On/Off Storage Systems with State Dependent Input, Output and Switching Rates

Onno Boxma* Haya Kaspi† Offer Kella‡§ David Perry ¶

November 10, 2003

Abstract

We consider a storage model which can be on or off. When on, the content increases at some state-dependent rate and the system can switch to the off state at a state-dependent rate as well. When off, the content decreases at some state-dependent rate (unless it is at zero) and the system can switch to the on position at a state-dependent rate. This process is a special case of a piecewise deterministic Markov process. We identify the stationary distribution and conditions for its existence and uniqueness.

Keywords: Fluid queue, on/off, storage system, piecewise deterministic Markov process

AMS 2000 Subject Classification: Primary 60K25

*EURANDOM and Department of Mathematics and Computer Science; Eindhoven University of Technology; P.O. Box 513; 5600 MB Eindhoven; The Netherlands and CWI; P.O. Box 94079; 1090 GB Amsterdam; The Netherlands (boxma@win.tue.nl)

†Faculty of Industrial Engineering and Management; Technion – Israel Institute of Technology, Haifa 32000; Israel (iehaya@techunix.technion.ac.il)

‡Department of Statistics; The Hebrew University of Jerusalem; Mount Scopus, Jerusalem 91905; Israel (mskella@mscc.huji.ac.il)

§Supported in part by grant 819/03 from the Israel Science Foundation.

¶Department of Statistics; University of Haifa; Haifa 31905; Israel (dperry@stat.haifa.ac.il)

1 Introduction

We consider a storage model which can be on or off. When on, the buffer content (storage level) increases at some state-dependent rate and the system can switch to the off state at a state-dependent rate as well. When off, the content decreases at some state-dependent rate (unless it is at zero) and the system can switch to the on position at a state-dependent rate. The two-dimensional Markov process $\mathcal{M} = \{(X_t, I_t) | t \geq 0\}$, where X_t denotes the buffer content and the "background state" I_t is alternatingly on and off, is a special case of a piecewise deterministic Markov process [5]. This type of model is often referred to as a fluid queue. Fluid queues have frequently been used to model production systems, as well as telecommunication systems at the burst level.

An interesting class of fluid queues is formed by the Markov-modulated fluid queues, in which $\{I_t | t \geq 0\}$ is an underlying Markov process that determines the rate at which the buffer content X_t increases or decreases; a key reference is [3]. An important recent generalization is the *feedback fluid queue* [1], in which not only is the buffer content determined by the background process, but also is the background process influenced by the buffer content. Feedback fluid queues may be used for studying the interaction between a communication network and its sources. E.g., feedback schemes in access communication networks were analyzed in [10, 11] via a feedback fluid queue.

The model under consideration in the present paper may be viewed as a feedback fluid queue with two background states and unlimited buffer content. The buffer content X_t decreases or increases with a rate which depends on I_t, while changes in I_t are determined by X_t. In addition, the rate at which X_t changes is also determined by its own level. In storage processes, there is a long tradition of studying storage systems, or *dams*, with a general release rate [6, 7]. Our main goals are to study conditions for the existence of a stationary distribution of the Markov process $\mathcal{M}$, and to determine that stationary distribution as well as some related quantities.

The paper is related to [8], which considers a fluid queue with i.i.d. (rather than state dependent) on and off time pairs and state dependent production and release rates. The paper also bears some relation to [14], which considers a feedback fluid queue with state-dependent release rates. Scheinhardt et al. [14] allow a finite number $N \geq 2$ of background states, but restrict themselves to a finite buffer content. The paper is also related to [4]. When restricting consideration of our model to the states in which the buffer content is decreasing, one obtains a *queue* with workload-dependent arrival and service rates, and with service requests that depend on the workload found at the time of their arrival. In [4] queues with workload-dependent arrival and service rates have been studied, mainly with the restriction of i.i.d. service requests. The close relation between fluid queues and ordinary queues has been extensively studied in [9].

The paper is organized as follows. Section 2 contains a model description. The stationary distribution of the Markov process under consideration is determined in Section 3. The conditions for the existence of a stationary distribution are quite intricate; they are discussed in Section 4. Sections 5, 6 and 7 are respectively devoted to the return time to a given buffer level, to the process restricted to down (off) or up (on) intervals, and to the stationary distribution of the buffer content at times when the process changes direction (from up to down or vice versa). In Section 8 we prove that, under some additional conditions, convergence to the stationary distribution is at an exponential rate.

2 The Model

We consider a storage system, or fluid queue, which may be modelled as a piecewise deterministic Markov process $\mathcal{M} = \{(X_t, I_t)|\ t \geq 0\}$ with state space $[0,\infty)\times\{0,1\}$. When in state $(x,1)$, the process decreases at rate $r_1(x)$ or switches to state $(x,0)$ with rate $\lambda_1(x)$. When in state $(x,0)$, the process increases at rate $r_0(x)$ or switches to state $(x,1)$ with rate $\lambda_0(x)$. We assume that $r_1(x)$ and $\lambda_1(x)$ are left continuous and that $r_0(x)$ and $\lambda_0(x)$ are right continuous. To avoid an infinite number of switches we assume that both $\lambda_0(\cdot)$ and $\lambda_1(\cdot)$ are bounded. We further assume that $r_0(x) > 0$ for all $x \geq 0$ and that $r_1(x) > 0$ for all $x > 0$ and $r_1(0) = 0$ and that $\lambda_1(0) > 0$ (so that the process does not get stuck in state $(0,1)$ forever). With these assumptions the extended generator of this process can be written as follows:

$$\begin{aligned} \mathcal{A}f(x,1) &= -r_1(x)f'(x,1) - \lambda_1(x)f(x,1) + \lambda_1(x)f(x,0), \\ \mathcal{A}f(x,0) &= \quad r_0(x)f'(x,0) - \lambda_0(x)f(x,0) + \lambda_0(x)f(x,1), \end{aligned} \tag{1}$$

for differentiable $f(x,i)$, where $f'(x,i)$ is the derivative with respect to x.

3 The Stationary Distribution

In this section we shall compute the unique stationary distribution for the Markov process $\mathcal{M} = \{(X_t, I_t)|\ t \geq 0\}$. To this end we notice that the process sampled at exponential times is Feller; i.e., $x \to U^1 f(x,j)$ where U^1 is the transition function of the process sampled at exp(1) distributed times (independent of our process), is continuous in x for all $f(x,j)$ that are continuous in x. When all states communicate, which will be the case under some natural conditions that are discussed in Section 4, $\mathcal{M}$ is an irreducible T-chain and the existence of a finite invariant measure for it will make it automatically Harris recurrent [12]. Furthermore, the equation $(I - \mathcal{A})U^1 = I$ implies that a measure μ that satisfies

$$\int \mathcal{A}f(x,0)\mu(\mathrm{d}x,0) + \int \mathcal{A}f(x,1)\mu(\mathrm{d}x,1) = 0,$$

is invariant for U^1 and by the results of [2] it is also invariant for the original process. If the measure μ is finite, $\mathcal{M}$ is immediately positive Harris recurrent. With that in mind we are now set to compute an invariant distribution for $\mathcal{M}$. Let us assume that a stationary distribution π exists and is of the form

$$\mathrm{E}_\pi f(X_t, I_t) = c_0 \int_0^\infty g_0(x) f(x,0)\mathrm{d}x + c_1 \int_0^\infty g_1(x) f(x,1)\mathrm{d}x + pf(0,1) \ , \tag{2}$$

where $g_i(\cdot)$ are densities (integrate to 1) and $c_0+c_1+p=1$. In particular, for every Borel set A, and with $1_A(0)$ denoting the indicator function of the event $\{0 \in A\}$, the stationary distribution π is given by

$$\begin{aligned} \pi(A,0) &= c_0 \textstyle\int_A g_0(x)\mathrm{d}x, \\ \pi(A,1) &= c_1 \textstyle\int_A g_1(x)\mathrm{d}x + p 1_A(0) \ . \end{aligned} \tag{3}$$

Theorem 1 *Assume that a stationary distribution π exists and is of the form (2).*
(i) If $\int_0^\epsilon \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right) du$ is finite for some $\epsilon > 0$ then

$$g_i(x) = \frac{1}{\alpha_i r_i(x)} e^{-\int_0^x \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right)\mathrm{d}u}, \tag{4}$$

where

$$\alpha_i = \int_0^\infty \frac{1}{r_i(x)} e^{-\int_0^x \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right)\mathrm{d}u} \mathrm{d}x \ ; \tag{5}$$

furthermore,

$$\begin{aligned} c_i &= \frac{\alpha_i}{\alpha_0 + \alpha_1 + \frac{1}{\lambda_1(0)}} \ , \qquad i = 0, 1, \\ p &= \frac{\frac{1}{\lambda_1(0)}}{\alpha_0 + \alpha_1 + \frac{1}{\lambda_1(0)}}. \end{aligned} \tag{6}$$

(ii) Otherwise,

$$g_i(x) = \frac{1}{\alpha_i r_i(x)} e^{-\int_\epsilon^x \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right)\mathrm{d}u}, \tag{7}$$

where $\int_\epsilon^x = -\int_x^\epsilon$ for $x < \epsilon$, and

$$\alpha_i = \int_0^\infty \frac{1}{r_i(x)} e^{-\int_\epsilon^x \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right)\mathrm{d}u} \mathrm{d}x \ ; \tag{8}$$

furthermore,

$$\begin{aligned} c_i &= \frac{\alpha_i}{\alpha_0 + \alpha_1} \ , \qquad i = 0, 1, \\ p &= 0. \end{aligned} \tag{9}$$

Proof: From the form of the generator we have that

$$
\begin{aligned}
0 &= c_0 \textstyle\int_0^\infty g_0(x)\mathcal{A}f(x,0)\mathrm{d}x + c_1 \int_0^\infty g_1(x)\mathcal{A}f(x,1)\mathrm{d}x + p\mathcal{A}f(0,1) \\
&= c_0 \textstyle\int_0^\infty [r_0(x)g_0(x)f'(x,0) - \lambda_0(x)g_0(x)f(x,0) + \lambda_0(x)g_0(x)f(x,1)]\mathrm{d}x \\
&\quad + c_1 \textstyle\int_0^\infty [-r_1(x)g_1(x)f'(x,1) - \lambda_1(x)g_1(x)f(x,1) + \lambda_1(x)g_1(x)f(x,0)]\mathrm{d}x \\
&\quad - p\lambda_1(0)f(0,1) + p\lambda_1(0)f(0,0)\ .
\end{aligned} \tag{10}
$$

Taking $f(x,1) \equiv f(x,0)$ we obtain

$$
c_0 \int_0^\infty r_0(x)g_0(x)f'(x,0)\mathrm{d}x = c_1 \int_0^\infty r_1(x)g_1(x)f'(x,0)\mathrm{d}x\ . \tag{11}
$$

Since (11) should hold for all differentiable functions $f(\cdot,0)$ for which the integrals are well defined, this yields the expected level crossings identity (rate up = rate down):

$$
c_0 r_0(x)g_0(x) = c_1 r_1(x)g_1(x). \tag{12}
$$

Taking either $f(x,0) \equiv 0$ and $f(x,1) \equiv 1$, or $f(x,0) \equiv 1$ and $f(x,1) \equiv 0$, gives

$$
c_0 \int_0^\infty \lambda_0(x)g_0(x)\mathrm{d}x = c_1 \int_0^\infty \lambda_1(x)g_1(x)\mathrm{d}x + p\lambda_1(0), \tag{13}
$$

which is also expected, as it implies that the stationary rate at which the second coordinate of the process moves from state 0 to state 1 is equal to the rate at which it moves from state 1 to state 0. This is also a type of level crossings identity.

We now substitute (13) in (10) to obtain

$$
\begin{aligned}
0 &= c_0 \textstyle\int_0^\infty \Big[r_0(x)g_0(x)f'(x,0) - \lambda_0(x)g_0(x)[f(x,0) - f(0,0)] \\
&\quad + \lambda_0(x)g_0(x)[f(x,1) - f(0,1)]\Big]\mathrm{d}x \\
&\quad + c_1 \textstyle\int_0^\infty \Big[- r_1(x)g_1(x)f'(x,1) - \lambda_1(x)g_1(x)[f(x,1) - f(0,1)] \\
&\quad + \lambda_1(x)g_1(x)[f(x,0) - f(0,0)]\Big]\mathrm{d}x\ .
\end{aligned} \tag{14}
$$

With $f(x,i) - f(0,i) = \int_0^x f'(u,i)\mathrm{d}u$ and a change of the order of integration, (14) becomes

$$
\begin{aligned}
0 &= \textstyle\int_0^\infty [c_0 r_0(x)g_0(x) - c_0 \int_x^\infty \lambda_0(u)g_0(u)\mathrm{d}u + c_1 \int_x^\infty \lambda_1(u)g_1(u)\mathrm{d}u]\, f'(x,0)\mathrm{d}x \\
&+ \textstyle\int_0^\infty [-c_1 r_1(x)g_1(x) - c_1 \int_x^\infty \lambda_1(u)g_1(u)\mathrm{d}u + c_0 \int_x^\infty \lambda_0(u)g_0(u)\mathrm{d}u]\, f'(x,1)\mathrm{d}x\ .
\end{aligned} \tag{15}
$$

Since (15) should hold for all differentiable functions for which the integrals are well defined, we obtain that necessarily

$$
0 = c_0 r_0(x)g_0(x) - c_0 \int_x^\infty \lambda_0(u)g_0(u)\mathrm{d}u + c_1 \int_x^\infty \lambda_1(u)g_1(u)\mathrm{d}u, \tag{16}
$$

as well as

$$0 = -c_1 r_1(x) g_1(x) - c_1 \int_x^\infty \lambda_1(u) g_1(u) \mathrm{d}u + c_0 \int_x^\infty \lambda_0(u) g_0(u) \mathrm{d}u \ . \tag{17}$$

Due to (12) the last two equations are identical.

Let $D_i(x)$ be such that $D_i'(x) = \lambda_i(x)/r_i(x)$. Then, with $h(x) = c_0 r_0(x) g_0(x) = c_1 r_1(x) g_1(x)$ we have that

$$0 = h(x) - \int_x^\infty D_0'(u) h(u) \mathrm{d}u + \int_x^\infty D_1'(u) h(u) \mathrm{d}u, \tag{18}$$

and thus

$$h'(x) = -(D_0'(x) - D_1'(x)) h(x). \tag{19}$$

This implies that

$$h(x) = c e^{-[D_0(x) - D_1(x)]}. \tag{20}$$

Since the $g_i(x)$ are densities, (4) follows. It remains to compute the constants c_0, c_1 and p. Firstly, recall that $c_0 + c_1 + p = 1$. Secondly, (12) implies that $c_1/c_0 = \alpha_1/\alpha_0$. Finally, letting $x \to 0$ in (16) and using (13), one obtains that $p\lambda_1(0) = c_0 r_0(0) g_0(0)$. If $D_0(x) - D_1(x) \to \infty$ as $x \to 0$ or $\lambda_1(0) = \infty$, then necessarily $p = 0$. Otherwise (4) with $x = 0$ yields:

$$p = \frac{c_0}{\lambda_1(0)\alpha_0} \ . \tag{21}$$

The theorem follows. ∎

Remark 1 *Several special cases of this result are contained in the literature. E.g., (i) the case $\lambda_i(x) \equiv \lambda_i$, $\mu_i(x) \equiv \mu_i$ gives a classical two-state fluid queue with exponential buffer content densities $g_0(\cdot)$ and $g_1(\cdot)$. And (ii) if $\lambda_0(x)$ and $r_0(x)$ go to infinity with $\frac{\lambda_0(x)}{r_0(x)} \equiv \mu$, one obtains an $M/M/1$ queue with $\exp(\mu)$ service time distribution, and with state-dependent service rate $r_1(x)$ and state-dependent arrival rate $\lambda_1(x)$. The stationary workload distribution of that model has already been derived in Section 4 of [4].*

Remark 2 *It should be noted that $g_i(x) r_i(x)$ depends on $\lambda_j(x)$ and $r_j(x)$, $j = 0, 1$, only via their ratio. Similar observations for related models have been made in [4], and have been explained there via a rescaling of time.*

It is evident that a necessary condition for the results of the theorem to be valid is that α_0 and α_1 are both finite. In the next section we study the conditions for the existence of a stationary distribution of $\mathcal{M}$ in detail.

4 Conditions for Irreducibility

Since we showed that whenever $\lambda_1(0) > 0$, $\alpha_0 < \infty$ and $\alpha_1 < \infty$ there is a stationary distribution, in order to show that it is unique it suffices to prove that the process $\mathcal{M}$ is irreducible and nonexplosive. This will also imply that the process is positive Harris recurrent. We shall show that the following set of conditions implies that $\mathcal{M}$ is indeed irreducible and nonexplosive.

Condition 1 $\int_0^\infty \frac{1}{r_i(x)} e^{-\int_1^x \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right) \mathrm{d}u} \mathrm{d}x < \infty$ *for* $i = 0, 1$.

Condition 2 $\int_x^y \frac{1}{r_i(u)} \mathrm{d}u < \infty$ *for all* $0 < x < y < \infty$ *and* $i = 0, 1$.

Condition 3 $\int_x^y \frac{\lambda_i(u)}{r_i(u)} \mathrm{d}u < \infty$ *for all* $0 < x < y < \infty$ *and* $i = 0, 1$.

Condition 4 $\int_x^\infty \frac{1}{r_0(u)} \mathrm{d}u = \infty$ *and* $\int_x^\infty \frac{\lambda_0(u)}{r_0(u)} \mathrm{d}u = \infty$ *for some (hence all)* $x > 0$.

Condition 5 *If* $\int_0^y \frac{1}{r_1(u)} \mathrm{d}u = \infty$ *then* $\int_0^y \frac{\lambda_1(u)}{r_1(u)} \mathrm{d}u = \infty$ *for some (hence all)* $y > 0$.

Remark 3 *Note that Conditions 1 and 4 together imply that* $\int_x^\infty \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right) \mathrm{d}u = \infty$ *for any given* x.

Condition 1 states that $\alpha_0 < \infty$ and $\alpha_1 < \infty$. Note that the lower limit of the integral in the exponent is 1 and not 0. This is not a mistake, since we want to allow for the case where $D_0(x) - D_1(x) \to \infty$ as $x \to 0$. When $x < 1$ then $\int_1^x = -\int_x^1$, as usual.

Next assume that given we are in state 0 (1), the time to get from x to y (y to x), where $0 < x < y < \infty$, is finite. The condition for this is Condition 2. Two types of explosion may be possible in such a process. One is where the content reaches ∞ in a finite time and the other where there may be an infinite number of switches between 0 and 1 in a finite time. To prevent the first it suffices to assume that the first part of Condition 4 holds. In order to prevent the second we have assumed that $\lambda_i(\cdot)$ are bounded. To make sure that all states are reachable, it suffices to assume that starting from any state $(x, 0)$ ($(x, 1)$), the time until a switch is made is greater than the time to reach any level $y > x$ ($y < x$) and is a.s. finite. Starting from $(x, 0)$ the time to reach $y > x$ (provided there is no switch) is $\theta(x, y) = \int_x^y \frac{1}{r_0(u)} \mathrm{d}u$. Denoting T to be the switching time and $w_{x,0}(t)$ the unique solution of $w_{x,0}(t) = x + \int_0^t r_0(w_{x,0}(u)) \mathrm{d}u$, we have that

$$P_{x,0}[T > \theta(x, y)] = e^{-\int_0^{\theta(x,y)} \lambda_0(w_{x,0}(s)) \mathrm{d}s} = e^{-\int_x^y \frac{\lambda_0(u)}{r_0(u)} \mathrm{d}u}.$$

Therefore, it is immediate that in order for $P_{x,0}[T > \theta(x, y)]$ to be positive for every finite $y > x$, we need to assume that Condition 3 holds for $i = 0$. Since the first part of Condition 4 implies that $\theta(x,y) \to \infty$ as $y \to \infty$, the second part assures us that T is a.s. finite.

Similar reasoning applies to the switching time from 1 to 0 with the possible exception when state $(0,1)$ is not reachable, in which case this state can be removed from the state space in order to maintain irreducibility. This happens when $\int_0^y \frac{1}{r_1(u)}\mathrm{d}u = \infty$. In the latter case, the condition that the switching time is almost surely finite is the necessary condition in Condition 5.

5 Expected Excursion Times

In this section we compute the expected time $\mathrm{E}R(y)$ that it takes for the buffer content to return to a given level y after an instant where it crossed it from below.

Theorem 2

$$\mathrm{E}R(y) = \alpha_0(y) + \alpha_1(y), \qquad y > 0, \tag{22}$$

where

$$\alpha_i(y) = \int_y^\infty \frac{1}{r_i(x)} e^{-\int_y^x \left(\frac{\lambda_0(u)}{r_0(u)} - \frac{\lambda_1(u)}{r_1(u)}\right)\mathrm{d}u} \mathrm{d}x \ , \qquad i = 0, 1. \tag{23}$$

If the conditions for $p > 0$ are satisfied, then it also holds for $y = 0$.

Proof: For $y = 0$, when $p > 0$, if μ_C is the expected regenerative epoch (cycle time), then it is clear that

$$p = \frac{\frac{1}{\lambda_1(0)}}{\mu_C},$$

where p is given by (6). So $\mu_C = \alpha_0 + \alpha_1 + \frac{1}{\lambda_1(0)}$. Hence, the expected return time to zero must be $\alpha_0 + \alpha_1$, which indeed equals $\alpha_0(0) + \alpha_1(0)$. By reflecting the process at some positive level y instead of zero (that is by setting $r_1(y) = 0$ and starting the process at $(y, 0)$) the result for any $y > 0$ is easily concluded. ■

The following intuitive argument also leads to the same formula for μ_C. The fraction of time the buffer content spends just above level zero in a given cycle can be thought of as the density at zero $c_0 g_0(0+) + c_1 g_1(0+)$. During a single cycle, the buffer content spends the infinitesimal time $\frac{1}{r_0(0+)} + \frac{1}{r_1(0+)}$ just above level zero. Therefore

$$c_0 g_0(0+) + c_1 g_1(0+) = \frac{\frac{1}{r_0(0+)} + \frac{1}{r_1(0+)}}{\mu_C} \ .$$

Applying Theorem 1 we see that this is consistent with $\mu_C = \alpha_0 + \alpha_1 + \frac{1}{\lambda_1(0)}$.

6 The Process Restricted to Down or Up Intervals

The process restricted to down intervals (off state) decreases at a rate $r_1(x)$ when at level $x > 0$ and has jumps up which are state-dependent. Standard regenerative arguments imply that for the case $p > 0$, the stationary distribution function of this restricted content process is given by

$$F(w) = \frac{\frac{1}{\lambda_1(0)} + \alpha_1 \int_0^w g_1(x)\mathrm{d}x}{\frac{1}{\lambda_1(0)} + \alpha_1} \ . \tag{24}$$

When $\lambda_1(x) = \lambda$, $r_1(x) = 1$ for $x > 0$, $\lambda_0(x)/r_0(x) = \mu$ for $x \geq 0$, and $\lambda < \mu$, it is easy to check that $1 - F(w) = \frac{\lambda}{\mu} e^{-(\mu-\lambda)w}$. For this case the restricted process is the workload process of an $M/M/1$ queue for which this formula is well known.

As for the process restricted to the up intervals (on state), the stationary distribution of the content level has the density $g_0(\cdot)$.

7 The Stationary Distribution of Peaks and Valleys

In this section we want to compute the distribution of local minima and maxima of $\mathcal{M}$. Each forms a discrete time Markov process. We call the local maxima *peaks* and the local minima *valleys*. Beginning with valleys and assuming that $p > 0$, Theorem 5.1 and the resulting equation (20) of [4] cover the model considered here, when we restrict $\mathcal{M}$ to down intervals. Therefore, if we denote by V_D a random variable that has the stationary distribution of the process restricted to down intervals and by W_D the stationary distribution of the discrete time process of valleys (which in the restricted process are the states right before jumps up) then for a given bounded Borel measurable function f,

$$Ef(W_D) = \frac{E\lambda_1(V_D)f(V_D)}{E\lambda_1(V_D)} \ . \tag{25}$$

Since, by (24), for a given bounded Borel measurable function h we have that

$$Eh(V_D) = \frac{\frac{1}{\lambda_1(0)}h(0) + \alpha_1 \int_0^\infty h(u)g_1(u)\mathrm{d}u}{\frac{1}{\lambda_1(0)} + \alpha_1} \ , \tag{26}$$

it follows that for a given bounded f,

$$Ef(W_D) = \frac{f(0) + \alpha_1 \int_0^\infty \lambda_1(u)g_1(u)f(u)\mathrm{d}u}{1 + \alpha_1 \int_0^\infty \lambda_1(u)g_1(u)\mathrm{d}u} \ . \tag{27}$$

Taking $f(u) = 1_{[0,x]}(u)$, this implies that the stationary distribution function of the valleys is given by

$$F_{W_D}(x) = \frac{1 + \alpha_1 \int_0^x \lambda_1(u)g_1(u)\mathrm{d}u}{1 + \alpha_1 \int_0^\infty \lambda_1(u)g_1(u)\mathrm{d}u} \ . \tag{28}$$

When $p = 0$ then W_D has the density

$$f_{W_D}(x) = \frac{\lambda_1(x)g_1(x)}{\int_0^\infty \lambda_1(u)g_1(u)\mathrm{d}u} \ . \tag{29}$$

Similarly, it can be shown by identical methods that if W_U denotes a random variable having the stationary distribution of the discrete time peak process, then it has the following density:

$$f_{W_U}(x) = \frac{\lambda_0(x)g_0(x)}{\int_0^\infty \lambda_0(u)g_0(u)\mathrm{d}u} \ . \tag{30}$$

Remark 4 *Equations (28)-(30) show that the densities of valleys and peaks are proportional to $\lambda_0(x)g_0(x)$ and $\lambda_1(x)g_1(x)$, respectively. A similar result, which may be viewed as a PASTA generalization, was derived in [4] for an $M/G/1$ queue with state-dependent arrival rate and service speed. It should also be noted that, cf. Remark 2, the densities of valleys and peaks depend on $\lambda_j(x)$ and $r_j(x)$, $j = 0, 1$, only via their ratio.*

8 Exponential Ergodicity

In this section we impose the following restrictions on the process parameters, that will ensure that the convergence to the stationary distribution π is at some exponential rate, i.e., that $\mathcal{M}$ is *f-exponentially ergodic* for some function $f > 1$:
Condition A.1. For some $\epsilon > 0$,

$$\sup_{x\geq 1} \int_x^\infty \frac{\lambda_0(u)}{r_0(u)} e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}\mathrm{d}v} [\lambda_1(x) \int_x^u (\frac{1}{r_0(v)} + \frac{1}{r_1(v)})\mathrm{d}v]\mathrm{d}u \leq 1 - \epsilon.$$

Condition A.2. There exists a $c^* > 0$ so that

$$\sup_{x>1} \lambda_1(x) \int_x^\infty \frac{\lambda_0(u)}{r_0(u)} e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}\mathrm{d}v} e^{c^* \int_x^u (\frac{1}{r_0(v)} + \frac{1}{r_1(v)})\mathrm{d}v} \mathrm{d}u = B < \infty.$$

We shall show that under those conditions the process $\mathcal{M}$ is f-exponentially ergodic as defined in [13], with $f = V + 1$, V to be defined below:

$$V(x,1) = \begin{cases} 1 & \text{if } x < 1, \\ e^{c\int_1^x \frac{1}{r_1(u)}\mathrm{d}u} & \text{if } x \geq 1, \end{cases}$$

$$V(x,0) = \int_x^\infty \frac{\lambda_0(u)}{r_0(u)} e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}dv} e^{\int_x^u \frac{c}{r_0(v)})\mathrm{d}v} V(u,1)\mathrm{d}u.$$

We do this by verifying that $\mathcal{A}V(x,i) \leq -k_1 V(x,i) + k_2$ for some $k_1 > 0$ and $k_2 < \infty$, thus verifying Condition (CD3) of [13]. The verification for $\mathcal{A}V(x,0)$ is easy:

$$\begin{aligned} \mathcal{A}V(x,0) &= r_0(x)(-\tfrac{\lambda_0(x)}{r_0(x)}V(x,1) + \tfrac{\lambda_0(x)}{r_0(x)}V(x,0) - \tfrac{c}{r_0(x)}V(x,0)) \\ &\quad -\lambda_0(x)V(x,0) + \lambda_0(x)V(x,1) \\ &= -cV(x,0). \end{aligned} \tag{31}$$

Now consider $\mathcal{A}V(x,1)$. With $R_1(x)=\int_1^x \frac{1}{r_1(u)}\mathrm{d}u$ for $x\geq 1$, we have

$$\mathcal{A}V(x,1)=-ce^{cR_1(x)}1_{\{x>1\}}-\lambda_1(x)V(x,1)+\lambda_1(x)V(x,0). \tag{32}$$

It is easy to see that for $x<1$, $\mathcal{A}V(x,1)$ is bounded as a product of a continuous function and a bounded function $\lambda_1(x)$. For $x\geq 1$,

$$\begin{aligned}\mathcal{A}V(x,1) &= \lambda_1(x)ce^{cR_1(x)}\int_x^\infty \frac{\lambda_0(u)}{r_0(u)}e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}\mathrm{d}v}\\ &\quad [c^{-1}(e^{c\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v}-1)-\frac{1}{\lambda_1(x)}]\mathrm{d}u.\end{aligned} \tag{33}$$

Using the Taylor expansion

$$\begin{aligned}e^{c\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v} &\leq 1+c\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v\\ &\quad +\frac{c^2}{2}(\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v)^2e^{c\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v},\end{aligned} \tag{34}$$

it follows that

$$\begin{aligned}\mathcal{A}V(x,1) &\leq ce^{cR_1(x)}\{\int_x^\infty \frac{\lambda_0(u)}{r_0(u)}e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}\mathrm{d}v}[\lambda_1(x)\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v-1]\\ &\quad +\lambda_1(x)e^{cR_1(x)}\frac{c}{2}(\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v)^2e^{c\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v}\}\mathrm{d}u.\end{aligned} \tag{35}$$

But by Condition A.1 above and Condition 4, for $x\geq x_0$,

$$\int_x^\infty \frac{\lambda_0(u)}{r_0(u)}e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}\mathrm{d}v}[\lambda_1(x)\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v-1]\mathrm{d}u<-\epsilon. \tag{36}$$

As for the second term in the expression for $\mathcal{A}V(x,1)$, we use the inequality $x^2\leq 2\Delta^{-2}e^{\Delta x}$ with $\Delta=\frac{c^*}{2}$ and $c<\Delta$ to show that

$$(\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v)^2e^{c\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v}<2\Delta^{-2}e^{c^*\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v}, \tag{37}$$

and according to Condition A.2 above,

$$\begin{aligned}&ce^{cR_1(x)}\{\int_x^\infty \frac{\lambda_0(u)}{r_0(u)}e^{-\int_x^u \frac{\lambda_0(v)}{r_0(v)}\mathrm{d}v}\lambda_1(x)\frac{c}{2}\int_x^u(\frac{1}{r_0(v)}+\frac{1}{r_1(v)})\mathrm{d}v)^2e^{c\int_x^u(\frac{1}{r_9(v)}+\frac{1}{r_1(v)})\mathrm{d}v}\mathrm{d}u\}\\ &<c^2e^{cR_1(x)}\Delta^{-2}B.\end{aligned} \tag{38}$$

It thus follows that with c as above and $x>1$,

$$\mathcal{A}V(x,1)\leq e^{cR_1(x)}[-c\epsilon+c^2\Delta^{-2}B]. \tag{39}$$

We now choose $c<\frac{\epsilon\Delta^2}{2B}$ and we get that for $x\geq x_0$,

$$\mathcal{A}V(x,1)\leq \frac{-c\epsilon}{2}V(x,1), \tag{40}$$

and it is bounded for $x<x_0$. It follows that the functions $V(x,i), i=0,1$, satisfy condition (CD3) of [13] and thus Theorem 6.1 of [13] implies that $\mathcal{M}$ is exponentially ergodic.

References

[1] Adan, I.J.B.F., Van Doorn, E.A., Resing, J.A.C., Scheinhardt, W.R.W. (1998). Analysis of a single-server queue interacting with a fluid reservoir. Queueing Systems 29, 313-336.

[2] Azéma, J., Kaplan-Duflo, M., Revuz, D. (1967). Mesure invariante sur les classes récurrentes des processus de Markov. Z. Wahrscheinlichkeitstheorie verw. Gebiete 8, 157-181.

[3] Anick, D., Mitra, D., Sondhi, M.M. (1982). Stochastic theory of a data-handling system with multiple sources. Bell Syst. Tech. J. 61, 1871-1894.

[4] Bekker, R., Borst, S.C., Boxma, O.J., Kella, O. (2003). Queues with workload-dependent arrival and service rates. SPOR-Report 2003-11, Eindhoven University of Technology; accepted for publication in *Queueing Systems*.

[5] Davis, M.H.A. (1984). Piecewise-deterministic Markov processes: A general class of non diffusion stochastic models. J. Roy. Statist. Soc. Series B 46, 353-388.

[6] Gaver, D.P., Miller, R.G. (1962). Limiting distributions for some storage problems. In: Arrow, K.J., Karlin, S., Scarf, H. (eds.). *Studies in Applied Probability and Management Science*. Stanford University Press, Stanford, California, 110-126.

[7] Harrison, J.M., Resnick, S.I. (1976). The stationary distribution and first exit probabilities of a storage process with general release rule. Math. Oper. Res. 1, 347-358.

[8] Kaspi, H., Kella, O., Perry, D. (1996). Dam processes with state dependent batch sizes and intermittent production processes with state dependent rates. Queueing Systems 24, 37-57.

[9] Kella, O., Whitt, W. (1992). A storage model with a two-state random environment. Oper. Res. 40, S257-S262.

[10] Mandjes, M., Mitra, D., Scheinhardt, W.R.W. (2002). A simple model of network access: feedback adaptation of rates and admission control. In: Proc. INFOCOM 2002, 3-12.

[11] Mandjes, M., Mitra, D., Scheinhardt, W.R.W. (2003). Models of network access using feedback fluid queues. Queueing Systems 44, 365-398.

[12] Meyn, S.P. and Tweedie, R.L.(1993). Markov Chains and Stochastic Stability. Springer Verlag.

[13] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes III: Foster Lyapunov criteria for continuous time processes. Adv. Appl. Prob. 25, 518-548.

[14] Scheinhardt, W.R.W., Van Foreest, N., Mandjes, M. (2003). Continuous feedback fluid queues. Report Department of Applied Mathematics, University of Twente.