A MATRIX-ANALYTIC APPROACH TO FLUID QUEUES WITH FEEDBACK CONTROL

A. DA SILVA SOARES & G. LATOUCHE

Département d'Informatique
Université Libre de Bruxelles
Blvd du Triomphe, CP 212, 1050 Bruxelles, Belgium
E-mail : andasilv@ulb.ac.be, latouche@ulb.ac.be

Abstract: Fluid queues consist of a buffer or reservoir controlled by a continuous-time Markov process evolving in the background. The content of the fluid buffer is a nonnegative real number, usually called the *level*, and the state of the underlying Markov process is called the *phase*. The renewal approach to fluid queues, originally developed by Ramaswami, is very efficient in that it allows for the analysis of different systems by a common set of tools. This is illustrated here with feedback fluid queues, that is, systems for which the rules of evolution of the phase process change when the buffer is either empty or full.

Key Words: Fluid queues, Markov models, performance modelling, communication systems.

INTRODUCTION

The purpose of this paper is to analyze feedback fluid queues using matrix analytic techniques, and to show how to derive performance measures of interest. A fluid queue consists of a buffer, or reservoir, controlled by a continuous time Markov process. We call it a feedback fluid queue if the control of the buffer content changes whenever it is empty or full, in order to regulate the amount of fluid in the buffer. For example, when a buffer with finite capacity becomes full, the behavior of the underlying Markov process might change so as to avoid, or at least reduce, fluid loss.

Our motivation for studying such a system is a model of an Internet congestion control protocol, namely the *Transport Control Protocol* (TCP), which is based on feedback (see van Foreest et al. [7, 8]). It roughly works as follows: a source sends packets into the network, the TCP receiver sends an acknowledgement to the source for each correctly received packet. As long as the source is informed that all the packets arrive at their destination, it increases its sending rate linearly in time. At some point in time, buffers along the network path start to fill and eventually overflow. When this happens, packets do not arrive correctly at the receiver, which, therefore, does not send acknowledgements to the source. The source reacts by reducing its sending rate exponentially fast in time, until congestion is cleared out. This scheme is called *linear increase/ multiplicative decrease*.

Fluid queues have a long tradition of being used to model the stochastic behaviour of telecommunication systems (see Anick et al. [1]). The authors in [7, 8] model the TCP protocol as a finite buffer fluid queue. As long as the buffer is not full, the sending rate increases according to a pure birth process; this corresponds to the 'linear increase' part of the protocol. When the buffer is full, the sending rate decreases according to a generator which implements a mechanism for the multiplicative decrease. The solution is obtained by spectral methods.

Our approach is based on renewal arguments and we treat separately the borders from the interior of the buffer, thereby being allowed to change in any way we want the evolution of the system when it reaches any of the two borders. The application of Markov renewal techniques in the analysis of fluid queues, as well as the connection with discrete time, discrete state space *Quasi Birth-and-Death* (QBD) processes, have shown their great efficiency; both were introduced and first exploited by Ramaswami [6]. In [4], we use renewal arguments to derive the stationary densities of the contents of both finite and infinite capacity fluid buffers; we show here that it is straightforward to adapt these results to feedback fluid queues.

The organization of this paper goes as follows. We define in the next section an infinite buffer model where the transition generator of the phase process changes whenever the buffer is empty, and we obtain its stationary density. In the third section, we deal with the finite buffer model and analyze a system where the generator of the phase process changes when the buffer is either empty or full. To simplify the equations, we temporarily only allow the input rates of fluid into the buffer to be equal to +1 or −1. This assumption is without loss of generality and we show later how to deal with the general setting. In addition to the distribution function of the buffer content, we compute the first and second moments of the buffer content. Finally, in the last section, we present an illustrative example.

Throughout the paper, vectors are denoted by lowercase boldface letters, and **0** and **1** denote vectors of zeros and ones, respectively.

INFINITE BUFFER

An infinite buffer fluid queue is a two-dimensional Markov process $\{(X(t), \varphi(t)) : t \in \mathbb{R}^+\}$ where $X(t)$ is called the level and denotes the content at time t of the fluid buffer, and $\varphi(t)$, called the phase, is the state at time t of the underlying Markov process which regulates the buffer content. The level $X(t)$ takes values in $\mathbb{R}^+$ and evolves as follows: at time t, if $\varphi(t) = i$, then

$$\frac{dX(t)}{dt} = \begin{cases} r_i, & \text{if } X(t) > 0, \\ \max(0, r_i), & \text{if } X(t) = 0, \end{cases}$$

where r_i can take any real value, including zero. The process $\{\varphi(t) : t \in \mathbb{R}^+\}$ is assumed to be irreducible and to have a finite state space $\mathcal{S}$.

We assume without loss of generality that the net input rates r_i are all equal to +1 or −1. We show in the section on general input rates how to return to the general setting. We decompose the state space of $\{\varphi(t)\}$ into two disjoint subsets $\mathcal{S}_+$ and $\mathcal{S}_-$, such that $\mathcal{S}_+ = \{i \in \mathcal{S} : r_i = 1\}$ and $\mathcal{S}_- = \{i \in \mathcal{S} : r_i = -1\}$, and we assume that neither $\mathcal{S}_+$ nor $\mathcal{S}_-$ are empty.

Traditional fluid queues are such that the environment process $\{\varphi(t)\}$ has, independently of the level of the buffer, the same generator T which we partition according to the decomposition of $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_-$; thus,

$$T = \begin{bmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{bmatrix}.$$

We denote by $\boldsymbol{\xi}$ its steady state probability vector, that is, the solution of the system $\boldsymbol{\xi} T = \mathbf{0}$, $\boldsymbol{\xi}\mathbf{1} = 1$, and we write that $\boldsymbol{\xi} = (\boldsymbol{\xi}_+, \boldsymbol{\xi}_-)$.

Consider now an infinite buffer fluid queue with feedback: $(X_f, \varphi_f) = \{(X_f(t), \varphi_f(t)) : t \in \mathbb{R}^+\}$, that is, suppose that the infinitesimal generator of $\varphi_f(t)$ remains T when $X_f(t) > 0$, but that it is $T^{(0)}$ when the buffer is empty, with $T^{(0)} = [T^{(0)}_{-+} \;\; T^{(0)}_{--}]$. Observe that we need not define $T^{(0)}_{ij}$ for $i \in \mathcal{S}_+$ since $X_f(t)$ cannot remain equal to zero when $r_i > 0$.

For $i \in \mathcal{S}$ and $x \geq 0$, let $F_i(x;t) = \Pr[X(t) \leq x, \varphi(t) = i]$ be the joint distribution of the level and the phase at time t. Define $f_i(x;t) = \partial/\partial x F_i(x;t)$ as the density of $(X(t), \varphi(t))$ evaluated at (x, i) for $i \in \mathcal{S}$, $x > 0$, and fixed t, with $f_i(0;t) = \lim_{x \to 0^+} f_i(x;t)$. The stationary density of the buffer content $\boldsymbol{\pi}(x) = (\pi_i(x) : i \in \mathcal{S})$ is the limit of $\boldsymbol{f}(x;t)$ as t goes to infinity. Let $\boldsymbol{p} = \lim_{t \to \infty} F(0;t)$ be the steady state probability mass vector of the empty buffer. Since it is not possible to have an empty buffer with a phase in $\mathcal{S}_+$, the sub-vector $\boldsymbol{p}_+ = \{p_i : i \in \mathcal{S}_+\}$ is equal to $\mathbf{0}$ and, therefore, $\boldsymbol{p} = (\mathbf{0}, \boldsymbol{p}_-)$.

The stationary density of the buffer content exists if and only if $\mu = \boldsymbol{\xi}_+\mathbf{1} - \boldsymbol{\xi}_-\mathbf{1} < 0$, which means that the mean drift of fluid into the buffer is negative. This condition will be referred to later on as the *stability condition.*

Theorem 1 *If $\mu < 0$, then the stationary density of the buffer content of the process (X_f, φ_f) for $x > 0$ is given by*

$$\boldsymbol{\pi}(x) = \boldsymbol{p}_- T^{(0)}_{-+} N_+(0, x)$$

where $(N_+(0,x))_{ij}$ is the expected number of visits to (x, j), under taboo of level zero, starting from $(0, i)$, for all j and for i in $\mathcal{S}_+$. This may also be written as

$$\boldsymbol{\pi}(x) = \boldsymbol{p}_- T^{(0)}_{-+} e^{Kx} [I \;\; \Psi] \tag{1}$$

where $K = T_{++} + \Psi T_{-+}$ and Ψ is the minimal nonnegative solution of the equation

$$T_{+-} + T_{++}\Psi + \Psi T_{--} + \Psi T_{-+} \Psi = 0.$$

The vector $\boldsymbol{p}_-$ is the unique solution of the system

$$\begin{cases} \boldsymbol{p}_- U^{(0)} &= \mathbf{0} \\ \boldsymbol{p}_-(\mathbf{1} - 2T^{(0)}_{-+} K^{-1}\mathbf{1}) &= 1 \end{cases} \tag{2}$$

where

$$U^{(0)} = T^{(0)}_{--} + T^{(0)}_{-+}\Psi \tag{3}$$

is the infinitesimal generator of the censored Markov process on the states $(0, \mathcal{S}_-)$.

Proof We omit the details because the proof is very similar to Theorems 2.1 and 2.2 in [4]: we repeat the same sequence of arguments and we only have to take into consideration the fact that the transition rates at level zero are not the same as elsewhere. □

To obtain the stationary distribution of the buffer content, one only needs to determine the matrix Ψ. This may be done by following the very efficient procedure introduced by Ramaswami [6] and for which the authors gave a probabilistic interpretation in [3]. This procedure is based on the Logarithmic-Reduction algorithm designed for discrete-level, discrete-time QBD processes (see Latouche and Ramaswami [5, Section 8.4]).

Since K and Ψ do not depend on $T^{(0)}$, we see from (1, 2) that one may change the behavior at level zero and obtain the new stationary distribution at no cost.

FINITE BUFFER

We now assume that the buffer has finite capacity b, and we denote by $(X^{(b)},\varphi) = \{(X^{(b)}(t),\varphi(t)) : t \in \mathbb{R}^+\}$ the resulting fluid queue. The level $X^{(b)}(t)$ takes values in the bounded interval $[0,b]$, and its evolution is governed as follows: if $\varphi(t) = i$,

$$\frac{dX^{(b)}(t)}{dt} = \begin{cases} r_i, & \text{if } 0 < X^{(b)}(t) < b, \\ \max(0,r_i), & \text{if } X^{(b)}(t) = 0, \\ \min(0,r_i), & \text{if } X^{(b)}(t) = b. \end{cases}$$

The fluid process is positive recurrent for any value of the drift μ.

The *feedback* fluid queue $(X_f^{(b)},\varphi_f)$ is a finite buffer fluid queue which differs from $(X^{(b)},\varphi)$ by the fact that the rates of the phase transitions change each time the buffer is empty or full. When it is empty $(X_f^{(b)}(t) = 0)$ and $\varphi_f(t)$ is in $\mathcal{S}_-$, the phase transition generator is $T^{(0)} = [T_{-+}^{(0)}\ T_{--}^{(0)}]$; when it is full $(X_f^{(b)}(t) = b)$ and $\varphi_f(t)$ is in $\mathcal{S}_+$, the generator is $T^{(b)} = [T_{++}^{(b)}\ T_{+-}^{(b)}]$.

We denote by $\boldsymbol{p}^{(0)}$ and $\boldsymbol{p}^{(b)}$ the probability vectors of levels zero and b for the process $(X_f^{(b)},\varphi_f)$. Like in the preceding section, $\boldsymbol{p}^{(0)} = (\boldsymbol{0},\boldsymbol{p}_-^{(0)})$; similarly, $\boldsymbol{p}^{(b)} = (\boldsymbol{p}_+^{(b)},\boldsymbol{0})$ since the buffer cannot be full with a phase in $\mathcal{S}_-$.

The next theorem gives the stationary density in terms of the boundary probability mass vectors $\boldsymbol{p}_-^{(0)}$ and $\boldsymbol{p}_+^{(b)}$, and of the matrices $N_+^{(b)}(x,y)$ and $N_-^{(b)}(x,y)$ of expected number of visits to level y, under taboo of the boundary levels, starting from a state in $(x,\mathcal{S}_+)$ or in $(x,\mathcal{S}_-)$, respectively.

Theorem 2 *For $0 < x < b$, the stationary density vector of the finite buffer fluid queue with feedback $(X_f^{(b)},\varphi_f)$ is given by*

$$\boldsymbol{\pi}^{(b)}(x) = (\boldsymbol{p}_-^{(0)}T_{-+}^{(0)},\boldsymbol{p}_+^{(b)}T_{+-}^{(b)}) \begin{bmatrix} N_+^{(b)}(0,x) \\ N_-^{(b)}(b,x) \end{bmatrix}. \quad (4)$$

If $\mu \neq 0$, this may also be written as

$$\boldsymbol{\pi}^{(b)}(x) = (\boldsymbol{v}_+,\boldsymbol{v}_-) \begin{bmatrix} e^{Kx} & e^{Kx}\Psi \\ e^{\hat{K}(b-x)}\hat{\Psi} & e^{\hat{K}(b-x)} \end{bmatrix} \quad (5)$$

where

$$(\boldsymbol{v}_+,\boldsymbol{v}_-) = (\boldsymbol{p}_-^{(0)}T_{-+}^{(0)},\boldsymbol{p}_+^{(b)}T_{+-}^{(b)}) \begin{bmatrix} I & e^{Kb}\Psi \\ e^{\hat{K}b}\hat{\Psi} & I \end{bmatrix}^{-1}, \quad (6)$$

the matrices K and Ψ are given in Theorem 1, $\hat{K} = T_{--} + \hat{\Psi}T_{+-}$ and $\hat{\Psi}$ is the minimal nonnegative solution of the equation

$$T_{-+} + T_{--}\hat{\Psi} + \hat{\Psi}T_{++} + \hat{\Psi}T_{+-}\hat{\Psi} = 0.$$

Proof To prove (4), we follow the steps in [4, Theorem 3.1]. The matrices $N_+^{(b)}(0,x)$ and $N_-^{(b)}(b,x)$ only deal with the system behavior between levels zero and b; therefore, they are the same as for the process $(X^{(b)},\varphi)$ and we apply [4, Lemmas 4.1 and 4.2] to conclude the proof. □

It is worth mentioning that $\hat{K}$ and $\hat{\Psi}$ play the same role as the matrices K and Ψ for the *level-reversed* fluid model. That process has the same characteristics as $(X^{(b)},\varphi)$, except for the net flow rates $\hat{r}_i$ which are equal to $-r_i$. Therefore, when the phase is in $\mathcal{S}_+$, the fluid level decreases, while when it is in $\mathcal{S}_-$, the level increases.

We shall need various matrices of first passage probabilities between levels zero and b, which we now define: $\Lambda_{ij}^{(b)}$ is the probability, starting from $(0,i)$ with i in $\mathcal{S}_+$, of reaching level b in phase j in $\mathcal{S}_+$, before returning to level zero, and $\Psi_{ik}^{(b)}$, with k in $\mathcal{S}_-$, is the probability of returning to level zero in phase k, without reaching level b. Similarly, $\hat{\Lambda}_{ik}^{(b)}$, with i and k in $\mathcal{S}_-$, is the probability, starting from (b,i), of reaching down to $(0,k)$ without returning to level b and $\hat{\Psi}_{ij}^{(b)}$ is the probability of returning to (b,j) before reaching down to level zero.

These probabilities only depend on the system behavior between the buffer boundaries. Therefore, they are the same as for the fluid model $(X^{(b)}, \varphi)$, and, for $\mu \neq 0$, are given by

$$\begin{aligned}
\Psi^{(b)}_{+-} &= (\Psi - e^{\hat{U}b}\Psi e^{Ub})(I - \hat{\Psi} e^{\hat{U}b}\Psi e^{Ub})^{-1} \\
\hat{\Psi}^{(b)}_{-+} &= (\hat{\Psi} - e^{Ub}\hat{\Psi} e^{\hat{U}b})(I - \Psi e^{Ub}\hat{\Psi} e^{\hat{U}b})^{-1} \\
\Lambda^{(b)}_{++} &= (I - \Psi\hat{\Psi})e^{\hat{U}b}(I - \Psi e^{Ub}\hat{\Psi} e^{\hat{U}b})^{-1} \\
\hat{\Lambda}^{(b)}_{--} &= (I - \hat{\Psi}\Psi)e^{Ub}(I - \hat{\Psi} e^{\hat{U}b}\Psi e^{Ub})^{-1},
\end{aligned}$$

as shown in [4], with $U = T_{--} + T_{-+}\Psi$ and $\hat{U} = T_{++} + T_{+-}\hat{\Psi}$.

The following theorem gives a procedure to determine $(\boldsymbol{p}^{(b)}_{+}, \boldsymbol{p}^{(0)}_{-})$.

Theorem 3 *The vector $(\boldsymbol{p}^{(b)}_{+}, \boldsymbol{p}^{(0)}_{-})$ is equal to $c(\boldsymbol{x}_+, \boldsymbol{x}_-)$, where $(\boldsymbol{x}_+, \boldsymbol{x}_-)$ is the unique solution of the system*

$$\left\{ \begin{array}{lcl} (\boldsymbol{x}_+, \boldsymbol{x}_-)W &=& (\mathbf{0}, \mathbf{0}) \\ \boldsymbol{x}_+\mathbf{1} + \boldsymbol{x}_-\mathbf{1} &=& 1 \end{array} \right. \quad (7)$$

with

$$W = \begin{bmatrix} T^{(b)}_{++} + T^{(b)}_{+-}\hat{\Psi}^{(b)}_{-+} & T^{(b)}_{+-}\hat{\Lambda}^{(b)}_{--} \\ T^{(0)}_{-+}\Lambda^{(b)}_{++} & T^{(0)}_{--} + T^{(0)}_{-+}\Psi^{(b)}_{+-} \end{bmatrix}.$$

The normalizing factor c is given by

$$\begin{aligned}
c &= \{\boldsymbol{x}_+\mathbf{1} + \boldsymbol{x}_-\mathbf{1} + \int_0^b \boldsymbol{y}(x)\mathbf{1}dx\}^{-1} \\
&= \{\boldsymbol{x}_+\mathbf{1} + \boldsymbol{x}_-\mathbf{1} + \boldsymbol{z}_+ \int_0^b e^{Kx}[I \ \Psi]dx\mathbf{1} \\
&\quad + \boldsymbol{z}_- \int_0^b e^{\hat{K}(b-x)}[\hat{\Psi} \ I]dx\mathbf{1}\}^{-1} \qquad (8)
\end{aligned}$$

where the vectors $\boldsymbol{y}(x)$ and $(\boldsymbol{z}_+, \boldsymbol{z}_-)$ are obtained by replacing $(\boldsymbol{p}^{(b)}_{+}, \boldsymbol{p}^{(0)}_{-})$ by $(\boldsymbol{x}_+, \boldsymbol{x}_-)$ in (5, 6).

Proof The vector $(\boldsymbol{p}^{(b)}_{+}, \boldsymbol{p}^{(0)}_{-})$ is proportional to the steady state probability vector of the restricted process obtained by observing the feedback fluid queue $(X^{(b)}_f, \varphi_f)$ only at those intervals of time where it is in $(0, \mathcal{S}_-)$ or in $(b, \mathcal{S}_+)$. It is easy to see, using the definition of the matrices $\Psi^{(b)}_{+-}$, $\hat{\Psi}^{(b)}_{-+}$, $\Lambda^{(b)}_{++}$ and $\hat{\Lambda}^{(b)}_{--}$, that the generator of the censored process is W. The remainder of the proof is immediate. □

The normalizing equation here is not as simple as in [4, Theorem 5.1]. In order to express the integrals in (8) in a simple manner, we will need to use generalized matrix inverses, because either K or $\hat{K}$ or both are singular as we show in Theorem 4 below. We use the group inverse, defined as follows: the group inverse of a matrix M, when it exists, is the unique matrix $M^{\#}$ such that $MM^{\#}M = M$, $M^{\#}MM^{\#} = M^{\#}$ and $MM^{\#} = M^{\#}M$. A useful property, which makes $M^{\#}$ easy to compute, is that $M^{\#}$, when it exists, is the unique solution of the system

$$\left\{ \begin{array}{lcl} M^{\#}M &=& I - \boldsymbol{v}\boldsymbol{u} \\ M^{\#}\boldsymbol{v} &=& \mathbf{0} \end{array} \right. \quad (9)$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ respectively denote the left and right eigenvectors of M for the eigenvalue 0, normalized by $\boldsymbol{u}\boldsymbol{v} = 1$, $\boldsymbol{u}\mathbf{1} = 1$. See Campbell and Meyer [2] for details.

Theorem 4 *If $\mu < 0$, the matrix K is non-singular and the matrix $\hat{K}$ is singular. The group inverse of $\hat{K}$ exists. Conversely, if $\mu > 0$, the matrix $\hat{K}$ is non-singular and the matrix K is singular. The group inverse of K exists.*

Proof We only prove the first assertion because the second is obtained in a similar manner.

Consider the QBD process obtained by restricting the infinite buffer fluid queue to those epochs when the level is a multiple of b. Its transition matrices are

$$A_0 = \begin{bmatrix} \Lambda^{(b)}_{++} & 0 \\ 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & \Psi^{(b)}_{+-} \\ \hat{\Psi}^{(b)}_{-+} & 0 \end{bmatrix} \quad (10)$$

$$\text{and} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & \hat{\Lambda}^{(b)}_{--} \end{bmatrix}.$$

Define R as the matrix which records the expected number of visits by the QBD to level 1, starting from level 0, before the first return to level 0. It is shown in [6] that e^{Kb} is the matrix of expected number of visits to the level b, for the fluid queue, starting from level 0 and before the first return to zero. Thus, we see that

$$R = \begin{bmatrix} e^{Kb} & e^{Kb}\Psi \\ 0 & 0 \end{bmatrix}.$$

Since $\mu < 0$, the drift is downwards and the spectral radius $\text{sp}(R)$ of R is strictly less than one [5, Corollary 7.1.2]; therefore, $\text{sp}(e^{Kb}) < 1$. This implies that all the eigenvalues of K must have a strictly negative real part, leading us to the conclusion that K is nonsingular.

To show that $\hat{K}$ is singular and that its group inverse exists, consider the QBD process obtained from (10) by reversing the levels. For this process, the matrix $\hat{R}$ of expected number of visits is given by

$$\hat{R} = \begin{bmatrix} 0 & 0 \\ e^{\hat{K}b}\hat{\Psi} & e^{\hat{K}b} \end{bmatrix}.$$

For the level-reversed process, the drift is upwards, which implies that $\text{sp}(\hat{R}) = 1$ and, by [5, Theorem 7.2.2], the eigenvalue 1 has multiplicity one. Therefore, $\text{sp}(e^{\hat{K}b}) = 1$, 0 is an eigenvalue of $\hat{K}$ and $\hat{K}$ is singular; all other eigenvalues of $\hat{K}$ have a strictly negative real part. Since the eigenvalue 0 has multiplicity one, the group inverse of $\hat{K}$ exists (see [2, Section 7.2]). □

The following lemma allows one to easily compute the integrals in (8).

Lemma 5 *If M is non-singular, then*

$$\int_0^b e^{Mx}dx = (I - e^{Mb})(-M)^{-1}. \qquad (11)$$

If M is singular and its group inverse exists, then

$$\int_0^b e^{M(b-x)}dx = (I - e^{Mb})(-M)^{\#} + b\boldsymbol{v}\boldsymbol{u} \qquad (12)$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ respectively denote the left and right eigenvectors of M for the eigenvalue 0, normalized by $\boldsymbol{u}\boldsymbol{v} = 1$, $\boldsymbol{u}\mathbf{1} = 1$.

Proof If M is invertible, equation (11) is easily proved [5, Equation 2.9]. Otherwise, we write that

$$\int_0^b e^{M(b-x)}dx = \int_0^b e^{Mx}dx = \sum_{n\geq 0} \frac{b^{n+1}}{(n+1)!}M^n.$$

By (9), $I = MM^{\#} + \boldsymbol{v}\boldsymbol{u}$ and we may write

$$\begin{aligned}\int_0^b e^{Mx}dx &= \sum_{n\geq 0} \frac{b^{n+1}}{(n+1)!}M^{n+1}M^{\#} \\ &\quad + \sum_{n\geq 0} \frac{b^{n+1}}{(n+1)!}M^n\boldsymbol{v}\boldsymbol{u} \\ &= \sum_{n\geq 1} \frac{(bM)^n}{n!}M^{\#} + b\boldsymbol{v}\boldsymbol{u}\end{aligned}$$

since $M\boldsymbol{v} = \mathbf{0}$. This finally reduces to

$$\int_0^b e^{Mx}dx = (I - e^{Mb})(-M)^{\#} + b\boldsymbol{v}\boldsymbol{u},$$

and the proof is complete. □

GENERAL INPUT RATES

We briefly indicate in this section how to obtain the solution for fluid queues with arbitrary input rates once we have the solution using input rates equal to 1 or −1.

Denote by $(\tilde{X}_f^{(b)}, \tilde{\varphi}_f)$ a finite buffer feedback fluid queue with arbitrary input rates $\tilde{r}_i$. The state space of $\tilde{\varphi}(t)$ is now decomposed in three disjoint subsets $\mathcal{S}_0 = \{i \in \mathcal{S} : r_i = 0\}$, $\mathcal{S}_+ = \{i \in \mathcal{S} : r_i > 0\}$ and $\mathcal{S}_- = \{i \in \mathcal{S} : r_i < 0\}$. The generators are denoted by $\tilde{T}$, $\tilde{T}^{(0)}$ and $\tilde{T}^{(b)}$.

We follow the procedure in the Appendix of [4] and obtain the following expressions for the probability mass vectors $\tilde{\boldsymbol{p}}^{(0)} = (\tilde{\boldsymbol{p}}_0^{(0)}, \mathbf{0}, \tilde{\boldsymbol{p}}_-^{(0)})$ and $\tilde{\boldsymbol{p}}^{(b)} = (\tilde{\boldsymbol{p}}_0^{(b)}, \tilde{\boldsymbol{p}}_+^{(b)}, \mathbf{0})$, and the density $\tilde{\boldsymbol{\pi}}^{(b)}(x) = (\tilde{\boldsymbol{\pi}}_0^{(b)}(x), \tilde{\boldsymbol{\pi}}_+^{(b)}(x), \tilde{\boldsymbol{\pi}}_-^{(b)}(x))$:

$$\tilde{\boldsymbol{p}}_-^{(0)} = \gamma \boldsymbol{x}_- |\tilde{C}_-|^{-1}, \quad \tilde{\boldsymbol{p}}_0^{(0)} = \tilde{\boldsymbol{p}}_-^{(0)} \tilde{T}_{-0}^{(0)} (-\tilde{T}_{00}^{(0)})^{-1},$$

$$\tilde{\boldsymbol{p}}_+^{(b)} = \gamma \boldsymbol{x}_+ |\tilde{C}_+|^{-1}, \quad \tilde{\boldsymbol{p}}_0^{(b)} = \tilde{\boldsymbol{p}}_+^{(b)} \tilde{T}_{+0}^{(b)} (-\tilde{T}_{00}^{(b)})^{-1},$$

$$\tilde{\boldsymbol{\pi}}_+^{(b)}(x) = \gamma \boldsymbol{y}_+(x) |\tilde{C}_+|^{-1},$$

$$\tilde{\boldsymbol{\pi}}_-^{(b)}(x) = \gamma \boldsymbol{y}_-(x) |\tilde{C}_-|^{-1},$$

and

$$\begin{aligned}\tilde{\boldsymbol{\pi}}_0^{(b)}(x) &= \tilde{\boldsymbol{\pi}}_+^{(b)}(x) \tilde{T}_{+0} (-\tilde{T}_{00})^{-1} \\ &\quad + \tilde{\boldsymbol{\pi}}_-^{(b)}(x) \tilde{T}_{-0} (-\tilde{T}_{00})^{-1},\end{aligned}$$

where $\tilde{C} = diag(\tilde{r}_i : i \in \mathcal{S}_+ \cup \mathcal{S}_-)$ and $\boldsymbol{x}_+$, $\boldsymbol{x}_-$ and $\boldsymbol{y}(x)$ are defined in Theorem 3.

The normalizing constant γ is given by

$$\begin{aligned}\gamma &= \{\boldsymbol{x}_+ \boldsymbol{w}_+^{(b)} + \boldsymbol{x}_- \boldsymbol{w}_-^{(0)} \\ &\quad + \int_0^b (\boldsymbol{y}_+(x)\boldsymbol{w}_+ + \boldsymbol{y}_-(x)\boldsymbol{w}_-)dx\}^{-1}\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{w}_+^{(b)} &= |\tilde{C}_+|^{-1}\{\mathbf{1} + \tilde{T}_{+0}^{(b)}(-\tilde{T}_{00}^{(b)})^{-1}\mathbf{1}\} \\ \boldsymbol{w}_-^{(0)} &= |\tilde{C}_-|^{-1}\{\mathbf{1} + \tilde{T}_{-0}^{(0)}(-\tilde{T}_{00}^{(0)})^{-1}\mathbf{1}\} \\ \boldsymbol{w}_+ &= |\tilde{C}_+|^{-1}\{\mathbf{1} + \tilde{T}_{+0}(-\tilde{T}_{00})^{-1}\mathbf{1}\} \\ \boldsymbol{w}_- &= |\tilde{C}_-|^{-1}\{\mathbf{1} + \tilde{T}_{-0}(-\tilde{T}_{00})^{-1}\mathbf{1}\}.\end{aligned}$$

Note the similarity to c in Theorem 3.

PERFORMANCE MEASURES

We are interested now in computing some performance measures for the marginal distribution of the fluid level, in a system with feedback and with arbitrary net input rates.

The probability masses $\tilde{m}_0$ and $\tilde{m}_b$ at levels zero and b are

$$\tilde{m}_0 = \tilde{\boldsymbol{p}}^{(0)}\mathbf{1} = \gamma \boldsymbol{x}_- \boldsymbol{w}_-^{(0)}$$

and
$$\tilde{m}_b = \tilde{\boldsymbol{p}}^{(b)}\mathbf{1} = \gamma \boldsymbol{x}_+ \boldsymbol{w}_+^{(b)},$$
and the density $\mu(x)$ for $0 < x < b$ is
$$\mu(x) = \tilde{\boldsymbol{\pi}}(x)\mathbf{1} = \gamma\{\boldsymbol{y}_+(x)\boldsymbol{w}_+ + \boldsymbol{y}_-(x)\boldsymbol{w}_-\}.$$
A first performance measure is the probability that the buffer overflows in the stationary regime, it is equal to $\tilde{m}_b$.

Remark 6 In the infinite buffer case, the density is $\mu(x) = \alpha\{\boldsymbol{\pi}_+(x)\boldsymbol{w}_+ + \boldsymbol{\pi}_-(x)\boldsymbol{w}_-\}$ and the probability mass at level zero is $\tilde{m}_0 = \alpha \boldsymbol{p}_- \boldsymbol{w}_-^{(0)}$, where $\boldsymbol{\pi}(x)$ and $\boldsymbol{p}_-$ are given in Theorem 1. The normalizing factor α is given by
$$\alpha = \Big(\boldsymbol{p}_- \boldsymbol{w}_-^{(0)} + \int_0^\infty (\boldsymbol{\pi}_+(x)\boldsymbol{w}_+ + \boldsymbol{\pi}_-(x)\boldsymbol{w}_-)dx\Big)^{-1}.$$

Next, we compute the stationary distribution function $\tilde{P}^{(b)}(x) = \lim_{t\to\infty} \Pr[\tilde{X}_f^{(b)}(t) \le x]$. We assume throughout that $\mu < 0$, so that K is nonsingular and $\hat{K}$ is singular; the equations have to be changed in an obvious manner in case $\mu > 0$.

Proposition 7 *If $\mu < 0$, the stationary distribution function of the buffer content of the process $(\tilde{X}_f^{(b)}, \tilde{\varphi}_f)$ is given by*
$$\tilde{P}^{(b)}(x) = \tilde{m}_0 + \gamma\left\{\boldsymbol{z}_+ A(x)[I, \Psi] + \boldsymbol{z}_- B(x)[\hat{\Psi}, I]\right\} \begin{bmatrix} \boldsymbol{w}_+ \\ \boldsymbol{w}_- \end{bmatrix}$$
for $0 \le x < b$, where $A(x) = (I - e^{Kx})(-K)^{-1}$ and $B(x) = e^{\hat{K}(b-x)}(I - e^{\hat{K}x})(-\hat{K})^{\#} + x\hat{\boldsymbol{v}}\hat{\boldsymbol{u}}$, and where $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ respectively denote the left and right eigenvectors of $\hat{K}$ for the eigenvalue 0, normalized by $\hat{\boldsymbol{u}}\hat{\boldsymbol{v}} = 1$, $\hat{\boldsymbol{u}}\hat{\mathbf{1}} = 1$. □

The proof is similar to the proof of Lemma 5 and is omitted.

The stationary mean and second moment of $(\tilde{X}_f^{(b)}, \tilde{\varphi}_f)$ are given next.

Proposition 8 *If $\mu < 0$, the mean M and second moment V of $\tilde{X}_f^{(b)}$ in stationary regime are given by*
$$M = b\tilde{m}_b + \gamma\left\{\boldsymbol{z}_+ C[I, \Psi] + \boldsymbol{z}_- D[\hat{\Psi}, I]\right\} \begin{bmatrix} \boldsymbol{w}_+ \\ \boldsymbol{w}_- \end{bmatrix} \quad (13)$$
and
$$V = b^2\tilde{m}_b + \gamma\left\{\boldsymbol{z}_+ E[I, \Psi] + \boldsymbol{z}_- F[\hat{\Psi}, I]\right\} \begin{bmatrix} \boldsymbol{w}_+ \\ \boldsymbol{w}_- \end{bmatrix} \quad (14)$$
where
$$C = (-K)^{-1}[(-K)^{-1}(I - e^{Kb}) - be^{Kb}] \quad (15)$$
$$D = (-\hat{K})^{\#}[bI - (I - e^{\hat{K}b})(-\hat{K})^{\#}] + \frac{b^2}{2}\hat{\boldsymbol{v}}\hat{\boldsymbol{u}} \quad (16)$$
$$E = 2(-K)^{-1}C - (-K)^{-1}b^2 e^{Kb}$$
$$F = 2\hat{K}^{\#}D + b^2(-\hat{K})^{\#} + \frac{b^3}{3}\hat{\boldsymbol{v}}\hat{\boldsymbol{u}},$$
and where $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ are defined in Proposition 7.

Proof The expected value of the stationary distribution is $M = b\tilde{m}_b + \int_0^b x\mu(x)dx$ and, by Theorem 3, is given by the right-hand side of (13), where $C = \int_0^b xe^{Kx}dx$ and $D = \int_0^b xe^{\hat{K}(b-x)}dx$.

By following the same argument as in Lemma 5, we show that
$$\int e^{Kx}dx = K^{-1}e^{Kx} + M_1$$
$$\int e^{\hat{K}x}dx = \hat{K}^{\#}e^{\hat{K}x} + x\hat{\boldsymbol{v}}\hat{\boldsymbol{u}} + M_2,$$
where M_1 and M_2 are arbitrary matrices. It is then easy to prove (15, 16) by part integration, keeping in mind that $\hat{K}^{\#}\hat{\boldsymbol{v}} = \mathbf{0}$.

One proves (14) in a similar manner. □

We finally turn to the Laplace-Stieltjes transform of the buffer content in equilibrium. We need to define the set $\hat{\sigma}$ of eigenvalues of $\hat{K}$ and the set $\hat{\Lambda} = \{-\lambda : \lambda \in \hat{\sigma}, \lambda \neq 0\}$.

Theorem 9 *If $\mu < 0$, the Laplace-Stieltjes transform $\phi(s)$ of $\tilde{P}^{(b)}$ is given by*
$$\phi(s) = \tilde{m}_0 + e^{-sb}\tilde{m}_b + \gamma\left\{\boldsymbol{z}_+ C(s)[I, \Psi] + \boldsymbol{z}_- D(s)[\hat{\Psi}, I]\right\} \begin{bmatrix} \boldsymbol{w}_+ \\ \boldsymbol{w}_- \end{bmatrix}, \quad (17)$$
for $\mathcal{R}(s) > 0$, where
$$C(s) = -(K - sI)^{-1}(I - e^{(K-sI)b}) \quad (18)$$
and
$$D(s) = (-\hat{K})^{\#}(e^{-sb}I - e^{\hat{K}b})(I + s\hat{K}^{\#})^{-1} + \frac{1}{s}(I - e^{-sb}I)\hat{\boldsymbol{v}}\hat{\boldsymbol{u}} \quad (19)$$
if $s \notin \hat{\Lambda}$, $D(s)$ being defined by continuity on $\hat{\Lambda}$.

Proof The Laplace-Stieltjes transform of $\tilde{P}^{(b)}$ is given by $\phi(s) = \tilde{m}_0 + e^{-sb}\tilde{m}_b + \int_0^b e^{-sx}\mu(x)dx$ which, by Theorem 3, is clearly seen to be equivalent to (17), where

$$C(s) = \int_0^b e^{(K-sI)x}dx$$
$$D(s) = \int_0^b e^{-s(b-x)}e^{\hat{K}x}dx.$$

We have seen in the proof of Theorem 4 that all the eigenvalues of K have a strictly negative real part, leading to the conclusion that the eigenvalues of $K - sI$ also have a strictly negative real part since $\mathcal{R}(s) > 0$. Therefore, $K - sI$ is nonsingular and (18) is proved.

We integrate by parts the expression for $D(s)$ and, keeping in mind the fact that $\hat{K}$ and $\hat{K}^{\#}$ commute, we find after some simple but tedious manipulations that

$$D(s)(I + s\hat{K}^{\#}) = (-\hat{K})^{\#}(e^{-sb}I - e^{\hat{K}b}) + \frac{1}{s}(I - e^{-sb}I)\hat{\boldsymbol{v}}\hat{\boldsymbol{u}}.$$

If s is not in $\hat{\Lambda}$, then $I + s\hat{K}^{\#}$ is invertible and, using the fact that $\hat{K}$ and $\hat{K}^{\#}$ commute and $\hat{\boldsymbol{u}}(I + s\hat{K}^{\#})^{-1} = \hat{\boldsymbol{u}}$, we finally obtain (19). □

NUMERICAL ILLUSTRATION

We now illustrate the results obtained in the preceding section. Consider a water reservoir of capacity b. There are N processes which consume water from the reservoir. Each process is either idle or it taps water at a constant rate c; it stays in the idle state for an exponentially distributed amount of time, with parameter β, then it enters in the active state and taps a quantity of water which is exponentially distributed, with parameter α/c. Under normal circumstances, therefore, a process leaves its active state at the constant rate α.

When the level of water is strictly between zero and b, the reservoir is normally filled at a constant rate R.

When the reservoir is full, there is a trigger which reacts after a random interval of time, and which reduces the rate at which the reservoir is filled, from R to $R/2$. If the reservoir remains full, a second trigger reacts, the input rate is reduced to zero and the reservoir stops being filled. The filling rate returns to R when another trigger reacts to the fact that the reservoir is not full anymore. The reaction time of each trigger is exponential with parameter γ.

For the reservoir to be empty, it is necessary that the number i of active processes should be such that $i > R/c$. At such a time, we assume that the incoming water is equally shared among the active processes. The consequence is that each process needs more time to accumulate the amount of water that it requests and the rate of transition to the idle state is $(\alpha R)/(ic)$ for each active process. When a probe detects this situation, the input rate is increased and becomes aR, where $a > 1$. At a later time, the level becomes positive again, and the filling rate returns to R when a trigger reacts to the fact that the reservoir has started filling.

We illustrate four different cases:

1. There is no feedback effect, the matrices $T^{(0)}$ and $T^{(b)}$ are equal to T.
2. Each process reduces its tapping rate at level zero, and there is no probe to detect that the reservoir is empty or full. In this case, $T^{(0)}$ is different but T and $T^{(b)}$ are still equal.
3. This is the same as Case 2, but now there are probes to detect that the reservoir is full. In this case, T, $T^{(0)}$ and $T^{(b)}$ are different.
4. This is the same as Case 3, with an additional probe to detect that the reservoir is empty.

The parameters chosen are the following: $\alpha = 1$ and $c = 1$, so that the unit of volume is fixed to be the expected quantity of water which is taken by a process and the unit of time is the expected duration of the active state, under normal conditions; the number of processes is $N = 20$; we set $R = 1.01N\beta c/(\alpha + \beta)$ so that the normal filling rate of the reservoir is slightly above what is required by the processes, on the average, and the drift μ is positive; finally, $\beta = 0.1$, $\gamma = 10$, $a = 2$ and $b = 2R$.

The probability masses $\tilde{m}_0$ and $\tilde{m}_b$ of levels zero and b, and the first two moments are given in Figure 1 below, and the distribution function of the content of the reservoir in stationary regime is given in Figure 2.
The difference between the first two cases is due to the fact that, at level zero, each process receives water at a lower rate and this creates a positive feedback loop: the processes remain active longer, this in turns give more opportunities for idle processes to become active, which further increases the time spent by each process in its active state, etc.

Case	1	2	3	4
$\tilde{m}_0$	0.1500	0.4171	0.4202	0.1040
$\tilde{m}_b$	0.1764	0.1210	0.0784	0.1126
M	1.7574	1.2053	1.0607	1.7460
V	5.1293	3.5177	2.9635	5.0432

Figure 1: Steady State Probability Masses of Levels 0 and b, Mean M and Second Moment V of the Reservoir Content

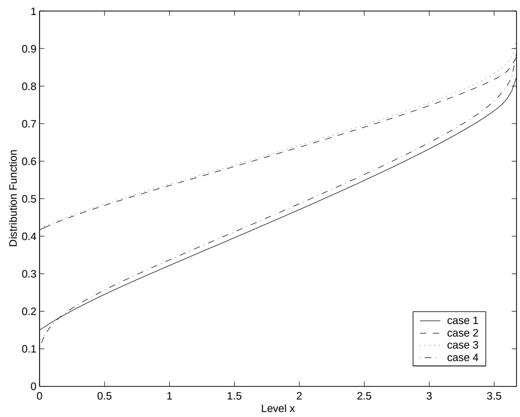

Figure 2: Stationary Distribution Function

In Case 4, the probe at level zero is fast ($\gamma = 10$) and the input rate to the reservoir is quickly doubled, so that the system spends little time at level zero. By increasing a, we may force $\tilde{m}_0$ to be even smaller.

The comparison of Cases 2 and 3 shows, as expected, that the probability of overflowing decreases when there are probes to detect that the reservoir is full. It appears that the idea of reducing the inflow by one half before cutting it altogether might not be very efficient.

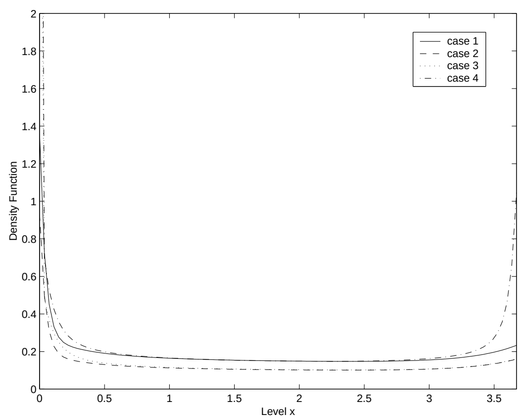

Figure 3: Stationary Density Function

The density function of the reservoir content in stationary regime is given in Figure 3. One clearly sees that the density is nearly uniform over most of the interval $(0, b)$. This is due to the fact that the drift is positive. If b were infinite, the fluid queue would be transient; with b finite, the stationary distribution tends to spread evenly over the whole state space. If we had chosen $\mu < 0$, the density would be more concentrated on the left of the interval $(0, b)$.

References

[1] D. Anick, D. Mitra, and M. M. Sondhi. Stochastic theory of a data handling system with multiple sources. *Bell System Tech. J.*, 61:1871–1894, 1982.

[2] S. L. Campbell and C. D. Meyer. *Generalized Inverses of Linear Transformations.* Dover Publications, New York, 1991. Republication.

[3] A. da Silva Soares and G. Latouche. Further results on the similarity between fluid queues and QBDs. In G. Latouche and P. Taylor, editors, *Matrix-Analytic Methods Theory and Applications – Proceedings of the 4th International Conference on Matrix-Analytic Methods*, pages 89–106. World Scientific, 2002.

[4] A. da Silva Soares and G. Latouche. Matrix-analytic methods for fluid queues with finite buffers. Submitted for publication, 2003.

[5] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling.* ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia PA, 1999.

[6] V. Ramaswami. Matrix analytic methods for stochastic fluid flows. In D. Smith and P. Hey, editors, *Teletraffic Engineering in a Competitive World (Proceedings of the 16th International Teletraffic Congress)*, pages 1019–1030. Elsevier Science B.V., Edinburgh, UK, 1999.

[7] N. van Foreest, M. Mandjes, and W. Scheinhardt. Performance analysis of heterogeneous interacting TCP sources. *Memorandum No. 1607, Faculty of Mathematical Sciences, University of Twente, Enschede, The Netherlands*, 2001.

[8] N. van Foreest, M. Mandjes, and W. Scheinhardt. Analysis of a feedback fluid model for heterogeneous TCP sources. *Comm. Statist. Stochastic Models*, 19:299–324, 2003.

Ana da Silva Soares studied mathematics at the Université Libre de Bruxelles and obtained her degree in 2000. She is now a PhD student at the Computer Science department of the same university. Her research interests include matrix-analytic and algorithmic methods in Markov models, fluid processes driven by a markovian environment and QBD processes. Her e-mail address is `andasilv@ulb.ac.be`.

Guy Latouche is Professor of Applied Probability. His research interests include various aspects of computational probability: matrix methods in Markov models, traffic models for telecommunication systems, fluid queues and nearly completely decomposable systems. He has published about 80 research papers and is editor of seven conference proceedings. He has written three books, among which "Introduction to Matrix Analytic Methods in Stochastic Modeling" with V. Ramaswami and "Numerical Methods for Structured Markov Chains" with D. Bini and B. Meini. His e-mail address is `latouche@ulb.ac.be`.