[2] J. Zaborzsky, G. Huang, B. Zheng, and T. C. Leung, "On the phase-portrait of a class of large nonlinear dynamic systems such as the power system," *IEEE Trans. Autom. Control*, vol. 33, no. 1, pp. 4–15, Jan. 1988.

[3] H. D. Chiang, C. C. Chu, and G. Cauley, "Direct stability analysis of electric power systems using energy functions: Theory, applications, and perspective," in *Proc. IEEE*, Nov. 1995, vol. 83, no. 11, pp. 1497–1529.

[4] S. Ushiki, "Analytic expressions of unstable manifolds," *Proc. Jpn Acad. Ser. A.-Math. Sci.*, vol. 56, no. 6, pp. 239–244, 1980.

[5] F. M. A. Salam, A. Arapostathis, and P. P. Varaiya, "Analytic expressions for the unstable manifold at equilibrium points in dynamical systems of differential equations," in *Proc. 22nd IEEE Conf. Decision Control*, Dec. 1983, vol. 3, pp. 1389–1392.

[6] S. Saha, A. A. Fouad, W. H. Kliemann, and V. Vittal, "Stability boundary approximation of a power system using the real normal form of vector fields," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 797–802, 1997.

[7] R. Qi, D. Cook, W. Kliemann, and V. Vittal, "Visualization of stable manifolds and multidimensional surfaces in the analysis of power system dynamics," *J. Nonlin. Sci.*, vol. 10, pp. 175–195, 2000.

[8] Z. Jing, Z. Jia, and Y. Gao, "Research of the stability region in a power system," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 2, pp. 298–304, Feb. 2003.

[9] P. A. Cook and A. M. Eskicioglu, "Transient stability analysis of electric power systems by the method of tangent hypersurfaces," in *Inst. Elect. Eng. Proc.-Gener. Transm. Distrib.*, Jul. 1983, vol. 130, no. 4, pp. 183–193.

[10] J. Zaborzsky, G. Huang, B. Zheng, and T. Leung, "New results on stability monitoring on the large electric power systems," in *Proc. 26th IEEE Conf. Decision Control*, 1987, pp. 30–40.

[11] M. Djukanovic, D. Sobajic, and Y.-H. Pao, "Neural-net based tangent hypersurfaces for transient security assessment of electric power systems," *Int. J. Electr. Power Energy Syst.*, vol. 16, no. 6, pp. 399–408, 1994.

[12] V. Venkatasubramanian and W. Ji, "Numerical approximation of $(n-1)$-dimensional stable manifolds in large systems such as the power system," *Automatica*, vol. 33, no. 10, pp. 1877–1883, Oct. 1997.

[13] D. Cheng, J. Ma, Q. Lu, and S. Mei, "Quadratic form of stable submanifold for power systems," *Int. J. Robust Nonlin. Contr.*, vol. 14, pp. 773–788, 2004.

[14] N. Kazantzis, "Singular PDEs and the problem of finding invariant manifold for nonlinear dynamical system," *Phys. Lett. A.*, vol. 272, no. 4, pp. 257–263, Jul. 2000.

IPA Derivatives for Make-to-Stock Production-Inventory Systems With Lost Sales

Yao Zhao and Benjamin Melamed

***Abstract*—This note applies the stochastic fluid model (SFM) paradigm to a class of single-stage, single-product make-to-stock (MTS) production-inventory systems with stochastic demand and random production capacity, where the finished-goods inventory is controlled by a continuous-time base-stock policy and unsatisfied demand is lost. This note derives formulas for infinitesimal perturbation analysis (IPA) derivatives of the sample-path time averages of the inventory level and lost sales with respect to the base-stock level and a parameter of the production rate process. These formulas are comprehensive in that they are exhibited for any initial inventory state, and include right and left derivatives (when they differ). The formulas are obtained via sample path analysis under very mild regularity assumptions, and are inherently nonparametric in the sense that no specific probability law need be postulated. It is further shown that all IPA derivatives under study are unbiased and fast to compute, thereby providing the theoretical basis for online adaptive control of MTS production-inventory systems.**

***Index Terms*—Infinitesimal perturbation analysis (IPA), lost sales, make-to-stock (MTS), production-inventory systems, stochastic fluid models (SFMs).**

Manuscript received March 1, 2006; revised November 16, 2006. Recommended by Associate Editor I. Paschalidis.

The authors are with the Department of Management Science and Information Systems, Rutgers Business School, Newark and New Brunswick, NJ 08854 USA (e-mail: yaozhao@andromeda.rutgers.edu; melamed@rbs.rutgers.edu).

Digital Object Identifier 10.1109/TAC.2007.902760

I. INTRODUCTION

This note derives infinitesimal perturbation analysis (IPA) derivatives of selected random variables for a class of make-to-stock (MTS) systems in stochastic fluid model (SFM) setting, where the traditional discrete arrival, service, and departure stochastic processes are replaced by corresponding stochastic fluid-flow rate processes. We henceforth refer to this approach as *IPA-over-SFM*. The IPA derivatives provide sensitivity information on system metrics with respect to control parameters of interest, and as such can serve as the theoretical underpinnings for online control algorithms. Comprehensive discussions of IPA derivatives and their applications can be found in Fu and Hu [3] and Cassandras and Lafortune [1].

The *IPA-over-SFM* approach has been successfully applied to theoretical studies of various queuing and production-inventory systems; see, e.g., [2], [6], [4], and [7]. These studies constrain the system to start from a prescribed initial inventory state, and only consider cases where the left and right IPA derivatives coincide. In contrast, Zhao and Melamed [8] considered any initial inventory state for MTS systems with backorders and derived sided IPA derivative formulas as needed. The goal of this note is to derive IPA derivatives for MTS systems with lost sales, and to show them to be unbiased. First, we derive IPA derivatives for the time averages of inventory level and lost sales with respect to the base-stock level for *all* initial inventory states, including sided derivatives when they differ. We are only aware of one note [6] addressing *IPA-over-SFM* queues with finite buffers, which can be used to model MTS systems with lost sales, though it constrains the initial condition to an empty buffer. Second, we derive IPA derivatives for the aforementioned metrics with respect to a production rate parameter, including sided derivatives when they differ. We point out that the assumptions in [6] preclude differing left and right IPA derivatives. As will become evident in the sequel, MTS systems with lost sales are also analytically more challenging than MTS systems with backorders, because the inventory state of the former has an extra boundary, a fact that results in more elaborate formulas.

The computation of IPA derivatives for *all* initial conditions is motivated by two reasons. The first reason is to enable potential applications of IPA derivatives to online control of MTS systems driven by nonstationary processes among others. The intent is to adjust the system parameters over time according to the changing statistics of the underlying processes, but not necessarily to optimize it. Clearly, IPA-based online control applications mandate the computation of IPA derivatives for *all* initial states, as well as all sided derivatives when they differ, since a control action can change the system parameters at any state (which is then considered as the new initial state). It makes little sense to wait for the system to return to certain selected inventory states as this could suspend control actions over extended periods of time. The second reason is that the transient IPA derivatives computed here are exact and unbiased, whereas their asymptotic counterparts may not provide adequate approximations. Furthermore, in order to compute asymptotic IPA derivatives, we still need to obtain their transient counterparts before sending time to infinity.

The rest of this note is organized as follows: Section II presents the MTS model. Sections III and IV derive IPA derivative formulas and prove their unbiasedness. Section V concludes the note. We will use the following notational conventions. Let the indicator function of set A be 1_A and $x^+ = \max\{x, 0\}$. A function $f(x)$ is said to be *locally differentiable at* x if it is differentiable in a neighborhood of x; it is said to be *locally independent of* x if it is constant in a neighborhood of x.

II. MTS Model With Lost Sales

Consider a single-stage, single-product MTS system, consisting of a production facility and an inventory facility. The two facilities are coupled: The latter sends orders to the former, while the former produces stock to replenish the latter. The production facility is comprised of a queue that houses a production server, preceded by an infinite buffer that holds incoming production orders. We assume that the production facility has an unlimited supply of raw material, so it never starves. The inventory facility satisfies incoming demands on a first-come–first-serve (FCFS) basis, and is controlled by a continuous-time base-stock policy with a base-stock level $S > 0$ (the case $S = 0$ corresponds to the just-in-time policy as a simple special case). Demands arrive at the inventory facility and are satisfied from inventory on hand (if available). Otherwise, when an inventory shortage is encountered, the incoming demand is satisfied by the amount of inventory on hand, and any shortage of inventory becomes a lost sale.

The MTS system can operate in one of the following two modes. In the *normal* operational mode, the inventory level does not exceed S. The production facility strives to replenish the inventory facility to its base-stock level, but no higher. In the *overage* operational mode, the inventory level exceeds S (for example, this could result from a control action that lowered S below the current inventory level). Production is then temporarily suspended until the inventory level reaches or crosses S from above, at which point normal operational mode resumes.

The MTS system with lost sales can be modeled as an SFM, where $I(t)$ is the (fluid) volume of inventory on-hand at time t, $X(t)$ is the (fluid) volume of outstanding orders at time t, $\alpha(t)$ represents the rate of incoming demands at time t, and $\mu(t)$ represents the production rate at time t. Finally, $\zeta(t)$ is the (fluid) loss rate of sales at time t. Let $[0, T]$ be a finite time interval; for example, T may designate the time period separating applications of control actions.

The notion of *sample path events* pertains to a property of a time point along a sample path (not to be confused with the ordinary notion of events as aggregates of sample paths); the distinction can be discerned by context. Similarly to [6], we define two types of sample path events: An *exogenous event* occurs either whenever a jump takes place in the sample path of $\{\alpha(t)\}$ or $\{\mu(t)\}$, or when the time horizon T is reached. An *endogenous event* occurs whenever a time interval is inaugurated, in which $X(t) = 0$ or $X(t) = S$. Throughout this note, we make the following mild regularity assumptions (cf. [6]).

Assumption 1:

1) The demand rate process $\{\alpha(t)\}$ and the production rate process $\{\mu(t)\}$ have right-continuous sample paths that are piecewise-constant with probability 1 (w.p.1.).
2) Each of the processes $\{\alpha(t)\}$ and $\{\mu(t)\}$ has a finite number of discontinuities in any finite time interval w.p.1, and the time points at which the discontinuities occur are independent of the parameters of interest.
3) No multiple sample path events occur simultaneously w.p.1.

In overage operational mode, the system satisfies the relations $(d/dt^+)I(t) = -\alpha(t), \zeta(t) = 0$, and $X(t) = 0$. In normal operational mode, the system satisfies the conservation relation

$$X(t) + I(t) = S. \tag{2.1}$$

The lost-sales rate process is given by $\zeta(t) = [\alpha(t) - \mu(t)]\,1_{\{I(t)=0,\alpha(t)>\mu(t)\}}, \quad t \geq 0$. The outstanding orders process is governed by the sided stochastic differential equation: $(d/dt^+)X(t) = 0$ if $X(t) = 0$ and $\alpha(t) \leq \mu(t)$ or $X(t) = S$ and $\alpha(t) \geq \mu(t)$; otherwise, $(d/dt^+)X(t) = \alpha(t) - \mu(t)$.

We consider the following *performance random variables* or simply *metrics*: the inventory time average $L_I(T) = (1/T)\int_0^T I(t)\,dt$ and the lost-sales time average $L_\zeta(T) = (1/T)\int_0^T \zeta(t)\,dt$. The control parameters of interest are the base-stock level at the inventory facility and a scaling parameter of the production rate at the production facility. Let $\theta \in \Theta$ denote a generic parameter of interest with a close and bounded domain Θ. We write $S(\theta), \mu(t,\theta), L_I(T,\theta), L_\zeta(T,\theta)$, and so on to indicate the dependence of a performance random variable on its parameter of interest. Our objective is to derive formulas for the IPA derivatives $(d/d\theta)L_I(T,\theta)$ and $(d/d\theta)L_\zeta(T,\theta)$ in SFM setting, using sample path analysis, and to show them to be unbiased.

Finally, we make the following assumption regarding the initial state.

Assumption 2: The initial inventory level does not depend on θ, i.e., $I(0,\theta) = I(0), \forall\theta \in \Theta$.

III. IPA Derivatives With Respect to the Base-Stock Level

This section derives IPA derivatives (including sided ones) for the inventory time average, $L_I(T,\theta)$ and the lost-sales time average $L_\zeta(T,\theta)$, both with respect to the base-stock level S. It will exhibit the requisite formulas for any initial inventory state.

Assumption 3:

1) $S(\theta) = \theta$, where $\theta \in \Theta$.
2) The processes $\{\alpha(t)\}$ and $\{\mu(t)\}$ are independent of the parameter θ.
3) For each $\theta \in \Theta$, the sided derivatives of $L_I(T,\theta)$ and $L_\zeta(T,\theta)$ exist w.p.1.

Let $(Q_j(\theta), R_j(\theta)), j = 1, \ldots, J(\theta)$, be the ordered extremal subintervals of $[0,\infty)$, such that $I(t,\theta) < S$ for all $t \in (Q_j, R_j)$. That is, the endpoints $Q_j(\theta)$ and $R_j(\theta)$ are obtained via $\inf$ and $\sup$ functions, respectively. By convention, if any of these endpoints does not exist, then it is set to ∞. Furthermore, let $Z_j(\theta) \in (Q_j(\theta), R_j(\theta))$ be the first time point in this interval at which $I(t,\theta) = 0$, provided such a point exists; otherwise, let $Z_j(\theta) = R_j(\theta)$. By [8, Observation 3], $Q_1(\theta) < R_1(\theta) < Q_2(\theta) < R_2(\theta) < \ldots < Q_{J(\theta)}(\theta) < R_{J(\theta)}(\theta)$.

We consider three initial states: $I(0) < S(\theta), I(0) > S(\theta)$, and $I(0) = S(\theta)$. The last state cannot be excluded because it may happen in applications where inventory levels are discrete. On the event $\{I(0) < S(\theta)\}\bigcup\{I(0) > S(\theta)\}$, we will make use of the hitting time $T_S(\theta) = \min\{t \in [0,\infty] : I(t,\theta) = S(\theta)\}$, if it exists; otherwise, define $T_S(\theta) = \infty$. On the event $\{I(0) = S(\theta)\}$, we make use of the hitting times $T_\mu(\theta)$ and $T_\alpha(\theta)$. On the event $\{Q_1(\theta) > 0\}$, define $T_\mu(\theta) = \min\{t \in [0, Q_1(\theta)) : \mu(t) > \alpha(t)\}$, if it exists; on the event $\{Q_1(\theta) = 0\}\bigcup[\{Q_1(\theta) > 0\}\bigcap\{\alpha(t) = \mu(t),\ t \in [0, Q_1(\theta))\}]$ define $T_\mu(\theta) = R_1(\theta)$, if $R_1(\theta)$ exists; otherwise, define $T_\mu(\theta) = \infty$. Finally, define $T_\alpha = \min\{t \in [0,T] : \alpha(t) > 0\}$, if it exists; otherwise, define $T_\alpha = \infty$. In words, $T_\mu(\theta)$ is a hitting time of $\{I(t,\theta)\}$, which corresponds to the first time that the inventory level changes in any perturbed process $\{I(t,\theta+\Delta\theta)\}$, while T_α plays an analogous role, but for a perturbed process $\{I(t,\theta-\Delta\theta)\}$. Note that T_α is independent of θ. We will also make use of horizon-dependent random indices, given by $J_S(T,\theta) = \max\{j \geq 1 : R_j(\theta) \leq T\}$, if it exists, and $J_S(T,\theta) = 0$, otherwise. These constitute restrictions of $J(\theta)$ to finite time horizons $[0,T]$.

Theorem 1: W.p.1, the IPA derivatives of the inventory time average with respect to the base-stock level are given for all $T > 0$ and $\theta \in \Theta$ as follows:

1) On the event $\{I(0) < S(\theta)\}$

$$\frac{d}{d\theta}L_I(T,\theta) = \frac{1}{T}\sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)].$$

2) On the event $\{I(0) > S(\theta)\} \bigcap \{T_S(\theta) < Q_1(\theta)\}$

$$\frac{d}{d\theta}L_I(T,\theta) = \frac{1}{T}\{1_{\{T_S(\theta)<T\}}[\min\{Z_1(\theta), T\} - T_S(\theta)] + \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]\}.$$

3) On the event $\{I(0) > S(\theta)\} \bigcap \{T_S(\theta) = Q_1(\theta)\}$

$$\frac{d}{d\theta}L_I(T,\theta) = \frac{1}{T}\{1_{\{T_S(\theta)<T\}}\frac{\mu(T_S(\theta))}{\alpha(T_S(\theta))}[\min\{Z_1(\theta), T\} - T_S(\theta)] + \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]\}.$$

4) On the event $\{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) = R_1(\theta)\}$

$$\frac{d}{d\theta^+}L_I(T,\theta) = \frac{1}{T}\left\{\sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]\right\}.$$

5) On the event $\{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) < R_1(\theta)\}$

$$\frac{d}{d\theta^+}L_I(T,\theta) = \frac{1}{T}\{1_{\{T_\mu(\theta)<T\}}[\min\{Z_1(\theta), T\} - T_\mu(\theta)] + \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]\}.$$

6) On the event $\{I(0) = S(\theta)\} \bigcap \{T_\alpha < Q_1(\theta)\}$

$$\frac{d}{d\theta^-}L_I(T,\theta) = \frac{1}{T}\{1_{\{T_\alpha<T\}}[\min\{Z_1(\theta), T\} - T_\alpha] + \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]\}.$$

7) On the event $\{I(0) = S(\theta)\} \bigcap \{T_\alpha = Q_1(\theta)\}$

$$\frac{d}{d\theta^-}L_I(T,\theta) = \frac{1}{T}\{1_{\{T_\alpha<T\}}\frac{\mu(T_\alpha)}{\alpha(T_\alpha)}[\min\{Z_1(\theta), T\} - T_\alpha] + \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]\}.$$

Proof: First note that in normal operational mode, an outstanding order (a lost sale, respectively) of the MTS system is equivalent to a parcel of workload (a lost volume, respectively) in the production facility's queue (cf. [6] which considers the initial state $X(0,\theta) = 0$, i.e., $I(0,\theta) = S(\theta)$).

To prove part 1), we write $L_I(T,\theta) = (1/T)[\int_0^{\min\{T_S(\theta),T\}} I(t,\theta)dt + \int_{\min\{T_S(\theta),T\}}^{\min\{Q_2,T\}} I(t,\theta)dt + \int_{\min\{Q_2,T\}}^{T} I(t,\theta)dt]$. By standard arguments (see [2], [6], and [8]), Q_2 is locally independent of θ, and in the time interval $[Q_2, T]$, the system always starts with full inventory in a neighborhood of θ. Taking the derivative of the previous equation and noting that $I(t,\theta)$ is independent of θ on $\{t < T_S(\theta)\}$, $I(t,\theta) = S(\theta)$ on $\{T_S(\theta) < t \le Q_2\}$, and that the terms associated with $(d/d\theta)T_S$ vanish, part 2) follows from (2.1) and [6, Proposition 3.2].

Part 2) follows from the proof of part 1) if we replace Q_2 with Q_1.

To prove part 3), we assume, without loss of generality, $T > R_1(\theta) \ge Z_1(\theta) > T_S(\theta) = Q_1(\theta)$. Thus, $L_I(T,\theta) = (1/T)[\int_0^{T_S(\theta)} I(t,\theta)dt + \int_{T_S(\theta)}^{Z_1(\theta)} I(t,\theta)dt + \int_{Z_1(\theta)}^{R_1(\theta)} I(t,\theta)dt + \int_{R_1(\theta)}^{T} I(t,\theta)dt]$. Clearly, $I(t,\theta)$ does not depend on θ on $\{0 \le t < T_S(\theta)\}$. By standard arguments (see [2], [6], and [8]), $I(t,\theta)$ does not depend on θ on $\{Z_1(\theta) < t < R_1(\theta)\}$. On $\{T_S(\theta) < t < Z_1(\theta)\}$, $I(t,\theta) = S(\theta) + \int_{T_S(\theta)}^{t}[\mu(\tau) - \alpha(\tau)]d\tau$. Differentiation with respect to θ yields $(d/d\theta)I(t,\theta) = 1 - [\mu(T_S(\theta)) - \alpha(T_S(\theta))]\,(d/d\theta)T_S(\theta)$. Moreover, $\int_0^{T_S(\theta)} \alpha(\tau)\,d\tau = I(0) - S(\theta)$ implies (by differentiation)

$$\frac{d}{d\theta}T_S(\theta) = \frac{-1}{\alpha(T_S(\theta))}. \tag{3.2}$$

Hence, $(d/d\theta)I(t,\theta) = \mu(T_S(\theta))/\alpha(T_S(\theta))$, and, after some algebra, part 3) is obtained.

To prove part 4), we write $L_I(T,\theta) = 1/T[\int_0^{\min\{T_\mu(\theta),T\}} I(t,\theta)dt + \int_{\min\{T_\mu(\theta),T\}}^{T} I(t,\theta)dt]$. By definition of $T_\mu(\theta)$, $I(t,\theta) = I(t,\theta+\Delta\theta)$ on $\{0 \le t \le T_\mu(\theta)\}$ for sufficiently small $\Delta\theta > 0$. Moreover, Q_2 is locally independent of θ, and $I(t,\theta) = S(\theta)$ on $\{T_\mu(\theta) < t \le Q_2\}$. Part 4) now holds by the proof of part 1). Part 5) holds by a similar proof if we replace Q_2 by Q_1. Finally, for parts 6) and 7), consider $\theta - \Delta\theta$. The system starts in overage operational mode because $I(0) = S(\theta) > S(\theta - \Delta\theta)$. By definition, T_α is the limiting time point at which $\{I(t,\theta - \Delta\theta)\}$ first hits $S(\theta - \Delta\theta)$ from above as $\Delta\theta \to 0$; so T_α is equivalent to the $T_S(\theta)$ of parts 2) and 3). Replacing $T_S(\theta)$ by T_α, and using arguments similar to those in the proof of parts 2) and 3) proves parts 6) and 7).■

We next derive the IPA derivatives for $L_\zeta(T,\theta)$. For any time interval $[a,b]$, let $N_{[a,b]}(\theta)$ be the number of intervals of the form $[Q_j(\theta), R_j(\theta)]$, such that $Z_j(\theta) < R_j(\theta)$ (i.e., lost sales actually occur) and $Z_j(\theta) \in [a,b]$.

Theorem 2: W.p.1, the IPA derivatives of the lost-sales time average with respect to the base-stock level are given for all $T > 0$ and $\theta \in \Theta$ as follows.

1) On the event $A(\theta) = \{I(0) < S(\theta)\}$ and $B(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) < Q_1(\theta)\}$

$$\frac{d}{d\theta}L_\zeta(T,\theta) = -\frac{N_{(T_S(\theta),T]}(\theta)}{T}.$$

2) On the event $C(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) = Q_1(\theta)\}$

$$\frac{d}{d\theta}L_\zeta(T,\theta) = -\frac{1}{T}\left[1_{\{Z_1(\theta)<R_1(\theta), Z_1(\theta)<T\}}\frac{\mu(T_S(\theta))}{\alpha(T_S(\theta))} + N_{(R_1(\theta),T]}(\theta)\right].$$

3) On the event $D(\theta) = \{I(0) = S(\theta)\}$

$$\frac{d}{d\theta^+}L_\zeta(T,\theta) = -\frac{N_{(T_\mu(\theta),T]}(\theta)}{T}.$$

4) On the event $E(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\alpha < Q_1(\theta)\}$

$$\frac{d}{d\theta^-}L_\zeta(T,\theta) = -\frac{N_{(T_\alpha,T]}(\theta)}{T}.$$

5) On the event $F(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\alpha = Q_1(\theta)\}$

$$\frac{d}{d\theta^-}L_\zeta(T,\theta) = -\frac{1}{T}\left[1_{\{Z_1(\theta)<R_1(\theta), Z_1(\theta)<T\}}\frac{\mu(T_\alpha)}{\alpha(T_\alpha)} + N_{(R_1(\theta),T]}(\theta)\right].$$

Proof: For part 1) and the event $A(\theta)$, we write $L_\zeta(T,\theta) = (1/T)[\int_0^{\min\{T_S(\theta),T\}} \zeta(t,\theta)dt + \int_{\min\{T_S(\theta),T\}}^{\min\{Q_2,T\}} \zeta(t,\theta)dt + \int_{\min\{Q_2,T\}}^{T} \zeta(t,\theta)dt]$. By the proof of part 1) in Theorem 1, Q_2 is locally independent of θ, and in the time interval $[Q_2, T]$, the system starts with full inventory in a neighborhood of θ. Differentiating the previous equation and noting that $\zeta(t,\theta)$ is independent of θ on $\{0 \le t < T_S(\theta)\}$ and $\zeta(t,\theta) = 0$ on $\{T_S(\theta) \le t \le Q_2\}$, part 1) follows from [6, Proposition 3.1]. The same proof applies to the event $B(\theta)$ of part 1), provided Q_2 is replaced by Q_1.

For part 2), we assume without loss of generality $R_1(\theta) < T$, and write $L_\zeta(T,\theta) = (1/T)[\int_0^{R_1(\theta)} \zeta(t,\theta)dt + \int_{R_1(\theta)}^{Q_2} \zeta(t,\theta)dt + \int_{Q_2}^{T} \zeta(t,\theta)dt]$. If $Z_1(\theta) = R_1(\theta)$, then there are no lost sales in $(Q_1(\theta), R_1(\theta))$, whereas if $Z_1(\theta) < R_1(\theta)$, then some lost sales occurred in $(Q_1(\theta), R_1(\theta))$. Let $[U_k(\theta), V_k(\theta)], k = 1,\dots,K(\theta)$, be extremal subintervals of $(Q_1(\theta), R_1(\theta))$ over which $I(t,\theta) = 0$. By standard arguments (see [6]), $U_k(\theta)$ is locally differentiable in θ, $V_k(\theta)$ is locally independent of θ, and

$$\frac{d}{d\theta}\int_0^{R_1(\theta)} \zeta(t,\theta)dt = \sum_{k=1}^{K(\theta)} \frac{d}{d\theta}\int_{U_k(\theta)}^{V_k} [\alpha(t) - \mu(t)]dt = -[\alpha(U_1(\theta)) - \mu(U_1(\theta))]\frac{d}{d\theta}U_1(\theta). \tag{3.3}$$

Because $T_S(\theta) = Q_1(\theta)$ by assumption, $\int_{T_S(\theta)}^{U_1(\theta)}(\alpha(t) - \mu(t))dt = S(\theta)$. Differentiating the above yields $[\alpha(U_1(\theta)) - \mu(U_1(\theta))](d/d\theta)U_1(\theta) - [\alpha(T_S(\theta)) - \mu(T_S(\theta))](d/d\theta)T_S(\theta) = 1$. By (3.2) and (3.3), $(d/d\theta)\int_0^{R_1(\theta)} \zeta(t,\theta)dt = -\mu(T_S(\theta))/\alpha(T_S(\theta))$. Since $\zeta(t,\theta) = 0$ on $\{R_1(\theta) \le t \le Q_2\}$, part 2) holds by [6, Proposition 3.1].

For part 3, let $T_\mu(\theta) < T$ for simplicity, and write $L_\zeta(T,\theta) = (1/T)[\int_0^{T_\mu(\theta)} \zeta(t,\theta)dt + \int_{T_\mu(\theta)}^{T} \zeta(t,\theta)dt]$. By definition of $T_\mu(\theta), I(t,\theta) = I(t,\theta+\Delta\theta)$ for $t \in [0, T_\mu(\theta))$ for sufficiently small $\Delta\theta > 0$. Furthermore, $\zeta(t,\theta) = 0$ on $\{T_\mu(\theta) \le t \le Q_2\}$, if $T_\mu(\theta) = R_1(\theta)$; similarly, $\zeta(t,\theta) = 0$ on $\{T_\mu(\theta) \le t \le Q_1\}$, if $T_\mu(\theta) < R_1(\theta)$. Thus, $(d/d\theta^+)\int_0^{T_\mu(\theta)} \zeta(t,\theta)dt = 0$ and part 3) holds by [6, Proposition 3.1].

For parts 4) and 5), consider $\theta - \Delta\theta$. Because $I(0) = S(\theta) > S(\theta - \Delta\theta)$, the system starts in overage operational mode. Parts 4) and 5) follow by the proof of event $B(\theta)$ of part 1), because Theorem 1 shows that in these cases, T_α plays a role analogous to $T_S(\theta)$. ∎

Theorem 3: Under Assumptions 1–3, the sided IPA derivatives with respect to the base-stock level $(d/d\theta^\pm)L_I(T,\theta)$ and $(d/d\theta^\pm)L_\zeta(T,\theta)$ are unbiased for all $T > 0$ and $\theta \in \Theta$.

Proof: Part 3) of Assumption 3 ensures that for all $T > 0$, [5, Assumption C4 of Lemma A2, p. 70] holds for both $L_I(T,\theta)$ and $L_\zeta(T,\theta)$. Because $\mu(T_S(\theta))/\alpha(T_S(\theta)) < 1$ and $(\mu(T_\alpha)/\alpha(T_\alpha)) < 1$ in Theorems 1 and 2, it follows that $0 \le (d/d\theta^\pm)L_I(T,\theta) \le 1$ by Theorem 1. Furthermore, by Theorem 2, $(d/d\theta^\pm)L_\zeta(T,\theta) \le 1 + N_{[0,T]}(\theta)/T$, where $E[N_{[0,T]}(\theta)]$ is finite because $N_{[0,T]}(\theta)$ is finite w.p.1 by part 2) of Assumption 1. Since the sample performance functions are continuous and piecewise differentiable, the one-sided derivatives exist for every θ. All IPA derivatives are bounded. Hence, [5, Assumption C3 of Lemma A2] is in force, which completes the proof. ∎

IV. IPA Derivatives With Respect to a Production Rate Parameter

This section derives sided IPA derivatives for the inventory time average $L_I(T,\theta)$ and the lost-sales time average $L_\zeta(T,\theta)$, both with respect to a production rate parameter θ of $\{\mu(t,\theta)\}$.

Assumption 4:

1) $(d/d\theta)\,\mu(t,\theta) = 1$, where $t \in [0,T]$ and $\theta \in \Theta$.
2) The process $\{\alpha(t)\}$ and the base-stock level S are independent of θ.
3) For each $\theta \in \Theta$, the sided derivatives of $L_I(T,\theta)$ and $L_\zeta(T,\theta)$ exist w.p.1.

We point out that unlike [6], [4], and [7], Assumption 4 admits the possibility that sided IPA derivatives do not coincide. Indeed, this could happen on events of the form $\{I(t,\theta) = S\}\bigcap\{\alpha(t) = \mu(t,\theta)\}$ and $\{I(t,\theta) = 0\}\bigcap\{\alpha(t) = \mu(t,\theta)\}$. These are generally not rare events, and in practice, their probabilities may well not vanish, because $I(t,\theta) = S$ or $I(t,\theta) = 0$ could hold for an extended period of time, and by part 1) of Assumption 1, $\{\alpha(t)\}$ and $\{\mu(t,\theta)\}$ have sample paths that are piecewise-constant w.p.1.

In this section, we may assume without loss of generality that $0 \le I(0) \le S$, since the production facility suspends replenishment in overage operational mode, so that the value of θ has no effect on the state of the system until it enters normal operational mode. Define $(U_m^+(\theta), V_m^+(\theta)), m = 1,\dots,M(\theta)$, to be the ordered extremal subintervals of $[0,\infty)$ such that for all $t \in (U_m^+, V_m^+)$, either $I(t,\theta) = S$ holds or both $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ hold. Define further $(U_n^-(\theta), V_n^-(\theta)), n = 1,\dots,N(\theta)$, to be the ordered extremal subintervals of $[0,\infty)$ such that for all $t \in (U_n^-, V_n^-)$, either $I(t,\theta) = 0$ holds or both $I(t,\theta) = S$ and $\alpha(t) < \mu(t,\theta)$ hold. By convention, if any of the aforementioned endpoints does not exist, then it is set to ∞. For notational convenience, define $V_0^+(\theta) = V_0^-(\theta) = 0$. By the continuity of $\{I(t,\theta)\}$ in t and part 3) of Assumption 1, $U_1^+(\theta) < V_1^+(\theta) < U_2^+(\theta) < V_2^+(\theta) < \cdots < U_{M(\theta)}^+(\theta) < V_{M(\theta)}^+(\theta)$ and $U_1^-(\theta) < V_1^-(\theta) < U_2^-(\theta) < V_2^-(\theta) < \cdots < U_{N(\theta)}^-(\theta) < V_{N(\theta)}^-(\theta)$.

We will need the following horizon-dependent random indices. The restriction of $M(\theta)$ to a finite time horizon $[0,T]$ is $M_I(T,\theta) = \max\{m \ge 1 : V_m^+(\theta) \le T\}$, if it exists, and zero, otherwise. The restriction of $N(\theta)$ to a finite time horizon $[0,T]$ is $N_I(T,\theta) = \max\{n \ge 1 : V_n^-(\theta) \le T\}$, if it exists, and zero, otherwise.

Theorem 4: W.p.1, the IPA derivatives of the inventory time average with respect to the production rate parameter are given for all $T > 0$ and $\theta \in \Theta$ as follows:

$$\frac{d}{d\theta^+}L_I(T,\theta) = \frac{1}{2T}\sum_{m=0}^{M_I(T,\theta)}[\min\{U_{m+1}^+(\theta), T\} - V_m^+(\theta)]^2 \tag{4.4}$$

$$\frac{d}{d\theta^-}L_I(T,\theta) = \frac{1}{2T}\sum_{n=0}^{N_I(T,\theta)}[\min\{U_{n+1}^-(\theta), T\} - V_n^-(\theta)]^2. \tag{4.5}$$

Proof: We only prove (4.4) since the proof of (4.5) is analogous. By part 3) of Assumption 4 and Leibniz's rule, $(d/d\theta^\pm)L_I(T,\theta) = (1/T)(d/d\theta^\pm)\int_0^T I(t,\theta)\,dt = (1/T)\int_0^T (d/d\theta^\pm)I(t,\theta)\,dt$. We next compute $(d/d\theta^+)I(t,\theta)$. By definition of $(U_m^+(\theta), V_m^+(\theta))$, one has $I(t,\theta) = I(t,\theta+\Delta\theta)$ on the events $\{U_m^+(\theta) < t < V_m^+(\theta)\}, m = 1,\dots,M(\theta)$, for sufficiently small $\Delta\theta$. On the events $\{V_m^+(\theta) < t < U_{m+1}^+(\theta)\}, m = 0,1,\dots,M(\theta)-1$, one has $I(t,\theta) = I(V_m^+(\theta),\theta) + \int_{V_m^+(\theta)}^t [\mu(\tau,\theta) - \alpha(\tau)]\,d\tau$. By standard arguments (see [2], [6], and [8]), $V_m^+(\theta)$ and, therefore, $I(V_m^+(\theta),\theta)$ are each locally independent of θ in a right neighborhood of θ. Consequently, $(d/d\theta^+)I(t,\theta) = -[\mu(V_m^+(\theta),\theta) - \alpha(V_m^+(\theta))](d/d\theta^+)V_m^+(\theta) + \int_{V_m^+(\theta)}^t d\tau = t - V_m^+(\theta)$. A modicum of algebra completes the proof. ∎

To derive IPA derivatives for $L_\zeta(T,\theta)$, we will need the following horizon-dependent random indices: $M_\zeta(T,\theta) = \max\{m \ge 1 : U_m^+(\theta) \le T\}$, if it exists, and zero, otherwise, as well as $N_\zeta(T,\theta) = \max\{n \ge 1 : U_n^-(\theta) \le T\}$, if it exists, and zero, otherwise. Let $\Phi(T,\theta)$ be the set of all indices $m \in \{1,2,\dots,M_\zeta(T,\theta)\}$ such that $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ on the event $\{U_m^+(\theta) < t < V_m^+(\theta)\}$. In a similar vein, let $\Psi(T,\theta)$ be the set of all indices $n \in \{1,2,\dots,N_\zeta(T,\theta)\}$ such that $I(t,\theta) = 0$ on the event $\{U_n^-(\theta) < t < V_n^-(\theta)\}$.

Theorem 5: W.p.1, the IPA derivatives of the lost-sales time average with respect to the production rate parameter are given for all $T > 0$ and $\theta \in \Theta$ as follows:

$$\frac{d}{d\theta^+} L_\zeta(T,\theta) = -\frac{1}{T} \sum_{m \in \Phi(T,\theta)} [\min\{V_m^+(\theta), T\} - V_{m-1}^+(\theta)] \quad (4.6)$$

$$\frac{d}{d\theta^-} L_\zeta(T,\theta) = -\frac{1}{T} \sum_{n \in \Psi(T,\theta)} [\min\{V_n^-(\theta), T\} - V_{n-1}^-(\theta)]. \quad (4.7)$$

Furthermore, letting $\leq_{\rm st}$ denote stochastic ordering, the sided IPA derivatives satisfy

$$\frac{d}{d\theta^+} L_\zeta(T,\theta) \geq_{\rm st} \frac{d}{d\theta^-} L_\zeta(T,\theta), \qquad T > 0; \qquad \theta \in \Theta. \quad (4.8)$$

Stochastic equality holds above, provided $\{I(t,\theta) = S\} \subseteq \{\alpha(t) < \mu(t,\theta)\}$ holds for all $t \in [0,T]$ and $\{I(t,\theta) = 0\} \subseteq \{\alpha(t) > \mu(t,\theta)\}$ holds for all $t \in [0,T]$.

Proof: We only prove (4.6) since the proof of (4.7) is analogous. By standard arguments (see [2], [6], and [8]), the set $\Phi(T,\theta)$ is locally independent of θ in a right neighborhood of θ, and $U_m^+(\theta)$ is locally differentiable with respect to θ in a right neighborhood of θ. Thus

$$\frac{d}{d\theta^+} L_\zeta(T,\theta) = \frac{1}{T} \sum_{m \in \Phi(T,\theta)} \frac{d}{d\theta^+} \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} \zeta(t,\theta)\, dt. \quad (4.9)$$

By part 1) of Assumption 4, for each $m \in \Phi(T,\theta)$

$$\frac{d}{d\theta^+} \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} \zeta(t,\theta)\, dt = -[\alpha(U_m^+(\theta)) - \mu(U_m^+(\theta),\theta)] \times \frac{d}{d\theta^+} U_m^+(\theta) - \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} dt. \quad (4.10)$$

To compute the term associated with $(d/d\theta^+)U_m^+(\theta)$, consider $\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt$ for $m \in \Phi(T,\theta)$. A modicum of algebra shows that $\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt$ is locally independent of θ. Taking the right derivative of this integral yields $-[\alpha(V_{m-1}^+(\theta)) - \mu(V_{m-1}^+(\theta),\theta)](d/d\theta^+)V_{m-1}^+(\theta) + [\alpha(U_m^+(\theta)) - \mu(U_m^+(\theta),\theta)](d/d\theta^+)U_m^+(\theta) - \int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} dt = 0$, because the first term vanishes, resulting in $-[\alpha(U_m^+(\theta)) - \mu(U_m^+(\theta),\theta)](d/d\theta^+)U_m^+(\theta) = -\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} dt$. Finally, substituting the previous equation into (4.8) yields $(d/d\theta^+)\int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} \zeta(t,\theta)\, dt = -\int_{V_{m-1}^+(\theta)}^{\min\{V_m^+(\theta),T\}} dt = -[\min\{V_m^+(\theta),T\} - V_{m-1}^+(\theta)]$. Equation (4.6) now follows by substituting the previous equation into (4.10).

Finally, inequality (4.8) follows directly from (4.6) and (4.7) and the definition of $V_m^+(\theta)$ and $V_n^-(\theta)$ by standard arguments. ■

Theorem 6: Under Assumptions 1 and 4, the IPA derivatives with respect to the production rate parameter, $(d/d\theta^\pm)L_I(T,\theta)$ and $(d/d\theta^\pm)L_\zeta(T,\theta)$, are unbiased for all $T > 0$ and $\theta \in \Theta$.

Proof: Part 3) of Assumption 4 ensures that for all $T > 0$, [5, Assumption C4 of Lemma A2] holds for both $L_I(T,\theta)$ and $L_\zeta(T,\theta)$. Moreover, Theorems 4 and 5 imply that $0 \leq (d/d\theta^\pm)L_I(T,\theta) \leq T/2$ and $|(d/d\theta^\pm)L_\zeta(T,\theta)| \leq 1$. Since the one-sided derivatives exist for every θ and all IPA derivatives are bounded, [5, Assumption C3 of Lemma A2] is in force, which completes the proof.

V. Discussion

Reference [8] and the current note can provide a theoretical basis for new online control algorithms of production-inventory systems, including those where the underlying stochastic processes (i.e., demand and production capacity processes) may be subject to nonstationary probability laws. One direction of future research is the extension of the current results to more general supply chains with multiple products, such as assemble-to-order systems (e.g., such as those implemented by Dell Computer Corporation, Round Rock, TX), where demand patterns fluctuate considerably over time.

References

[1] C. G. Cassandras and S. Lafortune, *Introduction to discrete event systems*. Boston, MA: Kluwer, 1999.

[2] C. G. Cassandras, Y. Wardi, B. Melamed, G. Sun, and C. G. Panayiotou, "Perturbation analysis for on-line control and optimization of stochastic fluid models," *IEEE Trans. Autom. Control*, vol. 47, no. 8, pp. 1234–1248, Aug. 2002.

[3] M. C. Fu and J. Q. Hu, *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*. Boston, MA: Kluwer, 1997.

[4] I. Paschalidis, Y. Liu, C. G. Cassandras, and C. Panayiotou, "Inventory control for supply chains with service level constraints: A synergy between large deviations and perturbation analysis," *Ann. Oper. Res.*, vol. 126, pp. 231–258, 2004.

[5] R. Y. Rubinstein and A. Shapiro, *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*. New York: Wiley, 1993.

[6] Y. Wardi, B. Melamed, C. G. Cassandras, and C. G. Panayiotou, "On-line IPA gradient estimators in stochastic continuous fluid models," *J. Optim. Theory Appl.*, vol. 115, no. 2, pp. 369–405, 2002.

[7] Y. Zhao and B. Melamed, "Make-to-stock systems with backorders: IPA gradients," presented at the Winter Simulation Conf. (WSC), Washington, D.C., Dec. 5–8, 2004.

[8] Y. Zhao and B. Melamed, "IPA derivatives for make-to-stock production-inventory systems with backorders," *Methodology Comput. Appl. Probability*, vol. 8, no. 2, pp. 191–222, 2005.

Minimal Communication for Essential Transitions in a Distributed Discrete-Event System

Feng Lin, Karen Rudie, and Stéphane Lafortune

***Abstract*—In a distributed discrete-event system with decentralized information, agents at the various sites (e.g., controllers or diagnosers) may be required to communicate in order to correctly perform some prescribed tasks. Bandwidth, power, or security constraints motivate the design of**

Manuscript received December 13, 2006; revised March 26, 2007. Recommended by Associate Editor S. Haar. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant 138887, the National Science Foundation (NSF) under Grants CCR-0082784, CCR-0325571, and INT-0213651, the National Aeronautics and Space Administration (NASA) under Grant NAG2-1279, the U.S. Office of Naval Research (ONR) under Grant N00014-03-1-0232, and the National Institute of Health (NIH) under Grant 1 R21 EB001529-01A1.

F. Lin is with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202 USA (e-mail: flin@ece.eng.wayne.edu).

K. Rudie is with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6 Canada (e-mail: karen.rudie@queensu.ca).

S. Lafortune is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA (e-mail: stephane@eecs.umich.edu).