

Fluid Models with Jumps

Elena I. Tzenova*, Ivo J.B.F. Adan†, Vidyadhar G. Kulkarni‡

August 2, 2004

Abstract

In this paper we study a general stochastic fluid model with a single infinite capacity buffer, where the buffer content can change continuously as well as by instantaneous upward jumps. The continuous as well as the instantaneous change is modulated by an external environment process modelled as a finite state continuous time Markov chain. The Laplace-Stieltjes transform of the steady-state joint distribution of the buffer content and the state of the environment is determined explicitly in terms of the solutions of a generalized eigenvalue problem. The methodology is illustrated by several well-known queueing problems.

1 Introduction

A stochastic fluid queueing system describes the input-output flow of a fluid in a storage device, called a buffer. The rates at which the fluid enters and leaves the buffer depend on a random environment process that is usually chosen to be an irreducible CTMC. The study of fluid flow models is motivated by their various applications. One of these are the real-world systems that deal with the processing of continuous entities such as the ones used in the petroleum and chemical industries. Fluid flow models also provide an important tool for the performance analysis of high-speed data networks, or large-scale production systems where a large number of relatively small jobs are processed. Fluid flow models are also used as models of the asymptotic behavior of queues in heavy traffic.

Most of the classical research on stochastic fluid systems in the area of telecommunications is based on the work of Anick, Mitra and Sondhi [2] which is an extension of the pioneering work of Kosten [4]. They study a model with several identical and independent Markov on-off sources that transmit fluid to an infinite-capacity buffer. The fluid is then

*Department of Mathematics, Tulane University, New Orleans, LA 70118-5698

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

‡Department of Operations Research, University of North Carolina, CB 3180, Chapel Hill, N.C. 27599

processed at a fixed rate. The limiting distribution of the buffer content process is computed as a solution of a set of ordinary differential equations. The identical input sources facilitate the analysis of the differential equations and the main result provides the system's eigenvalues in explicit form. In [8], [9] Mitra generalizes this model by introducing multiple on-off switching. Most of the work in fluid queues deals with the steady-state distribution of the buffer content. See the survey paper by Kulkarni [5] for an extensive overview of the research in this area.

In this paper we consider a modification of the classical fluid model, where in addition to the continuous changes, the buffer content can have instantaneous upward jumps. The rate of the continuous change and the size of the jumps both depend on the state of the environment. The aim of this paper is to determine the Laplace-Stieltjes transform of the steady-state joint distribution of the buffer content and state of the environment. This problem can be reduced to that of finding the boundary probabilities (of an empty buffer). We show that the boundary probabilities can be explicitly determined in terms of the solutions of a generalized eigenvalue problem. The developed technique is illustrated by some simple example of well-known queueing problems.

Fluid models with upward, respectively downward, jumps are natural continuous analogues of the $M/G/1$ -type and $G/M/1$ -type Markov chains studied by Neuts [10, 11]. Recently matrix-analytical results for the discrete problem have been extended to the continuous case. Takada [?] studies the model with upward jumps, and takes the following approach to find the boundary probabilities: instead of using the solutions of an eigenvalue problem, Takada first determines the fundamental matrix, similar to the one for the $M/G/1$ -type Markov chain, and then obtains the boundary probabilities, up to a multiplicative constant, as the steady-state distribution of the embedded Markov chain at visits to the boundary states. Sengupta [12] and Miyazawa and Takada [?] study the fluid model with downward jumps; they show that the steady-state joint distribution of the buffer content and state of the environment has a matrix-exponential form, where the exponent matrix is the solution of a non-linear matrix integral equation.

2 The Model

Consider a general stochastic fluid model with a single infinite capacity buffer where the buffer content $X(t)$ can increase continuously as well as by instantaneous jumps. The change of the fluid in the buffer depends on the state of an external random environment process $\{I(t), t \geq 0\}$ which is taken to be a stochastic process with a finite state space $S = \{1, \dots, N\}$. While $I(t) = i$ the buffer content increases continuously with rate $r_i \in (-\infty, \infty)$. The process $\{I(t), t \geq 0\}$ jumps to any state $j \in S$ (not necessarily different from i) with probability p_{ij} . In state $i \in S$ it will make a jump after an exponential amount of time with mean $1/q_i$. When the $I(t)$ process jumps from state i to state j the amount of fluid in the buffer can increase by a lump-sum *non-negative* random amount with a given

c.d.f. $G_{ij}(y)$, $y \geq 0$ and mean m_{ij} . Thus, the bivariate Markov process $\{(X(t), I(t)), t \geq 0\}$ can jump from state (x, i) to state $(x + y, j)$ with rate $q_i p_{ij} dG_{ij}(y)$, $y \geq 0$. Let R denote the diagonal $N \times N$ matrix with the net input rates r_i along the diagonal

$$R := \text{diag}[r_1, \dots, r_N],$$

and also define

$$Q_{ij}(x) := q_i p_{ij} G_{ij}(x), \quad x \geq 0, \quad i, j \in S, \quad i \neq j, \quad (2.1)$$

$$Q_{ii}(x) := q_i p_{ii} G_{ii}(x) - q_i, \quad x \geq 0, \quad i \in S. \quad (2.2)$$

Note that $\{I(t), t \geq 0\}$ itself is a CTMC on S with rate matrix $Q = Q(\infty)$. We assume that $\{I(t), t \geq 0\}$ is irreducible. Let $\pi_i := \lim_{t \rightarrow \infty} P(I(t) = i)$ be the limiting distribution of the $I(t)$ process. The following results states that the system is stable if the mean net input rate in steady state is negative (cf. Loynes [?]).

Theorem 2.1 *The system is stable if*

$$\sum_{i=1}^N \pi_i (r_i + \sum_{j=1}^N q_i p_{ij} m_{ij}) < 0. \quad (2.3)$$

Proof: Consider the sampled chain at jumps of the $I(t)$ process to state 1. Let X_n denote the buffer content at the n -th jump to state 1. Then it is readily verified that

$$\lim_{x \rightarrow \infty} E(X_{n+1} | X_n = x) = x + \frac{1}{\pi_1 q_1} \sum_{i=1}^N \pi_i (r_i + \sum_{j=1}^N q_i p_{ij} m_{ij}).$$

Hence, by virtue of (2.3), there exist positive constants ϵ, b, K such that for all $x \geq 0$,

$$E(X_{n+1} | X_n = x) - x \leq -\epsilon + b \mathbf{1}_{[x \leq K]},$$

where $\mathbf{1}_{[\cdot]}$ is the indicator function. From Foster's criterion (see e.g. Meyn and Tweedie [?]) we can now conclude that the sampled chain is stable. Since the times elapsing between successive jumps to state 1 are iid with finite mean, the result follows for the original chain. \diamond

Define the matrix $\Gamma = [\Gamma_{ij}]$ by

$$\Gamma_{ij} = q_i p_{ij} m_{ij}.$$

Then the stability condition (2.3) can be written in matrix form as follows

$$\pi (R + \Gamma) e < 0.$$

We shall assume from now on that this stability condition holds.

Now, denote

$$F_i(t, x) := P(X(t) \leq x, I(t) = i), \quad x, t \geq 0, \quad i \in S,$$

and

$$F_i(x) := \lim_{t \rightarrow \infty} P(X(t) \leq x, I(t) = i), \quad x \geq 0, \quad i \in S.$$

The next theorem shows that $F_i(x)$ is differentiable for $x > 0$ and gives the differential equations satisfied by $F_i(x)$. First we introduce the following notation:

$$\begin{aligned} F(x) &:= [F_1(x), \dots, F_N(x)], \\ \frac{dF}{dx} &:= \left[\frac{dF_1}{dx}, \dots, \frac{dF_N}{dx} \right], \\ S_- &:= \{i \in S : r(i) < 0\}, \quad N_- := |S_-|, \\ S_0 &:= \{i \in S : r(i) = 0\}, \quad N_0 := |S_0|, \\ S_+ &:= \{i \in S : r(i) > 0\}, \quad N_+ := |S_+|. \end{aligned}$$

Theorem 2.2 *The limiting distribution $F(x)$ satisfies*

$$\frac{dF}{dx} R = F * Q(x). \quad (2.4)$$

The boundary conditions are given by

$$F_i(0) = 0, \quad i \in S_+, \quad (2.5)$$

$$F(0)Q_{\cdot j}(0) = 0, \quad j \in S_0, \quad (2.6)$$

where $Q_{\cdot j}(0)$ is the j -th column of $Q(0)$.

Proof: First, consider $F_j(t, x)$, $x > 0, j \in S$ and condition on a small time interval of length $h > 0$ so that

$$\begin{aligned} F_j(t, x) &= \sum_{i=1}^N \int_{z=0^-}^{x-r_i h} P(X(t-h) \leq x - r_i h - z, I(t-h) = i) q_i p_{ij} h dG_{ij}(z) \\ &\quad + (1 - q_j h) P(X(t-h) \leq x - r_j h, I(t-h) = j) + o(h) \\ &= h \sum_{i=1}^N \int_{z=0^-}^{x-r_i h} F_i(t-h, x - r_i h - z) q_i p_{ij} dG_{ij}(z) + (1 - q_j h) F_j(t-h, x - r_j h) + o(h). \end{aligned}$$

Now let $t \rightarrow \infty$ and rearrange the last equation so that

$$\frac{F_j(x) - F_j(x - r_j h)}{h} = \sum_{i=1}^N \int_{z=0^-}^{x-r_i h} F_i(x - r_i h - z) q_i p_{ij} dG_{ij}(z) - q_j F_j(x - r_j h) + \frac{o(h)}{h} \quad (2.7)$$

From the definition of $Q(x)$ in (2.1) we have

$$dQ_{ij}(x) = q_i p_{ij} dG_{ij}(x), \quad x \geq 0, \quad i, j \in S, \quad i \neq j.$$

and from (2.2),

$$dQ_{ii}(x) = q_i p_{ii} dG_{ii}(x), \quad x > 0, i \in S.$$

Since $Q_{ii}(0-) = G_{ii}(0-) = 0$ we obtain for the jump of $Q_{ii}(x)$ at $x = 0$,

$$dQ_{ii}(0) = Q_{ii}(0) = q_i p_{ii} dG_{ii}(0) - q_i, \quad i \in S.$$

Now, we can write Eq. (??) in the following nice form:

$$\frac{F_j(x) - F_j(x - r_j h)}{h} = \sum_{i=1}^N \int_{z=0^-}^{x-r_i h} F_i(x - r_i h - z) dQ_{ij}(z) + \frac{o(h)}{h}.$$

After we let $h \rightarrow 0$ we get

$$r_j \frac{dF_j(x)}{dx} = \sum_{i=1}^N \int_{z=0^-}^x F_i(x - z) dQ_{ij}(z) = \sum_{i=1}^N F_i * Q_{ij}(x), \quad (2.8)$$

which in matrix form becomes Eq. (2.7), that is,

$$\frac{dF}{dx} R = F * Q(x).$$

It remains to establish the boundary conditions. From the definition of $F_i(x)$ we have

$$F_i(0) = \lim_{t \rightarrow \infty} P(X(t) = 0, I(t) = i), \quad i \in S.$$

Therefore, for states $i \in S_+$ with positive net input rates the long-run probabilities $F_i(0)$ are zero. Thus, the first set of boundary conditions (2.8) is given by

$$F_i(0) = 0, \quad i \in S_+.$$

To get the second set of boundary conditions (2.9) we can again apply conditioning on a small time interval of length $h > 0$, for $x = 0$, $j \in S_0$, as above to obtain

$$\sum_{i=1}^N \int_{z=0^-}^{(-r_i h)^+} F_i(t - h, (-r_i h)^+ - z) q_i p_{ij} h dG_{ij}(z) + (1 - q_j h) F_j(t - h, 0) = F_j(t, 0), \quad j \in S_0,$$

where $(\cdot)^+ = \max(\cdot, 0)$. After letting $t \rightarrow \infty$ and using the notation $dQ_{ij}(z)$ we get

$$\sum_{i=1}^N \int_{z=0^-}^{(-r_i h)^+} F_i((-r_i h)^+ - z) dQ_{ij}(z) = 0, \quad j \in S_0.$$

Now, as $h \rightarrow 0$ the boundary conditions (2.9) are obtained

$$\sum_{i=1}^N F_i(0) Q_{ij}(0) = 0, \quad j \in S_0.$$

◇

Remark: Equations (2.7)-(2.9) have a simple rate interpretation. For example, if $r_j > 0$, then the rate out of the set of states $\{(y, j), 0 \leq y \leq x\}$ is equal to

$$r_j \frac{dF_j(x)}{dx} - F_j * Q_{jj}(x)$$

and the rate into that set,

$$\sum_{i \neq j} F_i * Q_{ij}(x).$$

Equating the two rates yields Eq. (2.12).

We solve the differential equations (2.7) of the above theorem using Laplace Stieltjes Transforms (LST). By Eqs. (2.1) and (2.2), The LST of $Q(x)$ is given by

$$\tilde{Q}_{ij}(s) = q_i p_{ij} \tilde{G}_{ij}(s), \quad i, j \in S, \quad i \neq j, \quad (2.9)$$

$$\tilde{Q}_{ii}(s) = q_i p_{ii} \tilde{G}_{ii}(s) - q_i, \quad i \in S, \quad (2.10)$$

where the LSTs are defined as usual, i.e.

$$\tilde{Q}_{ij}(s) := \int_0^\infty e^{-sx} dQ_{ij}(x), \quad \tilde{G}_{ij}(s) := \int_0^\infty e^{-sx} dG_{ij}(x), \quad i, j \in S.$$

Note that one needs to account for the jump of $Q_{ii}(x)$ at $x = 0$ in deriving Eq. (2.14). The Eqs. (2.13) and (2.14) imply that $\tilde{Q}(0) = Q$ is the generator matrix of the CTMC $\{I(t), t \geq 0\}$. Taking the LST of both sides of Eq. (2.7) we get

$$\tilde{F}(s)(sR - \tilde{Q}(s)) = sF(0)R. \quad (2.11)$$

Thus, in order to find the LST $\tilde{F}(s)$ we need to know $F(0) = [F_1(0), \dots, F_N(0)]$. From the boundary conditions (2.8) we have that N_+ components of $F(0)$ are zero and N_0 components can be expressed in terms of the remaining N_- from the second set of boundary conditions (2.9). The next theorem is used to determine the remaining $N_- = N - N_+ - N_0$ components of $F(0)$ (cf. Theorem 5 in Loynes [6]).

Theorem 2.3 *Assume Condition (2.3) holds. Then the generalized eigenvalue problem*

$$(sR - \tilde{Q}(s))\phi = 0 \quad (2.12)$$

has exactly N_- solutions $(s_1, \phi_1), \dots, (s_{N_-}, \phi_{N_-})$, with $s_1 = 0, \operatorname{Re}(s_i) > 0, i = 2, \dots, N_-$ and $\phi_i \neq 0$. Furthermore, these zeros s_1, \dots, s_{N_-} lie on or inside the circle in the complex plane with center at $\lambda = \max_{i:r_i \neq 0} -\frac{q_i}{r_i}$ and radius λ .

Proof: is given in the Appendix. ◇

The next theorem is the main result of this paper. It summarizes the entire set of equations satisfied by $F(0)$.

Theorem 2.4 *Assume Condition (2.3) holds. Then the LST row vector $\tilde{F}(s)$ satisfies*

$$\tilde{F}(s)(sR - \tilde{Q}(s)) = sF(0)R, \quad (2.13)$$

where the unknowns $F(0) = [F_1(0), F_2(0), \dots, F_N(0)]$ are given by the unique solution to the N equations

$$F_i(0) = 0, \quad i \in S_+, \quad (2.14)$$

$$F(0)Q_{\cdot, i}(0) = 0, \quad i \in S_0,$$

$$F(0)R\phi_i = 0, \quad \text{for } i = 2, \dots, N_-, \quad (2.15)$$

$$F(0)Re = \pi(R + \Gamma)e. \quad (2.16)$$

Proof: The equation for $\tilde{F}(s)$ is as derived above, Eq. (2.15). Also, recall the first two sets of equations for $F_i(0)$ when $i \in S_+$ or $i \in S_0$, given in Theorem 2.1, Eqs. (2.8) and (2.9). Since $\tilde{F}(s)$ is analytic for $\text{Re}(s) \geq 0$, it must be the case that for every i ,

$$\tilde{F}(s_i)(s_i R - \tilde{Q}(s_i))\phi_i = s_i F(0)R\phi_i = 0.$$

However, for $s_1 = 0$ the last equation is trivially satisfied and we need one additional condition to determine $F(0)$. Thus, we get Eqs. (2.20). Now, let

$$M(s) := sR - \tilde{Q}(s)$$

and $M'(s)$ be the derivative of $M(s)$. To derive the normalization equation (2.21) first recall that from (2.17) we have

$$\tilde{F}(s)M(s) = sF(0)R. \quad (2.17)$$

After differentiating equation (2.23) and setting $s = 0$ we get

$$\tilde{F}(0)M'(0) + \tilde{F}'(0)M(0) = F(0)R.$$

Post-multiplying by e and noting that $M(0)e = -\tilde{Q}(0)e = 0$, yields

$$F(0)Re = \tilde{F}(0)M'(0)e.$$

Since $\tilde{F}(0) = F(\infty) = \pi$ and $M'(0) = R + \Gamma$, we get Eq. (2.21). The uniqueness of the solution follows from the assumed stability and by following similar arguments as in the proof of Theorem 3 of Gail et al. [?]. ◇

Remark: Theorem 2.3 characterizes the LST $\tilde{F}(s)$ as the solution to Eq. (2.17), where the boundary vector $F(0)$ is determined by the solutions to the generalized eigenvalue problem 2.16 through the Eqs. (??)-(2.21). The equivalent result to Theorem 2.3 is given in Takada [?], Proposition 2.1, where, as already mentioned in the Introduction, the boundary vector $F(0)$ is determined through the fundamental matrix.

3 Examples

We illustrate the above methodology with the help of several examples.

3.1 The $M|G|1$ Queue.

Consider a standard $M|G|1$ queue with arrival rate λ and service time distribution G with mean τ . Let $X(t)$ be the work content in this system at time t . The $\{X(t), t \geq 0\}$ process decreases at rate 1 and jumps up by a random amount with cdf G whenever an arrival occurs. If $X(t)$ becomes zero, it stays zero until an upward jump occurs. Thus we can model this as a fluid process with jumps with a single state environment process. The parameters are as follows: $S = \{1\}$, $q_1 = \lambda$, $p_{1,1} = 1$, $G_{1,1} = G$, $m_{1,1} = \tau$, $r_1 = -1$. Hence $\pi_1 = 1$, and the condition of stability (2.3) reduces to $-1 + \lambda\tau < 0$, or $\rho = \lambda\tau < 1$. This is the standard condition of stability for the $M|G|1$ queue. We get $\tilde{Q}(s) = \lambda(\tilde{G}(s) - 1)$, and Equation (2.17) reduces to the scalar equation

$$\tilde{F}(s)(s - \lambda(1 - \tilde{G}(s))) = sF(0).$$

There is only one boundary condition, given by Eq. (2.21). It reduces to $F(0) = 1 - \lambda\tau = 1 - \rho$. Using that we get

$$\tilde{F}(s) = \frac{s(1 - \rho)}{s - \lambda(1 - \tilde{G}(s))}.$$

This matches the LST of the queueing time in an $M|G|1$ queue, as it must.

3.2 The $BMAP|G|1$ Queue.

Consider a single server queue whose arrival process is given by a $BMAP$ defined by the sequence $\{D_k, k \geq 0\}$. Here D_0 has negative diagonal elements and nonnegative off-diagonal elements, $D_k, k \geq 1$, are nonnegative, and D , defined as

$$D = \sum_{k=0}^{\infty} D_k,$$

is the generator an irreducible CTMC on state space $S = \{1, 2, \dots, N\}$. Transitions according to D_k correspond to batch arrivals of size k . The service times are iid random variables with cdf G and mean τ . Let $I(t)$ be the phase of the arrival process at time t . It is a CTMC on S with generator matrix D . Let π be the limiting distribution of $I(t)$.

Let $X(t)$ be the work content in the system at time t . The $\{X(t), t \geq 0\}$ process decreases at rate 1 and jumps up by a random amount with cdf G^{*k} whenever a batch arrival of size k occurs. If $X(t)$ becomes zero, it stays zero until an upward jump occurs. Thus we can model the work content process as a fluid process with jumps with the following parameters: The state space of the environment process (the phase of the arrival process) is $S = \{1, 2, \dots, N\}$. We get

$$\tilde{Q}(s) = \sum_{k=0}^{\infty} D_k \tilde{G}^k(s) = D(\tilde{G}(s)),$$

where

$$D(z) = \sum_{k=0}^{\infty} D_k z^k, \quad |z| \leq 1.$$

We also have $R = -I$, since the net rate is -1 in every state. The condition of stability is

$$\pi \sum_{k=1}^{\infty} k D_k e \tau < 1.$$

Since $S_- = S$, Theorem 2.2 implies that there are N (eigenvalue, eigenvector) pairs (s_i, ϕ_i) ($i = 1, 2, \dots, N$) satisfying

$$(sI + D(\tilde{G}(s)))\phi = 0,$$

with $s_1 = 0$ and $Re(s_i) > 0$ for $i = 2, 3, \dots, N$. Finally, Theorem 2.3 yields the following result for the LST of F :

$$\tilde{F}(s) = sF(0)(sI + D(\tilde{G}(s)))^{-1}, \quad (3.18)$$

where the unknown vector $F(0)$ is obtained from the N equations:

$$F(0)\phi_i = 0, \quad i = 2, 3, \dots, N.$$

$$F(0)e = 1 - \pi \sum_{k=1}^{\infty} k D_k e \tau.$$

This matches with the known results; see Eq. (44) in Lucantoni [7].

3.3 The $PH|G|1$ Queue.

The PH arrival process with parameters (α, T) is a special case of the $BMAP$. Here $D_0 = T$, $D_1 = -Te\alpha$ and $D_k = 0$ for all $k > 1$, where $T = [t_{i,j}]_{i,j=1,2,\dots,N}$ is an invertible generator matrix with non-positive row sums of an irreducible CTMC on state space $S = \{1, 2, \dots, N\}$, and $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]$ is a non-negative vector whose elements add up to one. Eq. (3.25) for the LST of F now simplifies to

$$\tilde{F}(s) = sF(0)(sI + T - Te\alpha\tilde{G}(s))^{-1},$$

which matches with Eq. (5.2.4), p. 257 of Neuts [10].

3.4 The $M|G|1$ Queue with an Up-Down Server.

Consider the queueing system of the first example, with the following modification: The server can be up or down. The up times are $\text{Exp}(\beta)$ random variables, and the down times are $\text{Exp}(\alpha)$ random variables. Let $I(t)$ be the state of the server (1 if up, and 0 if down) at time t . Thus $\{I(t), t \geq 0\}$ is a CTMC. Let $X(t)$ be the work content at time t in this system. We model $\{(X(t), I(t)), t \geq 0\}$ as a fluid queue with jumps with the following parameters:

$$q_0 = \lambda + \alpha, \quad q_1 = \lambda + \beta$$

$$p_{01} = \frac{\alpha}{\alpha + \lambda}, \quad p_{00} = \frac{\lambda}{\alpha + \lambda}, \quad p_{10} = \frac{\beta}{\lambda + \beta}, \quad p_{11} = \frac{\lambda}{\lambda + \beta}$$

$$G_{01}(x) = G_{10}(x) = 1, \quad G_{00}(x) = G_{11}(x) = G(x)$$

so that

$$Q(x) = \begin{bmatrix} -(\alpha + \lambda(1 - G(x))) & \alpha \\ \beta & -(\beta + \lambda(1 - G(x))) \end{bmatrix}, \quad (3.19)$$

and

$$R = \begin{bmatrix} 0 & \\ & -1 \end{bmatrix}.$$

This implies that $S_0 = \{0\}$ and $S_- = \{1\}$. The stability condition is given by

$$\rho = \lambda\tau < \frac{\alpha}{\alpha + \beta} = \pi_1.$$

We have

$$\tilde{Q}(s) = \begin{bmatrix} -(\alpha + \lambda(1 - \tilde{G}(s))) & \alpha \\ \beta & -(\beta + \lambda(1 - \tilde{G}(s))) \end{bmatrix}. \quad (3.20)$$

Theorem 2.3 yields the following result for the LST of $F(x) = [F_0(x), F_1(x)]$:

$$\tilde{F}(s) = sF(0)(sR - \tilde{Q}(s))^{-1}.$$

The boundary conditions yield the following explicit expression for $F(0) = [F_0(0), F_1(0)]$:

$$F_0(0) = \frac{\beta}{\alpha + \lambda} \left(\frac{\alpha}{\alpha + \beta} - \rho \right), \quad F_1(0) = \frac{\alpha}{\alpha + \beta} - \rho.$$

This can be simplified to get

$$\tilde{F}_0(s) = \frac{s\beta(\pi_1 - \rho)}{\alpha\beta - (\alpha + \lambda(1 - \tilde{G}(s)))(\beta + \lambda(1 - \tilde{G}(s)) - s)},$$

$$\tilde{F}_1(s) = \frac{s(\pi_1 - \rho)(\alpha + \lambda(1 - \tilde{G}(s)))}{\alpha\beta - (\alpha + \lambda(1 - \tilde{G}(s)))(\beta + \lambda(1 - \tilde{G}(s)) - s)}.$$

This matches the results in Gaver [3]. As a special case, when $\beta = 0$, the server is permanently up, and we get the standard queue of the first example. The results match for this special case.

3.5 An Inventory Model

Consider the following inventory model. Demand occurs continuously at rate μ . A machine produces continuously at rate $r > \mu$. It is subject to failures and repairs; the up times are $\text{Exp}(\beta)$ random variables and the down times $\text{Exp}(\alpha)$ random variables. In addition, an external supplier delivers the product according to a compound Poisson process with batch arrival rate λ and batch size distribution G with mean $\tau < \infty$. All unsatisfied demand is lost.

Let $X(t)$ be the inventory at time t and let $I(t)$ be the state of the machine at time t (1 if up, and 0 if down). Then $\{(X(t), I(t)), t \geq 0\}$ is a fluid model with jumps with $r_0 = -\mu < 0$ and $r_1 = r - \mu > 0$. Clearly, $S_- = \{0\}$ and $S_+ = \{1\}$; so in this example we have a non-empty S_+ . The parameters $q_i, p_{ij}, G_{ij}(x), i, j = 0, 1$, are the same as in Example ???. Hence the matrices $Q(x)$ and $\tilde{Q}(s)$ are as in Eq. (??) and Eq. (??) respectively. The system is stable when

$$\lambda\tau < \mu - \frac{\alpha}{\alpha + \beta}r.$$

By Theorem 2.3, $F_1(0) = 0$ from Eq.(??) and $F_0(0)$ is determined from Eq.(2.21)

$$F_0(0) = \frac{\mu - \frac{\alpha}{\alpha + \beta}r - \lambda\tau}{\mu},$$

and the LST is

$$\tilde{F}(s) = sF(0)(sR - \tilde{Q}(s))^{-1} = \frac{sF_0(0)}{\Delta}(s(r - \mu) + \beta + \lambda(1 - \tilde{G}(s)), \alpha),$$

where

$$\Delta = \det(sR - \tilde{Q}(s)) = (-s\mu + \alpha + \lambda(1 - \tilde{G}(s)))(s(r - \mu) + \beta + \lambda(1 - \tilde{G}(s))) - \alpha\beta.$$

4 Appendix

Proof of Theorem 2.2: We follow the method used in a similar proof in [1]. Note that it follows from Eqs. (2.13) and (2.14) that we can rewrite the matrix $\tilde{Q}(s)$ in the form

$$\tilde{Q}(s) = \bar{Q}(s) - Q_d,$$

where

$$\bar{Q}(s) := [q_i p_{ij} \tilde{G}_{ij}(s)], \text{ and } Q_d := \text{diag}[q_1, \dots, q_N].$$

Then Eq. (2.16) has a solution with $\phi \neq 0$ if

$$\det(\bar{Q}(s) - Q_d - sR) = 0,$$

which is equivalent to

$$\det(Q_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I) = 0.$$

We first assume that for some $\epsilon > 0$ the transformations $\tilde{G}_{ij}(s)$ are analytic for all s with $\operatorname{Re}(s) > -\epsilon$. For example, this holds for distributions with an exponential tail or distributions with finite support.

Let

$$\lambda := \max_{i:r_i \neq 0} \left(-\frac{q_i}{r_i} \right),$$

and let

$$C_\delta := \{s : |s - \lambda| = \lambda + \delta\},$$

denote the circle with center at λ and radius $\lambda + \delta$, where $0 < \delta < \epsilon$. Next, we show that for $0 \leq u \leq 1$ and small $\delta > 0$,

$$\det(uQ_d^{-1}\tilde{Q}(s) - sQ_d^{-1}R - I) \neq 0, \quad s \in C_\delta. \quad (4.21)$$

First, in the case $s \in C_\delta$ with $\operatorname{Re}(s) \geq 0$ we prove that the matrix $uQ_d^{-1}\tilde{Q}(s) - sQ_d^{-1}R - I$ is diagonally dominant (with strict dominance in at least one row). This, plus the fact that $P = [p_{ij}]$ is irreducible, imply (see [13]) that it has a non-zero determinant.

1. Let $i : r_i \neq 0$ and denote by $\lambda_i := -\frac{q_i}{r_i}$ so that $\lambda := \max_{i:r_i \neq 0} \lambda_i$. Clearly,

$$\left| up_{ii}\tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1 \right| = \left| up_{ii}\tilde{G}_{ii}(s) + \frac{s}{\lambda_i} - 1 \right| \geq \left| \frac{s}{\lambda_i} - 1 \right| - up_{ii}\tilde{G}_{ii}(0).$$

- $i : r_i < 0$ or equivalently $i : \lambda_i > 0$

Then we have

$$\left| \frac{s}{\lambda_i} - 1 \right| = \frac{|s - \lambda_i|}{\lambda_i} \geq \frac{|s - \lambda|}{\lambda_i} = \frac{|\lambda + \delta|}{\lambda_i} \geq \frac{|\lambda_i + \delta|}{\lambda_i} = 1 + \frac{\delta}{\lambda_i},$$

and therefore

$$\begin{aligned} \left| up_{ii}\tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1 \right| &\geq 1 + \frac{\delta}{\lambda_i} - up_{ii}\tilde{G}_{ii}(0) > u\tilde{G}_{ii}(0) - up_{ii}\tilde{G}_{ii}(0) = \sum_{j \neq i} up_{ij} \\ &\geq \sum_{j \neq i} |up_{ij}\tilde{G}_{ij}(s)|. \end{aligned}$$

- $i : r_i > 0$ or equivalently $i : \lambda_i < 0$

Then it is clear that for $\operatorname{Re}(s) \geq 0$ and $s \in C_\delta$

$$\left| \frac{s}{\lambda_i} - 1 \right| > 1,$$

and therefore

$$\begin{aligned} \left| up_{ii}\tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1 \right| &> 1 - up_{ii}\tilde{G}_{ii}(0) \geq u\tilde{G}_{ii}(0) - up_{ii}\tilde{G}_{ii}(0) = \sum_{j \neq i} up_{ij} \\ &\geq \sum_{j \neq i} |up_{ij}\tilde{G}_{ij}(s)|. \end{aligned}$$

2. Let $i : r_i = 0$.

Then

$$|up_{ii}\tilde{G}_{ii}(s) - 1| \geq 1 - up_{ii}\tilde{G}_{ii}(0) \geq u\tilde{G}_{ii}(0) - up_{ii}\tilde{G}_{ii}(0) = \sum_{j \neq i} up_{ij} \geq \sum_{j \neq i} |up_{ij}\tilde{G}_{ij}(s)|.$$

Note that the stability condition (2.3) implies that at least one of the rates $r_i < 0$. Thus we have strict dominance in at least one row. Further, the arguments above apply to any circle C_δ with $\delta > 0$. This implies that the determinant in (4.31) has no zeros in the right half plane outside C_0 .

Next, consider the case $s \in C_\delta$ with $Re(s) < 0$. Note that the determinant of (4.31) is nonzero if and only if 0 is not an eigenvalue. Therefore, next we study the eigenvalues of $uQ_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I$ in a neighborhood of $s = 0$. If we write

$$uQ_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I = P - I - sQ_d^{-1}R + (u - 1)Q_d^{-1}\bar{Q}(s) + Q_d^{-1}\bar{Q}(s) - P,$$

we can see that the above matrix is a perturbation of $P - I$ when (s, u) is close to $(0, 1)$. Since P is assumed to be irreducible, $P - I$ has a simple eigenvalue 0. Thus, in a neighborhood of $(0, 1)$, there exist differentiable $x(s, u)$ and $\mu(s, u)$ such that

$$(uQ_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I)x(s, u) = \mu(s, u)x(s, u), \quad \mu(0, 1) = 0, \quad x(0, 1) = e$$

Differentiating the last equation with respect to s and setting $s = 0$, $u = 1$ in the result gives the following:

$$\left(Q_d^{-1} \frac{d\bar{Q}(0)}{ds} - Q_d^{-1}R \right) e + (P - I) \frac{\partial}{\partial s} x(0, 1) = \frac{\partial}{\partial s} \mu(0, 1) e,$$

After pre-multiplying this equation by πQ_d , where $\pi := [\pi_1, \dots, \pi_N]$, $\pi_i = \lim_{t \rightarrow \infty} P(I(t) = i)$ we get

$$\pi \left(\frac{d\bar{Q}(0)}{ds} - R \right) e + \pi Q_d (P - I) \frac{\partial}{\partial s} x(0, 1) = \frac{\partial}{\partial s} \mu(0, 1) \pi Q_d e.$$

Note that π is the solution to

$$\pi Q_d (P - I) = 0,$$

and also from the definition of $\bar{Q}(s)$ above we have

$$\frac{d\bar{Q}(s)}{ds} = \left[q_i p_{ij} \frac{d\tilde{G}_{ij}(s)}{ds} \right],$$

and therefore

$$\frac{d\bar{Q}(0)}{ds} = [-q_i p_{ij} m_{ij}] = -Q_d P \cdot M,$$

where $P \cdot M$ denotes the Hadamard matrix multiplication. Hence we have

$$\pi(-Q_d P \cdot M - R)e = \frac{\partial}{\partial s} \mu(0, 1) \pi Q_d e. \quad (4.22)$$

In a similar way differentiation with respect to u leads to the following:

$$\pi Q_d P e = \frac{\partial}{\partial u} \mu(0, 1) \pi Q_d e,$$

and since

$$\pi Q_d P e = \sum_{i=1}^N \sum_{j=1}^N \pi_i q_i p_{ij} = \sum_{j=1}^N \pi_j q_j = \pi Q_d e,$$

we have

$$\frac{\partial}{\partial u} \mu(0, 1) = 1.$$

Therefore, by Taylor series expansion of $\mu(s, u)$ in a neighbourhood of $(0, 1)$ from the last equation and Eq. (4.42) it follows

$$\mu(s, u) = s \frac{\pi(-Q_d P \cdot M - R)e}{\pi Q_d e} + u - 1 + o(\sqrt{s^2 + (u - 1)^2}).$$

From the stability condition (2.3) we have

$$\pi(-Q_d P \cdot M - R)e > 0,$$

so we can conclude that $\mu(s, u) \neq 0$ for small δ , $s \in C_\delta$ with $\text{Re}(s) < 0$, and u close to 1, say $1 - \hat{\delta} \leq u \leq 1$. For $0 \leq u < 1 - \hat{\delta}$, it can be shown that $u Q_d^{-1} \tilde{Q}(s) - s Q_d^{-1} R - I$ is diagonally dominant (with strict dominance in at least one row) for $s \in C_\delta$ with $\text{Re}(s) < 0$, provided that δ is taken small enough so that

$$1 - \frac{\delta}{\lambda_i} > (1 - \hat{\delta}) \max_{i,j \in S} \tilde{G}_{ij}(-\delta).$$

For $i : r_i < 0$ we have as before

$$\begin{aligned} |up_{ii} \tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| &\geq 1 + \frac{\delta}{\lambda_i} - up_{ii} \tilde{G}_{ii}(s) > 1 - \frac{\delta}{\lambda_i} - up_{ii} \tilde{G}_{ii}(s) > \\ &> (1 - \hat{\delta}) \max_{i,j \in S} \tilde{G}_{ij}(-\delta) - up_{ii} \tilde{G}_{ii}(s) > \\ &> u \max_{i,j \in S} \tilde{G}_{ij}(-\delta) - up_{ii} \max_{i,j \in S} \tilde{G}_{ij}(-\delta) \geq \sum_{j \neq i} |up_{ij} \tilde{G}_{ij}(s)|. \end{aligned}$$

For $i : r_i \geq 0$ it is obvious that

$$|up_{ii} \tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| \geq 1 - up_{ii} \tilde{G}_{ii}(s) > 1 - \frac{\delta}{\lambda_i} - up_{ii} \tilde{G}_{ii}(s),$$

and the rest of the argument is the same as above. This completes the proof of (4.31).

Let $f(u)$ denote the number of zeros of $\det(uQ_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I)$ inside the circle C_δ . Then from the Cauchy Theorem of Complex Analysis we have

$$f(u) = \frac{1}{2\pi i} \int_{C_\delta} \frac{\frac{\partial}{\partial s} \det(uQ_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I)}{\det(uQ_d^{-1}\bar{Q}(s) - sQ_d^{-1}R - I)} ds.$$

It is clear that $f(0) = N_-$, since

$$\det(-sQ_d^{-1}R - I) = - \prod_{i=1}^N \left(\frac{sr_i}{q_i} + 1 \right),$$

and therefore the zeros are $s_i = -\frac{q_i}{r_i}$, $r_i \neq 0$ and exactly N_- of them have $\text{Re}(s) \geq 0$ and thus are inside the circle C_δ . Since $f(u)$ is an integer-valued continuous function on $[0, 1]$ it follows that it is constant. Hence, $f(1) = N_-$. As $\delta \rightarrow 0$, we can conclude that $\det(sR - \tilde{Q}(s))$ has exactly N_- zeros inside or on C_0 .

It is clear that $s = 0$ satisfies

$$\det(sR - \tilde{Q}(s)) = 0.$$

It can be shown that $s = 0$ is a simple root of this equation using the irreducibility of P and the stability condition (2.3). The arguments are the same as in [1] and therefore we skip them.

To finally complete the proof of Theorem 2.2 we have to remove the initial assumption that for some $\epsilon > 0$ the transforms $\tilde{G}_{ij}(s)$ are analytic for all s with $\text{Re}(s) > -\epsilon$. To this end, first consider the truncated distributions $G_{ij}^K(x)$ defined as $G_{ij}^K(x) = G_{ij}(x)$ for $0 \leq x < K$ and $G_{ij}^K(x) = 1$ for $x \geq K$. Then Theorem 2.2 holds for the distributions $G_{ij}^K(x)$; by letting K tend to infinity, the result also follows for the original distributions.

References

- [1] I.J.B.F. ADAN AND V. KULKARNI (2003) Single server queue with Markov dependent inter-arrival and service times. *Queueing Systems* **45**, 113-134.
- [2] D. ANICK, D. MITRA AND M.M. SONDHI (1982) Stochastic theory of a data handling system with multiple sources. *Bell Syst. Tech. J.* **61**, 1871-1984.
- [3] H.R. GAIL, S.L. HANTLER, B.A. TAYLOR (1992) On a preemptive Markovian queue with multiple servers and two priority classes. *Math. Oper. Res.* **17**, 365-391.
- [4] D.P. GAVER (1962) A waiting line with interrupted service, including priorities. *J. R. Statist. Soc., Series B*, **24(1)**, 73-90.

- [5] L. KOSTEN (1974) Stochastic theory of a multi-entry buffer, part 1. *Delft Progress Report*, 10-18.
- [6] V.G. KULKARNI (1997) Fluid models for single buffer systems. *Frontiers in Queueing; Models and Applications in Science and Engineering*, 321-338.
- [7] R.M. LOYNES (1962) The stability of a queue with non-independent inter-arrival and service times. *Proc. Camb. Philos. Soc.* **58**, 497–520.
- [8] R.M. LOYNES (1962) Stationary waiting time distributions for single-server queues. *Ann. Math. Statist.* **33**, 1323-1339.
- [9] D.M. LUCANTONI (1991) New results on the single server queue with a batch Markovian arrival process. *Stochastic Models* **7(1)**, 1-46.
- [10] S.P. MEYN, R. TWEEDIE (1994) Markov Chains and Stochastic Stability. *Springer-Verlag, New York*.
- [11] D. MITRA (1988) Stochastic fluid models. *Performance* **87**, Elsevier Science Publishers, North Holland, 39-51.
- [12] D. MITRA (1988) Stochastic theory of a fluid model of producers and consumers coupled by a buffer. *Adv. Appl. Prob.* **20**, 646-676.
- [13] M. MIYAZAWA, H. TAKADA (2002) A matrix exponential form for hitting probabilities and its application to a Markov-modulated fluid queue with downward jumps. *J. Appl. Prob.* **39**, 604-618.
- [14] M. NEUTS (1989) Structured stochastic matrices of $M|G|1$ type and their applications. *Marcel Dekker, New York*.
- [15] M. NEUTS (1981) Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach. *Dover Publications, Inc., New York*.
- [16] B. SENGUPTA (1989) Markov processes whose steady state distribution is matrix-exponential with an application to the $GI/PH/1$ queue. *Adv. Appl. Prob.* **21(1)**, 159-180.
- [17] JOS H. A. DE SMIT (1983) The queue $GI/M/s$ with customers of different types or the queue $GI/H_m/s$. *Adv. Appl. Prob.* **15**, 392-419.
- [18] H. TAKADA (2001) Markov modulated fluid queues with batch fluid arrivals. *J. Operat. Res. Soc. Japan* **44**, 344-365.
- [19] M. MIYAZAWA, H. TAKADA (2002) A matrix exponential form for hitting probabilities and its application to a Markov-modulated fluid queue with downward jumps. *J. Appl. Prob.* **39**, 604-618.