Output Analysis of Multiclass Fluid Models with Static Priorities

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

Abstract

We consider a stochastic fluid flow model with a single server and K infinite capacity buffers. The input to the k-th buffer is a Markovian on-off process that transmits fluid at a constant rate p_k while it is on and at rate 0 while it is off. The fluid is emptied from the buffers by a single server at a constant rate μ according to a *static priority service discipline* in which class 1 fluid has the highest and class K fluid has the lowest priority. The output process of class k is defined to be *on* if fluid of class k is leaving the buffer at a positive rate and *off* otherwise. In this paper we derive an exact method for computing the mean on-time and the mean off-time of the output process of class k. We illustrate the method by a numerical example.

Keywords: stochastic fluid model, static priority service discipline, output process

1 Introduction

We study a fluid-flow model with a single server and incoming fluid flows generated by K independent on-off Markovian input sources. The k-th source has $\exp(\alpha_k)$ on-times and $\exp(\beta_k)$ off-times, $k = 1, \ldots, K$. It generates fluid at rate p_k while it is on and at rate 0 while it is off. Thus, the input process of class k is completely described by three parameters (α_k, β_k, p_k). The fluid generated by the K on-off sources is stored in K separate infinite capacity buffers from where it is removed according to a *static priority* service discipline under which the highest priority fluid always takes precedence over any of the lower priority fluids. We assume that class k enjoys complete priority over fluid of class $j > k$, at all times. Thus class 1 fluid has the highest priority and class K has the lowest priority. The leftover service capacity after serving the fluids of classes $1, 2, \ldots, k$ is available to serve fluids of class $k + 1$ and above.

[*]EURANDOM, Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands; email: tzenova@eurandom.tue.nl

[†]Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands; email: iadan@win.tue.nl

[‡]Department of Statistics and Operations Research, University of North Carolina, CB 3180, Cahpel Hill, N.C. 27599; email: vkulkarn@email.unc.edu

The output process under the static priority service discipline is rather complex. Several input bursts of a given class may combine in one output burst. Similarly, a single input burst may get split into several output bursts due to interruptions by higher priority fluid coming to the buffer. The output processes of different classes of fluid are neither independent, nor on-off. The rate during an output burst is not constant due to the variations in the leftover capacity available for that class. The idea is to approximate the class k output process by a three-parameter Markovian on-off process with parameters $(\alpha_k^o, \beta_k^o, p_k^o)$. The output process of class k is defined to be *on* if fluid of class k is leaving the buffer at a positive rate and *off* otherwise. Thus we need to find the mean on- and off-times, $1/\alpha_k^o$ and $1/\beta_k^o$, and we approximate the non-constant output rate of class k by the mean peak rate p_k^o. Kulkarni and Glazebrook [5] study the output process of a single-buffer multi-class queue with First-Come-First-Serve (FCFS) discipline, and determine the mean on- and off-times of a given class by constructing an appropriate reward processes. However, this approach does not readily extend to the case of static priority service discipline.

The motivation behind this analysis arises from the study of telecommunication networks as multi-class fluid networks (MFN) described by a set of nodes and a set of different classes of fluid. Each node has an infinite capacity buffer for storing each class of fluid that enters this node. The input of each node may consist of fluid generated by an external environment process and/or fluid coming from other nodes within the network. The different classes of fluid in a given node are served according to a predefined service discipline, in our case static priority. The external fluid inputs to each node in the network are assumed to be generated by independent Markovian on-off sources. We approximate the output processes from a given node as three-parameter independent on-off processes by using the approach developed here. The analysis of the network can then proceed iteratively by using these approximated output on-off processes as inputs to other nodes in the same spirit as in Whitt [7]. Thus, each node acts as a non-linear mapping of the input parameters to the output parameters. Hirasawa [1] adopts a similar approach to study an MFN with FCFS discipline and develops an algorithm based on the parametric-decomposition method of Kuehn [2]. Hirasawa characterizes the network traffic in terms of four parameters - mean rate, effective peak rate, mean burst length, and mean squared burst length. In this paper we develop the exact transformation of $\{(\alpha_k, \beta_k, p_k), k = 1, 2, ..., K\}$ into $\{(\alpha_k^o, \beta_k^o, p_k^o), k = 1, 2, ..., K\}$ for the single-node with multi-class, static priority traffic. The calculations involved are quite intricate, and can be used to test approximations that will be needed to do the network analysis. However, the main contribution of this paper is the exact transformation, and not the network analysis. That will be a topic for future research.

2 Problem Description

Denote the state of source k at time t by $I_k(t)$, where

$$I_k(t) = \begin{cases} 0, & \text{if source } k \text{ is off at time } t, \\ 1, & \text{if source } k \text{ is on at time } t. \end{cases}$$

The combined state of the K sources at time t (called the environment) is given by

$$I(t) = (I_1(t), \ldots, I_K(t)),\ t \geq 0.$$

Thus, $\{I(t),\ t \geq 0\}$ is an irreducible CTMC on the finite state space

$$S = \{i = (i_1, \ldots, i_K) : i_k = 0, 1,\ k = 1, \ldots, K\},$$

with rate matrix $Q = [q_{ij}]$. Define the combined input rate $p(i)$ of all sources in state $i = (i_1, \ldots, i_K) \in S$ as

$$p(i) = \sum_{k=1}^{K} i_k p_k.$$

The server operates at a constant rate μ, called its capacity. The net input rate $r(i)$ to the buffer, when the environment is in state $i = (i_1, \ldots, i_K) \in S$, is given by

$$r(i) = \sum_{k=1}^{K} i_k p_k - \mu = p(i) - \mu$$

and let $R = \text{diag}[r(i),\ i \in S]$ denote the net input rate matrix. Let $X_k(t), k = 1, \ldots, K$, denote the amount of fluid of class k in the buffer at time t and $X(t) = \sum_{k=1}^{K} X_k(t)$ be the total amount of fluid in the buffer at time t. Then $\{(I(t), X(t)), t \geq 0\}$ is a Markov process. The rate of change of the fluid level in the buffer $\{X(t), t \geq 0\}$ is given by

$$\frac{d}{dt}X(t) = \begin{cases} r(i) & \text{if } I(t) = i,\ X(t) > 0, \\ \\ \max(r(i), 0) & \text{if } I(t) = i,\ X(t) = 0. \end{cases}$$

Now let

$$\pi(i) = \lim_{t \to \infty} P(I(t) = i),\ i \in S$$

be the limiting distribution of the governing CTMC $\{I(t), t \geq 0\}$. The system is stable if and only if the expected net input rate is negative in steady state, i.e.,

$$\sum_{i \in S} \pi(i) r(i) < 0. \tag{2.1}$$

Henceforth, we assume that the system is stable, so that the limiting distribution of the bivariate process $\{(I(t), X(t)), t \geq 0\}$ exists. We denote it by

$$\pi(i, x) = \lim_{t \to \infty} P(I(t) = i, X(t) \leq x),\ i \in S, \tag{2.2}$$

see, e.g., Kulkarni [4] for methods of computing $\pi(i, x)$.

In this paper we study the output process of class k. Recall that the output process of class k is defined to be *on* if fluid of class k is leaving the buffer at a positive rate and *off* otherwise. Let

$$S_k(t) = \begin{cases} 1, & \text{if the output process of class } k \text{ is on at time } t, \\ 0, & \text{if the output process of class } k \text{ is off at time } t. \end{cases} \tag{2.3}$$

Let $B_k,\ C_k$ denote the mean sojourn times of the S_k process in states $1,\ 0$, respectively, i.e., B_k is the mean on-time, and C_k is the mean off-time of the output process of class k. In this paper we derive methods of computing B_k and C_k. The output process of class k can then be approximated by a Markovian on-off process with $\exp(\alpha_k^o)$ on-times and $\exp(\beta_k^o)$ off-times, where $\alpha_k^o = 1/B_k$ and $\beta_k^o = 1/C_k$, and constant transmission rate p_k^o derived from the conservation of fluid principle, i.e., the mean output rate of class k equals the mean input rate of class k, if the system is stable. This yields

$$\frac{p_k^o \beta_k^o}{\alpha_k^o + \beta_k^o} = \frac{p_k \beta_k}{\alpha_k + \beta_k},$$

and therefore

$$p_k^o = \frac{p_k \beta_k}{\alpha_k + \beta_k} \frac{(\alpha_k^o + \beta_k^o)}{\beta_k^o},\ \ k = 1, 2, \ldots, K.$$

For each priority class we can act as if the lower priority classes do not exist at all, and hence, it suffices to show how the mean on- and off-times of the output process of the lowest priority class K can be determined. To do so, we construct from the *original* K-class system a new 2-class *aggregated* system as follows. Class 1 fluid input process in the aggregated system is identical to the superposition of the input processes of the first $K-1$ on-off sources in the original system. Class 2 fluid input process in the aggregated system is identical to the input process of the K-th on-off source in the original system. Thus, the class 1 input in the aggregated system is modulated by a CTMC $\{J_1(t) = (I_1(t), \ldots, I_{K-1}(t)), t \geq 0\}$ that can be in 2^{K-1} different states $\{(i_1, \ldots, i_{K-1}), i_k = 0, 1,\ k = 1, \ldots, K-1\}$ and the class 2 input is modulated by the two state CTMC $\{J_2(t) = I_K(t), t \geq 0\}$. Thus $J(t) = (J_1(t), J_2(t)) = I(t)$. Let $Y_k(t)$, $k = 1, 2$, be the amount of class k fluid in this aggregated 2-class sysstem. Clearly

$$Y_1(t) = \sum_{k=1}^{K-1} X_k(t),\ \ Y_2(t) = X_K(t),\ \ t \geq 0.$$

We continue with the analysis of this two-source priority model. It is convenient to introduce the input rates

$$p_1(i) := \sum_{k=1}^{K-1} i_k p_k,\ \text{and}\ p_2(i) := i_K p_K,\ i \in S,$$

and the net input rate of class 1 fluid to the buffer

$$r_1(i) := p_1(i) - \mu,$$

when the environment is in state $i = (i_1, \ldots, i_K) \in S$. Then $r_1(i)$ determines the following partitioning of the state space S,

$$\begin{aligned} S_- &:= \{i \in S : r_1(i) < 0\},\ N_- := |S_-|, \\ S_+ &:= \{i \in S : r_1(i) \geq 0\},\ N_+ := |S_+|. \end{aligned}$$

First, we consider the easier case of $S_+ = \emptyset$ for which the solution can be found directly, by applying the theory of Alternating Renewal Processes. Then we study the more complicated case of $S_+ \neq \emptyset$.

3 Output analysis if S_+ is empty

It is clear that S_+ is empty if and only if

$$\sum_{k=1}^{K-1} p_k < \mu.$$

Under this condition there is always some leftover capacity of the server at which the lowest priority class of fluid K is served. In other words, the moment source K turns on, fluid of class K starts immediately leaving the buffer with a rate that depends on the state of the environment at that moment. Consider the output process S_K of class K as defined in Equation (2.3) and assume that $S_K(0) = 0$. Then, the output process stays off for an $\exp(\beta_K)$ and then it turns on as soon as source K turns on. Then it stays on for some random amount of time (as long as there is class K fluid passing through the buffer) which depends on the state of the environment, and thus depends on the off time of the S_K process. Therefore, the S_K process is an alternating renewal process and we obtain

$$\lim_{t\to\infty} P(S_K(t) = 0) = \frac{\frac{1}{\beta_K}}{\frac{1}{\beta_K} + B_K},$$

where B_K is the mean on-time of the output process of class K. The left-hand side of this equation can be computed as

$$\lim_{t\to\infty} P(S_K(t) = 0) = \sum_{i:i_K=0} \pi(i, 0),$$

where $\pi(i, 0)$ is defined in Equation (2.2). Hence we obtain

$$B_K = \frac{1 - \sum_{i:i_K=0} \pi(i, 0)}{\beta_K \sum_{i:i_K=0} \pi(i, 0)}.$$

Clearly we also have

$$C_K = \frac{1}{\beta_K}.$$

Note that these calculations simplify even more if we assume that

$$\sum_{k=1}^{K} p_k < \mu.$$

Then the output processes of all the classes are identical to the input processes, and thus

$$B_k = \frac{1}{\alpha_k}, \quad C_k = \frac{1}{\beta_k},$$

for all $k = 1, 2, ..., K$.

4 Output Analysis if S_+ is non-empty

In this section we assume that S_+ is non-empty, i.e.,

$$\sum_{k=1}^{K-1} p_k \geq \mu.$$

Recall that we consider the 2-class aggregated system in which the fluid is generated by only two independent input sources (with the second source being on-off and the first one being in any of the 2^{K-1} possible states as described above). Let $\gamma_k,\ k = 1, 2$, be the long-run fraction of time class k fluid is not being served, i.e., is not leaving the buffer. Denote by ν_k the long-run rate at which class k output off-periods are generated. Then from the classic theory of Markov-Regenerative processes it follows that

$$B_k + C_k = \frac{1}{\nu_k},\ \ k = 1, 2,$$

and

$$\gamma_k = \frac{C_k}{C_k + B_k},\ \ \ 1 - \gamma_k = \frac{B_k}{C_k + B_k},\ \ k = 1, 2.$$

Hence

$$C_k = \frac{\gamma_k}{\nu_k},$$

and

$$B_k = \frac{1 - \gamma_k}{\nu_k}.$$

Thus, to find B_k and C_k we need to determine γ_k and ν_k. We calculate them first for $k = 1$, and then for $k = 2$. Let, for $i \in S$ and $x \geq 0$,

$$\pi_1(i, x) = \lim_{t \to \infty} P(J(t) = i, Y_1(t) \leq x),$$

which can be computed in the same manner as $\pi(i, x)$. Then we have the following result, the proof of which is straightforward and hence is omitted.

Theorem 4.1

$$\gamma_1 = \sum_{i:p_1(i)=0} \pi_1(i, 0)$$

and

$$\nu_1 = \sum_{i:p_1(i)=0} \pi_1(i, 0) \sum_{j:p_1(j)>0} q_{ij}.$$

Next we study the buffer content process $Y_2 = \{Y_2(t), t \geq 0\}$ of the class 2 fluid. We say that this process is in "active" phase at time t if $J(t) \in S_-$ and $Y_1(t) = 0$. Thus the low priority fluid can be served in the active phase. If the Y_2 process is not in active phase we say it is in "inactive" phase. Under the assumption of this section, $S_+ \neq \emptyset$, there are inactive-periods alternating with active-periods for class 2 fluid. We evaluate γ_2 and ν_2 by studying the Y_2 process restricted to the active-periods. Let $\tau(t)$ the time spent by the Y_2 process in the active-periods over $[0,t]$ and define the restricted processes $\{Y_2^{ac}(t),\ t \geq 0\} := \{Y_2(\tau(t)+),\ t \geq 0\}$ and $\{J^{ac}(t),\ t \geq 0\} := \{J(\tau(t)+),\ t \geq 0\}$. Clearly, $J^{ac}(t) \in S_-,$ for all $t \geq 0$. A typical sample path of the Y_2 and Y_2^{ac} processes are shown in Figure 4.1. Note that the restricted Y_2^{ac} process is a fluid process with upward jumps.

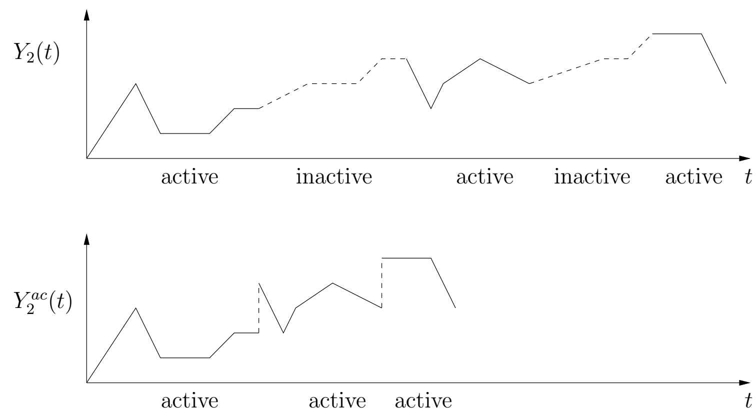

Figure 4.1: Sample paths of the Y_2 and Y_2^{ac} process; the solid lines correspond to active-periods and the dashed lines to inactive-periods.

Let

$$\pi_2^{ac}(t,i,x) := P(J^{ac}(t) = i,\ Y_2^{ac}(t) \leq x), \quad t \geq 0,\ x \geq 0, i \in S_-,$$

$$\pi_2^{ac}(i,x) := \lim_{t\to\infty} \pi_2^{ac}(t,i,x), \quad x \geq 0, i \in S_-, \qquad \pi_2^{ac}(x) := [\pi_2^{ac}(i,x),\ i \in S_-],$$

$$T := \inf\{t \geq 0 :\ J(t) \in S_- \text{ and } Y_1(t) = 0\}.$$

Thus T represents the inactive-period. Let $A_2(T)$ be the total amount of class 2 fluid that comes in the buffer during $[0,T]$. Furthermore, let

$$g(j,x) := P(A_2(T) > 0 | J(0) = j, Y_1(0) = x),\ j \in S,\ x \geq 0,$$

and recall that $Q = [q_{ij}]$ is the generator matrix of the $J\ (= I)$ process. The computation of γ_2 and ν_2 uses $\pi_2^{ac}(i,0), \pi_2^{ac}(i,\infty), i \in S_-$, and $g(j,0), j \in S_+$ as shown in the next theorem.

Theorem 4.2 *Let* $\gamma = \sum_{i \in S_-} \pi_1(i,0)$ *, and*

$$\begin{aligned} A \;:=\; & \sum_{i\in S_-:p_2(i)=0} \pi_2^{ac}(i,0) \sum_{j\in S_-:p_2(j)>0} q_{ij} + \sum_{i\in S_-:p_2(i)>0} \pi_2^{ac}(i,\infty) \sum_{j\in S_+} q_{ij} + \\ & + \sum_{i\in S_-:p_2(i)=0} (\pi_2^{ac}(i,\infty) - \pi_2^{ac}(i,0)) \sum_{j\in S_+} q_{ij} + \sum_{i\in S_-:p_2(i)=0} \pi_2^{ac}(i,0) \sum_{j\in S_+:p_2(j)=0} q_{ij} g(j,0). \end{aligned}$$

Then

$$\gamma_2 = \gamma \sum_{i \in S_-: p_2(i)=0} \pi_2^{ac}(i,0) + 1 - \gamma,$$

and

$$\nu_2 = \gamma A.$$

Remark: γ is the long-run fraction of time the Y_2 process is in an active-period, and A is the long-run rate at which class 2 output on-periods (and hence off-periods) are generated given Y_2 is in an active-period.

Proof: During periods of time with $Y_1(t) > 0$ or $Y_1(t) \geq 0$ and $I(t) \in S_+$ (referred to as inactive-periods) class 2 is not served, since there is no leftover service capacity available. The long-run fraction of time spent in inactive-periods is given by $1 - \gamma$. Class 2 is also not served during active-periods, i.e., during periods of time $Y_1(t) = 0$ and $J(t) \in S_-$, when there is no class 2 fluid in the buffer and there is no inflow of class 2. Hence

$$\gamma_2 = 1 - \gamma + \gamma \sum_{i \in S_-: p_2(i)=0} \pi_2^{ac}(i,0).$$

To derive the expression for ν_2 we compute the rate at which output on-periods are generated; since on- and off-periods alternate, this is the same as the rate at which output off-periods are generated. Note that output of class 2 is only possible while the Y_2 process is in an active-period, i.e., there is leftover service capacity to serve class 2. The long-run fraction of time the Y_2 process is in active-periods is γ. Given that the Y_2 process is in an active-period, the output on-periods of class 2 can start somewhere within an active-period or right at the beginning of an active-period. The first term of A accounts for output on-periods of class 2 that start within an active-period; this is only possible due to jumps from a state $i \in S_-$ in which $Y_2 = 0$ and $p_2(i) = 0$ to a state $j \in S_-$ with $p_2(j) > 0$. The other three terms in A account for output on-periods that start right at the beginning of an active-period. An inactive-period always starts by a jump from a state $i \in S_-$ into a state $j \in S_+$; note that such jumps correspond to state changes in the environment process J_1, so the input rate of class 2 fluid remains unaltered, i.e., $p_2(j) = p_2(i)$. Then, at the beginning of the next active-period, an output on-period starts if the amount of class 2 fluid in the buffer is positive. This is the case when (i) the state j at the start of the inactive-period is such that $p_2(j) > 0$; (ii) the state j is such that $p_2(j) = 0$, but the amount of class 2 fluid is already positive at the start of the inactive-period; or (iii) the state j is such that $p_2(j) = 0$, there is no class 2 fluid in the buffer at the start of the inactive-period, but during the inactive-period a positive amount of class 2 fluid will be added to the buffer. Thus, we obtain the expression for the long-run rate at which class 2 output on periods are generated, $\nu_2 = \gamma A$. $\diamondsuit$

The remainder of this section shows how to compute $\pi_2^{ac}(i,0), \pi_2^{ac}(i,\infty), i \in S_-$, and $g(j,0), j \in S_+$. Let for $x \geq 0$, $Re(s) \geq 0$,

$$\tilde{\psi}_{ji}(x,s) := E(e^{-sA_2(T)}; J(T) = i | J(0) = j, Y_1(0) = x), \quad j \in S, i \in S_-, \tag{4.4}$$

It is easy to see from Figure 4.1 that $\{Y_2^{ac}(t),\ t \geq 0\}$ is a *fluid process with jumps*, as analyzed in Tzenova et al. [6], where the jump sizes correspond to the total amount $A_2(T)$ of class 2 fluid accumulated in the buffer during the skipped off-periods. Clearly, the Laplace-Stieltjes Transform (LST) of the jump sizes is given by $\tilde{\psi}_{ji}(0,s)$. For the purposes of the following results we define the column vector

$$\tilde{\psi}_i(x,s) := [\tilde{\psi}_{ji}(x,s), j \in S], \quad i \in S_-$$

and the (net) input matrices for class 1 and 2 fluid,

$$R_1 := \text{diag}[r_1(i),\ i \in S], \qquad P_2 := \text{diag}[p_2(i), i \in S].$$

Let $\lambda_k(s),\ \phi_k(s) = [\phi_{jk}(s), j \in S]$ denote the eigenvalues and eigenvectors to the system $(\lambda_k(s)R_1 + Q - sP_2)\phi_k(s) = 0$. It is assumed that all eigenvalues $\lambda_k(s)$ are distinct. Under stability condition (2.1) it is known that (see, e.g., Kulkarni [3])

$$|\{k : Re\{\lambda_k(s)\} \leq 0\}| = |S_-|.$$

Lemma 4.3 *The column vector $\tilde{\psi}_i(x,s)$ satisfies the system of differential equations*

$$R_1 \frac{\partial \tilde{\psi}_i}{\partial x}(x,s) + (Q - sP_2)\tilde{\psi}_i(x,s) = 0 \tag{4.5}$$

with boundary conditions

$$\tilde{\psi}_{ji}(0,s) = \delta_{ji}, \quad j,i \in S_-, \tag{4.6}$$

where δ_{ji} is the Kronecker symbol. The solution to (4.5) and (4.6) is given by

$$\tilde{\psi}_i(x,s) = \sum_{k:Re\{\lambda_k(s)\}\leq 0} a_k^i \phi_k(s) e^{\lambda_k(s)x},$$

where the coefficients a_k are determined as the solution to the linear system

$$\tilde{\psi}_{ji}(0,s) = \sum_{k:Re\{\lambda_k(s)\}\leq 0} a_k^i \phi_{jk}(s) = \delta_{ji}, \quad j,i \in S_-.$$

Proof: Let $x > 0,\ j \in S$ and $i \in S_-$. After conditioning on a small time interval of length $h > 0$ we have

$$\tilde{\psi}_{ji}(x,s) = \sum_{k \neq j} q_{jk} h E(e^{-s(A_2(T)+p_2(j)h)}; J(T) = i | J(0) = k,\ Y_1(0) = x + r_1(j)h)$$
$$+(1 + q_{jj}h + o(h))E(e^{-s(A_2(T)+p_2(j)h)}; J(T) = i | J(0) = j,\ Y_1(0) = x + r_1(j)h) + o(h).$$

Rearranging the last equation we get

$$\frac{e^{sp_2(j)h}\tilde{\psi}_{ji}(x,s) - \tilde{\psi}_{ji}(x + r_1(j)h, s)}{h} = \sum_{k\in S} q_{jk}\tilde{\psi}_{ki}(x + r_1(j)h, s) + o(1).$$

Next, we substitute $e^{sp_2(j)h} = 1 + sp_2(j)h + o(h)$ and let h tend to 0 to get

$$-r_1(j)\frac{\partial \tilde{\psi}_{ji}}{\partial x}(x,s) + sp_2(j)\tilde{\psi}_{ji}(x,s) = \sum_{k\in S} q_{jk}\tilde{\psi}_{ki}(x,s).$$

In matrix notation this equation is equivalent to (4.5). The boundary conditions (4.6) follow from the definition of $A_2(T)$. Given $Y_1(0) = 0$ and $J(0) \in S_-$ it is clear that the length of the inactive-period is 0 and therefore $A_2(T) = 0$. The solution to (4.5) with boundary conditions (4.6) follows by well-known results from the classical theory of linear differential equations. $\diamondsuit$

Using the quantities $\tilde{\psi}_{ji}(x,s)$, we can model the Y_2^{ac} process as a fluid process with jumps as in Tzenova et al. [6] as follows: the state space of Y_2^{ac} is S_-, and the LST of the transition matrix representing the jumps is

$$\tilde{Q}_{ij}^{ac}(s) = q_{ij} + \sum_{k\in S_+} q_{ik}\tilde{\psi}_{kj}(0,s), \quad i,j \in S_-. \tag{4.7}$$

Note that the rate matrix of the environment process J^{ac} is given by $\tilde{Q}^{ac}(0)$ and that $\pi_2^{ac}(\infty)$ is the limiting distribution of the J^{ac} process, i.e., $\pi_2^{ac}(\infty)$ is the unique solution to

$$\pi_2^{ac}(\infty)\tilde{Q}^{ac}(0) = 0, \quad \pi_2^{ac}(\infty)e = 1.$$

For a given matrix M and subsets of indices A, B, we denote the sub-matrix

$$M_{A,B} := [M_{ij}, i \in A, j \in B].$$

Let s_k, ϕ_k^{ac} denote the eigenvalues and eigenvectors to the system $(s_k R_{S_-,S_-} - \tilde{Q}^{ac}(s_k))\phi_k = 0$, and assume that all eigenvalues s_k are distinct. It is known that (see, e.g., Kulkarni [3])

$$|\{k : Re\{s_k\} > 0\}| = |\{i \in S_- : p_1(i) + p_2(i) < \mu\}|.$$

Let $\tilde{\pi}_2^{ac}(i,s)$ denote the LST of $\pi_2^{ac}(i,x)$, defined as usual

$$\tilde{\pi}_2^{ac}(i,s) := \int_{0-}^{\infty} e^{-sx} d\pi_2^{ac}(i,x), \quad i \in S_-, Re(s) \geq 0.$$

Using this notation, Theorem 2.3 of Tzenova et al. [6] yields the following:

Remark: By (4.7) the quantities $\tilde{Q}_{ij}^{ac}(\infty)$ and $\frac{d}{ds}\tilde{Q}_{ij}^{ac}(0)$ appearing in Theorem 4.4 satisfy

$$\begin{aligned}\tilde{Q}_{ij}^{ac}(\infty) &= q_{ij} + \sum_{k\in S_+} q_{ik}\tilde{\psi}_{kj}(0,\infty),\\ \frac{d}{ds}\tilde{Q}_{ij}^{ac}(0) &= \sum_{k\in S_+} q_{ik}\frac{\partial \tilde{\psi}_{kj}}{\partial s}(0,0),\end{aligned}$$

where $\tilde{\psi}_{kj}(x,\infty)$ and $\frac{\partial \tilde{\psi}_{kj}}{\partial s}(x,0)$ can be interpreted as

$$\begin{aligned}\tilde{\psi}_{kj}(x,\infty) &= P(A_2(T)=0, J(T)=j|J(0)=k, Y_1(0)=x),\\ \frac{\partial \tilde{\psi}_{kj}}{\partial s}(x,0) &= E(A_2(T); J(T)=j|J(0)=k, Y_1(0)=x).\end{aligned}$$

Both $\tilde{\psi}_{kj}(x,\infty)$ and $\frac{\partial \tilde{\psi}_{kj}}{\partial s}(x,0)$ can be directly obtained as a solution to a system of linear ordinary differential equations similar to the equations (4.8)-(4.9) in Lemma 4.5 below.

Theorem 4.4 *The row vector of LSTs $\tilde{\pi}_2^{ac}(s) = [\tilde{\pi}_2^{ac}(i,s), i \in S_-]$ satisfies*

$$\tilde{\pi}_2^{ac}(s)(sR_{S_-,S_-} - \tilde{Q}^{ac}(s)) = s\pi_2^{ac}(0)R_{S_-,S_-},$$

where the unknowns $\pi_2^{ac}(0) = [\pi_2^{ac}(i,0), i \in S_-]$ are given by the solution to:

$$\pi_2^{ac}(i,0) = 0, \quad i \in S_- : p_1(i) + p_2(i) > \mu,$$

$$\begin{aligned}\pi_2^{ac}(0)\tilde{Q}^{ac}_{\cdot,i}(\infty) = 0, &\quad i \in S_- : p_1(i) + p_2(i) = \mu,\\ \pi_2^{ac}(0)R_{S_-,S_-}\phi_k^{ac} = 0, &\quad k : Re(s_k) > 0\\ \pi_2^{ac}(0)R_{S_-,S_-}e = \pi_2^{ac}(\infty)&(R_{S_-,S_-} - \frac{d}{ds}\tilde{Q}^{ac}(0))e.\end{aligned}$$

Using the above theorem we get the vector $\pi_2^{ac}(0)$ needed in Theorem 4.2. Next we compute the quantities $g(j,x)$ defined earlier. Clearly $g(j,x) = 1$ if $j \in S_+$ and $p_2(j) > 0$. Therefore, we only need to compute $g(j,x)$ if j is such that $p_2(j) = 0$. We need the following notation:

$$S_{20} := \{i \in S : p_2(i) = 0\},\ N_{20} := |S_{20}|, \quad \text{and } S_{2+} := \{i \in S : p_2(i) > 0\},\ N_{2+} := |S_{2+}|,$$

$$g(x) := [g(j,x),\ j \in S_{20}],\quad g'(x) := \left[\frac{dg}{dx}(j,x),\ j \in S_{20}\right].$$

Let $\lambda_k, \phi_k = [\phi_{jk}, j \in S_{20}]$ denote the eigenvalues and eigenvectors to the system

$$(\lambda_k (R_1)_{S_{20},S_{20}} + Q_{S_{20},S_{20}})\phi_k = 0,$$

and assume that all eigenvalues λ_k are distinct. Then (see, e.g., Kulkarni [3])

$$|\{k : Re\{\lambda_k\} \leq 0\}| = |S_{20} \cap S_-|.$$

The following lemma gives $g(x)$ as a solution to a system of linear ordinary differential equations. We omit the proof, since it follows along similar lines to the one of Lemma 4.3.

Lemma 4.5 *The column vector $g(x)$ satisfies*

$$(R_1)_{S_{20},S_{20}}g'(x) + Q_{S_{20},S_{20}}g(x) + Q_{S_{20},S_{2+}}e = 0, \tag{4.8}$$

with boundary conditions

$$g(j,0) = 0, \ j \in S_{20} \cap S_-. \tag{4.9}$$

The solution is given by

$$g(x) = \sum_{k:Re\{\lambda_k\} \leq 0} a_k e^{\lambda_k x} \phi_k + e,$$

where the coefficients a_k are the solution to

$$g(j,0) = \sum_{k:Re\{\lambda_k\} \leq 0} a_k \phi_{jk} + 1 = 0, \ j \in S_{20} \cap S_-.$$

5 Numerical Example

We illustrate the approach in case of four independent input sources with input rates

$$p = [p_1, p_2, p_3, p_4] = [8, 3, 10, 2],$$

and identical exponential on-times and off-times with respective parameters

$$\alpha_k = 4, \beta_k = 1, \quad k = 1, \ldots, 4.$$

In Figure 5.2 we vary the service capacity μ and plot the corresponding mean on-time B_k of the output process of class k, $k = 1, \ldots, 4$. The system becomes unstable for

$$\mu \leq \sum_{i=1}^{4} \frac{p_i \beta_i}{\alpha_i + \beta_i} = \frac{23}{5} = 4.6.$$

Thus, we vary μ from 4.75 to $23.02 > \sum_{k=1}^{4} p_k$.

By increasing μ, B_1 decreases and it becomes equal to the mean on-time of source 1, $1/\alpha_1 = 0.25$ for values of $\mu > p_1 = 8$. Thus, the observed pattern of B_1 is not surprising. The pattern of B_2 is different; it is *discontinuous* at $\mu = p_1 = 8$. When μ exceeds $p_1 = 8$, there is a significant jump of B_2 after which it decreases. The explanation is that for $\mu < p_1 = 8$, even though $p_2 = 3$ is relatively small, class 2 fluid gets interrupted by class 1, every time source 1 turns on when there is no class 1 fluid in the buffer. This does not happen for values of $\mu > p_1 = 8$ that can handle class 1 and always provide leftover capacity for class 2. Class 3 is getting interrupted more often. The leftover capacity for its service depends on class 1 and class 2 service requirements. This leads to two jumps of B_3, when μ exceeds $p_1 = 8$ and again when μ exceeds $p_1 + p_2 = 11$. Overall, larger values of B_3 are observed, since it arrives at a much larger rate $p_3 = 10$, leading to a significant accumulation in the buffer. The observed values of B_4 show the complexity of the problem: being the lowest priority, class 4 fluid is affected by all other flows. It is accumulated in the buffer for values of $\mu \leq \sum_{k=1}^{3} p_k = 21$ after which as μ comes closer to $\sum_{k=1}^{4} p_k = 23$ it approaches the mean input on-time of source 4, $1/\alpha_4 = 0.25$. Two differences in the behavior of B_4 in comparison to that of B_2 and B_3 can be observed. First, the values of B_4 stay below 0.25 from one point on and second, there are no such big jumps observed as in the cases of B_2 and B_3. A possible explanation is in the particular values of the parameters.

6 Conclusions

In this paper we considered a single node that serves multiple on-off fluid sources according to a static priority. The output process of a given source is said to be on if the fluid of that source is exiting the node at a positive rate, otherwise it is said to be off. The main contribution of this paper is the exact method of computing the mean on and off periods of the output processes for each source, although the motivation for these calculations comes from the desire to build networks of multi-class fluid queues. The efficacy of approximating the output processes by exponential on-off processes using the mean on and off times will

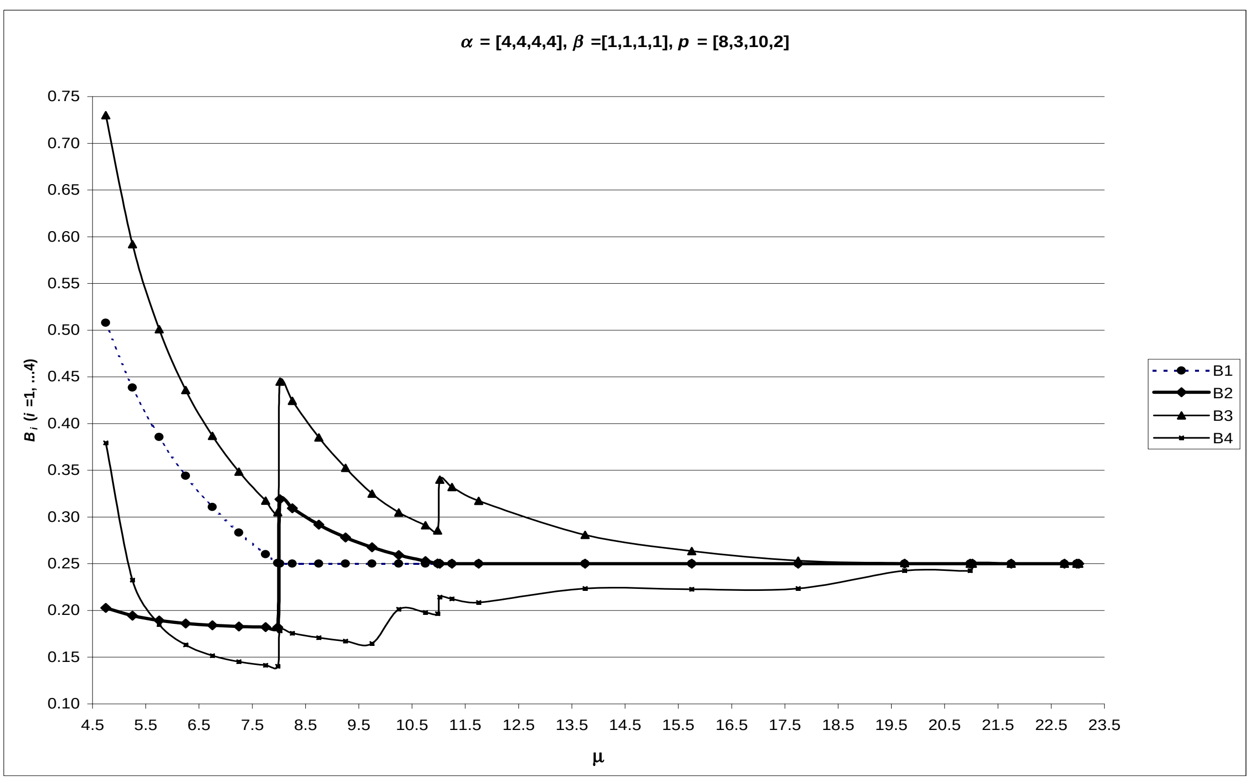

Figure 5.2: Mean on-times B_k , $k = 1, \ldots, 4$, as the service rate μ varies from 4.75 to 23.02

require a detailed study and is a topic of further research. The exact methods developed here can also provide an accuracy check for approximate methods that will need to be developed for computing the output mean on and off periods in order to make the network analysis tractable for large networks.

References

[1] Y. Hirasawa, (2000). Approximating Traffic Parameters in Multi-Class Fluid Networks. *Ph.D. Thesis, Department of Operations Research, University of North Carolina, Chapel Hill, N.C. 27599-3180*

[2] P. J. Kuehn, (1979) Approximate analysis of general queueing networks by decomposition. *IEEE Trans. Commun.,* COM-27, pp 113-126.

[3] V.G. Kulkarni (1995). *Modeling and Analysis of Stochastic Systems.* CRC Press.

[4] V. G. Kulkarni (1997). Fluid models for single buffer systems. *Frontiers in Queueing; Models and Applications in Science and Engineering*, 321-338, Ed. J.H. Dshalalow, CRC Press.

[5] V. G. Kulkarni and K. D. Glazebrook (2002). Output analysis of a single-buffer multiclass queue: FCFS service. *J. Appl. Probab.*, 39, pp 341-358.

[6] E. Tzenova, I.J.B.F. Adan, and V. Kulkarni (2005). Fluid models with jumps, *Sochastic Models,* 21, pp 37-56.

[7] W. Whitt (1983). The queueing network analyzer. *The Bell Syst. Tech. J.,* 62, pp 2779-2815.