European Journal of Control (2010)6:642–661

DOI:10.3166/EJC.16.642–661

European Journal of Control

Perturbation Analysis and Optimization of Stochastic Hybrid Systems

Christos G. Cassandras[1,*], Yorai Wardi[2,**], Christos G. Panayiotou[3,***], Chen Yao[1,****]

[1] Division of Systems Engineering, and Center for Information and Systems Eng., Boston University, Brookline, MA 02446, USA;
[2] School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA, USA;
[3] KIOS Research Center for Intelligent, Systems and Networks, and Dept. of Electrical and Computer Eng., University of Cyprus, Cyprus

We present a general framework for carrying out perturbation analysis in Stochastic Hybrid Systems (SHS) of arbitrary structure. In particular, Infinitesimal Perturbation Analysis (IPA) is used to provide unbiased gradient estimates of performance metrics with respect to various controllable parameters. These can be combined with standard gradient-based algorithms for optimization purposes and implemented on line with little or no distributional information regarding the stochastic processes involved. We generalize an earlier concept of "induced events" for this framework to include system features such as delays in control signals or modeling multiple user classes sharing a resource. We apply this generalized IPA to two SHS with different characteristics. First, we develop a gradient estimator for the performance of a linear switched system with control signal delays and a safety constraint and show that it is independent of the random delay's distributional characteristics. Second, we derive closed-form unbiased IPA estimators for a Stochastic Flow Model (SFM) of systems executing tasks subject to either hard or soft real-time constraints. These estimators are incorporated in a gradient-based algorithm to optimize performance by controlling a task admission threshold parameter. Simulation results are included to illustrate this optimization approach.

**Correspondence to:* C.G. Cassandras, E-mail: cgc@bu.edu
** E-mail: wardi@ee.gatech.edu
*** E-mail: christosp@ucy.ac.cy
**** E-mail: cyao@bu.edu

Keywords: Stochastic Hybrid System, Stochastic Flow Model, Perturbation Analysis

1. Introduction

The study of hybrid systems is based on a combination of modeling frameworks originating in both time-driven and event-driven dynamic systems and resulting in *hybrid automata*. In a hybrid automaton, discrete events (either controlled or uncontrolled) cause transitions from one discrete state (or "mode") to another. While operating at a particular mode, the system's behavior is described by differential equations. In a stochastic setting, such frameworks are augmented with models for random processes that affect either the time-driven dynamics or the events causing discrete state transitions or both. A general-purpose stochastic hybrid automaton model may be found in [5] along with various classes of *Stochastic Hybrid Systems* (SHS) which exhibit different properties or suit different types of applications. The motivation for SHS models sometimes comes from time-driven systems whose behavior changes as a result of switching events and sometimes from the perspective of a Discrete Event System (DES) where the time between events may depend on one or more time-driven components. In some cases, DES become prohibitively complex and one resorts to SHS models through which the system dynamics are

Received 14 May 2009; Accepted 13 November 2009
Recommended by M. Prandini, J. Lygeros

abstracted to an appropriate level that retains essential features enabling effective and accurate control and optimization.

The performance of an SHS is generally hard to estimate because of the absence of closed-form expressions capturing the dependence of interesting performance metrics on design or control parameters. Consequently, we lack the ability to systematically adjust such parameters for the purpose of improving – let alone optimizing – performance. In the domain of DES, it was discovered in the early 1980s that event-driven dynamics give rise to state trajectories (sample paths) from which one can very efficiently and nonintrusively extract sensitivities of various performance metrics with respect to at least certain types of design or control parameters. This has led to the development of a theory for perturbation analysis in DES [4, 9, 12], the most successful branch of which is *Infinitesimal Perturbation Analysis* (IPA) due to its simplicity and ease of implementation. Using IPA, one obtains unbiased gradient estimates of performance metrics that can be incorporated into standard gradient-based algorithms for optimization purposes. However, IPA estimates become biased (hence unreliable for control purposes) when dealing with various aspects of DES that cause significant discontinuities in sample functions of interest. Such discontinuities normally arise when a parameter perturbation causes the order in which events occur to be affected and this event-order change may violate a basic "commuting condition" [9]. When this happens, one must resort to significantly more complicated methods for deriving unbiased gradient estimates [4].

In recent years, it was shown that IPA can also be applied to at least some classes of SHS and yield simple unbiased gradient estimators that can be used for optimization purposes. In particular, *Stochastic Flow Models* (SFMs), as introduced in [6], are a class of SHS where the time-driven component captures general-purpose flow dynamics and the event-driven component describes switches, controlled or uncontrolled, that alter the flow dynamics. Flow or fluid models are an example of an abstraction process applied to a large class of DES and especially useful in analyzing communication networks with large traffic volumes. Introduced in [1], fluid models have been shown to be very useful in simulating various kinds of high speed networks [16], manufacturing systems [7] and, more generally, settings where users compete over different sharable resources. It should be stressed that fluid models may not always provide accurate representations for the purpose of analyzing the performance of the underlying DES. What we are interested in, however, is control and optimization, in which case the value of a fluid model lies in capturing only those system features needed to design an effective controller that can potentially optimize performance without any attempt at estimating the corresponding optimal performance value with accuracy. While in most traditional fluid models the flow rates involved are treated as fixed parameters, an SFM has the extra feature of treating flow rates as *stochastic processes*. With virtually no limitations imposed on the properties of such processes, the use of IPA has been shown to provide simple gradient estimators for stochastic resource contention systems that include blocking phenomena and a variety of feedback control mechanisms [10, 11, 22, 23, 27–29].

Thus far, the use of IPA has been limited to the class of SFMs and normally exploits the special structure of specific systems investigated. Recently, a unified framework was introduced in [23], which provides the basis for a general theory of perturbation analysis for SHS. In this framework, the types of events causing discrete state transitions are classified as exogenous or endogenous. This, however, excludes interesting system features such as delays in control signals or modeling multiple user classes sharing a resource. The inclusion of such features requires a new class of "induced events" which can result in a (potentially infinite) event chain, a new phenomenon in the study of perturbation analysis. In [24], induced events were added to the original framework, albeit under some restrictive conditions. In this paper, we extend the framework in [24] to allow induced events of arbitrary generality, augmenting the continuous state with appropriate "timer" variables. Furthermore, the framework we present is placed in the more general context of stochastic hybrid automata with arbitrary structure.

In Section 2, we present the general framework for IPA in the setting of stochastic hybrid automata. We then apply IPA to two SHS. The first (Section 3), primarily used to illustrate the theory, is a linear switched system with a control signal issued at a specific system mode signaling a change in dynamics. The interesting feature is a random delay incurred before this control signal can take effect. The performance of the system is defined to represent a highly simplified tracking problem with a safety constraint and we derive an IPA estimator for an appropriate cost function. We will show that this estimator is in fact independent of the random delay distributional characteristics. The second SHS (Section 4) falls in the class of SFMs and addresses a crucial problem in time-critical systems. We consider a setting where tasks are executed subject to real-time constraints, i.e., tasks must not exceed a given "deadline" or must minimize the amount of time by which they exceed it. We derive IPA estimators in closed form for both types of objectives and also incorporate them into a gradient-based optimization algorithm to demonstrate their use.

2. A General Framework for Perturbation Analysis of Stochastic Hybrid Systems

We consider an SHS and adopt a standard hybrid automaton formalism to model its operation [5]. Thus, let $q \in Q$ (a countable set) denote the discrete state (or mode) and $x \in X \subseteq \mathbb{R}^n$ denote the continuous state. Let $\upsilon \in \Upsilon$ (a countable set) denote a discrete control input and $u \in U \subseteq \mathbb{R}^m$ a continuous control input. Similarly, let $\delta \in \Delta$ (a countable set) denote a discrete disturbance input and $d \in D \subseteq \mathbb{R}^p$ a continuous disturbance input. The state evolution is determined by means of (*i*) a vector field $f : Q \times X \times U \times D \to X$, (*ii*) an invariant (or domain) set $Inv : Q \times \Upsilon \times \Delta \to 2^X$, (*iii*) a guard set $Guard : Q \times Q \times \Upsilon \times \Delta \to 2^X$, and (*iv*) a reset function $r : Q \times Q \times X \times \Upsilon \times \Delta \to X$.

A sample path of such a system consists of a sequence of intervals of continuous evolution followed by a discrete transition. The system remains at a discrete state q as long as the continuous (time-driven) state x does not leave the set $Inv(q, \upsilon, \delta)$. If x reaches a set $Guard(q, q', \upsilon, \delta)$ for some $q' \in Q$, a discrete transition can take place. If this transition does take place, the state instantaneously resets to (q', x') where x' is determined by the reset map $r(q, q', x, \upsilon, \delta)$. Changes in υ and δ are discrete events that either *enable* a transition from q to q' by making sure $x \in Guard(q, q', \upsilon, \delta)$ or *force* a transition out of q by making sure $x \notin Inv(q, \upsilon, \delta)$. We will also use $\mathcal{E}$ to denote the set of all events that cause discrete state transitions and will classify events in a manner that suits the purposes of perturbation analysis.

In what follows, we describe the general framework for IPA presented in [24], casting it in the setting of stochastic hybrid automata, and enhance it by generalizing the definition of induced events used in [24]. Let $\theta \in \mathbb{R}^l$ be a global variable, henceforth called the *control parameter*, and suppose that $\theta \in \Theta$ for a given compact, convex set $\Theta \subset \mathbb{R}^l$. This may represent a system design parameter, a parameter of an input process, or a parameter that characterizes a policy used in controlling this system. The disturbance input $d \in D$ encompasses various random processes that affect the evolution of the state (q, x). We will assume that all such processes are defined over a common probability space, $(\Omega, \mathcal{F}, P)$. Let us fix a particular value of the parameter $\theta \in \Theta$ and study a resulting sample path of the SHS. Over such a sample path, let $\tau_k(\theta)$, $k = 1, 2, \ldots$, denote the occurrence times of the discrete events in increasing order, and define $\tau_0(\theta) = 0$ for convenience. We will use the notation τ_k instead of $\tau_k(\theta)$ when no confusion arises. The continuous state is also generally a function of θ, as well as of t, and is thus denoted by $x(\theta, t)$. Over an interval $[\tau_k(\theta), \tau_{k+1}(\theta))$, the system is at some mode during which the time-driven state satisfies:

$$\dot{x} = f_k(x, \theta, t) \tag{1}$$

where $\dot{x}$ denotes the time derivative $\frac{\partial x}{\partial t}$. Note that we suppress the dependence of f_k on the inputs $u \in U$ and $d \in D$ and stress instead its dependence on the parameter θ which may generally affect either u or d or both. The purpose of perturbation analysis is to study how changes in θ influence the state $x(\theta, t)$ and the event times $\tau_k(\theta)$ and, ultimately, how they influence interesting performance metrics which are generally expressed in terms of these variables.

The following assumption guarantees that (1) has a unique solution w.p.1 for a given initial boundary condition $x(\theta, \tau_k)$ at time $\tau_k(\theta)$:

Assumption: *W.p.1, there exists a finite non-empty set of points* $t_j \in [\tau_k(\theta), \tau_{k+1}(\theta)]$, $j = 1, 2, \ldots$, *which are independent of* θ, *such that, the function* f_k *is continuously differentiable on* $\mathbb{R}^n \times \Theta \times ([\tau_k(\theta), \tau_{k+1}(\theta)] \setminus \{t_1, t_2 \ldots\})$. *Moreover, there exists a random variable* $K > 0$ *such that* $E[K] < \infty$ *and the norm of the first derivative of* f_k *on* $\mathbb{R}^n \times \Theta \times ([\tau_k(\theta), \tau_{k+1}(\theta)] \setminus \{t_1, t_2 \ldots\})$ *is bounded from above by K.*

An event occurring at time $\tau_{k+1}(\theta)$ triggers a change in the mode of the system. A change in mode may also result in new dynamics represented by f_{k+1}, although this may not always be the case; for example, two modes may be distinct because the state $x(\theta, t)$ enters a new region where the system's performance is measured differently without altering its time-driven dynamics (i.e., $f_{k+1} = f_k$). The identity of f_{k+1} depends on the kind of event that occurs at time $t = \tau_{k+1}(\theta)$, and the associated event times play an important role in defining the interactions between the time-driven and event-driven dynamics of the system. In particular, the linearized system (defined below in detail) will be shown to depend on the derivatives of the event times with respect to the control parameters, i.e., the terms $\frac{d\tau_k}{d\theta}$.

We now classify events that define the set $\mathcal{E}$ as follows:

1. **Exogenous events.** An event is *exogenous* if it causes a discrete state transition at time τ_k independent of the controllable vector θ and satisfies $\frac{d\tau_k}{d\theta} = 0$. Exogenous events typically correspond to uncontrolled random changes in input processes.
2. **Endogenous events.** An event occurring at time τ_k is *endogenous* if there exists a continuously differentiable function $g_k : \mathbb{R}^n \times \Theta \to \mathbb{R}$ such that

$$\tau_k = \min\{t > \tau_{k-1} : g_k(x(\theta, t), \theta) = 0\} \tag{2}$$

The function g_k normally corresponds to a guard condition in a hybrid automaton model.

3. **Induced events.** An event at time τ_k is *induced* if it is triggered by the occurrence of another event at time $\tau_m \leq \tau_k$. The triggering event may be exogenous, endogenous, or itself an induced event. The events that trigger induced events are identified by a subset of the event set, $\mathcal{E}_I \subseteq \mathcal{E}$.

We point out that the event times are generally not continuous in θ for all $\theta \in \Theta$, since variations in θ may alter the sequence of events. However, we shall see that, under mild assumptions, for a given $\theta \in \Theta$, the derivatives $\frac{d\tau_k}{d\theta}$ exist w.p.1.

Consider a performance function of the control parameter θ:

$$J(\theta; x(\theta, 0), T) = E\left[\mathcal{L}(\theta; x(\theta, 0), T)\right]$$

where $\mathcal{L}(\theta; x(\theta, 0), T)$ is a sample function of interest evaluated in the interval $[0, T]$ with initial conditions $x(\theta, 0)$. For simplicity, we write $J(\theta)$ and $\mathcal{L}(\theta)$ and assume that $\mathcal{L}(\theta)$ has the form of a cost functional over the interval $[0, T]$. Suppose that there are N events (independent of θ) occurring during the time interval $[0, T]$ and define $\tau_0 = 0$ and $\tau_{N+1} = T$. Let $L_k : \mathbb{R}^n \times \Theta \times \mathbb{R}^+ \to \mathbb{R}$ be a function satisfying Assumption 1, henceforth labeled the *cost function*, and define $\mathcal{L}(\theta)$ by

$$\mathcal{L}(\theta) = \sum_{k=0}^{N} \int_{\tau_k}^{\tau_{k+1}} L_k(x, \theta, t)dt \qquad (3)$$

where we reiterate that $x = x(\theta, t)$ is a function of θ and t. We point out that the restriction of the definition of $J(\theta)$ to a finite horizon T is made merely for the sake of simplicity of exposition.

Given that we do not wish to impose any limitations (other than mild technical conditions) on the random processes that characterize the discrete or continuous disturbance inputs in our hybrid automaton model, it is infeasible to obtain closed-form expressions for $J(\theta)$. Therefore, for the purpose of optimization, we resort to iterative methods such as stochastic approximation algorithms (e.g., [15]) which are driven by estimates of the cost function gradient with respect to the parameter vector of interest. Thus, we are interested in estimating $\partial J/\partial\theta$ based on sample path data, where a sample path of the system may be directly observed or it may be obtained through simulation. We then seek to obtain θ^* minimizing $J(\theta)$ through an iterative scheme of the form

$$\theta_{n+1} = \theta_n - \eta_n H_n(\theta_n; x(\theta, 0), T, \omega_n), \quad n = 0, 1, \ldots \qquad (4)$$

where $H_n(\theta_n; x(0), T, \omega_n)$ is an estimate of $dJ/d\theta$ evaluated at θ and based on information obtained from a sample path denoted by ω_n and $\{\eta_n\}$ is an appropriately selected step size sequence. In order to execute an algorithm such as (4), we need the estimate $H_n(\theta_n)$ of $dJ/d\theta$. The IPA approach is based on using the sample derivative $d\mathcal{L}/d\theta$ as an estimate of $dJ/d\theta$. The strength of the approach is that $d\mathcal{L}/d\theta$ can be obtained from observable sample path data alone and, usually, in a very simple manner that can be readily implemented on line. Moreover, it is often the case that $d\mathcal{L}/d\theta$ is an *unbiased* estimate of $dJ/d\theta$, a property that allows us to use (4) in obtaining θ^*. We will return to this issue later, and concentrate first on deriving the IPA estimates $d\mathcal{L}/d\theta$.

The computation of the IPA derivative $d\mathcal{L}/d\theta$ can be obtained by taking formal derivatives with respect to θ in (1)–(3) as shown in [23]. We will review these results next and extend them by including the class of induced events as defined earlier. We point out that in the case of SFMs (mentioned in the introduction) flow dynamics render the computation of $d\mathcal{L}/d\theta$ very simple, enabling the derivation of results reported in the literature as in [6, 22, 27, 29]. We will also adhere to the convention that for a function $F : \mathbb{R}^n \to \mathbb{R}^m$, the derivative $\frac{dF}{dx}$ is an $m \times n$ matrix. In the particular case where $m = 1$, $\frac{dF}{dx}$ is a row, n-dimensional vector.

2.1. Infinitesimal Perturbation Analysis

Let us fix $\theta \in \Theta$, consider a particular sample path, and assume for the time being that all derivatives mentioned in the sequel do exist. To simplify notation, we define the following for all state and event time sample derivatives:

$$x'(t) \equiv \frac{\partial x(\theta, t)}{\partial \theta}, \qquad \tau_k' \equiv \frac{\partial \tau_k}{\partial \theta}, \quad k = 0, \ldots, N \qquad (5)$$

In addition, we will write $f_k(t)$ instead of $f_k(x, \theta, t)$ whenever no ambiguity arises. By taking derivatives with respect to θ in (1) on the interval $[\tau_k(\theta), \tau_{k+1}(\theta))$ we get

$$\frac{d}{dt}x'(t) = \frac{\partial f_k(t)}{\partial x}x'(t) + \frac{\partial f_k(t)}{\partial \theta} \qquad (6)$$

The boundary (initial) condition of this linear equation is specified at time $t = \tau_k$, and by writing (1) in an integral form and taking derivatives with respect to θ when $x(\theta, t)$ is continuous in t at $t = \tau_k$, we obtain for $k = 1, \ldots, N$:

$$x'(\tau_k^+) = x'(\tau_k^-) + \left[f_{k-1}(\tau_k^-) - f_k(\tau_k^+)\right]\tau_k' \qquad (7)$$

We note that whereas $x(\theta, t)$ is continuous in t, $x'(t)$ may be discontinuous in t at the event times τ_k, hence the left limit and right limit above are generally different. In addition, if $x(\theta, t)$ is not continuous in t at $t = \tau_k$ and the value of

$x(\tau_k^+)$ is determined by the reset function $r(q, q', x, \upsilon, \delta)$ discussed earlier, then

$$x'(\tau_k^+) = \frac{dr(q, q', x, \upsilon, \delta)}{d\theta} \tag{8}$$

Furthermore, once the initial condition $x'(\tau_k^+)$ is given, the linearized state trajectory $\{x'(t)\}$ can be computed in the interval $t \in [\tau_k(\theta), \tau_{k+1}(\theta))$ by solving (6) to obtain:

$$x'(t) = e^{\int_{\tau_k}^t \frac{\partial f_k(u)}{\partial x} du} \left[\int_{\tau_k}^t \frac{\partial f_k(v)}{\partial \theta} e^{-\int_{\tau_k}^t \frac{\partial f_k(u)}{\partial x} du} dv + \xi_k \right] \tag{9}$$

with the constant ξ_k determined from $x'(\tau_k^+)$ which is obtained from (7) since $x'(\tau_k^-)$ is the final-time boundary condition of the linearized system in the interval $[\tau_{k-1}(\theta), \tau_k(\theta))$; alternatively, it is obtained from (8).

Clearly, to complete the description of the trajectory of the linearized system (6)–(7), we have to specify the derivative τ_k' which appears in (7). Since τ_k, $k = 1, 2, \ldots$, are the mode-switching times, these derivatives explicitly depend on the interaction between the time-driven dynamics and the event-driven dynamics, and specifically on the type of event occurring at time τ_k. Using the event classification given earlier, we have the following.

1. **Exogenous events.** By definition, such events are independent of θ, therefore $\tau_k' = 0$.
2. **Endogenous events.** In this case, (2) holds and taking derivatives with respect to θ we get:

$$\frac{\partial g_k}{\partial x} \left[x'(\tau_k^-) + f_{k-1}(\tau_k^-) \tau_k' \right] + \frac{\partial g_k}{\partial \theta} = 0 \tag{10}$$

which can be rewritten as

$$\tau_k' = -\left[\frac{\partial g_k}{\partial x} f_{k-1}(\tau_k^-) \right]^{-1} \left(\frac{\partial g_k}{\partial \theta} + \frac{\partial g_k}{\partial x} x'(\tau_k^-) \right) \tag{11}$$

with $\frac{\partial g_k}{\partial x} f_{k-1}(\tau_k^-) \neq 0$.

3. **Induced events.** If an induced event occurs at $t = \tau_k$, the value of τ_k' depends on the derivative τ_m' where $\tau_m \leq \tau_k$ is the time when the associated triggering event takes place. The event induced at τ_m will occur at some time $\tau_m + \omega(\tau_m)$, where $\omega(\tau_m)$ is a random variable which is generally dependent on the continuous and discrete states $x(\tau_m)$ and $q(\tau_m)$ respectively. This implies the need for additional state variables, denoted by $y_m(\theta, t)$, $m = 1, 2, \ldots$, associated with events occurring at times τ_m, $m = 1, 2 \ldots$ The role of each such state variable is to provide a "timer" activated when a triggering event occurs. Recalling that triggering events are identified as belonging to a set $\mathcal{E}_I \subseteq \mathcal{E}$, let e_k denote the event occurring at τ_k and define

$$F_k = \{m : e_m \in \mathcal{E}_I,\ m \leq k\} \tag{12}$$

to be the set of all indices with corresponding triggering events up to τ_k. Omitting the dependence on θ for simplicity, the dynamics of $y_m(t)$ are then given by

$$\dot{y}_m(t) = \begin{cases} -C(t) & \tau_m \leq t < \tau_m + \omega(\tau_m),\ m \in F_m \\ 0 & \text{otherwise} \end{cases}$$

$$y_m(\tau_m^+) = \begin{cases} y_0 & y_m(\tau_m^-) = 0,\ m \in F_m \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where y_0 is an initial value for the timer $y_m(t)$ which decreases at a "clock rate" $C(t) > 0$ until $y_m(\tau_m + \omega(\tau_m)) = 0$ and the associated induced event takes place. Clearly, these state variables are only used for induced events, so that $y_m(t) = 0$ unless $m \in F_m$. The value of y_0 may depend on θ or on the continuous and discrete states $x(\tau_m)$ and $q(\tau_m)$, while the clock rate $C(t)$ may depend on $x(t)$ and $q(t)$ in general, and possibly θ. However, in most simple cases where we are interested in modeling an induced event to occur at time $\tau_m + \omega(\tau_m)$, we have $y_0 = \omega(\tau_m)$ and $C(t) = 1$, i.e., the timer simply counts down for a total of $\omega(\tau_m)$ time units until the induced event takes place. An example where y_0 in fact depends on the state $x(\tau_m)$ and the clock rate $C(t)$ is not necessarily constant arises in the case of multiclass resource contention systems as recently described in [25].

Henceforth, we will consider $y_m(t)$, $m = 1, 2, \ldots$, as part of the continuous state of the SHS and, similar to (5), we set

$$y_m'(t) \equiv \frac{\partial y_m(t)}{\partial \theta}, \quad m = 1, \ldots, N. \tag{14}$$

For the common case where y_0 is independent of θ and $C(t)$ is a constant $c > 0$ in (13), the following lemma facilitates the computation of τ_k' for an induced event occurring at τ_k.

Lemma 2.1: *If in (13) y_0 is independent of θ and $C(t) = c > 0$ (constant), then $\tau_k' = \tau_m'$.*

Proof: If an event occurs at $t = \tau_k$ induced by some event at $\tau_m \leq \tau_k$, then (2) holds with the switching function $y_m = 0$. Using (11) with $x = y_m, f_{k-1}(\tau_k^-) = -c$, we get $\tau_k' = y_m'(\tau_k^-)/c$.

At the triggering event time τ_m, we have $y_m(\tau_m^+) = y_0$. Differentiating with respect to θ gives

$$y_m'(\tau_m^+) + \frac{\partial y_m}{\partial t}(\tau_m^+)\tau_m' = 0$$

and since $\frac{\partial y_m}{\partial t} = -c$, it follows that $\tau_m' = y_m'(\tau_m^+)/c$. Since, from (13), no event in the interval $[\tau_m, \tau_k]$ has any effect on $y_m'(t)$, we conclude that $\tau_k' = \tau_m'$. ■

We point out that in [24] $\tau_k' = \tau_m'$ was used as the defining condition for induced events, thus limiting them to the special case satisfying the conditions of Lemma 2.1. The use of $y_m(t)$, $m = 1, \ldots, N$, allows us to generalize the class of induced events, imposing no limits on the ultimate relationship between τ_k' and τ_m'.

With the inclusion of the state variables $y_m(t)$, $m = 1, \ldots, N$, the derivatives $x'(t)$, τ_k', and $y_m'(t)$ can be evaluated through (6)–(11). In general, this evaluation is recursive over the event (mode switching) index $k = 0, 1, \ldots$. In some cases, however, it can be reduced to simple expressions, as seen in the analysis of many SFMs, e.g., [6].

Now the IPA derivative $d\mathcal{L}/d\theta$ can be obtained by taking derivatives in (3) with respect to θ:

$$\frac{d\mathcal{L}(\theta)}{d\theta} = \sum_{k=0}^{N} \frac{d}{d\theta} \int_{\tau_k}^{\tau_{k+1}} L_k(x, \theta, t)dt, \tag{15}$$

and applying the Leibnitz rule we obtain, for every $k = 0, \ldots, N$,

$$\begin{aligned}\frac{d}{d\theta} \int_{\tau_k}^{\tau_{k+1}} & L_k(x, \theta, t)dt \\ &= \int_{\tau_k}^{\tau_{k+1}} \left[\frac{\partial L_k}{\partial x}(x, \theta, t)x'(t) + \frac{\partial L_k}{\partial \theta}(x, \theta, t)\right] dt \\ &\quad + L_k(x(\tau_{k+1}), \theta, \tau_{k+1})\tau_{k+1}' - L_k(x(\tau_k), \theta, \tau_k)\tau_k'\end{aligned} \tag{16}$$

where $x'(t)$ and τ_k' are determined through (6)–(11). What makes IPA appealing, especially in the SFM setting, is the simple form the right-hand-side above often assumes.

We close this section with a comment on the unbiasedness of the IPA derivative $d\mathcal{L}/d\theta$. The IPA derivative of a sample function $\mathcal{L}(\theta)$ is statistically unbiased [4, 12] if, for every $\theta \in \Theta$,

$$E\left[\frac{d\mathcal{L}(\theta)}{d\theta}\right] = \frac{d}{d\theta}E[\mathcal{L}(\theta)] = \frac{dJ(\theta)}{d\theta}. \tag{17}$$

Obviously, without unbiasedness the use of IPA would come under question, and the main motivation for studying IPA in the SHS setting is that it yields unbiased derivatives for a large class of systems and performance metrics compared to the traditional DES setting [4]. The following conditions have been established in [21] (Lemma A2, page 70) as sufficient for the unbiasedness of IPA:

Proposition 1: *Suppose that the following conditions are in force: (i) For every $\theta \in \Theta$, the derivative $\frac{d\mathcal{L}(\theta)}{d\theta}$ exists w.p.1. (ii) W.p.1, the function $\mathcal{L}(\theta)$ is Lipschitz continuous on Θ, and the Lipschitz constant has a finite first moment. Fix $\theta \in \Theta$. Then, the derivative $\frac{dJ(\theta)}{d\theta}$ exists, and the IPA derivative $\frac{d\mathcal{L}(\theta)}{d\theta}$ is unbiased.*

The crucial assumption for Proposition 1 is the continuity of the sample performance function $\mathcal{L}(\theta)$; in fact, it is when discontinuities in $\mathcal{L}(\theta)$ arise that IPA becomes biased in many DES. In contrast, in many SHS (and SFMs in particular), such continuity is guaranteed in a straightforward manner. Differentiability w.p. 1 at a given $\theta \in \Theta$ often follows from mild technical assumptions on the probability law underlying the system, such as the exclusion of co-occurrence of multiple events. Lipschitz continuity of $\mathcal{L}(\theta)$ generally follows from upper boundedness of $|\frac{d\mathcal{L}(\theta)}{d\theta}|$ by an absolutely integrable random variable, generally a weak assumption. In light of these observations, the proofs of unbiasedness of IPA have become standardized and the assumptions in Proposition 1 can be verified fairly easily from the context of a particular problem.

In the next two sections, we apply the general IPA setting for SHS, and (6)–(11) in particular, to two problems. The first is a simple switched linear system with a parameterized controller to trigger one of the mode switches. The interesting feature is the presence of a random delay before this actual switch can take effect. Therefore, we are interested in controlling a parameter to ensure that some performance metric is minimized taking into account the statistical effects of this delay. As we will see, this involves the presence of induced events. We will also see that the resulting IPA estimator is independent of any distributional information regarding the random delays.

The second problem applies to an SFM for resource contention systems where time-critical tasks are executed and the objective is to minimize either the fraction of tasks that violate a given deadline constraint or the average "tardiness" resulting from such violations. This is a difficult problem that has attracted considerable attention in the fields of computer science, communication networks, and manufacturing, e.g., [2, 3, 8, 13, 14, 17–20, 26].

We will omit a discussion of unbiasedness of the IPA estimator derived for the first problem above. However, to demonstrate the use of Proposition 1, we shall explicitly study unbiasedness in the second problem (which is of considerable practical interest) by identifying the technical conditions required in that case and providing explicit proofs.

3. A Switched System with Control Action Delays

Consider a simple scalar linear system operating in two modes. In the first mode, the continuous state $x(t)$ evolves according to the dynamics

$$\dot{x} = ax(t) + \lambda(t) \tag{18}$$

where a is a given constant and $\lambda(t)$ is an input process, either known or some random process (noise) with unknown characteristics. In the second mode, the system dynamics become

$$\dot{x} = -u + \lambda(t) \tag{19}$$

where $u > 0$ is fixed and $\lambda(t)$ is the same process as before. The switch from the first to the second mode is intended to occur when the state first reaches from below a given value $b > 0$. However, because of a random delay following the switch command, the controller issues this command when $x(t)$ first reaches from below some θ such that $0 < \theta < b$. When this event occurs at time $t = \tau_k$, the associated delay is denoted by $\omega(\tau_k)$.

The objective of the system is to maintain $x(t)$ as close as possible to some target value $c > b$; however, there is a cost incurred whenever $x(t)$ is too close to c in the sense that $x(t) \geq b$. This is expressed as a cost function

$$\begin{aligned} J(\theta) &= E[\mathcal{L}(\theta)] \\ &= E\left[\int_0^T [(c - x(\theta,t)) + C \cdot \mathbf{1}[x(\theta,t) \geq b]]\, dt\right] \end{aligned} \tag{20}$$

where $C > 0$ is a given cost parameter and $\mathbf{1}[\cdot]$ is the usual indicator function. Note that we have explicitly written $x(\theta, t)$ to stress the dependence of the continuous state on the controllable parameter θ. Lastly, the system operates so that a switch from (19) to (18) takes place whenever $x(\theta, t)$ reaches 0 from above. We will assume that the initial condition is given as $x(\theta, 0) = 0$. i.e., the system starts in the first mode. We point out that this is a highly simplified version of a tracking problem (if c is replaced by a target trajectory $c(t)$) with a "safety requirement" to maintain a distance from the target trajectory which in this simple case is captured by penalizing the system whenever $x(\theta, t) \geq b$. We can think of a particle that repeatedly tries to get close to a point c while maintaining a safe distance from it and then returns to a rest position at $x = 0$ before repeating the process.

Figure 1 shows a hybrid automaton model for this system, consisting of four discrete states (modes). Invariant sets for each mode are indicated by conditions in brackets. The guard conditions causing discrete state transitions are shown on the transition arrows and ω is used to represent the event occurring after some random delay following the condition $x = \theta$. Note that the time-driven dynamics satisfy (18) in modes 1,2,3; the distinction is used to identify mode 3 when $[x \geq b]$, during which a cost is incurred as shown in (20), and also to identify mode 2 in which there is a pending control action which has yet to take effect

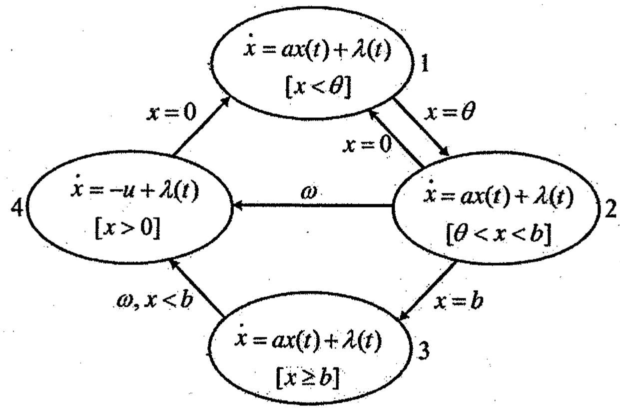


Fig. 1. Hybrid automaton for SHS with random control delays.

while $[\theta < x < b]$. Due to the possibly random process $\lambda(t)$, the ω event might take place before or after $[x \geq b]$ becomes true. If it happens after $[x \geq b]$ is true, then it is also possible that the guard condition $x < b$ is satisfied before ω occurs. The controller is designed so that if the transition due to the guard condition $x < b$ occurs prior to ω, then the ω event is disabled. This implies that no ω event is feasible in mode 4, in which the time-driven dynamics (19) are in effect. However, it is still possible that, while in mode 4, $x(\theta, t) \geq b$ due to $\lambda(t)$, which would incur some additional cost. For simplicity, we assume that $\lambda(t)$ is bounded (w.p. 1 if it is random) and that u is selected so that $-u + \lambda(t) < 0$, thus ensuring that no transition from 4 to 3 is feasible. Along the same lines, it is also possible that a transition from mode 2 to mode 1 takes place so that the ω event would occur in that mode; however, the controller is designed to disable the ω event if that were to happen.

Recalling the event classification of the previous section, we first observe that there are no exogenous events. On the other hand, there are three endogenous events: (*i*) The event corresponding to the guard condition $x = \theta$ being satisfied, (*ii*) The event corresponding to the guard condition $x = b$ being satisfied, and (*iii*) The event corresponding to the guard condition $x = 0$ being satisfied. Finally, we observe that an ω event is induced by the first of these three endogenous events. Therefore, we need to introduce additional state variables $y_m(t)$ (strictly speaking, $y_m(\theta, t)$) as in (13) with $C(t) = 1$ and $y_0 = \omega(\tau_m)$. The complete hybrid automaton model is then shown in Fig. 2, where the label ω is replaced by the guard condition $y_m = 0$ indicating that $y_m = 0$ for some $m = 1, 2, \ldots$ which was previously such that $y_m > 0$. Note, however, that when the transition from mode 3 to mode 4 occurs due to the condition $x < b$, the state variable $y_m(t)$ is reset from a positive value to 0. Similarly, if a transition from mode 2 to mode 1 occurs due to the condition $x = 0$, the state variable $y_m(t)$ is reset from a positive value to 0. This

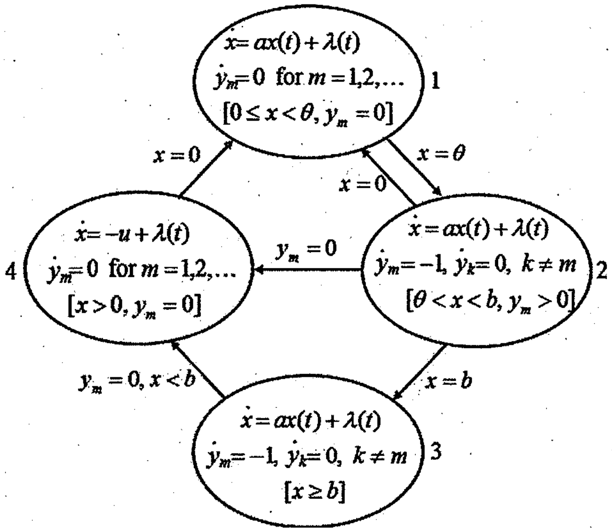


Fig. 2. Modified hybrid automaton for SHS with random control delays.

means that in this case (13) must be modified to include the reset function

$$r(q,q',y_m(\tau_k)) = \begin{cases} 0 & q=3,\ q'=4 \text{ or } q=2,\ q'=1 \\ y_m(\tau_k^-) & \text{otherwise} \end{cases} \tag{21}$$

The reset condition is not explicitly shown in Fig. 2, but is implicit in the invariant condition $y_m = 0$ active in modes 4 and 1.

Despite the operational simplicity of this system, we can see that the actual hybrid automaton that describes it is quite elaborate. We will see next, however, that the evaluation of the IPA derivatives $x'(t)$, $y_m'(t)$ and τ_k' for the state variables and events times as defined in (5) and (14) is relatively simple.

3.1. Infinitesimal Perturbation Analysis

We begin with the observation that a sample path of this SHS consists of a sequence of "cycles": each cycle starts at mode 1 and consists of one of three possible mode sequences: $(1,2)$, $(1,2,4)$ or $(1,2,3,4)$ before returning to mode 1. Consider a typical cycle and let τ_1, τ_2, τ_3, τ_4 denote the times when modes $1,\ldots,4$ respectively are entered (with the understanding that modes 3, 4 may not be entered at all in a cycle). The initial condition is $x(\tau_1^+)$, the state when the cycle starts; note that if $\tau_1 = 0$, then $x(\tau_1^+) = 0$ and $x'(\tau_1^+) = 0$. The cycle will complete at some time τ_f when a transition from mode 4 to 1 takes place, at which point the next cycle will start.

While in mode 1, (6) holds over the interval $[\tau_1, \tau_2)$ and we have

$$\frac{d}{dt}x'(t) = ax'(t) \tag{22}$$

which implies that

$$x'(t) = x'(\tau_1^+)e^{a(t-\tau_1)},\ t\in[\tau_1,\tau_2) \tag{23}$$

and, in particular,

$$x'(\tau_2^-) = x'(\tau_1^+)e^{a(\tau_2-\tau_1)} \tag{24}$$

Let us now consider all ensuing state transitions in the cycle.

1. At time τ_2 a transition from mode 1 to mode 2 takes place. This is due to an endogenous event with $g_2(x(\theta,\tau 2),\theta) = x(\tau_2) - \theta = 0$. Applying (11) we get

$$\tau_2' = \frac{1-x'(\tau_2^-)}{ax(\tau_2^-)+\lambda(\tau_2^-)} = \frac{1-x'(\tau_2^-)}{a\theta+\lambda(\tau_2^-)}$$

and using (24):

$$\tau_2' = \frac{1-x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta+\lambda(\tau_2^-)} \tag{25}$$

Since the $x(t)$ dynamics remain unchanged in this transition, using (7), we have $x'(\tau_2^+) = x'(\tau_2^-)$ and from (24):

$$x'(\tau_2^+) = x'(\tau_1^+)e^{a(\tau_2-\tau_1)} \tag{26}$$

On the other hand, using (13) we have $y_2(\tau_k^+) = \omega(\tau_2)$ and differentiating with respect to θ gives

$$y_2'(\tau_2^+) + \frac{\partial y_2}{\partial t}(\tau_2^+)\tau_2' = 0$$

where $\frac{\partial y_2}{\partial t}(\tau_2^+) = -1$ and we get $y_2'(\tau_2^+) = \tau_2'$. Combining this with (25) results in

$$y_2'(\tau_2^+) = \frac{1-x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta+\lambda(\tau_2^-)} \tag{27}$$

Finally, (22) again holds over the interval $[\tau_2,\tau_3]$ and we get, similar to (24),

$$x'(\tau_3^-) = x'(\tau_1^+)e^{a(\tau_3-\tau_1)} \tag{28}$$

In addition, applying (6) to $y_2'(t)$ over $[\tau_2,\tau_3)$, we have

$$y_2'(\tau_3^-) = y_2'(\tau_2^+) = \frac{1-x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta+\lambda(\tau_2^-)} \tag{29}$$

2. At time τ_3 a transition from mode 2 to mode 3 takes place. If this transition occurs, it is due to an endogenous event with $g_3(x(\theta,\tau_3),\theta) = x(\tau_3) - b = 0$. Applying (11) we get

$$\tau_3' = \frac{-x'(\tau_3^-)}{ax(\tau_3^-) + \lambda(\tau_3^-)} = \frac{-x'(\tau_3^-)}{ab + \lambda(\tau_3^-)}$$

and using (28):

$$\tau_3' = \frac{-x'(\tau_1^+)e^{a(\tau_3-\tau_1)}}{ab + \lambda(\tau_3^-)} \tag{30}$$

Since both the $x(t)$ and $y_2(t)$ dynamics remain unchanged in this transition, using (7) we have $x'(\tau_3^+) = x'(\tau_3^-)$ and $y_2'(\tau_3^+) = y_2'(\tau_3^-)$ and using (28) and (29):

$$x'(\tau_3^+) = x'(\tau_1^+)e^{a(\tau_3-\tau_1)} \tag{31}$$

$$y_2'(\tau_3^+) = \frac{1 - x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta + \lambda(\tau_2^-)}$$

Finally, (22) again holds over the interval $[\tau_3, \tau_4)$ and we get

$$x'(\tau_4^-) = x'(\tau_1^+)e^{a(\tau_4-\tau_1)} \tag{32}$$

and

$$y_2'(\tau_4^-) = y_2'(\tau_3^+) = \frac{1 - x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta + \lambda(\tau_2^-)}$$

3. At time τ_4 a transition into mode 4 takes place. This can happen in one of two ways, as described next.

First, a transition from mode 3 to mode 4 may take place due to the guard condition $x < b$. This is an endogenous event with $g_4(x(\theta,\tau_4),\theta) = x(\tau_4) - b = 0$, so, just as in the previous case, applying (11) we get:

$$\tau_4' = \frac{-x'(\tau_4^-)}{ax(\tau_4^-) + \lambda(\tau_4^-)} = \frac{-x'(\tau 4)}{ab + \lambda(\tau_4^-)}$$

and using (32):

$$\tau_4' = \frac{-x'(\tau_1^+)e^{a(\tau_4-\tau_1)}}{ab + \lambda(\tau_4^-)} \tag{33}$$

In this case, using (7), we get:

$$\begin{aligned} x'(\tau_4^+) &= x'(\tau_4^-) + \left[ax(\tau_4^-) + \lambda(\tau_4^-) + u - \lambda(\tau_4^+)\right]\tau_4' \\ &= x'(\tau_4^-) + \left[ab + \lambda(\tau_4^-) + u - \lambda(\tau_4^+)\right]\tau_4' \end{aligned}$$

Combining this with (32) and (33) we get

$$x'(\tau_4^+) = \frac{-u + \lambda(\tau_4^+)}{ab + \lambda(\tau_4^-)}x'(\tau_1^+)e^{a(\tau_4-\tau_1)} \tag{34}$$

In addition, due to the reset condition (21), we get $y_2(\tau_k^+) = 0$ and $y_2'(\tau_k^+) = 0$. Finally, while in mode 4, applying (6) over the interval $[\tau_4, \tau_f)$ we have $x'(\tau_4^+) = x'(\tau_f^-)$ and using (34):

$$x'(\tau_f^-) = \frac{-u + \lambda(\tau_4^+)}{ab + \lambda(\tau_4^-)}x'(\tau_1^+)e^{a(\tau_4-\tau_1)} \tag{35}$$

On the other hand, suppose the transition into mode 4 takes place due to the guard condition $y_2(\tau_4) = 0$ from either mode 2 or mode 3. This is an induced event which satisfies $g_4(y_4(\tau_4),\theta) = y_2(\tau_k) = 0$. Lemma 1 applies in this case and we immediately get

$$\tau_4' = \tau_2' = \frac{1 - x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta + \lambda(\tau_2^-)} \tag{36}$$

In this case, using (7), we get:

$$\begin{aligned} x'(\tau_4^+) &= x'(\tau_4^-) + \left[ax(\tau_4^-) + \lambda(\tau_4^-) + u - \lambda(\tau_4^+)\right]\tau_4' \\ &= x'(\tau_4^-) + \left[ax(\tau_4) + \lambda(\tau_4^-) + u - \lambda(\tau_4^+)\right]\tau_4' \end{aligned}$$

where the value of $x(\tau_4)$ is not specified by the guard condition as in previous cases. Using (32) and (36) and setting $\Delta\lambda(\tau_4) \equiv \lambda(\tau_4^-) - \lambda(\tau_4^+)$ (=0 when $\lambda(t)$ is continuous), we get

$$\begin{aligned} x'(\tau_4^+) &= x'(\tau_1^+)e^{a(\tau_4-\tau_1)} + [ax(\tau_4) + u + \Delta\lambda(\tau_4)] \\ &\quad \times \frac{1 - x'(\tau_1^+)e^{a(\tau_2-\tau_1)}}{a\theta + \lambda(\tau_2^-)} \\ &= \frac{[ax(\tau_4) + u + \Delta\lambda(\tau_4)]}{a\theta + \lambda(\tau_2^-)} + x'(\tau_1^+)e^{a(\tau_2-\tau_1)} \\ &\quad \times \left[e^{a(\tau_4-\tau_2)} - \frac{[ax(\tau_4) + u + \Delta\lambda(\tau_4)]}{a\theta + \lambda(\tau_2^-)}\right] \end{aligned} \tag{37}$$

In addition, $y_2(\tau_k^+) = 0$ and $y_2'(\tau_k^+) = 0$ still apply as before and $x'(\tau_4^+) = x'(\tau_f^-)$ which gives

$$\begin{aligned} x'(\tau_f^-) &= \frac{[ax(\tau_4) + u + \Delta\lambda(\tau_4)]}{a\theta + \lambda(\tau_2^-)} + x'(\tau_1^+)e^{a(\tau_2-\tau_1)} \\ &\quad \times \left(e^{a(\tau_4-\tau_2)} - \frac{[ax(\tau_4) + u + \Delta\lambda(\tau_4)]}{a\theta + \lambda(\tau_2^-)}\right) \end{aligned} \tag{38}$$

5. At time τ_f a transition into mode 1 takes place and the cycle is complete. This is an endogenous event with $g_f(x(\theta,\tau_f),\theta) = x(\tau_f) = 0$ that can occur in either mode 4 or mode 2. Applying (11) we get

$$\tau_f' = \frac{-x'(\tau_f^-)}{-u + \lambda(\tau_f^-)} \tag{39}$$

If the transition is from mode 2 to mode 1, then clearly $x'(\tau_f^+) = x'(\tau_f^-) = x'(\tau_2^+)e^{a(\tau_f-\tau_2)}$ and, recalling (26),

$$x'(\tau_f^+) = x'(\tau_f^-) = x'(\tau_1^+)e^{a(\tau_f-\tau_1)} \tag{40}$$

If the transition is from mode 4 to mode 1, then (7) implies that

$$\begin{aligned} x'(\tau_f^+) &= x'(\tau_f^-) + \left[-u + \lambda(\tau_f^-) - ax(\tau_f^+) - \lambda(\tau_f^+)\right]\tau_f' \\ &= x'(\tau_f^-) - \left[u + \lambda(\tau_f^+) - \lambda(\tau_f^-)\right]\tau_f' \end{aligned}$$

which, in view of (39), gives

$$x'(\tau_f^+) = \frac{\lambda(\tau_f^+)}{-u + \lambda(\tau_f^-)} x'(\tau_f^-) \tag{41}$$

where the value of $x'(\tau_f^-)$ depends on the actual mode sequence in the cycle and is given by either (35) or (38). Thus, in case the induced event never took place, i.e., mode sequence $(1, 2, 3, 4)$ took place and the transition into mode 4 was the result of the guard condition $x < b$, the next cycle completes with

$$x'(\tau_f^+) = \frac{\lambda(\tau_f^+)}{-u + \lambda(\tau_f^-)} \cdot \frac{-u + \lambda(\tau_4^+)}{ab + \lambda(\tau_4^-)} x'(\tau_1^+)e^{a(\tau_4-\tau_1)} \tag{42}$$

Otherwise, when the induced event takes place, the next cycle completes with

$$\begin{aligned} x'(\tau_f^+) &= \frac{\lambda(\tau_f^+)}{-u + \lambda(\tau_f^-)} \\ &\times \left[\frac{ax(\tau_4) + u + \Delta\lambda(\tau_4)}{a\theta + \lambda(\tau_2^-)} + x'(\tau_1^+)e^{a(\tau_2-\tau_1)}\right. \\ &\times \left.\left(e^{a(\tau_4-\tau_2)} - \frac{ax(\tau_4) + u + \Delta\lambda(\tau_4)}{a\theta + \lambda(\tau_2^-)}\right)\right] \end{aligned} \tag{43}$$

Given that the initial condition is $x'(0) = 0$, we can immediately see that until an induced event occurs in a cycle we have $x'(\tau_f^+) = 0$ and $\tau_f' = 0$. Once, however, the presence of an induced event results in some $x'(\tau_f^+) \neq 0$, the IPA derivatives for the continuous state and the event times remain nonzero in general. In addition, it is not hard to show that $x'(\tau_f^-)$ and τ_f' are non-negative as long as the cycle starts with $x'(\tau_1^+) \geq 0$. However, $x'(\tau_f^+)$ can become negative if $\lambda(\tau_f^+)$ is negative as seen in (41). It is also possible that, in the absence of any restriction on $\lambda(t)$, after entering mode 1 the state is such that $x(t) < 0$ and remains at that mode thereafter (if desired, this behavior can be prevented by modifying the dynamics in (18) so that $\dot{x} = 0$ if $x(t) = 0$ and $\lambda(t) < 0$).

It is also worth pointing out that the evaluation of the IPA derivatives τ_k' and $x'(t)$ above does not involve any knowledge of the random delay $\omega(\tau_k)$. It requires event time information, an observation $x(\tau_4)$ whenever mode 4 in entered, and observations of the process $\lambda(t)$ only at mode transition times τ_2, τ_3, τ_4, τ_f. Moreover, the use of the state variables $y_m(t)$ does not affect τ_k' and $x'(t)$ in this particular SHS because the conditions of Lemma 2.1 are satisfied.

Now let us return to the cost function $J(\theta)$ in (20) in order to determine the IPA derivative $d\mathcal{L}/d\theta$. Decomposing a sample path into cycles indexed by $k = 1, 2, \ldots$, let $\tau_{k,i}$ denote the ith event in the kth cycle. Let us also define

$$\begin{aligned} \Omega = \{k : x(\theta, t) \geq b,\ t \in [\tau_{k,3}, \tau_{k,4}], \\ \tau_{k,4} > \tau_{k,3},\ k = 1, \ldots, N\} \end{aligned}$$

i.e., the subset of cycles which contain a strictly positive interval during which the system is in mode 3. Then, observe that a sample function $\mathcal{L}(\theta)$ can be written as

$$\mathcal{L}(\theta) = \sum_{k=1}^{N}\sum_{i=1}^{n_k}\int_{\tau_{k,i}}^{\tau_{k,i+1}} (c - x(\theta, t))dt + C\sum_{k\in\Omega}(\tau_{k,4} - \tau_{k,3})$$

where N is the (sample path dependent) number of cycles in the interval $[0, T]$, including a possibly incomplete last one. Differentiating with respect to θ we get

$$\begin{aligned} \frac{d\mathcal{L}(\theta)}{d\theta} &= \sum_{k=1}^{N}\sum_{i=1}^{n_k}\left[(c - x(\tau_{k,i+1}))\tau_{k,i+1}'\right. \\ &\quad \left. - (c - x(\tau_{k,i}))\tau_{k,i}' - \int_{\tau_{k,i}}^{\tau_{k,i+1}} x'(t)dt\right] \\ &\quad + C\sum_{k\in\Omega}(\tau_{k,4}' - \tau_{k,3}') \end{aligned} \tag{44}$$

In the second term, $\tau_{k,3}'$ is given by (30) with the value of $x'(\tau_{k,1}^+)$ obtained at the start of the kth cycle. On the other hand, $\tau_{k,4}'$ is given by either (33) if the transition to mode 4 is caused by the guard condition $x < b$ or by (36) if it is caused by an induced event.

In the first term of (44), the values of $\tau_{k,i+1}'$ and $\tau_{k,i}'$ are given by the event time derivatives (25), (30), (33), (36), (39) depending on the event observed. The values of $x(\tau_{k,i+1})$ and $x(\tau_{k,i})$ are directly observable and, except for induced events, coincide with 0, θ, or b. Finally, the integral terms are evaluated as follows. For $(\tau_{k,1}, \tau_{k,2})$, (23) applies and we have

$$\int_{\tau_{k,1}}^{\tau_{k,2}} x'(t)dt = x'(\tau_{k,1}^+) \int_{\tau_{k,1}}^{\tau_{k,2}} e^{a(t-\tau_{k,1})}dt$$

$$= \frac{x'(\tau_{k,1}^+)}{a}\left[e^{a(\tau_{k,2}-\tau_{k,1})} - 1\right]$$

Since the $x(t)$ dynamics are unaffected in $(\tau_{k,2}, \tau_{k,3})$ and $(\tau_{k,3}, \tau_{k,4})$ (if mode 3 is entered), the expression above remains the same until mode 4 is entered, with $\tau_{k,1}, \tau_{k,2}$ replaced by the corresponding event times. For $(\tau_{k,4}, \tau_{k,5})$ where $x'(t)$ remains unchanged from $x'(\tau_{k,4}^+)$,

$$\int_{\tau_{k,4}}^{\tau_{k,5}} x'(t)dt = x'(\tau_{k,4}^+)(\tau_{k,5} - \tau_{k,4})$$

where $x'(\tau_{k,4}^+)$ is given by either (34) or (37) depending on whether mode 4 is entered as a result of an induced event or not.

We note that the cost term $(c - x(\theta, t))$ may be replaced by a quadratic cost $\frac{1}{2}(c - x(\theta, t))^2$ without affecting any of our analysis except for the fact that the integral terms in (44) would require integrating the process $\lambda(t)$. If this is a random process, then the evaluation of these integrals would require sample path observations of $\lambda(t)$ and some numerical integration.

4. A System with Time-Critical Tasks

In this section we consider a resource contention system, modeled as a finite capacity queue, where time critical tasks are executed. The objective is to optimally trade off a cost due to tasks that are blocked because there are no resources available (they find the queue full) against the delay that an accepted task will experience. Specifically, we assume that each task has a deadline by which it has to be completed and investigate two types of deadlines. A hard deadline where if the delay exceeds the deadline then a fixed penalty is paid and a soft deadline where the penalty is proportional to the duration of the deadline violation. The control parameter of interest is the size of the queue denoted by θ. We point out that increasing θ decreases the cost due to blocked tasks; however, it also allows the buffer to grow larger which increases the delay experienced by the tasks. For this problem, we adopt the SFM framework [6] to derive the IPA estimators; however, we emphasize that the resulting IPA algorithm is computed using data observed from the actual DES (see Section 4.3 for more details).

Next, we present the adopted SFM which is also shown in Fig. 3. In this context, $\alpha(t)$ is the fluid arrival rate and $\beta(t)$ is the maximum rate by which fluid is discharged.

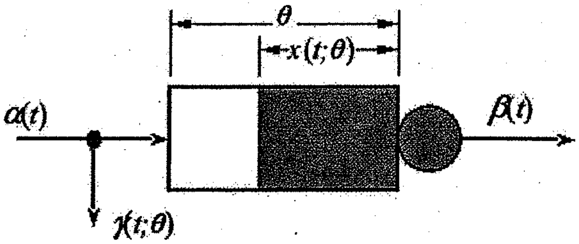


Fig. 3. The SFM System Model.

These processes are very general except for two mild technical assumptions that will be described later. The amount of fluid in the buffer is denoted by $x(\theta, t) \in [0, \theta], \theta \in \mathbb{R}^+$ and its dynamics are shown below:

$$\frac{dx(\theta, t)}{dt^+} = \begin{cases} 0, & \text{if } x(\theta,t) = 0 \text{ and } \alpha(t) - \beta(t) \leq 0, \\ 0, & \text{if } x(\theta,t) = \theta \text{ and } \alpha(t) - \beta(t) \geq 0, \\ \alpha(t) - \beta(t), & \text{otherwise} \end{cases} \tag{45}$$

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

$$\gamma(\theta, t) = \begin{cases} \alpha(t) - \beta(t), & \text{if } x(\theta,t) = \theta, \\ 0, & \text{if } x(\theta,t) < \theta. \end{cases} \tag{46}$$

Finally, we point out that a particle that arrives in the system at time t will depart from the system at time $t + \omega(\theta, t)$, where $\omega(\theta, t)$ is the time that the system needs to service all fluid that arrived up until t. If $x(\theta, t) = 0$, and $\alpha(t) - \beta(t) < 0$ then the particle will depart immediately, i.e., $\omega(\theta, t) = 0$. On the other hand, if $x(t) > 0$, then the particle will depart as soon as all fluid that has accumulated in the buffer exits. In other words, the following relation should hold

$$\int_t^{t+\omega(\theta,t)} \beta(\tau)d\tau = x(\theta, t). \tag{47}$$

Next, we define the objective function that we aim to optimize

$$J(\theta) = E[L(\theta) + C \cdot D(\theta)] \tag{48}$$

where $L(\theta)$ is the average blocked fluid due to overflow, $D(\theta)$ is the average delay deadline violation cost and C is an appropriate weight between the two. The loss volume is given by

$$L(\theta) = \frac{1}{T}\int_0^T \gamma(\theta, t)dt.$$

For the delay deadline violation cost we consider two possibilities; a *hard* deadline ($D_1(\theta)$) and a *soft* deadline ($D_2(\theta)$) as shown below.

$$D_1(\theta) = \frac{1}{T}\int_0^T \mathbf{1}[\omega(\theta,t) > d(t)]\,dt, \tag{49}$$

$$D_2(\theta) = \frac{1}{T}\int_0^T [\omega(\theta,t) - d(t)]^+\,dt, \tag{50}$$

where $d(t)$ denotes the deadline applicable to particles that arrive at time t (independent of the parameter θ) and $[y]^+ = \max\{0,y\}$.

To complete the overall system dynamics observe that both $D_1(\theta)$ and $D_2(\theta)$ are functions of $\omega(\theta,t)$ which suggests that $\omega(\theta,t)$ must be treated as an additional state variable. To derive its dynamics, we will use the definition of $\omega(\theta,t)$ in (47). We will adopt the same notation as in previous sections and set $\dot{\omega} = \frac{\partial\omega}{\partial t}$. We will also write $\omega(t)$ instead of $\omega(\theta,t)$ for ease of notation. Differentiating both sides of (47) with respect to t we get:

$$(\dot{\omega}+1)\cdot\beta\,(t+\omega(t)) - \beta\,(t) = \dot{x}$$

If $0 < x(t) < \theta$, then $\dot{x} = \alpha(t) - \beta(t)$ and it follows that

$$\dot{\omega} = \frac{\alpha(t) - \beta\,(t+\omega(t))}{\beta\,(t+\omega(t))}$$

On the other hand, if $x(t) = \theta$, then $\dot{x} = 0$ and we get $\dot{\omega} = \beta(t+\omega(t))$. Finally, if $x(t) = 0$, then $\omega(t) = 0$ and $\dot{\omega} = 0$. To summarize

$$\dot{\omega} = \begin{cases} 0 & \text{if } x(t) = 0 \\ \dfrac{\beta(t) - \beta(t+\omega(t))}{\beta(t+\omega(t))} & \text{if } x(t) = \theta \\ \dfrac{\alpha(t) - \beta\,(t+\omega(t))}{\beta\,(t+\omega(t))} & \text{otherwise.} \end{cases} \tag{51}$$

Even though the dynamics of $\omega(t)$ are complicated, we will see that we can still derive simple expressions for the IPA derivatives $dD_1(\theta)/d\theta$ and $dD_2(\theta)/d\theta$.

The problem as described above fits the general SHS framework introduced in Section 2. Figure 4 shows a hybrid automaton model for this system, consisting of four discrete states (modes). Invariant sets for each mode are indicated by conditions in brackets and the guard conditions are shown on the transition arrows. In this SFM, there are several exogenous events corresponding to changes in the external processes $\{\alpha(t)\}$, $\{\beta(t)\}$, and $\{d(t)\}$. In particular, the transition from mode 1 to mode 2 when the buffer

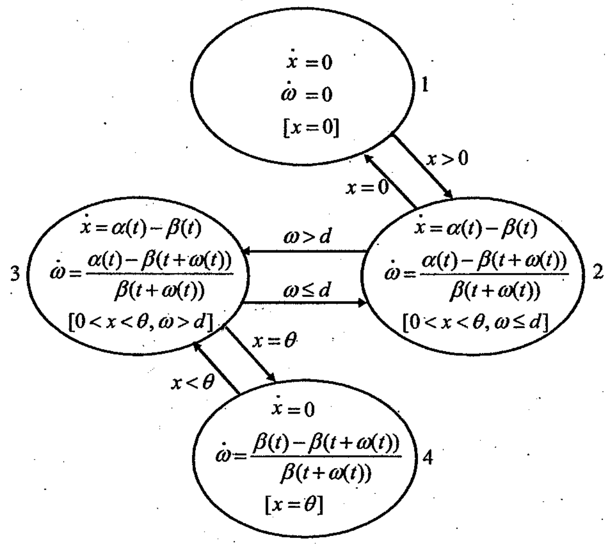


Fig. 4. Hybrid automaton model for the system of Fig. 3.

ceases to be empty is an exogenous event corresponding to a change in the sign of $\alpha(t) - \beta(t)$ from non-positive to positive; similarly, the transition from mode 4 to mode 3 when the buffer ceases to be full is an exogenous event corresponding to a change in the sign of $\alpha(t) - \beta(t)$ from non-negative to negative. The transition from mode 2 to mode 3 (or mode 3 to mode 2) may be exogenous if it is caused by a discontinuity in $d(t)$ causing $\omega(t) - d(t)$ to switch from non-positive to positive (or from positive to non-positive). On the other hand, it may be caused by the evolution of $\omega(t)$ hitting the switching function $\omega(t) = d(t)$, in which case it is an endogenous event. Additional endogenous events are: (*i*) The event corresponding to the guard condition $x = \theta$ being satisfied, (*ii*) The event corresponding to the guard condition $x = 0$ being satisfied. We also note that this SFM contains no induced events.

A typical sample path of this system is shown in Fig. 5 (when $\beta(t) = \beta$ and $d(t) = d$). As already pointed out, the events "buffer ceases to be empty or full" seen in this sample path are exogenous; these correspond to transitions from mode 1 to 2 and from mode 4 to 3 in Fig. 4. On the other hand, the events "buffer becomes empty or full" are endogenous and depend on θ; these correspond to transitions from mode 3 to 1 and from mode 3 to 4 in Fig. 4. It is convenient to partition a sample path into B_T Non-Empty Periods (NEP), i.e., periods where $x(\theta,t) > 0$. The kth NEP (denoted as $\mathcal{P}_k$) starts with the event "buffer ceases to be empty" at time $t = v_{k,0}$ and ends with the event "buffer becomes empty" at time $t = v_{k,S_k}$, i.e., $\mathcal{P}_k = [v_{k,0}, v_{k,S_k})$. In the kth NEP we also identify another $(S_k - 1)$ points $v_{k,j}, j = 1,\cdots,S_k-1$ where the buffer either becomes full or ceases to be full. Note that $v_{k,j}$ identifies the event "buffer becomes full", if $j = 2i+1$,

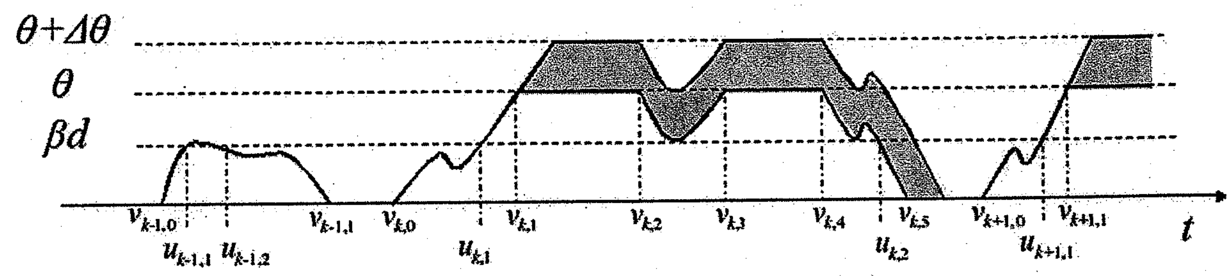


Fig. 5. Typical system sample path.

where $i = 0, \cdots, \frac{S_k - 1}{2}$ i.e., j is odd. Also, if $j = 2i$, i.e., j is even, $v_{k,j}$ identifies the event "buffer ceases to be full". For completeness, we also define the empty periods as the maximal intervals such that $x(\theta, t) = 0$. In addition, we define boundary periods as the intervals during which $x(\theta, t)$ is constant and equal to a boundary value, i.e., either empty ($x(\theta, t) = 0$) which correspond to the intervals $(v_{k,S_k}, v_{k+1,0})$ or full ($x(\theta, t) = \theta$) which correspond to the intervals $(v_{k,2i-1}, v_{k,2i})$, where $i = 1, \cdots, \frac{S_k - 1}{2}$. Similarly, non-boundary periods are all periods such that $0 < x(\theta, t) < \theta$.

Next, we define the threshold function $y(t)$,

$$y(t) = \int_t^{t+d(t)} \beta(\tau) d\tau. \tag{52}$$

Thus, a fluid molecule that entered the buffer at t will experience a delay greater than its deadline $d(t)$ if $x(\theta, t) > y(t)$ (see (47)). Figure 5 shows an example where $\beta(t) = \beta$ and $d(t) = d$ are constant for all t, thus $y(t) = \beta d$, constant for all t. Observe that the condition "$\omega > d$" in Fig. 4 is equivalent to the condition "$x > y$".

In addition, we define the sequence of event times $\{u_{k,j}\}$ such that the automaton of Fig. 4 switches between modes $\mathcal{M} = 2$ and $\mathcal{M} = 3$ for the jth time, during the kth NEP (see Fig. 5). During the kth NEP, this event may occur M_k times, i.e., $j = 1, \cdots, M_k$ and by definition we let $u_{k,0} = v_{k,0}$ and $u_{k,M_k+1} = v_{k,S_k}$, with the understanding that it is possible to have $M_k = 0$. Note that during the intervals $(u_{k,2n-1}, u_{k,2n})$, $n = 1, \cdots, \frac{M_k}{2}$ the automaton is either in mode 3 or 4 and the condition "$\omega(t) > d(t)$" is satisfied. Using the above notation, we can rewrite sample functions $D_1(\theta)$ and $D_2(\theta)$ as follows

$$D_1(\theta) = \frac{1}{T} \sum_{k=1}^{B_T} \sum_{n=1}^{(M_k-1)/2} \left(u_{k,2n} - u_{k,2n-1} \right). \tag{53}$$

$$D_2(\theta) = \frac{1}{T} \sum_{k=1}^{B_T} \sum_{n=1}^{(M_k-1)/2} \int_{u_{k,2n-1}}^{u_{k,2n}} (\omega(t) - d(t))\, dt. \tag{54}$$

Finally, we make the following assumptions regarding the processes $\alpha(t)$, $\beta(t)$ and $d(t)$.

Assumption A1: *With probability 1, $\alpha(t)$, $\beta(t)$ and $d(t)$ are piecewise constant functions independent of θ with $\alpha(t) < \infty$, $\beta(t) < \infty$, and $d(t) < \infty$.*

Assumption A2: *With probability 1, the processes $\alpha(t)$, $\beta(t)$, and $d(t)$ do not experience any discontinuity at exactly the same time.*

By Assumption **A1**, we represent a flow rate as a piecewise continuous process, which allows us to approximate it with an arbitrarily large number of constant segments. Assumption **A2** is a purely technical one allowing two events to occur concurrently with probability 0. Even if it is violated, one can perform IPA using one-sided derivatives. Given **A1** and **A2**, the process $y(t)$ is piecewise continuous with $0 \le y(t) < \infty$. The discontinuities of $\beta(t)$ are not inherited by $y(t)$ due to the integral in (52); however, $y(t)$ may inherit the discontinuities of $d(t)$.

4.1. Infinitesimal Perturbation Analysis

As in the previous section, we shall use the notation:

$$x'(t) \equiv \frac{\partial x(\theta, t)}{\partial \theta}, \quad \omega'(t) \equiv \frac{\partial \omega(\theta, t)}{\partial \theta},$$

$$u'_{k,j} \equiv \frac{du_{k,j}}{d\theta}, \quad v'_{k,j} \equiv \frac{dv_{k,j}}{d\theta}.$$

Lemma 4.1: *At time $t = u_{k,j}$, the sample derivative of the delay with respect to θ, $\omega'(t)$, is given by*

$$\omega'(t) = \frac{x'(t)}{\beta(t + \omega(t))}.$$

Proof: At $t = u_{k,j}$, (47) holds with $0 < x(t) < \theta$. Taking derivatives with respect to θ in (47), we get

$$\omega'(t)\beta\left(t + \omega(t)\right) + \int_t^{t+\omega(\theta,t)} \frac{\partial \beta}{\partial \theta} d\tau = x'(t)$$

and since $\beta(t)$ is independent of θ, it follows that

$$\omega'(t) = \frac{x'(t)}{\beta(t+\omega(t))}$$

yielding the desired result. ∎

Let us now consider all possible transitions in the kth NEP of a typical sample path. Clearly, exogenous events can have no effect on state variable or event time derivatives, so we concentrate on the three possible endogenous events.

1. At time $u_{k,j}$ a transition from mode 2 to mode 3 (or from mode 3 to mode 2) takes place. If the transition is not due to a discontinuity in $d(t)$ (which would be an exogenous event), then it is due to an endogenous event with switching function $\omega(u_{k,j})-d=0$. Applying (11) we get

$$u'_{k,j} = -\left[\frac{\alpha(u^-_{k,j}) - \beta\left(u^-_{k,j}+d(u^-_{k,j})\right)}{\beta\left(u^-_{k,j}+d(u^-_{k,j})\right)}\right]^{-1}\omega'(u^-_{k,j}).$$

Using Lemma 4.1 and recalling Assumption **A2**, we get

$$u'_{k,j} = \frac{-x'(u^-_{k,j})}{\alpha(u_{k,j}) - \beta\left(u_{k,j}+d(u_{k,j})\right)}. \tag{55}$$

Since the $x(t)$ and $\omega(t)$ dynamics remain unchanged in this transition, using (7), we have $x'(u^+_{k,j}) = x'(u^-_{k,j})$ and $\omega'(u^+_{k,j}) = \omega'(u^-_{k,j})$.

2. At time $v_{k,j}$ with $j=2i+1, i=0,\cdots,\frac{S_k-1}{2}$ a transition from mode 3 to mode 4 takes place with switching function $x(v_{k,j})-\theta=0$. Applying (11) we get

$$v'_{k,j} = -\frac{x'(v^-_{k,j})-1}{\alpha(v^-_{k,j})-\beta(v^-_{k,j})}.$$

In addition, applying (7) gives

$$x'(v^+_{k,j}) = x'(v^-_{k,j}) + \left[\alpha(v^-_{k,j})-\beta(v^-_{k,j})\right]v'_{k,j}$$

and combining the two equations implies that

$$x'(v^+_{k,j}) = 1. \tag{56}$$

3. At time v_{k,S_k} a transition from mode 2 to mode 1 takes place with switching function $x(v_{k,S_k})=0$. Applying (11) we get

$$v'_{k,S_k} = \frac{-x'(v^-_{k,S_k})}{\alpha(v^-_{k,S_k})-\beta(v^-_{k,S_k})}.$$

In addition, applying (7) gives

$$x'(v^+_{k,S_k}) = x'(v^-_{k,S_k}) + \left[\alpha(v^-_{k,S_k})-\beta(v^-_{k,S_k})\right]v'_{k,S_k}$$

and combining the two equations implies that

$$x'(v^+_{k,S_k}) = 0. \tag{57}$$

It follows from (56) and (57) that $x'(t)$ always resets to 0 at the end of a NEP and then switches to 1 with every "buffer becomes full" event if one occurs during a NEP.

Let us define

$$u^*_{k,j} = \max\{v_{k,2i} : v_{k,2i} < u_{k,j},\quad k=1,2,\cdots, i=0,1,\cdots, j=1,2,\cdots\}$$

to be the most recent event "buffer ceases to be either full or empty" that occurred just before $t=u_{k,j}$. We can then establish the following result.

Lemma 4.2: *Let S_T be the finite set of all time instants where $d(t)$ experiences a discontinuity in the interval $[0,T]$. Then*

$$u'_{k,j} = \begin{cases} \dfrac{-1}{\alpha(u_{k,j})-\beta(u_{k,j}+d(u_{k,j}))} & \text{if } x(u^*_{k,j})=\theta,\ u_{k,j}\notin S_T \\ 0 & \text{otherwise} \end{cases} \tag{58}$$

Proof: First, as already pointed out, if $u_{k,j}\in S_T$ then the event at $u_{k,j}$ is exogenous, hence $u'_{k,j}=0$. Next, consider (55). There are two cases: (*i*) If $u^*_{k,j}$ is such that $x(u^*_{k,j})=0$, then by (57) we have $x'(u^*_{k,j})=0$, therefore, $u'_{k,j}=0$, and (*ii*) If $u^*_{k,j}$ is such that $x(u^*_{k,j})=\theta$, then by (56) we have $x'(u^*_{k,j})=1$, therefore, $u'_{k,j}=-1/[\alpha(u_{k,j})-\beta(u_{k,j}+d(u_{k,j}))]$ and the result follows. ∎

Now we are ready to derive the IPA estimators for the performance metrics of interest $D_1(\theta)$ and $D_2(\theta)$. For the IPA estimator of $L(\theta)$ the reader is referred to [6].

Lemma 4.3: *The sample derivative of $D_1(\theta)$ with respect to θ is given by*

$$D'_1(\theta) = \sum_{k=1}^{B_T}\sum_{n=1}^{(M_k-1)/2}\left[-\frac{\mathbf{1}[x(u^*_{k,2n})=\theta]}{\alpha(u_{k,2n})-\beta(u_{k,2n}+d(u_{k,2n}))} + \frac{\mathbf{1}[x(u^*_{k,2n-1})=\theta]}{\alpha(u_{k,2n-1})-\beta(u_{k,2n-1}+d(u_{k,2n-1}))}\right]$$

The proof follows easily by differentiating (53) and then substituting the result from Lemma 4.2. Next, the question that arises is how one can compute $D_1'(\theta)$. Assuming that one has a way of estimating the instantaneous arrival and service rates, the only information missing is the time instants $u_{k,n}$ which are the arrival times of certain "molecules" that will violate the delay threshold. As described in detail in Section 4.3, this is easily doable in a DES setting.

Lemma 4.4: *Assuming $\beta(t) = \beta$ constant, the sample derivative of $D_2(\theta)$ is given by*

$$D_2'(\theta) = \frac{1}{\beta}\sum_{k=1}^{B_T}\sum_{n=1}^{(M_k-1)/2}\int_{u_{k,2n-1}}^{u_{k,2n}} x'(\theta,t)dt$$

where each integral $I_{u_{k,2n-1}}^{u_{k,2n}} = \int_{u_{k,2n-1}}^{u_{k,2n}} x'(t)dt$ is evaluated depending on the value of $v_{k,1}$ as follows.

- *If $u_{k,2n} < v_{k,1}$, then $I_{u_{k,2n-1}}^{u_{k,2n}} = 0$.*
- *If $u_{k,2n-1} < v_{k,1} < u_{k,2n}$, then $I_{u_{k,2n-1}}^{u_{k,2n}} = u_{k,2n} - v_{k,1}$.*
- *If $v_{k,1} < u_{k,2n-1}$, then $I_{u_{k,2n-1}}^{u_{k,2n}} = u_{k,2n} - u_{k,2n-1}$.*

The proof follows easily by differentiating (54) and then substituting the results from Lemma 4.1 (by definition, for all points $u_{k,j}$, $\omega(u_{k,j}) - d(u_{k,j}) = 0$). Note that this sample derivative is again very easy to implement. It simply consists of an accumulator (double sum) of time intervals as indicated by the three considered cases. This result holds also for time varying $d(t)$, however, if $\beta(t)$ is also time varying, then some numerical integration may be needed (see Lemma 4.1).

An important observation from both Lemmas 4.3 and 4.4 is that the IPA derivatives for the two metrics $D_1(\theta)$ and $D_2(\theta)$ do not require any knowledge of delay $\omega(t)$ which was used for modeling purposes but ultimately it is not involved in the IPA evaluation. The IPA evaluation depends only on detecting "buffer full" and "buffer empty" events and their times, $y = \beta d$ events and their times $u_{k,j}$ and also $\alpha(u_{k,j}) - \beta(u_{k,j} + d(u_{k,j}))$. The latter may be a little difficult to obtain in an SFM setting, however, as mentioned earlier, in a DES all these quantities can be easily evaluated (see Section 4.3).

4.2. Unbiasedness

In this section we show that the IPA estimators derived in the previous section are unbiased. To show the unbiasedness property we make the following mild assumptions.

Assumption A3: *Let $\tau_1, \cdots, \tau_M$ be the time instants when the net inflow process $\alpha(t) - \beta(t)$ changes value, and let $R_i = \alpha(\tau_i) - \beta(\tau_i)$, $i = 1, \cdots, M$. Assume that $R_{min} \le R_i \le R_{max}$ and that there exists $\mathcal{R} < \infty$ such that the pdf of R_i satisfies $f_{R_i}(r) \le \mathcal{R}$ for all $R_{min} < r < R_{max}$.*

Assumption A4: *Let $X_i = x(\theta, \tau_i)$, $i = 1, \cdots, M$ where τ_i is defined in Assumption A3. Assume that the pdf of X_i satisfies $f_{X_i}(X_i) \le P < \infty$ for all $0 \le X_i \le \theta$. Also, assume that, for some $\epsilon > 0$, $f_{X_i}(X_i) = 0$ for all X_i such that $|X_i - y(\tau_i)| < \epsilon$.*

Assumption **A3** is a mild condition requiring that rates do not become infinite, which is expected in practice and can obviously be verified. Assumption **A4** is a technical condition that does not allow the net inflow process to change rate exactly (or arbitrarily close) to points where $x(\theta, t) = y(t)$.

Theorem 4.1: *Under Assumptions A1–A4, the sample derivative of Lemma 4.3 is unbiased with respect to the objective function (49).*

Proof: Existence is guaranteed by Assumptions **A1** and **A2**. Therefore, we only need to show Lipschitz continuity of $D_1(\theta)$. Lipschitz continuity can be shown using the mean value theorem. For $D_1(\theta)$, let $\theta < \bar{\theta} < \theta + \Delta\theta$ such that

$$\frac{dD_1(\bar{\theta})}{d\theta} = \frac{D_1(\theta + \Delta\theta) - D_1(\theta)}{\Delta\theta}.$$

Thus we need to show that $\frac{dD_1(\bar{\theta})}{d\theta}$ is bounded by a random variable Q with $E[Q] < \infty$. From Lemma 4.3, by simply re-indexing all terms, we get

$$D_1'(\theta) = \sum_j \frac{\mathbf{1}\left[x(u_j^*;\theta) = \theta\right]}{|\alpha(u_j) - \beta(u_j + d(u_j)|}$$

During any interval $I_i = [\tau_i, \tau_{i+1})$, $i = 0, \cdots, M-1$, the net inflow rate R_i is constant for all $t \in I_i$. Thus, we can rewrite

$$D_1'(\theta) = \sum_{k=1}^{M} \frac{\mathbf{1}[C_k]}{R_k} \equiv \sum_{k=1}^{M} G_k \qquad (59)$$

where C_k is the event that $x(t_0;\theta) = y(t_0)$ for some $t_0 \in I_k$, and $\mathbf{1}[C_k]$ is the indicator function that takes the value 1 if the event occurred during I_k and 0 otherwise. Note that during any I_k only a single C_k event can occur since during the interval the net inflow is constant. The random variable G_k is defined so that $G_k = \frac{\mathbf{1}[C_k]}{R_k}$. Then,

$$E[G_k] = E_{X_k}\left[E_{R_k}\left[E_{G_k}[G_k|R_k, X_k]\right]|X_k\right] \qquad (60)$$

and each term is computed next assuming l_k is the length of the intervals I_k obviously bounded by $l_k \leq T$.

$$\begin{aligned} E_{G_k}[G_k|R_k, X_k] &= \frac{1}{R_k}\Pr\{\mathbf{1}[C_k] = 1 \mid R_k, X_k\} \\ &= \frac{1}{R_k}\Pr\{|X_k - y(\tau_k)| < R_k * l_k \mid R_k, X_k\} \\ &= \frac{1}{R_k}\Pr\left\{R_k > \frac{|X_k - y(\tau_k)|}{l_k} \,\middle|\, R_k, X_k\right\} \\ &= \begin{cases} \frac{1}{R_k} & \text{If } R_k > \frac{|X_k - y(\tau_k)|}{l_k} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

$$\begin{aligned} &E_{R_k}\left[E_{G_k}[G_k|R_k, X_k]|X_k\right] \\ &= \int_{\frac{|X_k - y(\tau_k)|}{l_k}}^{R_{max}} \frac{1}{r} f_{R_k}(r)dr \\ &\leq \mathcal{R}\left(\ln(R_{max}) - \ln\left(\frac{|X_k - y(\tau_k)|}{l_k}\right)\right) \\ &\leq Z(\epsilon) < \infty \end{aligned}$$

where $Z(\epsilon)$ is some constant that depends only on T and ϵ and we have used Assumptions **A3** and **A4**. Finally, again by **A4**,

$$\begin{aligned} E_{X_k}\left[E_{R_k}\left[E_{G_k}[G_k|R_k, X_k]\right]|X_k\right] &\leq \int_0^{\theta} Z(\epsilon) f_{X_k}(x)dx \\ &\leq \theta \cdot P \cdot Z(\epsilon) < \infty. \end{aligned}$$

Thus, from (59), $D_1'(\theta)$ is bounded by a random variable Q with $E[Q] < M \cdot \theta \cdot P \cdot Z(\epsilon) < \infty$. ■

Theorem 4.2: *Under assumptions **A1–A2**, the sample derivative of Lemma 4.4 is unbiased with respect to the objective function (50).*

The proof is similar to the proof of Theorem 4.1. Note that the sample derivative terms in Lemma 4.4 do not have the R_k at the denominator, thus one can easily show that the sample derivative is bounded without requiring Assumptions **A3** and **A4**.

4.3. Simulation Results

In this section we consider an "actual" DES and present simulation results where we use the IPA estimators obtained using the "equivalent" SFM. However, the estimators are computed using information observed from the sample path of the DES. In this real DES, tasks arrive according to a generally unknown stochastic process and are placed in the queue that operates according to a First In First Out (FIFO) policy. Associated with every task is a deadline and a service time, each taken from some distribution. Note that, since this is a DES, the state $x(\cdot)$ takes only integer values and the concept of sample derivative may not be well defined (or it may be zero). To distinguish them from the SFM quantities, in this section we denote the state by $X(\cdot)$ and the buffer capacity by K and allow them to take real values. In the DES setting, the objective functions of interest are given below

$$\tilde{L}(K) = \frac{1}{N_T}\sum_{n=1}^{N_T} \mathbf{1}[X(A_n; K) = K] \tag{61}$$

$$\tilde{D}_1(K) = \frac{1}{N_T}\sum_{n=1}^{N_T} \mathbf{1}[B_n - A_n > \mathcal{B}_n] \tag{62}$$

$$\tilde{D}_2(K) = \frac{1}{N_T}\sum_{n=1}^{N_T} [B_n - A_n - \mathcal{B}_n]^+ \tag{63}$$

where we used the $\tilde{\cdot}$ notation to emphasize that these functions are with respect to the DES. A_n and B_n are, respectively, the arrival and departure times of task n and $\mathcal{B}_n$ is the associated deadline. N_T is the number of tasks that arrived in the interval $[0, T]$. In the DES context, one needs to find K^* that minimizes an objective function of the form (48).

Since there are no closed form solutions for the required derivatives, we compare the obtained sample derivatives with the corresponding finite differences $\Delta W(K) = W(K+1) - W(K)$ where W corresponds to the performance metric of interest ($D_1(K)$ or $D_2(K)$[1]). We point out that the finite difference is not necessarily an accurate estimate of the required derivatives and it may have significant error at points where the objective function is non-linear. However we use it as an approximate estimate since there are no analytical results. Another important point is that, in order to obtain a single estimate using finite differences, one needs to observe the system under two different parameters K and $K+1$. The great advantage of the IPA approach is that the estimate is obtained by observing the system under a *single* parameter K. Thus the IPA approach is more appropriate for on-line parameter optimization.

Recall that the SFM is used to derive performance sensitivity estimates which cannot be obtained in the actual DES, in order to drive the optimization scheme (4) for the original DES. To evaluate the IPA estimators in Lemma 4.3 and Lemma 4.4, we need to identify all events defined

[1] $\tilde{L}(K)$ is not treated here since it was covered extensively in [6].

in the SFM with observable events in the DES sample path. For exogenous events, we monitor changes in arrival rates $\alpha(t)$ and $\beta(t)$ through simple rate estimators; if $|\alpha(t) - \alpha(t-\delta)| > \epsilon$ for some small δ and adjustable ϵ, we detect such changes and similarly for $\beta(t)$. For endogenous events, we simply observe when the queue becomes full or empty. One difficulty in identifying endogenous events is that we have to recognize that in a DES the queue content may "chatter" near K. For example, suppose at time t the queue is full, i.e., $X(t) = K$, although the interarrival time mean is less than the interdeparture time mean; this means the arrival rate is larger than the output rate, yet it is still possible that for some sample path the departure event will happen first at time $t_1 > t$, so that $X(t_1) = K - 1 < K$. However, from the comparison of input and output rates, we know that in the SFM the system does not actually leave θ at t_1. Moreover, suppose at $t_2 > t_1$ an arrival event occurs which brings X back to K; then, based on our previous analysis, since the system *does not* actually leave mode 4 at t_1, the event at t_2 should not be considered as an endogenous event. Thus, during the interval $[t, t_2]$, the queue length of the actual DES "chatters" around the threshold level. However, in the SFM, the system remains in mode 4 during this period. Then, the question is how to identify the start and end of a "full buffer period". We resolve this issue by comparing the service rate $\beta(t)$ and arrival rate $\alpha(t)$ as measured on the actual DES sample path. Then, we can identify a "full buffer period" as starting at τ_k, if $X(\tau_{k-1}) < K, X(\tau_k) = K$, and $\beta(\tau_k) < \alpha(\tau_k)$; similarly, we detect the end of this period at τ_k, if $X(\tau_{k-1}) = K$, $X(\tau_k) < K$, and $\beta(\tau_k) > \alpha(\tau_k)$. Finally, as mentioned in the last section, events at times $u_{k,j}$ can be identified by applying timestamps on tasks such that event time $u_{k,2n+1}$, $k = 1, 2, \cdots, n = 0, 1, \cdots$ is the arrival time of the task that violated the delay threshold constraint following a task that did not violate its constraint; whereas event time $u_{k,2n}$, $n = 1, 2, \cdots$, is the arrival time of a task that arrived after $u_{k,2n-1}$ but did not violate the deadline constraint.

Finally, even though the estimator in Lemma 4.2 is provably unbiased (for SFM-based objective functions), in practice there could be numerical problems introduced by the fact that the difference in its denominator may be close to zero over some time intervals. To avoid such numerical issues, we use the following heuristic for the computation of $\tilde{D}_1'(\theta)$ at the expense of a small bias.

$$\tilde{D}_1'(\theta) = \frac{1}{N_T} \sum_k \sum_n \left[\frac{\mathbf{1}\left[x(u_{k,2n}^*) = K\right]}{c + |\alpha(u_{k,2n}) - \beta|} + \frac{\mathbf{1}\left[x(u_{k,2n-1}^*) = K\right]}{c + |\alpha(u_{k,2n-1}) - \beta|} \right] \tag{64}$$

where c is a small constant and for the purposes of the results presented next, $c = 0.2$. This constant is used to prevent numerical problems when **A3** is violated in practice, thus it does not allow the denominator to become very close to 0. Note that in some cases, it is possible that the estimator of $\alpha(u) - \beta(u)$ may become very small and, as a result, the estimated u' may dominate the value of the double sum (especially when the observation interval is not very long). The introduction of c may introduce a small bias in the estimation however, it does prevent a single term of the double sum to dominate its value. Also note that the use of absolute values is also justified because at points $u_{k,2n}$, $\alpha(u_{k,2n}) < \beta$ and at points $u_{k,2n-1}$, $\alpha(u_{k,2n-1}) > \beta$, $n = 1, 2, \cdots$ (for simplicity, for all scenarios we use a constant β).

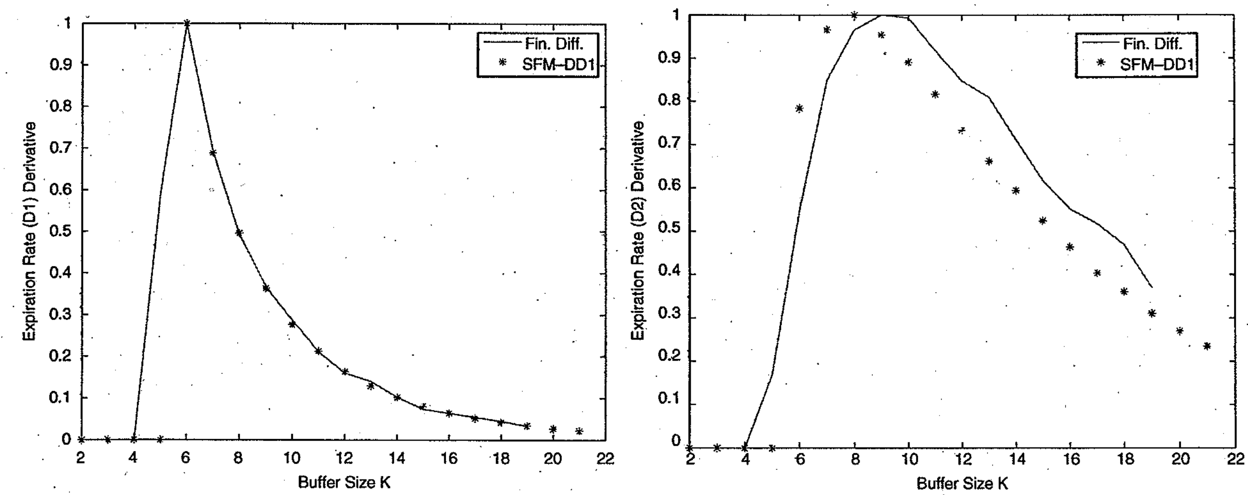
In the first scenario, we assume that a source is generating tasks according to a Poisson process with rate 50 tasks per second. The service time of each task is 18 ms fixed. Furthermore, each task is marked with a deadline which is also deterministic and equal to 90 ms.

Figure 6(a) presents the normalized estimate $\tilde{D}_1'(\theta)$ as computed from data observed from the sample path of the system described in scenario 1. Furthermore, the figure shows the normalized estimate of $D_1'(K)$ computed using forward finite differences. Figure 6(b) presents the normalized estimate $\tilde{D}_2'(\theta)$ again for scenario 1 together with the corresponding normalized estimate of $D_2'(K)$ computed using finite differences. The results presented in Fig. 6 correspond to the average of 30 sample paths of 1 hour simulated time. Notice that for small buffer sizes, no task experiences any deadline violation and therefore both objectives (and their corresponding derivatives/finite differences) are 0.

In the second scenario, it is assumed that the source is generating tasks according to a Markov Modulated Poisson Process. At each state, the source generates tasks according to a Poisson process with rate λ where λ is a random variable uniformly distributed between 33 and 100 tasks per second. The times when the source switches its rate are exponentially distributed with mean 1 second. The service time of each task is again deterministic and equal to 20 ms. Finally, the completion deadline of each task is a random variable uniformly distributed between 100 and 300 ms.

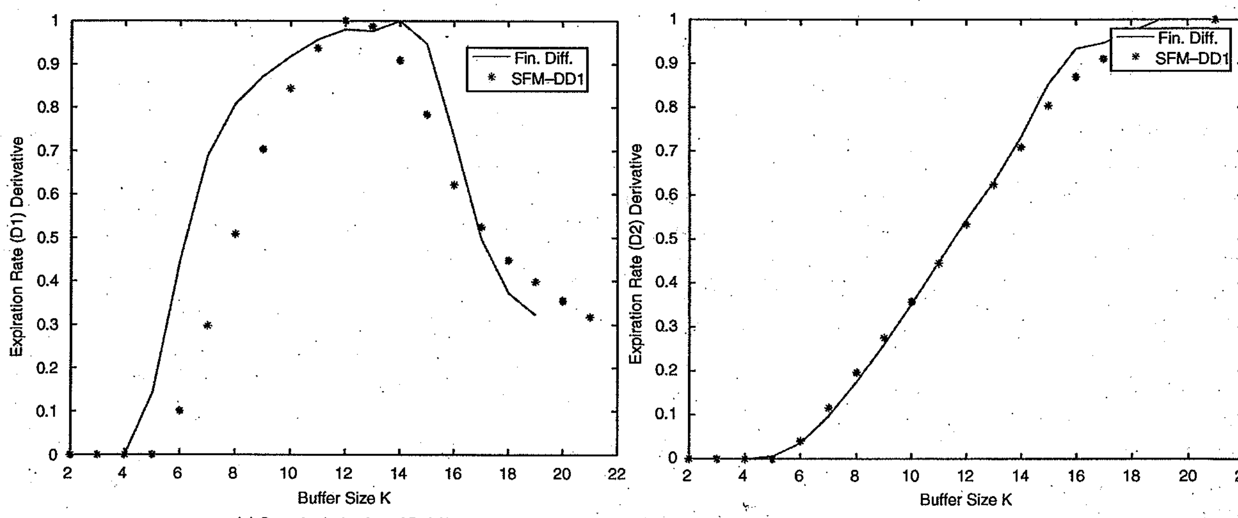
Figure 7(a) presents the normalized estimate $\tilde{D}_1'(\theta)$ as computed from data observed from the sample path of the system described in scenario 2. The figure also shows the normalized estimate of $D_1'(K)$ computed using finite differences. Figure 7(b) presents the normalized estimate $\tilde{D}_2'(\theta)$ for scenario 2 together with corresponding finite difference estimate. The results presented in Fig. 7 are from a single sample path, 5 hours of simulation time long.

Figure 8 shows an example of applying our IPA estimates and (4) in optimizing the buffer threshold of a



(a) Sample derivative of $D_1(\theta)$ (b) Sample derivative of $D_2(\theta)$

Fig. 6. Sample derivatives with respect to the buffer size for Scenario 1.



(a) Sample derivative of $D_1(\theta)$ (b) Sample derivative of $D_2(\theta)$

Fig. 7. Sample derivatives with respect to the buffer size for Scenario 2.

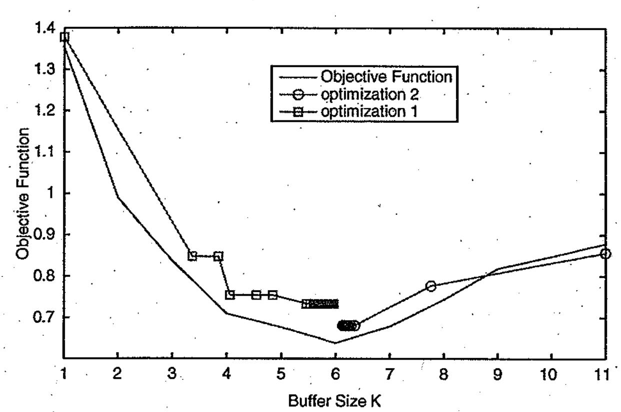


Fig. 8. Optimization of Actual DES using SFM-based gradient estimation.

queueing system (again, in the "actual" DES setting). The objective function is of the form (48), and its accurate estimate is obtained by exhaustive simulation averaged over 100 sample paths for about 4 hours. Note that this estimate is obtained separately beforehand and it is not used in the optimization process. This gives an optimal point $K^* = 6$. For this example, tasks arrive according to a Markov Modulated Poisson Process, with mean interarrival time uniformly distributed on [0.6sec, 1.4sec]. Service times and delay thresholds are also piecewise constant random processes with means 0.9sec and 3.5sec respectively. The two trajectories labeled "optimization" are results of implementing (4) using gradient estimates obtained through IPA on a single sample path with different starting points. We can see that each converges to a point sufficiently close to the "true" optimum, illustrating the effectiveness of our method. We also point out that the modification (64) is not used in this case: the numerical effect of an occasional estimate with very small denominator is negligible during an iterative optimization process. Note that although we have used this relatively simple one-dimensional problem simply to illustrate the use of our IPA algorithms, this framework has been successfully applied to higher-dimensional problems as well [22],[29].

5. Conclusions

We have presented a general framework for carrying out perturbation analysis in SHS of arbitrary structure. By analyzing sample paths of an SHS, we can obtain derivatives of event (mode switching) times and of state variables with respect to various controllable parameters. An attractive property of these derivatives is that they are computable exclusively from directly-observable data and they are generally simple to compute. These derivatives ultimately translate into gradient estimates of performance metrics with respect to the controllable parameters which are then used to drive standard gradient-based optimization algorithms implementable on line with little or no distributional information regarding stochastic processes involved. We applied this framework to two SHS with different characteristics, thus illustrating its generality. We also addressed the issue of establishing the unbiasedness of the gradient estimators derived and explicitly showed how to prove it in one of the two SHS studied.

One of the contributions of the paper is the generalization of the concept of "induced" events which extends the range of SHS we can study. In recent work [25], induced events were seen to cause potentially infinite event chains giving rise to interesting phenomena with occasionally counterintuitive effects.

Acknowledgements

Supported in part by NSF under Grants DMI-0330171 and EFRI-0735974, by AFOSR under grants FA9550-04-1-0133 and FA9550-09-1-0095, and by DOE under grant DE-FG52-06NA27490, and by ONR under grant N00014-09-1-1051 (C.G. Cassandras and C. Yao). This research was supported in part by the EU project CON4COORD (INFSO-ICT-223844) (C.G. Panayiotou).

References

1. Anick D, Mitra D, Sondhi MM. Stochastic theory of a data-handling system with multiple sources. *Bell Syst Techn J*, 1982; 61: 1871–1894.
2. Aydin H, Melhem R, Mossé D, Mejia-Alvarez P. Power-aware scheduling for periodic real-time tasks. *IEEE Trans. Computers*, 2004; 53(5): 584–600.
3. Buttazzo GC. *Hard Real-Time Computing Systems: Predictable Scheduling Algorithms and Applications*. Kluwer Academic Publishers, Norwell, MA, 1997.
4. Cassandras CG, Lafortune S. *Introduction to Discrete Event Systems, Second Edition*. Springer, 2008.
5. Cassandras CG, Lygeros J, editors. *Stochastic Hybrid Systems*. Taylor and Francis, 2006.
6. Cassandras CG, Wardi Y, Melamed B, Sun G, Panayiotou CG. Perturbation analysis for on-line control and optimization of stochastic fluid models. *IEEE Trans Automatic Control*, 2002; AC-47(8): 1234–1248.
7. Connor D, Feigin G, Yao DD. Scheduling semiconductor lines using a fluid network model. *IEEE Trans Robotics Autom*, 1994; 10: 88–98.
8. Gamal AE, Nair C, Prabhakar B, Uysal-Biyikoglu E, Zahedi S. Energy-efficient scheduling of packet transmissions over wireless networks. In *Proc IEEE INFO-COM*, volume 3, 23-27, pages 1773–1782, New York City, USA, 2002.
9. Glasserman P. *Gradient Estimation via Perturbation Analysis*. Kluwer Academic Pub., 1991.
10. Haurie A, L'Ecuyer P, Van Delft Ch. Monte Carlo optimization of parametrized policies in a class of piecewise deterministic control systems arising in manufacturing flow control. In *Proc. 29th IEEE Conference on Decision Control*, 1990, pp. 587–589.
11. Haurie A, L'Ecuyer P, Van Delft Ch. Monte-Carlo optimization of parametrized policies in a class of piecewise deterministic control systems arising in manufacturing flow control. *J Discrete Event Dynamic Syst: Theory Appl*, 1994; 4(1): 87–111.
12. Ho YC, Cao XR. *Perturbation Analysis of Discrete Event Dynamic Systems*. Kluwer Academic Publishers, Boston, 1991.
13. Jeffay K, Stanat DF, Martel CU. On non-preemptive scheduling of periodic and sporadic tasks. In *Proc. IEEE Real-Time Syst Symp*, 1991, pp. 129–139.
14. Jonsson J, Lonn H, Shin KG. Non-preemptive scheduling of real-time threads on multi-level-context architectures. In *Proc. IEEE Workshop Parallel Distributed Real-Time Syst*, volume 1586, pp. 363–374. Springer Verlag, 1999.
15. Kushner HJ, Yin GG. *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, New York, NY, 1997.
16. Liu B, Guo Y, Kurose J, Towsley D, Gong WB. Fluid simulation of large scale networks: Issues and tradeoffs. In *Proc Intl Conf Parallel Distributed Proc Techn Appl*, 1999, pp. 2136–2142.
17. Liu JWS. *Real-Time System*. Prentice Hall Inc., 2000.
18. Mao J, Cassandras CG, Zhao QC. Optimal dynamic voltage scaling in power-limited systems with real-time constraints. *IEEE Trans Mobile Comput*, 2007; 6(6): 678–688.
19. Miao L, Cassandras CG. Optimal transmission scheduling for energy-efficient wireless networks. In *Proc. INFOCOM*, 2006.
20. Pepyne DL, Cassandras CG. Optimal control of hybrid systems in manufacturing. In *Proc. IEEE*, 2000; 88: 1108–1123.
21. Rubinstein RY, Shapiro A. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*. John Wiley and Sons, New York, New York, 1993.

22. Sun G, Cassandras CG, Panayiotou CG. Perturbation analysis and optimization of stochastic flow networks. *IEEE Trans Autom Control*, 2004; 49(12): 2113–2128.
23. Wardi Y, Adams R, Melamed B. A unified approach to infinitesimal perturbation analysis in stochastic flow models: the single-stage case. *IEEE Trans Autom Control*, 2010; 55(1): 89–103.
24. Wardi Y, Riley G. Infinitesimal perturbation analysis in networks of stochastic flow models: General framework and case study of tandem networks with flow control. *J Discrete Event Dynamic Syst*, 2010; 20(2): 275–305.
25. Yao C, Cassandras CG. Perturbation analysis and optimization of multiclass multi-objective stochastic flow models. In *Proc. 48th IEEE Conf of Decision Control*, 2009, pp. 914–919.
26. Yao F, Demers A, Shenker S. A scheduling model for reduced CPU energy. In *Proc. 36th Annual Symp Foundations Comput Sci (FOCS'95)*, pp. 374–382. IEEE Computer Society, 1995.
27. Yu H, Cassandras CG. Perturbation analysis of feedback-controlled stochastic flow systems. *IEEE Trans Autom Control*, AC-2004; 49: 1317–1332.
28. Yu H, Cassandras CG. *A New Paradigm for On-Line Management of Communication Networks with Multiplicative Feedback Control* (Girard A, Sanso B, Vazquez-Abad F, Eds.). Springer-Verlag, 2005, pp. 297–332.
29. Yu H, Cassandras CG. Perturbation analysis and feedback control of communication networks using stochastic hybrid models. *J Nonlinear Anal*, 2006; 65: 1251–1280.