Two-buffer fluid models with multiple ON-OFF inputs and threshold assistance

Guy Latouche
Université Libre de Bruxelles
Département d’Informatique
1050 Brussels, Belgium
latouche@ulb.ac.be

Giang T. Nguyen
Université Libre de Bruxelles
Département d’Informatique
1050 Brussels, Belgium
giang.nguyen@ulb.ac.be

Zbigniew Palmowski
University of Wrocław
Mathematical Institute
50-384 Wrocław, Poland
zbigniew.palmowski@gmail.com

ABSTRACT
We consider a two-buffer fluid model with $N$ ON-OFF inputs and threshold assistance, which is an extension of the same model with $N = 1$ in [18]. While the rates of change of both buffers are piecewise constant and dependent on the underlying Markovian phase of the model, the rates of change for Buffer 2 are also dependent on the specific level of Buffer 1. This is because both buffers share a fixed output capacity, the precise proportion of which depends on Buffer 1. The generalization of the number of ON-OFF inputs necessitates slight modifications in the original rules of output-capacity sharing from [18], and considerably complicates both the theoretical analysis and numerical computation of various performance measures.

Here, we give a short explanation on how to derive the marginal probability distribution of Buffer 1, and bounds for that of Buffer 2. In an upcoming paper, we describe the procedures in more details. Furthermore, restricting Buffer 1 to a finite size, we determine its marginal probability distribution in the specific case of $N = 1$, thus providing numerical comparisons to the corresponding results in [18] where Buffer 1 is assumed to be infinite. We also demonstrate how this imposed restriction effects the bounds of marginal probabilities for Buffer 2.

Categories and Subject Descriptors
I.6.5 [Simulation and Modelling]: Model development; G.3 [Probability and Statistics]: [queueing theory, stochastic processes]

General Terms
Performance, Measurement

1. INTRODUCTION
Stochastic fluid models have a wide range of real-life applications, such as industrial and computer engineering, actuarial science, environmental modeling and telecommunication. A Markov-modulated single-buffer fluid model is a two-dimensional Markov process $\{X(t), \varphi(t) : t \in \mathbb{R}^+\}$, where $X(t)$ is the continuous level of the buffer, and $\varphi(t)$ is the discrete phase of the underlying irreducible Markov chain that governs the rates of change. A practical and well-studied case is piecewise constant rates; the fluid is assumed to have a constant rate $c_i$, when $\varphi(t) = i$, for $i$ in a finite state space $\mathcal{S}$. The traditional approach for obtaining performance measures of Markov-modulated single-buffer fluids with piecewise constant rates is to use spectral analysis (see, among others, [17, 19, 25, 18]). Over the last two decades, matrix analytic methods have gained a lot of attention as an alternative and algorithmically effective approach for analyzing these standard fluids (see, for instance, [22, 24, 1, 2, 6, 10, 5, 7, 11, 8]).

In this paper, we consider a two-buffer fluid model $(X(t), Y(t), \varphi_1(t), \varphi_2(t) : t \in \mathbb{R}^+)$, where $X(t) \geq 0$ and $Y(t) \geq 0$ represent the levels of Buffers 1 and 2, respectively. At a given time $t \geq 0$, the rates of change of Buffer 1 depend only on the underlying Markovian phase $\varphi_1(t)$; however, the rates of change of Buffer 2 depend on both $\varphi_2(t)$ and $X(t)$. This is because while each buffer receives its own input sources, both buffers share a fixed output capacity $c$, in proportion dependent on the level of Buffer 1. More specifically, Buffer $j$ receives $N$ ON-OFF input sources, each has exponentially distributed ON- and OFF-intervals at corresponding rates $\alpha_j$ and $\beta_j$, and continuously generates fluid at rate $R_j$ during ON-intervals, for $j = 1, 2$. When the fluid level $X(t)$ of Buffer 1 is above a certain threshold $X^*$, Buffer 1 is allocated the total shared output capacity $c$, leaving Buffer 2 without any; when $0 < X(t) < x^*$, Buffer $j$ has output capacity $c_j$, $c_1 + c_2 = c$ and when $X(t) = 0$, Buffer 1 has output capacity $\min\{c, R_1, c_1\}$, and Buffer 2 $c - \min\{1, R_1, c_1\}$, where $i$ is the number of inputs of Buffer 1 being on at the time $t$.

This theoretical model is an extension of the same model with $N = 1$ in [18], where the reader can find a comprehensive account of practical applications in communication networks. We note that when $X(t) = 0$, the rule for output-capacity allocation in our general $N$ ON-OFF input model differs to that in the single ON-OFF input model in [18], which is to allocate the total capacity $c$ to Buffer 2. The totality rule is logical for the single ON-OFF input: when there is only one ON-OFF input for each buffer, Buffer 1 is empty only when its input is off; in that case, Buffer 2 can receive the whole output capacity $c$, until the moment the input of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SMCTOOLS 2011, May 16, Paris, France
Copyright © 2011 ICST 978-1-936968-09-1
DOI 10.4108/icst.valuetools.2011.245848
Buffer 1 is on again. In our model where each buffer receives \( N \) ON-OFF inputs for \( N \geq 1 \), it is possible for Buffer 1 to be empty while \( i \) inputs are on, for \( 0 \leq i \leq \left[ \frac{R_1}{c} \right] \). Under these circumstances, assigning the total output capacity \( c \) to Buffer 2 would immediately cause Buffer 1 to try to increase from level 0, consequently grabbing back \( c_1 \) amount of output capacity. However, as \( i \leq \left[ \frac{R_1}{c} \right] \), the output capacity \( c_1 \) would be sufficient to empty Buffer 1, forcing it to give away the whole output capacity \( c \) to Buffer 2, etc. Therefore, applying the original totality rule at \( X(t) = 0 \) for the generalized \( N \) ON-OFF input model would potentially lead to inconsistency.

The behavior described above at level 0 for Buffer 1 when \( 0 \leq i \leq \left[ \frac{R_1}{c} \right] \) is referred to as being sticky [11], a property arisen when net rates of the buffer for the same Markovian phase but different levels are different in a particular way that makes it unable to go up or down, thus remaining stuck at a level until the background Markov chain switches to a non-sticky phase. In our model, by allocating \( iR_1 \) output capacity to Buffer 1 and \( c - iR_1 \) to Buffer 2 when \( X(t) = 0 \) and \( 0 \leq i \leq \left[ \frac{R_1}{c} \right] \), we let Buffer 1 remain at level zero, while eliminating potential uncertainty and utilizing the total output capacity in the most effective way. For the same reason, when \( X(t) = x^+ \) and \( \left[ \frac{R_1}{c} \right] \leq i \leq \left[ \frac{R_2}{c} \right] \), the output capacity is \( iR_1 \) for Buffer 1, and \( c - iR_1 \) for Buffer 2. While the stickiness, borne in the generalization of the number of IN-OUT inputs, necessitates only slight modifications in the output-capacity allocation policy, it considerably complicates the analysis and numerical computation of performance measures of the model. To deal with this complication, we employ a mixture of tools from both dominant approaches: spectral analysis and matrix analytic methods.

The rest of the paper is organized as follows: in Section 2, we formulate the model mathematically. Assuming that both buffer sizes are infinite, we derive the marginal probability distribution of Buffer 1 in Section 3.1, and bounds for those of Buffer 2 in Section 3.2.

2. REFERENCE MODEL

Consider a four-dimensional Markov process \( \{X(t), Y(t), \varphi_1(t), \varphi_2(t) : t \in \mathbb{R}^+ \} \), where \( X(t) \geq 0 \) and \( Y(t) \geq 0 \) are the levels in buffers 1 and 2, respectively, and for \( j = 1, 2 \) and \( N \geq 1 \), \( \varphi_j(t) \) represents the phase of the background irreducible Markov chain for Buffer \( j \) with finite state space \( \mathcal{S} = \{0, \ldots, N\} \); state \( i \in \mathcal{S} \) indicates that \( i \) ON-OFF inputs are on. The generator \( T_j \) for \( \{\varphi_j(t)\} \) is

\[
T_j = \begin{bmatrix}
\alpha_j & N\beta_j & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\gamma_j & \cdots & \cdots & \cdots & \cdots \\
\end{bmatrix},
\]

with each diagonal element \( \ast \) defined appropriately such that each row sum of \( T_j \) is 0. For \( i_1, i_2 \in \mathcal{S} \), we denote by \( \dot{x}_{i_1} \) and \( \dot{y}_{i_2} \) the respective net rates for Buffer 1 in phase \( i_1 \) and Buffer 2 in phase \( i_2 \). For \( X(t) > x^+ \) and \( Y(t) > 0 \)

\[
\dot{x}_{i_1} = i_1R_1 - c,
\dot{y}_{i_2} = i_2R_2,
\]

for \( X(t) = x^+ \) and \( Y(t) > 0 \):

\[
\dot{x}_{i_1} = \begin{cases} 
0 & \text{for } \left[ \frac{c_1}{R_1} \right] \leq i_1 \leq \left[ \frac{c}{R_1} \right], \\
= i_1R_1 - c & \text{otherwise}, \\
\end{cases}
\]

\[
\dot{y}_{i_2} = i_2R_2 - (c - i_1R_1) \quad \text{for } \left[ \frac{c_1}{R_1} \right] \leq i_1 \leq \left[ \frac{c}{R_1} \right],
\]

\[
\dot{y}_{i_2} = i_2R_2 - c_2 \quad \text{otherwise}.
\]

For \( 0 < X(t) < x^+ \) and \( Y(t) > 0 \):

\[
\dot{x}_{i_1} = i_1R_1 - c,
\dot{y}_{i_2} = i_2R_2 - c_2,
\]

and for \( X(t) = 0 \) and \( Y(t) > 0 \):

\[
\dot{x}_{i_1} = 0 \quad \text{for } 0 \leq i_1 \leq \left[ \frac{c_1}{R_1} \right],
\]

\[
\dot{y}_{i_2} = i_2R_2 - (c - i_1R_1) \quad \text{for } 0 \leq i_1 \leq \left[ \frac{c_1}{R_1} \right],
\]

\[
\dot{y}_{i_2} = i_2R_2 - c_2 \quad \text{otherwise}.
\]

For \( Y(t) = 0 \), \( \dot{y}_{i_2} \) is the maximum between 0 and the net rate of Buffer 2 in \( i_2 \in \mathcal{S} \) when \( Y(t) > 0 \).

We assume that \( NR_j > c \), that \( \frac{c_1}{R_1}, \frac{c_2}{R_2} \notin \mathbb{N} \), and that the system is positive recurrent.

3. INFINITE BUFFERS WITH MULTIPLE ON-OFF INPUTS

3.1 Analysis for Buffer 1

To analyze Buffer 1 when \( N = 1 \), Mahabhashyam et al. [18] consider an equivalent system of two standard single sub-buffers, each with a single ON-OFF input, one sub-buffer with constant output capacity \( c_1 \) and the other with constant output capacity \( c \). Decomposing Buffer 1 in this fashion, the authors show that the marginal probability distribution of Buffer 1 can be obtained by appropriately combining the average time of going up from \( x^+ \) and then going down to \( x^+ \) in Sub-buffer 1, and the average time of going down from \( x^+ \) and then going up to \( x^+ \) in Sub-buffer 2. The authors determine analytic expressions for the former average time by using, from [20], the busy period distribution of a standard single buffer with one exponential ON-OFF input and constant output capacity, and for the latter by establishing a pair of partial differential equations, transferred into ordinary differential equations and then solved by a spectral decomposition technique.

In this paper, for general \( N \geq 1 \), we analyze Buffer 1 by modeling it as a level-dependent fluid and applying matrix analytic methods. With this approach, while it is not simple to obtain closed-form expressions for \( N \geq 2 \), we can obtain various performance measures numerically using fast convergent algorithms (see, most relevantly, [5, 11] and the references therein). The focus of this section is the marginal probability distribution for Buffer 1.

We refer to \( X(t) = 0 \) and \( X(t) = x^+ \) as boundaries \( \circ \) and \( \ast \), and \( 0 < X(t) < x^+ \) and \( X(t) > x^+ \) as bands 1 and 2. While \( T_1 \) governs the transitions of \( \{\varphi_1(t)\} \) for all \( X(t) \geq 0 \), the
rate of Buffer 1 in the same phase varies between boundaries and bands. Therefore, we partition $S$ differently for each boundary and each band. We denote, respectively, by $S^o(s)$, $S^o(k)$ and $S^o(s)$ the sets of states with negative, zero and positive net rates when Buffer 1 is at boundary $s$, for $s \in \{0, \ast\}$, and by $S^o(k)$ and $S^o(s)$ the sets of states with negative and positive net rates when Buffer 1 is in band $k$, for $k = 1, 2$.

Then, $S = S^{(o)} \cup S^{(1)} = S^{(1)} \cup S^{(2)} = S^{(2)} \cup S^{(2)}$, with

$$S^{(o)} = S^{(1)} = S^{(2)} = \{0, \ldots, \left\lfloor \frac{c_1}{R_1} \right\rfloor \},$$

$$S^{(u)} = S^{(1)} = S^{(2)} = \{\left\lfloor \frac{c_1}{R_1} \right\rfloor, \ldots, N\},$$

$$S^{(s)} = \{\left\lfloor \frac{c}{R_1} \right\rfloor, \ldots, \left\lfloor \frac{c}{R_1} \right\rfloor \}, S^{(a)} = \{\left\lfloor \frac{c}{R_1} \right\rfloor, \ldots, N\},$$

For each band $k$, we partition $T_1$ into sub-matrices $T_{jm}^{(k)}$, of which each element $T_{jm}^{(k)}$, recording transitions going from $i \in S^{(k)}_i$ to $j \in S^{(m)}_j$, and denote by $C^{(k)}_i$ a diagonal absolute net rate matrix for $i \in S^{(k)}_i$:

$$C_0^{(1)} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

$$C_1^{(1)} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

$$C_2^{(1)} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

We illustrate the relationships between the large cast of characters in Figure 1.

Exploiting Markov-renewal arguments, da Silva Soares and Latouche [11, Theorem 4.2] prove that the stationary density vector of a Markov-modulated level-dependent single-buffer fluid queue can be obtained by properly combining limiting densities from above and below each boundary (when possible) and steady state probability masses at these boundaries. To obtain the necessary limiting densities and probability masses, we consider the jump chain $\{J_n : n \geq 0\}$ of the process $\{X(t), \varphi(t)\}$ restricted to the set of boundary states $B = \{\bullet, i : 0 \in \{0, \ast\}, \ast \in S\}$. We note that this jump chain will also be useful for obtaining bounds on marginal probabilities of Buffer 2, as described in Section 3.2. By [11, Theorem 4.4] and block-partitioned according to $B = (0, S^{(o)}) \cup (0, S^{(1)}) \cup (0, S^{(2)}) \cup (0, S^{(2)}) \cup (0, S^{(2)}) \cup (0, S^{(2)})$, the $(2N+1) \times (2N+1)$ transition matrix $\Omega$ of $\{J_n\}$ is

$$\Omega = \begin{bmatrix} \Phi_1 & P_{uu}^{(1)} & \cdots & P_{ud}^{(1)} \\ P_{su}^{(1)} & \Phi_2 & \cdots & P_{sd}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{u_1 u_2}^{(1)} & P_{u_1 u_3}^{(1)} & \cdots & \Phi_{uu}^{(1)} \end{bmatrix},$$

where $\Phi^{(1)}_{uu}, \Phi^{(2)}_{uu}, \Lambda^{(0)}, \Lambda^{(1)}$ and $\Lambda^{(2)}$ denote various first passage probability matrices, with

$$[\Phi_{uu}]_{ij} = \text{probability of returning to } \bullet \text{ and in } j \in S^{(o)},$$

$$[\Lambda_{uu}]_{ij} = \text{probability of reaching level } 0 \text{ and in } i \in S^{(o)},$$

$$[\Lambda_{uu}]_{ij} = \text{probability of reaching level } 0 \text{ and in } j \in S^{(o)},$$

$$[\Lambda_{uu}]_{ij} = \text{probability of reaching level } 0 \text{ and in } i \in S^{(o)},$$

$$[\Lambda_{uu}]_{ij} = \text{probability of reaching level } 0 \text{ and in } j \in S^{(o)},$$

$$[\Lambda_{uu}]_{ij} = \text{probability of reaching level } 0 \text{ and in } i \in S^{(o)},$$

For the remainder of the paper, we denote by $I$ the identity matrix and by $1$ the vector of ones of the appropriate size. Let $\Delta$ be a diagonal matrix with $\Delta_i = \{T_1\}_i$ and define

$$P = I - \Delta^{-1}T_1.$$

Then, clearly each of $P_{uu}^{(1)}, P_{uu}^{(2)}, P_{uu}^{(3)}, P_{uu}^{(4)}$ and $P_{uu}^{(5)}$ is a sub-matrix of $P$:

$$P = \begin{bmatrix} P_{uu}^{(1)} & P_{uu}^{(2)} \\ P_{uu}^{(3)} & P_{uu}^{(4)} \\ P_{uu}^{(5)} & P_{uu}^{(6)} \end{bmatrix} = \begin{bmatrix} P_{uu}^{(1)} & P_{uu}^{(2)} \\ P_{uu}^{(3)} & P_{uu}^{(4)} \\ P_{uu}^{(5)} & P_{uu}^{(6)} \end{bmatrix}.$$

Respectively, $\Phi^{(1)}_{uu}, [\Lambda^{(0)}, \Lambda^{(1)}, \Lambda^{(2)}]$ and $[\Phi^{(1)}_{uu}, \Lambda^{(0)}, \Lambda^{(1)}, \Lambda^{(2)}]$ equal to $\Psi^{(1)}_{1,1}, \Lambda^{(1)}_{1,1}, \Lambda^{(2)}_{1,1}$ and $\Phi^{(1)}_{uu}$, the corresponding first passage probability matrices for the level-independent fluid queue $\{M[t], \mu(t) : t \in \mathbb{R}^+\}$ with finite size $x^\ast$, state space $S^{(1)} \cup S^{(2)}$, generator $T_1$ and rate matrices $C^{(1)}$ and $C^{(2)}$. By [10, Theorem 5.2],

$$\begin{bmatrix} \Lambda^{(1)}_{1,1} & \Phi^{(1)}_{uu} \\ \Phi^{(1)}_{uu} & \Lambda^{(1)}_{1,1} \end{bmatrix} = \begin{bmatrix} \Psi_{1}^{(1)} & \Phi^{(1)}_{uu} \\ \Phi^{(1)}_{uu} & \Lambda^{(1)}_{1,1} \end{bmatrix} = \begin{bmatrix} \Psi_{1}^{(1)} & \Phi^{(1)}_{uu} \\ \Phi^{(1)}_{uu} & \Lambda^{(1)}_{1,1} \end{bmatrix},$$

Figure 1: Buffer 1

458
where $\Psi_1$ is the minimum nonnegative solution to the Riccati equation
\[
(C^{(1)}_+)^{-1}T^{(1)}_+ + (C^{(1)}_+)^{-1}T^{(2)}_+ \Psi_1 \\
+ \Psi_1(C^{(1)}_+)^{-1}T^{(1)}_+ + \Psi_1(C^{(1)}_+)^{-1}T^{(2)}_+ \Psi_1 = 0,
\]
and $\hat{\Psi}_1$ is the minimum nonnegative solution to the Riccati equation
\[
(C^{(1)}_+)^{-1}T^{(1)}_+ + (C^{(1)}_+)^{-1}T^{(2)}_+ \hat{\Psi}_1 \\
+ \hat{\Psi}_1(C^{(1)}_+)^{-1}T^{(1)}_+ + \hat{\Psi}_1(C^{(1)}_+)^{-1}T^{(2)}_+ \hat{\Psi}_1 = 0,
\]
and
\[
U_1 = (C^{(1)}_+)^{-1}T^{(1)}_+ + (C^{(1)}_+)^{-1}T^{(2)}_+ \Psi_1,
\]
and
\[
U_1 = (C^{(1)}_+)^{-1}T^{(1)}_+ + (C^{(1)}_+)^{-1}T^{(2)}_+ \hat{\Psi}_1.
\]
Similarly, $[\Psi^{(1)}_d, \Psi^{(2)}_d] = \Psi_2$, which is the first passage probability matrix for the infinite level-dependent fluid queue $\{M_t(t), \rho(t) : t \in \mathbb{R}^+\}$ with state space $S^{(2)} \cup S^{(1)}$, the generator $T_1$ and rate matrices $C^{(2)}_+$ and $C^{(2)}_+$. By [22], the matrix $\Psi_2$ is the minimum nonnegative solution to the Riccati equation
\[
(C^{(2)}_+)^{-1}T^{(2)}_+ + (C^{(2)}_+)^{-1}T^{(2)}_+ \Psi_2 \\
+ \Psi_2(C^{(2)}_+)^{-1}T^{(2)}_+ + \Psi_2(C^{(2)}_+)^{-1}T^{(2)}_+ \Psi_2 = 0.
\]
Applying fast convergent algorithms described in [9, 4], we can solve Riccati equations (4), (5), and (6) to obtain $\Psi_1, \Psi_2$ and $\Psi_2$, and consequently $\Omega$.

We denote by $m = [p^{(o)}_+, p^{(s)}_+]$ the probability mass vector of Buffer 1 at the set of boundary sticky states $K = \{(0, \zeta) : \zeta \in S^{(1)}(\star), \zeta \in S^{(s)}\}$, and define $E^{(o)} = \{(\star, \zeta) : \zeta \in S^{(1)}(\star) \cup S^{(s)}\}$ and $E^{(s)} = \{(0, \zeta) : \zeta \in S^{(s)}\}$. Note that $K = B - (E^{(s)} \cup E^{(o)})$. Then, the transition matrix $\Omega^{(s)}$ of the censored fluid queue on $B = E^{(o)}$ is
\[
\Omega^{(s)} = \begin{bmatrix}
\Psi^{(s)}_{ua} & 0 & 0 & \cdots \\
0 & \Psi^{(s)}_{sa} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & \Psi^{(s)}_{sa}
\end{bmatrix}
\]

Consequently, the transition matrix of the censored fluid queue on $K$ is
\[
\Omega^{(s)} = \Omega^{(o)}_K + \Omega^{(s)}_{\Omega^{(o)}_K} \left(I - \Omega^{(o)}_{E^{(s)}(\star)}(\star)\right)^{-1} \Omega^{(s)}_{E^{(o)}(\star)}K
\]

and its generator matrix is
\[
\Theta = \Delta^{(s)}(I - \Omega^{(o)}_K)
\]

where $\Delta^{(s)}$ is the diagonal matrix with $[\Delta^{(s)}]_{ii} = [T_1]_i$ for $0 \leq i \leq [\tau]_1$. By [11, Theorem 4.5], $\Theta$ is the generator of the censored fluid queue on $S^{(o)} \cup S^{(s)}$, and
\[
m = k[x^{(o)}_1, x^{(s)}_1].
\]
function of Buffer 1 is

\[ P(X_1 \leq x) = \begin{cases} \pi(x) dx & \text{for } 0 < x < x^*, \\ \left[ p_1^{(i)} + p_2^{(i)} \right] 1 + \int_0^x \pi(x) dx & \text{for } x \geq x^*. \end{cases} \]

\[ = \begin{bmatrix} \frac{\lambda^{(1)}(s)}{\Psi^{(1)}(s)} \\ \frac{\lambda^{(2)}(s)}{\Psi^{(2)}(s)} \\ \frac{\lambda^{(3)}(s)}{\Psi^{(3)}(s)} \\ \frac{\lambda^{(4)}(s)}{\Psi^{(4)}(s)} \end{bmatrix} \]

\[ = \begin{bmatrix} e^{-\psi_1(s)x} & \psi_1(s) & I \\ \psi_1(s) e^{-\psi_1(s)x} & e^{-\psi_1(s)x} & I \\ I & I & I \end{bmatrix}^{-1}, \]

where \( \psi_1(s) \) is the minimum nonnegative solution to the Riccati equation

\[ (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI) + (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI)\psi_1(s) \]

\[ + \psi_1(s)(C^{(1)}_+)^{-1}(T^{(1)}_+ - sI) \psi_1(s) = 0, \]

\[ \psi_1(s) \text{ is the minimum nonnegative solution to the Riccati equation} \]

\[ (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI) + (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI)\psi_1(s) \]

\[ + \psi_1(s)(C^{(1)}_+)^{-1}(T^{(1)}_+ - sI) \psi_1(s) = 0, \]

\[ \bar{U}_1(s) = (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI) + (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI)\psi_1(s), \]

\[ \text{and} \]

\[ \bar{U}_1(s) = (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI) + (C^{(1)}_+)^{-1}(T^{(1)}_+ - sI)\psi_1(s). \]

Similarly, \( \tilde{\Psi}_{2a}(s), \tilde{\Psi}_{2d}(s) = \tilde{\Psi}_2(s), \) which is the matrix of the LST of first passage times for \( \{M_2(t), \rho_2(t)\} \). By [6, Theorem 1], the matrix \( \tilde{\Psi}_2(s) \) is the minimum nonnegative solution to the Riccati equation

\[ (C^{(2)}_+)^{-1}(T^{(2)}_+ - sI) + (C^{(2)}_+)^{-1}(T^{(2)}_+ - sI)\psi_2(s) \]

\[ + \psi_2(s)(C^{(2)}_+)^{-1}(T^{(2)}_+ - sI) \psi_2(s) = 0, \]

\[ \bar{U}_2(s) = (C^{(2)}_+)^{-1}(T^{(2)}_+ - sI) + (C^{(2)}_+)^{-1}(T^{(2)}_+ - sI)\psi_2(s), \]

\[ \text{and} \]

\[ \bar{U}_2(s) = (C^{(2)}_+)^{-1}(T^{(2)}_+ - sI) + (C^{(2)}_+)^{-1}(T^{(2)}_+ - sI)\psi_2(s). \]

\[ \text{Bean et al. [7] give efficient algorithms for solving (22), (23) and (24) to obtain } \tilde{\Psi}_1(s), \tilde{\Psi}_1(s) \text{ and } \tilde{\Psi}_2(s), \]}

\[ \text{and consequently } \Omega(s) \]

\[ \text{Before we state the bounds for Buffer 2, we need to define effective bandwidths and failure rate functions. For } v > 0, \text{ the effective bandwidth } \epsilon_b(v) \text{ of an input that generates } F(t) \text{ amount of fluid in time } t \text{ (see, for example, [12, 16]) is defined to be} \]

\[ \epsilon_b(v) = \lim_{t \to \infty} \frac{1}{t} \log E[e^{vF(t)}]. \]

\[ \text{By [3, 12], the effective bandwidth } \epsilon_b(v) \text{ of a single exponential ON-OFF source for fixed } v \]

\[ \epsilon_b(v) = \frac{R_2 - \alpha_2 - \beta_2 + \sqrt{(R_2 - \alpha_2 - \beta_2)^2 + 4\beta_2 R_2}}{2v}. \]

\[ \text{To obtain the effective bandwidth } \epsilon_b(v) \text{ for the compensating source, we begin by defining } \Phi(v, u) \text{ [15, equation (86)]}\]

\[ \text{to be the matrix with sub-matrices} \]

\[ \Phi(v, u) \in \mathbb{C}^{4} \]
Denote by $\epsilon(D)$ the maximal real eigenvalue of a matrix $D$, then by [15, Sections 4 and 5], the effective bandwidth $eb_i(v)$ for fixed $v$ is the unique positive solution to the equation

$$\epsilon(\Phi(v, eb_i(v))) = 1.$$  \hfill (27)

Let $\eta$ be the minimum positive solution to

$$eb_{i}(\eta) + Neb_{j}(\eta) = c.$$  \hfill (28)

By [15, Section 6.2], for $i, j \in B$, the failure rate function $\lambda_{ij}(x)$ of the compensating source is

$$\lambda_{ij}(x) = \frac{[\Omega(t)]_{ij}}{[\Omega]_{ij} - [\Omega(t)]_{ij}}.$$  \hfill (29)

The failure rate $\lambda_{ij}(x)$ is said to be an increasing failure rate (IFR) if $\lambda_{ij}(x) \uparrow x$, and a decreasing failure rate (DFR) if $\lambda_{ij}(x) \downarrow x$. For $i \in B$, we denote by $\tau_i$ the expected sojourn time in $i$ of $A(t)$

$$\tau_i = -\sum_{j \in B} \bar{h}_{ij}(0).$$  \hfill (30)

by $\omega$ the stationary vector associated with $\Omega$, $\omega \Omega = 1$ and $\omega 1 = 1$, by $p$ the vector with elements

$$p_i = \frac{\omega_i \tau_i}{\sum_{j \in B} \omega_j \tau_j},$$  \hfill (31)

and by $h$ the left eigenvector of $\Phi(\eta, eb_{i}(\eta))$ corresponding to eigenvalue one, $h \Phi(\eta, eb_{i}(\eta)) = h$. We define $H_e$ as

$$H_e = \sum_{i \in B} \frac{h_i}{\eta(\bar{a}_i - eb_{i}(\eta))} \left( \sum_{j \in B} \Phi(\eta, eb_{i}(\eta))_{ij} - 1 \right),$$  \hfill (32)

and $\Psi_{\max}(i, j)$ and $\Psi_{\min}(i, j)$ for $i, j \in B$, depending on whether $\lambda_{ij}(x)$ is IFR or DFR and whether $\bar{a}_i > eb_i(\eta)$, as in Table 1, using [15, Theorem 7].

Applying [15, Theorem 6] and then simplifying using [14, Section 4.2.4], we obtain the following result.

**Theorem 1.** For $x > 0$,

$$K_x e^{-nx} \leq P(X > x) \leq K^* e^{-nx},$$  \hfill (33)

where

$$K_x = \frac{R_2}{eb_{i}(\eta) \alpha_2} \left( \frac{h_i}{\eta(\bar{a}_i - eb_{i}(\eta))} \sum_{j \in B} \Phi(\eta, eb_{i}(\eta))_{ij} - 1 \right)^N,$$  \hfill (34)

and

$$K^* = \frac{R_2}{eb_{i}(\eta) \alpha_2} \left( \frac{h_i}{\eta(\bar{a}_i - eb_{i}(\eta))} \sum_{j \in B} \Phi(\eta, eb_{i}(\eta))_{ij} - 1 \right)^N,$$  \hfill (35)

with $i, j \in B$ and $1 \leq s \leq N$ such that $\bar{a}_i + sR_2 > c$ and $[\Omega]_{i,j} > 0$.

**4. ACKNOWLEDGEMENT**

This work has been subsidized by the ARC grant AUWB-08/13-ULB 5 financed by the Ministère de la Communauté française de Belgique. The second author also gratefully acknowledges the hospitality of the Mathematical Institute at University of Wroclaw, where part of this work has been done.

**5. REFERENCES**


Table 1: Values of $\Psi_{\max}(i,j)$ and $\Psi_{\min}(i,j)$.

<table>
<thead>
<tr>
<th>$\Psi_{\max}(i,j)$</th>
<th>IFR, $\dot{a}_i &gt; \text{eb}_c(\eta)$</th>
<th>IFR, $\dot{a}_i \leq \text{eb}_c(\eta)$</th>
<th>DFR, $\dot{a}_i &gt; \text{eb}_c(\eta)$</th>
<th>DFR, $\dot{a}_i \leq \text{eb}_c(\eta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi_{\max}(i,j)$</td>
<td>$\frac{\Phi(\eta, \text{eb}<em>c(\eta))}{\Omega</em>{ij}p_i}$</td>
<td>$\frac{\tau_{i,j}}{p_i(\lambda_{ij}(\infty)-\eta(\dot{a}_i-\text{eb}_c(\eta)))}$</td>
<td>$\frac{\tau_{i,j}}{p_i(\lambda_{ij}(\infty)-\eta(\dot{a}_i-\text{eb}_c(\eta)))}$</td>
<td>$\frac{\Phi(\eta, \text{eb}<em>c(\eta))}{\Omega</em>{ij}p_i}$</td>
</tr>
<tr>
<td>$\Psi_{\min}(i,j)$</td>
<td>$\frac{\tau_{i,j}}{p_i(\lambda_{ij}(\infty)-\eta(\dot{a}_i-\text{eb}_c(\eta)))}$</td>
<td>$\frac{\Phi(\eta, \text{eb}<em>c(\eta))}{\Omega</em>{ij}p_i}$</td>
<td>$\frac{\Phi(\eta, \text{eb}<em>c(\eta))}{\Omega</em>{ij}p_i}$</td>
<td>$\frac{\tau_{i,j}}{p_i(\lambda_{ij}(\infty)-\eta(\dot{a}_i-\text{eb}_c(\eta)))}$</td>
</tr>
</tbody>
</table>


