Smoothed Perturbation Analysis Derivative Estimation for Markov Chains

Michael C. Fu
College of Business and Management, University of Maryland, College Park, MD 20742, USA

Jian-Qiang Hu
Dept. of Manufacturing Engineering, Boston University, Boston, MA 02115, USA

August 1991; revised May 1993

Abstract

Using the technique of smoothed perturbation analysis, we consider steady-state performance measures for continuous-time Markov chains and derive a derivative estimator that can be estimated from a single sample path of the chain under consideration. The estimator is applicable to multi-class queueing networks, for which previous infinitesimal perturbation analysis estimators failed. A simple multi-class queueing network is used to illustrate the application of the estimator.

Keywords: sensitivity analysis; derivative estimation; perturbation analysis; Markov chains

1 Introduction

Let $\{Z(t)\}$ be a positive recurrent, irreducible Markov chain, and let $\bar{f}(\theta)$ be a steady-state performance measure defined on $\{Z(t)\}$, where θ is a parameter (or vector of parameters) in the infinitesimal generator (rate) matrix. We wish to estimate $d\bar{f}(\theta)/d\theta$ from a single sample path of the system. For example, $\{Z(t)\}$ could define the state of a multi-class queueing network in terms of the customers at each station at time t, $\bar{f}(\theta)$ could be the throughput of the system, and θ could be the service rate at one of the stations for a particular class of customers. Then $d\bar{f}(\theta)/d\theta$ gives the sensitivity of throughput to the particular service rate. Furthermore, the estimate could also be used to improve performance by adjusting the rate appropriately.

The topic of derivative estimation for Markov chains was also addressed in Glasserman (1992), where an infinitesimal perturbation analysis (IPA) estimator was constructed for Markov chains satisfying a certain structural condition, and in Gong (1988), where the technique of smoothed perturbation analysis (SPA) was applied to a multi-class queueing network. In our work, we also apply SPA, but unlike the estimators derived in Gong (1988), our estimators do not require additional simulation, i.e., they can be estimated from a *single* simulation. The basic idea behind SPA, formulated by Gong and Ho (1987), is the use of conditional expectation to smooth a discontinuous sample function. The advantage of the estimator we derive over that of Glasserman (1992) is that it need not satisfy the structural condition, and thus can handle very general multi-class queueing systems that cannot be handled by IPA. We illustrate this result with a simple multi-class queueing network example.

In Fu and Hu (1992), derivative estimators for transient performance measures of much more general discrete-event systems were derived using SPA, and unbiasedness of the estimators was proved. However, the generality comes at a cost, in that the estimator contains terms whose estimation may require prohibitively high additional simulation. In this work, we show that by specializing to steady-state performance measures of continuous-time Markov chains, we obtain an estimator that can be easily estimated from a single sample path (or simulation) of the system.

The remainder of the paper is organized as follows. Section 2 contains a brief descriptive summary of the results for the general form of the estimator derived in Fu and Hu (1992). Section 3 derives the estimator for continuous-time Markov chains. Section 4 contains the illustrative multi-class queueing network example.

2 Summary of Previous Results

In this section, we present the general estimator in Fu and Hu (1992), and then use it in the next section to derive the estimator for continuous-time Markov chains. We assume that the reader has some familiarity with generalized semi-Markov processes (GSMPs) (see, e.g., Whitt, 1980, or Glasserman, 1991). Recall that the evolution of a GSMP, say $\{Z(t)\}$, is characterized by a sequence of triples consisting of a state, the holding time spent in that state, and the event triggering the transition to the next state. The underlying timing is driven by clocks that keep track of the times until the next occurrence of the different possible events. The triggering event is the event with the shortest clock time remaining among all possible events. For example, in a closed queueing network, the events would be potential service completions at the various stations, with the triggering event being the next service completion. We define

$$\begin{aligned}
t_k &= \text{epoch of the } k\text{th state transition,} \\
S_k &= \text{the } k\text{th state visited,} \\
\tau_{k+1} &= \text{the time spent in } S_k \\
&= t_{k+1} - t_k, \\
e_{k+1} &= \text{the event while in state } S_k \text{ that causes the transition to } S_{k+1}, \\
\mathcal{E}(S) &= \text{the set of possible events in state } S, \\
F_e(\cdot) &= \text{the distribution (with density } f_e\text{) function of the clock time for event } e, \\
X_k(e) &= \text{total clock time for event } e \in \mathcal{E}(S_k)\text{, distributed as } F_e, \\
C_k(e) &= \text{remaining clock time for event } e \in \mathcal{E}(S_k) \text{ at } t_k.
\end{aligned}$$

The parameter is assumed to enter through the underlying distributions $\{F_e\}$. In our specialization to Markov chains, each F_e will be exponential. Let NP denote the original sample path of the GSMP, i.e., the so-called nominal path, let (DNP, k) denote the original sample path under the condition $\tau_{k+1} = 0$, and let (PP, k) denote the original sample path under the conditions $\tau_{k+1} = 0$ **and** the events e_k and e_{k+1} change order. Note that these conditions would usually result in (DNP, k) and

(PP,k) being substantially different from NP and from each other after visiting state S_{k-1}. For performance measure f, let $f^{DNP,k}$ denote f for the path (DNP,k), and let $f^{PP,k}$ denote f for the path (PP,k). These two quantities are crucial elements of our estimator. In addition, we need two other quantities, for which we first define

$$
\begin{aligned}
L_k(e) &= \text{the age of event } e \in \mathcal{E}(S_k) \text{ at } t_k \\
&= X_k(e) - C_k(e), \\
T_k(e) &= \text{most recent time before } t_k \text{ that a new clock time for event } e \in \mathcal{E}(S_k) \text{ was generated} \\
&= t_k - L_k(e), \\
y_k &= \text{the second shortest clock time at } t_k \\
&= \begin{cases} \infty & \text{if } |\mathcal{E}(S_k)| = 1 \\ \min_{e \in \mathcal{E}(S_k)\setminus\{e_{k+1}\}}\{C_k(e)\} & \text{otherwise} \end{cases}, \\
X^*_{k+1} &= X_k(e_{k+1}) - \tau_{k+1}.
\end{aligned}
$$

(Note that τ_{k+1} is by definition the shortest clock time at t_k.) Now, we can define

$$
f_{\tau_{k+1}}(x) = \frac{f_{e_{k+1}}(L_k(e_{k+1}) + x)}{F_{e_{k+1}}(L_k(e_{k+1}) + y_k) - F_{e_{k+1}}(L_k(e_{k+1}))}. \tag{1}
$$

Intuitively, $f_{\tau_{k+1}}(\cdot)$ represents the conditional (on certain clock times) density of τ_{k+1}.

Lastly, we define (see Figure 1)

$$
\left(\frac{d\tau_{k+1}}{d\theta}\right)^+_{IPA} = \left[\frac{dt_k}{d\theta} - \left(\frac{dT_{k+1}(e_{k+1})}{d\theta} + \left.\frac{dX_k(e_{k+1})}{d\theta}\right|_{X^*_{k+1}}\right)\right]^+, \tag{2}
$$

where $\cdot|_{X^*_{k+1}}$ indicates that the corresponding quantity is evaluated at X^*_{k+1}. Note that in Figure 1, we have $L_k(\beta) = X^*_{k+1}$. Intuitively, $(d\tau_{k+1}/d\theta)^+_{IPA}$ represents the difference in the accumulated IPA estimator at t_k and the accumulated IPA estimator at t_{k+1} under the condition $\tau_{k+1} = 0$. The quantities defining $(d\tau_{k+1}/d\theta)^+_{IPA}$ are calculated by the usual IPA techniques (e.g., Glasserman, 1991). Note that if $e_{k+1} = e_k$, then $(d\tau_{k+1}/d\theta)^+_{IPA}$ is zero, since in that case the two terms in the difference are identical; this simply means that two identical events will never change order.

Now we present the estimator derived in Fu and Hu (1992). For f a transient performance measure, the following is an unbiased estimator for $dE[f(\theta)]/d\theta$ (Fu and Hu, 1992, Theorem 3):

$$
\left(\frac{\partial f}{\partial \theta}\right)_{IPA} + \sum_{k \in L^*} (f^{PP,k} - f^{DNP,k}) f_{\tau_{k+1}}(0) \left(\frac{d\tau_{k+1}}{d\theta}\right)^+_{IPA}. \tag{3}
$$

The first term is the usual IPA estimator, which we shall call the IPA contribution of the estimator. It is well-known that IPA suffices for sample path derivative estimation when the performance measure of interest is almost surely continuous (e.g., Cao, 1988; Glasserman, 1991); intuitively, small changes in the parameter θ cause only small changes in the performance measure in a sample-pathwise sense. The second term is what we will refer to as the SPA contribution, and it estimates the additional effects when the conditions described in the previous sentence do not prevail. The set L^* represents such situations along the sample path, formally defined as critical adjacent event pairs in Fu and Hu

(1992). This concept will be discussed in more detail in the next section. The index k represents the kth state visited by the process. Thus, if the kth state visited involves a critical adjacent event pair, an appropriate quantity is calculated in the estimator. That quantity is the product of a jump rate – the rate at which the adjacent events change order as a function of the parameter – and the resultant expected effect of the event order change. The former is given by $f_{\tau_{k+1}}(0)(d\tau_{k+1}/d\theta)^+_{IPA}$ defined by Equations (1) and (2), whereas the latter is given by $(f^{PP,k} - f^{DNP,k})$, and these two quantities are independent. It is the quantity $(f^{PP,k} - f^{DNP,k})$ that may require prohibitively large amounts of additional simulation for general systems.

3 Derivation of the Estimator for Markov Chains

Since we will be using the estimator given by Equation (3), we will view continuous-time Markov chains as special cases of GSMPs, where all the clock times characterized by $\{F_e\}$ are exponential. Thus, when we consider sample paths and performances, we will retain the concept of an event triggering a state transition. In this paper, we consider steady-state performance measures of the following form:

$$\bar{f}(\theta) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(S_{i-1}, e_i, \tau_i),$$

which we take as the limit of the transient performance measure

$$f_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} g(S_{i-1}, e_i, \tau_i).$$

In Fu and Hu (1992), the estimator (3) introduced in the previous section was proved to be unbiased for estimating the derivative of the expected value of this transient performance measure. By specializing to steady-state performance measures of Markov chains, we shall see that exploiting the Markov property and taking the limit $n \to \infty$ yields a nice explicit estimator.

Since the chain is positive recurrent and irreducible, $\bar{f}$ is independent of its initial distribution. We first simplify the estimator for the steady-state performance measure. An adjacent event pair is defined by a state and the two subsequent events, e.g., $(e_k, e_{k+1}|S_{k-1})$ in Figure 1. Intuitively, adjacent events are **critical** if an order change in a sample path results in a significant change in the performance measure, which means that it is not included in the IPA contribution of the estimator.

First, we derive an efficient way to estimate the term $(f_n^{PP,k} - f_n^{DNP,k})$ in Equation (3), given that $(e_k, e_{k+1}|S_{k-1})$ is critical. For notational simplification, we will use $e_k = \alpha, e_{k+1} = \beta, S_{k-1} = S$, i.e., the critical adjacent event pair will be written as $(\alpha, \beta|S)$. We define the following quantities over the sample paths DNP and PP:

$$\begin{aligned} G_n^{DNP,k} &= \sum_{i=1}^{n} g(S_{i-1}^{DNP}, e_i^{DNP}, \tau_i^{DNP}), \\ G_n^{PP,k} &= \sum_{i=1}^{n} g(S_{i-1}^{PP}, e_i^{PP}, \tau_i^{PP}), \end{aligned}$$

where the superscripts over the states, events, and holding times indicate the corresponding quantities in that sample path. The key point to note is that the paths DNP and PP are identical to each other and to NP up to state S_{k-1}, i.e.,

$$S_{i-1}^{PP} = S_{i-1}^{DNP} = S_{i-1}, \quad e_i^{PP} = e_i^{DNP} = e_i, \quad \tau_i^{PP} = \tau_i^{DNP} = \tau_i, \quad \text{for } i \le k.$$

We can construct PP from DNP by inserting the piece of sample path that starts at S_{k+1}^{PP} and ends at S_{k+1}^{DNP}. Figure 2 demonstrates the idea, where the notation"$||$" following S_{k-1}^{DNP} indicates the simultaneity of the events α and β, with the time spent in the kth state visited being zero, i.e., $\tau_{k+1} = 0$. For Markov chains, due to the memoryless property of the exponential, this procedure is exact. This idea was first proposed by Gong (1988). However, what we propose is not to actually simulate the extra inserted path, but to take advantage of some statistical properties for large n. Let $G^{I,k}$ denote the following summation for a piece of sample path that starts at $S_0^I = S_{k+1}^{PP}$ and ends at the first hitting time of $S_{n_I}^I = S_{k+1}^{DNP}$:

$$G^{I,k} = \sum_{i=1}^{n_I} g(S_{i-1}^I, e_i^I, \tau_i^I),$$

where n_I denotes the number of events in that sample path. We also define

$$\Delta G^k = g(S_{k-1}, \tau_k, \beta) + g(S_k^{PP}, 0, \alpha) - g(S_{k-1}, \tau_k, \alpha) - g(S_k, 0, \beta),$$

which denotes the difference in the performance measure between the DNP and PP sample paths up to the $(k+1)$th state visited, due to the exchanged event order of α and β. For notational convenience, we drop the subscript or superscript k when there is no confusion. For n large, we have

$$\begin{aligned} f_n^{PP} - f_n^{DNP} &\approx f_{n+n_I}^{PP} - f_n^{DNP} = \frac{G_{n+n_I}^{PP}}{n+n_I} - \frac{G_n^{DNP}}{n} = \frac{G_n^{DNP} + G^I + \Delta G}{n+n_I} - \frac{G_n^{DNP}}{n} \\ &\approx \frac{G^I + \Delta G}{n} - \frac{n_I}{n}\frac{G_n^{DNP}}{n} \approx \frac{1}{n}\left(G^I + \Delta G - n_I \bar{f}\right). \end{aligned}$$

The first two $\approx'$s in the above derivation indicate where terms of $o(1/n)$ have been neglected. Thus,

$$E[f^{PP} - f^{DNP}] \approx \frac{1}{n}(E\left[G^I\right] + E[\Delta G] - E[n_I]\bar{f}).$$

Note that $E[n_I]$ is the expected length and $E\left[G^I\right]$ is the expected summed performance of a sample path that starts at S_{k+1}^{PP} and ends at S_{k+1}^{DNP}.

Labelling all possible critical adjacent event pairs as $L_m^* = (\alpha(m), \beta(m)|S(m))$, with corresponding quantities also indexed by m, we have for large n that the SPA contribution of Equation (3) is approximately equal in expectation to

$$\frac{1}{n}\sum_{L_m^*} P\{L_m^*\}\left(E\left[G^{I(m)}\right] + E[\Delta G(m)] - E[n_{I(m)}]\bar{f}\right) E\left[f_{\tau(m)}(0)\left(\frac{d\tau(m)}{d\theta}\right)_{IPA}^{+}\right], \qquad (4)$$

where the probabilities and expectations correspond to steady-state quantities associated with the occurrence of the critical adjacent event pair L_m^*.

Next, we consider the last expectation term in Equation (4). We define

$$\mu_e \quad = \quad \text{rate of exponential distribution for event } e.$$

Again, fix a critical adjacent event pair, $(\alpha, \beta | S)$. Let S' denote the state reached from state S with event α occurring. Substituting the exponential distribution into the definition of $f_{\tau(m)}(0)$ given by Equation (1), we obtain

$$f_\tau(0) = \lambda/(1 - e^{\lambda Y}),$$

where we have defined $\lambda = \mu_\beta$, and Y is a r.v. representing the second shortest clock at the time of occurrence of α in state S. By the memoryless property of the exponential distribution, the two terms in the expectation $E[f_{\tau(m)}(0)(d\tau(m)/d\theta)^+_{IPA}]$ are independent, because the second term involves only quantities before the occurrence of α. The expectation of the former turns out to be the sum of the rates of all possible events in state S:

Lemma. $E[f_\tau(0)] = \sum_{e \in \mathcal{E}(S')} \mu_e$.

Proof. The random variable Y is the minimum of the clocks of all events other than the next event, conditioned on being greater than the clock for the next event. In terms of Figure 1, at t_k it is the minimum of all event clocks other than β, under the condition that it is greater than τ. Let $X(e)$ denote the random variable representing the clock time of event e. By the memoryless property of the exponential distribution, at t_k, $X(e)$ has the distribution F_e, i.e., its original distribution. Then Y is given by $\min_{e \in \mathcal{E}(S') \setminus \{\beta\}} X(e)$. Since the clocks are all exponential, the (unconditional) distribution of Y is exponential with rate

$$\mu = \sum_{e \in \mathcal{E}(S') \setminus \{\beta\}} \mu_e.$$

The conditional density of Y is given by

$$\begin{aligned} f_Y(x) &= \frac{d}{dx} P(Y \le x | Y > \tau) = \frac{d}{dx} \frac{P(Y \le x, Y > \tau)}{P(Y > \tau)} = \frac{\frac{d}{dx} P(\tau < Y \le x)}{\lambda/(\lambda + \mu)} \\ &= \frac{\lambda + \mu}{\lambda} \frac{d}{dx} \int_0^x P(y < Y \le x) dF_\beta(y) = \frac{\lambda + \mu}{\lambda} \int_0^x \frac{\partial}{\partial x} [e^{-\mu y} - e^{-\mu x}] dF_\beta(y) \\ &= \frac{\lambda + \mu}{\lambda} \mu e^{-\mu x} (1 - e^{-\lambda x}) \end{aligned}$$

Thus,

$$E\left[\frac{\lambda}{1 - e^{-\lambda Y}}\right] = \int_0^\infty \frac{\lambda}{1 - e^{-\lambda x}} \frac{\lambda + \mu}{\lambda} \mu e^{-\mu x} (1 - e^{-\lambda x}) dx = \lambda + \mu = \sum_{e \in \mathcal{E}(S')} \mu_e. \qquad \Box$$

To calculate the probability term $P\{L^*_m\}$ in Equation (4), we again use the Markov property to note that it is given by $P(S)P(\alpha|S)P(\beta|S')$, where $P(S)$ is the long-run proportion of visits to state S, i.e., the stationary probability for the embedded discrete-time Markov chain, and $P(\alpha|S)$ and $P(\beta|S')$ denote the probabilities of the next event being α from state S and β from state S', respectively. These probabilities are given by $P(\alpha|S) = \mu_\alpha / \sum_{e \in \mathcal{E}(S)} \mu_e$ and $P(\beta|S') = \mu_\beta / \sum_{e \in \mathcal{E}(S')} \mu_e$. By definition, we have

$$P(S) = \lim_{n \to \infty} \frac{N_n(S)}{n},$$

where $N_n(S)$ is the number of visits to state S (out of n transitions), which we rewrite as

$$P(S) = \lim_{n \to \infty} \frac{N_n(S)}{T_n(S)} \frac{T_n(S)}{t_n} \frac{t_n}{n},$$

where $T_n(S)$ is the total time spent in state S (over n transitions). The first term is given by $\sum_{e\in\mathcal{E}(S)}\mu_e$, and the second term is the stationary probability of being in state S, which we denote by p_S.

Substituting all these results into Equation (4), we get the expected value of the SPA contribution:

$$\lim_{n\to\infty}\frac{t_n}{n}\sum_{L_m^*}p_{S(m)}\mu_{\alpha(m)}\mu_{\beta(m)}\left(E\left[G^{I(m)}\right]+E[\Delta G(m)]-E[n_{I(m)}]\bar{f}\right)E\left[\left(\frac{d\tau(m)}{d\theta}\right)^+_{IPA}\right].$$

The total estimator over n transitions is thus:

$$\left(\frac{\partial f}{\partial\theta}\right)_{IPA}+\frac{1}{n}\sum_{L_m^*}T_n(S(m))\mu_{\alpha(m)}\mu_{\beta(m)}\left(\widehat{G}^{I(m)}+\widehat{\Delta G}(m)-\widehat{n}_{I(m)}f_n\right)\left(\widehat{\frac{d\tau(m)}{d\theta}}\right)^+_{IPA}, \qquad (5)$$

where the "^" indicates an estimator of the corresponding expectation. Summarizing, for each critical adjacent pair, $L_m^*=(\alpha(m),\beta(m)|S(m))$, we accumulate the time spent in the state $S(m)$, the exponential rates of $\alpha(m)$ and $\beta(m)$ (which are known), and the IPA difference defined by Equation (2). The remaining expectations involving the subpath $I(m)$ can be directly estimated from a single simulation of the chain. The subpath is defined by the state reached from $S(m)$ after the occurrence of $\beta(m)$, then $\alpha(m)$, and an ending state defined by the state reached from $S(m)$ after the occurrence of $\beta(m)$, then $\alpha(m)$. The commuting condition of Glasserman (1991) requires these two states to be the same, so the terms involving $I(m)$ vanish. Furthermore, for the class of performance measures he considers, the $\Delta G(m)$ term also vanishes, so the SPA contribution is zero, i.e., the estimator requires only the IPA contribution.

Finally, we note that the expectations involving $I(m)$ can also in principle be solved by use of Poisson's equation, which calculates expected rewards acquired before hitting an absorbing set of states; here, the "final" state in $I(m)$ would serve as the absorbing state. However, this is in general probably an impractical solution for the same situations that one would simulate the chain in the first place: if the size of the state space is enormous. In the next section, we consider an example that illustrates an application of the estimator. For a simplified version of the example, we prove the strong consistency of our estimator by explicit analytic calculation of all the terms in the estimator. A sample path proof of strong consistency using regenerative techniques as in Fu and Hu (1993) can also be done, but will not be carried out here.

4 An Example: A Simple Multi-Class Queueing Network

To illustrate how one would apply the estimator, we consider a simple example: a closed queueing network with two servers and two classes of jobs, with N_1 class 1 jobs and N_2 class 2 jobs. The configuration of the system is as follows (refer to Figure 3). A class 1 job receives service at server 1 only (immediately) rejoining the queue at station 1 upon service completion at server 1, whereas a class 2 job receives service at both stations 1 and 2, going to station 2 upon service completion at station 1 and going to station 1 upon service completion at station 2. All service times are exponentially distributed, with λ the service rate (of both classes of jobs) at station 1 and μ the service rate at station 2. We consider the parameter $\theta = 1/\mu$ and the performance measure

$$\bar{f}(\theta) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \tau_i = \lim_{n\to\infty} \frac{t_n}{n}, \tag{6}$$

which is the average time between two (consecutive) service completions, i.e., the reciprocal of the throughput of the system.

For this particular performance measure, the IPA contribution is zero and $\Delta G = 0$, so Equation (5) becomes

$$\frac{1}{n}\sum_{L_m^*} T_n(S(m))\mu_{\alpha(m)}\mu_{\beta(m)} \left(\widehat{G}^{I(m)} - \widehat{n}_{I(m)} f_n\right) \left(\frac{d\widehat{\tau(m)}}{d\theta}\right)^+_{IPA}, \tag{7}$$

where $G^{I(m)} = \sum_{I(m)} \tau_i$, since $g(\cdot,\cdot,\tau_i) = \tau_i$ in Equation (6). To do the estimation, we need to determine the critical adjacent event pairs. There are only three distinct types of events – a class 1 or class 2 service completion at server 1 and a class 2 service completion at server 2 – and only at most two of them are possible at any one time, so there at most two potential critical adjacent event pairs associated with every state. For our performance measure, only one of the two is actually critical, the pair corresponding to a class 1 service completion at server 1 and a class 2 service completion at server 2. For the right-hand derivative represented by Equation (7) (corresponding to a positive perturbation in θ), the class 2 service completion at server 2 will be perturbed later in time, so the critical adjacent event pairs are of the form $(\alpha, \beta|S)$, where α is a service completion at station 2 and β is a service completion at station 1. Thus, our final estimator is

$$\frac{\lambda\mu}{n}\sum_{S(m)\in\tilde{S}} T_n(S(m)) \left(\widehat{G}^{I(m)} - \widehat{n}_{I(m)} f_n\right) \left(\frac{d\widehat{\tau(m)}}{d\theta}\right)^+_{IPA}, \tag{8}$$

where $\tilde{S} = \{$states with ≥ 1 customer at server 2 and a class 1 customer in service at server 1$\}$.

We now specialize to the case of one class 1 job and two class 2 jobs, i.e., $N_1 = 1$ and $N_2 = 2$, and prove that our estimator is correct. This particular network was also considered in Cao (1987), for which it was shown that IPA fails. We define a Markov chain representation for the system having six states:

$$s_1 = (1), s_2 = (1,2), s_3 = (2,1), s_4 = (1,2,2), s_5 = (2,1,2), s_6 = (2,2,1),$$

where the ith component of the row vector indicates the class of the ith job at server 1, the first job being in service and the others in queue. The transition rate diagram for this Markov chain representation is given in Figure 4.

As already noted, the IPA contribution is equal to zero. (Since $dt_n/d\theta$ is bounded by the length of the regenerative cycle, which has regenerative point s_6, divided by θ, this also follows directly from the proof of Lemma 1 in Glasserman et al., 1991.) Therefore, we can focus on the SPA contribution given by Equation (8). By definition, we have $\tilde{S} = \{S_1, S_2\}$, i.e., the only two critical adjacent event pairs are $(\alpha, \beta|s_1)$ and $(\alpha, \beta|s_2)$. Also, by definition, $I(1)$ is a path starting from s_2 and ending at the first hitting time of s_3, since s_2 is the state reached from s_1 after β, then α, and s_3 is the state reached from s_1 after α, then β (see Figure 4, where μ arcs correspond to event α and λ arcs correspond to event β). Similarly, $I(2)$ is a path from s_5 to s_6. Thus, the explicit estimator for this example is given by

$$\lambda\mu f_n \sum_{m=1}^{2} p_m \left(\left[\widehat{\sum_{I(m)} \tau_i} \middle| (\alpha,\beta|s_m) \right] - [\widehat{n}_I(m)| (\alpha,\beta|s_m)] f_n \right) \left(\frac{d\widehat{\tau(m)}}{d\theta} \right)^{+}_{IPA}. \tag{9}$$

We now proceed to show that this estimator is consistent.

Letting $p_i, i = 1,\ldots,6$, denote the stationary probabilities, we have

$$p_1 = \frac{\lambda^2}{\lambda^2+2\lambda\mu+3\mu^2}, \quad p_2 = p_3 = \frac{\lambda\mu}{\lambda^2+2\lambda\mu+3\mu^2}, \quad p_4 = p_5 = p_6 = \frac{\mu^2}{\lambda^2+2\lambda\mu+3\mu^2}. \tag{10}$$

The throughput of the system, defined as $\sum_n p_n \gamma_n$, where γ_n is the output rate in state n, is given by

$$\frac{\lambda(\lambda^2+3\lambda\mu+5\mu^2)}{\lambda^2+2\lambda\mu+3\mu^2},$$

and thus taking the reciprocal, we have

$$\bar{f}(\theta) = \frac{\lambda^2+2\lambda\mu+3\mu^2}{\lambda(\lambda^2+3\lambda\mu+5\mu^2)}.$$

Taking the derivative with respect to θ, we get

$$\frac{d\bar{f}(\theta)}{d\theta} = \frac{\mu^2(\lambda^2+4\lambda\mu+\mu^2)}{(\lambda^2+3\lambda\mu+5\mu^2)^2}. \tag{11}$$

We now show that our SPA estimator converges to the same result.

Equation (9) converges to

$$\lambda\mu\bar{f} \sum_{m=1}^{2} p_m \left(E\left[\sum_{I(m)} \tau_i \middle| (\alpha,\beta|s_m) \right] - E[n_I(m)| (\alpha,\beta|s_m)] \bar{f} \right) E\left[\left(\frac{d\tau(m)}{d\theta} \right)^{+}_{IPA} \right]. \tag{12}$$

We will need the transition probabilities for the embedded (discrete-time) Markov chain, which are given by the matrix

$$\mathbf{P} = [p_{ij}] = \begin{pmatrix} \lambda/(\lambda+\mu) & \mu/(\lambda+\mu) & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda/(\lambda+\mu) & \mu/(\lambda+\mu) & 0 & 0 \\ \lambda/(\lambda+\mu) & 0 & 0 & 0 & \mu/(\lambda+\mu) & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

We first calculate $E[\sum_I \tau_i]$ and $E[n_I]$ for $I(1)$, the path that starts at s_2 and ends at s_3. Starting at s_2, a sample path can end at s_3 following two different paths: s_2s_3 and $s_2s_4s_6s_3$. The probability that it follows the first path is $\lambda/(\lambda+\mu)$ and the second one is $\mu/(\lambda+\mu)$. Hence, we have

$$E\left[\sum_{i=1}^{n_I} \tau_i \middle| (\alpha,\beta|s_1) \right] = \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\frac{2}{\lambda} = \frac{\lambda+2\mu}{\lambda(\lambda+\mu)} \tag{13}$$

and

$$E[n_I| (\alpha,\beta|s_1)] = 1 + \frac{\mu}{\lambda+\mu}2 = \frac{\lambda+3\mu}{\lambda+\mu}. \tag{14}$$

For $I(2)$, in which the sample path starts at s_5 and ends at s_6, the calculation is more complicated than the previous one, since there are many different ways a sample path starting at s_5 can end

s_6. Let y_i and z_i denote $E[\sum_{i=1}^{n_I} \tau_i]$ and $E[n_I]$ for the sample path that starts at s_i and ends at s_6 ($i = 1, \ldots, 5$), where we are interested in finding y_5 and z_5. Let $\mathbf{y} = [\ y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6\ ]^T$ and $\mathbf{z} = [\ z_1 \quad z_2 \quad z_3 \quad z_4 \quad z_5 \quad z_6\ ]^T$. Then, both satisfy a Poisson's equation,

$$\begin{aligned} \mathbf{y} &= \mathbf{Qy} + \mathbf{a}, \\ \mathbf{z} &= \mathbf{Qz} + \mathbf{b}, \end{aligned}$$

where $\mathbf{Q}$ is the matrix obtained by deleting the 6th row and column from $\mathbf{P}$ (i.e., with states 1 through 5 taken as the set of transient states and state 6 as the absorbing set),

$$\mathbf{a} = [\ 1/(\lambda+\mu) \quad 1/(\lambda+\mu) \quad 1/(\lambda+\mu) \quad 1/\lambda \quad 1/\lambda\ ],$$

$$\text{and} \quad \mathbf{b} = [\ 1 \quad 1 \quad 1 \quad 1 \quad 1\ ].$$

Solving, we obtain

$$E[\sum_{I(2)} \tau_i] \ = \ y_5 \ = \ \frac{\lambda^3 + 2\lambda^2\mu + 4\lambda\mu^2 + 2\mu^3}{\lambda\mu^2(\lambda+\mu)} \tag{15}$$

$$E[n_{I(2)}] \ = \ z_5 \ = \ \frac{\lambda^3 + 3\lambda^2\mu + 6\lambda\mu^2 + 3\mu^3}{\mu^2(\lambda+\mu)} \tag{16}$$

Now we calculate the $E[(d\tau(m)/d\theta)^+_{IPA}]$ terms. Note that the perturbation at server 1 is always equal to zero since perturbations at server 2 never propagate to server 1 (recall that θ is a parameter in server 2's service time), because server 1 is always busy. Since θ is a scale parameter, the steady-state IPA accumulation at server 2 in $s_i, i = 1, \cdots, 6$, can be written as x_i/θ, where x_i has the dimension of time. We have $x_4 = x_5 = x_6 = 0$, because server 2 is idle in these states, so the perturbation is always lost. (NB: The IPA analyses contained in the previous three sentences are just an application of the original IPA for queueing networks introduced by Ho and Cao, 1983). Now, we condition on the previous state to derive recursive equations for the expected values of the remaining x_i:

$$\begin{aligned} E[x_1] &= p^*_{11}(E[x_1] + E[T_1]) + p^*_{13}(E[x_3] + E[T_1]) \\ E[x_2] &= p^*_{21}(E[x_1] + E[T_2]) \\ E[x_3] &= p^*_{32}(E[x_2] + E[T_3]), \end{aligned}$$

where T_i is the time spent in state s_i and p^*_{ij} is the probability that the previous state was j, given that the present state is i. For the former, we have $E[T_1] = E[T_2] = E[T_3] = 1/(\lambda+\mu)$, whereas the latter are simply the transition probabilities of the reversed chain given by $p^*_{ij} = p_{ji}p_j/p_i$, and thus, $p^*_{11} = p^*_{32} = p^*_{21} = \lambda/(\lambda+\mu), p^*_{12} = \mu/(\lambda+\mu)$; so solving the equations, we obtain

$$\frac{1}{\theta}(E[x_1] + \frac{1}{\lambda+\mu}) \ = \ \frac{\lambda^2 + 4\lambda\mu + 2\mu^2}{\mu(2\lambda+\mu)}; \tag{17}$$

$$\frac{1}{\theta}(E[x_2] + \frac{1}{\lambda+\mu}) \ = \ \frac{\lambda^2 + 3\lambda\mu + \mu^2}{\mu(2\lambda+\mu)}. \tag{18}$$

Substituting (10), (13), (14), (15), (16), (17) and (18), into Equation (12), we get

$$\bar{f}\lambda\mu \left[p_1(\frac{\lambda+2\mu}{\lambda(\lambda+\mu)} - \frac{\lambda+3\mu}{\lambda+\mu}\bar{f})\frac{1}{\theta}(E[x_1] + \frac{1}{\lambda+\mu}) + p_2(y_5 - z_5\bar{f})\frac{1}{\theta}(E[x_2] + \frac{1}{\lambda+\mu}) \right] = \frac{\mu^2(\lambda^2 + 4\lambda\mu + \mu^2)}{(\lambda^2 + 3\lambda\mu + 5\mu^2)^2},$$

which matches (11). Note that although the proof was done for a special case, the method of proof is applicable for any values of N_1 and N_2, as well as for more complicated multi-class queueing networks.

Acknowledgments

Jian-Qiang Hu was supported in part by the National Science Foundation (NSF) under grants DDM-9215369 and EID-9212122.

References

[1] Cao, X.R., "First Order Perturbation Analysis of a Simple Multiclass Finite Source Queue", *Performance Evaluation*, **7**, 31-41, 1987.

[2] Cao, X.R., "A Sample Performance Function of Closed Jackson Queueing Networks", *Operations Research*, **36**, 128-136, 1988.

[3] Fu, M.C. and Hu, J.Q., "Extensions and generalizations of smoothed perturbation analysis in a generalized semi-Markov process framework", *IEEE Transactions Automatic Control*, **37**, 1483-1500, 1992.

[4] Fu, M.C. and Hu, J.Q., "Smoothed perturbation analysis for queues with finite buffers", *Queueing Systems: Theory and Applications*, Vol.14, 57-78, 1993.

[5] Glasserman, P., "Structural Conditions For Perturbation Analysis Derivative Estimation: Finite-Time Performance Indices", *Operations Research*, **39**, 724-738, 1991.

[6] Glasserman, P., "Derivative Estimates from Simulation of Continuous-Time Markov Chains", *Operations Research*, **40**, 292-308, 1992.

[7] Glasserman, P., Hu, J.Q., and Strickland, S.G., "Strongly Consistent Steady-State Derivative Estimates", *Probability in the Engineering and Information Sciences,* **5**, 391-413, 1991.

[8] Gong, W.B., "Smoothed Perturbation Analysis of Markovian Queueing Networks", *Proceedings of the American Control Conference*, 456-461, 1988.

[9] Gong, W.B. and Ho, Y.C., "Smoothed Perturbation Analysis of Discrete-Event Dynamic Systems", *IEEE Transactions on Automatic Control*, **AC-32**, 858-867, 1987.

[10] Ho, Y.C. and Cao, X.R., "Optimization and Perturbation Analysis of Queueing Networks", *Journal of Optimization Theory and Applications,* **40**, 559-582, 1983.

[11] Whitt, W., "Continuity of Generalized Semi-Markov Processes", *Mathematics of Operations Research*, **5**, 494-501, 1980.