Adaptive Control of Stochastic Manufacturing Systems with Hidden Markovian Demands and Small Noise

T. E. Duncan, B. Pasik-Duncan, and Q. Zhang

Abstract—The adaptive production planning of failure-prone manufacturing systems is considered in this paper. In real manufacturing systems, the product demand is usually not known *a priori*. One of the major tasks in production scheduling is to estimate and predict the demand. In this paper, the authors consider the demand to be either the sum of an unknown rate and a small white noise or the sum of a hidden Markov chain and a small white noise. An algorithm is given to define a family of estimates for the unknown demand processes. Based on this family of estimates, adaptive controls are constructed, which are shown to be nearly optimal.

Index Terms—Hidden Markov chain, nearly optimal control, parameter identification, production planning.

I. INTRODUCTION

The study of flexible manufacturing systems has attracted much attention in the recent years because of the capability of these systems to describe today's increasingly unpredictable market place. One of the key factors that determines the production strategy of a manufacturing firm is the market demand for its products. There are various ways of formulating the demand processes. Typically, the rate of demand is modeled as a finite-state Markov chain (see [7]). In practice, the demand process is usually not directly observable by the controller of the production. The presence of the unobservable demand processes makes the problem very difficult to solve. It is the purpose of this paper to investigate a class of production planning problems with incomplete information of demand. In particular, the demand process is described as a hidden Markov chain plus a small noise perturbation. The problem is to find the rate of production that minimizes the overall costs of inventory/shortage and production.

To solve the problem, the unobservable demand process has to be estimated. In this paper, a family of estimates is given to identify the unknown process. Using these estimates of the unknown parameters, an adaptive control is constructed for the problem. The adaptive control is shown to be asymptotically optimal as the noise tends to zero.

Failure-prone manufacturing systems are considered in this paper. The production availability (or the machine state) process is modeled as a finite-state Markov chain; see [7] for discussions and more details on these models.

There is a sizable amount of work on stochastic adaptive control. However, the work on continuous time stochastic adaptive control (e.g., [1], [3], [4], and [6]) is significantly less than the work on discrete-time stochastic adaptive control. The use of hidden Markov models is described in [5].

This paper is organized as follows. In Section II, the problem of the adaptive production planning with unobservable constant

demand rates is formulated. In Section III, we give an algorithm for identifying the unknown parameters. An error estimate of the algorithm is also given. In Section IV, an adaptive control based on the identification algorithm is obtained. The adaptive control is shown to be nearly optimal as the noise in the demand process tends to zero. To emphasize the main idea in parameter identification and to simplify the exposition, only the models with constant demand rates are considered in these sections. In Section V, these results are extended to models involving hidden Markov chains. Some related technical lemmas are given in the Appendix.

II. PROBLEM FORMULATION

Consider a manufacturing system that produces n distinct part types using m identical machines that are subject to break down and repair. Let $u(t) \in R^n$ denote the vector of production rates, $x(t) \in R^n$ the vector of total inventories/backlogs, and $z(t) \in R^n$ the vector of demands. These processes are related by the following differential equation:

$$dx(t) = u(t)\, dt - dz(t), \qquad x(0) = x \in R^n. \tag{1}$$

The demand $z(\cdot)$ is given by the following differential equation:

$$dz(t) = \bar{z}\, dt + \sqrt{\varepsilon}\sigma\, dw(t), \qquad z(0) = z_0 \tag{2}$$

where $\bar{z}$ is a vector of unknown constants, $\varepsilon > 0$ is a small parameter, σ is a given $n \times n$ matrix, and $(w(t), t \geq 0)$ is a standard R^n-valued Brownian motion, defined on a complete probability space $(\Omega, \mathcal{F}, P)$. For simplicity of exposition, initially we consider the case where $\bar{z} \in \mathcal{Z} = \{z_1, z_2, \cdots, z_N\}$. An extension of the models involving hidden Markov demand is considered in Section V.

Let $\mathcal{M} = \{0, 1, \cdots, m\}$ denote the set of machine total capacity states and let the process $(\alpha(t), t \geq 0)$ where $\alpha(t) \in \mathcal{M}$ denote the total capacity process for the manufacturing system.

Note that $(x(\cdot), \alpha(\cdot))$ is observable up to time t. Since $u(\cdot)$ depends on $(x(\cdot), \alpha(\cdot))$ up to t, $z(\cdot)$ is also available up to time t. The cost function J is defined by

$$J(u(\cdot)) = E \int_0^\infty e^{-\rho t} G(x(t), u(t))\, dt \tag{3}$$

where G is the running cost of inventory/backlog and production and $\rho > 0$ is the discount rate. The problem is to find a control $u(\cdot)$ that minimizes $J(u(\cdot))$.

Now the production (or control) constraints are specified. For each $i \in \mathcal{M} = \{0, 1, 2, \cdots, m\}$, let

$$\mathcal{U}(i) = \{u = (u_1, \cdots, u_n) : u_j \geq 0, \ j = 1, \cdots, n,$$
$$\text{and } p_1 u_1 + \cdots + p_n u_n \leq i\} \tag{4}$$

where $p_j \geq 0$, $j = 1, \cdots, n$, are given constants with p_j representing the amount of capacity needed to produce part type j at rate one. With this definition, the production constraint at time t is $u(t) \in \mathcal{U}(\alpha(t))$.

Assumptions:

A1) There exist constants C and $k \in N = \{1, 2, \cdots\}$ such that for all x, x_1, u and $u_1 \in R^n$, $0 \leq G(x, u) \leq C(1 + |x|^{2k})$, and $|G(x, u) - G(x_1, u_1)| \leq C((1 + |x|^{2k} + |x_1|^{2k})|x - x_1| + |u - u_1|)$.

A2) $\alpha(t) \in \mathcal{M}$ is a finite-state Markov process governed by the $(m+1) \times (m+1)$ matrix generator $Q = (q_{ij})$ with $q_{ij} \geq 0$ if $i \neq j$ and $q_{ii} = -\sum_{j \neq i} q_{ij}$.

A3) The random variable $\bar{z}$ and the processes $(\alpha(t), t \geq 0)$ and $(w(t), t \geq 0)$ are independent.

Definition: A control $u(\cdot) = \{u(t, \alpha(t), x(t)) : t \geq 0\}$ is *admissible* if: 1) $u : R_+ \times \mathcal{M} \times R^n \to R^n$ is a Borel measurable function such that (1) has a unique strong solution and 2) $u(t) \in \mathcal{U}(\alpha(t))$ for all $t \geq 0$. The set of all admissible controls is denoted by $\mathcal{A}$.

The control problem can be expressed as follows:

$$\mathcal{P} : \begin{cases} \min J(u(\cdot)) = E \int_0^\infty e^{-\rho t} G(x(t), u(t)) \, dt \\ \text{s.t. } dx(t) = (u(t) - \bar{z}) \, dt - \sqrt{\varepsilon} \sigma \, dw(t), \quad x(0) = x. \end{cases} \quad (5)$$

III. PARAMETER IDENTIFICATION

First of all, for a given $u(\cdot) \in \mathcal{A}$, we define a $\mathcal{Z}$-valued family of estimates $(\hat{z}(t), t \geq 0)$ of the unknown parameter $\bar{z} \in \mathcal{Z}$ and show that it is strongly consistent. Let $(\hat{z}(t), t \geq 0)$ be given by

$$\hat{z}(t) = \arg \min_{i \in \{1, \cdots, N\}} |\check{z}(t) - z_i| \quad (6)$$

$$\check{z}(t) = \frac{-1}{t} \left(x(t) - x(0) - \int_0^t u(s) \, ds \right) \quad (7)$$

where $(x(t), t \geq 0)$ satisfies (1) and $\check{z}(0) = 0$.

Proposition 1: Let $\bar{z} \in \mathcal{Z}$ in (2) be fixed. If $(\hat{z}(t), t \geq 0)$ is the family of estimates of $\bar{z}$ given by (6) then

$$\lim_{t \to \infty} \hat{z}(t) = \bar{z} \quad \text{a.s.} \quad (8)$$

Proof: It follows from (5) that for $t > 0$

$$\bar{z} = \frac{-1}{t} \left(x(t) - x(0) - \int_0^t u(s) \, ds + \int_0^t \sqrt{\varepsilon} \sigma \, dw(s) \right).$$

Apply the Strong Law of Large Numbers to the family of random variables $\{-t^{-1} \int_0^t \sqrt{\varepsilon} \sigma \, dw(s), \, t > 0\}$ to verify (8). $\square$

The verification of next proposition follows directly from the Law of the Iterated Logarithm for Brownian motion (e.g., [8]).

Proposition 2: Let $(\hat{z}(t), t > 0)$ be given by (6). For each $\tilde{\varepsilon} > 0$ there is a $t_0 > 0$ such that

$$P(|\hat{z}(t) - z_j| < |\hat{z}(t) - z_i| \text{ for some } j \neq i$$
$$\text{and some } t \geq t_0 \mid \bar{z} = z_i) < \tilde{\varepsilon}. \quad (9)$$

Proposition 3: For each $\varepsilon > 0$ the family of estimates $(\hat{z}(t), t > 0)$ given by (6) satisfies, for some K

$$P \left(\limsup_{t \to \infty} \frac{|\hat{z}(t) - z_i|}{t^{-1/2 + \varepsilon}} \leq K \mid \bar{z} = z_i \right) = 1. \quad (10)$$

The verification of this proposition follows from the proof of [Th. 2.1, 2] that uses the Law of the Iterated Logarithm.

Lemma 1: For any $t > 0$, $P(\hat{z}(t) \neq z_i \mid \bar{z} = z_i) \leq K \varepsilon / t$. In particular, let $t = \delta_\varepsilon = \varepsilon |\log \varepsilon|$. Then

$$P(\hat{z}(\delta_\varepsilon) \neq z_i \mid \bar{z} = z_i) \leq K / |\log \varepsilon| \to 0, \quad \text{as } \varepsilon \to 0.$$

Proof: Note that $\check{z}(t) = \bar{z} + \sqrt{\varepsilon} \sigma w(t) / t$. We have, for $j \neq i$

$$\{\hat{z}(t) = z_j\} = \{|\check{z}(t) - z_j| < |\check{z}(t) - z_i|\}$$
$$= \{|z_i - z_j + \sqrt{\varepsilon} \sigma w(t)/t| < |\sqrt{\varepsilon} \sigma w(t)/t|\}$$
$$\subset \{|\sqrt{\varepsilon} \sigma w(t)/t| \geq |z_i - z_j|/2\}.$$

Thus, in view of the independence of $\bar{z}$ and $w(\cdot)$ in A3)

$$P(\hat{z}(t) = z_j \mid \bar{z} = z_i)$$
$$\leq P(|\sqrt{\varepsilon} \sigma w(t)/t| \geq |z_i - z_j|/2 \mid \bar{z} = z_i)$$
$$= P(|\sqrt{\varepsilon} \sigma w(t)/t| \geq |z_i - z_j|/2)$$
$$\leq 4E(\sqrt{\varepsilon} \sigma w(t)/t)^2/|z_i - z_j|^2 \leq K \varepsilon / t. \quad \square$$

IV. ASYMPTOTIC OPTIMAL CONTROLS

Recall that $\mathcal{Z} = \{z_1, z_2, \cdots, z_N\}$. For $i \in \Theta = \{1, \cdots, N\}$, let $u^*(\alpha, z_i, x)$ denote the optimal control for

$$\mathcal{P}_i : \begin{cases} \min J(u(\cdot)) \\ \text{s.t. } dx(t) = (u(t) - z_i) \, dt, \quad x(0) = x. \end{cases} \quad (11)$$

For each $i \in \Theta$, it can be shown (cf. [7]) that the value function $v^i(\alpha, x)$, defined as the minimum cost over $\mathcal{A}$ with $\alpha(0) = \alpha$ and $x(0) = x$, is the unique viscosity solution to the following Hamilton–Jacobi–Bellman (HJB) equation:

$$\rho v^i(\alpha, x) = \min_{u \in \mathcal{U}(\alpha)} [(u - z_i) \nabla_x v^i(\alpha, z_i, x) + G(x, u)]$$
$$+ Q v^i(\cdot, x)(\alpha) \quad (12)$$

for any $\alpha \in \mathcal{M}$, where $Q v^i(\cdot, x)(\alpha) = \sum_{\beta \neq \alpha} q_{\alpha \beta}(v^i(\beta, x) - v^i(\alpha, x))$. Let an optimal feedback control $u^*(\alpha, z_i, x)$ be obtained by minimizing the right-hand side of (12), i.e.,

$$u^*(\alpha, z_i, x) \nabla_x v^i(\alpha, x) + c(u^*(\alpha, z_i, x))$$
$$= \min\{u \nabla_x v^i(\alpha, x) + c(u) : u \in \mathcal{U}(\alpha)\}. \quad (13)$$

To obtain a Lipschitz optimal control, an additional assumption is made.

A4) The function G given in (3) can be expressed as $G(x, u) = h(x) + c(u)$ where $h(x)$ is differentiable and $c(u)$ is twice differentiable with $c_{uu}(u) \geq c_0 > 0$. Furthermore, there exist a constant C and a integer $k \geq 0$ such that
$$|h(x + y) - h(x) - \nabla_x h(x) y| \leq C(1 + |x|^{2k})|y|^2.$$

A proof of the following lemma is given in [7].

Lemma 2: Assume A1)–A4). For $\mathcal{P}_i$ in (11) let $u^*(\alpha, z_i, x)$ be an optimal feedback control determined from (13). Then there exist constants C and $k \geq 0$ such that for $x, x_1 \in R^n$

$$|u^*(\alpha, z_i, x) - u^*(\alpha, z_i, x_1)| \leq C(1 + |x|^{2k} + |x_1|^{2k})|x - x_1|.$$

As in Lemma 1, let $\delta_\varepsilon = \varepsilon |\log \varepsilon|$. Define the certainty equivalence adaptive control $(u^\varepsilon(t, \alpha, x), t > 0)$ as

$$u^\varepsilon(t, \alpha, x) = \begin{cases} u^*(\alpha, z_1, x), & \text{if } t < \delta_\varepsilon \\ u^*(\alpha, \hat{z}(\delta_\varepsilon), x), & \text{if } t \geq \delta_\varepsilon \end{cases} \quad (14)$$

for the system given by (1) and (2). It is easy to see that such a control is admissible.

Theorem 1: For the control of the manufacturing system given by (1) and (2) with the cost functional (3) if A1)–A4) are satisfied, then the adaptive control $(u^\varepsilon(t, \alpha, x), t > 0)$ given by (14) is asymptotically optimal, i.e.,

$$\lim_{\varepsilon \to 0} \left| J(u^\varepsilon(\cdot)) - \inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot)) \right| = 0. \quad (15)$$

Proof: Note that $J(u^\varepsilon(\cdot)) \geq \inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot))$. It suffices to show $\limsup_{\varepsilon \to 0} (J(u^\varepsilon(\cdot)) - \inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot))) = 0$. For each $i \in \Theta$, given $\bar{z} = z_i$, it is easy to show that

$$\inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot))$$
$$= \inf_{u(\cdot) \in \mathcal{A}} E \left[\int_0^\infty e^{-\rho t} G(x(t), u(t)) \, dt \mid \bar{z} = z_i \right]$$
$$\geq \inf_{\bar{u}(\cdot) \in \mathcal{A}} E \left[\int_0^\infty e^{-\rho t} G(\bar{x}(t), \bar{u}(t)) \, dt \mid \bar{z} = z_i \right] + O(\sqrt{\varepsilon})$$
$$= E \left[\int_0^\infty e^{-\rho t} G(\bar{x}^*(t), \bar{u}_i^*(t)) \, dt \mid \bar{z} = z_i \right] + O(\sqrt{\varepsilon}) \quad (16)$$

where $d\bar{x}(t) = (\bar{u}(t) - z_i) \, dt$, $\bar{x}(0) = x$, $d\bar{x}^*(t) = (\bar{u}_i^*(t) - z_i) \, dt$, $\bar{x}^*(0) = x$, and $\bar{u}_i^*(t) = u^*(\alpha(t), z_i, \bar{x}^*(t))$. Let $\eta(T) =$

$\sup_{u(\cdot)\in\mathcal{A}} E \int_T^\infty e^{-\rho t} G(x(t), u(t))\, dt$. Then, Lemma A.1 (see the Appendix) yields $\eta(T) \leq C \int_T^\infty e^{-\rho t}(1+e^{\rho t/2})\, dt$ for some constant C.

On the other hand, for each $T > 0$

$$E \int_0^\infty e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt$$

$$= E \int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt$$

$$+ E \int_T^\infty e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt$$

$$\leq E \int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt + \eta(T).$$

By Lemma 1, it follows that

$$E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt\right]$$

$$= \sum_{i=1}^N E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt \,\Big|\, \bar{z} = z_i\right] P(\bar{z} = z_i).$$

Note that

$$E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt \,\Big|\, \bar{z} = z_i\right]$$

$$= E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt\, I_{\{\hat{z}(\delta_\varepsilon)=z_i\}} \,\Big|\, \bar{z} = z_i\right]$$

$$+ E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u^\varepsilon(t))\, dt\, I_{\{\hat{z}(\delta_\varepsilon)\neq z_i\}} \,\Big|\, \bar{z} = z_i\right]$$

where I_F is the indicator function of a set F. The first term above is less than or equal to

$$E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u_i^*(t))\, dt \,\Big|\, \bar{z} = z_i\right].$$

The second term is bounded above by $KTP(\hat{z}(\delta_\varepsilon) \neq z_i \mid \bar{z} = z_i) \leq KT/|\log\varepsilon| \to 0$ as $\varepsilon \to 0$. For $i \in \Theta$ let

$$R_i(\varepsilon, T) = E\left[\int_0^T e^{-\rho t} G(x^\varepsilon(t), u_i^*(t))\, dt \,\Big|\, \bar{z} = z_i\right]$$

$$- E\left[\int_0^T e^{-\rho t} G(\bar{x}^*(t), u_i^*(t))\, dt \,\Big|\, \bar{z} = z_i\right].$$

Then, Lemma A.3 and (16) yield that, for each fixed $T > 0$, $R_i(\varepsilon, T) \to 0$ as $\varepsilon \to 0$. Moreover, $J(u^\varepsilon(\cdot)) \leq \inf_{u(\cdot)\in\mathcal{A}} J(u(\cdot)) + \sup_{i\in\Theta} R_i(\varepsilon, T) + \eta(T) + O(\varepsilon^2)$. By this inequality, it follows that $\limsup_{\varepsilon\to 0}(J(u^\varepsilon(\cdot)) - \inf_{u(\cdot)\in\mathcal{A}} J(u(\cdot))) \leq \eta(T) \to 0$ as $T \to \infty$. This implies (15). ☐

V. HIDDEN MARKOVIAN DEMAND

In this section the demand process is described as the sum of a hidden Markov chain and a small white noise, that is, more precisely

$$dz(t) = z_h(t)\, dt + \sqrt{\varepsilon}\sigma\, dw(t), \qquad z(0) = 0 \qquad (17)$$

where $(z_h(t), t \geq 0)$ is a Markov chain, $\mathcal{Z}$-valued and generated by Q_h, that is hidden to the controller of the system. Moreover, we assume as in A3) that $z_h(\cdot)$, $\alpha(\cdot)$, and $w(\cdot)$ are independent.

Note that if $\varepsilon = 0$, then the value of $z_h(s)$ is observable for $s \leq t$. The HJB equation for the problem is

$$\rho v(\alpha, z, x) = \min_{0\leq u\leq\alpha}\left[(u - z)\nabla_x v(\alpha, z, x)\right.$$

$$\left. + G(x, u)\right] + Q v(\cdot, z, x)(\alpha) + Q_h v(\alpha, \cdot, x)(z) \qquad (18)$$

for any $\alpha \in \mathcal{M}$ and $z \in \mathcal{Z}$. It can be shown that the value function $v(\alpha, z, x)$ is the unique viscosity solution to (18). Using this and Assumption A4), it can be shown the value function is continuously differentiable (cf. [Ch. 5, 7]). An optimal control $u^*(\alpha, z, x)$ for this problem can be obtained by minimizing the right-hand side of (18), i.e.,

$$u^*(\alpha, z, x)\nabla_x v(\alpha, z, x) + c(u^*(\alpha, z, x))$$

$$= \min\{u\nabla_x v(\alpha, z, x) + c(u) : 0 \leq u \leq \alpha\}. \qquad (19)$$

Let $(Y(t), t \geq 0)$ be the process given by $Y(t) = x - x(t) + \int_0^t u(s)\, ds$ and for $\delta > 0$ let $\tilde{z}_h(t, \delta)$ be given by $\tilde{z}_h(t, \delta) = (Y(t) - Y(t-\delta))/\delta$. Note that

$$\tilde{z}(t, \delta) = \frac{1}{\delta}\int_{t-\delta}^t z_h(s)\, ds + \sqrt{\varepsilon}\sigma\left(\frac{w(t) - w(t-\delta)}{\delta}\right).$$

For each $k = 1, 2, \cdots$ and $t = k\delta_\varepsilon$, define $\hat{z}_h(k\delta_\varepsilon)$ as $\hat{z}_h(k\delta_\varepsilon) = z_i$ if

$$\min_{j=\{1,\cdots,N\}} |\tilde{z}(k\delta_\varepsilon, \delta_\varepsilon) - z_j| = |\tilde{z}(k\delta_\varepsilon, \delta_\varepsilon) - z_i|.$$

Define $\hat{z}_h^\varepsilon(t) = z_1$ for $t \in [0, \delta_\varepsilon)$ and

$$\hat{z}_h^\varepsilon(t) = \hat{z}_h(k\delta_\varepsilon), t \in [k\delta_\varepsilon, (k+1)\delta_\varepsilon), \qquad k = 1, 2, \cdots. \qquad (20)$$

Since the estimate $\hat{z}_h^\varepsilon(t), t > 0$ is required to estimate the process $(z_h(t), t \geq 0)$, the performance measure for these estimates is weaker than for $(\hat{z}(t), t > 0)$ given by (6).

Lemma 3: For each $T > 0$, we have

$$\lim_{\varepsilon\to 0} E \int_0^T |\hat{z}_h^\varepsilon(t) - z_h(t)|\, dt = 0. \qquad (21)$$

Proof: The verification of (21) follows from the expectation of reflected Brownian motion and the fact that given $\delta > 0$ there is a $\tilde{\delta} > 0$ such that for each interval I of length $\tilde{\delta}$ in $[0, T]$ the probability that $(z_h(t), t \in I)$ changes state is less than δ. ☐

Theorem 2: Consider the control problem (1) and (3) where $z(t)$ is determined by (17). If A1)–A4) hold, then

$$u^\varepsilon(\cdot) = u^*\big(\alpha(t), \hat{z}_h^\varepsilon(t)x^\varepsilon(t)\big)$$

is asymptotically optimal, that is

$$\lim_{\varepsilon\to 0}\left| J(u^\varepsilon(\cdot)) - \inf_{u(\cdot)\in\mathcal{A}} J(u(\cdot))\right| = 0. \qquad (22)$$

Proof: Let $(\bar{x}^\varepsilon(t), t \in [0, T])$ and $(x^\varepsilon(t), t \in [0, T])$ denote the processes that satisfy, with $\bar{x}^*(0) = x$ and $x^\varepsilon(0) = x$

$$d\bar{x}^*(t) = (u^*(\alpha(t), z_h(t), \bar{x}^*(t)) - z_h(t))\, dt,$$

$$dx^\varepsilon(t) = (u^*(\alpha(t), \hat{z}_h^\varepsilon(t), x^\varepsilon(t)) - z_h(t))\, dt - \sqrt{\varepsilon}\sigma\, dw(t)$$

respectively, where u^* is determined from (19). Then it easily follows that for n sufficiently large

$$E|x^\varepsilon(t) - \bar{x}^*(t)| \leq E \int_0^t \big| u^*\big(\alpha(s), \hat{z}_h^\varepsilon(s), x^\varepsilon(s)\big)$$

$$- u^*\big(\alpha(s), \hat{z}_h^\varepsilon(s), \bar{x}^*(s)\big)\big|\, ds$$

$$+ E \int_0^t \big| u^*\big(\alpha(s), \hat{z}_h^\varepsilon(s), \bar{x}^*(s)\big)$$

$$- u^*\big(\alpha(s), z_h(s), \bar{x}^*(s)\big)\big|\, ds + O(\sqrt{\varepsilon}). \qquad (23)$$

Note that $I_{\{\hat{z}_h^\varepsilon(t)\neq z_h(t)\}} \leq K|\hat{z}_h^\varepsilon(t) - z_h(t)|$, where $K = \max\{1/|z_i - z_j| : i \neq j\}$. It follows from Lemma 3 that

$$E \int_0^t \big| u^*\big(\alpha(s), \hat{z}_h^\varepsilon(s), \bar{x}^*(s)\big) - u^*(\alpha(s), z_h(s), \bar{x}^*(s))\big|\, ds$$

$$\leq CE \int_0^t |\hat{z}_h^\varepsilon(s) - z_h(s)|\, ds \to 0 \quad \text{as } \varepsilon \to 0. \qquad (24)$$

In view of the Lipschitz property of $u^*(\alpha, z, x)$ with respect to x, (23), (24), and the Gronwall's inequality, it follows that for each $0 \leq t \leq T$, $E|x^\varepsilon(t) - \bar{x}^*(t)| \to 0$ as $\varepsilon \to 0$. Thus, by Lemma A.1, it follows that $J(u^\varepsilon(\cdot)) - J(\bar{u}^*(\cdot)) \to 0$ as $\varepsilon \to 0$, where $\bar{u}^*(t) = u^*(\alpha(t), z_h(t), \bar{x}^*(t))$.

On the other hand, it is easy to see that $\inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot)) \geq \inf_{u(\cdot) \in \mathcal{A}_{z_h}} J(u^\delta(\cdot))$, where $\mathcal{A}_{z_h} = \{u(\cdot) : u(t) \text{ is adapted to } \sigma\{(\alpha(s), z_h(s), x(s) : s \leq t\}\}$ because the control in $\mathcal{A}_{z_h}$ contains more information. Following similar procedure as in the proof of Theorem 1, we can show $\inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot)) \geq J(\bar{u}^*(\cdot))$. Hence, $\limsup_{\varepsilon \to 0} (J(u^\varepsilon(\cdot)) - \inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot))) \to 0$ as $T \to \infty$. $\square$

APPENDIX

In this section we state and prove several technical lemmas.

Lemma A.1: Let $(x(t), t \geq 0)$ satisfy (1), let $u \in \mathcal{A}$, and let $\bar{z} = z_i$. There exists a constant C such that

$$\eta(T) := \sup_{i \in \Theta, u(\cdot) \in \mathcal{A}} E\left[\int_T^\infty e^{-\rho t} G(x(t), u(t)) \, dt\right]$$
$$\leq C \int_T^\infty e^{-\rho t} (1 + \exp(\rho t / 2)) \, dt. \tag{25}$$

Proof: For $\bar{z} = z_i$ and $k \geq 1$, we can show by using Ito's formula and Gronwall's inequality that

$$E|x(t)|^{2k} \leq C \exp(\rho t / 2). \tag{26}$$

By Assumption A1), it follows that

$$E \int_T^\infty e^{-\rho t} G(x(t), u(t)) \, dt$$
$$\leq C \int_T^\infty e^{-\rho t} (1 + E[|x(t)|^{2k}]) \, dt$$
$$\leq C \int_T^\infty e^{-\rho t} (1 + \exp(\rho t / 2)) \, dt. \qquad \square$$

Lemma A.2: Let $(x(t), t \geq 0)$ satisfy (1), let $u \in \mathcal{A}$, and let $\bar{z} = z_i$. There is a constant C such that

$$\sup_{i \in \Theta, u(\cdot) \in \mathcal{A}} E\left[\int_0^\infty e^{-\rho t} G(x(t), u(t)) \, dt\right] \leq C.$$

Proof: By (26), the following inequality is satisfied:

$$E \int_0^\infty e^{-\rho t} G(x(t), u(t)) \, dt$$
$$\leq C \int_0^\infty e^{-\rho t} (1 + E[|x(t)|^{2k}]) \, dt$$
$$\leq C \int_0^\infty e^{-\rho t} \left(1 + \exp\left(\frac{\rho t}{2}\right)\right) dt = \frac{3C}{\rho}. \qquad \square$$

Lemma A.3: Let $\varepsilon > 0$ be fixed, $t_\varepsilon = c_0 \varepsilon |\log \varepsilon|$, and $(x(t), t \geq 0)$ and $(\bar{x}(t), t \geq 0)$ satisfy

$$dx(t) = (u_i(t) - z_i) \, dt - \sqrt{\varepsilon} \sigma \, dw(t), \quad x(0) = x$$

where

$$u_i(t) = \begin{cases} u_i^0(t), & \text{if } 0 \leq t < t_\varepsilon \\ u^*(\alpha(t), z_i, x(t)), & \text{if } t_\varepsilon \leq t \leq T, \end{cases}$$
$$d\bar{x}(t) = (\bar{u}_i^*(t) - z_i) \, dt, \qquad \bar{x}(0) = x$$

where

$$\bar{u}_i^*(t) = \begin{cases} \bar{u}_i^0(t), & \text{if } 0 \leq t < t_\varepsilon \\ u^*(\alpha(t), z_i, \bar{x}(t)), & \text{if } t_\varepsilon \leq t \leq T \end{cases}$$

where u_i^0 and $\bar{u}_i^0$ are arbitrary controls and $u^*(\alpha, z_i, x)$ is determined by (13). Then, for some C and $c > 0$

$$E|x(t) - \bar{x}(t)| \leq C(\varepsilon |\log \varepsilon| + \sqrt{\varepsilon} \sqrt{T}) \exp(cT).$$

Proof: It is elementary that

$$E|x(t_\varepsilon) - \bar{x}(t_\varepsilon)| \leq C(t_\varepsilon + \sqrt{\varepsilon} E|w(t_\varepsilon)|) \leq C \varepsilon |\log \varepsilon|.$$

Moreover, by the Lipschitz continuity of $u^*(\alpha, z_i, x)$ in x, it follows that, for $0 \leq t \leq T$

$$E|x(t) - \bar{x}(t)| \leq E|x(t_\varepsilon) - \bar{x}(t_\varepsilon)|$$
$$+ C \int_{t_\varepsilon}^t E|x(s) - \bar{x}(s)| \, ds + C\sqrt{\varepsilon}.$$

By Gronwall's inequality, we have

$$E|x(t) - \bar{x}(t)| \leq C(\varepsilon |\log \varepsilon| + \sqrt{\varepsilon} \sqrt{T}) \exp(cT). \qquad \square$$

REFERENCES

[1] H. F. Chen, T. E. Duncan, and B. Pasik-Duncan, "Stochastic adaptive control for continuous time linear systems with quadratic cost," *J. Appl. Math. Optim.*, vol. 34, pp. 113–138, 1996.
[2] T. E. Duncan and B. Pasik-Duncan, "Rate of convergence for an estimator in a portfolio and consumption model," *J. Optim. Th. Appl.*, vol. 61, pp. 53–59, 1989.
[3] ——, "Adaptive control of continuous time linear stochastic systems," *Math. Contr., Signals Syst.*, vol. 3, pp. 45–60, 1990.
[4] T. E. Duncan, B. Pasik-Duncan, and L. Stettner, "Almost self-optimizing strategies for the adaptive control of diffusion processes," *J. Optim. Th. Appl.*, vol. 81, pp. 479–507, 1994.
[5] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control.* New York: Springer-Verlag, 1994.
[6] U. G. Haussmann and Q. Zhang, "Stochastic adaptive control with small observation noise," *Stochastics Stochastics Rep.*, vol. 32, pp. 109–144, 1990.
[7] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems.* Boston, MA: Birkhäuser, 1994.
[8] S. R. S. Varadhan, *Lectures on Diffusion Problems and Partial Differential Equations.* New York: Springer-Verlag, 1980.