta} t \right);$$

here $\mathbb{I}_1(x)$ is the modified Bessel function of the first kind. By differentiating $f_+(s)$ and inserting $s = 0$ we can now find all moments of the busy-period P . The first moment is

$$\mathbb{E}_+(P \mid Q(0) = 0) = -f'_+(0) = \frac{\alpha + \beta}{\beta\lambda - \alpha\mu}.$$

The asymptotics of the density and the tail distribution are given by, as $t \rightarrow \infty$,

$$\begin{aligned} \frac{d}{dt}\varrho_+(t) &\sim \frac{(\beta\lambda)^{1/4}}{(\alpha\mu)^{3/4}} \frac{\sqrt{\alpha + \beta}}{2\sqrt{\pi}} \cdot \frac{1}{t\sqrt{t}} \exp \left(-\frac{(\sqrt{\beta\lambda} - \sqrt{\alpha\mu})^2}{\alpha + \beta} t \right), \\ \varrho_+(t) &\sim \frac{(\beta\lambda)^{1/4}}{(\alpha\mu)^{3/4}} \frac{(\alpha + \beta)^{3/2}}{2\sqrt{\pi}(\sqrt{\beta\lambda} - \sqrt{\alpha\mu})^2} \cdot \frac{1}{t\sqrt{t}} \exp \left(-\frac{(\sqrt{\beta\lambda} - \sqrt{\alpha\mu})^2}{\alpha + \beta} t \right); \end{aligned} \quad (27)$$

here ' \sim ' means that the ratio of both sides tends to 1 as $t \rightarrow \infty$.

Now consider a more specific example. Taking $\alpha = \lambda = \mu = 1$ and $\beta = 2$, we are in the setting of [3, Section 9]. Then

$$f_+(s) = \frac{3(1 + s) - \sqrt{9(1 + s)^2 - 8}}{2}; \quad \mathbb{E}_+(P) = 3,$$

in agreement with the findings of [3]. We find, however, a number of new results:

$$\frac{d}{dt}\varrho_+(t) = \sqrt{2} \cdot \frac{1}{t} e^{-t} \mathbb{I}_1 \left(\frac{2\sqrt{2}}{3} t \right), \quad t > 0; \quad \varrho_+(t) \sim \frac{\sqrt{3}(3 + 2\sqrt{2})}{\sqrt{2\sqrt{2}} \cdot \pi} \cdot \frac{1}{t\sqrt{t}} e^{-\frac{3-2\sqrt{2}}{3}t}, \quad t \rightarrow \infty.$$

Now consider the logarithmic asymptotics of $\mathbb{P}_+(P > t)$. The cumulant function is given by $\Gamma(\vartheta) = \log \mathbf{sp}(\mathbf{A} + \vartheta\mathbf{R})$, where $\mathbf{sp}(M)$ is the largest eigenvalue of the matrix M , see [11]. In our example,

$$\Gamma(\vartheta) = -\frac{(\lambda + \mu + (\beta - \alpha)\vartheta)}{2} + \frac{\sqrt{((\beta + \alpha)\vartheta + (\mu - \lambda))^2 + 4\mu\lambda}}{2}.$$

Furthermore we have under our stability condition that $\Gamma'(0) < 0$, and $\Gamma(\cdot)$ attains its minimum at

$$\vartheta^* = \frac{\lambda - \mu}{\beta + \alpha} + \sqrt{\frac{\lambda\mu}{\alpha\beta} \frac{\beta - \alpha}{\alpha + \beta}} = \frac{(\sqrt{\lambda\alpha} + \sqrt{\mu\beta})(\sqrt{\beta\lambda} - \sqrt{\alpha\mu})}{\sqrt{\alpha\beta}(\alpha + \beta)} > 0,$$

so that

$$\Gamma(\vartheta^*) = -\frac{(\sqrt{\beta\lambda} - \sqrt{\alpha\mu})^2}{\alpha + \beta} < 0.$$

Hence, by virtue of Thm. 4.1, the decay rate of $\mathbb{P}_+(P > t)$ is $\Gamma(\vartheta^*)$, which agrees with the asymptotics given in (27).

Covariance function. Equations (16) are written as

$$\begin{cases} -(\lambda + \vartheta - \alpha t)\ell_+(\vartheta, s, t) + \lambda\ell_-(\vartheta, s, t) = \alpha\eta_+(\vartheta, s | 0) - \frac{\vartheta}{s+t}; \\ \mu\ell_+(\vartheta, s, t) - (\mu + \vartheta + \beta t)\ell_-(\vartheta, s, t) = -\beta\eta_-(\vartheta, s | 0) - \frac{\vartheta}{s+t}, \end{cases}$$

whereas (18) reads

$$\eta_-(\vartheta, s | 0) = \frac{\mu}{\mu + \vartheta}\eta_+(\vartheta, s | 0) + \frac{\vartheta}{\mu + \vartheta}.$$

The vector $\mathbf{w}(s, t)$ is given by

$$\mathbf{w}(s, t) = \left(\alpha\eta_+(\vartheta, s | 0) - \frac{\vartheta}{s+t}, -\beta\eta_-(\vartheta, s | 0) - \frac{\vartheta}{s+t} \right)^T.$$

Let us first compute the zeros of the determinant of $(\mathbf{\Lambda} + t\mathbf{R} - \vartheta\mathbf{I})$, for given ϑ . We find

$$\tau_{\pm}(\vartheta) = \frac{1}{2\alpha\beta} \left(\beta(\lambda + \vartheta) - \alpha(\mu + \vartheta) \pm \sqrt{[\beta(\lambda + \vartheta) + \alpha(\mu + \vartheta)]^2 - 4\alpha\beta\lambda\mu} \right);$$

let $\tau(\vartheta)$ be the positive root. The procedure described in Section 5 now yields

$$\eta_+(\vartheta, s | 0) = \frac{\vartheta}{s + \tau(\vartheta)} \frac{\lambda\beta(s + \tau(\vartheta)) + (\vartheta + \mu)(\beta\tau(\vartheta) + \lambda + \mu + \vartheta)}{(\alpha(\vartheta + \mu)(\beta\tau(\vartheta) + \mu + \vartheta) - \lambda\mu\beta)}.$$

Then, due to Equation (17),

$$\eta_-(\vartheta, s | 0) = \frac{\vartheta}{\mu + \vartheta} \frac{(\beta\tau(\vartheta) + \mu + \vartheta)(\alpha(s + \tau(\vartheta)) + \mu) + \lambda\mu}{(s + \tau(\vartheta))(\alpha(\vartheta + \mu)(\beta\tau(\vartheta) + \mu + \vartheta) - \lambda\mu\beta)}.$$

It now follows that

$$\begin{aligned} \ell_+(\vartheta, s, t) &= \frac{\left(\frac{\lambda\beta\mu}{\mu + \vartheta} - \alpha(\beta t + \mu + \vartheta) \right) \eta_+(\vartheta, s | 0) + \frac{\vartheta}{s+t}(\beta t + \lambda + \mu + \vartheta) + \frac{\lambda\beta\vartheta}{\mu + \vartheta}}{-\alpha\beta t^2 + [\beta\lambda - \alpha\mu + \vartheta(\beta - \alpha)]t + \vartheta(\vartheta + \lambda + \mu)} \\ \ell_-(\vartheta, s, t) &= \frac{(\beta(\lambda + \vartheta - \alpha t) - \alpha(\mu + \vartheta)) \eta_-(\vartheta, s | 0) + \frac{\vartheta}{s+t}(\alpha s + \lambda + \mu + \vartheta)}{-\alpha\beta t^2 + [\beta\lambda - \alpha\mu + \vartheta(\beta - \alpha)]t + \vartheta(\vartheta + \lambda + \mu)}. \end{aligned}$$

If the distribution of $(Q(0), X(0))$ is given we can then compute the Laplace transform of the covariance function, by relying on Thm. 5.2. In the remainder of this example we consider the situation that the system is in stationarity at time 0, and hence $p^0(x)$ is given by (26).

Since the formulae of the above functions are long and cumbersome, we prefer to treat a more specific example: as in [3], we choose $\alpha = \lambda = \mu = 1$ and $\beta = 2$. Then we have

$$\ell_+^{(s,t)}(\vartheta, 0, \xi_+) = 2 \frac{-8 - 77\vartheta + 392\vartheta^2 + 549\vartheta^3 + 232\vartheta^4 + 32\vartheta^5 + (8 + 5\vartheta + \vartheta^2)\sqrt{9(1 + \vartheta)^2 - 8}}{\vartheta^2(\vartheta + 2)(2\vartheta + 5)^2},$$

$$\ell_-^{(s,t)}(\vartheta, 0, \xi_+) = 2 \frac{-4 - 39\vartheta + 267\vartheta^2 + 432\vartheta^3 + 208\vartheta^4 + 32\vartheta^5 + (4 + 3\vartheta)\sqrt{9(1 + \vartheta)^2 - 8}}{\vartheta^2(\vartheta + 2)(2\vartheta + 5)^2}.$$

The Laplace transform $\rho(\vartheta)$ is given by

$$\rho(\vartheta) = \frac{-4 - 37\vartheta + 15\vartheta^2(\vartheta + 3)(2\vartheta + 3) + (\vartheta + 4)\sqrt{9(1 + \vartheta)^2 - 8}}{15\vartheta^3(\vartheta + 2)(2\vartheta + 5)}.$$

Clearly $\rho(\vartheta)$ is a well-defined function for $\vartheta \geq 0$, but from its expression we observe that $\rho(\vartheta)$ can be continued analytically to the left up to the point $\bar{\vartheta} = -(\sqrt{\beta\lambda} - \sqrt{\alpha\mu})^2/(\alpha + \beta) = -\frac{1}{3}(\sqrt{2} - 1)^2$.

Around this branching point,

$$\rho(\vartheta) \sim \frac{\sqrt[4]{2}(4\sqrt{2} + 18)}{15\sqrt{3} \left(\frac{2\sqrt{2}}{3} - 1\right)^3 \left(\frac{2\sqrt{2}}{3} + 1\right)^3 \left(\frac{4\sqrt{2}}{3} + 3\right)} \sqrt{\vartheta + \frac{1}{3}(\sqrt{2} - 1)^2}, \text{ as } \vartheta \downarrow \bar{\vartheta}.$$

Relying on standard techniques, we have, for $t \rightarrow \infty$,

$$r(t) \sim \frac{\sqrt[4]{2}(2\sqrt{2} + 9)}{15\sqrt{3\pi} \left(1 - \frac{2\sqrt{2}}{3}\right)^3 \left(\frac{2\sqrt{2}}{3} + 1\right) \left(\frac{4\sqrt{2}}{3} + 3\right)} \cdot \frac{1}{t\sqrt{t}} \exp\left(-\frac{3 - 2\sqrt{2}}{3}t\right).$$

Notice that in this example the asymptotics of the busy-period distribution and the correlation function coincide up to a constant factor; we have encountered the same proportionality property for queues with spectrally-positive Lévy input in [9].

7 Concluding remarks

In this paper we have considered transient characteristics of a Markov-fluid-driven queue, viz., the busy period and the covariance function of the workload process, by studying their Laplace transforms. In the case of the busy period we used sample-path large deviations to obtain its logarithmic asymptotics.

We conclude by listing a number of open issues. In Thm. 4.1 we found the logarithmic asymptotics of tail distribution of the busy period. The results for the two-state case in Section 6, however, lead to the conjecture that, for $i \in \mathcal{E}^+$, there is a constant ω such that

$$\mathbb{P}_i(P > t \mid Q(0) = 0) \sim \frac{\omega}{t\sqrt{t}} e^{t\Gamma(\vartheta^*)}.$$

A similar relation can be conjectured for $\text{Cov}(Q(0), Q(u))$, as $u \rightarrow \infty$, in view of the findings of Section 6. In this case, however, not even the logarithmic asymptotics are known. It is not clear, for instance if (and, if yes, how) sample-path large deviations [6] are of any help here. In fact, it is not even clear *a priori* that $\text{Cov}(Q(0), Q(u))$ is positive and decreasing; in case of spectrally-positive Lévy input these properties were shown relying on the concept of completely monotone functions [9, 16].

References

- [1] S. Ahn and V. Ramaswami. Efficient algorithms for transient analysis of stochastic fluid flow models. *J. Appl. Probab.*, 42:531–549, 2005.
- [2] D. Anick, D. Mitra, and M. Sondhi. Stochastic theory of data-handling system with multiple sources. *Bell Syst. Tech. J.*, 61(8):1871–1894, 1977.
- [3] S. Asmussen. Busy period analysis, rare events and transient behavior in fluid models. *J. Appl. Math. Stoch. Anal.*, 7(3):269–299, 1994.
- [4] S. Asmussen and T. Rolski. Risk theory in a periodic environment: the Cramer-Lundberg approximation and Lundberg’s inequality. *Math. Oper. Res.*, 19:410–433, 1994.
- [5] N. Barbot, B. Sericola, and M. Telek. Distribution of the busy period in stochastic fluid models. *Stoch. Models*, 17:407–427, 2001.
- [6] C.-S. Chang. Sample path large deviations and intree networks. *Queueing Syst.*, 20:7–36, 1995.
- [7] A. da Silva Soares and G. Latouche. Matrix-analytic methods for fluid queues with finite buffers. *Perf. Eval.*, 63:295–314, 2006.
- [8] A. Dembo and O. Zeitouni. *Large deviations techniques and applications, 2nd edition*. Springer, New York, NY, USA, 1998.
- [9] A. Es-Saghouani and M. Mandjes. On the correlation structure of a Lévy-driven queue. *Submitted*, 2008.
- [10] S. Geršgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk. SSSR Ser. Mat.*, 1:749–454, 1931.
- [11] G. Kesidis, J. Walrand, and C.-S. Chang. Effective bandwidths for multiclass Markov fluids and other ATM sources. *IEEE/ACM Trans. Netw.*, 1:424–428, 1993.

- [12] L. Kosten. Stochastic theory of a multi-entry buffer (I). *Delft Progress Report, Series F*, 1:10–18, 1974.
- [13] L. Kosten. Stochastic theory of data-handling systems with groups of multiple sources. In H. Rudin and W. Bux, editors, *Performance of Computer-Communication Systems*, pages 321–331. Elsevier, 1984.
- [14] M. Mandjes and W. Scheinhardt. A fluid model for a relay node in an ad hoc network: evaluation of resource sharing policies. *To appear in J. Appl. Math. Stoch. Anal.*, 2008.
- [15] A. Narayanan and V. G. Kulkarni. First passage times in fluid models with an application to two priority fluid systems. In *Proceedings of the 2nd IPDS '96*, pages 166–175, 1996.
- [16] T. Ott. The covariance function of the virtual waiting-time process in an M/G/1 queue. *Adv. Appl. Prob.*, 9:158–168, 1977.
- [17] N. Prabhu. *Stochastic Storage Processes: Queues, Insurance Risk, Dams and Data Communication*, 2nd edition. Springer, New York, NY, USA, 1998.
- [18] Q. Ren and H. Kobayashi. Transient solution for the buffer behavior in statistical multiplexing. *Perf. Eval.*, 21:65–87, 1995.
- [19] W. Scheinhardt. *Markov-modulated and feedback fluid queues*. PhD thesis, University of Twente, The Netherlands, 1998.
- [20] P. Sonneveld. Some properties of the generalized eigenvalue problem $Mx = \lambda(\Gamma - cI)x$, where M is the infinitesimal generator of a Markov process, and Γ is a real diagonal matrix. *Delft University of Technology Report 04-02*, 2004.