Perturbation Analysis of Stochastic Fluid Models with Respect to the Fluid Arrival Process

C.G. Panayiotou, C.G. Cassandras, and G. Sun

Abstract—**In this paper we study a Stochastic Fluid Model (SFM) and derive sensitivity estimators for two performance measures of interest (workload and loss volume) with respect to the fluid inflow process. The derived estimators are proved unbiased and are very easy implement. Such estimators can be computed based on observable information from a *single* sample path, thus, they can be used in the control and optimization of queueing systems such as communication networks and manufacturing systems.**

Keywords— **Stochastic Fluid Models (SFM), Infinitesimal Perturbation Analysis (IPA), network management and control.**

I. Introduction

In this paper we adopt a Stochastic Fluid Model (SFM) for the control and optimization of queueing systems, such as communication networks and manufacturing systems, and analyze two performance measures of interest (workload and loss volume) as functions of the fluid inflow process. SFMs have been used extensively in various contexts, e.g., analyzing manufacturing or communication systems (see [1], [2], [3] and references therein).

In this paper we follow the approach used in [4], [5] and derive sample derivatives of the two measures (loss volume and workload) with respect to a parameter θ that controls the fluid inflow process (For more information on sample derivatives and Infinitesimal Perturbation Analysis (IPA) please refer to [6], [7]). The same model is also investigated in [4], [5]. In [4] we derive derivatives of the performance measures with respect to the buffer size. In [5] we generalize those results and derive sample derivatives with respect to various control parameters that affect the buffer capacity, as well as the fluid arrival and departure processes. An assumption made in [5] is that any discontinuities of the arrival process (as a function of time) are independent of the control parameters. In this paper, we derive one-sided derivatives for the same performance measures but with respect to a control parameter θ that *affects* the jump points of the inflow process; for example in a computer network setting with ON/OFF sources, the parameter θ may affect the instants when the ON to OFF or OFF to ON transitions occur. More specifically, we assume that the duration of each ON period is given by a distribution with a location parameter θ and derive the sensitivities of the loss volume and workload with respect to θ. Similar sensitivities with respect to the arrival process are also investigated in [8] for multiclass systems. However, in [8] the authors assume that θ is a scale rather than a location parameter of the distribution. Furthermore, [8] deals only with the sensitivity of the workload while in this paper we also derive the sensitivity of the loss volume.

The contribution of this paper consists of the derivation of sample derivative estimates of the performance measures of interest (workload and loss volume) with respect to a location parameter of the distribution of the duration of the ON period. These estimates are distribution invariant and can be computed based on observable information from a *single* sample path, thus they can be used for on-line control.

C.G. Panayiotou is with the Dept. of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. Email:christosp@ucy.ac.cy, corresponding author.

C.G. Cassandras and G. Sun are with the Dept. of Manufacturing Engineering, Boston University, Boston, MA. Email:cgc@bu.edu, gsun@bu.edu. They are supported in part by the National Science Foundation under Grant EEC-00-88073, by AFOSR under contract F49620-01-0056, and by ARO under under grant DAAD19-01-0610.

II. The Stochastic Fluid Model (SFM)

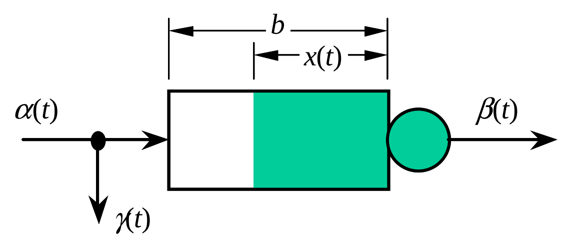

Fig. 1. System Model

The SFM setting is based on the fluid-flow world-view where “liquid molecules” flow in a continuous fashion. The basic SFM, used in [4] and shown in Fig. 1, consists of a single-server preceded by a fluid storage tank and it is characterized by five stochastic processes, all defined on a common probability space $(\Omega, \mathcal{F}, P)$ and labelled as follows:
$\{\alpha(t)\}$*:* the input flow rate (inflow) to the SFM.
$\{\beta(t)\}$*:* the service rate, i.e., the maximum fluid discharge rate from the server.
$\{\delta(t)\}$*:* the actual fluid discharge rate from the server,
$\{x(t)\}$*:* the buffer occupancy or buffer content, i.e., the amount of fluid in the buffer,
$\{\gamma(t)\}$*:* the overflow rate due to a full buffer.

The above processes evolve over a given time interval $[0, T]$ for a given fixed $0 < T < \infty$. We assume that the inflow process $\{\alpha(t)\}$ is described by an ON/OFF source as shown in Fig. 2. During the ON period the inflow rate $\alpha(t) = \alpha$ while during the OFF period $\alpha(t) = 0$. The duration of the ON and OFF periods are random variables from an arbitrary distribution and we denote by ξ_k and η_k the beginning and end of the kth ON period. Also, the service-rate process $\{\beta(t)\}$ is assumed to be constant, i.e.,

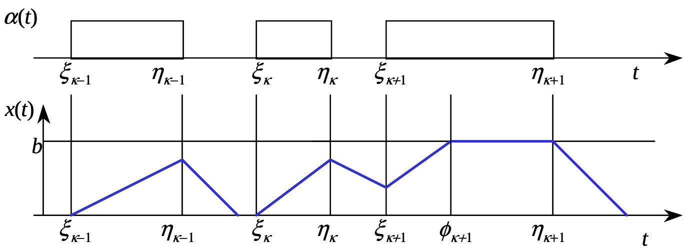

Fig. 2. Typical sample path

$\beta(t) = \beta$ for all $t \in [0, T]$. We assume that $\alpha > \beta$ otherwise no buffering will ever occur. Furthermore, we assume that the buffer has a finite capacity b.

For the purposes of this paper, the duration of the ON period is given by some distribution with a location parameter θ e.g., $f_\tau(\tau;\theta)$. The processes $\{\alpha(t;\theta)\}$ and $\{\beta(t)\}$, along with the buffer size b, define the behavior of the SFM. In particular, they define the buffer content, $x(t;\theta)$, the overflow rate $\gamma(t;\theta)$, and the output flow $\delta(t;\theta)$. Also, θ is the variable parameter we will concentrate on for the purpose of IPA. We will assume that if θ is perturbed by $\Delta\theta$, then, the duration of each ON period in $\alpha(t;\theta)$ will also change by $\Delta\theta$. This is precisely the definition of a location parameter which implies symmetric distributions like uniform, normal, etc. [7]. In this model, the instant when the ON period starts does not depend on θ; only the end of the period depends on θ which implies that if the length on the ON period increases by $\Delta\theta$ then the length of the OFF period also decreases by $\Delta\theta$. Note that the assumption that θ is a location parameter of the ON period distribution is not very restrictive since the analysis can be readily extended to scale parameters as well; simply replace $\Delta\theta$ by $(X/\theta)\Delta\theta$, where X is the random variable describing the duration of the ON period.

Let $A(t;\theta)$ denote the net inflow process

$$A(t;\theta) = \alpha(t;\theta) - \beta(t) \tag{1}$$

The buffer content is defined by the following one-sided differential equation,

$$\frac{dx(t;\theta)}{dt^+} = \begin{cases} 0, & \text{if } x(t;\theta) = 0 \text{ and } A(t;\theta) \le 0, \\ 0, & \text{if } x(t;\theta) = b \text{ and } A(t;\theta) \ge 0, \\ A(t;\theta), & \text{otherwise} \end{cases} \tag{2}$$

whose initial condition will be set to $x(0;\theta) = x_0$; for simplicity, we set $x_0 = 0$ throughout the paper. The overflow rate $\gamma(t;\theta)$ is given by the following equation,

$$\gamma(t;\theta) = \begin{cases} \alpha(t;\theta) - \beta(t), & \text{if x(t;}\theta\text{) = b,} \\ 0, & \text{if x(t;}\theta\text{) < b.} \end{cases} \tag{3}$$

Since the input function $\{\alpha(t;\theta)\}$ is piecewise constant and $\{\beta(t)\}$ is assumed constant (β), the state trajectory $x(t;\theta)$ is piecewise linear and continuous in t, and the output function $\gamma(t;\theta)$ is piecewise constant. Moreover, the state trajectory can be decomposed into two kinds of intervals: *boundary periods* (BP) and *non-boundary periods* (NBP). Boundary periods are maximal intervals during which the buffer is either empty ($x(t;\theta) = 0$) or full ($x(t;\theta) = b$). Non-boundary periods are supremal intervals during which the buffer is neither empty nor full ($0 < x(t;\theta) < b$).

We will assume that the real-valued parameter θ is confined to a closed and bounded (compact) interval Θ. Let $\mathcal{L}(\theta) : \Theta \to \mathbb{R}$ be a random function defined over an appropriate probability space $(\Omega, \mathcal{F}, P)$. Strictly speaking, we write $\mathcal{L}(\theta, \omega)$ to indicate that this sample function depends on the sample point $\omega \in \Omega$, but will drop ω unless it is necessary to stress this fact. In what follows, we will consider two performance metrics, the *Loss Volume* $L_T(\theta)$ and the *Work* $Q_T(\theta)$, both defined over the interval $[0, T]$ by the following equations:

$$L_T(\theta) = \int_0^T \gamma(t;\theta)dt, \tag{4}$$

$$Q_T(\theta) = \int_0^T x(t;\theta)dt. \tag{5}$$

where, as already mentioned, we assume that $x(\theta; 0) = 0$ at the start of the interval $[0, T]$. Observe that $\frac{1}{T}E[L_T(\theta)]$ is the *Expected Loss Rate* over the interval $[0, T]$, a common performance metric of interest (from which related metrics such as *Loss Probability* can also be derived). Similarly, $\frac{1}{T}E[Q_T(\theta)]$ is the *Expected Buffer Content* over $[0, T]$.

Our objective in the next section is the estimation of the sample derivatives $dL_T(\theta)/d\theta$ and $dQ_T(\theta)/d\theta$. Furthermore, we show that these estimates are also unbiased.

III. Infinitesimal Perturbation Analysis (IPA)

As already mentioned, the sample path of this system is partitioned in boundary (BP) and Non-Boundary (NBP) periods. Let (μ_n, ν_n) denote the nth NBP where μ_n denotes the event *buffer ceases to be empty* or *full* and ν_n denotes the event that *buffer becomes empty* or *full*. Similarly, $[\nu_n, \mu_{n+1}]$ denotes the nth BP. Note that in general a NBP starts either when $x(t) = 0$ and $\alpha(t) - \beta(t)$ becomes positive or when $x(t) = b$ and $\alpha(t) - \beta(t)$ becomes negative. In this case, since $\beta(t) = \beta$, a NBP starts at the beginning of either an ON or OFF period at instants ξ_k or η_k for some k. Also, a NBP may end during either an ON or OFF period; it ends during an ON period if the buffer becomes full at point ϕ_j or during an OFF period if the buffer becomes empty at point e_j for some j. Furthermore, we define

$$L_n(\theta) = \int_{\mu_n}^{\nu_n} \gamma(t;\theta)dt \quad \text{and} \quad \bar{L}_n(\theta) = \int_{\nu_n}^{\mu_{n+1}} \gamma(t;\theta)dt$$

$$Q_n(\theta) = \int_{\mu_n}^{\nu_n} x(t;\theta)dt \quad \text{and} \quad \bar{Q}_n(\theta) = \int_{\nu_n}^{\mu_{n+1}} x(t;\theta)dt$$

Using the above partitioning, we can rewrite the objectives as

$$L_T(\theta) = \sum_{n=1}^{N} L_n(\theta) + \sum_{n=1}^{\bar{N}} \bar{L}_n(\theta) \tag{6}$$

$$Q_T(\theta) = \sum_{n=1}^{N} Q_n(\theta) + \sum_{n=1}^{\bar{N}} \bar{Q}_n(\theta) \tag{7}$$

where N is the number of NBPs and $\bar{N}$ is the number of BPs in the observation interval $[0,T]$. Clearly, $L_n(\theta)=0$ for all $n=1,\cdots,N$ since there is no overflow during a NBP. Next we perturb the length of the ON period by increasing the end point of each ON period η_k by $\Delta\theta \geq 0$ to $\eta'_k = \eta_k + \Delta\theta$, $k=1,2,\cdots$ and derive the change in $L_T(\theta)$ and $Q_T(\theta)$, $\Delta L_T(\theta)$ and $\Delta Q_T(\theta)$ respectively.

$$\Delta L_T(\theta) = \sum_{n=1}^{N} \Delta L_n(\theta) + \sum_{n=1}^{\bar{N}} \Delta \bar{L}_n(\theta) \tag{8}$$

$$\Delta Q_T(\theta) = \sum_{n=1}^{N} \Delta Q_n(\theta) + \sum_{n=1}^{\bar{N}} \Delta \bar{Q}_n(\theta) \tag{9}$$

where

$$\Delta L_n(\theta) = \int_{\mu_n}^{\nu_n} \Delta\gamma(t;\theta)dt, \quad \Delta\bar{L}_n(\theta) = \int_{\nu_n}^{\mu_{n+1}} \Delta\gamma(t;\theta)dt$$
$$\Delta Q_n(\theta) = \int_{\mu_n}^{\nu_n} \Delta x(t;\theta)dt, \quad \Delta\bar{Q}_n(\theta) = \int_{\nu_n}^{\mu_{n+1}} \Delta x(t;\theta)dt$$

and

$$\Delta\gamma(t;\theta) = \gamma(t;\theta+\Delta\theta) - \gamma(t;\theta) \tag{10}$$
$$\Delta x(t;\theta) = x(t;\theta+\Delta\theta) - x(t;\theta) \tag{11}$$

To evaluate the above differences, we re-index all points η_k according to the NBP they belong to and write $\eta_{n,i}$ to indicate the ith occurrence of the source's transition from ON to OFF during the nth NBP; $i=1,\cdots,H_n$ where H_n is the number of such transitions. We reiterate that a NBP can start with either a *buffer ceases to be empty* or *full* event. In the case that the nth NBP starts with a *buffer ceases to be full* event, the beginning of the NBP coincides with a source transition from ON to OFF. In this case, $\nu_n = \eta_{n,1}$ and H_n includes this event as well (see Fig. 3).

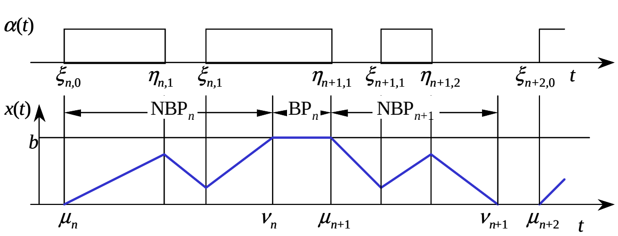

Fig. 3. Boundary and non-boundary periods

Next we divide a NBP into a set of intervals

$$\left\{(\mu_n,\eta_{n,1}) \bigcup \left\{ \bigcup_{i=1}^{H_n-1} [\eta_{n,i},\eta_{n,i+1}) \right\} \bigcup [\eta_{n,H_n},\nu_n) \right\}$$

and we will use the notation $I_{n,0}=(\mu_n,\eta_{n,1})$, $I_{n,i}=[\eta_{n,i},\eta_{n,i+1})$, $i=1,\cdots,H_n-1$, and $I_{n,H_n}=(\eta_{n,H_n},\nu_n)$. Also, for notational convenience we will use $\eta_{n,0}=\mu_n$, and, $\eta_{n,H_n+1}=\nu_n$. In addition, we point out that $I_{n,0}=(\mu_n,\eta_{n,1})$ exists only if the NBP starts after a *ceased to be empty* event. If the NBP starts after a *ceased to be full* event then $I_{n,0}$ is undefined. Furthermore, it is possible that $H_n=0$ and in this case the NBP constitutes a single interval (μ_n,ν_n). Finally, in every interval $I_{n,i}$, $i=1,\cdots,H_n-1$ there is a point $\xi_{n,i}$ which indicates the instant that the inflow process switches from OFF to ON.

Next we point out that any $\eta_{n,i}=\eta_{n,i}(\theta)$ is actually a function of θ, while, $\xi_{n,i}$, $i=1,\cdots,H_n$, $n=1,2,\cdots$, does not depend on θ. In the sequel, to simplify the notation, we drop the argument θ unless it is needed to make a point, and use the "prime" notation to indicate quantities in the perturbed sample, i.e., the sample path where θ is increased to $\theta+\Delta\theta$. For example, $\eta_{n,i}(\theta+\Delta\theta)=\eta'_{n,i}$, $x(t;\theta+\Delta\theta)=x'(t)$ and so on. Furthermore, we define the following quantities:

$\lambda_{n,i}(t;\theta)=\lambda_{n,i}(t)$*:* Loss volume during the ith interval $t\in I_{n,i}$ and by definition $\lambda_{n,i}(\eta_{n,i})=0$, $i=0,\cdots,H_n$, $n=1,2,\cdots$.

and the corresponding perturbations:

$$\begin{aligned}\Delta\lambda_{n,i}(t) &= \lambda_{n,i}(t;\theta+\Delta\theta)-\lambda_{n,i}(t;\theta)\\ &= \lambda'_{n,i}(t)-\lambda_{n,i}(t) \quad \text{for} \quad t\in I_{n,i},\end{aligned} \tag{12}$$

$$\Delta\bar{\lambda}_{n,i} = \lambda'_{n,i}(\eta_{n,i+1}) - \lambda_{n,i}(\eta_{n,i+1}) \tag{13}$$

$i=0,\cdots,H_n$. An example of the nominal and perturbed sample paths are shown in Fig. 4.

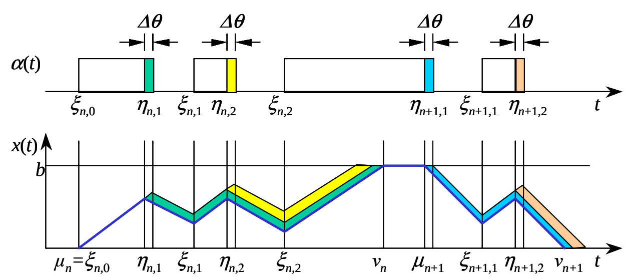

Fig. 4. Nominal and perturbed sample paths

A. Perturbation in a Non-Boundary Period

Next we focus on the nth NBP and determine ΔL_n and ΔQ_n in the interval (μ_n,ν_n). In the remaining of this section, for notational convenience we drop the subscript n unless it is needed avoid confusion. The interval (μ,ν) is divided into a set of intervals $I_0,\cdots,I_H$, and we assume that $\Delta x(\mu)=0$ (this is not always the case but as we will show later, it is true in expectation and it doesn't matter). Furthermore, during a NBP $x(t)<b$ and $\gamma(t)=0$ therefore $L_n=0$ and as a result $\Delta L_n = L'_n$.

We start first from the interval $I_0=(\mu,\eta_1)$ where it is clear that $\Delta x(t)=\Delta x(\mu)=0$ for all $t\in I_0$ since $x(t)$ is a continuous function of t and the sample paths of the nominal and perturbed sample paths are identical. Consequently, $\Delta\bar{\lambda}_0=0$.

Remark 1: At this point notice that for NBPs where $H=0$, $\Delta x(t)=0$ for all $t\in(\mu,\nu)$, thus such NBPs will not contribute any perturbation to either the loss volume or workload.

Next, we investigate the interval I_1 which we break into three subintervals; (a) $[\eta_1, \eta_1')$ where perturbation is generated, (b) $[\eta_1', \xi_1)$ the remaining OFF period and (c) $[\xi_1, \eta_2)$ the following ON period. In our analysis we show how perturbations are *generated* in subinterval (a) and how they are *propagated* in subintervals (b) and (c). For the subinterval $[\eta_1, \eta_1') = [\eta_1, \eta_1 + \Delta\theta)$ we identify two cases,

Case I: If $x(\eta_1) = b$, then $x'(\eta_1) = b$ and $\Delta x(\eta_1) = 0$. Since $\alpha > \beta$ then $x'(t) = b$ for all $t \in [\eta_1, \eta_1')$ while $x(t) < b$, and $\Delta x(t)$ for $t \in [\eta_1, \eta_1')$ is given by

$$\Delta x(t) \;=\; b - x(\eta_1) + \int_{\eta_1}^{t} \beta d\tau \;=\; \beta(t - \eta_1)$$

Note that this case is possible *only* at the beginning of a NBP that starts with a *cease to be full* event. Furthermore, the loss in this interval is

$$\lambda_1'(t) = \int_{\eta_1}^{t} (\alpha - \beta)dt \;\text{ and }\; \lambda_1'(\eta_1') = (\alpha - \beta)\Delta\theta. \tag{14}$$

Case II: If $x(\eta_1) < b$, then there are two subcases as follows

II.a If $b - x(\eta_1) \geq (\alpha - \beta)\Delta\theta$ then neither the nominal nor the perturbed sample paths will become full and therefore,

$$\begin{aligned}\Delta x(t) &= x(\eta_1) + \int_{\eta_1}^{t} (\alpha - \beta)d\tau - x(\eta_1) + \int_{\eta_1}^{t} \beta d\tau \\ &= \alpha(t - \eta_1)\end{aligned}$$

and $\lambda_1'(t) = 0$ for all $t \in (\eta_1, \eta_1')$.

II.b If $b - x(\eta_1) < (\alpha - \beta)\Delta\theta$ then, in the perturbed sample path, the buffer will become full in this interval, i.e., $\eta_1 < \phi_1' \leq \eta_1'$. Where ϕ_j' indicates the time when the perturbed sample path becomes full for the jth time, $j = 1, 2, \cdots$. In this case, for all $t \in [\eta_1, \phi_1')$

$$\Delta x(t) = \alpha(t - \eta_1)$$

while for all $t \in [\phi_1', \eta_1')$

$$\begin{aligned}\Delta x(t) &= b - x(\phi_1') + \int_{\phi_1'}^{t} \beta d\tau \\ &= \alpha(t - \eta_1) + b - x(\eta_1) - (\alpha - \beta)(t - \eta_1).\end{aligned}$$

where we used $x(\phi_1') = x(\eta_1) - \beta(\phi_1' - \eta_1)$. Thus combining the results for **II.a** and **II.b** above, we get

$$\Delta x(t) = \alpha(t - \eta_1) - [x(\eta_1) + (\alpha - \beta)(t - \eta_1) - b]^{+} \tag{15}$$

At this point, for notational convenience we define the quantity

$$Y(t) = x(t) + \Delta x(t) - b \leq 0 \tag{16}$$

Note that $Y(t) \leq 0$ for any $t \in [0, T]$ because $x'(t) = x(t) + \Delta x(t) \leq b$ since both nominal and perturbed sample paths have the same buffer capacity b.

Combining the two cases together we get

$$\Delta x(t) = \left\{ \begin{array}{ll} \alpha(t - \eta_1) - [Y(\eta_1) + (\alpha - \beta)(t - \eta_1)]^{+} & \text{if } x(\eta_1) < b \\ \beta(t - \eta_1) & \text{if } x(\eta_1) = b \end{array} \right. \tag{17}$$

for all $t \in [\eta_1, \eta_1')$ (note that $\Delta x(\eta_1) = 0$). Thus

$$\Delta x(\eta_1') = \left\{ \begin{array}{ll} \alpha\Delta\theta - [Y(\eta_1) + (\alpha - \beta)\Delta\theta]^{+} & \text{if } \; x(\eta_1) < b \\ \beta\Delta\theta & \text{if } \; x(\eta_1) = b \end{array} \right. \tag{18}$$

Furthermore,

$$\begin{aligned}\Delta\lambda_1(\eta_1') &= \lambda_1'(\eta_1') \\ &= \left\{ \begin{array}{ll} (\alpha - \beta)\Delta\theta & \text{if } \; x(\eta_1) = b \\ [Y(\eta_1) + (\alpha - \beta)\Delta\theta]^{+} & \text{if } \; x(\eta_1) < b \end{array} \right.\end{aligned} \tag{19}$$

Next, consider the interval $[\eta_1', \xi_1)$. During this interval,

$$\begin{aligned}x(t) &= \left[x(\eta_1') - \int_{\eta_1'}^{t} \beta d\tau \right]^{+} \\ x'(t) &= \left[x(\eta_1') + \Delta x(\eta_1') - \int_{\eta_1'}^{t} \beta d\tau \right]^{+}\end{aligned}$$

If this interval is fully contained in the NBP (i.e., the buffer does not become empty), then $x(\xi_1) > 0$. Furthermore, $x'(\xi_1) > 0$ (This can be easily seen from (18) since $\Delta\theta \geq 0$ and $Y(\xi_1) \leq 0$ due to (16)) and thus, in this interval, there is no perturbation generated or cancelled, therefore $\Delta x(t)$ for $t \in [\eta_1', \xi_1)$ is given by

$$\begin{aligned}\Delta x(t) &= \Delta x(\eta_1') \\ &= \left\{ \begin{array}{ll} \alpha\Delta\theta - [Y(\eta_1) + (\alpha - \beta)\Delta\theta]^{+} & \text{if } x(\eta_1) < b \\ \beta\Delta\theta & \text{if } x(\eta_1) = b \end{array} \right.\end{aligned} \tag{20}$$

Remark 2: If the buffer empties before the next ξ instant implies that the NBP has ended with an empty period at ν with $\Delta x(\nu) = \Delta x(\eta_1')$.

Finally, we investigate $t \in [\xi_1, \eta_2)$ and identify two cases where we exclude the case $x(\eta_2) = b$ because it implies that the NBP has ended earlier. This case will be covered later when we investigate BPs.

Case I: If $b - x(\eta_2) \geq \Delta x(\eta_1')$ then the perturbed sample path will not overflow during this interval and therefore

$$\Delta x(t) = \Delta x(\xi_1) = \Delta x(\eta_1') \quad \text{for all } \xi_1 \leq t \leq \eta_2$$

and $\lambda_1'(t) = \lambda_1'(\eta_1')$, $\xi_1 \leq t \leq \eta_2$.

Case II: If $b - x(\eta_2) < \Delta x(\eta_1')$ then $\xi_1 \leq \phi_i' < \eta_2$, for some $i = 1, 2$ and therefore

$$\begin{aligned}\Delta x(t) &= \left\{ \begin{array}{ll} \Delta x(\xi_1) & \text{for } \; \xi_1 \leq t < \phi_i' \\ \Delta x(\xi_1) - (\alpha - \beta)(t - \phi_i') & \text{for } \; \phi_i' \leq t < \eta_2 \end{array} \right. \\ &= \Delta x(\xi_1) - (\alpha - \beta) \left[t - \eta_2 + \frac{x(\eta_2) + \Delta x(\xi_1) - b}{\alpha - \beta} \right]^{+}\end{aligned}$$

where the term $[\cdot]^{+}$ is the loss volume in the perturbed sample path. Combining the two cases for $t \in [\xi_1, \eta_2)$ we get

$$\Delta x(t) = \Delta x(\eta_1') - (\alpha - \beta) \left[t - \eta_2 + \frac{x(\eta_2) + \Delta x(\eta_1') - b}{\alpha - \beta} \right]^{+}$$

and at the end of the interval we get

$$\Delta x(\eta_2) = \begin{cases} \beta\Delta\theta - [x(\eta_2) + \beta\Delta\theta - b]^+ & \text{if } x(\eta_1) = b \\ \alpha\Delta\theta - [Y(\eta_1) + (\alpha-\beta)\Delta\theta]^+ & \\ \quad - [x(\eta_2) + \Delta x(\eta_1') - b]^+ & \text{if } x(\eta_1) < b \end{cases} \tag{21}$$

and $\Delta\lambda_1' = \lambda_1'(\eta_1') + [x(\eta_2) + \Delta x(\eta_1') - b]^+$. Substituting $\lambda_1'(\eta_1')$ from (19) we get

$$\Delta\lambda_1' = \begin{cases} (\alpha-\beta)\Delta\theta + [x(\eta_2) + \beta\Delta\theta - b]^+ & \text{if } x(\eta_1) = b \\ [Y(\eta_1) + (\alpha-\beta)\Delta\theta]^+ & \\ \quad + [x(\eta_2) + \Delta x(\eta_1') - b]^+ & \text{if } x(\eta_1) < b \end{cases} \tag{22}$$

So far we have determined the perturbation in the buffer content in the interval $[\mu, \eta_2)$. The perturbation $\Delta x(\eta_2)$ will propagate to the remaining NBP so our objective is to find the perturbation at any t in any interval I_k in the NBP, $k = 1, \cdots, H$. So, next we investigate the kth interval I_k which, as before, is broken into three subintervals; (a) $[\eta_k, \eta_k')$ where perturbation is generated, (b) $[\eta_k', \xi_k)$ the remaining OFF period and (c) $[\xi_k, \eta_{k+1})$ the following ON period. Due to space limitations we only present the final results for each interval. For the subinterval $[\eta_k, \eta_k') = [\eta_k, \eta_k + \Delta\theta)$, we get

$$\Delta x(t) = \Delta x(\eta_k) + \alpha(t-\eta_k) - [Y(\eta_k) + (\alpha-\beta)(t-\eta_k)]^+ \tag{23}$$

where the term in $[\cdot]^+$ is the possible loss volume in the perturbed sample path. Note that (23) is similar to (17) with the addition of the $\Delta x(\eta_k)$ which corresponds to the perturbation propagated due to earlier perturbations. Therefore, at the end of the interval

$$\Delta x(\eta_k') = \Delta x(\eta_k) + \alpha\Delta\theta - [Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+ \tag{24}$$

and

$$\lambda_k'(\eta_k') = [Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+ .$$

Equation (24) suggests that the perturbation at η_k' consists of three parts. (a) $\Delta x(\eta_k)$ is the perturbation that has been *accumulated* earlier. (b) $\alpha\Delta\theta$ is the perturbation *generated* at η_k and (c) $[Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+$ is the possible overflow in the perturbed sample path.

Next, consider the interval $[\eta_k', \xi_k)$. Following earlier arguments (see (20)) we find that for all $t \in [\eta_k', \xi_k)$

$$\begin{aligned} \Delta x(t) &= \Delta x(\eta_k') \\ &= \Delta x(\eta_k) + \alpha\Delta\theta - [Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+ \end{aligned} \tag{25}$$

Finally, we investigate the subinterval $[\xi_k, \eta_{k+1})$. For $t \in [\xi_k, \eta_{k+1})$ we get

$$\begin{aligned} \Delta x(t) &= \Delta x(\eta_k) + \alpha\Delta\theta - [Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+ \\ &\quad - (\alpha-\beta)\left[t - \eta_{k+1} + \frac{x(\eta_{k+1}) + \Delta x(\eta_k') - b}{\alpha-\beta}\right]^+ \end{aligned} \tag{26}$$

and at the end of the interval we get

$$\begin{aligned} \Delta x(\eta_{k+1}) &= \Delta x(\eta_k) + \alpha\Delta\theta - [Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+ \\ &\quad - [x(\eta_{k+1}) + \Delta x(\eta_k') - b]^+ \end{aligned} \tag{27}$$

$$\Delta\lambda_k = [Y(\eta_k) + (\alpha-\beta)\Delta\theta]^+ + [x(\eta_{k+1}) + \Delta x(\eta_k') - b]^+ \tag{28}$$

Now we are ready to determine the change in workload ΔQ_n and loss volume ΔL_n.

Lemma 1: During the nth NBP, the change in the workload is given by

$$\begin{aligned} \Delta Q_n &= (\beta-\alpha)\Delta\theta(\nu_n - \eta_{n,1})\mathbf{1}[x(\mu_n) = b] \\ &\quad + \alpha\Delta\theta \sum_{i=1}^{H_n}(\nu_n - \eta_{n,i}) \\ &\quad - \sum_{i=1}^{H_n}[Y(\eta_{n,i}) + (\alpha-\beta)\Delta\theta]^+ (\nu_n - \eta_{n,i}) \\ &\quad - \sum_{i=1}^{H_n-1}\left[x(\eta_{n,i+1}) + \Delta x(\eta_{n,i}') - b\right]^+ (\nu_n - \eta_{n,i}) \end{aligned} \tag{29}$$

where $\Delta x(\eta_{n,i}')$ is given by (24).

Proof: Due to space limitations, all proofs are omitted.

Lemma 2: During the nth NBP, the change in the loss volume is given by

$$\begin{aligned} \Delta L_n &= (\alpha-\beta)\Delta\theta\mathbf{1}_b + \sum_{i=1}^{H_n-1}\left[x(\eta_{n,i+1}) + \Delta x(\eta_{n,i}') - b\right]^+ \\ &\quad + [Y(\eta_{n,1}) + (\alpha-\beta)\Delta\theta]^+ \mathbf{1}_0 + \sum_{i=2}^{H_n}[Y(\eta_{n,i}) + (\alpha-\beta)\Delta\theta]^+ \\ &\quad + \Bigg((\beta-\alpha)\Delta\theta\mathbf{1}_b + \alpha\Delta\theta H_n - \sum_{i=1}^{H_n-1}[Y(\eta_{n,i}) + (\alpha-\beta)\Delta\theta]^+ \\ &\quad - \sum_{i=1}^{H_n-1}\left[x(\eta_{n,i+1}) + \Delta x(\eta_{n,i}') - b\right]^+\Bigg)\mathbf{1}[x(\nu_n) = b] \end{aligned} \tag{30}$$

where $\mathbf{1}_b = \mathbf{1}[x(\mu_n) = b]$ and $\mathbf{1}_0 = \mathbf{1}[x(\mu_n) = 0]$.

B. Perturbation in a Boundary Period

In this section we investigate the perturbation generated in the boundary period $[\nu_n, \mu_{n+1}]$. Depending on the type of boundary period we identify the following two cases.

Case I: $x(t) = 0$ for all $t \in [\nu_n, \mu_{n+1}]$. This period will start with the event *buffer becomes empty* which can occur only during an OFF period, i.e., at some time instant $e_n = \nu_n$, $\eta_{n,H_n} < e_n < \xi_{n,H_n}$. From the analysis of the NBP (see (25)), we get that

$$\Delta x(\nu_n) = \Delta x(\eta_{n,H_n}) + \alpha\Delta\theta - [Y(\eta_{n,H_n}) + (\alpha-\beta)\Delta\theta]^+ \tag{31}$$

Furthermore, since $x(\nu_n) = 0$, by definition $x'(\nu_n) = \Delta x(\nu_n)$ therefore for $t \in [\nu_n, \mu_{n+1}]$

$$\Delta x(t) = \left[\Delta x(\nu_n) - \int_{\nu_n}^{t} \beta dt\right]^+ = [\Delta x(\nu_n) - \beta(t-\nu_n)]^+$$

and for intervals that are long enough, $\Delta x(\mu_{n+1}) = 0$.

Thus,

$$\Delta \bar{Q}_n = \int_{\nu_n}^{\nu_n + \frac{\Delta x(\nu_n)}{\beta}} \left(\Delta x(\nu_n) - \beta(t - \nu_n) \right) dt \qquad (32)$$

Furthermore, for $b > 0$, neither sample path will experience any loss and therefore, $\Delta \bar{L}_n = 0$.

Case II: $x(t) = b$ for all $t \in [\nu_n, \mu_{n+1}]$. In this case, since $\Delta x(t) \geq 0$, $x'(t) = b$ and thus $\Delta x(t) = 0$ for all $t \in [\nu_n, \mu_{n+1}]$. Furthermore, $\gamma(t) = \gamma'(t) = \alpha - \beta$ and therefore $\Delta \bar{L}_n = 0$.

IV. Unbiasedness

First we present the following lemma (see [4] for the proof) which is useful in establishing the fact that terms involving $[\cdot]^+$ (as in Lemmas 1 and 2) do not contribute to the derivative estimates of $\mathbb{E}[Q_T(\theta)]$ and $\mathbb{E}[L_T(\theta)]$.

Lemma 3: Let Δ and I be non-negative random variables and let $f_\Delta(x)$ be the pdf of I conditioned on Δ with $f_\Delta(x) \leq C < \infty$ for all x in the support of $f_\Delta(x)$. Then

$$\mathbb{E}[\Delta - I]^+ \leq C\mathbb{E}\left[\Delta^2\right] \qquad (33)$$

Lemma 4:

$$[(\alpha - \beta)\Delta\theta + Y(\eta_{n,H_n-1})]^+ \leq (\alpha - \beta)C\Delta\theta^2 \qquad (34)$$

for all $n = 1, \cdots, N$.

Lemma 5:

$$\left[\Delta x(\eta'_{n,i}) - (b - x(\eta_{n,i+1}))\right]^+ \leq (\alpha - \beta)\bar{C}\Delta\theta^2 \qquad (35)$$

for all $i = 1, \cdots, H_n$ and $n = 1, \cdots, N$.

Theorem 3: The (right) derivative of $L_T(\theta)$ with respect to the θ is given by

$$\frac{d\mathbb{E}[Q_T(\theta)]}{d\theta} = \mathbb{E}\left[\frac{dQ_T(\theta)}{d\theta}\right]$$
$$= \mathbb{E}\left[\sum_{n=1}^{N}\left((\beta - \alpha)(\nu_n - \eta_{n,1})\mathbf{1}_b + \alpha\sum_{i=1}^{H_n}(\nu_n - \eta_{n,i})\right)\right]. \qquad (36)$$

Theorem 4: The (right) derivative of $L_T(\theta)$ with respect to θ is given by

$$\frac{d\mathbb{E}[L_T(\theta)]}{d\theta} = \mathbb{E}\left[\frac{dL_T(\theta)}{d\theta}\right]$$
$$= \mathbb{E}\left[(\alpha - \beta)\sum_{n=1}^{N}\mathbf{1}_b + \sum_{n=1}^{N}\left(\alpha H_n + (\beta - \alpha)\mathbf{1}_b\right)\mathbf{1}[x(\nu_n) = b]\right] \qquad (37)$$

V. IPA Algorithms

Initialize: $dL = 0$, $dQ = 0$, $dx = 0$, $\tau = 0$.

If at t NBP ends: buffer becomes full or empty

- If *buffer becomes full*: $dL \leftarrow dL + dx$. `%notes 1 and 5`
- $dQ \leftarrow dQ + dx(t - \tau)$ `%see note 3`
- $dx = 0$. `%see note 2`

If at t source transitions from ON to OFF:

- If *buffer ceases to be full*:
  – $dL \leftarrow dL + (\alpha - \beta)$ `%see note 4`
  – $dx = \beta$ `%see note 1`

Else
  – $dQ \leftarrow dQ + dx(t - \tau)$ `%see note 3`
  – $dx \leftarrow dx + \alpha$ `%see note 1`
- $\tau = t$

At the end of the observation interval,

$$\frac{dQ_T(\theta)}{d\theta} = dQ \quad \text{and} \quad \frac{dL_T(\theta)}{d\theta} = dL$$

Notes:

1. Perturbation generation. If the nth NBP starts with an ON to OFF transition, then the perturbation in $x(\cdot)$ is β. This can be seen from the second sum term of (37) where one α due to the first ON to OFF transition is cancelled by the $(\beta - \alpha)$ term. Furthermore, we arrive to the same conclusion from (36) where the first term of the outer sum is $(\beta - \alpha)(\nu_n - \eta_{n,1})$ while from the inner sum we get a term $\alpha(\nu_n - \eta_{n,1})$. As a result, the net contribution is a term $\beta(\nu_n - \eta_{n,1})$. On the other hand, if the ON to OFF transition does not occur at the beginning of the NBP, then the perturbation is simply α (see the inner sum of (36)).
2. At the end of the NBP the perturbation is reset to zero.
3. Rather than saving *all* points $\eta_{n,i}$, $i = 1, \cdots, H_n$ until the end of the NBP in order to evaluate the inner sum of (36) we recognize that

$$\alpha\sum_{i=1}^{H_n}(\nu_n - \eta_{n,i}) = \alpha\sum_{i=1}^{H_n} i(\eta_{n,i+1} - \eta_{n,i})$$

and by convention we set $\eta_{n,H_n+1} = \nu_n$.
4. If a NBP starts with a *cease to be full* event, then there is a term $\alpha - \beta$ contributed to the loss derivative (see the first sum term in (37)).
5. If a NBP ends with a *buffer becomes full* event then *all* generated contribution (accumulated due to the second sum of (37) or the accumulator dx from the algorithm above) will result in additional losses.

References

[1] G. Kesidis, A. Singh, D. Cheung, and W. Kwok, "Feasibility of fluid-driven simulation for ATM network," in *Proc. IEEE Globecom*, vol. 3, pp. 2013–2017, 1996.

[2] B. Liu, Y. Guo, J. Kurose, D. Towsley, and W. Gong, "Fluid simulation of large scale networks: Issues and tradeoffs," in *Proceedings of the Intl. Conf. on Parallel and Distributed Processing Techniques and Applications*, June 1999. Las Vegas, Nevada.

[3] A. Yan and W. Gong, "Fluid simulation for high-speed networks with flow-based routing," *IEEE Transactions on Information Theory*, vol. 45, pp. 1588–1599, 1999.

[4] C. Cassandras, G. Sun, C. Panayiotou, and Y. Wardi, "Perturbation analysis and control of two-class stochastic fluid models for communication networks," *IEEE Transactions on Automatic Control*, 2003. Accepted for publication; a preprint can be found at http://vita.bu.edu/cgc/fullpubs/TwoClassSFM.pdf.

[5] Y. Wardi, B. Melamed, C. Cassandras, and C. Panayiotou, "IPA gradient estimators in single-node stochastic fluid models," *Journal of Optimization Theory and Applications*, vol. 115, no. 2, pp. 369–406, 2002.

[6] Y. Ho and X. Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*. Boston, MA: Kluwer, 1991.

[7] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, 1999.

[8] Y. Liu and W. Gong, "Perturbation analysis for stochastic fluid queuing systems," *Journal of Discrete Event Dynamic Systems: Theory and Applications*, vol. 12, pp. 391–416, 2002.