Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Partially Observed Inventory Systems

Alain Bensoussan, Metin Çakanyıldırım and Suresh P. Sethi

***Abstract*— In some inventory control contexts, such as Vendor Managed Inventories, inventory with spoilage, misplacement, or theft, inventory levels may not always be observable to the decision makers. However, when shortages occur, inventory levels receive more attention and they may become completely observed. We study such an inventory control context where the unmet demand is lost and orders must be decided on the basis of partial information to minimize the total discounted costs over an infinite horizon. This problem has an infinite-dimensional state space, and for it we establish the existence of a feedback policy when one period costs are bounded or when the discount factor is sufficiently small.**

I. Introduction

Inventory control is among the most important topics in operations research. One of the critical assumptions in the vast inventory literature, dating back to at least the Harris lot size model of 1913 [7], has been that the level of inventory at any given time is fully observed. Some of the most celebrated results, such as the optimality of the base-stock policy, have been obtained under the assumption of full observation. Yet the inventory level is never fully observed in practice. In such an environment, most of the well-known inventory policies are not only not optimal, but are also not applicable. A main reason for why the analysis of inventory problems under partial observations has been neglected lies in its mathematical difficulty. Whereas one works with a finite dimensional state space in the full observation case, one usually has to deal with an infinite dimensional state space in the partial observation setting. More specifically, the inventory level at a given time is no longer a system state in $\Re^n$, it must now be represented by its conditional probability given some limited information available at that time. Thus, the analysis takes place in the space of probability distributions. This is, of course, inevitable, and simplifies only in particular situations, when for instance the separation principle applies; see [1] for example.

Concerning controls of dynamic systems in general, a great step forward was achieved in the applied mathematics and engineering control literature, when the Zakai equation [9] was discovered. Prior to that, the evolution of the conditional probability had been studied with the highly nonlinear Kushner equation [8]. The Zakai equation uses a transformation that changes the Kushner equation into a pair of linear equations. This transformation corresponds to the concept of "change of measure" [6]. While it does not remove the infinite dimensionality, the linearity has permitted a number of important control problems with partial observations to be solved. Of course, there remain numerical difficulties due to the infinite dimensionality of the state. Nevertheless, a sound theory is available.

The key idea in going from the Kushner equation to the Zakai equation is in introducing unnormalized conditional probabilities in place of conditional probabilities. This linearizes the state equation, and the problem becomes much simpler to study. Ideas of this kind have not been introduced yet in the context of solving partial observation control problems in management. While the standard Zakai setup cannot be directly applied to inventory problems, we show that unnormalized conditional probabilities can be introduced and are indeed quite appropriate.

II. The Zero Balance Walk Model

We study a periodic review inventory problem with partially observable inventory levels. In our model, the inventory levels are not automatically observed by the Inventory Manager (IM) who decides on order quantities. We first construct a finite horizon model with T periods. The order of events in any given period t is as follows: The IM observes the event when the inventory level falls to zero, but he does not observe the inventory level when it is positive. The manager determines how much to order and the order is delivered instantaneously. Next the customer demand occurs, but it is not observed by the IM unless the inventory level drops to zero. In each period, the IM incurs inventory related costs, but he does not observe these costs immediately. Lastly the state defining the inventory level is updated for the next period.

In classical inventory settings, the inventory level I_t at the beginning of period t is observed, and is used to determine the order quantity q_t in period t. Each period t has a random demand D_t defined on the probability space $(\Omega, \mathcal{F}, P)$. The demand is met, to the extent possible, from the on-hand stock $I_t + q_t$. We suppose that the demand that is not immediately met from the on-hand stock is lost. Then the evolution of inventory dynamics is given as follows:

$$I_{t+1} = (I_t + q_t - D_t)^+ \quad \text{for } t \geq 1. \tag{1}$$

We assume demand D_t to be i.i.d. A generic demand is denoted by D, which is i.i.d. with each D_t. Let f denote the density and F denote the cumulative distribution of D. Let $\bar{F} = 1 - F$.

When the demand is met entirely, inventory holding costs apply to the remaining inventory. Otherwise, there are lost sales costs. It is well known that the base stock policy is optimal for this setting. It is interesting to investigate the

All three authors are with the School of Management, University of Texas at Dallas, Richardson, TX 75083-0688. {alain.bensoussan, metin, sethi}@utdallas.edu

validity of the optimality of the base stock policy, or lack of it, for the zero-balance walk model.

In the zero-balance walk model, the inventory levels are partially observed by the IM as follows.

$$I_1 \text{ is either 0 or its distribution is known.} \tag{2}$$

In general, the IM does not observe the demand or the inventory level. However, looking at empty shelves and concluding $I_t = 0$ does not take much effort, and constitutes a free observation. Thus, we allow I_t to be observed only when the inventory shelves are empty (i.e. $[I_t = 0]$). To study such partial observations of the inventory level, we introduce a signal (message) random variable

$$z_t := \mathbb{1}_{I_t=0}, \quad t \geq 0. \tag{3}$$

The signal z_t is a discrete-time Markov Chain with the state space $\{0, 1\}$: 1 means an empty shelf and 0 means a nonempty shelf.

When the inventory levels are fully observed, the order q_t is adapted to the sigma field $\mathcal{F}_t := \sigma(\{I_j : 1 \leq j \leq t\})$ generated by the inventory levels observed by period t. Note that the demand observations up to the beginning of period t also generate the same field, i.e., $\mathcal{F}_t = \sigma(\{I_1, D_j : 1 \leq j \leq t-1\})$. With our partial observations model, q_t is adapted to $\mathcal{Z}_t := \sigma(\{z_j : 1 \leq j \leq t\})$. Clearly $\mathcal{Z}_t \subset \mathcal{F}_t$, so our partial observations model must decide on order quantities on the basis of less than full information.

Given a stationary cost function $c(I_t, q_t)$ that depends on the beginning inventory level I_t and the order size q_t in period t, and with $\tilde{q}$ defining the admissible sequence of actions $\tilde{q} = \{q_1, q_2, \ldots\}$, the total discounted cost is defined by

$$J(\zeta, \pi, \tilde{q}) := \mathrm{E} \sum_{t=1}^{T} \alpha^t c(I_t, q_t), \tag{4}$$

where $\alpha < 1$ is the discount factor. The initial conditions are a pair $(\zeta, \pi(x))$, where ζ is 1 or 0. If ζ is 1, then $I_1 = 0$. If ζ is 0, then $I_1 > 0$ and $\pi(\cdot)$ is the probability distribution of I_1. We look for q_t, adapted to $\mathcal{Z}_t$, $t \geq 0$, to minimize $J(\zeta, \pi, \tilde{q})$.

A. Evolution of State Probabilities

We now develop the conditional probability density $\pi_t(.)$ of I_t given $\mathcal{Z}_{t-1}$ and $I_t > 0$. By definition,

$$\int_0^x \pi_t(y)dy = \mathrm{P}(I_t \leq x | \mathcal{Z}_{t-1}, I_t > 0).$$

Since the event $[I_t = 0]$ is observable, conditional probabilities are needed only when $I_t > 0$.

For any real and bounded test function $\varphi(.)$, we can use the conditional Bayes theorem (e.g. [6]) to obtain

$$\begin{aligned} \int_0^\infty \varphi(x)\pi_t(x)dx &= \mathrm{E}[\varphi(I_t)|\mathcal{Z}_{t-1}, I_t > 0] \\ &= \frac{\mathrm{E}[\varphi(I_t)\mathbb{1}_{I_t>0}|\mathcal{Z}_{t-1}]}{\mathrm{E}[\mathbb{1}_{I_t>0}|\mathcal{Z}_{t-1}]} = \frac{\mathrm{E}[\varphi(I_t)\mathbb{1}_{I_t>0}|\mathcal{Z}_{t-1}]}{\mathrm{P}(I_t > 0|\mathcal{Z}_{t-1})}. \end{aligned} \tag{5}$$

In order to obtain a recursive expression for π_t in terms of π_{t-1}, we begin with expressing $\mathrm{E}(\varphi(I_t)|\mathcal{Z}_t)$ in terms of conditional expectations with respect to $\mathcal{Z}_{t-1}$ in the next lemma.

Lemma 1.

$$\begin{aligned} E(\varphi(I_t)|\mathcal{Z}_t) &= \mathbb{1}_{I_t=0}\varphi(0) + \mathbb{1}_{I_t>0}\frac{E(\varphi(I_t)\mathbb{1}_{I_t>0}|\mathcal{Z}_{t-1})}{P(I_t > 0|\mathcal{Z}_{t-1})} \\ &= \mathbb{1}_{I_t=0}\varphi(0) + \mathbb{1}_{I_t>0}E(\varphi(I_t)|\mathcal{Z}_{t-1}, I_t > 0) \end{aligned} \tag{6}$$

Instead of the conditional expectations in Lemma 1, the left-hand side in (6) can also be expressed by using the conditional density function π_t. Using (5) on the right-hand side of (6) gives

$$\mathrm{E}(\varphi(I_t)|\mathcal{Z}_t) = \mathbb{1}_{I_t=0}\varphi(0) + \mathbb{1}_{I_t>0}\int_0^\infty \varphi(z)\pi_t(z)dz. \tag{7}$$

The density π_t is obtained by setting (6) and (7) to be equal. For $I_t = 0$, this equality yields $\pi_t = \delta$ which is the Dirac delta function taking the value of zero everywhere except at 0 where it is infinite. For the more interesting case of $I_t > 0$, the next lemma molds (6) into a convenient form to set (7) equal to (6) and solve for π_t.

Lemma 2.

$$\begin{aligned} &E(\varphi(I_t)|\mathcal{Z}_t)\mathbb{1}_{I_t>0} = \\ &\mathbb{1}_{I_{t-1}=0}\frac{\int_0^\infty \varphi(z)f(q_{t-1}-z)\mathbb{1}_{q_{t-1}\geq z}dz}{F(q_{t-1})} + \mathbb{1}_{I_{t-1}>0} \times \\ &\quad \frac{\int_0^\infty \varphi(z)\int_{(z-q_{t-1})^+}^\infty f(y+q_{t-1}-z)\pi_{t-1}(y)dydz}{\int_0^\infty F(y+q_{t-1})\pi_{t-1}(y)dy}. \end{aligned} \tag{8}$$

Having obtained the conditional expectation in Lemma 2, we go back to the conditional probability π_t as defined in (7) for $I_t > 0$. Setting the second term on the right-hand side of (7) equal to (8),

$$\begin{aligned} \pi_t(x) = \mathbb{1}_{I_{t-1}=0}&\left\{\frac{f(q_{t-1}-x)\mathbb{1}_{x\leq q_{t-1}}}{F(q_{t-1})}\right\} + \mathbb{1}_{I_{t-1}>0} \times \\ &\left\{\frac{\int_{(x-q_{t-1})^+}^\infty f(y+q_{t-1}-x)\pi_{t-1}(y)dy}{\int_0^\infty F(y+q_{t-1})\pi_{t-1}(y)dy}\right\}. \end{aligned} \tag{9}$$

This expression specializes to the conditional probabilities stated in the next theorem.

Theorem 1. *The conditional probability π_t can be expressed recursively as follows:*

$$\pi_t(x) = \begin{cases} \mathbb{1}_{x\leq q_{t-1}}\dfrac{f(q_{t-1}-x)}{F(q_{t-1})} & \text{if } I_{t-1} = 0 \\ \dfrac{\int_{(x-q_{t-1})^+}^\infty \pi_{t-1}(y)f(y+q_{t-1}-x)dy}{\int_0^\infty \pi_{t-1}(y)F(y+q_{t-1})dy} & \text{if } I_{t-1} > 0 \end{cases}$$

The conditional probability evolves according to a highly nonlinear equation

$$\pi_t(x) = z_{t-1}\frac{f(q_{t-1}-x)\mathbb{I}_{x<q_{t-1}}}{F(q_{t-1})} + (1-z_{t-1}) \times \frac{\int_{(x-q_{t-1})^+}^{\infty} f(y+q_{t-1}-x)\pi_{t-1}(y)dy}{\int_0^{\infty} F(q_{t-1}+y)\pi_{t-1}(y)dy} \qquad t \geq 2, \quad (10)$$

$$\pi_1(x) = \pi(x),$$

which corresponds to the Kushner equation [8] in our inventory context.

We can linearize (10) as follows. Set

$$p_t(x) := \lambda_t \pi_t(x), \tag{11}$$

where λ_t is a weighting factor to be defined shortly. On account of this weighting, $p_t(x)$ can be viewed as unnormalized probability. Furthermore, it evolves according to the linear equation

$$\begin{aligned} p_t(x) &= z_{t-1}f(q_{t-1}-x)\mathbb{I}_{x<q_{t-1}} + (1-z_{t-1}) \times \\ &\qquad \int_{(x-q_{t-1})^+}^{\infty} f(y+q_{t-1}-x)p_{t-1}(y)dy \\ p_1(x) &= \pi(x). \end{aligned} \tag{12}$$

This equation corresponds to the Zakai equation for systems with diffusions in [9] and [1]. By integrating both sides of (11),

$$\begin{aligned} \lambda_t &= \int_0^{\infty} p_t(x)dx \\ &\overset{(12)}{=} z_{t-1}F(q_{t-1}) \\ &\quad +(1-z_{t-1})\int_0^{\infty} F(q_{t-1}+y)p_{t-1}(y)dy \\ &\overset{(11)}{=} z_{t-1}F(q_{t-1}) \\ &\quad +(1-z_{t-1})\lambda_{t-1}\int_0^{\infty} F(q_{t-1}+y)\pi_{t-1}(y)dy. \end{aligned}$$

The last equation defines λ_t recursively starting with $\lambda_1 = 1$. However, note that λ_t depends on π_{t-1} on the right-hand side. The normalized probabilities can easily be computed from the unnormalized probabilities as follows:

$$\pi_t(x) = \frac{p_t(x)}{\int_0^{\infty} p_t(x)dx}. \tag{13}$$

These equations can be written in the operator form in the space

$$\mathcal{H} := \left\{ p \in L^1(\Re^+) \ : \ \int_0^{\infty} x|p(x)|dx < \infty \right\},$$

where $L^1(\Re^+)$ is the space of integrable functions whose domain is the set of nonnegative real numbers. If we define regular addition and multiplication by a scalar on $\mathcal{H}$ and include negative valued functions in $\mathcal{H}$, then $\mathcal{H}$ becomes a subspace of $L^1(\Re^+)$. Working with the subspace $\mathcal{H}$ is convenient for some of our arguments. However, we are ultimately interested in unnormalized probabilities, which are nonnegative. Let us equip the subspace $\mathcal{H}$ with the norm

$$||p|| = \int_0^{\infty} |p(x)|dx + \int_0^{\infty} x|p(x)|dx. \tag{14}$$

The dual space of $\mathcal{H}$ is denoted by $\mathcal{H}_*$, and it is the space of functions ϕ with linear growth, i.e.,

$$\mathcal{H}_* = \left\{ \phi \ : \ \sup_{x>0} \frac{|\phi(x)|}{1+x} < \infty \right\}.$$

Furthermore, we have the duality product

$$\langle p, \phi \rangle = \int_0^{\infty} p(x)\phi(x)dx \quad \text{ for } p \in \mathcal{H},\ \phi \in \mathcal{H}_*.$$

For any scalar $q > 0$, define the linear operator ρ from $\mathcal{H}$ to $\mathcal{H}$ as

$$\rho(q,p)(x) = \int_{(x-q)^+}^{\infty} f(y+q-x)p(y)dy.$$

Note that $\rho(q,\delta)(x) = f(q-x)\mathbb{I}_{x<q}$ so $\rho(0,\delta)(x) = 0$ for the Dirac delta function δ. Define the nonlinear operator θ as

$$\theta(q,p) = \frac{\rho(q,p)}{\langle \rho(q,p), 1 \rangle} \ . \tag{15}$$

With these notations, we can write (10) and (12) in the operator form:

$$\pi_t = z_{t-1}\theta(q_t,\delta) + (1-z_{t-1})\theta(q_t,\pi_{t-1}) \ , \tag{16}$$

$$p_t = z_{t-1}\rho(q_t,\delta) + (1-z_{t-1})\rho(q_t,p_{t-1}) \ , \tag{17}$$

with the initial conditions

$$\pi_1 = p_1 = \pi. \tag{18}$$

Once again, we emphasize that (17) is a linear equation, while (16) is nonlinear.

B. The Bellman Equation

We write $p_t(\tilde{q})$ and $\pi_t(\tilde{q})$ to emphasize the dependence of the states p_t or π_t on the control policy. We assume that $c(I_t, q_t)$ has linear growth in I_t for every fixed q_t, i.e., $c(., q_t) \in \mathcal{H}_*$. The cost function can be written as follows:

$$\begin{aligned} J(\zeta,\pi,\tilde{q}) &= \sum_{t=1}^{\infty} \alpha^t \mathrm{E}[\mathrm{E}[c(I_t,q_t)|\mathcal{Z}_t]] \\ &= \sum_{t=1}^{\infty} \alpha^t \mathrm{E}\{z_t c(0,q_t) \\ &\quad +(1-z_t)\langle c(I_t,q_t), \pi_t(\tilde{q})\rangle\} \end{aligned}$$

where $\pi_t(\tilde{q})$ is the solution of (10). Recall that the initial conditions $\zeta_1 = \zeta \in \{0,1\}$ and $\pi_1 = \pi$ are given. In the sequel, we study only the discounted infinite horizon costs, so the time index t is suppressed. We define the value function

$$V(\zeta,\pi) := \inf_{\tilde{q}} J(\zeta,\pi,\tilde{q}).$$

If we write $v := V(1,\pi)$ which, in fact, is not dependent on π, and $V(\pi) := V(0,\pi)$, then we obtain the following system:

$$\begin{aligned} V(\pi) &= \inf_q \left\{ \langle c(.,q), \pi(.)\rangle \ + \alpha v \int_0^\infty \bar{F}(y+q)\pi(y)dy \right. \\ &\qquad \left. + \alpha V(\theta(q,\pi)) \int_0^\infty F(y+q)\pi(y)dy \right\}, \quad (19) \\ v &= \inf_q \left\{ c(0,q) + \alpha v \bar{F}(q) + \alpha V(\theta(q,\delta)) F(q) \right\} \quad (20) \end{aligned}$$

All the remaining integrals in this paper are over $[0,\infty)$ so we remove the limits from here on.

A direct study of the system in (19)-(20) is not very easy. The matters simplify considerably when working with the unnormalized probability $p \in \mathcal{H}^+$. The unnormalized probability evolves in accordance with the linear operator ρ. To make ideas concrete, we define a new value function $Z(.)$ as follows:

$$Z(p) := V\left(\frac{p}{\lambda}\right)\lambda\,, \quad \lambda := \int p(x)dx.$$

Eventually, we obtain the following new system of equations:

$$\begin{aligned} Z(p) &= \inf_q \left\{ \langle c(.,q), p(.)\rangle \ + \alpha v \int \bar{F}(y+q)p(y)dy \right. \\ &\qquad \left. + \alpha Z(\rho(q,p)) \right\}, \quad (21) \\ v &= \inf_q \left\{ c(0,q) + \alpha v \bar{F}(q) + \alpha Z(\rho(q,\delta)) \right\}. \quad (22) \end{aligned}$$

The pair $(v, Z(p))$ is the solution of (21)-(22). Moreover,

$$Z(\mu p) = \mu Z(p) \quad \text{for every } \mu > 0.$$

Thus, $Z(0) = 0$.

Unlike the operator θ, ρ is a linear operator. Thus, it is easier to study the system in (21)-(22) than that in (19)-(20). The linearity facilitates our arguments dealing with the existence of an optimal feedback control and our discussion when it is finite. Furthermore, it helps in studying finite approximations of the infinite dimensional state space as well as in building associated approximate solutions to (21)-(22).

III. Existence of a Solution to the Bellman Equation

A. Bounded Costs

For the existence results, we bound the single period cost. Suppose that positive constants c, c_0, c_1, and h are such that

$$cq < c(x,q) \le c_0 + c_1 q + hx \quad \text{for } x \ge 0, \tag{23}$$

where c_0 can be interpreted as the maximum expected lost sales cost that can be incurred in a period. Indeed, we set $c_0 = c(0,0)$. Since there will be a positive order q when there is a stock out, c_0 will be bounded by the cost of losing the sale of E(D) units. Let $a_0 := \max\{c_0/(1-\alpha), h\}$.

To accommodate our unnormalized conditional probabilities, we define the functional space

$$\mathcal{B} := \left\{ \phi(p) : \mathcal{H}^+ \to \Re : \sup_{p\in\mathcal{H}^+} \frac{|\phi(p)|}{||p||} < \infty \right\} \tag{24}$$

equipped with the norm

$$||\phi||_{\mathcal{B}} := \sup_{p\in\mathcal{H}^+} \frac{|\phi(p)|}{||p||}, \tag{25}$$

where $||p||$ still refers to the norm that we initially defined in $\mathcal{H} \supseteq \mathcal{H}^+$. For any $\phi \in \mathcal{B}$, we must have $\phi(0) = 0$. $\mathcal{B}$ includes all Z functions that solve (21)-(22). This fact is formalized in the next lemma.

Lemma 3. *Each Z that solves* (21)-(22) *is in $\mathcal{B}$.*

We need some short-hand notation. Define the function $K : \Re \times \mathcal{H} \to \Re$ as

$$\begin{aligned} K(q,p;v,Z) &:= \langle c(.,q), p(.)\rangle + \alpha v \int \bar{F}(y+q)p(y)dy \\ &\qquad + \alpha Z(\rho(q,p)). \end{aligned}$$

For $p = \delta$, $K(q,\delta;v,Z) = c(0,q) + \alpha v \bar{F}(q) + \alpha Z(\rho(q,\delta))$. Define the map $T : \Re \times \mathcal{B} \to \Re \times \mathcal{B}$ as

$$T\begin{pmatrix} v \\ Z(p) \end{pmatrix} := \begin{pmatrix} \inf_q K(q,\delta;v,Z) \\ \inf_q K(q,p;v,Z) \end{pmatrix}. \tag{26}$$

Define $Z_0(p)$ as the value function that solves the Bellman equations when $q = 0$. Then, it must solve

$$\begin{aligned} \langle c(.,0), p(.)\rangle + \alpha v_0 \int \bar{F}(y)p(y)dy + \alpha Z_0(\rho(0,p)) \\ = Z_0(p). \quad (27) \end{aligned}$$

Also define $v_0 := Z_0(p = \delta)$. By (22) and $Z(0) = 0$, $v_0 = c_0 + \alpha v_0$. Existence of $Z_0(p)$ is established with the next lemma.

Lemma 4. *Z_0 exists and is uniquely defined.*

If $v \le v_0$, $Z(p) \le Z_0(p)$ and

$$\begin{pmatrix} \tilde{v} \\ \tilde{Z}(p) \end{pmatrix} := T\begin{pmatrix} v \\ Z(p) \end{pmatrix},$$

then $\tilde{v} \le c(0,0) + \alpha v \le c_0 + \alpha v_0 = v_0$. Also

$$T\begin{pmatrix} v \\ Z(p) \end{pmatrix} \le \begin{pmatrix} v \\ Z(p) \end{pmatrix}. \tag{28}$$

This inspires the next result, where we prove the existence of a solution of the system (21)- (22) by using a value iteration scheme. The solution is named as $(\bar{v}, \bar{Z})$.

Theorem 2. *A solution of* (21)-(22) *exists.*

Proof: Let

$$\begin{pmatrix} v_{n+1} \\ Z_{n+1}(p) \end{pmatrix} := T\begin{pmatrix} v_n \\ Z_n(p) \end{pmatrix}.$$

Starting with v_0 and Z_0 defined by (27), we first claim that

$$\begin{pmatrix} v_{n+1} \\ Z_{n+1}(p) \end{pmatrix} \le \begin{pmatrix} v_n \\ Z_n(p) \end{pmatrix}.$$

Since $(v_n,\ Z_n)$ is a nonincreasing sequence with a lower bound of $(0,0)$, it has a limit $(\bar{v},\ \bar{Z})$: $(v_n,\ Z_n) \downarrow (\bar{v},\ \bar{Z})$, and we establish that $(\bar{v}, \bar{Z})$ is a fixed point. The fixed point is not necessarily unique, but it is the maximum solution in the following sense. Any solution (v, Z) that satisfies $(v, Z) \leq (v_0, Z_0)$ also satisfies $(v, Z) \leq (\bar{v}, \bar{Z})$. $\Box$

Bounding the optimal order quantity: Start by setting $q = 0$ in (21) we can check

$$q \leq \frac{a_0}{c(1-\alpha)} \left\{ 1 + \frac{\int xp(x)dx}{\int p(x)dx} \right\}. \tag{29}$$

Note that the bound depends on the unnormalized probability p and can be arbitrarily large as $p \to 0$. Because of this observation, we choose to assume a bound on the order quantity in the next subsection.

B. Bounded Order Quantities

In this section we assume that there is a finite bound on the order quantity q in addition to the cost bounds in the previous section. The finite bound can be due to the supplier's limited production or transportation capacity, or the storage capacity IM can use. Let the capacity be m and let the corresponding Z and v be denoted by Z^m and v^m. Then (21)-(22) is written as

$$\begin{aligned} Z^m(p) &= \inf_{q \leq m} \Big\{ \langle c(.,q), p(.) \rangle \\ &\quad + \alpha v^m \int \bar{F}(y+q)p(y)dy + \alpha Z^m(\rho(q,p)) \Big\} \\ v^m &= \inf_{q \leq m} \left\{ c(0,q) + \alpha v^m \bar{F}(q) + \alpha Z^m(\rho(q,\delta)) \right\}. \end{aligned} \tag{30}$$

We can check that constants A^m and B^m exist such that

$$\begin{aligned} |Z^m(p) - Z^m(p')| &\leq A^m \int |p(y) - p'(y)|dy \\ &\quad + B^m \int y|p(y) - p'(y)|dy, \end{aligned} \tag{31}$$

for any two $p, p' \in \mathcal{H}$. Therefore, Z^m is Lipschitz continuous on $\mathcal{H}$. This additional smoothness property allows us to establish the uniqueness of a solution to the system in (30)

We next establish that $(\bar{v}^m, \bar{Z}^m)$ converges to $(\bar{v}, \bar{Z})$, which is the maximal solution of the system in (21)-(22).

So far, we studied the existence and the convergence of (v^m, Z^m). The next theorem validates the monotone iterative process, that is (v^m, Z^m) minimizes the total discounted cost. As a side product of the theorem, v^m and Z^m turn out to be unique because they are equal to the minimum costs, which are unique by definition.

Theorem 3. *The solution (v^m, Z^m) of (30) is the minimum total discounted cost, i.e.,*

$$\begin{aligned} Z^m(\pi) &= \inf_{\tilde{q}: q_t \leq m} J(0, \pi, \tilde{q}), \\ v^m &= \inf_{\tilde{q}: q_t \leq m} J(1, \delta, \tilde{q}). \end{aligned}$$

Since $Z^m(\pi)$ and v^m are defined as a solution of (30) and they are given by the infima in Theorem 3, both $Z^m(\pi)$ and v^m are unique. As m increases, we have

$$\inf_{\tilde{q}: q_t \leq m} J(0, \pi, \tilde{q}) \downarrow \inf_{\tilde{q}} J(0, \pi, \tilde{q}),$$

$$\inf_{\tilde{q}: q_t \leq m} J(1, \pi, \tilde{q}) \downarrow \inf_{\tilde{q}} J(1, \pi, \tilde{q}).$$

These imply

$$Z(\pi) = \inf_{\tilde{q}} J(0, \pi, \tilde{q}),$$

$$v = \inf_{\tilde{q}} J(1, \pi, \tilde{q}).$$

Thus, $Z(\pi)$ and v are interpreted as the infima of the costs even when m disappears. However, a corresponding feedback solution that yields $Z(\pi)$ and v may not exist unless m is finite.

In this section, we establish that $Z^m(p)$ is continuous and converges to $Z(p)$. However, these results do not guarantee the continuity of $Z(p)$. Instead we have the weaker form of continuity as presented in the next lemma.

Lemma 5. *$Z(p)$ is upper semicontinuous:*

$$\limsup_{p_k \to p} Z(p_k) \leq Z(p).$$

IV. A Sufficiently Small Discount Rate

We argue that T is a contraction map for a sufficiently small α. Namely, we let $M := 1 + a_0/(c(1-\alpha))$ and require that $\alpha(1+M) < 1$. In this case, a solution to the system (21)-(22) exists.

Consider the difference

$$\begin{aligned} K(q,p;v,Z) - K(q,p;v',Z') &= \\ &\alpha(v - v') \int \bar{F}(y+q)p(y)dy \\ &+ \alpha[Z(\rho(q,p)) - Z'(\rho(q,p))]. \end{aligned}$$

Let

$$\eta := \max \left\{ \sup_{p \in \mathcal{H}^+} \frac{Z(p) - Z'(p)}{||p||},\ |v - v'| \right\}.$$

we obtain

$$|K(q,p;v,Z) - K(q,p;v',Z')| \leq \eta\alpha(1+M)||p||.$$

Hence, it follows that

$$|\inf_q K(q,p;v,Z) - \inf_q K(q,p;v',Z')| \leq \eta\alpha(1+M)||p||.$$

Recalling $||\delta|| = 1$ and specializing to $p = \delta$, we have

$$|\inf_q K(q,\delta;v,Z) - \inf_q K(q,\delta;v',Z')| \leq \eta\alpha(1+M).$$

We can consider η as a distance d in the space $\Re \times \mathcal{B}$. That is,

$$\eta = d\left(\begin{pmatrix} v \\ Z(p) \end{pmatrix}, \begin{pmatrix} v' \\ Z'(p) \end{pmatrix} \right).$$

In summary, we have proved

$$\begin{aligned} d\left(T\begin{pmatrix} v \\ Z(p) \end{pmatrix}, T\begin{pmatrix} v' \\ Z'(p) \end{pmatrix} \right) &\leq \\ \alpha(1+M) &\left(\begin{pmatrix} v \\ Z(p) \end{pmatrix}, \begin{pmatrix} v' \\ Z'(p) \end{pmatrix} \right). \end{aligned}$$

In addition, if $\alpha(1+M) < 1$ as required at the beginning of this section, T is a contraction map on $\Re \times \mathcal{B}$. It is also a contraction on $\Re \times \mathcal{B}_c$, where $\mathcal{B}_c$ denotes the closed subset containing the continuous functions in $\mathcal{B}$. Therefore, when $\alpha(1+m) < 1$, (21)-(22) have a unique fixed point in $\Re\times\mathcal{B}_c$.

A. Relaxing the Condition on the Discount Rate

We relax the condition $\alpha(1+M) < 1$ to $\alpha M < 1$ by measuring the distance d in $\Re\times\mathcal{B}$ for a fixed λ. Namely, we consider the projection of $\Re \times \mathcal{B}$ onto $\{\lambda\} \times \mathcal{B}$. For any λ, we define the projected value function $Z^\lambda(p)$ as the solution of

$$\begin{aligned} Z^\lambda(p) \quad := \quad & \inf_q \left\{ \langle c(.,q), p(.)\rangle \ + \alpha\lambda \int \bar{F}(y+q)p(y)dy \right. \\ & \left. +\alpha Z^\lambda(\rho(q,p)) \right\}. \end{aligned} \tag{32}$$

then we check that that

$$||T^\lambda(Z_1) - T^\lambda(Z_2)||_{\mathcal{B}} \leq \alpha M||Z_1 - Z_2||_{\mathcal{B}}.$$

If $\alpha M < 1$, then $T^\lambda(Z)$ is a contraction mapping. Thus, $T^\lambda(Z)$ has a fixed point in $\Re \times \mathcal{B}_c$. Consequently, $Z^\lambda(p)$ is uniquely defined. It is shown to be nondecreasing in λ with the next lemma.

Lemma 6. *$Z^\lambda(p)$ is nondecreasing in λ.*

Now consider the function $g(\lambda)$ for $\lambda \geq 0$ defined by

$$g(\lambda) := \inf_q \left\{c(0,q) + \alpha\lambda\bar{F}(q) + \alpha Z^\lambda(\rho(q,\delta))\right\}, \tag{33}$$

where Z^λ is given by (32). By Lemma 6, the map $g(\lambda)$ is nondecreasing. It can be proved that it is concave with a rate of increase of at most α .

Theorem 4. *The system in* (21)-(22) *has a unique solution if $c(I,q)$ is nondecreasing in q.*

It is worth noting that the condition in the statement of Theorem 4 is very weak. Without loss of generality, we can use the convention that the cost $c(I,q)$ is charged at the beginning of a period. In that case, holding and lost sales costs depend on only I, while the ordering costs increase in q as required by Theorem 4.

V. Concluding Remarks

This paper has provided a rigorous treatment of a class of inventory problems with partial observations. The observation process is a binary valued Markov chain, which arises from the "zero balance walk" approach to inventory management. Since the inventory level is often not observed, its conditional distribution given the observation represents the state of the system. This approach immediately results in a dynamic program in a functional space. The dynamic programming equations are simplified by using unnormalized probabilities. Doing so, a Zakai-type system of equations are derived for our inventory problems. This simplification has allowed us to prove the existence of a value function under various assumptions. For small discount factor α, we show the uniqueness of a solution to the system of equations in (21)-(22). Then this solution must be the value function.

Unnormalized probabilities and the corresponding Zakai-type system of equations can also be used to simplify the analysis of other problems. A case in point is the partially observed demands in the context of the newsvendor problem. In this case, the demand is observed if it is less than the inventory. Otherwise, only the event that it is larger than or equal to the inventory is observed. These observations are used to update the demand distribution. This gives rise to a dynamic programming equation, whose state is the current demand distribution, which is infinite dimensional. An equivalent dynamic program, whose state is the unnormalized demand distribution, can be constructed. The equivalent dynamic program facilitates the existence and the structural property arguments in Bensoussan et al. [2].

Information delays also lead to partially observed inventory problems. When there is information delay, the current inventory level is not observed by the IM. Instead, he observes the exact inventory level of a prior period. In this case, the state of the system can be summarized by a sufficient statistic (finite-dimensional vector). Bensoussan et al. [3] and [4] suppose that the unmet demand is backordered. The former paper establishes that a reference inventory position is a sufficient statistic. The latter generalizes the former by allowing for information delays given to be a Markov process. In addition to the reference inventory position, [4] shows that the value of the latest delay observation and the age of this observation must be included in the sufficient statistic vector. Our results on the partially observed inventories are announced in Bensoussan et al. [5].

References

[1] A. Bensoussan, *Stochastic control of partially observable systems*, Cambridge University Press, England, 1992.

[2] A. Bensoussan, M. Çakanyıldırım and S.P. Sethi (2005). *A multiperiod newsvendor problem with partially observed demand.* Working paper SOM 200550, School of Management, University of Texas at Dallas, TX.

[3] A. Bensoussan, M. Çakanyıldırım and S.P. Sethi (2005). *Optimality of base stock and (s,S) policies for inventory problems with information delays*. Working paper SOM 200547, School of Management, University of Texas at Dallas, TX.

[4] A. Bensoussan, M. Çakanyıldırım and S.P. Sethi (2005). *Optimal ordering policies for inventory problems with dynamic information delays*. Working paper 200560, School of Management, University of Texas at Dallas, TX.

[5] A. Bensoussan, M. Çakanyıldırım and S.P. Sethi (2005). *On the optimal control of partially observed inventory systems*. To appear in the Comptes Rendus de l'Academié des Sciences.

[6] R.J. Elliott, L. Aggoun and J.B. Moore, *Hidden Markov Models: Estimation and Control,* Springer-Verlag New York, 1995.

[7] F.W. Harris, *Operations and Cost*, in Factory Management Series, A.-W. Shaw Company, Chicago, Il, Chapter IV, 1913.

[8] H.J. Kushner, *Dynamic equations for nonlinear filtering,* Journal of Differential Equations, Vol.3 (1967), pp. 179-190.

[9] M. Zakai, *On the optimal filtering of diffusion processes,* Z. Wahrsch. verw. Gebiete, Vol.11 (1969), pp. 230-243.