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

Yao Zhao
Benjamin Melamed

Rutgers University
Rutgers Business School – Newark and New Brunswick
Department of MSIS
94 Rockafeller Rd.
Piscataway, NJ 08854
yaozhao@andromeda.rutgers.edu, melamed@rbs.rutgers.edu

Abstract

A stochastic fluid model (SFM) is a queueing model in which workload flow is modeled as fluid flow. More specifically, the traditional discrete arrival, service and departure stochastic processes are replaced by corresponding stochastic fluid-flow rate processes in an SFM. This paper applies the 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. The paper derives formulas for IPA (Infinitesimal Perturbation Analysis) 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 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 very fast to compute, thereby providing the theoretical basis for on-line adaptive control of MTS production-inventory systems.

Keywords and Phrases: Infinitesimal Perturbation Analysis, IPA, IPA derivatives, IPA gradients, Lost Sales, Make-to-Stock, production-inventory systems, stochastic fluid models, SFM.

1 Introduction

Production-inventory systems consist of production facilities that feed replenishment product to inventory facilities, driven by random demand and possibly random production processes, as well as feedback information from inventory to production facilities. An important instance of production-inventory systems is the *Make-to-Stock (MTS)* class, where the inventory facility sends its state information to the production facility as a control signal, which modulates production with the aim of maintaining the inventory level at a prescribed level, called *base-stock level*. Such systems can admit backorders when stock is depleted, or suffer lost sales. This paper treats MTS systems with lost sales (see Section 2), and is a sequel to Zhao and Melamed (2005), which treats MTS systems with backorders.

Economic considerations in supply chains call for effective control of inventory levels and production rates, in order to optimize some prescribed performance metrics. In many real world applications, the underlying demand and production processes may be subject to time varying probability laws. This motivates on-line algorithms that can adaptively control such systems over time with the objective of minimizing inventory on-hand without compromising customer service metrics. To this end, we propose to use IPA (Infinitesimal Perturbation Analysis) derivatives of selected random variables [for comprehensive discussions of IPA derivatives and their applications, refer to Glasserman (1991), Ho and Cao (1991) and Fu (1994a, 1994b)]. 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 on-line control algorithms. Specifically, let $L(\theta)$ be a random variable, parameterized by a generic real-valued parameter θ chosen from a closed and bounded set Θ. The IPA derivative (gradient) of $L(\theta)$ with respect to θ is the random variable $L^{'}(\theta) = \frac{d}{d\theta}L(\theta)$, provided that it exists almost surely. Furthermore, $L^{'}(\theta)$ is said to be *unbiased*, if the expectation and differentiation operators commute, namely, $E[\frac{d}{d\theta}L(\theta)] = \frac{d}{d\theta}E[L(\theta)]$; otherwise, it is said to be *biased*. Sufficient conditions for unbiased IPA derivatives are given in the following lemma

Fact 1 *(see Rubinstein and Shapiro (1993), Lemma A2, p. 70)*
An IPA derivative $L'(\theta)$ is unbiased, if

(a) *For each $\theta \in \Theta$, the IPA derivatives $L'(\theta)$ exist w.p.1 (with probability 1).*

(b) *W.p.1, $L(\theta)$ is Lipschitz continuous in Θ, and the (random) Lipschitz constants have finite first moments.* □

Comprehensive discussions of IPA derivatives and their applications can be found in Glasserman (1991), Ho and Cao (1991) and Fu (1994).

Most papers on production-inventory systems (and MTS systems in particular) postulate specific probability laws that govern the underlying stochastic processes (e.g., Poisson demand arrivals and exponential service times). For simple systems, such as the one-stage MTS variety, closed-form formulas of key performance metrics (e.g., statistics of inventory levels and lost sales or backorders) have been derived as functions of control parameters. For example, Zipkin (1986) and Karmarkar (1987) obtain the optimal control of these systems with respect to batch sizes and re-order points by standard optimization techniques. For more complex MTS systems, such as the multi-stage serial variety, closed-form formulas are not available. A sample path analysis is carried out by Buzacott, Price and Shanthikumar (1991) for a 2-stage production system which is governed by the continuous-time base-stock policy. Diffusion models and deterministic fluid models have been proposed in order to mitigate the analytical and computational complexity of performance evaluation and optimal control. For example, Wein (1992) used a diffusion process to model a multi-product, single-server MTS system, while Veatch (2002) discussed diffusion and fluid-flow models of serial MTS systems. Note, however, that diffusion models require a heavy traffic condition in order to be valid approximations (Wein 1992). In a similar vein, while deterministic fluid-flow models provide valuable insights into the control rules of such systems, deterministic modeling may well result in substantial numerical errors (Veatch 2002).

Simulation has been widely used to study the performances of complex production-inventory systems under uncertainty. Glasserman and Tayur (1995) considered a class of production-inventory systems under the so-called periodic-review, modified base-stock policy, and estimated its performance metrics and IPA derivatives using simulation. While periodic-review policies evaluate system performance at discrete review times, discrete-event simulation, in contrast, can track system performance continuously, but this can be overly time consuming for large-scale systems, due to the large number of events that need to be processed (e.g., arrivals and service completions). All in all, most papers on stochastic production-inventory systems postulate a specific underlying probability law, and focus on off-line control and optimization algorithms.

Recent work has sought to address these shortcomings in the context of fluid-flow queueing systems, and especially, the *stochastic fluid model (SFM)* setting, where transactions carry fluid workload, random discrete arrivals become random arrival rates and random discrete services become random service rates. SFM-like settings represent an alternative (continuous or fluid-flow) queueing paradigm, which differs from the traditional (discrete) queueing paradigm in the way workload is transported in the system[1]. Both paradigms are set in a network of nodes, each of which houses a server and a buffer, where network sources and sinks are viewed as exogenous nodes, and all others as endogenous nodes. Transactions representing parcels of workload arrive at the network from some source, traverse the network according to some itinerary, and then depart the network at some sink. The two queueing paradigms differ, however, in the way workload moves in the system. In the discrete queueing paradigm, transaction workload moves "abruptly" among nodes following a service time, while in the continuous queueing paradigm, transaction workload moves "gradually" (i.e., flows like fluid) for the duration of its service time.

A heuristic modeling rationale underlying SFM systems is the assumption that individual transactions carry miniscule workload as compared to the entire transaction flow, so the effect of individual transactions is infinitesimal and akin to "molecules" in a fluid flow. Furthermore, in many cases, a transaction workload *does* move gradually from one node to another, rather than abruptly (e.g., a conveyor belt carrying bulk material, loading and unloading a truck, train, etc.) In fact, discrete queueing systems can be abstracted as "limiting cases" of continuous queueing systems, where the flow rate is zero when a transaction is still, but at the moment of motion the flow rate becomes momentarily infinite; in other words, the flow rate is akin to a Dirac function. Pursuing this line of reasoning, the "Dirac pulses" of flow rates in a discrete queueing system can be approximated by high flow rates of short duration in a continuous queueing system. Whichever reasoning is used, the modeler can often choose to model a queueing system using either paradigm on equal footing. Finally, we point out that ceteris paribus, SFM systems enjoy an important advantage over their discrete counterparts: IPA derivatives in SFM setting are *unbiased*, while their counterparts in discrete queueing systems are by and large *biased* (Heidelberger et al. 1988). Thus, the local shape of sample paths in the fluid-flow paradigm confers technical advantages on

[1]For simplicity we address only open networks in this discussion.

them. IPA derivatives, derived in SFM setting, can provide important information and insights for their *discrete counterparts*, by applying derivative formulas obtained in SFM setting to queueing systems that have been traditionally viewed as belonging to the discrete queueing paradigm. While preliminary unpublished work by one of the authors suggests that this approach is viable, more work is needed to establish its broad applicability.

Motivated by the considerations above, Wardi et al. (2002) derived IPA derivatives in SFM setting; we henceforth refer to this approach as *IPA-over-SFM*. Wardi et al. (2002) considered two performance metrics: loss volume and buffer-workload time average; each of these metrics was differentiated with respect to buffer size, a parameter of the arrival rate process and a parameter of the service rate process. The paper showed the IPA derivatives to be unbiased, easily computable and nonparametric. Consequently, these derivatives can be computed in simulations, or in the field, and the values can have potential applications to on-line control and stochastic optimization. Paschalidis et al. (2004) treated multi-stage MTS production-inventory systems with backorders in SFM setting. Assuming that inventory at each stage is controlled by a continuous-time base-stock policy, the paper computed the right IPA derivatives of the time averaged inventory level and service level with respect to base-stock levels, and used them to determine the optimal base-stock levels at each stage. Zhao and Melamed (2004) applied the IPA-over-SFM approach to a class of single-product, single-stage MTS systems with backorders, and derived IPA formulas for the time averages of inventory level and backorder level with respect to the base-stock level, as well as a parameter of the production rate process. It should be pointed out that Wardi et al. (2002), Paschalidis et al. (2004) and Zhao and Melamed (2004) assume that systems start with certain initial inventory states, and only consider cases where the left and right IPA derivatives coincide. In contrast, Zhao and Melamed (2005) considered any initial inventory state and derived sided IPA derivative formulas where needed, thereby providing the theoretical basis for IPA-based on-line control of MTS systems with backorders.

The goal of this paper is to derive IPA derivatives for MTS systems with lost-sales, and to show them to be unbiased. The paper makes the following contributions. First, we derive IPA derivative formulas for two metrics, the inventory-level time average and lost-sales time average, with respect to the base-stock level for *all* initial inventory states, including sided derivatives when they differ. We are only aware of one paper [Wardi et al. (2002)] addressing *IPA-over-SFM* queues with finite buffers, which can be used to model MTS systems with lost sales. But unlike the current paper, Wardi et al. (2002) limits the initial condition to an empty buffer. Second, we derive IPA derivative formulas for the aforementioned metrics with respect to a production rate parameter, including sided derivatives when they differ. In contrast, Wardi et al. (2002) only considers cases where the left and right IPA derivatives coincide (in fact, the left and right IPA derivatives may differ). The computation of the general IPA derivatives for any initial inventory state and for cases where the left and right IPA derivatives may differ requires major extensions of the current results in the literature. As will become evident in the sequel, MTS systems with lost-sales are also analytically more challenging than MTS systems with backorders, a fact that results in more elaborate formulas.

The merit of our contribution stems from potential applications of IPA derivatives to on-line control of MTS systems. Clearly, IPA-based on-line control applications mandate the computation of IPA derivatives for all initial inventory states, as well as all sided derivatives when they differ, since a control action can change system parameters at a variety of system states (which are then considered as new initial states). Moreover, it obviously makes little or no sense to wait for the system to return to selected inventory states for which IPA derivatives are known, as this could suspend control actions over extended periods of time. To summarize, for IPA-based applications to be general and efficacious, it is necessary that the requisite IPA derivative formulas satisfy the following requirements:

1. For usability, they should be *comprehensive* in the sense that they are valid for any initial condition of the system. In addition, if a left-derivative does not coincide with its right-derivative counterpart, then both should be exhibited.

2. For statistical accuracy, they should be *unbiased*.

3. For generality, they should be *nonparametric* in the sense that they are solely computable from the sample path observed without making any distributional assumptions on the underlying probability law.

4. To enable on-line applications, they should be *fast* to compute.

To this end, this paper derives all sided IPA derivatives for MTS systems with lost-sales for any initial inventory state. It further shows these IPA derivatives to be unbiased, nonparametric, and easy to compute, which facilitates on-line control applications.

Throughout the paper, we use the following notational conventions and terminology. The indicator function of set A is denoted by 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 is constant in a neighborhood of x.

The rest of the paper is organized as follows. Section 2 presents the production-inventory models under study. Section 3 provides variational bounds for system metrics. Section 4 derives IPA derivative formulas and shows them to be unbiased. Finally, Section 5 concludes the paper.

2 The Make-to-Stock Model With Lost Sales

Consider the traditional single-stage, single-product MTS system, consisting of a production facility and an inventory facility. The two facilities interact: the latter sends back 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 (a single machine, a group of machines or a production line), 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 some base-stock level $S > 0$ (the case $S = 0$ corresponds to a just-in-time system and its treatment is a simple special case.) More specifically, the inventory and production facilities are coupled, and operate in two modes as follows:

Normal operational mode. While the inventory level does not exceed S, the inventory facility places the orders of incoming demands as discrete production jobs in the production facility's buffer according to some operational rule (to be detailed below). The production facility fills these outstanding orders and replenishes the inventory facility back to its base-stock level, but no higher. We refer to this operational mode as *normal operation*, because the system strives to reach an inventory level S, and in so doing, it maintains an inventory level not exceeding S.

Overage operational mode. While the inventory level exceeds S (this could happen, for example, as a result of a control action that lowered S), the production facility buffer is empty,

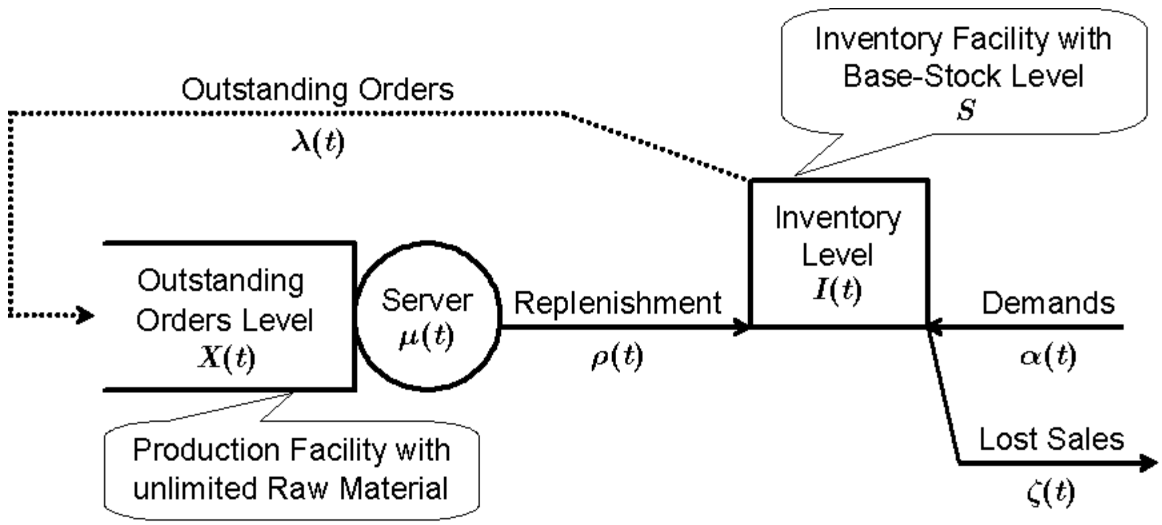

Figure 1: The Make-to-Stock production-inventory system with lost sales

so production is temporarily suspended until the inventory level reaches or crosses S from above, at which point normal operation is resumed. We refer to this operational mode as *overage operation*, because it allows the system to adapt to a lower base-stock level, S, aiming to enter normal operation.

The demand process consists of an interarrival-time process of demands and their random magnitude. Demands arrive at the inventory facility and are satisfied from inventory on hand (if available). Otherwise, when an inventory shortage is encountered, the behavior of the MTS queue is governed by the *lost-sales rule* as follows: The incoming demand is satisfied by the amount of inventory on hand, and any shortage of inventory becomes a lost sale. Thus, the system's overall actions aim to move the inventory level to the base-stock level, S.

2.1 Mapping MTS Systems to SFM Versions

We next proceed to map the traditional discrete MTS system with lost sales into an SFM version, as depicted in Figure 1. Level-related stochastic processes are mapped into fluid versions of their traditional counterparts in a natural way, as follows:

Inventory level. The traditional jump process of the level of inventory on hand at the inventory facility is mapped to a fluid-level counterpart, $\{I(t)\}$, where $I(t)$ is the (fluid) volume of inventory on-hand at time t.

Outstanding orders. The traditional jump process of the level of outstanding orders in the buffer of the production facility is mapped to a fluid-level counterpart, $\{X(t)\}$, where $X(t)$ is the (fluid) volume of outstanding orders at time t.

Traffic-related stochastic processes in Figure 1 are mapped into fluid versions of their traditional counterparts, as follows:

Arrival rate. The traditional arrival process of discrete demands at the inventory facility is mapped to a fluid-flow stochastic process, $\{\alpha(t)\}$, where $\alpha(t)$ is the rate of incoming demands at time t.

Production rate. The traditional service (production) process of discrete product at the production facility is mapped to a fluid-flow stochastic process, $\{\mu(t)\}$, where $\mu(t)$ is the production rate at time t.

Loss rate. The traditional loss process of discrete sales at the inventory facility is mapped to a fluid-flow stochastic process, $\{\zeta(t)\}$, where $\zeta(t)$ is the (fluid) loss rate of sales at time t.

Outstanding order rate. The traditional arrival process of signals for placing discrete outstanding orders at the production facility is mapped to a fluid-flow stochastic process, $\{\lambda(t)\}$, where $\lambda(t)$ is the rate of incoming outstanding orders at time t.

Replenishment rate. The traditional traffic process of discrete replenished product from the production facility to the inventory facility is mapped to a fluid-flow stochastic process, $\{\rho(t)\}$, where $\rho(t)$ is the traffic rate of product at time t.

We now proceed to exhibit the formal definitions of all fluid-model components of the MTS system with lost sales.

During overage operation, the inventory process is governed by the one-side stochastic differential equation

$$\frac{d}{dt^+}I(t) \;=\; -\alpha(t), \tag{2.1}$$

and

$$\zeta(t) \;=\; 0, \tag{2.2}$$

$$X(t) \;=\; 0. \tag{2.3}$$

During normal operation, the model satisfies the conservation relation,

$$X(t) + I(t) \;=\; S. \tag{2.4}$$

The outstanding orders process is governed by the sided stochastic differential equation,

$$\frac{dX(t)}{dt^+} \;=\; \begin{cases} 0, & X(t) = 0 \;\text{ and }\; \alpha(t) \le \mu(t) \\ 0, & X(t) = S \;\text{ and }\; \alpha(t) \ge \mu(t) \\ \alpha(t) - \mu(t), & \text{otherwise} \end{cases} \tag{2.5}$$

The lost-sales rate process is given by

$$\zeta(t) \;=\; [\alpha(t) - \mu(t)]\, 1_{\{I(t)=0, \alpha(t)>\mu(t)\}}, \quad t \ge 0. \tag{2.6}$$

The arrival-rate process of outstanding orders is given by

$$\lambda(t) \;=\; \begin{cases} 0, & \text{if I(t)} > \text{S} \\ \mu(t), & \text{if I(t)} = 0 \text{ and } \alpha\text{(t)} > \mu\text{(t)} \\ \alpha(t), & \text{otherwise} \end{cases} \tag{2.7}$$

and the replenishment-rate process is given by

$$\rho(t) \;=\; \begin{cases} \mu(t), & \text{if X(t)} > 0 \\ \min\{\mu(t), \lambda(t)\}, & \text{if X(t)} = 0. \end{cases} \tag{2.8}$$

2.2 Performance Metrics and Parameters

Let $[0, T]$ be a finite time interval, where T is pre-defined constant, determines the time period during which system performances are evaluated before a control action regarding the inventory policy and/or production rate is taken. We should not confuse T with the review period of a periodic-review inventory policy.

In this paper, we will be interested in the following random variables, to be henceforth referred to as *performance random variables* or simply *metrics*.

Inventory time average. The time average of the inventory on-hand (fluid volume) over the interval $[0, T]$, given by

$$L_I(T) \; = \; \frac{1}{T} \int_0^T I(t)\, dt. \tag{2.9}$$

Lost-sales time average. The time average of fluid rate of lost sales over the interval $[0, T]$, given by

$$L_\zeta(T) \; = \; \frac{1}{T} \int_0^T \zeta(t)\, dt. \tag{2.10}$$

Observe that the metrics $L_I(T)$ and $L_\zeta(T)$ are random variables for each T.

Let $\theta \in \Theta$ denotes 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 explicitly display the dependence of a performance random variable on its parameter of interest. Our objective is to derive formulas for the IPA derivatives $\frac{d}{d\theta} L_I(T, \theta)$, and $\frac{d}{d\theta} L_\zeta(T, \theta)$ in the SFM setting, using sample path analysis, and to show them to be unbiased.

The parameters of interest in this section are listed below:

Base-stock level. The base-stock level of the inventory facility,

$$S(\theta) = \theta, \quad \theta \in \Theta. \tag{2.11}$$

Production rate parameter. A parameter of the production rate process, such that

$$\frac{d}{d\theta}\, \mu(t, \theta) = 1, \quad t \in [0, T], \quad \theta \in \Theta, \tag{2.12}$$

interpreted as a scaling parameter of the production rate.

2.3 Assumptions

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 Wardi et al. (2002), we define two types of sample path events:

Exogenous events. An *exogenous event* occurs either whenever a jump occurs in the sample path of $\{\alpha(t)\}$ or $\{\mu(t)\}$, or when the time horizon T, is reached.

Endogenous events. An *endogenous event* occurs whenever a time interval is inaugurated, in which $X(t) = 0$ or $X(t) = S$.

Throughout this paper, we assume the following mild regularity conditions (cf. Wardi et al. (2002)).

Assumption 1

(a) *The demand rate process, $\{\alpha(t)\}$, and the production rate process, $\{\mu(t)\}$, have right-continuous sample paths that are piecewise-constant w.p.1.*

(b) *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.*

(c) *No multiple events occur simultaneously w.p.1.* □

The following observations follow from Assumption 1.

Observation 1

1. *W.p.1, there exists a finite integer $N \geq 0$ and a sequence of (random) time points $0 = T_0 < T_1 < \cdots < T_N < T_{N+1} = T$, such that the process $\{\alpha(t) - \mu(t)\}$ is constant over each interval (T_n, T_{n+1}), $n = 0, \cdots, N$, and each time point T_n, $1 \leq n \leq N$, is a jump point of the process.*
2. *The process $\{\alpha(t) - \mu(t)\}$ is constant over each time interval (T_n, T_{n+1}), $n = 0, \cdots, N$.*

Proof. See Observation 1 in Zhao and Melamed (2005). □

Finally, we shall be interested in pairs of systems, the original system (indexed by θ) and a perturbed system (indexed by $\theta \pm \Delta\theta$), both starting at the same initial conditions. To simplify the notation in the sequel, we shall also make the following assumption, without any loss of practical generality.

Assumption 2 *The initial inventory level does not depend on θ, namely, $I(0, \theta) = I(0)$ for all $\theta \in \Theta$.* □

3 Variational Bounds

In this section, we derive variational bounds for various parameterized stochastic processes and performance metrics in the MTS model with lost sales. These results will be used in subsequent sections to simplify the derivation of IPA derivatives and to establish their unbiasedness. The variational bounds will be shown to hold with respect to the control parameters of interest at each time point, starting from an arbitrary initial inventory level, $I(0)$.

It follows from Eqs. (2.1), (2.4) and (2.5) that the time derivative of $I(t)$ satisfies

$$\frac{dI(t)}{dt^+} = \begin{cases} -\alpha(t), & \text{if } I(t) > S \\ 0, & \text{if } I(t) = S \text{ and } \alpha(t) \leq \mu(t) \\ 0, & \text{if } I(t) = 0 \text{ and } \alpha(t) \geq \mu(t) \\ \mu(t) - \alpha(t), & \text{otherwise.} \end{cases} \tag{3.1}$$

3.1 Variational Bounds With Respect to the Base-Stock Parameter

In this section, the IPA parameter of interest is $S(\theta) = \theta$ for $\theta \in \Theta$. Let $\{I(t, \theta)\}$ be the inventory level process in an MTS system with lost-sales, where $\theta \in \Theta$ and $I(0, \theta) = I(0)$. Then, Eq. (3.1) induces a (random) partition of the interval $[0, T]$, given by

$$\mathcal{R}(\theta) \;=\; \{\mathcal{R}_1(\theta), \mathcal{R}_2(\theta), \mathcal{R}_3(\theta), \mathcal{R}_4(\theta)\}, \tag{3.2}$$

where each region, $\mathcal{R}_k(\theta)$, $1 \le k \le 4$, is defined as follows,

$$\begin{aligned}
\mathcal{R}_1(\theta) \;&=\; \{t \in [0,T] :\; I(t,\theta) = 0 \text{ and } \alpha(t) \ge \mu(t)\}, \\
\mathcal{R}_2(\theta) \;&=\; \{t \in [0,T] :\; [I(t,\theta) = 0 \text{ and } \alpha(t) < \mu(t)] \\
&\qquad \text{or } [I(t,\theta) = S(\theta) \text{ and } \alpha(t) > \mu(t)] \text{ or } 0 < I(t,\theta) < S(\theta)\}, \\
\mathcal{R}_3(\theta) \;&=\; \{t \in [0,T] :\; I(t,\theta) = S(\theta) \text{ and } \alpha(t) \le \mu(t)\}, \\
\mathcal{R}_4(\theta) \;&=\; \{t \in [0,T] :\; I(t,\theta) > S(\theta)\}.
\end{aligned}$$

We first prove the variational bounds for the inventory level process, $\{I(t, \theta)\}$.

Proposition 1 *For an MTS system with the lost-sales rule, let $\theta_1, \theta_2 \in \Theta$. Then,*

$$0 \;\le\; |I(t,\theta_1) - I(t,\theta_2)| \;\le\; |\theta_1 - \theta_2|, \quad t \in [0,T]. \tag{3.3}$$

Proof. Recall that by Assumption 2,

$$I(0,\theta_1) \;=\; I(0,\theta_2) \;=\; I(0). \tag{3.4}$$

Clearly, Eqs. (3.4) and (3.1) imply the result trivially for $\theta_1 = \theta_2$. It remains to show the result for the case $\theta_1 \neq \theta_2$. Without loss of generality, we assume that $\theta_1 < \theta_2$, and show that

$$0 \;\le\; I(t,\theta_2) - I(t,\theta_1) \;\le\; \theta_2 - \theta_1, \quad t \in [0,T]. \tag{3.5}$$

To this end, we first prove the lefthand side of inequality (3.5) by showing that whenever $I(t, \theta_1) = I(t, \theta_2)$ for any $t \in [0, T]$, one has $\dfrac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] \ge 0$.

An examination of Eq. (3.1) reveals that the equality $I(t, \theta_1) = I(t, \theta_2)$ can take place only in the following cases.

Case 1: $t \in \mathcal{R}_k(\theta_1) \bigcap \mathcal{R}_k(\theta_2)$ for some $1 \le k \le 4$. In this case, we immediately have $\dfrac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = 0$.

Case 2: $t \in \mathcal{R}_3(\theta_1) \bigcap \mathcal{R}_2(\theta_2)$. In this case, $I(t,\theta_1) = I(t,\theta_2) = S(\theta_1) = \theta_1$ and $\dfrac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = \mu(t) - \alpha(t) \ge 0$, where the inequality follows from the definition of $\mathcal{R}_3(\theta_1)$.

Case 3: $t \in \mathcal{R}_4(\theta_1) \bigcap \mathcal{R}_2(\theta_2)$. In this case, $\dfrac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = \mu(t) \ge 0$ by the definition of $\mathcal{R}_2(\theta_2)$ and $\mathcal{R}_4(\theta_1)$.

Case 4: $t \in \mathcal{R}_4(\theta_1) \bigcap \mathcal{R}_3(\theta_2)$. In this case, $\dfrac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = \alpha(t) \ge 0$ by the definition of $\mathcal{R}_3(\theta_2)$ and $\mathcal{R}_4(\theta_1)$.

The lefthand side of inequality (3.5) follows from Eq. (3.4) and by the continuity of the realizations of the inventory level process.

To prove the righthand side of inequality (3.5), we examine the behavior of $\{I(t,\theta_2) - I(t,\theta_1)\}$ in the four regions of the partition (3.2). Informally, the proof characterizes $\{I(t,\theta_2) - I(t,\theta_1)\}$ for all pairs of regions in the partitions associated with each θ, such that $I(t,\theta_1)$ is in one region and $I(t,\theta_2)$ is in the other. More formally, the characterization covers t in all intersections of the form $\mathcal{R}_i(\theta_1) \bigcap \mathcal{R}_j(\theta_2)$, $1 \le i,j \le 4$. Note that the intersections partition the interval $[0,T]$ and their number is finite w.p.1 by Part (b) of Assumption 1. The proof proceeds in two steps. In the first step, we consider the extremal open set (a,b) of any such intersection. We then show that if

$$I(a,\theta_2) - I(a,\theta_1) \le \theta_2 - \theta_1, \tag{3.6}$$

then

$$I(t,\theta_2) - I(t,\theta_1) \le \theta_2 - \theta_1, \quad a < t < b. \tag{3.7}$$

By continuity of the inventory level process, the inequality (3.7) will then extend to the interval $[a,b]$. In the second step, we order the intervals (a_k, b_k) contiguously, and prove the inequality (3.5) throughout $[0,T]$, by a straightforward induction on k, where the induction basis holds by Eq.(3.4), and the induction step is immediate from the contiguity of the ordered intersections.

The following observation reduces substantially the number of region-pair cases (intersections) to be checked. There is no need to check for pairs of regions with the same subscript, $i = j$, since in their intersection $\frac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = 0$ trivially, which implies that $I(t,\theta_2) - I(t,\theta_1)$ is constant in the intersection. It remains to check the following list of cases.

Case 1: $t \in \mathcal{R}_1(\theta_1) \bigcap \mathcal{R}_2(\theta_2)$. In this case, $\frac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = \mu(t) - \alpha(t) \le 0$, where the inequality follows from the definition of $\mathcal{R}_1(\theta_1)$. In view of (3.6), inequality (3.7) immediately follows.

Case 2: $t \in \mathcal{R}_1(\theta_1) \bigcap \mathcal{R}_3(\theta_2)$. In this case, $\frac{d}{dt^+}[I(t,\theta_2) - I(t,\theta_1)] = 0$.

Case 3: $t \in \mathcal{R}_3(\theta_1) \bigcap \mathcal{R}_2(\theta_2)$. In this case, $I(t,\theta_2) \le \theta_2$ and $I(t,\theta_1) = \theta_1$, which implies that $I(t,\theta_2) - I(t,\theta_1) \le \theta_2 - \theta_1$.

Case 4: $t \in \mathcal{R}_4(\theta_1) \bigcap \mathcal{R}_2(\theta_2)$. In this case, $I(t,\theta_2) \le \theta_2$ and $\theta_1 < I(t,\theta_1)$, which implies that $I(t,\theta_2) - I(t,\theta_1) \le \theta_2 - \theta_1$.

Case 5: $t \in \mathcal{R}_4(\theta_1) \bigcap \mathcal{R}_3(\theta_2)$. In this case, $I(t,\theta_2) = \theta_2$ and $\theta_1 < I(t,\theta_1)$, which implies that $I(t,\theta_2) - I(t,\theta_1) \le \theta_2 - \theta_1$.

The proof is complete. □

We next derive variational bounds for the time average of lost sales. To this end, we define $K(T,\theta)$ to be the number of extremal subintervals of $[0,T]$ in which $I(t,\theta) = 0$.

Proposition 2 *For an MTS system with the lost-sales rule, let $\theta_1, \theta_2 \in \Theta$. Then,*

$$\int_0^T |\zeta(t,\theta_1) - \zeta(t,\theta_2)|\, dt \le \max\{K(T,\theta_1), K(T,\theta_2)\}|\theta_1 - \theta_2|. \tag{3.8}$$

Proof. The case of $\theta_1 = \theta_2$ is trivial, so it remains to prove the case $\theta_1 \neq \theta_2$, and assume $\theta_1 < \theta_2$ without loss of generality. We prove the following inequality,

$$0 \;\leq\; \int_0^T [\zeta(t,\theta_1) - \zeta(t,\theta_2)]\, dt \;\leq\; K(T,\theta_1)[\theta_2 - \theta_1], \tag{3.9}$$

which is stronger than the requisite result.

Since $\theta_1 < \theta_2$, the proof of Proposition 1 implies that $I(t,\theta_1) \leq I(t,\theta_2)$ for all $t \in [0,T]$. Consequently, $\zeta(t,\theta_2) \leq \zeta(t,\theta_1)$ for all $t \in [0,T]$, which establishes the lefthand side of (3.9).

We next prove the righthand side of inequality (3.9). In view of the inequality $I(t,\theta_1) \leq I(t,\theta_2)$ for all $t \in [0,T]$, it suffices to show that for any extremal subinterval $[U,V]$ of $[0,T]$ in which $I(t,\theta_1) = 0$, one has

$$\int_U^t [\zeta(\tau,\theta_1) - \zeta(\tau,\theta_2)]\, d\tau \;\leq\; I(U,\theta_2) - I(U,\theta_1), \quad t \in [U,V]. \tag{3.10}$$

Define $W \in [U,V]$ to be the first time point at which $I(t,\theta_2) = 0$, if it exists; otherwise, define $W = V$. Since $I(t,\theta_1) = I(t,\theta_2) = 0$ for $t \in [W,V)$, it follows from Eq. (2.6) that $\int_W^V [\zeta(\tau,\theta_1) - \zeta(\tau,\theta_2)]\, d\tau = 0$, so it remains to consider the interval $[U,W)$. But for any $t \in [U,W)$, $\zeta(t,\theta_2) = 0$ and $\zeta(t,\theta_1) = \alpha(t) - \mu(t)$. Hence, for every $t \in [U,W)$,

$$\int_U^t [\zeta(\tau,\theta_1) - \zeta(\tau,\theta_2)]\, d\tau \;=\; \int_U^t [\alpha(\tau) - \mu(\tau)]\, d\tau.$$

We conclude that for every $t \in [U,W)$,

$$\int_U^t [\alpha(\tau) - \mu(\tau)]\, d\tau \;=\; I(U,\theta_2) - I(t,\theta_2) \;\leq\; I(U,\theta_2) - I(U,\theta_1) \;\leq\; \theta_2 - \theta_1,$$

where the equality is due to the dynamics of Eq. (3.1), the first inequality follows from the relation $I(t,\theta_2) \geq I(U,\theta_1) = 0$, and the second inequality follows from (3.5). The result now follows by applying inequality (3.10) to all extremal subintervals of the form $[U,V]$ and summing the corresponding inequalities. □

3.2 Variational Bounds With Respect to a Production Rate Parameter

In this section, the IPA parameter of interest is a parameter, θ, of the production rate process, $\{\mu(t,\theta)\}$, satisfying Eq. (2.12). Our results build upon prior results in Wardi and Melamed (2001), which assume a special initial condition for the workload.

Observation 2 *For an MTS system with the lost-sales rule, the stochastic differential equations governing the outstanding orders process $\{X(t)\}$ in normal operation, the loss-rate process, $\{\zeta(t)\}$, and the replenishment rate process, $\{\rho(t)\}$, are identical to those governing the buffer workload, overflow and outflow processes, respectively, in the SFM queuing system studied in Wardi et al. (2001).*

Proof. Follows from the fact that we can identify the demand arrival rate process, production rate process, and base-stock level parameter, respectively, with the inflow rate process, service rate process and buffer capacity parameter in Wardi and Melamed (2001). □

For notational convenience, we define an auxiliary process, called the *extended outstanding orders process*, $\{Y(t,\theta)\}$, by

$$Y(t,\theta) \;=\; \begin{cases} S - I(t,\theta), & \text{if } I(t,\theta) > S \quad \text{(overage operation)} \\ X(t,\theta), & \text{if } I(t,\theta) \le S \quad \text{(normal operation)} \end{cases} \tag{3.11}$$

Observe that $Y(t)$ is negative during overage operation and non-negative during normal operation. Furthermore, Eq. (2.4) implies the conservation relation,

$$I(t,\theta) + Y(t,\theta) \;=\; S, \quad t \ge 0, \tag{3.12}$$

valid for each operational mode (overage and normal).

Proposition 3 *For an MTS system with the lost-sales rule, let* $\theta_1, \theta_2 \in \Theta$. *Then,*

$$\max\{|Y(t,\theta_1) - Y(t,\theta_2)| : \; t \in [0,T]\} \;\le\; T\,|\theta_1 - \theta_2|$$

and

$$\int_0^T |\zeta(t,\theta_1) - \zeta(t,\theta_2)|\,dt \;\le\; 2\,T\,|\theta_1 - \theta_2|.$$

Proof. By Assumption 2 and Eq. (3.12), $Y(0,\theta_1) = Y(0,\theta_2)$. In view of the fact that $\{Y(t,\theta_1)\}$ and $\{Y(t,\theta_2)\}$ coincide during overage operation, it suffices to assume that the system starts in normal operation, namely, $Y(0,\theta_1) = Y(0,\theta_2) \ge 0$. The results follow immediately from Proposition 3.2 of Wardi and Melamed (2001), since the proof there is readily seen to hold for any initial state in normal operation. □

Corollary 1 *For an MTS system with the lost-sales rule, let* $\theta_1, \theta_2 \in \Theta$. *Then,*

$$|I(t,\theta_1) - I(t,\theta_2)| \;\le\; T\,|\theta_1 - \theta_2|, \quad t \in [0,T].$$

Proof. Eq. (3.12) and Proposition 3 imply that

$$|I(t,\theta_1) - I(t,\theta_2)| \;=\; |[S - Y(t,\theta_1)] - [S - Y(t,\theta_2)]| \;\le\; T\,|\theta_1 - \theta_2|.$$

□

4 IPA Derivatives

We are now in a position to derive IPA derivatives for various parameterized stochastic processes and performance metrics in the MTS model subject to lost sales rule. We mention in passing that such systems are analytically more challenging than MTS systems with backorders, because the inventory state of the former has an extra boundary. More specifically, while the inventory state of both systems is bounded from above by S, that of MTS systems with lost sales is also bounded from below by 0.

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, we let $Z_j(\theta) \in (Q_j(\theta), R_j(\theta))$ be the first time point in this interval at which $I(t,\theta) = 0$ if such a point exists; otherwise, let $Z_j(\theta) = R_j(\theta)$.

Observation 3

$$Q_1(\theta) < R_1(\theta) < Q_2(\theta) < R_2(\theta) < \ldots < Q_{J(\theta)}(\theta) < R_{J(\theta)}(\theta). \tag{4.1}$$

Proof. See Observation 3 in [19]. □

4.1 IPA Derivatives with Respect to the Base-Stock Level

This section treats 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, and exhibits their formulas for any initial inventory state. We first prove a number of useful lemmas that simplify the proofs of the main results later in this section. We then proceed to obtain the IPA derivatives for $L_I(T, \theta)$ by first obtaining those for the inventory process, $\{I(t, \theta)\}$, following which we obtain the IPA derivatives for $L_\zeta(T, \theta)$. Finally, we establish the unbiasedness of all the IPA derivatives above.

Assumption 3

(a) *$S(\theta) = \theta$, where $\theta \in \Theta$.*

(b) *The processes $\{\alpha(t)\}$ and $\{\mu(t)\}$ are independent of the parameter θ.* □

The following lemma provides basic properties for the inventory level process.

Lemma 1

(a) *For every $j \geq 1$,*

$$\frac{d}{d\theta} I(t, \theta) = 1, \quad t \in [R_j(\theta), Z_{j+1}(\theta)).$$

(b) *For every $j \geq 1$,*

$$\frac{d}{d\theta} I(t, \theta) = 0, \quad t \in (Z_j(\theta), R_j(\theta)).$$

Proof. To prove part (a), note that each $R_j(\theta)$, $j \geq 1$ is locally differentiable with respect to θ by part (c) of Assumption 1. By Observation 3, $R_j(\theta) < Q_{j+1}(\theta)$, where $Q_{j+1}(\theta)$ is a jump point of $\{\alpha(t) - \mu(t)\}$, and therefore $Q_{j+1}(\theta)$ is locally independent of θ. Consequently, $I(t, \theta) = S(\theta)$ for $t \in (R_j(\theta), Q_{j+1}(\theta)]$ and

$$I(t, \theta) \;=\; S(\theta) + \int_{Q_{j+1}(\theta)}^{t} [\mu(t) - \alpha(t)] \, dt, \quad t \in (Q_{j+1}(\theta), Z_{j+1}(\theta)).$$

Part (a) now follows by differentiating $I(t, \theta)$ with respect to θ for $t \in (R_j(\theta), Z_{j+1}(\theta)]$.

To prove part (b), consider first the extremal time interval $[Z_j(\theta), \tilde{Z}_j(\theta)]$ in which $I(t, \theta) = 0$. By part (c) of Assumption 1, $Z_j(\theta)$ is locally differentiable with respect to θ, $\tilde{Z}_j(\theta)$ is a jump point of $\{\alpha(t) - \mu(t)\}$ and $Z_j(\theta) < \tilde{Z}_j(\theta)$. Therefore, $\tilde{Z}_j(\theta)$ is locally independent of θ and $I(t, \theta) = 0$ is locally independent of θ for $t \in (Z_j(\theta), \tilde{Z}_j(\theta)]$. Finally, note that by Eq. (3.1), $\dfrac{d}{dt^+} I(t, \theta)$ is independent of θ for $t \in (\tilde{Z}_j(\theta), R_j(\theta))$. The proof is now complete. □

The following lemma provides basic properties for the time average of lost sales.

Lemma 2 *Let $0 \le u \le T$ be a time point, independent of θ.*

(a) *For every $j = 1, \ldots, J(\theta)$, on the event $\{Z_j(\theta) = R_j(\theta)\}$,*

$$\frac{d}{d\theta} \int_{Q_j(\theta)}^{u} \zeta(t,\theta)\, dt \;=\; 0, \qquad \text{for } Q_j(\theta) < u \le R_j(\theta) \tag{4.2}$$

(b) *For every $j = 2, \ldots, J(\theta)$, on the event $\{Z_j(\theta) < R_j(\theta)\}$,*

$$\frac{d}{d\theta} \int_{Q_j(\theta)}^{u} \zeta(t,\theta)\, dt \;=\; \begin{cases} 0, & \text{for } Q_j(\theta) < u < Z_j(\theta) \\ -1, & \text{for } Z_j(\theta) < u \le R_j(\theta) \end{cases} \tag{4.3}$$

(c) *For every $j = 2, \ldots, J(\theta)$ and for $u = Z_j(\theta)$, on the event $\{Z_j(\theta) < R_j(\theta)\}$,*

$$\frac{d}{d\theta^+} \int_{Q_j(\theta)}^{u} \zeta(t,\theta)\, dt \;=\; 0 \tag{4.4}$$

$$\frac{d}{d\theta^-} \int_{Q_j(\theta)}^{u} \zeta(t,\theta)\, dt \;=\; -1 \tag{4.5}$$

Proof. To prove part (a), we show that the integral in Eq. (4.2) is locally independent of θ. To see that, observe that

$$\{Z_j(\theta) = R_j(\theta)\} = \{I(t,\theta) > 0, t \in [(Q_j(\theta), R_j(\theta)]\} \subset \{\zeta(t,\theta) = 0, t \in [(Q_j(\theta), R_j(\theta)]\}$$

and each $I(t,\theta)$ is continuous in θ by Proposition 1. The result now follows for this part since the integral clearly vanishes.

To prove part (b) for $Q_j(\theta) < u < Z_j(\theta)$, the proof of part (a) is applicable. To prove part (b) for $Z_j(\theta) < u \le R_j(\theta)$, note that Eq. (2.6) implies

$$\int_{Q_j(\theta)}^{u} \zeta(t,\theta)\, dt \;=\; \int_{Z_j(\theta)}^{u} [\alpha(t) - \mu(t)]\, 1_{\{I(t,\theta)=0, \alpha(t)>\mu(t)\}}\, dt \qquad \text{on } \{Z_j(\theta) < R_j(\theta)\}. \tag{4.6}$$

By part (c) of Assumption 1, $Z_j(\theta)$ is locally differentiable with respect to θ. Furthermore, from part (b) of Lemma 1, we conclude that $\{I(t,\theta)\}$ is locally independent of θ for $t \in (Z_j(\theta), R_j(\theta)]$. It follows from Leibnitz's rule that differentiating Eq. (4.6) with respect to θ yields

$$\frac{d}{d\theta} \int_{Z_j(\theta)}^{u} \zeta(t,\theta)\, dt \;=\; -[\alpha(Z_j(\theta)) - \mu(Z_j(\theta))]\, \frac{d}{d\theta} Z_j(\theta). \tag{4.7}$$

To compute the right-hand side of Eq. (4.7), note first that the proof of part (a) of Lemma 1, implies that $Q_j(\theta)$ is locally independent of θ for $j \ge 2$. Since

$$\int_{Q_j(\theta)}^{Z_j(\theta)} [\alpha(t) - \mu(t)]\, dt \;=\; I(Q_j(\theta)) - I(Z_j(\theta)) \;=\; S(\theta),$$

differentiating this equation with respect to θ yields

$$[\alpha(Z_j(\theta)) - \mu(Z_j(\theta))]\, \frac{d}{d\theta} Z_j(\theta) \;=\; 1.$$

The result now follows by substituting the above into Eq. (4.7).

Part (c) follows from Eq. (4.3), by noting that $u = Z_j(\theta)$ satisfies $Z_j(\theta - \Delta\theta) < u < Z_j(\theta - \Delta\theta)$.
□

Remark. The event $\{Z_j(\theta) = u\}$ often has probability 0, so the brief proof of part (c) above is included just for completeness.

Lemma 3

(a) *For every* $j = 1, \ldots, J(\theta)$,

$$\frac{d}{d\theta} \int_{Q_j(\theta)}^{R_j(\theta)} \zeta(t,\theta)\, dt \;=\; 0, \quad \textit{on} \;\; \{Z_j(\theta) = R_j(\theta)\}. \tag{4.8}$$

(b) *For every* $j = 2, \ldots, J(\theta)$,

$$\frac{d}{d\theta} \int_{Q_j(\theta)}^{R_j(\theta)} \zeta(t,\theta)\, dt \;=\; -1 \quad \textit{on} \;\; \{Z_j(\theta) < R_j(\theta)\}. \tag{4.9}$$

Proof. Part (a) follows by an argument similar to that in the proof of part (a) in Lemma 2.

To prove part (b), note that by part (c) of Assumption 1, both $Z_j(\theta)$ and $R_j(\theta)$ is locally differentiable with respect to θ. Combining these facts with $\zeta(R_j(\theta), \theta) = 0$, it follows from Leibnitz's rule that

$$\int_{Q_j(\theta)}^{R_j(\theta)} \zeta(t,\theta)\, dt \;=\; \int_{Z_j(\theta)}^{R_j(\theta)} [\alpha(t) - \mu(t)]\, 1_{\{I(t,\theta)=0, \alpha(t)>\mu(t)\}}\, dt \quad \text{on } \{Z_j(\theta) < R_j(\theta)\}. \tag{4.10}$$

The rest of the proof is similar to that of part (b) in Lemma 2. □

We first derive the IPA derivatives for the inventory process $\{I(t,\theta)\}$. In the next two lemmas we make use of the hitting time, $T_S(\theta)$, defined by

$$T_S(\theta) \;=\; \begin{cases} \min\{t \in [0,\infty] : \; I(t,\theta) = S(\theta)\}, & \text{if the minimum exists} \\ \infty, & \text{otherwise} \end{cases} \tag{4.11}$$

Lemma 4 *Consider an MTS system with the lost sales rule on the event* $\{I(0) < S(\theta)\}$ *(that is, the system starts in normal operation with partial inventory). Then, for any* $t \geq 0$ *and* $\theta \in \Theta$,
(a) *On the event* $A(\theta) = \{I(0) < S(\theta)\} \bigcap \{t < T_S(\theta)\}$,

$$\frac{d}{d\theta} I(t,\theta) \;=\; 0.$$

(b) *On the events* $B_j(\theta) = \{I(0) < S(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$,

$$\frac{d}{d\theta} I(t,\theta) \;=\; 1.$$

(c) *On the events* $C_j(\theta) = \{I(0) < S(\theta)\} \bigcap \{Z_j(\theta) < t < R_j(\theta)\}$, $j \geq 2$,

$$\frac{d}{d\theta} I(t,\theta) \;=\; 0.$$

Proof. By Observation 3,

$$0 \;=\; Q_1(\theta) \;<\; T_S(\theta) \;=\; R_1(\theta) \;<\; Q_2(\theta) \quad \text{on } \; \{I(0) < S(\theta)\}, \tag{4.12}$$

and this holds for all cases of this lemma.

To prove part (a), note that by Eq. (3.1) and Assumption 2, the time derivative $\frac{d}{dt^+} I(t,\theta)$ is locally independent of θ on $A(\theta)$. Consequently,

$$I(t,\theta) \;=\; I(0) + \int_0^t \frac{d}{d\tau^+} I(\tau,\theta)\, d\tau \qquad \text{on} \;\; \mathrm{A}(\theta)$$

is independent of θ on $A(\theta)$, which proves part (a).

Finally, part (b) follows immediately from part (a) of Lemma 1, while part (c) follows immediately from part (b) of Lemma 1. □

Lemma 5 *Consider an MTS system with the lost-sales rule on the event $\{I(0) > S(\theta)\}$ (that is, the system starts in overage operation). Then, for any $t \geq 0$ and $\theta \in \Theta$,*
(a) *On the event $A(\theta) = \{I(0) > S(\theta)\} \bigcap \{t < T_S(\theta)\}$,*

$$\frac{d}{d\theta} I(t,\theta) \;=\; 0.$$

(b) *On either of the events $B_1(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) < Q_1(\theta)\} \bigcap \{T_S(\theta) < t < Z_1(\theta)\}$ or $B_{2,j}(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) < Q_1(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$ or $B_{3,j}(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) = Q_1(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$,*

$$\frac{d}{d\theta} I(t,\theta) \;=\; 1.$$

(c) *On the events $C_j(\theta) = \{I(0) > S(\theta)\} \bigcap \{Z_j < t < R_j(\theta)\}$, $j \geq 1$,*

$$\frac{d}{d\theta} I(t,\theta) \;=\; 0.$$

(d) *On the event $D(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) = Q_1(\theta)\} \bigcap \{T_S(\theta) < t < Z_1(\theta)\}$,*

$$\frac{d}{d\theta} I(t,\theta) \;=\; \frac{\mu(Q_1(\theta))}{\alpha(Q_1(\theta))}.$$

Proof. To prove part (a), note that by Eq. (3.1) that $\frac{d}{dt^+} I(t,\theta) = -\alpha(t)$ on $A(\theta)$. Therefore,

$$I(t,\theta) \;=\; I(0) - \int_0^t \alpha(\tau)\, d\tau$$

is independent of θ on $A(\theta)$, whence the result follows.

To prove part (b) on the event $B_1(\theta)$, note that by Observation 3,

$$T_S(\theta) \;<\; Q_1(\theta) \;<\; R_1(\theta) \qquad \text{on} \;\; B_1(\theta). \tag{4.13}$$

Clearly, $T_S(\theta)$ is locally differentiable with respect to θ. In view of Eq. (4.13), $Q_1(\theta)$ corresponds to a jump in $\{\alpha(t) - \mu(t)\}$, and part (b) on the event $B_1(\theta)$ follows by a proof similar to that of part (a) in Lemma 1. Part (b) on the events $B_{2,j}$ and $B_{3,j}(\theta)$, $j \geq 1$ follows immediately from part (a) of Lemma 1.

Part (c) follows immediately from part (b) of Lemma 1.

To prove part (d), note that by Observation 3,

$$T_S(\theta) \;=\; Q_1(\theta) \;<\; R_1(\theta) \;< Q_2(\theta) \qquad \text{on} \;\; D(\theta). \tag{4.14}$$

Furthermore,

$$\int_0^{Q_1(\theta)} \alpha(\tau)\, d\tau \;=\; I(0) - S(\theta) \;=\; I(0) - \theta, \qquad \text{on} \;\; D(\theta).$$

Differentiating the above equation with respect to θ yields,

$$\frac{d}{d\theta} Q_1(\theta) \;=\; \frac{-1}{\alpha(Q_1(\theta))}, \qquad \text{on} \;\; D(\theta). \tag{4.15}$$

next, write $I(t,\theta) = S(\theta) + \int_{Q_1(\theta)}^{t} [\mu(\tau) - \alpha(\tau)]\, d\tau$ on the event $D(\theta)$, and then differentiate it with respect to θ, yielding the desired result

$$\frac{d}{d\theta} I(t,\theta) \;=\; 1 - [\mu(Q_1(\theta)) - \alpha(Q_1(\theta))] \,\frac{d}{d\theta} Q_1(\theta) \;=\; \frac{\mu(Q_1(\theta))}{\alpha(Q_1(\theta))} \qquad \text{on} \;\; D(\theta), \tag{4.16}$$

where the second equality is obtained by substituting Eq. (4.15), and noting the inequalities $\alpha(Q_1(\theta)) > \mu(Q_1(\theta)) \geq 0$ on the event $\{I(0) > S(\theta)\} \bigcap \{Q_1(\theta) = T_S(\theta)\}$. □

On the event $\{I(0) = S(\theta)\}$, the situation is more complex, because the left and right derivatives of $I(t,\theta)$ with respect to θ do not coincide and must be computed separately. This event cannot be excluded because it may happen in applications where inventory levels are discrete.

We first derive the right-derivatives, $\frac{d}{d\theta^+} I(t,\theta)$ by borrowing from Lemma 4 and making use of the hitting time $T_\mu(\theta)$, given by

$$T_\mu(\theta) \;=\; \begin{cases} \min\{t \in [0, Q_1(\theta)) : \;\mu(t) > \alpha(t)\}, & \text{if the minimum exists on the event } \{Q_1(\theta) > 0\} \\ R_1(\theta), & \text{if } R_1(\theta) \text{ exists on the event} \\ & \{Q_1(\theta) = 0\} \bigcup [\{Q_1(\theta) > 0\} \bigcap \{\alpha(t) = \mu(t),\; t \in [0, Q_1(\theta))\}] \\ \infty, & \text{otherwise} \end{cases} \tag{4.17}$$

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, $\{I(t,\theta + \Delta\theta)\}$ for any $\Delta\theta > 0$.

Lemma 6 *Consider an MTS system with the lost-sales rule on the event $\{I(0) = S(\theta)\}$ (that is, the system starts in normal operation with full inventory). Then, for any $t \geq 0$ and $\theta \in \Theta$,*
(a) *On the event $A(\theta) = \{I(0) = S(\theta)\} \bigcap \{t < T_\mu(\theta)\}$,*

$$\frac{d}{d\theta^+} I(t,\theta) \;=\; 0.$$

(b) *On the events $B_1(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) < R_1(\theta)\} \bigcap \{T_\mu(\theta) < t < Z_1(\theta)\}$ or $B_{2,j}(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) < R_1(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$ or $B_{3,j}(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) = R_1(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$,*

$$\frac{d}{d\theta^+} I(t,\theta) \;=\; 1.$$

(c) *On the events $C_{1,j}(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) = R_1(\theta)\} \bigcap \{Z_j(\theta) < t < R_j(\theta)\}$, $j \geq 2$ or $C_{2,j}(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\mu(\theta) < R_1(\theta)\} \bigcap \{Z_j(\theta) < t < R_j(\theta)\}$, $j \geq 1$,*

$$\frac{d}{d\theta^+} I(t,\theta) \;=\; 0.$$

Proof. Consider a perturbed system with $S(\theta+\Delta\theta) = \theta + \Delta\theta$, where $\Delta\theta > 0$. Since $I(0) = S(\theta) < S(\theta+\Delta\theta)$, it follows that the perturbed system starts in normal operation. Denote $\Delta S = S(\theta+\Delta\theta) - S(\theta)$. By Observation 3,

$$0 \;=\; Q_1(\theta+\Delta\theta) \;\leq\; T_\mu(\theta) \;<\; R_1(\theta+\Delta\theta) \;<\; Q_2(\theta+\Delta\theta) \quad \text{on } \{I(0) = S(\theta)\}, \tag{4.18}$$

and this holds for all cases of this lemma.

To prove part (a), note first that the event $\{T_\mu(\theta) = 0\}$ can be precluded, since it implies $A(\theta) = \emptyset$. Otherwise, the definition of $T_\mu(\theta)$ and Eq. (3.1) imply that $\frac{d}{dt^+} I(t,\theta) = \frac{d}{dt^+} I(t, \theta+\Delta\theta)$ on $A(\theta)$. By Assumption 2, we conclude that $I(t,\theta+\Delta\theta) = I(t,\theta)$ are independent of θ on $A(\theta)$, and therefore, part (a) follows immediately.

To prove part (b), observe that part (b) of Assumption 1 implies that there exists $\epsilon > 0$, such that for any $\Delta\theta \leq \epsilon$,

$$R_1(\theta+\Delta\theta) \;=\; T_\mu(\theta) + \frac{\Delta S}{\mu(T_\mu(\theta)) - \alpha(T_\mu(\theta))} \qquad \text{on } \{I(0) = S(\theta)\}, \tag{4.19}$$

where the inequality $\mu(T_\mu(\theta)) - \alpha(T_\mu(\theta)) > 0$ follows from the definition of $T_\mu(\theta)$. We now proceed with the proof on the event $B_1(\theta)$, by considering two cases.

Case 1: On the event $B_1(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$, it follows from the definition of $T_\mu(\theta)$ and part (c) of Assumption 1 that $T_\mu(\theta) < Q_1(\theta)$ and $Q_1(\theta)$ is a jump point of $\{\alpha(\theta) - \mu(\theta)\}$. Therefore, $Q_2(\theta+\Delta\theta) = Q_1(\theta)$ for sufficiently small $\Delta\theta$. Note that in this case, $I(t,\theta+\Delta\theta) = I(t,\theta)$ on $\{0 < t < T_\mu(\theta)\}$, and then it increases to $S(\theta+\Delta\theta)$ and stays there until $Q_2(\theta+\Delta\theta)$. Furthermore, over the interval $[R_1(\theta+\Delta\theta), Z_1(\theta)]$ and for sufficiently small $\Delta\theta$, both the original system and the perturbed system operate in normal mode and are driven by identical dynamics. Consequently, the difference process $\{I(t,\theta+\Delta\theta) - I(t,\theta)\}$ is constant and equals $\Delta\theta$, over that interval. By part (c) of Assumption 1, we can choose sufficiently small $\Delta\theta$, such that

$$Z_2(\theta+\Delta\theta) \;=\; Z_1(\theta) + \frac{\Delta S}{\alpha(Z_1(\theta)) - \mu(Z_1(\theta))} \qquad \text{on } B_1 \bigcap \{Z_1(\theta) < R_1(\theta)\}, \tag{4.20}$$

where the inequality $\alpha(Z_1(\theta)) - \mu(Z_1(\theta)) > 0$ follows from the definition of $Z_1(\theta)$. But because in this case, $R_1(\theta+\Delta\theta) \to T_\mu(\theta)$ and $Z_2(\theta+\Delta\theta) \to Z_1(\theta)$ as $\Delta\theta \to 0$ by Eqs.(4.19) and (4.20), we conclude that $\frac{d}{d\theta^+} I(t,\theta) = 1$ on the event $B_1(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$.

Case 2: on the event $B_1(\theta) \bigcap \{Z_1(\theta) = R_1(\theta)\}$ the proof is similar, except that $Z_2(\theta+\Delta\theta)$ need not be considered.

Finally, Part (b) on events $B_{2,j}(\theta)$ and $B_{3,j}(\theta)$, $j \geq 1$ follows immediately from part (a) of Lemma 1, and part (c) follows immediately from part (b) of Lemma 1. □

We next derive the left-derivatives, $\frac{d}{d\theta^-} I(t,\theta)$, by borrowing from Lemma 5, and making use of the hitting time T_α, given by

$$T_\alpha \;=\; \begin{cases} \min\{t \in [0,T] : \; \alpha(t) > 0\}, & \text{if the minimum exists} \\ \infty, & \text{otherwise} \end{cases} \tag{4.21}$$

Note that T_α is independent of θ.

Lemma 7 *Consider an MTS system with the lost-sales rule on the event* $\{I(0) = S(\theta)\}$ *(that is, the system starts in normal operation with full inventory). Then, for any* $t \geq 0$ *and* $\theta \in \Theta$,
(a) *On the event* $A(\theta) = \{I(0) = S(\theta)\} \bigcap \{t < T_\alpha\}$,

$$\frac{d}{d\theta^-} I(t, \theta) \; = \; 0.$$

(b) *On either of the events* $B_1(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\alpha < Q_1(\theta)\} \bigcap \{T_\alpha < t < Z_1(\theta)\}$ *or* $B_{2,j}(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\alpha < Q_1(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$ *or* $B_{3,j}(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\alpha = Q_1(\theta)\} \bigcap \{R_j(\theta) < t < Z_{j+1}(\theta)\}$, $j \geq 1$,

$$\frac{d}{d\theta^-} I(t, \theta) \; = \; 1.$$

(c) *On the events* $C_j(\theta) = \{I(0) = S(\theta)\} \bigcap \{Z_j(\theta) < t < R_j(\theta)\}$, $j \geq 1$,

$$\frac{d}{d\theta^-} I(t, \theta) \; = \; 0.$$

(d) *On the event* $D(\theta) = \{I(0) = S(\theta)\} \bigcap \{T_\alpha = Q_1(\theta)\} \bigcap \{Q_1(\theta) < t < Z_1(\theta)\}$,

$$\frac{d}{d\theta^-} I(t, \theta) \; = \; \frac{\mu(T_\alpha)}{\alpha(T_\alpha)}.$$

Proof. Consider a perturbed system with $S(\theta - \Delta\theta) = \theta - \Delta\theta$, where $\Delta\theta > 0$. Since $I(0) = S(\theta) > S(\theta - \Delta\theta)$ by assumption, it follows that the perturbed system starts in overage operation. Denote $\Delta S = S(\theta) - S(\theta - \Delta\theta)$.

To prove part (a), note first that the event $\{T_\alpha = 0\}$ can be precluded, since it implies $A(\theta) = \emptyset$. Otherwise, on the event $A(\theta)$, the perturbed system is in overage operation with no demand arrivals, so that $I(t, \theta - \Delta\theta) = I(0) = I(t, \theta)$ on the event $A(\theta)$, and the result follows immediately.

To prove part (b), observe that part (b) of Assumption 1 implies that there exists $\epsilon > 0$, such that for any $\Delta\theta \leq \epsilon$,

$$T_S(\theta - \Delta\theta) \; = \; T_\alpha + \frac{\Delta S}{\alpha(T_\alpha)} \qquad \text{on } \{I(0) = S(\theta)\}, \tag{4.22}$$

where the inequality $\alpha(T_\alpha) > 0$ follows from the definition of T_α. Note that Observation 3 implies

$$T_\alpha \; < T_S(\theta - \Delta\theta) \; < \; Q_1(\theta) \; < \; R_1(\theta) \quad \text{ on } \; B_1(\theta). \tag{4.23}$$

We now proceed with the proof on the event $B_1(\theta)$, by considering two cases.

Case b.1: On the event $B_1(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$, Eq. (4.23) implies that $Q_1(\theta)$ is a jump point of $\{\alpha(t) - \mu(t)\}$. It follows by a proof similar to that of part (a) in Lemma 1 that $I(t, \theta - \Delta\theta) = I(t, \theta) - \Delta\theta$ on the event $\{I(0) = S(\theta)\} \bigcap \{T_\alpha < Q_1(\theta)\} \bigcap \{T_S(\theta - \Delta\theta) < t < Z_1(\theta - \Delta\theta)\}$, where

$$Z_1(\theta - \Delta\theta) \; = \; Z_1(\theta) - \frac{\Delta S}{\alpha(Z_1(\theta)) - \mu(Z_1(\theta))} \qquad \text{on } B_1(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}, \tag{4.24}$$

and the inequality $\alpha(P_1(\theta)) - \mu(Z_1(\theta)) > 0$ follows from the definition of $Z_1(\theta)$. To see that, observe that $\{I(t, \theta - \Delta\theta)\}$ starts in overage mode and hits its base-stock level, $S(\theta - \Delta\theta)$ at time $T_S(\theta - \Delta\theta)$, and then stays there until time $Q_1(\theta - \Delta\theta) = Q_1(\theta)$. But because $T_S(\theta - \Delta\theta) \to T_\alpha$

and $Z_1(\theta - \Delta\theta) \to Z_1(\theta)$ on $B_1(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$ as $\Delta\theta \to 0$ by Eqs.(4.22) and (4.24), we conclude that $\frac{d}{d\theta^-} I(t,\theta) = 1$ on this event.

Case b.2: On the event $B_1(\theta) \bigcap \{Z_1(\theta) = R_1(\theta)\}$, the proof is similar, except that except that $Z_1(\theta - \Delta\theta)$ need not be considered.

Part (b) on the remaining events, $B_{2,j}$ and $B_{3,j}$, $j \geq 1$, follows immediately from part (a) of Lemma 1, while part (c) follows immediately from part (b) of Lemma 1.

Finally, we prove part (d) by considering two separate cases. Here, the process $\{I(t,\theta)\}$ stays at $S(\theta)$ until time T_α, at which point the arrival rate jumps, such that $\alpha(T_\alpha) > \mu(T_\alpha)$.

Case d.1: Consider the event $D(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$. Then,

$$Q_1(\theta) \;=\; T_\alpha \;<\; T_S(\theta - \Delta\theta) \;=\; Q_1(\theta - \Delta\theta) \;<\; R_1(\theta) \quad \text{on} \;\; D(\theta), \tag{4.25}$$

where the first inequality is a consequence of Eq. (4.22), and both the second equality and second inequality follow from part (b) of Assumption 1. It follows that

$$I(T_S(\theta - \Delta\theta), \theta) \;=\; S(\theta) \;-\; \frac{\Delta S}{\alpha(T_\alpha)}\,[\alpha(T_\alpha) - \mu(T_\alpha)] \quad \text{on} \;\; D(\theta). \tag{4.26}$$

But over the interval $[T_S(\theta - \Delta\theta), Z_1(\theta - \Delta\theta)]$, both the original system and the perturbed system operate in normal mode and are driven by identical dynamics, given by Eq. (2.5). Therefore, over this interval,

$$I(t,\theta) - I(t,\theta - \Delta\theta) \;=\; \frac{\Delta S\,\mu(T_\alpha)}{\alpha(T_\alpha)} \qquad \text{on } D(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\} \tag{4.27}$$

is constant, and by part (c) of Assumption 1, we can choose sufficiently small $\Delta\theta$, such that

$$Z_1(\theta - \Delta\theta) \;=\; Z_1(\theta) \;-\; \frac{\Delta S\,\mu(T_\alpha)}{\alpha(T_\alpha)\,[\mu(Z_1(\theta)) - \alpha(Z_1(\theta))]} \quad \text{on} \;\; D(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}. \tag{4.28}$$

Next, send $\Delta S = \Delta\theta \to 0$ in Eqs. (4.22) and (4.28), yielding respectively, $T_S(\theta - \Delta\theta) \to T_\alpha$ and $Z_1(\theta - \Delta\theta) \to Z_1(\theta)$. The requisite result on the event $D(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$ now follows from Eq. (4.27).

Case d.2: Consider the event $D(\theta) \bigcap \{Z_1(\theta) = R_1(\theta)\}$. The proof for this case is identical to the proof of part (d) in Lemma 6 of Zhao and Melamed (2004). □

In the next proof we shall make use of horizon-dependent random indices, $J_S(T,\theta)$, which constitute restrictions of $J(\theta)$ to finite time horizons $[0,T]$ as follows,

$$J_S(T,\theta) \;=\; \begin{cases} \max\{j \geq 1 : \;\; R_j(\theta) \leq T\}, & \text{if it exists} \\ 0, & \text{otherwise.} \end{cases} \tag{4.29}$$

We are now in a position to derive the IPA derivatives for the inventory time average, $L_I(T,\theta)$.

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:*

(a) *On the event $\{I(0) < S(\theta)\}$,*

$$\frac{d}{d\theta} L_I(T,\theta) \;=\; \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta), T\} - R_j(\theta)]. \tag{4.30}$$

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

$$\frac{d}{d\theta} L_I(T,\theta) \;=\; 1_{\{T_S(\theta)<T\}}[\min\{Z_1(\theta),T\} - T_S(\theta)] \;+\; \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta),T\} - R_j(\theta)]. \qquad (4.31)$$

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

$$\frac{d}{d\theta} L_I(T,\theta) = 1_{\{T_S(\theta)<T\}} \frac{\mu(T_S(\theta))}{\alpha(T_S(\theta))} [\min\{Z_1(\theta),T\} - T_S(\theta)] + \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta),T\} - R_j(\theta)]. \quad (4.32)$$

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

$$\frac{d}{d\theta^+} L_I(T,\theta) \;=\; \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta),T\} - R_j(\theta)]. \qquad (4.33)$$

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

$$\frac{d}{d\theta^+} L_I(T,\theta) \;=\; 1_{\{T_\mu(\theta)<T\}}[\min\{Z_1(\theta),T\} - T_\mu(\theta)] \;+\; \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta),T\} - R_j(\theta)]. \qquad (4.34)$$

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

$$\frac{d}{d\theta^-} L_I(T,\theta) \;=\; 1_{\{T_\alpha<T\}}[\min\{Z_1(\theta),T\} - T_\alpha] \;+\; \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta),T\} - R_j(\theta)]. \qquad (4.35)$$

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

$$\frac{d}{d\theta^-} L_I(T,\theta) \;=\; 1_{\{T_\alpha<T\}} \frac{\mu(T_\alpha)}{\alpha(T_\alpha)} [\min\{Z_1(\theta),T\} - T_\alpha] \;+\; \sum_{j=1}^{J_S(T,\theta)} [\min\{Z_{j+1}(\theta),T\} - R_j(\theta)]. \quad (4.36)$$

Proof. We show that Leibniz's rule can be applied to Eq. (2.9) yielding

$$\frac{d}{d\theta^\pm} L_I(T,\theta) \;=\; \frac{1}{T}\frac{d}{d\theta^\pm}\int_0^T I(t,\theta)\,dt \;=\; \frac{1}{T}\int_0^T \frac{d}{d\theta^\pm} I(t,\theta)\,dt. \qquad (4.37)$$

To this end, note that Assumption 1 and Lemma 4 - 7 ensure that w.p.1., the sided derivatives $\frac{d}{d\theta^\pm} I(t,\theta)$ exist and are finite over the interval $[0,T]$, except possibly for a finite number of time points. Furthermore, since the end points of the integration interval of Eq. (4.37) are independent of θ, it follows from Proposition 1 that the differentiation and the integration operations commute there. The theorem now follows by substituting the values of the derivatives computed in Lemmas 4 - 7 into Eq. (4.37). □

We next derive the IPA derivatives for the lost sales time average, $L_\zeta(T,\theta)$. Let $N_{[a,b]}(\theta)$ be the number of intervals of the form $[Q_j(\theta), R_j(\theta)]$, such that $Z_j(\theta) < R_j(\theta)$ 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:*
(a) *On the event* $A(\theta) = \{I(0) < S(\theta)\}$,

$$\frac{d}{d\theta} L_\zeta(T,\theta) \;=\; -\frac{N_{(T_S(\theta),T]}(\theta)}{T}. \qquad (4.38)$$

(b) *On the event* $B(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_S(\theta) < Q_1(\theta)\}$,

$$\frac{d}{d\theta} L_\zeta(T,\theta) \;=\; -\frac{N_{(T_S(\theta),T]}(\theta)}{T}. \tag{4.39}$$

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

$$\frac{d}{d\theta} L_\zeta(T,\theta) \;=\; -\frac{1}{T}\left[1_{\{Z_1(\theta)<R_1(\theta),Z_1(\theta)<T\}}\frac{\mu(T_S(\theta))}{\alpha(T_S(\theta))} \;+\; N_{(R_1(\theta),T]}(\theta)\right]. \tag{4.40}$$

(d) *On the event* $D(\theta) = \{I(0) = S(\theta)\}$,

$$\frac{d}{d\theta^+} L_\zeta(T,\theta) \;=\; -\frac{N_{(T_\mu(\theta),T]}(\theta)}{T}. \tag{4.41}$$

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

$$\frac{d}{d\theta^-} L_\zeta(T,\theta) \;=\; -\frac{N_{(T_\alpha,T]}(\theta)}{T}. \tag{4.42}$$

(f) *On the event* $F(\theta) = \{I(0) > S(\theta)\} \bigcap \{T_\alpha = Q_1(\theta)\}$,

$$\frac{d}{d\theta^-} L_\zeta(T,\theta) \;=\; -\frac{1}{T}\left[1_{\{Z_1(\theta)<R_1(\theta),Z_1(\theta)<T\}}\frac{\mu(T_\alpha)}{\alpha(T_\alpha)} \;+\; N_{(R_1(\theta),T]}(\theta)\right]. \tag{4.43}$$

Proof. We prove each case separately.

Case (a): $\frac{d}{d\theta} L_\zeta(T,\theta)$ can be written as

$$\begin{aligned}\frac{d}{d\theta} L_\zeta(T,\theta) \;&=\; \frac{d}{d\theta}\frac{1}{T}\int_0^{\min\{T_S(\theta),T\}} \zeta(t,\theta)\,dt \;+\; \frac{d}{d\theta}\frac{1}{T}\int_{\min\{T_S(\theta),T\}}^{T} \zeta(t,\theta)\,dt \\ &=\; \frac{d}{d\theta}\frac{1}{T}\int_{\min\{T_S(\theta),T\}}^{T} \zeta(t,\theta)\,dt \quad \text{on} \;\; A(\theta).\end{aligned} \tag{4.44}$$

To see that, note that by Lemma 4, the process $\{I(t,\theta)\}$ is locally independent of θ on the event $A(\theta) \bigcap \{t < T_S(\theta)\}$, and consequently, $\int_0^{T_S(\theta)} \zeta(t,\theta)\,dt$ is also independent of θ on the same event, in view of $\zeta(T_S(\theta),\theta) = 0$. We next compute the right hand side of Eq. (4.44) by partitioning the interval $[\min\{T_S(\theta),T\},T]$ into subintervals of the form $[R_j(\theta), \min\{Q_{j+1}(\theta),T\}]$ and $[Q_j(\theta), \min\{R_j(\theta),T\}]$, and computing the requisite derivative of the integral over each sub-interval. Over intervals of the form $[R_j(\theta), \min\{Q_{j+1}(\theta),T\}]$, the process $\{\zeta(t,\theta)\}$ vanishes identically, and consequently,

$$\frac{d}{d\theta}\int_{R_j(\theta)}^{\min\{Q_{j+1}(\theta),T\}} \zeta(t,\theta)\,dt \;=\; 0.$$

For each non-empty interval of the form $[Q_j(\theta), \min\{R_j(\theta),T\}]$, consider the corresponding $Z_j(\theta)$. It follows from Lemma 2 and 3 that

$$\frac{d}{d\theta}\int_{Q_j(\theta)}^{\min\{R_j(\theta),T\}} \zeta(t,\theta)\,dt \;=\; 0 \quad \text{on } \{Z_j(\theta) = R_j(\theta)\}, \tag{4.45}$$

and

$$\frac{d}{d\theta}\int_{Q_j(\theta)}^{\min\{R_j(\theta),T\}} \zeta(t,\theta)\,dt \;=\; \begin{cases} -1 & \text{on } \{Z_j(\theta) < R_j(\theta)\} \bigcap \{Z_j(\theta) < T\} \\ 0 & \text{on } \{Z_j(\theta) < R_j(\theta)\} \bigcap \{T < Z_j(\theta)\}. \end{cases} \tag{4.46}$$

Since the event $\{Z_j(\theta) = T\}$ has probability 0, the result for this case now follows by applying Eq. (4.45) or Eq. (4.46) to each non-empty interval of the form $[Q_j(\theta), \min\{R_j(\theta), T\}]$.

Case (b): Note that by part (c) of Assumption 1, $T_S(\theta)$ is locally differentiable with respect to θ, and $Q_1(\theta)$ is a jump point of $\{\alpha(t) - \mu(t)\}$. Furthermore, $\frac{d}{d\theta} L_\zeta(T, \theta)$ can be written as

$$\begin{aligned}\frac{d}{d\theta} L_\zeta(T, \theta) &= \frac{d}{d\theta} \frac{1}{T} \int_0^{\min\{T_S(\theta), T\}} \zeta(t, \theta)\, dt \;+\; \frac{d}{d\theta} \frac{1}{T} \int_{\min\{T_S(\theta), T\}}^{T} \zeta(t, \theta)\, dt \\ &= \frac{d}{d\theta} \frac{1}{T} \int_{\min\{T_S(\theta), T\}}^{T} \zeta(t, \theta)\, dt \quad \text{on} \quad B(\theta). \end{aligned} \tag{4.47}$$

since the first integral on the right is identically zero in θ. The result for this case now follows by a proof similar to that of Case (a).

Case (c): Note that $\frac{d}{d\theta} L_\zeta(T, \theta)$ can be written as

$$\frac{d}{d\theta} L_\zeta(T, \theta) \;=\; \frac{d}{d\theta} \frac{1}{T} \int_0^{\min\{R_1(\theta), T)\}} \zeta(t, \theta)\, dt \;+\; \frac{d}{d\theta} \frac{1}{T} \int_{\min\{R_1(\theta), T)\}}^{T} \zeta(t, \theta)\, dt \quad \text{on} \quad C(\theta). \tag{4.48}$$

We first show that the first integral on the right of Eq. (4.48) evaluates to

$$\frac{d}{d\theta} \frac{1}{T} \int_0^{\min\{R_1(\theta), T)\}} \zeta(t, \theta)\, dt \;=\; -\frac{1}{T} 1_{\{Z_1(\theta) < R_1(\theta), Z_1(\theta) < T\}} \frac{\mu(T_S(\theta))}{\alpha(T_S(\theta))} \quad \text{on} \quad C(\theta). \tag{4.49}$$

To this end, we consider two events that partition $C(\theta)$. On the event $C(\theta) \bigcap \{Z_1(\theta) = R_1(\theta)\}$ (i.e., the inventory process does not hit zero in the interval $[Q_1(\theta), R_1(\theta)]$), part (a) of Lemma 2 and part (a) of Lemma 3 imply that $\frac{d}{d\theta} \int_0^{\min\{R_1(\theta), T)\}} \zeta(t, \theta)\, dt = 0$. On the complementary event $C(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$ (i.e., the inventory process does hit zero in the interval $[Q_1(\theta), R_1(\theta)]$), a proof similar to that of part (b) of Lemma 2 and and part (b) of Lemma 3 yields

$$\frac{d}{d\theta} \int_{Z_1(\theta)}^{\min\{R_1(\theta), T\}} \zeta(t, \theta)\, dt \;=\; -1_{\{Z_1(\theta) < T\}} \left[\alpha(Z_1(\theta)) - \mu(Z_1(\theta))\right] \frac{d}{d\theta} Z_1(\theta). \tag{4.50}$$

To compute the right-hand side above, we differentiate the relation

$$\int_{Q_1(\theta)}^{Z_1(\theta)} [\alpha(t) - \mu(t)]\, dt = S(\theta)$$

with respect to θ obtaining

$$[\alpha(Z_1(\theta)) - \mu(Z_1(\theta))] \frac{d}{d\theta} Z_1(\theta) \;-\; [\alpha(Q_1(\theta)) - \mu(Q_1(\theta))] \frac{d}{d\theta} Q_1(\theta) \;=\; 1.$$

Substituting the above into Eq. (4.50), and then substituting Eq. (4.15) for $\frac{d}{d\theta} Q_1(\theta)$ results in the expression

$$\frac{d}{d\theta} \int_{Z_1(\theta)}^{\min\{R_1(\theta), T\}} \zeta(t, \theta)\, dt \;=\; -1_{\{Z_1(\theta) < T\}} \frac{\mu(Q_1(\theta))}{\alpha(Q_1(\theta))} \quad \text{on} \;\; C(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\},$$

which completes the computation of the requisite result in Eq. (4.49).

Finally, to compute the second integral on the right-hand side of Eq. (4.48), we merely note that the proof of part (a) implies

$$\frac{d}{d\theta} \int_{\min\{R_1(\theta),T\}}^{T} \zeta(t,\theta)\, dt \;=\; -N_{(R_1(\theta),T]}(\theta) \quad \text{on} \;\; C(\theta),$$

which completes the proof for this case.

Case (d): Consider a perturbed system with $S(\theta+\Delta\theta) = \theta + \Delta\theta$, where $\Delta\theta > 0$, and denote $\Delta S = S(\theta+\Delta\theta) - S(\theta)$. Since $I(0) = S(\theta) < S(\theta+\Delta\theta)$ on $D(\theta)$, it follows that the perturbed system starts in normal operation. We prove this case on three events that partition $D(\theta)$.

Consider the first event, given by $D_1(\theta) = D(\theta) \bigcap \{T_\mu(\theta) < R_1(\theta)\} \bigcap \{Z_1(\theta) < R_1(\theta)\}$. By definition of $T_\mu(\theta)$, one has $0 = Q_1(\theta+\Delta\theta) < Q_1(\theta)$ on $D_1(\theta)$. From the proof of part (b) in Lemma 6 it follows that

$$I(t,\theta+\Delta\theta) \;=\; I(t,\theta) + \Delta\theta \quad \text{on} \;\; D_1(\theta) \bigcap \{R_1(\theta+\Delta\theta) \le t \le Z_1(\theta)\}, \tag{4.51}$$

where $R_1(\theta+\Delta\theta)$ is given by in Eq. (4.19). Since over the interval $[0, Z_2(\theta+\Delta\theta))$, where $Z_2(\theta+\Delta\theta)$ is given by Eq. (4.20), the perturbed system has no lost sales while the original system has lost sales only over the sub-interval $[Z_1(\theta), Z_2(\theta+\Delta\theta))$, we can write,

$$\begin{aligned}
\int_0^{Z_2(\theta+\Delta\theta)} [\zeta(t,\theta+\Delta\theta) - \zeta(t,\theta)]\, dt \;&=\; \int_{Z_1(\theta)}^{Z_2(\theta+\Delta\theta)} [\zeta(t,\theta+\Delta\theta) - \zeta(t,\theta)]\, dt \\
&=\; -\int_{Z_1(\theta)}^{Z_2(\theta+\Delta\theta)} [\alpha(t) - \mu(t)]\, dt \\
&=\; I(Z_1(\theta),\theta) - I(Z_1(\theta),\theta+\Delta\theta) \\
&=\; -\Delta\theta \quad \text{on} \;\; D_1(\theta),
\end{aligned} \tag{4.52}$$

where the next to last equality follows from the fact that the inventory difference at $Z_1(\theta)$ is consumed by the perturbed system in the interval $[Z_1(\theta), Z_2(\theta)+\Delta\theta)$, and the last equality follows from Eq. (4.51). Next, decompose $L_\zeta(T,\theta)$ into two terms as in Eq. (4.48), and evaluate each term on $D_1(\theta)$. In view of Eq. (4.52), the first term evaluates to

$$\frac{d}{d\theta^+} \frac{1}{T} \int_0^{\min\{R_1(\theta),T\}} \zeta(t,\theta)\, dt \;=\; -1_{\{Z_1(\theta)<T\}} \quad \text{on} \;\; D_1(\theta).$$

The evaluation of the second term is similar to the evaluation of the second term in the decomposition of Case (a), since the hitting time $T_S(\theta)$ in Case (a) corresponds to the hitting time $R_1(\theta)$ in this case. Since the two evaluations result in identical expressions, this completes the proof of this case on the event $D_1(\theta)$.

Consider the second event, given by $D_2(\theta) = D(\theta) \bigcap \{T_\mu(\theta) < R_1(\theta)\} \bigcap \{Z_1(\theta) = R_1(\theta)\}$. The proof on $D_2(\theta)$ is identical to that of the proof on $D_1(\theta)$, except that there are no lost sales in the interval $[0, R_1(\theta)]$.

Finally, consider the third event, given by $D_3(\theta) = D(\theta) \bigcap \{T_\mu(\theta) = R_1(\theta)\}$. The proof on $D_3(\theta)$ is essentially identical to that of Case (a) and leads to the same result. The proof of this case is now complete.

Case (e): Consider a perturbed system with $S(\theta-\Delta\theta) = \theta - \Delta\theta$, where $\Delta\theta > 0$, and denote $\Delta S = S(\theta) - S(\theta-\Delta\theta)$. Since $I(0) = S(\theta) > S(\theta-\Delta\theta)$ on $E(\theta)$, it follows that the perturbed system starts in overage operation. We prove this case on two events that partition $E(\theta)$.

Consider the first event, given by $E_1(\theta) = E(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$. By the proof of part (b) in Lemma 7,

$$I(t, \theta - \Delta\theta) \;=\; I(t, \theta) - \Delta\theta \quad \text{on} \;\; E_1(\theta) \bigcap \{T_S(\theta - \Delta\theta) \le t \le Z_1(\theta - \Delta\theta)\}, \tag{4.53}$$

where $T_S(\theta - \Delta\theta)$ is given in Eq. (4.22) and $Z_1(\theta - \Delta\theta)$ is given in Eq. (4.24). Since over the interval $[0, Z_1(\theta))$, the original system has no lost sales while the perturbed system has lost sales only over the sub-interval $[Z_1(\theta - \Delta\theta), Z_1(\theta))$, we can write,

$$\begin{aligned}
\int_0^{Z_1(\theta)} [\zeta(t,\theta) - \zeta(t, \theta - \Delta\theta)]\, dt \;&=\; \int_{Z_1(\theta-\Delta\theta)}^{Z_1(\theta)} [\zeta(t,\theta) - \zeta(t, \theta - \Delta\theta)]\, dt \\
&=\; - \int_{Z_1(\theta-\Delta\theta)}^{Z_1(\theta)} [\alpha(t) - \mu(t)]\, dt \\
&=\; I(Z_1(\theta - \Delta\theta), \theta - \Delta\theta) - I(Z_1(\theta - \Delta\theta), \theta) \\
&=\; -\Delta\theta \quad \text{on} \;\; E_1(\theta),
\end{aligned} \tag{4.54}$$

where the next to last equality follows from the fact that the inventory difference at $Z_1(\theta - \Delta\theta)$ is consumed by the original system in the interval $[Z_1(\theta - \Delta\theta), Z_1(\theta))$, and the last equality follows from Eq. (4.53). Next, decompose $L_\zeta(T, \theta)$ into two terms as in Eq. (4.48), and evaluate each term on $E_1(\theta)$. In view of Eq. (4.54), the first term evaluates to

$$\frac{d}{d\theta^-} \frac{1}{T} \int_0^{\min\{R_1(\theta), T\}} \zeta(t, \theta)\, dt \;=\; -1_{\{Z_1(\theta) < T\}} \quad \text{on} \;\; E_1(\theta).$$

The evaluation of the second term is similar to the evaluation of the second term in the decomposition of Case (a), since the hitting time $T_S(\theta)$ in Case (a) corresponds to the hitting time $R_1(\theta)$ in this case. Since the two evaluations result in identical expressions, this completes the proof of this case on the event $E_1(\theta)$.

Consider the second event, given by $E_2(\theta) = E(\theta) \bigcap \{Z_1(\theta) = R_1(\theta)\}$. The proof on $E_2(\theta)$ is identical to that of the proof on $E_1(\theta)$, except that there are no lost sales in the interval $[0, R_1(\theta)]$. The proof of this case is now complete.

Case (f): The setting of this case is the same as in Case (e). We prove this case on two events that partition $F(\theta)$.

Consider the first event, given by $F_1(\theta) = F(\theta) \bigcap \{Z_1(\theta) < R_1(\theta)\}$. By part (d) in Lemma 7,

$$I(t, \theta - \Delta\theta) \;=\; I(t, \theta) - \Delta\theta \, \frac{\mu(T_\alpha)}{\alpha(T_\alpha)} \quad \text{on} \;\; F_1(\theta) \bigcap \{T_S(\theta - \Delta\theta) \le t \le Z_1(\theta - \Delta\theta)\}, \tag{4.55}$$

where $T_S(\theta - \Delta\theta)$ is given in Eq. (4.22) and $Z_1(\theta - \Delta\theta)$ is given in Eq. (4.28). Since over the interval $[0, Z_1(\theta))$, the original system has no lost sales while the perturbed system has lost sales only over the sub-interval $[Z_1(\theta - \Delta\theta), Z_1(\theta))$, we can write,

$$\begin{aligned}
\int_0^{Z_1(\theta)} [\zeta(t,\theta) - \zeta(t, \theta - \Delta\theta)]\, dt \;&=\; \int_{Z_1(\theta-\Delta\theta)}^{Z_1(\theta)} [\zeta(t,\theta) - \zeta(t, \theta - \Delta\theta)]\, dt \\
&=\; - \int_{Z_1(\theta-\Delta\theta)}^{Z_1(\theta)} [\alpha(t) - \mu(t)]\, dt \\
&=\; I(Z_1(\theta - \Delta\theta), \theta - \Delta\theta) - I(Z_1(\theta - \Delta\theta), \theta) \\
&=\; -\Delta\theta \, \frac{\mu(T_\alpha)}{\alpha(T_\alpha)}, \quad \text{on} \;\; F_1(\theta),
\end{aligned} \tag{4.56}$$

where the next to last equality follows from the fact that the inventory difference at $Z_1(\theta - \Delta\theta)$ is consumed by the original system in the interval $[Z_1(\theta - \Delta\theta), Z_1(\theta))$, and the last equality follows from Eq. (4.55). Next, decompose $L_\zeta(T, \theta)$ into two terms as in Eq. (4.48), and evaluate each term on $F_1(\theta)$. In view of Eq. (4.56), the first term evaluates to

$$\frac{d}{d\theta^-} \frac{1}{T} \int_0^{\min\{R_1(\theta), T\}} \zeta(t, \theta)\, dt \;=\; -1_{\{Z_1(\theta) < T\}} \frac{\mu(T_\alpha)}{\alpha(T_\alpha)} \quad \text{on} \quad F_1(\theta).$$

The evaluation of the second term is similar to the evaluation of the second term in the decomposition of Case (a), since the hitting time $T_S(\theta)$ in Case (a) corresponds to the hitting time $R_1(\theta)$ in this case. Since the two evaluations result in identical expressions, this completes the proof of this case on the event $F_1(\theta)$.

Consider the second event, given by $F_2(\theta) = F(\theta) \bigcap \{Z_1(\theta) = R_1(\theta)\}$. The proof on $F_2(\theta)$ is identical to that of the proof on $F_1(\theta)$, except that there are no lost sales in the interval $[0, R_1(\theta)]$. The proof of this case is now complete. □

Finally, we show that the IPA derivatives of Theorem 1 and Theorem 2 are unbiased.

Theorem 3 *Under Assumptions 1, 2 and 3, the sided IPA derivatives with respect to the base level parameter,* $\frac{d}{d\theta^\pm} L_I(T, \theta)$ *and* $\frac{d}{d\theta^\pm} L_\zeta(T, \theta)$*, are unbiased for all* $T > 0$ *and* $\theta \in \Theta$*.*

Proof. Theorem 1 and Theorem 2 ensure that for all $T > 0$, Condition (a) of Fact 1 is satisfied for both $L_I(T, \theta)$ and $L_\zeta(T, \theta)$. Now, for any $\theta_1, \theta_2 \in \Theta$,

$$\begin{aligned} |L_I(T, \theta_1) - L_I(T, \theta_2)| &= \left| \frac{1}{T} \int_0^T [I(t, \theta_1) - I(t, \theta_2]\, dt \right| \\ &\leq \frac{1}{T} \int_0^T |I(t, \theta_1) - I(t, \theta_2)|\, dt \;\leq\; |\theta_1 - \theta_2|, \end{aligned} \tag{4.57}$$

where the second inequality is a consequence of Proposition 1. Furthermore, by Proposition 2,

$$|L_\zeta(T, \theta_1) - L_\zeta(T, \theta_2)| \;\leq\; \frac{\max\{K(\theta_1), K(\theta_2)\}}{T} |\theta_1 - \theta_2|, \tag{4.58}$$

where $E[K(\theta)]$ is finite because $K(\theta)$ is finite w.p.1 for any finite time horizon $[0, T]$ and any $\theta \in \Theta$ by part (b) of Assumption 1.

Eqs. (4.57) and (4.58) establish that Condition (b) of Fact 1 holds for both $L_I(T, \theta)$ and $L_\zeta(T, \theta)$, thereby completing their proof of unbiasedness. □

4.2 IPA Derivatives with Respect to the Production Rate Parameter

This section treats 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 the production rate parameter, θ, in $\{\mu(t, \theta)\}$. We first prove a number of useful lemmas on local differentiability and local independence, which simplify the proofs of the main results later in this section. We then proceed to obtain the sided IPA derivatives for $L_I(T, \theta)$ by first obtaining those for the inventory process, $\{I(t, \theta)\}$, following which we obtain the sided IPA derivatives for $L_\zeta(T, \theta)$. Finally, we establish the stochastic ordering of the sided derivatives for $L_\zeta(T, \theta)$, and the unbiasedness of all the IPA derivatives above.

Assumption 4

(a) *The production rate process* $\{\mu(t,\theta)\}$ *is subject to Eq. (2.12).*

(b) *The process* $\{\alpha(t)\}$ *and the base-stock level, S, are independent of* θ. □

We point out that unlike Wardi et al. (2002), Paschalidis et al. (2004) and Zhao and Melamed (2004), 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 (a) 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 replenishment process, $\{\rho(t)\}$, vanishes during overage operation, so that the value of θ has no effect on the state of the system until it enters normal mode.

Define $(U_m^+(\theta), V_m^+(\theta))$, $m=1,\ldots,M(\theta)$, to be the ordered extremal subintervals of $[0,\infty)$, such that either $I(t,\theta)=S$ for all $t\in(U_m^+,V_m^+)$ or $I(t,\theta)=0$ and $\alpha(t)>\mu(t,\theta)$ for all $t\in(U_m^+,V_m^+)$. Define further $(U_n^-(\theta), V_n^-(\theta))$, $n=1,\ldots,N(\theta)$, to be the ordered extremal subintervals of $[0,\infty)$, such that either $I(t,\theta)=0$ for all $t\in(U_n^-,V_n^-)$ or $I(t,\theta)=S$ and $\alpha(t)<\mu(t,\theta)$ for all $t\in(U_n^-,V_n^-)$. By convention, if any of the endpoints above, $U_m^+(\theta)$, $V_m^+(\theta)$, $U_n^-(\theta)$ or $V_n^-(\theta)$, does not exist, then it is set to ∞. For notational convenience, define $V_0^+(\theta)=V_0^-(\theta)=0$.

Observation 4

$$U_1^+(\theta) < V_1^+(\theta) < U_2^+(\theta) < V_2^+(\theta) < \cdots < U_{M(\theta)}^+(\theta) < V_{M(\theta)}^+(\theta) \tag{4.59}$$

$$U_1^-(\theta) < V_1^-(\theta) < U_2^-(\theta) < V_2^-(\theta) < \cdots < U_{N(\theta)}^-(\theta) < V_{N(\theta)}^-(\theta) \tag{4.60}$$

Proof. To prove Inequality (4.59), we show the following inequalities

$$U_m^+(\theta) \neq V_m^+(\theta), \quad \text{for all} \quad m \ge 1 \tag{4.61}$$

$$V_m^+(\theta) \neq U_{m+1}^+(\theta), \quad \text{for all} \quad m \ge 1 \tag{4.62}$$

To prove inequality (4.61), we consider the following two cases.

Case 1: $I(t,\theta)=S$ on the event $\{U_m^+(\theta)<t<V_m^+(\theta)\}$. Inequality (4.61) holds due to part (c) of Assumption 1, because $U_m^+(\theta)$ is a point at which a time interval is inaugurated over which $I(t,\theta)=S$, while $V_m^+(\theta)$ is a jump point of $\{\alpha(t)-\mu(t,\theta)\}$ from a non-positive value to a positive value.

Case 2: $I(t,\theta)=0$ and $\alpha(t)>\mu(t,\theta)$ on the event $\{U_m^+(\theta)<t<V_m^+(\theta)\}$. Inequality (4.61) holds due to part (c) of Assumption 1, because either $U_m^+(\theta)$ is a point at which a time interval is inaugurated over which $I(t,\theta)=0$ or $U_m^+(\theta)$ is a jump point of $\{\alpha(t)-\mu(t,\theta)\}$ from zero to a positive value, while $V_m^+(\theta)$ is a jump point of $\{\alpha(t)-\mu(t,\theta)\}$ from a positive value to a non-positive value.

To prove inequality (4.62), note that by the extremality of the intervals $(U_m^+(\theta), V_m^+(\theta))$ it is impossible to have $I(t,\theta)=S$ on events $\{U_m^+(\theta)<t<V_m^+(\theta)\}$ and $\{U_{m+1}^+(\theta)<t<V_{m+1}^+(\theta)\}$, or $I(t,\theta)=0$ and $\alpha(t)>\mu(t,\theta)$ on the events $\{U_m^+(\theta)<t<V_m^+(\theta)\}$ and $\{U_{m+1}^+(\theta)<t<V_{m+1}^+(\theta)\}$. It remains to consider the cases where $I(t,\theta)=S$ on the event $\{U_m^+(\theta)<t<V_m^+(\theta)\}$ and $I(t,\theta)=0$ and $\alpha(t)>\mu(t,\theta)$ on the event $\{U_{m+1}^+(\theta)<t<V_{m+1}^+(\theta)\}$, or $I(t,\theta)=0$ and $\alpha(t)>\mu(t,\theta)$ on the event $\{U_m^+(\theta) < t < V_m^+(\theta)\}$ and $I(t,\theta) = S$ on the event $\{U_{m+1}^+(\theta) < t < V_{m+1}^+(\theta)\}$. Inequality (4.62) now follows immediately by the continuity of $\{I(t,\theta)\}$ with respect to t.

Finally, the proof of (4.60) is analogous to that of (4.59). □

We shall need the following horizon-dependent random indices. The restriction of $M(\theta)$ to a finite time horizon $[0,T]$ is

$$M_I(T,\theta) = \begin{cases} \max\{m \geq 1: \; V_m^+(\theta) \leq T\}, & \text{if it exists} \\ 0, & \text{otherwise} \end{cases} \tag{4.63}$$

and the restriction of $N(\theta)$ to a finite time horizon $[0,T]$ is

$$N_I(T,\theta) = \begin{cases} \max\{n \geq 1: \; V_n^-(\theta) \leq T\}, & \text{if it exists} \\ 0, & \text{otherwise.} \end{cases} \tag{4.64}$$

The next two lemmas provide properties of horizon dependent indices $M_I(T,\theta)$ and $N_I(T,\theta)$, and the end points of the intervals $(U_m^+(\theta), V_m^+(\theta))$ and $(U_m^-(\theta), V_m^-(\theta))$.

Lemma 8

(a) *$M_I(T,\theta)$ is locally independent of θ in a right neighborhood of θ.*

(b) *For all $m = 1, \cdots, M(\theta)$, $U_m^+(\theta)$ is locally differentiable with respect to θ in a right neighborhood of θ.*

(c) *For all $m = 1, \cdots, M(\theta)$, $V_m^+(\theta)$ is locally independent of θ in a right neighborhood of θ.*

Proof. To prove part (a), note that if $I(t,\theta) = S$ on $\{U_m^+(\theta) < t < V_m^+(\theta)\}$, then it follows from Corollary 1 that in a right neighborhood of θ, $I(t,\theta+\Delta\theta) = S$ on $\{U_m^+(\theta) < t < V_m^+(\theta)\}$. Consider next the case where $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ on $\{U_m^+(\theta) < t < V_m^+(\theta)\}$. By Eq. (2.12), there exists a right neighborhood of θ, such that $\alpha(t) > \mu(t,\theta+\Delta\theta)$. It follows from Corollary 1 and Observation 4 that in a right neighborhood of θ, the interval $(U_m^+(\theta), V_m^+(\theta))$ remains non-null. Suppose that part (a) does not hold. Then there must exist a $\tau \in [0,T]$ such that $I(\tau,\theta) = S$ and $I(t,\theta) < S$ for any t in a neighborhood of τ. But this is precluded by part (c) of Assumption 1, so (a) must hold.

To prove part (b), we consider four cases.

Case 1: $I(t,\theta) = S$ on $\{0 < U_m^+(\theta) < t < V_m^+(\theta)\}$. Follows immediately from Corollary 1.

Case 2: $I(t,\theta) = S$ on $\{0 = U_1^+(\theta) < t < V_1^+(\theta)\}$. Follows from the inequalities

$$\alpha(0) \leq \mu(0,\theta) < \mu(0,\theta+\Delta\theta) \quad \text{for any } \Delta\theta > 0,$$

where the first inequality is a consequence of the definition of the intervals $(U_m^+(\theta), V_m^+(\theta))$, and the second inequality is a consequence of Eq. (2.12).

Case 3: $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ on $\{0 < U_m^+(\theta) < t < V_m^+(\theta)\}$. Follows immediately from Corollary 1.

Case 4: $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ on $\{0 = U_1^+(\theta) < t < V_1^+(\theta)\}$. Follows by the fact that there exists a right neighborhood of θ such that $\alpha(t) > \mu(t,\theta+\Delta\theta)$.

To prove part (c), we consider two cases.

Case 5: $I(t,\theta) = S$ on $\{U_m^+(\theta) < t < V_m^+(\theta)\}$. Note that the process $\{\alpha(t) - \mu(t,\theta)\}$ jumps at $V_m^+(\theta)$ from a non-positive value to a positive value. The result now follows, since by part (b) of Assumption 1, such jumps are independent of θ.

Case 6: $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ on $\{U_m^+(\theta) < t < V_m^+(\theta)\}$. Note that the process $\{\alpha(t) - \mu(t,\theta)\}$ jumps at $V_m^+(\theta)$ from a positive value to a non-positive value. The rest of the argument is identical to that of Case 5. □

Lemma 9

(a) *$N_I(T,\theta)$ is locally independent of θ in a left neighborhood of θ.*

(b) *For all $n = 1, \ldots, N(\theta)$, $U_n^-(\theta)$ is locally differentiable with respect to θ in a left neighborhood of θ.*

(c) *For all $n = 1, \ldots, N(\theta)$, $V_n^-(\theta)$ is locally independent of θ in a left neighborhood of θ.*

Proof. Analogous to the proof of Lemma 8. □

The next two lemmas compute the sided derivatives, $\dfrac{d}{d\theta^\pm} I(t,\theta)$.

Lemma 10 *Consider an MTS system with the lost-sales rule. Then, for any $t \geq 0$ and $\theta \in \Theta$,*
(a) *On the events $A_m(\theta) = \{U_m^+(\theta) < t < V_m^+(\theta)\}$, $m = 1, \ldots, M(\theta)$,*

$$\frac{d}{d\theta^+} I(t,\theta) \;=\; 0.$$

(b) *On the events $B_m(\theta) = \{V_m^+(\theta) < t < U_{m+1}^+(\theta)\}$, $m = 0, 1, \ldots, M(\theta) - 1$,*

$$\frac{d}{d\theta^+} I(t,\theta) \;=\; t - V_m^+(\theta).$$

Proof. To prove parts (a), we consider two cases.

Case 1: $I(t,\theta) = S$ on the event $A_m(\theta)$ for given m. In this case, $\alpha(t) - \mu(t, \theta + \Delta\theta) < 0$ on the event $A_m(\theta)$ for sufficiently small $\Delta\theta > 0$ by Eq. (2.12). It follows that $I(t,\theta) = S$ on $A_m(\theta)$ in a right neighborhood of θ, whence Part (a) holds.

Case 2: $I(t,\theta) = 0$ and $\alpha(t) > \mu(t,\theta)$ on the event $A_m(\theta)$ for given m. In this case, $\alpha(t) - \mu(t, \theta + \Delta\theta) > 0$ on the event $A_m(\theta)$ for sufficiently small $\Delta\theta > 0$ by Eq. (2.12). It follows that $I(t,\theta) = 0$ on $A_m(\theta)$ in a right neighborhood of θ, whence Part (b) holds.

To prove part (b), write

$$I(t,\theta) \;=\; I(V_m^+(\theta), \theta) + \int_{V_m^+(\theta)}^{t} [\mu(\tau,\theta) - \alpha(\tau)]\, d\tau, \qquad \text{on } B_m(\theta). \tag{4.65}$$

Note that $I(V_m^+(\theta),\theta) = S$ or $I(V_m^+(\theta),\theta) = 0$ in a right neighborhood of θ by part (c) of Lemma 8, implying that $I(V_m^+(\theta),\theta)$ is locally independent of θ in a right neighborhood of θ. Consequently, taking right derivatives with respect to θ in Eq. (4.65) yields

$$\frac{dI(t,\theta)}{d\theta^+} \;=\; -[\mu(V_m^+(\theta),\theta) - \alpha(V_m^+(\theta))]\, \frac{d}{d\theta^+} V_m^+(\theta) \;+\; \int_{V_m^+(\theta)}^{t} d\tau \;=\; t - V_m^+(\theta), \qquad \text{on } B_m(\theta),$$

where the first term on the right is due to part (c) of Lemma 8 and the second term is due to Eq. (2.12). The proof is now complete. □

Lemma 11 *Consider an MTS system with the lost-sales rule. Then, for any $t \geq 0$ and $\theta \in \Theta$,*
(a) *On the events $A_n(\theta) = \{U_n^-(\theta) < t < V_n^-(\theta)\}$, $n = 1, \ldots, N(\theta)$,*

$$\frac{d}{d\theta^-} I(t,\theta) \;=\; 0.$$

(b) *On the events $B_n(\theta) = \{V_n^-(\theta) < t < U_{n+1}^-(\theta)\}$, $n = 0, 1, \ldots, N(\theta) - 1$,*

$$\frac{d}{d\theta^-} I(t,\theta) \;=\; t - V_n^-(\theta).$$

Proof. Analogous to the proof of Lemma 10. □

We are now in a position to derive the IPA derivatives for the inventory time average $L_I(T,\theta)$.

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:*

$$\frac{d}{d\theta^+} L_I(T,\theta) \;=\; \frac{1}{2T} \sum_{m=0}^{M_I(T,\theta)} [\min\{U_{m+1}^+(\theta), T\} - V_m^+(\theta)]^2, \tag{4.66}$$

$$\frac{d}{d\theta^-} L_I(T,\theta) \;=\; \frac{1}{2T} \sum_{n=0}^{N_I(T,\theta)} [\min\{U_{n+1}^-(\theta), T\} - V_n^-(\theta)]^2. \tag{4.67}$$

Proof. We show that Leibniz's rule can be applied to Eq. (2.9) yielding

$$\frac{d}{d\theta^\pm} L_I(T,\theta) \;=\; \frac{1}{T} \frac{d}{d\theta^\pm} \int_0^T I(t,\theta)\, dt \;=\; \frac{1}{T} \int_0^T \frac{d}{d\theta^\pm} I(t,\theta)\, dt. \tag{4.68}$$

To this end, note that Assumption 1 and Lemmas 10 and 11 ensure that w.p.1., the sided derivatives $\frac{d}{d\theta^\pm} I(t,\theta)$ exist and are finite over the interval $[0,T]$, except possibly for a finite number of time points. Furthermore, since the end points of the integration interval of Eq. (4.68) are independent of θ, it follows from Corollary 1 that the differentiation and integration operations commute there. The theorem now follows by substituting the values of the derivatives computed in Lemmas 10 and 11 into Eq. (4.68). □

We next derive IPA derivatives for the lost-sales time average $L_\zeta(T,\theta)$. We shall need the following horizon-dependent random indices.

$$M_\zeta(T,\theta) \;=\; \begin{cases} \max\{m \geq 1 : \; U_m^+(\theta) \leq T\}, & \text{if it exists} \\ 0, & \text{otherwise} \end{cases} \tag{4.69}$$

$$N_\zeta(T,\theta) \;=\; \begin{cases} \max\{n \geq 1 : \; U_n^-(\theta) \leq T\}, & \text{if it exists} \\ 0, & \text{otherwise} \end{cases} \tag{4.70}$$

Let $\Phi(T,\theta)$ be the set of all indices $m \in \{1, 2, \ldots, M_\zeta(T,\theta)\}$ 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 $n \in \{1, 2, \ldots, N_\zeta(T,\theta)\}$ such that $I(t,\theta) = 0$ on the event $\{U_n^-(\theta) < t < V_n^-(\theta)\}$. It follows from the proof of part (a) of Lemma 8 that the set $\Phi(T,\theta)$ is locally independent of θ in a right neighborhood of θ, and the set $\Psi(T,\theta)$ is locally independent of θ in a left neighborhood of θ.

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:*

$$\frac{d}{d\theta^+} L_\zeta(T,\theta) \;=\; -\frac{1}{T} \sum_{m \in \Phi(T,\theta)} [\min\{V_m^+(\theta), T\} - V_{m-1}^+(\theta)] \tag{4.71}$$

$$\frac{d}{d\theta^-} L_\zeta(T,\theta) \;=\; -\frac{1}{T} \sum_{n \in \Psi(T,\theta)} [\min\{V_n^-(\theta), T\} - V_{n-1}^-(\theta)] \tag{4.72}$$

Proof. To prove Eq. (4.71), note that

$$\frac{d}{d\theta^+} L_\zeta(T,\theta) \;=\; \frac{1}{T} \sum_{m \in \Phi(T,\theta)} \frac{d}{d\theta^+} \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} \zeta(t,\theta)\, dt, \tag{4.73}$$

because the set $\Phi(T,\theta)$ is locally independent of θ in a right neighborhood of θ. For each $m \in \Phi(T,\theta)$, part (c) of Lemma 8 and Eq. (2.12) imply

$$\frac{d}{d\theta^+} \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} \zeta(t,\theta)\, dt \;=\; -[\alpha(U_m^+(\theta)) - \mu(U_m^+(\theta),\theta)]\, \frac{d}{d\theta^+} U_m^+(\theta) \;-\; \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} dt. \tag{4.74}$$

We next compute the first term on the right-hand side of Eq. (4.74), and then take its right derivative with respect to θ. To this end, we first compute the integrals $\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt$ for two cases of $m \in \Phi(T,\theta)$.

Case 1: $m = 1$. On the event $\{U_1^+(\theta) > 0\}$,

$$\int_{V_0^+(\theta)}^{U_1^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; I(V_0^+(\theta)) - I(U_1^+(\theta)) \;=\; I(0),$$

since $I(U_1^+(\theta)) = 0$. Furthermore, on the complementary event $\{U_1^+(\theta) = 0\}$,

$$\int_{V_0^+(\theta)}^{U_1^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt = 0,$$

since $V_0^+(\theta) = U_1^+(\theta) = 0$.

Case 2: $m > 1$. On the event $\{I(t,\theta) = S,\; U_{m-1}^+(\theta) < t < V_{m-1}^+(\theta)\}$,

$$\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; I(V_{m-1}^+(\theta)) - I(U_m^+(\theta)) \;=\; S,$$

since $I(V_{m-1}^+(\theta)) = S$ and $I(U_m^+(\theta)) = 0$. Furthermore, on the complementary event $\{I(t,\theta) = 0 \text{ and } \mathrm{\alpha(t) > \mu(t,\theta),\; U_{m-1}^+(\theta) < t < V_{m-1}^+(\theta)}\}$

$$\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; I(V_{m-1}^+(\theta)) - I(U_m^+(\theta)) \;=\; 0,$$

since $I(V_{m-1}^+(\theta)) = I(U_m^+(\theta)) = 0$.

It follows that in all cases above, the integral $\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt$ equals either $I(0)$ or 0 or S, all of which are independent of θ. Taking the right derivative of this integral with respect to θ yields

$$-[\alpha(V_{m-1}^+(\theta)) - \mu(V_{m-1}^+(\theta),\theta)] \frac{d}{d\theta^+} V_{m-1}^+(\theta) + [\alpha(U_m^+(\theta)) - \mu(U_m^+(\theta),\theta)] \frac{d}{d\theta^+} U_m^+(\theta) - \int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} dt = 0,$$

in view of Eq. (2.12). But the first term in the above equation varnishes by part (c) of Lemma 8, resulting in

$$-[\alpha(U_m^+(\theta)) - \mu(U_m^+(\theta),\theta)] \frac{d}{d\theta^+} U_m^+(\theta) \;=\; -\int_{V_{m-1}^+(\theta)}^{U_m^+(\theta)} dt.$$

Finally, substitute the above equation into Eq. (4.74), yielding

$$\frac{d}{d\theta^+} \int_{U_m^+(\theta)}^{\min\{V_m^+(\theta),T\}} \zeta(t,\theta)\, dt \;=\; -\int_{V_{m-1}^+(\theta)}^{\min\{V_m^+(\theta),T\}} dt \;=\; -[\min\{V_m^+(\theta),T\} - V_{m-1}^+(\theta)].$$

Eq. (4.71) now follows by substituting the above equation into Eq. (4.73).

We next prove Eq. (4.72), using a symmetric proof. Note that

$$\frac{d}{d\theta^-} L_\zeta(T,\theta) \;=\; \frac{1}{T} \sum_{n\in\Psi(T,\theta)} \frac{d}{d\theta^-} \int_{U_n^-(\theta)}^{\min\{V_n^-(\theta),T\}} \zeta(t,\theta)\, dt, \tag{4.75}$$

because the set $\Psi(T,\theta)$ is locally independent of θ in a left neighborhood of θ. For each $n \in \Psi(T,\theta)$, part (c) of Lemma 9 and Eq. (2.12) imply

$$\frac{d}{d\theta^-} \int_{U_n^-(\theta)}^{\min\{V_n^-(\theta),T\}} \zeta(t,\theta)\, dt \;=\; -[\alpha(U_n^-(\theta)) - \mu(U_n^-(\theta),\theta)] \frac{d}{d\theta^-} U_n^-(\theta) \;-\; \int_{U_n^-(\theta)}^{\min\{V_n^-(\theta),T\}} dt. \tag{4.76}$$

We next compute the first term on the right-hand side of Eq. (4.76), and then take its left derivative with respect to θ. To this end, we first compute the integrals $\int_{V_{n-1}^-(\theta)}^{U_n^-(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt$ for two cases of $n \in \Psi(T,\theta)$.

Case 3: $n = 1$. On the event $\{U_1^-(\theta) > 0\}$,

$$\int_{V_0^-(\theta)}^{U_1^-(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; I(V_0^-(\theta)) - I(U_1^-(\theta)) \;=\; I(0),$$

since $I(U_1^-(\theta)) = 0$. Furthermore, on the complementary event $\{U_1^-(\theta) = 0\}$,

$$\int_{V_0^-(\theta)}^{U_1^-(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; 0,$$

since $V_0^-(\theta) = U_1^-(\theta) = 0$.

Case 4: $n > 1$. On the event $\{I(t,\theta) = 0,\; U_{n-1}^-(\theta) < t < V_{n-1}^-(\theta)\}$,

$$\int_{V_{n-1}^-(\theta)}^{U_n^-(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; I(V_{n-1}^-(\theta)) - I(U_n^-(\theta)) \;=\; 0,$$

since $I(V_{n-1}^-(\theta)) = 0$ and $I(U_n^-(\theta)) = 0$. Furthermore, on the complementary event , $\{I(t,\theta) = S \text{ and } \alpha(t) < \mu(t,\theta),\ U_{n-1}^-(\theta) < t < V_{n-1}^-(\theta)\}$,

$$\int_{V_{n-1}^-(\theta)}^{U_n^-(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt \;=\; I(V_{n-1}^-(\theta)) - I(U_n^-(\theta)) \;=\; S,$$

since $I(V_{n-1}^-(\theta)) = S$ and $I(U_n^-(\theta)) = 0$.

It follows that in all cases above, the integral $\int_{V_{n-1}^-(\theta)}^{U_n^-(\theta)} [\alpha(t) - \mu(t,\theta)]\, dt$ equals either $I(0)$ or 0 or S, all of which are independent of θ. Taking the left derivative of this integral with respect to θ yields

$$-[\alpha(V_{n-1}^-(\theta)) - \mu(V_{n-1}^-(\theta),\theta)]\frac{d}{d\theta^-}V_{n-1}^-(\theta) + [\alpha(U_n^-(\theta)) - \mu(U_n^-(\theta),\theta)]\frac{d}{d\theta^-}U_n^-(\theta) - \int_{V_{n-1}^-(\theta)}^{U_n^-(\theta)} dt \;=\; 0.$$

in view of Eq. (2.12). But the first term in the above equation varnishes by part (c) of Lemma 9, resulting in

$$-[\alpha(U_n^-(\theta)) - \mu(U_n^-(\theta),\theta)]\frac{d}{d\theta^-}U_n^-(\theta) \;=\; -\int_{V_{n-1}^-(\theta)}^{U_n^-(\theta)} dt.$$

Finally, substitute the above equation into Eq. (4.76), yielding

$$\frac{d}{d\theta^-}\int_{U_n^-(\theta)}^{\min\{V_n^-(\theta),T\}} \zeta(t,\theta)\, dt \;=\; -\int_{V_{n-1}^-(\theta)}^{\min\{V_n^-(\theta),T\}} dt \;=\; -[\min\{V_n^-(\theta),T\} - V_{n-1}^-(\theta)].$$

Eq. (4.72) now follows by substituting the above equation into Eq. (4.75). □

Let $\leq_{st}$ denote stochastic ordering, and let $=_{st}$ denote stochastic equality [Shaked and Shanthikumar (1994)].

Corollary 2 *In any MTS system with the lost-sales rule, the sided IPA derivatives satisfy the inequality*

$$\frac{d}{d\theta^+}L_\zeta(T,\theta) \;\geq_{st}\; \frac{d}{d\theta^-}L_\zeta(T,\theta), \quad T > 0, \quad \theta \in \Theta. \tag{4.77}$$

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

Proof. Consider the aforementioned sided IPA derivatives in the same probability space. For any $m \in \Phi(T,\theta)$, it follows from the definitions of $\Phi(T,\theta)$ and $\Psi(T,\theta)$ that there exists $n \in \Psi(T,\theta)$ so that $(U_m^+(\theta), V_m^+(\theta)) \subseteq (U_n^-(\theta), V_n^-(\theta))$, and

$$\min\{V_m^+(\theta),T\} \;\leq\; \min\{V_n^-(\theta),T\}. \tag{4.78}$$

Define the following random indices,

$$m^*(m) \;=\; \begin{cases} \max\{1 \leq i < m:\ I(V_i^+(\theta),\theta) = S\}, & \text{if it exists} \\ 0, & \text{otherwise} \end{cases} \tag{4.79}$$

$$n^*(n) \;=\; \begin{cases} \max\{1 \leq j < n:\ I(V_j^-(\theta),\theta) = S\}, & \text{if it exists} \\ 0, & \text{otherwise} \end{cases} \tag{4.80}$$

In another word, $m^*(m)$ is the largest index prior to m such that $I(V^+_{m^*(m)}(\theta), \theta) = S$, while $n^*(n)$ is the largest index prior to n such that $I(V^-_{n^*(n)}(\theta), \theta) = S$

By definition, for each event $\{I(t,\theta) = S,\ U^+_i(\theta) < t < V^+_i(\theta)\}$, there exists $1 \le j \le N_\zeta(T,\theta)$ such that $(U^-_j(\theta), V^-_j(\theta)) \subseteq (U^+_i(\theta), V^+_i(\theta))$. In a similar vein, for each event $\{I(t,\theta) = S,\ U^-_j(\theta) < t < V^-_j(\theta)\}$, there exists $1 \le i \le M_\zeta(T,\theta)$ such that $(U^-_j(\theta), V^-_j(\theta)) \subseteq (U^+_i(\theta), V^+_i(\theta))$. It follows that $(U^-_{n^*(n)}(\theta), V^-_{n^*(n)}(\theta)) \subseteq (U^+_{m^*(m)}(\theta), V^+_{m^*(m)}(\theta))$ and

$$V^+_{m^*(m)}(\theta) \ \ge \ V^-_{n^*(n)}(\theta). \tag{4.81}$$

Combining this inequality with (4.78) yields,

$$\min\{V^+_m, T\} - V^+_{m^*(m)}(\theta) \ \le \ \min\{V^-_n, T\} - V^-_{n^*(n)}(\theta). \tag{4.82}$$

Since (4.82) holds for every $m \in \Phi(T,\theta)$, the requisite inequality (4.77) now follows from Theorem 5.

Finally, if $I(t,\theta) = S$ implies $\alpha(t) < \mu(t,\theta)$ and $I(t,\theta) = 0$ implies $\alpha(t) > \mu(t,\theta)$, then $M_\zeta(T,\theta) = N_\zeta(T,\theta)$, and the sequence of intervals $(U^+_m(\theta), V^+_m(\theta))$, $m = 1, \ldots, M_\zeta(T,\theta)$, is identical to the sequence of intervals $(U^-_n(\theta), V^-_n(\theta))$, $n = 1, \ldots, N_\zeta(T,\theta)$. The requisite result now follows immediately from Theorem 5. □

Finally, we show that the IPA derivatives of Theorem 4 and Theorem 5 are unbiased.

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

Proof. Theorems 4 and 5 ensure that for all $T > 0$, Condition (a) of Fact 1 is satisfied for both $L_I(T,\theta)$ and $L_\zeta(T,\theta)$. For any $\theta_1, \theta_2 \in \Theta$,

$$\begin{aligned} |L_I(T,\theta_1) - L_I(T,\theta_2)| \ &= \ \left| \frac{1}{T} \int_0^T [I(t,\theta_1) - I(t,\theta_2]\, dt \right| \\ &\le \ \frac{1}{T} \int_0^T |I(t,\theta_1) - I(t,\theta_2)|\, dt \ \le \ T|\theta_1 - \theta_2|, \end{aligned} \tag{4.83}$$

where the second inequality is a consequence of Corollary 1. Furthermore, by Proposition 3,

$$|L_\zeta(T,\theta_1) - L_\zeta(T,\theta_2)| \ \le \ 2|\theta_1 - \theta_2|. \tag{4.84}$$

Eqs. (4.83) and (4.84) establish that Condition (b) of Fact 1 holds for both $L_I(T,\theta)$ and $L_\zeta(T,\theta)$, thereby completing their proof of unbiasedness. □

5 Discussion

This paper formulates Make-to-Stock (MTS) production-inventory systems with lost sales in stochastic fluid model (SFM) setting, and derives IPA derivatives of time averages of inventory level and lost sales with respect to the base-stock level and a production rate parameter. These IPA derivative formulas are comprehensive because they are derived for any initial inventory state and without the assumption that the left and right derivatives must coincide. All IPA derivatives obtained are shown to be unbiased, nonparametric and fast to compute, which holds out the promise of broad applications to on-line control of MTS production-inventory systems.

This paper, together with Zhao and Melamed (2005) can provide a theoretical basis for new on-line control algorithms of production-inventory systems, including those where the underlying stochastic processes (e.g., the demand and production capacity processes) may be subject to non-stationary probability laws. One direction of future research is the extension of the current results to more general supply networks with multiple facilties and multiple products, such as multi-stage assembly systems. A case in point is Assemble-to-Order systems (such as those implemented by Dell Computer Corporation), which face high demand volumes and are still required to provide high service levels within tight committed service times (Perman 2001). An important characteristic of such systems is that their demand patterns fluctuate considerably over time.

Clearly, a key issue is how to control production speeds at workstations and the base-stock levels of component inventories so that the production-inventory system may quickly adapt to changing demand without sacrificing service levels. While off-line algorithms can determine the optimal control parameters for systems in steady state, IPA-based on-line control algorithms can sample the system's transient state and compute the requisite (nonparametric) transient IPA derivatives, and then use them to quickly predict system metrics under hypothetically changed parameters of interest. This will be the subject of future work to be reported elsewhere.

Acknowledgments

This research was done under the auspices of the Rutgers Center of Supply Chain Management at the Rutgers Business School - Newark and New Brunswick.

References

[1] Bashyam, S. and M. C. Fu (1994) "Application of perturbation analysis to a class of periodic review (s, S) inventory systems" *Naval Research Logistics*, **41**, 47-80.

[2] Buzacott, J., S. Price and J. Shanthikumar (1991) "Service level in multi-stage MRP and base-stock controlled production systems", in *New Directions for Operations Research in Manufacturing* (G. Fandel, T. Gulledge, A. Hones, eds.), 445-463. Springer, Berlin, Germany.

[3] Cassandras, C.G., Y. Wardi, B. Melamed, G. Sun and C.G. Panayiotou (2002) "Perturbation Analysis for On-Line Control and Optimization of Stochastic Fluid Models", *IEEE Trans. On Automatic Control*, Vol. AC-47, No. 8, 1234–1248.

[4] Fu, M. C. (1994a) "Optimization via simulation: A review", *Annals of Operations Research*, **53**, 199-247.

[5] Fu, M. C. (1994b) "Sample path derivatives for (s, S) inventory systems", *Operations Research*, **42**, 351-364.

[6] Glasserman P. (1991) *Gradient estimation via perturbation analysis*, Kluwer Academic Publishers, Boston, MA.

[7] Glasserman P. and S. Tayur (1995) "Sensitivity analysis for base-stock levels in multiechelon production-inventory systems", *Management Science* **41**, 263-281.

[8] Heidelberger, H., X.R. Cao, M. Zazanis, and R. Suri (1988) "Convergence properties of infinitesimal analysis estimates", *Management Science* **34**, 1281-1302.

[9] Ho, Y. C. and X. R. Cao (1991) *Perturbation Analysis of Discrete Event Simulation*, Kluwer Academic Publishers, Boston, MA.

[10] Karmarkar, U. (1987) "Lot sizes, lead-times and in-process inventories", *Management Science* **33**, 409-419.

[11] Paschalidis, I., Y. Liu, C.G. Cassandras, and C. Panayiotou (2004) "Inventory Control for Supply Chains with Service Level Constraints: A Synergy between Large Deviations and Perturbation Analysis", *Annals of Operations Research* **126**, 231–258.

[12] Perman, S. (2001) "Automate or die", *Business 2.0*, July.

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

[14] Veatch, M. H. (2002) "Using fluid solutions in dynamic scheduling", in *Analysis and Modeling of Manufacturing Systems* (S. B. Gershwin, Y. Dallery, C. T. Papadopoulos, J. M. Smith, eds.), 399-426, Kluwer, New York.

[15] Wardi, Y., B. Melamed (2001) "Variational Bounds and Sensitivity Analysis of Continuous Flow Models", *J. of Discrete Event Dynamic Systems*, **11(3)**, 249-282.

[16] Wardi, Y., B. Melamed, C. G. Cassandras & C. G. Panayiotou (2002) "On-line IPA gradient estimators in stochastic continuous fluid models", *J. of Optimization Theory and Applications* **115(2)**, 369-405.

[17] Wein, L. M. (1992) "Dynamic scheduling of a multiclass Make-to-Stock queue", *Operations Research*, **40**, 724-735.

[18] Zhao, Y. and B. Melamed (2004) "Make-to-Stock Systems With Backorders: IPA Gradients", Winter Simulation Conference (WSC 04), Washington D.C., December 5 - 8, 2004.

[19] Zhao, Y. and B. Melamed (2005) "IPA Derivatives for Make-to-stock Production-Inventory Systems with Backorders", Working Paper, Department of Management Science and Information Systems, Rutgers University.

[20] Zipkin, P. (1986) "Models for design and control of stochastic, multi-item batch production systems", *Operations Research* **34**, 91-104.