IPA for Continuous Petri Nets with Threshold-Based Flow Control

Alessandro Giua, Carla Seatzu, Yorai Wardi

Abstract

This paper considers an application of the Infinitesimal Perturbation Analysis (IPA) gradient-estimation technique to a class of continuous Petri nets. In particular, it proposes a systematic approach for computing the derivatives of the sample performance functions with respect to structural and control parameters. The resulting algorithms are recursive in both time and network flows, and their steps are computed in response to the occurrence and propagation of certain events in the network. Such events correspond to discontinuities in the network flow-rates, and their special characteristics are due to the properties of continuous transitions and fluid places. Following a general outline of the framework we focus on a simple yet canonical example, and investigate throughput and workload-related performance criteria as functions of a threshold control variable. Simulation experiments support the analysis and testify to the potential viability of the proposed approach.

Published as:

A. Giua, C. Seatzu, Y. Wardi, "IPA for Continuous Petri Nets with Threshold-Based Flow Control," *13th IFAC World Congress* (Milan, Italy), August-September 2011.

This work has been partially supported by the European Communitys Seventh Framework Programme under project DISC (Grant Agreement n. INFSO-ICT-224498).

A. Giua and C. Seatzu are with the Dept. of Electrical and Electronic Engineering, University of Cagliari, Italy, e-mail: {giua, seatzu}@diee.unica.it.

Y. Wardi is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, e-mail: yorai.wardi@ece.gatech.edu.

I. INTRODUCTION AND PROBLEM STATEMENT

In the past decade much of the research on Infinitesimal Perturbation Analysis (IPA) has focused on fluid queueing networks, whose dynamics are characterized by flow rates rather than by the movement and storage of discrete jobs. Consequently, certain traffic processes have an inherent continuity which renders them amenable to the application of IPA (Cassandras (2006)). More recently, the investigation of IPA was extended to the setting of continuous Petri nets, whose concept of the continuous transition sets them apart from continuous queueing networks and poses additional challenges. Xie (2002) derived IPA algorithms for networks with piecewise-constant flow rates, and Giua et-al. (2010) developed a general algorithmic framework for a class of networks with piecewise-continuous flow rates. However, this framework (as well as the algorithms in Xie (2002)) assume that the performance perturbations are generated by exogenous processes, thereby excluding many forms of feedback control laws. The purpose of this paper is to initiate an effort to extend that framework to a class of parameterized feedback laws consisting of threshold-based flow control. Since this is but an initial study, we lay out a general framework for algorithms but focus the detailed analysis on the specific example shown in Fig. 1; though simple, it captures many of the salient features of IPA in the setting of continuous Petri nets. A followup publication will extend the analysis to a more general class of networks and systems.

IPA is a general technique for computing gradients (derivatives) of sample performance functions defined on stochastic Discrete Event Dynamic Systems (DEDS) (Cassandras and Lafortune (1999)). Its principal application areas has been in queueing networks and more recently in fluid queues, due to their special structure yielding simple algorithms for the sample gradients. These gradients, called the *IPA gradients*, can be used in sensitivity analysis and optimization of the related expected-value function via stochastic approximation. Formally, let $J(\theta)$ be a random function whose realization is defined on the sample path of a DEDS and let $\ell(\theta) := E[J(\theta)]$ be its expected value, where $\theta \in R^n$ is the variable parameter. IPA computes the sample gradient $\nabla J(\theta)$ which, under some circumstances, can be used in optimization of $\ell(\theta)$. In a class of queueing models and performance functions, the IPA gradient $\nabla J(\theta)$ can be computed by simple algorithms, and this has provided a major factor motivating research and development of IPA. For a detailed discussion on IPA and its potential applications, please see Cassandras and Lafortune (1999).

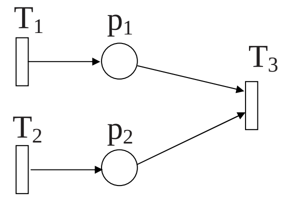

Fig. 1. Example of a Continuous Petri Net

Continuous Petri nets are stochastic Petri nets with "fluidized" tokens flowing on their arcs instead of discrete tokens. Introduced in Alla and David (1998), their algebraic properties subsequently were investigated in Silva and Recalde (2004). A typical transition T is characterized by a given maximum firing (flow) rate, $V_T(t) \geq 0$, which can be a function of time as the notation indicates. The actual firing rate, denoted by $v_T(t)$, satisfies the inequalities $0 \leq v_T(t) \leq V_T(t)$. A typical place is denoted by p, and its fluid level (occupancy) is denoted by $m_p(t)$. As in earlier studies of IPA in the Petri-net setting (Xie (2002); Giua et-al. (2010)), we consider a class of Petri nets called event graphs, namely networks whose places have each a single input transition and a single output transition. No further restrictions are made on the topology of the networks, and they may be closed, open, or neither closed not open.

Following the notation in Giua et-al. (2010) we define, for a transition T, the terms $in(T)$ and $out(T)$ to be the set of input places to T and the set of output places from T, respectively. If $in(T) = \emptyset$ then T is a source transition, and if $out(T) = \emptyset$ then T is a sink transition. Furthermore, for every place p we denote by $in(p)$ and $out(p)$ the input transition to p and the output transition from p, respectively.

Continuous Petri nets can be viewed as stochastic DEDS whose state variable is comprised of all transition-flow rates $v_T(t)$ and the place-fluid levels $m_p(t)$. Define $\varepsilon_T(t) := \{p \in in(T) : m_p(t) = 0\}$, namely the input places to T which are empty at time t. Suppose that the network evolves in a time-interval $[0, \mathcal{T}]$ for some given $\mathcal{T} > \prime$. The state equations are, for every transition T,

$$v_T(t) = \begin{cases} V_T(t), & \text{if } \varepsilon_T(t) = \emptyset \\ \min\{v_{in(p)}(t) : p \in \varepsilon_T(t)\}, & \text{if } \varepsilon_T(t) \neq \emptyset; \end{cases} \tag{1}$$

and for every place p,

$$\dot{m}_p(t) = v_{in(p)}(t) - v_{out((p)}(t), \tag{2}$$

with some given initial condition $m_p(0)$. These equations must be satisfied simultaneously for all of the network's nodes (transitions and places), and sufficient conditions for the existence of unique solutions were mentioned in Giua et-al. (2010).

In the IPA setting let $\theta \in R^n$ be a variable parameter of the traffic processes, which are therefore denoted by $V_T(\theta,t)$, $v_T(\theta,t)$, and $m_p(\theta,t)$. With this notation, Equations (1) and (2) are to be understood in the following way: fix $\theta \in R^n$, and let the state evolve in the time-interval $[0,\mathcal{T}]$ according to these equations. Sample performance functions of frequent interest in applications, like throughput and delay, are related to the following two functions defined, respectively, for transitions T and places p (see Giua et-al. (2010)):

$$J_T(\theta) \ := \ \int_0^{\mathcal{T}} v_T(\theta,t)dt, \tag{3}$$

and

$$J_p(\theta) \ := \ \int_0^{\mathcal{T}} m_p(\theta,t)dt. \tag{4}$$

Their sample gradients, $\nabla J_T(\theta)$ and $\nabla J_p(\theta)$, are the targets of the IPA algorithms, and hence are called the *IPA gradients*. We investigate them under the general network-structure considered in Giua et-al. (2010) except that we assume that the variable θ is not a parameter of an exogenous process but rather a threshold-control parameter, and this necessitates a new line of analysis. As mentioned earlier, this being an initial study, we first present our analysis in general and abstract terms and then apply it to an example involving the particular system shown in Fig. 1.

Section 2 presents a general framework for computing the IPA derivatives, and Section 3 analyzes the aforementioned example. Section 4 contains simulation results, and Section 5 concludes the paper.

II. General Framework for IPA

This section considers the IPA gradients of the sample performance functions $J_T(\theta)$ and $J_p(\theta)$ defined by Equations (3) and (4). To somewhat simplify the exposition we assume that $\theta \in R$ so that the IPA gradient is called the *IPA derivative* and denoted by $\frac{dJ}{d\theta}(\theta)$. Throughout the discussion therein we assume that all of the mentioned derivatives exist and the standard rules of calculus apply. We also assume that for a given $\theta \in R$, transition T, and place p, the function $v_T(\theta,\cdot)$ is piecewise continuous and piecewise continuously differentiable, and the function $m_p(\theta,\cdot)$ is continuous and piecewise continuously differentiable. Furthermore, the discontinuity (jump) time-points of $v_T(\theta,\cdot)$ are functions of θ and hence are denoted by $t_{k,T}(\theta)$, $k=1,\ldots,K$ for some (random) K, and we assume that, at a given θ, their derivatives $\frac{dt_{k,T}}{d\theta}(\theta)$ exist w.p.1. These assumptions could be verifyable for particular systems as will be seen in the next section.

Consider first the IPA derivative $\frac{dJ_T}{d\theta}(\theta)$. Taking derivatives in (3) we obtain its following general form,

$$\frac{dJ_T}{d\theta}(\theta) \;=\; \int_0^{\mathcal{T}} \frac{\partial v_T}{\partial \theta}(\theta,t)dt \;+$$
$$\sum_{k=1}^{K}\Big(v_T(\theta,t_{k,T}(\theta)^-) - v_T(\theta,t_{k,T}(\theta)^+)\Big)\frac{dt_{k,T}}{d\theta}(\theta). \tag{5}$$

The terms $\frac{\partial v_T}{\partial\theta}(\theta,t)$ typically can be computed directly and easily from the sample path, and hence the main challenge is to compute the sum-terms in the Right-Hand Side (RHS) of Equation (5). These terms also arise in the IPA derivative $\frac{dJ_p}{d\theta}(\theta)$. Indeed, taking derivatives in (4) we obtain,

$$\frac{dJ_p}{d\theta}(\theta) \;=\; \int_0^T \frac{\partial m_p}{\partial \theta}(\theta,t)dt, \tag{6}$$

and the integrant in this equation has the following form. If t lies in the interior of an empty period at p then $\frac{\partial m_p}{\partial\theta}(\theta,t)=0$. On the other hand, if $m_p(\theta,t)>0$, let $\xi(\theta) := \max\{\tau \leq t : m_p(\theta,\tau)=0\}$, then by (2)

$$m_p(\theta,t) \;=\; \int_{\xi(\theta)}^{t}\Big(v_{in(p)}(\theta,\tau) - v_{out(p)}(\theta,\tau)\Big)d\tau. \tag{7}$$

Let $t_{j,in(p)}(\theta)$, $j=j_1,\ldots,j(t)$ denote the jump-points of the function $v_{in(p)}(\theta,\cdot)$ in the interval $(\xi(\theta),t)$, and let $t_{\ell,out(p)}(\theta)$, $\ell=\ell_1,\ldots,\ell(t)$ denote the jump-points of the function $v_{out(p)}(\theta,\cdot)$ in the interval $(\xi(\theta),t)$. Suppose that $v_{in(p)}(\theta,\cdot)$ and $v_{out(p)}(\theta,\cdot)$ are continuous at t. Then (7) implies that

$$\begin{aligned}\frac{\partial m_p}{\partial\theta}(\theta,t) \;=\; & \int_{\xi(\theta)}^{t}\Big(\frac{\partial v_{in(p)}}{\partial\theta}(\theta,\tau)-\frac{\partial v_{out(p)}}{\partial\theta}(\theta,\tau)\Big)d\tau \\ & +\sum_{j=j_1}^{j(t)}\Big(v_{in(p)}(\theta,t_{j,in(p)}(\theta)^-)-v_{in(p)}(\theta,t_{j,in(p)}(\theta)^+)\Big) \\ & \times\frac{dt_{j,in(p)}}{d\theta}(\theta)- \\ & \sum_{\ell=\ell_1}^{\ell(t)}\Big(v_{out(p)}(\theta,t_{\ell,out(p)}(\theta)^-)-v_{out(p)}(\theta,t_{\ell,out(p)}(\theta)^+)\Big) \\ & \times\frac{dt_{\ell,out(p)}}{d\theta}(\theta) \\ & -\Big(v_{in(p)}(\theta,\xi(\theta)^+)-v_{out(p)}(\theta,\xi(\theta)^+)\Big)\frac{d\xi}{d\theta}(\theta). \end{aligned} \tag{8}$$

We see that the terms $\Big(v_T(\theta,t_{k,T}(\theta)^-)-v_T(\theta,t_{k,T}(\theta)^+)\Big)$ $\times\frac{dt_{k,T}}{d\theta}(\theta)$ play a role in both the IPA derivatives $\frac{dJ_T}{d\theta}(\theta)$ and $\frac{dJ_p}{d\theta}(\theta)$, and we next indicate a recursive way to compute them along a sample path.

To simplify the notation, we define the terms $\Delta v_T(\theta,t):=v_T(\theta,t^-)-v_T(\theta,t^+)$ and $\Delta V_T(\theta,t):=V_T(\theta,t^-)-V_T(\theta,t^+)$, and we observe that $\Delta v_T(\theta,t)\neq 0$ ($\Delta V_T(\theta,t)\neq 0$, resp.) only if t is a jump point of the function $v_T(\theta,\cdot)$ ($V_T(\theta,\cdot)$, resp.). Equations (5) and (8) require the computation of $\Delta v_T(\theta,t_T(\theta))\frac{dt_T}{d\theta}(\theta)$, where $t_T(\theta)$ serves as a generic notation for a jump point of $v_T(\theta,\cdot)$. The computation of these terms is at the heart of the IPA algorithm, and it is based on the notion of discrete events in the following way. Associated with each transition there is a sequence of events occurring at random times. Every jump in the function $v_T(\theta,\cdot)$ is an event associated with T, and every event either is such a jump or causes a jump possibly at another transition. When an event occurs in transition T at time $t_T(\theta)$, we compute the term $\Delta v_T(\theta,t_T(\theta))\frac{dt_T}{d\theta}(\theta)$. Furthermore, this term may be used in the computation of similar terms associated with other events, possibly occurring at other transitions, and the specific event at T may indicate how this can be done.

The various events are classified according to the following definition, where the timing of an event at transition T is indicated by $t_T(\theta)$.

***Definition** 2.1:* 1) An exogenous event at transition T is a jump in $V_T(\theta,\cdot)$ such that, conditioned on $V_T(\theta,t_T(\theta)^-)$ and $t_T(\theta)$, the term $\Delta V_T(\theta,t_T(\theta))\frac{dt_T}{d\theta}(\theta)$ is independent of any other past network processes.

2) An event at transition T is a null event if $\frac{dt_T}{d\theta}(\theta) = 0$.
3) Let $\gamma(\theta)$ be a continuously-differentiable, non-negative-valued function. A Type-1 endogenous event at transition T, with respect to the function $\gamma(\theta)$, is the event that, at some $p \in in(T)$, $m_p(\theta, t_T(\theta)) = \gamma(\theta)$ while $m_p(\theta, t_T(\theta)^-) \neq \gamma(\theta)$. A Type-2 endogenous event at transition T, with respect to the function $\gamma(\theta)$, is the event that, at some $p \in in(T)$, $m_p(\theta, t_T(\theta)) = m_p(\theta, t_T(\theta)^-) = \gamma(\theta)$ while $m_p(\theta, t_T(\theta)^+) \neq \gamma(\theta)$.
4) A pair of events at transitions U and T is called triggering-induced if the event at U causes the event at T, $t_U(\theta) = t_T(\theta)$, and $\frac{dt_U}{d\theta}(\theta) = \frac{dt_T}{d\theta}(\theta)$.

***Remark** 2.2:* 1) The processes $V_T(\theta, t)$ often are exogenous and hence their jumps are exogenous events.

2) Null events typically arise as a result of a jump in an exogenous function like $V_T(t)$ which does not depend on θ. Such events contribute nothing to the IPA derivative and hence need not be considered by an algorithm.
3) By Definition 2.1(3), endogenous events are defined as occurring at an input place p of a transition T. We will refer to such events as occurring at either T or $p \in in(T)$, as convenient.
4) Type-1 endogenous events mean that $m_p(\theta, t)$ is becoming equal to $\gamma(\theta)$ at the time $t = t_T(\theta)$, and it may stay at that value for a positive amount of time ($m_p(\theta, t_T(\theta)^+) = \gamma(\theta)$) or just cross that value ($m_p(\theta, t_T(\theta)^+) \neq \gamma(\theta)$). On the other hand, a type-2 endogenous event signifies that $m_p(\theta, t)$ departs from the value of $\gamma(\theta)$ after being there for a positive amount of time ($m_p(\theta, t_T(\theta)^-) = \gamma(\theta)$).
5) A special case of endogenous events is when $\gamma(\theta) = 0$; in this case a type-1 endogenous event is the start of an empty period while a type-2 endogenous event is the end of an empty period.
6) Definition 2.1(4) specifies that an induced event occurs at the same time as its triggering event. It could be extended to include the case where the induced event occurs after its triggering event, but its present form suffices for the purpose of his paper.
7) It is possible to have a chain of events, $e_1, \ldots, e_n$, such that e_1 is a triggering event; for all $i = 1, \ldots, n-1$, e_i triggers e_{i+1}; and e_n is an induced event. Such a chain is called an induced chain. General sufficient conditions for the finiteness of induced chains were

mentioned in Giua et-al. (2010).

This classification of events extends the one in Giua et-al. (2010) by having a far-more general notion of endogenous events, which in Giua et-al. (2010) is restricted to the boundaries of empty periods. Consequently the variable θ must be a parameter of an exogenous process in Giua et-al. (2010), while here it can be a tunable parameter of a control policy as well.

The analysis below is carried out under the following assumption, variants of which are often made in the literature on IPA in fluid queues; see Cassandras (2006) and references therein.

***Assumption** 2.3:* 1) For a fixed θ, w.p.1 no two events occur at the same time unless they are part of an induced chain.

2) No type-1 endogenous event can be induced. Furthermore, if such event occurs in transition T then for every $p \in in(T)$ the function $v_{in(p)}(\theta, \cdot)$ is continuous at its occurrence time.

3) All the derivatives mentioned in the sequel exist and the standard rules of calculus apply.

The computation and propagation of the terms $\Delta v_T(\theta, t_T(\theta))$ in time and network flows are carried out according to the following guidelines and rules associated with the various events.

- *Exogenous events.* Typically $V_T(\theta, t)$ is an exogenous process and hence, at any its jump points $t_T(\theta)$, the term $\Delta V_T(\theta, t_T(\theta))\frac{dt_T}{d\theta}(\theta)$ is easily computable from the sample path. This paper does not have exogenous events and hence they will not be discussed further.
- *Null events.* By Definition 2.1(2) $\Delta V_T(\theta, t_T(\theta))\frac{dt_T}{d\theta}(\theta)$ $= 0$, and hence such events contribute nothing to the IPA derivative and need not be considered by an algorithm.
- *Triggering-induced events.* Consider an event at transition U that triggers an event at another transition T. Then, by Definition 2.1(4), $t_U(\theta) = t_T(\theta)$ and $\frac{dt_U}{d\theta}(\theta) = \frac{dt_T}{d\theta}(\theta)$. Now the relationships between the terms $\Delta v_U(\theta, t_U(\theta))$ and $\Delta v_T(\theta, t_T(\theta))$ may have to be determined by ad-hoc ways according to the specific characteristics of the two events. One case that frequently arises is when T is immediately downstream from U, namely there exists a place $p \in out(U) \cap in(T)$, and $t_U(\theta)$ lies in the interior of an empty period at p. Then a jump in $v_U(\theta, \cdot)$ triggers a similar jump in $v_T(\theta, \cdot)$, and by (1), $\Delta v_T(\theta, t_T(\theta)) = \Delta v_U(\theta, t_U(\theta))$ and hence $\Delta v_T(\theta, t_T(\theta))\frac{dt_T}{d\theta}(\theta) = \Delta v_U(\theta, t_U(\theta))\frac{dt_U}{d\theta}(\theta)$.
- *Type-1 endogenous events.* Suppose that $m_p(\theta, t)$ becomes equal to $\gamma(\theta)$ for some $p \in in(T)$, namely $m_p(\theta, t_T(\theta)) = \gamma(\theta)$ while $m_p(\theta, t_T(\theta)^-) \neq \gamma(\theta)$. Such events naturally occur when

an empty period begins at p, i.e. $\gamma(\theta)=0$, and in this case the function $v_T(\theta,t)$ may have a jump at the time $t=t_T(\theta)$. In this paper we are also interested in the case where $\gamma(\theta)>0$ is a threshold variable which controls the maximum flow rate at another transition, W. In this case the function $v_T(\theta,\cdot)$ may or may not have a jump at $t=t_T(\theta)$, but this event triggers an induced event in some transition W, namely a jump in $v_W(\theta,\cdot)$. The relationship between $\Delta v_W(\theta,t_W(\theta))$ and $\Delta v_T(\theta,t_T(\theta))$ has to be determined according to each specific case, but $\frac{dt_W}{d\theta}(\theta)=\frac{dt_T}{d\theta}(\theta)$, and this term has the following form. Define $\xi_p(\theta):=\min\{\tau<t_T(\theta):m_p(\theta,\tau)=0\}$ (we assume, for the sake of notational consistency, that $m_p(\theta,0)=0$). Furthermore, let $U:=in(p)$, let $\tau_{\ell,U}(\theta)$, $\ell=1,\ldots,L$ denote the jump-times of the function $v_U(\theta,\cdot)$ in the interval $(\xi_p(\theta),t_T(\theta))$, and let $\tau_{m,T}(\theta)$, $m=1,\ldots,M$ denote the jump times of the function $v_T(\theta,\cdot)$ in the same interval.

***Proposition** 2.4:* The following relation holds:

$$\begin{aligned}\frac{dt_T}{d\theta}(\theta) \;=\; & \\ \frac{1}{\Big(v_U(\theta,t_T(\theta))-v_T(\theta,t_T(\theta)^-)\Big)} \times \Big[\frac{d\gamma}{d\theta}(\theta) & \\ -\int_{\xi_p(\theta)}^{t_T(\theta)}\Big(\frac{\partial v_U}{\partial\theta}(\theta,\tau)-\frac{\partial v_T}{\partial\theta}(\theta,\tau)\Big)d\tau & \\ -\sum_{\ell=1}^{L}\Delta v_U(\theta,\tau_{\ell,U}(\theta))\frac{d\tau_{\ell,U}}{d\theta}(\theta) & \\ +\sum_{m=1}^{M}\Delta v_T(\theta,\tau_{m,T}(\theta))\frac{d\tau_{m,T}}{d\theta}(\theta) & \\ +\Big(v_U(\theta,\xi_p(\theta)^+)-v_T(\theta,\xi_p(\theta)^+)\Big)\frac{d\xi_p}{d\theta}(\theta)\Big]. & \qquad (9)\end{aligned}$$

Proof. Since place p is nonempty throughout the interval $(\xi_p(\theta),t_T(\theta))$, (2) implies that

$$m_p(\theta,t_T(\theta)) \;=\; \int_{\xi_p(\theta)}^{t_T(\theta)}\Big(v_U(\theta,\tau)-v_T(\theta,\tau)\Big)d\tau. \qquad (10)$$

Moreover, $m_p(\theta,t_T(\theta))=\gamma(\theta)$. Plugging this in (10) and taking derivatives with respect to

θ we obtain,

$$\begin{aligned}
\frac{d\gamma}{d\theta}(\theta) \;=\; & \Big(v_U(\theta, t_T(\theta)) - v_T(\theta, t_T(\theta)^-)\Big)\frac{dt_T}{d\theta}(\theta) \\
& + \int_{\xi_p(\theta)}^{t_T(\theta)} \Big(\frac{\partial v_U}{\partial \theta}(\theta,\tau) - \frac{\partial v_T}{\partial \theta}(\theta,\tau)\Big) d\tau \\
& + \sum_{\ell=1}^{L} \Delta v_U(\theta, \tau_{\ell,U}(\theta)) \frac{d\tau_{\ell,U}}{d\theta}(\theta) \\
& - \sum_{m=1}^{M} \Delta v_T(\theta, \tau_{m,T}(\theta)) \frac{d\tau_{m,T}}{d\theta}(\theta) \\
& - \Big(v_U(\theta, \xi_p(\theta)^+) - v_T(\theta, \xi_p(\theta)^+)\Big)\frac{d\xi_T}{d\theta}(\theta),
\end{aligned} \tag{11}$$

where we recall that $v_U(\theta, \cdot)$ is continuous at $t = t_T(\theta)$ by Assumption 2.3.2. Now Equation (9) follows from (11) after some algebra. □

We point out that for the case where the endogenous event is the start of an empty period at a place $p \in in(T)$, we have that $v_U(\theta, t_T(\theta)) - v_T(\theta, t_T(\theta)^-) = -\Delta v_T(\theta, t_T(\theta))$ (since $v_U(\theta, t_T(\theta)^+) = v_T(\theta, t_T(\theta)^+)$), and plugging this in (9), we can compute the term $\Delta v_T(\theta, t_T(\theta))$.

All of these formulas indicate how the terms $\Delta v_T(\theta, t_T(\theta))$ can be computable recursively in the network once we specify the relevant laws associated with endogenous events. Rather than describe them in general terms, we present an example that illustrates the general principle.

III. Example: Simple Manufacturing-System Model

Consider the Petri net shown in Fig. 1, representing a manufacturing process whose inventories are controlled by the backorders. Transition T_1 represents the source of product orders, T_2 represents the source of raw material (parts), and T_3 stands for the production operation. The maximum flow rates through these transitions represent the product-order rate, parts' arrival rate, and production capacity rate, respectively. Place p_1 contains the amounts of backorders while p_2 represents a storage facility for inventory parts. The product-order rate is assumed to be an exogenous process denoted by $\{V_1(t)\}$. The parts' arrival rate is denoted by $V_2(\theta, t)$, where θ will be defined shortly. The production capacity rate is assumed to be a constant $V_3 > 0$ for the sake of simplicity of exposition. The contents at the places p_1 and p_2 are denoted by $m_1(\theta, t)$ and $m_2(\theta, t)$, respectively.

Our objective in this example is to investigate the application of IPA to threshold-based flow control, and therefore we consider only the case of controlling $V_2(\theta,t)$ by a threshold parameter at the place p_1. We are cognizant of the fact that typically additional controls are implemented as well (e.g., controlling $V_2(\theta,t)$ by the inventory level $m_2(\theta,t)$), and they will be considered in a later, more comprehensive study.

In the system under study here, $V_2(\theta,t)$ is controlled by $m_1(\theta,t)$ in the following way. Let $\theta>0$ be a threshold control parameter of p_1, and suppose that $V_2(\theta,t)$ is equal to a given low value $V_{2,1}>0$ if $m_1(\theta,t)<\theta$, and to a given higher value $V_{2,2}>V_{2,1}$ as long as $m_1(\theta,t)>\theta$. We make the (reasonable) assumption that $V_{2,1}<V_3\leq V_{2,2}$. Now the idea is to have V_2 switch from $V_{2,1}$ to $V_{2,2}$ whenever m_1 crosses the threshold value θ in the upward direction, and vice versa if m_1 crosses θ downwards. However, it may happen that m_1 rises to θ and then it attempts to decline due to the resulting increase in V_2; consequently it tries to rise again, etc. This causes a jitter over some time-interval, or a sliding mode, and it is due to the fact that the traffic flows are characterized by rates as opposed to the movement of discrete entities such as parts. This situation can arise only whenever $m_1=\theta$ while $m_2=0$ and $V_{2,1}\leq V_1\leq V_3$. In this case, of course, $V_2=V_1$ as long as the jitter continues, and m_1 remains equal to θ. To put all of this formally, $V_2(\theta,t)$ is defined via the following threshold-control law,

$$V_2(\theta,t) \;=\; \begin{cases} V_{2,1}, & \text{if } m_1(\theta,t)<\theta \\ V_{2,2}, & \text{if } m_1(\theta,t)>\theta \\ V_1(t), & \text{if } m_1(\theta,t)=\theta, \quad m_2(\theta,t)=0, \\ & \text{and } V_{2,1}\leq V_1(t)\leq V_3 \\ V_{2,1}, & \text{under all other circumstances.} \end{cases} \tag{12}$$

Cost functions of interest are $J_i(\theta) \;:=\; \int_0^{\mathcal{T}} m_i(\theta,t)dt$, $i=1,2$, and we will analyze their IPA derivatives in the following paragraphs.

In the simulation example that we consider it is assumed that the product orders arrive in batches according to a point process, and hence $\{V_1(t)\}$ is modeled as a sequence of impulses. Regarding the parts' arrival rates, we assume that $V_{2,2}=V_3$ (as well as that $V_{2,1}<V_3$), thereby checking the growth of the parts' inventories during periods of large backorders. We make the following assumption about $V_1(t)$.

***Assumption** 3.1:* The product-order process has the form

$$V_1(t) = \sum_{n=1}^{\infty} \alpha_n \delta(t - s_n), \tag{13}$$

where $\delta(\cdot)$ is the Dirac delta function, α_n, $n = 1, \ldots$, are positive-valued, independent and identically - distributed (iid) random variables, and $\eta_n := s_n - s_{n-1}$, $n = 2, \ldots$, are iid. Furthermore, all of the random variables α_n and η_n are mutually independent, and their distributions have finite first moments and bounded probability-density functions.

This assumption serves as an adequate approximation to bulk arrivals. The presence of impulses implies that $V_1(t)$ can have only the point-values of 0 or ∞, thereby precluding the possibility of the third case in the RHS of Equation (12). Now the assumptions about mutual independence of α_n and η_n is made here to ensure that Assumption 2.3 is satisfied, as will be evident from the discussion in the sequel. Of course it can be relaxed to allow various dependencies while achieving the same goal, and this will be done in a followup paper which will extend the results in this one in several ways. Finally, we mention that while the product-arrival process is defined over an infinite time-horizon, we consider of course only the impulses that occur during the interval $[0, \mathcal{T}]$.

The following proposition summarizes the events that may occur and points out related quantities; its proof is straightforward and hence will be omitted.

Note that we denote by $t_i(\theta)$ the timing of the described event which is associated with transition T_i.

***Proposition** 3.2:* The following list exhausts the possible events in the system and the related quantities are computed accordingly.

1) *Exogenous and null events.* All jumps in $v_1(\cdot)$ are null events, and hence $\frac{dt_1}{d\theta}(\theta) = 0$.
2) *Type-1 endogenous events*. There are three possibilities, as listed below.

   (2.1) *Start of an empty period at* p_1. $v_1(t_3(\theta)) = 0$, and

$$v_3(\theta, t_3(\theta)^-) = \begin{cases} V_{2,1}, & \text{if } m_2(\theta, t_3(\theta)) = 0 \\ V_3, & \text{if } m_2(\theta, t_3(\theta)) > 0. \end{cases} \tag{14}$$

   Next, in the RHS of (9), $\frac{d\gamma}{d\theta}(\theta) = 0$, and all of the other terms there can be assumed to be known (having been computed) by time $t = t_3(\theta)$. Furthermore, if $m_2(\theta, t_3(\theta)) = 0$, this event triggers the end of an empty period at p_2, and in this case $v_3(\theta, t_3(\theta)^+) = 0$ and

$v_2(\theta, t_3(\theta)^+) = V_{2,1}$. (These quantities can serve as the term $v_T(\theta, \xi_p(\theta)^+)$ at the end of Equation (9).)

(2.2) *Start of an empty period at* p_2. Then

$v_2(\theta, t_3(\theta)) = V_{2.1}$, and $v_3(\theta, t_3(\theta)^-) = V_3$. It is impossible to have $m_1(\theta, t_3(\theta)) = 0$. In the RHS of (9), $\frac{d\gamma}{d\theta}(\theta) = 0$.

(2.3) $m_1(\theta, \cdot)$ *crosses* θ *downwards, namely*

$m_1(\theta, t_3(\theta)^-) > \theta$ while $m_1(\theta, t_3(\theta)^+) < \theta$. Then $v_1(\theta, t_3(\theta)) = 0$, and $v_3(\theta, t_3(\theta)^-) = V_3$. Furthermore, this event triggers an induced event at T_2, where $v_2(\theta, t_2(\theta)^-) = V_3$ and $v_2(\theta, t_2(\theta)^+) = V_{2,1}$. In the RHS of (9), $\frac{d\gamma}{d\theta}(\theta) = 1$.

3) *Type-2 endogenous events.* There are two possibilities, as listed below.

(3.1) *End of an empty period at* p_1. This is the result of a jump in $V_1(\cdot)$ which is a null event, and hence $\frac{dt_3}{d\theta}(\theta) = 0$.

(3.2) *End of an empty period at* p_2. This must be triggered by the start of an empty period at p_1, as described in case (2.1), above.

4) *Induced events.* Only the following events are possible.

(4.1) *Jump in* $v_3(\theta, \cdot)$ *triggered by a jump in* $v_1(\cdot)$ *while* p_1 *is empty.* This is a null event and hence $\frac{dt_3}{d\theta}(\theta) = 0$.

(4.2) *Jump in* $v_3(\theta, \cdot)$ *triggered by a jump in* $v_2(\theta, \cdot)$ *while* p_2 *is empty.* A jump down in $v_2(\theta, \cdot)$ is triggered by the type-1 endogenous event described in case (2.3), above. This must happen in the interior of an empty period at p_2, and hence, $v_3(\theta, t_3(\theta)^-) = V_3$ while $v_3(\theta, t_3(\theta)^+) = V_{2,1}$. On the other hand, a jump up in $v_2(\theta, \cdot)$ must be triggered by a jump in $m_1(t)$ upward across θ; it is a null event and hence $\frac{dt_3}{d\theta}(\theta) = 0$.

(4.3) *Jump in* $v_2(\theta, \cdot)$ *induced by* $m_1(\theta, \cdot)$ *crossing the value of* θ. A jump up in $v_2(\theta, \cdot)$ must be triggered by a jump up in $m_1(\theta, t)$ across θ, is a null event and hence $\frac{dt_2}{d\theta}(\theta) = 0$. On the other hand, a jump down in $V_2(\theta, t)$ is described in case (2.3), above. $\square$

Consider now the IPA derivatives $\frac{dJ_{p_i}}{d\theta}(\theta)$, $i = 1, 2$, defined by Equation (6), in whose RHS the integrant is given by Equation (8). In the RHS of (8) the integral term is 0, and the other terms can be computed (recursively) according to the guidelines established in Proposition 3.2.

IV. NUMERICAL SIMULATIONS

To illustrate the effectiveness of the presented approach we now illustrate the results of an optimization problem carried out on the simple Petri net system defined above, assuming as cost function

$$J(\theta) \ := \ \int_0^{\mathcal{T}} (m_1(\theta,t)+m_2(\theta,t))\,dt.$$

In order to maximize $E[J(\theta)]$ we use a stochastic approximation algorithm of the Robbins-Monro type (see Kushner and Clark (1978)) that computes an iteration-sequence $\theta(k) \in R$. The algorithm has the following form:

***Algorithm** 4.1: Data:* $\theta_{\max} > 0$, $\theta(1) \in R$ such that $\theta(1) \in [0, \theta_{\max}]$, a small $\varepsilon > 0$, an upper bound $IPA_{\max}$ on the absolute value of $\frac{dJ}{d\theta}(\theta(k))$, and a positive step-size sequence $\{\lambda_k\}_{k=1}^{\infty}$ satisfying the convergence conditions for Robbins-Monro algorithms, namely $\sum_{k=1}^{\infty} \lambda_k = \infty$ and $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$.
Step 1: Set $k = 1$.
Step 2: Simulate the system for a $\mathcal{T}$-second horizon, and compute the sample derivative $\frac{dJ}{d\theta}(\theta(k))$, by using equations (8)-(9) and Proposition 3.2.
Step 3: If $\left|\frac{dJ}{d\theta}(\theta(k))\right| < IPA_{\max}$, set

$$\theta(k+1) \ = \ \theta(k) - \lambda_k \frac{dJ}{d\theta}(\theta(k)); \tag{15}$$

otherwise, set

$$\theta(k+1) \ = \ \theta(k) - \lambda_k \ IPA_{\max} \ \text{sign}\left(\frac{dJ}{d\theta}(\theta(k))\right). \tag{16}$$

If $\theta(k+1) \notin (\varepsilon, \theta_{\max} - \varepsilon)$, set $\theta(k+1) = \theta(k)$.
Step 4: Set $k = k+1$, and go to Step 1. ■

We remark that Step 3 ensures that $\theta(k)$ remains feasible for all $k = 1, 2, \ldots$, namely $\theta(k) \in [0, \theta_{\max}]$.

Simulations were run with the following parameters. The maximum transition firing rates are equal to $V_{2,1} = 2$ and $V_{2,2} = V_3 = 6$. The product-order process is defined according to eq. (13) assuming that batches arrive at time intervals having constant spacing of 10 time units, i.e., $s_n = 10n$, $n = 1, \ldots,$. The batch-sizes α_n are assumed to be random variables with exponential distribution and average value equal to 50. The step size in Step 3 of the algorithm was chosen to be $\lambda_k = 0.5/k^{0.6}$. Finally, the time-horizon $\mathcal{T}$ has been taken equal to $\mathcal{T} = 1000$ time units.

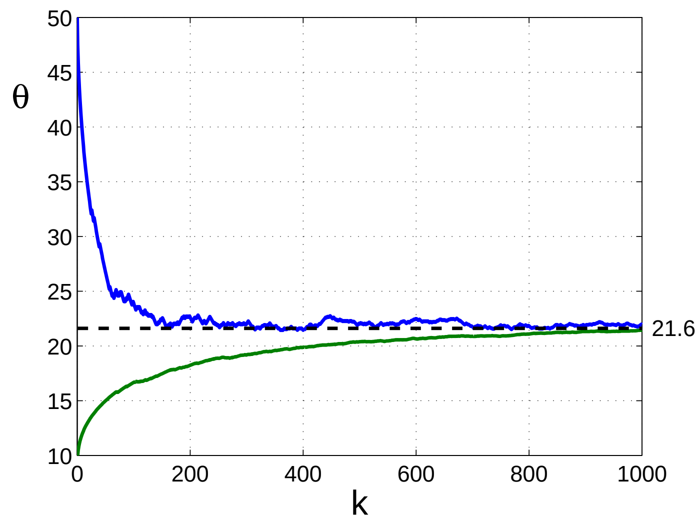

Fig. 2. The graphs of θ vs. k.

The values of θ for $k = 1, \ldots, k_{\max}$ are shown in Fig. 2 assuming two different values of $\theta(1)$, namely $\theta(1) = 50$ and $\theta(1) = 10$. As it can be observed in both cases the value of θ converges to about 21.6. This was supported by extensive simulations and plots $J(\theta)$, not shown here.

V. Conclusions

The main contribution of this paper consists of the application of IPA to a class of continuous Petri nets. A systematic approach for computing the derivatives of the sample performance functions with respect to structural and control parameters has been proposed. An example of a threshold-control variable in a simple yet canonical example has been investigated in detail. Simulation experiments support the theoretical developments and suggest a potential viability of our proposed approach. Future work will consist of extending the results to more general net structures and control variables, and providing a systematic approach for computing the IPA derivatives.

References

Alla, H., and David, R. (1998). Continuous and Hybrid Petri Nets. *Journal of Circuits, Systems, and Computers,* Vol. 8, pp. 159-188.

Cassandras, C., and Lafortune, S. (1999). *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, Boston, Massachusetts.

Cassandras, C. (2006). Stochastic flow systems: Modeling and sensitivity analysis. In *Stochastic Hybrid Systems: Recent Developments and Research Trends,* Eds. C.G. Cassandras and J. Lygeros, CRC Press, New York, New York, pp. 137–165.

Giua, A., Seatzu, C., and Wardi, Y. (2010). IPA for Continuous Petri Nets. In *Proc. 10th Workshop on Discrete Event Systems (WODES'10)*, Berlin, Germany, August 30 – September 1.

Kushner, H., and Clark, D. (1978). *Stochastic Approximation for Constrained and Unconstrained Systems*. Berlin, Germany: Springer-Verlag, 1978.

Silva, M. and Recalde, L. (2004). On fluidification of Petri net models: from discrete to hybrid and continuous models. *Annual Reviews in Control,* Vol. 28, pp. 253-266.

Xie, X. (2002) Fluid-Stochastic-Event Graphs for Evaluation and Optimization of Discrete-Event Systems with Failures *IEEE Transactions on Robotics and Automation*, Vol. 18, pp. 360-367.