

OPTIMAL CONTROL OF A HYBRID SYSTEM
WITH PATHWISE AVERAGE COST*

Mrinal K. Ghosh†, Aristotle Arapostathis‡ and Steven I. Marcus§

Abstract. We study the ergodic control problem of switching diffusions representing a typical hybrid system that arises in numerous applications such as fault tolerant control systems, flexible manufacturing systems, etc. Under certain conditions, we establish the existence of a stable Markov nonrandomized policy which is almost surely optimal for the pathwise long-run average cost criterion. We then study the corresponding Hamilton-Jacobi-Bellman (HJB) equation and establish the existence of a unique solution in a certain class. Using this, we characterize the optimal policy as a minimizing selector of the Hamiltonian associated with the HJB equations. We apply these results to a failure prone manufacturing system and show that the optimal production rate is of the hedging point type.

1. Introduction. We address the problem of controlling switching diffusions by continually monitoring the drift and jump rates of the continuous and discrete components, respectively. The objective is to minimize, almost surely, the pathwise long-run average (ergodic) cost over all admissible policies. A controlled switching diffusion is a typical example of a hybrid system which arises in numerous applications of systems with multiple modes or failure modes, such as fault tolerant control systems, multiple target tracking, flexible manufacturing systems etc. [9], [10], [12]. The state of the system at time t is given by a pair $(X(t), S(t)) \in \mathbb{R}^d \times \mathcal{S}$, $\mathcal{S} = \{1, 2, \dots, N\}$. The continuous component $X(t)$ is governed by a "controlled diffusion process" with a drift vector which depends on the discrete component $S(t)$. The discrete component $S(t)$ is a "controlled Markov chain" with a transition matrix depending on the continuous component. The evolution of the process $(X(t), S(t))$ is governed by the following equations:

$$dX(t) = b(X(t), S(t), u(t))dt + \sigma(X(t), S(t))dW(t), \quad (1.1)$$

$$P(S(t + \delta t) = j \mid S(t) = i, X(s), S(s), s \leq t) \\ = \lambda_{ij}(X(t), u(t))\delta t + o(\delta t), \quad i \neq j, \quad (1.2)$$

for $t \geq 0$, $X(0) = X_0$, $S(0) = S_0$, where b , σ , λ are suitable functions, $\lambda_{ij} \geq 0$, $i \neq j$, $\sum_{j=1}^N \lambda_{ij} = 0$, $W(\cdot)$ is a standard Brownian motion and $u(\cdot)$ is a nonanticipative control process (admissible policy). The latter is called a Markov policy if $u(t) = v(X(t), S(t))$ for a suitable function v . Our aim is to minimize almost surely (a.s.) over all admissible policies the quantity

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T c(X(t), S(t), u(t))dt, \quad (1.3)$$

where c is the running cost function. Under certain conditions, we will show that there exists a Markov policy v and constant ρ^* such that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T c(X(t), S(t), v(X(t), S(t)))dt = \rho^* \quad \text{a.s.}$$

and for any other admissible policy $u(\cdot)$

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T c(X(t), S(t), u(t))dt \geq \rho^* \quad \text{a.s.}$$

This will establish that v is optimal in a much stronger sense; viz., the most "pessimistic" average cost under v is no worse than the most "optimistic" average cost under any other admissible policy. Also, under the conditions assumed in this paper, the optimal pathwise average cost coincides with the optimal expected average cost. So we will not distinguish between the two optimality criteria.

Our paper is organized as follows. In Section 2 we present a concise description of the problem. Section 3 is devoted to the study of recurrence and ergodicity of switching diffusions. The existence of an optimal policy is established in Section 4. The Hamilton-Jacobi-Bellman (HJB) equations are studied in Section 5. Using the results of Sections 2-5, a failure prone manufacturing system is analyzed in Section 6. Proofs are omitted due to length limitations.

2. Problem Description. We first exhibit that the switching diffusion (1.1), (1.2) can be constructed on a given probability space. Our presentation follows [9], [10]; we repeat it here for the sake of clarity and completeness. Let U be a compact metric space and $\mathcal{S} = \{1, 2, \dots, N\}$. Let

$$\bar{b} = [\bar{b}_1, \dots, \bar{b}_d]' : \mathbb{R}^d \times \mathcal{S} \times U \rightarrow \mathbb{R}^d \\ \sigma = [\sigma_{ij}(\cdot, \cdot)] : \mathbb{R}^d \times \mathcal{S} \rightarrow \mathbb{R}^{d \times d}, \\ \bar{\lambda}_{\ell m} : \mathbb{R}^d \times \mathcal{S} \rightarrow \mathbb{R}, \quad \ell, m = 1, \dots, N \\ \bar{\lambda}_{\ell m} \geq 0, \quad \text{for } \ell \neq m, \quad \sum_{m=1}^N \bar{\lambda}_{\ell m} = 0, \quad \text{for any } \ell \in \mathcal{S}.$$

We make the following assumptions which will be in effect throughout the paper.

(A1) (i) For each $i \in \mathcal{S}$, $\bar{b}(\cdot, i, \cdot)$ is bounded, continuous and Lipschitz in its first argument uniformly with respect to the third.

(ii) $\sigma_{ij}(\cdot, \cdot)$ is bounded, Lipschitz and for each k , $\sigma(\cdot, k)\sigma'(\cdot, k)$ is uniformly elliptic, i.e., there exists a constant $\mu > 0$ such that $\sigma(\cdot, k)\sigma'(\cdot, k) \geq \mu I$.

(iii) For each $\ell, m \in \mathcal{S}$, $\bar{\lambda}_{\ell, m}$ is continuous and Lipschitz in its first argument uniformly with respect to the second. Also, there exist constants Λ_0 , λ_0 , $0 < \lambda_0 < \Lambda_0$ such that for $\ell, m \in \mathcal{S}$, $\ell \neq m$, $\lambda_0 \leq \bar{\lambda}_{\ell m} \leq \Lambda_0$.

For any Polish space Y , $\mathfrak{B}(Y)$ will denote its Borel σ -field and $\mathcal{P}(Y)$ the space of probability measures endowed with the Prohorov topology, i.e., the topology of weak convergence. Let

† Department of Mathematics, Indian Institute of Science, Bangalore 560012, India.

‡ Department of Electrical and Computer Engineering, University of Texas, Austin, Texas 78712-1084.

§ Electrical Engineering Department and Systems Research Center, University of Maryland, College Park, MD 20742.

*This work was supported in part by the Texas Advanced Research Program under Grant No. 003658-186, in part by the Air Force Office of Scientific Research under Grants F49620-92-J-0045 and F49620-92-J-0083, and in part by the National Science Foundation under Grant CDR-8803012.

$\mathfrak{M}(Y)$ be the set of all nonnegative integer-valued σ -finite measures on $\mathfrak{B}(Y)$. Let $\mathfrak{M}_\sigma(Y)$ be the smallest σ -field on $\mathfrak{M}(Y)$ with respect to which all the maps from $\mathfrak{M}(Y)$ to $\mathbf{N} \cup \{\infty\}$ of the form $\mu \mapsto \mu(B)$ with $B \in \mathfrak{B}(Y)$ are measurable. $\mathfrak{M}(Y)$ will always be assumed to be endowed with this measurability structure. Let $\mathcal{V} = \mathcal{P}(U)$ and $b = [b_1, \dots, b_d]^T : \mathbf{R}^d \times \mathcal{S} \times \mathcal{V} \rightarrow \mathbf{R}^d$ be defined by

$$b_i(\cdot, \cdot, v) = \int_U \bar{b}_i(\cdot, \cdot, u)v(du). \quad (2.1)$$

Similarly, for $i, j \in \mathcal{S}$, $v \in \mathcal{V}$, λ_{ij} is defined as

$$\lambda_{ij}(\cdot, \cdot, v) = \int_U \bar{\lambda}_{ij}(\cdot, \cdot, u)v(du). \quad (2.2)$$

For $i, j \in \mathcal{S}$, $x \in \mathbf{R}^d$, $v \in \mathcal{V}$, we construct the intervals $\Delta_{ij}(x, v)$ of the real line in the following manner.

$$\begin{aligned} \Delta_{12}(x, v) &= [0, \lambda_{12}(x, v)) \\ \Delta_{13}(x, v) &= [\lambda_{12}(x, v), \lambda_{12}(x, v) + \lambda_{13}(x, v)) \\ &\vdots \\ \Delta_{1N}(x, v) &= \left[\sum_{j=2}^{N-1} \lambda_{1j}(x, v), \sum_{j=2}^N \lambda_{1j}(x, v) \right) \\ \Delta_{21}(x, v) &= \left[\sum_{j=2}^N \lambda_{1j}(x, v), \sum_{j=2}^N \lambda_{1j}(x, v) + \lambda_{21}(x, v) \right) \\ &\vdots \end{aligned}$$

and so on. For fixed x and v , these are disjoint intervals, and the length of $\Delta_{ij}(x, v)$ is $\lambda_{ij}(x, v)$. Now, define a function

$$h : \mathbf{R}^d \times \mathcal{S} \times \mathcal{V} \times \mathbf{R} \rightarrow \mathbf{R}$$

by

$$h(x, i, v, z) = \begin{cases} j - i & \text{if } z \in \Delta_{ij}(x, v) \\ 0 & \text{otherwise.} \end{cases} \quad (2.3)$$

Let $(X(t), S(t))$ be the $(\mathbf{R}^d \times \mathcal{S})$ -valued controlled switching diffusion process given by the following stochastic differential equations:

$$\begin{aligned} dX(t) &= b(X(t), S(t), v(t)) dt + \sigma(X(t), S(t)) dW(t) \\ dS(t) &= \int_{\mathbf{R}} h(X(t), S(t-), v(t), z) p(dt, dz), \end{aligned} \quad (2.4)$$

for $t \geq 0$ with $X(0) = X_0$, $S(0) = S_0$, where

- (i) X_0 is a prescribed \mathbf{R}^d -valued random variable.
- (ii) S_0 is a prescribed \mathcal{S} -valued random variable.
- (iii) $W(\cdot) = [W_1(\cdot), \dots, W_d(\cdot)]^T$ is a d -dimensional standard Wiener process.
- (iv) $p(dt, dz)$ is an $\mathfrak{M}(\mathbf{R}_+ \times \mathbf{R})$ -valued Poisson random measure with intensity $dt \times m(dz)$, where m is the Lebesgue measure on \mathbf{R} .
- (v) $p(\cdot, \cdot)$, $W(\cdot)$, X_0 and S_0 are independent.
- (vi) $v(\cdot)$ is a \mathcal{V} -valued process with measurable sample paths satisfying the following nonanticipativity property.

Let

$$\begin{aligned} \mathcal{F}_t^v &= \sigma\{v(s), s \leq t\}, \\ \mathcal{F}_{[t, \infty)}^{W, p} &= \sigma\{W(s) - W(t), p(A, B) : A \in \mathfrak{B}([s, \infty)), \\ &\quad B \in \mathfrak{B}(\mathbf{R}), s \geq t\}. \end{aligned}$$

Then \mathcal{F}_t^v and $\mathcal{F}_{[t, \infty)}^{W, p}$ are independent.

Such a process $v(\cdot)$ will be called an admissible (control) pol-

icy. If $v(\cdot)$ is a Dirac measure, i.e., $v(\cdot) = \delta_{u(\cdot)}$, where $u(\cdot)$ is U -valued, then it is called an admissible nonrandomized policy. An admissible policy is called feedback if $v(\cdot)$ is progressively measurable with respect to the natural filtration of $(X(\cdot), S(\cdot))$. A particular subclass of feedback policies is of special interest. A feedback policy $v(\cdot)$ is called a (homogeneous) Markov policy if $v(\cdot) = \tilde{v}(X(\cdot), S(\cdot))$ for a measurable map $\tilde{v} : \mathbf{R}^d \times \mathcal{S} \rightarrow \mathcal{V}$. With an abuse in notation the map \tilde{v} itself is called a Markov policy. Let Π , Π_M and Π_{MD} denote the sets of all admissible, Markov and Markov nonrandomized policies respectively.

If $(W(\cdot), p(\cdot, \cdot), X_0, S_0, v(\cdot))$ satisfying the above are given on a prescribed probability space (Ω, \mathcal{F}, P) , then under (A1), the equation (2.4) will admit an a.s. unique strong solution [11, Chap. 3], and $X(\cdot) \in C(\mathbf{R}_+; \mathbf{R}^d)$, $S(\cdot) \in D(\mathbf{R}_+; \mathcal{S})$, where $D(\mathbf{R}_+; \mathcal{S})$ is the space of right continuous functions on \mathbf{R}_+ with left limits taking values in \mathcal{S} . However, if $v(\cdot)$ is a feedback policy, then there exists a measurable map

$$f : \mathbf{R}_+ \times C(\mathbf{R}_+; \mathbf{R}^d) \times D(\mathbf{R}_+; \mathcal{S}) \rightarrow \mathcal{V}$$

such that for each $t \geq 0$, $v(t) = f(t, X(\cdot), S(\cdot))$ and is measurable with respect to the σ -field generated by $\{X(s), S(s), s \leq t\}$. Thus, $v(\cdot)$ cannot be specified a priori in (2.4). Instead, one has to replace $v(t)$ in (2.4) by $f(t, X(\cdot), S(\cdot))$, and (2.4) takes the form

$$\begin{aligned} dX(t) &= b(X(t), S(t), f(t, X(\cdot), S(\cdot))) dt + \sigma(X(t), S(t)) dW(t) \\ dS(t) &= \int_{\mathbf{R}} h(X(t), S(t-), f(t, X(\cdot), S(\cdot)), z) p(dt, dz) \end{aligned} \quad (2.5)$$

for $t \geq 0$ with $X(0) = X_0$, $S(0) = S_0$. In general, (2.5) will not even admit a weak solution. However, if the feedback policy is Markov, then the existence of a unique strong solution can be established. We now introduce some notation which will be used throughout the paper. Define $L^1(\mathbf{R}^d \times \mathcal{S}) = \{f : \mathbf{R}^d \times \mathcal{S} \rightarrow \mathbf{R} : \text{for each } i \in \mathcal{S}, f(\cdot, i) \in L^1(\mathbf{R}^d)\}$. $L^1(\mathbf{R}^d \times \mathcal{S})$ is endowed with the product topology of $(L^1(\mathbf{R}^d))^N$. Similarly, we define $C_0^\infty(\mathbf{R}^d \times \mathcal{S})$, $W_{loc}^{2,p}(\mathbf{R}^d \times \mathcal{S})$, etc. For $f \in W_{loc}^{2,p}(\mathbf{R}^d \times \mathcal{S})$, $u \in U$, we write

$$L^u f(x, k) = L_k^u f(x, k) + \sum_{j=1}^N \bar{\lambda}_{kj}(x, u) f(x, j), \quad (2.6)$$

where

$$\begin{aligned} L_k^u f(x, k) &= \frac{1}{2} \sum_{i, j, \ell=1}^d \sigma_{i\ell}(x, k) \sigma_{j\ell}(x, k) \frac{\partial^2 f(x, k)}{\partial x_i \partial x_j} \\ &\quad + \sum_{j=1}^d \bar{b}_j(x, k, u) \frac{\partial f(x, k)}{\partial x_j} \end{aligned} \quad (2.7)$$

and more generally, for $v \in \mathcal{V}$

$$L^v f(x, k) = \int_U L^u f(x, k) v(du). \quad (2.8)$$

The following result is proved in [10].

Theorem 2.1. Under a Markov policy v , (2.4) admits an a.s. unique strong solution such that $(X(\cdot), S(\cdot))$ is a Feller process with differential generator L^v .

A Markov policy v is called stable if the corresponding process $(X(\cdot), S(\cdot))$ is positive recurrent. In this case, the process will have a unique invariant probability measure, denoted by $\eta_v \in \mathcal{P}(\mathbf{R}^d \times \mathcal{S})$. The uniqueness of η_v is guaranteed by (A1) (ii)

and (iii). We assume that the set of stable Markov policies is nonempty.

The Optimization Problem. Let $\bar{c} : \mathbb{R}^d \times S \times U \rightarrow \mathbb{R}_+$ be the cost function. We make the following assumption on \bar{c} .

(A2) For each $i \in S$, $\bar{c}(\cdot, i, \cdot)$ is continuous.

We define $c : \mathbb{R}^d \times S \times \mathcal{V} \rightarrow \mathbb{R}_+$ by

$$c(x, i, v) = \int_U \bar{c}(x, i, u)v(du). \quad (2.9)$$

Let $v(\cdot)$ be an admissible policy and $(X(\cdot), S(\cdot))$ the corresponding process. The pathwise (long-run) average cost incurred under $v(\cdot)$ is

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T c(X(t), S(t), v(t)) dt. \quad (2.10)$$

We wish to a.s. minimize (2.10) over all admissible policies. Our goal is to establish the existence of a stable Markov policy which is a.s. optimal. We will carry out our study under two alternate sets of hypotheses: (a) a condition on the cost which penalizes unstable behavior, (b) a blanket stability condition which implies that all Markov policies are stable. We will describe these conditions in Section 4.

3. Recurrence and Ergodicity of Switching Diffusions. Due to the interaction between the continuous and discrete components, the study of recurrence and ergodicity of switching diffusions is quite involved. Let v be a Markov policy which will be fixed throughout this section unless explicitly mentioned otherwise. Let $P^v : \mathbb{R}_+ \times \mathbb{R}^d \times S \rightarrow \mathcal{P}(\mathbb{R}^d \times S)$ denote the transition function of the corresponding process $(X(\cdot), S(\cdot))$. Also $P_{x,i}^v$ and $E_{x,i}^v$ denote the probability measure and the expectation operator, respectively, on the canonical space, of the process $(X(\cdot), S(\cdot))$ starting at $(x, i) \in \mathbb{R}^d \times S$. The following result plays a crucial role in recurrence.

Lemma 3.1. For any $(t, x, i) \in \mathbb{R}_+ \times \mathbb{R}^d \times S$, the support of $P^v(t, x, i; \cdot)$ is $\mathbb{R}^d \times S$.

Let τ_{ii}, τ_j be the stopping times defined as follows.

$$\begin{aligned} \tau_{ii} &= \inf\{t > 0 : S(t) = i \text{ and } S(u) \neq i, \text{ for some } 0 < u < t\} \\ \tau_j &= \inf\{t > 0 : S(t) = j\}. \end{aligned}$$

Let $D \subset \mathbb{R}^d$ be a bounded open set. Define

$$\begin{aligned} \tau_{D,i} &= \inf\{t \geq 0 : (X(t), S(t)) \notin D \times \{i\}\} \\ \tau_D &= \inf\{t \geq 0 : X(t) \notin D\}. \end{aligned}$$

Using (3.2) and well known arguments in Markov processes [8, Vol. I, p. 111] the following results can be proved.

Lemma 3.2. Under (A1), the following holds.

- (i) $\sup_{v \in \Pi_M, x \in \mathbb{R}^d} E_{x,i}^v \tau_{ii} < \infty$.
- (ii) $\sup_{v \in \Pi_M, x \in \mathbb{R}^d} E_{x,i}^v \tau_j < \infty$.
- (iii) $\sup_{v \in \Pi_M, x \in \mathbb{R}^d} E_{x,i}^v \tau_{D,i} < \infty$.
- (iv) $\sup_{v \in \Pi_M, x \in \mathbb{R}^d} E_{x,i}^v \tau_D < \infty$.

It is well known that the properties of harmonic functions of a Markov process play an important role in recurrence and ergodicity [2]. Therefore we will study some properties of the harmonic functions of the process $(X(\cdot), S(\cdot))$ under the Markov policy v . Let $D \subset \mathbb{R}^d$ be an open set and $f : D \times S \rightarrow \mathbb{R}$. The function f is called L^v -harmonic on D if it is bounded on compact subsets of D , and for all $x \in D, i \in S$,

$$f(x, i) = E_{x,i}^v f(X(\tau_{V,i}), S(\tau_{V,i})) \quad (3.1)$$

for every neighborhood V of x having compact closure \bar{V} in D . It is clear that if f is L^v -harmonic then

$$f(x, i) = E_{x,i}^v f(X(\tau_V), S(\tau_V)). \quad (3.2)$$

Lemma 3.3. Let $D \subset \mathbb{R}^d$ be open. Then, under (A1):

- (i) Every L^v -harmonic function in D is continuous in D .
- (ii) If $L^v f = 0, f \in W^{2,p}(D \times S)$, then f is L^v -harmonic. Conversely, if f is L^v -harmonic and $f \in W_{loc}^{2,p}(D \times S)$, then $L^v f = 0$ in D .
- (iii) (Maximum Principle) Let D be connected and $f \geq 0$ and L^v -harmonic in D . Then f is either strictly positive on $D \times S$ or identically zero.

Remark 3.1. (i) The condition $\bar{\lambda}_{ij}(x) \geq \lambda_0$ plays a crucial role in the above results. For example, the maximum principle does not hold without this, as the following counterexample shows. Let $d = 1, S = \{1, 2\}, \bar{\lambda}_{11}(\cdot) \equiv 0, \bar{\lambda}_{21}(\cdot) \equiv 1, \bar{b} = 0, \sigma = \sqrt{2}, f(x, 1) \equiv 0, f(x, 2) = \cosh x$. Then it is easily verified that $L^v f(x, i) \equiv 0, i = 1, 2$.

(ii) To the best of our knowledge, this maximum principle is not known in the literature on partial differential equations.

We next establish Harnack's inequality for L^v -harmonic functions, which is a very important result in partial differential equations. This result will not play any role in recurrence or ergodicity, but it will be crucial in deriving a certain estimate in Section 5.

Theorem 3.1. Let $\Omega \subset \mathbb{R}^d$ be a bounded open domain and D an open set whose closure is a compact subset of Ω . Let $f \geq 0$ and $L^v f = 0$ in $W^{2,p}(\Omega \times S)$. Then for any $x, y \in D$ and $i, j \in S$, we have

$$f(x, i) \leq C f(y, j),$$

where C is a constant which depends only on d, N , the diameter of D , the Hausdorff distance between $\bar{D}, \bar{\Omega}$, the bounds on $\bar{b}, \sigma, \bar{\lambda}$ in (A1), the ellipticity constant of $\sigma\sigma'$, and is independent of the L^v -harmonic function f or the Markov policy v .

Remark 3.2. The counterexample in Remark 3.1 again shows that Theorem 3.1 does not hold if we drop the condition $\bar{\lambda}_{ij} \geq \lambda_0 > 0$.

We now discuss the recurrence properties of switching diffusions. Our treatment closely follows [2]. The point $(x, i) \in \mathbb{R}^d \times S$ is said to be recurrent if given any $\epsilon > 0$,

$$P_{x,i}^v(X(t_n) \in B(x, \epsilon), S(t_n) = i, \text{ for a sequence } t_n \uparrow \infty) = 1.$$

A point (x, i) is transient if

$$P_{x,i}^v(|X(t)| \rightarrow \infty, \text{ as } t \rightarrow \infty) = 1.$$

If all points of the switching diffusion are recurrent, then it is called recurrent. A transient switching diffusion is similarly defined. Under (A1) (ii), (iii), we show that a switching diffusion is either recurrent or transient.

Lemma 3.4. Under (A1), the following statements are equivalent. (i) The switching diffusion is recurrent. (ii) $P_{x,i}^v(X(t) \in D, \text{ for some } t \geq 0) = 1$, for all $x \in \mathbb{R}^d, i \in S$ and any non-empty open set D . (iii) There exists a compact set $K \subset \mathbb{R}^d$ such that $P_{x,i}^v(X(t) \in K, \text{ for some } t \geq 0) = 1$ for all $(x, i) \in \mathbb{R}^d \times S$. (iv) $P_{x,i}^v(X(t_n) \in D, \text{ for a sequence } t_n \uparrow \infty) = 1$, for all $x \in \mathbb{R}^d, i \in S$ and any non-empty open set D . (v) There exists a point $z \in \mathbb{R}^d$, a pair of numbers $r_0, r_1, 0 < r_0 < r_1$, and a $y \in \partial B(z, r_1)$ such that $P_{y,i}^v(\tau_{\bar{B}(z, r_0)} < \infty) = 1$, for any $i \in S$.

Theorem 3.2. Under (A1), for any Markov policy, the switching diffusion is either recurrent or transient.

A recurrent switching diffusion will admit a unique (up to a constant multiple) σ -finite invariant measure. The switching diffusion is called *positive recurrent* if it is recurrent and admits a finite invariant measure.

A Markov policy v is called *stable* if the corresponding process is positive recurrent; the corresponding invariant probability measure is denoted by η_v .

Theorem 3.3. Let z, r_0, r_1 be as in Lemma 3.4(v). Then under (A1), the switching diffusion is positive recurrent if

$$\sup_{y \in \partial B(z, r_1), i \in \mathcal{S}} E_{y,i}^v [\tau_{\bar{B}(z, r_0)}^c] < \infty. \quad (3.3)$$

Note that it may be very difficult to verify (3.3) for general b, σ, λ . One usually verifies (3.3) by constructing a Liapunov function [2]. For switching diffusions such a construction seems difficult, since it involves solving a system of ordinary differential equations in closed form. However, we present some criteria for positive recurrence and discuss some implications.

(A3) There exists a $w \in C^2(\mathbb{R}^d \times \mathcal{S})$, $w \geq 0$ such that $L^v w(x, i) \rightarrow \infty$ as $|x| \rightarrow \infty$ for each i , $E_{x,i}^v w(X(t), S(t))$ and $E_{x,i}^v |L^v w(X(t), S(t))|$ are locally bounded, and

$$L^v w(x, i) \leq p - qw(x, i)$$

for some $p > 0$ and $q > 0$.

Theorem 3.4. Under (A1) and (A3), the process $(X(\cdot), S(\cdot))$ under the Markov policy v is positive recurrent.

(A4) There exists a C^2 function $w : \mathbb{R}^d \times \mathcal{S} \rightarrow \mathbb{R}_+$ such that

(i) $\lim_{|x| \rightarrow \infty} w(x, i) = +\infty$ uniformly in $|x|$.

(ii) There exists $a > 0, \varepsilon > 0$ such that for $|x| > a$, $L^u w(x, i) < -\varepsilon$, for all $u \in U, i \in \mathcal{S}$, and $|\nabla w(x, i)|^2 \geq \mu^{-1}$, where μ is the constant in (A1) (iii).

(iii) $w(x, i)$ and $|\nabla w(x, i)|$ have polynomial growth. The proof of [4, Lemma 6.2.2, p. 150] can be closely paralleled to yield the following.

Theorem 3.5. Under (A1) and (A4), the process $(X(\cdot), S(\cdot))$ under any Markov policy v is positive recurrent. Thus, all Markov policies are stable.

4. Existence of an Optimal Policy. In this section we will establish the existence of a stable Markov nonrandomized policy under certain conditions. We will follow the methodology developed in [4], [5], [6], [7] for controlled diffusions. For switching diffusions, similar techniques carry through with some extra technical details.

Let Π_{SM} and Π_{SMD} denote the set of stable Markov and stable Markov nonrandomized policies respectively. Since we look for an optimal policy in Π_{SMD} , it is natural to assume that it is nonempty. Let $v \in \Pi_{SMD}$. Then

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T c(X(s), S(s), v(X(s), S(s))) ds \\ = \sum_i \int_{\mathbb{R}^d} c(x, i, v(x, i)) \eta_v(dx, i) := \rho_v \quad \text{a.s.} \end{aligned} \quad (4.1)$$

Let

$$\rho^* := \inf_{v \in \Pi_{SM}} \rho_v. \quad (4.2)$$

We assume that $\rho^* < \infty$. Consider the following conditions.

(A5) Assume that for each $i \in \mathcal{S}$,

$$\liminf_{|x| \rightarrow \infty} \inf_{u \in U} \bar{c}(x, i, u) > \rho^*.$$

(A6) There exists a $w \in C^2(\mathbb{R}^d \times \mathcal{S})$, $w \geq 0$ such that

(i) $w(x, i) \rightarrow \infty$ as $|x| \rightarrow \infty$ uniformly in $|x|$ for each i .

(ii) For each $v \in \Pi_M$, $E_{x,i}^v w(X(t), S(t))$

and $E_{x,i}^v |L^v w(X(t), S(t))|$ are locally bounded.

(iii) There exists $p > 0, q > 0$ such that $L^u w(x, i) \leq p - qw(x, i)$ for each $u \in U$.

Our main result is the following.

Theorem 4.1. Let (A1), (A2) hold. Under any one of the conditions (A4), (A5) or (A6), there exists a $v^* \in \Pi_{SMD}$ which is a.s. optimal.

5. Hamilton-Jacobi-Bellman Equations. In this section, we will study the HJB equations and characterize the optimal policy in terms of their solution. We will work under the following assumption.

(A7) The cost function \bar{c} is bounded, continuous and Lipschitz in its first argument uniformly with respect to the third.

We will follow the vanishing discount approach, i.e., derive the HJB equations for the ergodic criterion as a vanishing limit of the HJB equations for the discounted criterion as the discount factor approaches zero. The results follow those of [6]. However, they differ in important technical details. For $\alpha > 0, x \in \mathbb{R}^d, i \in \mathcal{S}$, let $V_\alpha(x, i)$ denote the discounted value function with discount factor $\alpha > 0$, i.e.,

$$V_\alpha(x, i) = \inf_{u \in U} E_{x,i}^u \left[\int_0^\infty e^{-\alpha t} c(X(t), S(t), u(t)) dt \right]. \quad (5.1)$$

The following result is proved in [10].

Theorem 5.1. Under (A1), (A7), V_α is the unique solution in $C^2(\mathbb{R}^d \times \mathcal{S}) \cap C_b(\mathbb{R}^d \times \mathcal{S})$ of

$$\inf_{u \in U} \{ L^u V_\alpha(x, i) + \bar{c}(x, i, u) \} = \alpha V_\alpha(x, i). \quad (5.2)$$

For $i \in \mathcal{S}$, set

$$\begin{aligned} G_i &= \{ x \in \mathbb{R}^d : \inf_{u \in U} \bar{c}(x, i, u) \leq \rho^* \} \\ G &= \bigcup_{i=1}^N G_i \end{aligned} \quad (5.3)$$

By (A5) and (A7), G is compact.

The following result plays a very crucial role.

Lemma 5.1. Under (A1), (A5), (A7), there exists a $\alpha_0 \in (0, 1)$ such that for $\alpha \in (0, \alpha_0]$, $\min_{x,i} V_\alpha(x, i)$ is attained on the set G as defined in (5.3). Also, $|V_\alpha(x, i) - V_\alpha(y, j)|$ is bounded on compacta, uniformly over $\alpha \in (0, \alpha_0)$.

Theorem 5.2. Under (A1), (A5), (A7), there exists a function $V : \mathbb{R}^d \times \mathcal{S} \rightarrow \mathbb{R}$ and a scalar $\rho \in \mathbb{R}$ such that for some fixed $i_0 \in \mathcal{S}$,

$$\rho \leq \rho^*, \quad V(0, i_0) = 0, \quad \inf_x V(x, i) > -\infty \quad (5.4)$$

$V \in C^2(\mathbb{R}^d \times \mathcal{S})$ and (V, ρ) satisfies the HJB equations given by

$$\inf_{u \in U} [L^u V(x, i) + \bar{c}(x, i, u)] = \rho. \quad (5.5)$$

Based on Lemma 5.1 and Theorem 3.1, the following results can be derived by closely following the methods in [6].

Lemma 5.2. Assume (A1), (A5), (A7). Let $v \in \Pi_{MD}$ be such that for each i

$$\begin{aligned} \inf_{u \in U} & \left[\sum_{j=1}^d \bar{b}_j(x, i, u) \frac{\partial V(x, i)}{\partial x_j} + \sum_{j=1}^N \bar{\lambda}_{ij}(x, u) V(x, j) + \bar{c}(x, i, u) \right] \\ & = \sum_{j=1}^d \bar{b}_j(x, i, v(x, i)) \frac{\partial V(x, i)}{\partial x_j} \\ & \quad + \sum_{j=1}^N \bar{\lambda}_{ij}(x, v(x, i)) V(x, j) + \bar{c}(x, i, v(x, i)) \quad \text{a.e.} \end{aligned} \quad (5.6)$$

Then $v \in \Pi_{SMD}$. The scalar ρ in (5.5) equals ρ^* and v in (5.6) is a.s. optimal.

Theorem 5.3. Among all pairs $(\phi, \rho) \in W_{loc}^{2,p}(\mathbb{R}^d \times S) \times \mathbb{R}$, $2 \leq p < \infty$, satisfying (5.5), (V, ρ^*) , where V is the function in Theorem 5.2, is the unique one satisfying the properties (5.4).

Theorem 5.4. Under (A1), (A5), (A7), $v \in \Pi_{SMD}$ is a.s. optimal if and only if (5.6) holds.

Remark 5.1. The bounded condition on the cost function \bar{c} can be dropped by mimicking the arguments in [6, p. 202].

We will now study the HJB equation under (A1), (A6) and (A7). Recall that under (A1), (A6), $\Pi_M = \Pi_{SM}$. We say that a function $f: \mathbb{R}^d \times S \rightarrow \mathbb{R}$ is in the class $O(w)$ if for each $i \in S$

$$\limsup_{|x| \rightarrow \infty} |f(x, i)|/w(x, i) < \infty. \quad (5.7)$$

Theorem 5.5. Under (A1), (A6) and (A7), the equation (5.5) admits a unique solution (V, ρ) in the class $W_{loc}^{2,p}(\mathbb{R}^d \times S) \cap O(w)$, $2 \leq p < \infty$ satisfying $V(0, i_0) = 0$ for a fixed $i_0 \in S$.

Remark 5.2. (i) The statement of Theorem 5.5 holds under (A1), (A4) and (A7). (ii) For the stable case we have carried out our analysis under the Liapunov condition (A6). Analogous results can be derived under the weaker condition (A4).

6. An Application to a Manufacturing Model. We now use the results of the previous sections to analyze the manufacturing model studied in [1], [3], [10]. Suppose there is one machine producing a single commodity. We assume that the demand rate is a constant $d > 0$. Let the machine state $S(t)$ take values in $\{0, 1\}$, $S(t) = 0$ or 1 , according as the machine is down or functional. Let $S(t)$ be a continuous time Markov chain with generator

$$\begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix}$$

where $\lambda_0 > 0$, $\lambda_1 > 0$, λ_0 and λ_1 are infinitesimal repair and failure rates respectively. The inventory $X(t)$ is governed by the Ito equation

$$dX(t) = (u(t) - d)dt + \sigma dW(t) \quad (6.1)$$

where $\sigma > 0$, $u(t)$ is the production rate, $W(t)$ is a one-dimensional Wiener process independent of $S(t)$. The last term in (6.1) can be interpreted as "sales return", "inventory spoilage", "sudden demand fluctuations", etc. A negative value of $X(t)$ represents backlogged demand. The production rate is constrained by

$$u(t) = \begin{cases} 0 & \text{if } S(t) = 0 \\ \in [0, r] & \text{if } S(t) = 1. \end{cases}$$

Let $c: \mathbb{R} \rightarrow \mathbb{R}_+$ be the cost function which is assumed to be convex and Lipschitz. Also $c(x) = c'(|x|)$ for some $c': \mathbb{R}_+ \rightarrow \mathbb{R}_+$ which is increasing. Thus, c satisfies (A5). We will show that a

certain hedging point policy is stable. Therefore, by the results of Section 4 there exists an a.s. optimal Markov nonrandomized policy for the cost criterion

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T c(X(t)) dt.$$

The HJB equations in this case are

$$\begin{aligned} & \left(\frac{\sigma^2}{2} V''(x, 0) - dV'(x, 0) \right) \\ & \left(\frac{\sigma^2}{2} V''(x, 1) - \min_{u \in [0, r]} \{ (u - d)V'(x, 1) \} \right) \\ & + \begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix} \begin{pmatrix} V(x, 0) \\ V(x, 1) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} c(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rho. \end{aligned} \quad (6.2)$$

The results of Section 5 ensure existence of a C^2 solution (V, ρ^*) of (6.2), where ρ^* is the optimal cost. Using the convexity of $c(\cdot)$, it can be shown that $V(\cdot, i)$ is convex for each i . Hence there exists an x^* such that

$$\begin{aligned} V'(x, 1) & \leq 0 & \text{for } x \leq x^* \\ & \geq 0 & \text{for } x \geq x^*. \end{aligned} \quad (6.3)$$

From (6.2), it follows that the value of u which minimizes $(u - d)V'(x, 1)$ is

$$u = \begin{cases} r & \text{if } x \leq x^* \\ 0 & \text{if } x \geq x^*. \end{cases}$$

At $x = x^*$, $V'(x^*, 1) = 0$ and therefore any $u \in [0, r]$ minimizes $(u - d)V'(x, 1)$. Thus, in view of Theorem 5.4, we can choose any $u \in [0, r]$ at $x = x^*$. To be specific, we choose $u = d$ at x^* , i.e., we just produce to meet the demand exactly. Thus, the following $v^* \in \Pi_{SMD}$ is optimal

$$v^*(x, 0) = 0, \quad v^*(x, 1) = \begin{cases} r & \text{if } x < x^* \\ d & \text{if } x = x^* \\ 0 & \text{if } x > x^*. \end{cases} \quad (6.4)$$

The stability of the policy (6.4) follows from Lemma 5.2 provided we show that Π_{SM} is nonempty. We show that the zero-inventory policy v given by

$$v(x, 0) = 0, \quad v(x, 1) = \begin{cases} r & \text{if } x \leq 0 \\ 0 & \text{if } x > x^*. \end{cases} \quad (6.5)$$

is stable if and only if

$$\frac{(r - d)}{\lambda_1} > \frac{d}{\lambda_0}. \quad (6.6)$$

The condition (6.6) is very appealing from an intuitive point of view. Note that λ_0^{-1} and λ_1^{-1} are mean sojourn time of the chain in states 0 and 1 respectively. In state 0 the mean inventory depletes at a rate d while in state 1 it builds up at the rate $(r - d)$. Thus, if (6.6) is satisfied, one would expect the zero-inventory policy to stabilize the system. Our analysis confirms this intuition. We first show that under v the process $(X(\cdot), S(\cdot))$ has an invariant measure η_v with a strictly positive "density-mass". In view of Lemma 3.1, it would follow from the ergodic theory of Markov processes [13, Chap. 1] that $(X(\cdot), S(\cdot))$ is positive recurrent. Thus, v would be stable. To this end, we attempt to solve the adjoint system

$$\begin{aligned} (L^v)^* \phi(x, i) & = 0 \\ \phi(x, i) & > 0, \quad \sum_i \int_{\mathbb{R}} \phi(x, i) dx = 1. \end{aligned} \quad (6.7)$$

Without any loss of generality, assume that $\sigma = \sqrt{2}$. Then for $x > 0$,

$$\begin{aligned}\phi''(x, 0) + d\phi'(x, 0) - \lambda_0\phi(x, 0) + \lambda_1\phi(x, 1) &= 0 \\ \phi''(x, 1) + d\phi'(x, 1) - \lambda_1\phi(x, 1) + \lambda_0\phi(x, 0) &= 0\end{aligned}\quad (6.8)$$

The solutions of (6.7) in $L^1[0, \infty)$ take the parameterized form

$$\phi^+(x) = \begin{pmatrix} \phi^+(x, 0) \\ \phi^+(x, 1) \end{pmatrix} = \alpha \begin{pmatrix} \lambda_1 \\ \lambda_0 \end{pmatrix} e^{-s_1 x} + \beta \begin{pmatrix} -\lambda_1 \\ \lambda_0 \end{pmatrix} e^{-s_2 x}, \quad (6.9)$$

$\alpha, \beta \in \mathbb{R}$, $s_1 = d$, $s_2 = (d + \sqrt{d^2 + 4\lambda_0})/2$. For $x < 0$,

$$\begin{aligned}\phi''(x, 0) + d\phi'(x, 0) - \lambda_0\phi(x, 0) + \lambda_1\phi(x, 1) &= 0 \\ \phi''(x, 1) - (r-d)\phi'(x, 1) - \lambda_1\phi(x, 1) + \lambda_0\phi(x, 0) &= 0\end{aligned}\quad (6.10)$$

Under (6.6), all the solutions of (6.10) in $L^1(-\infty, 0]$ can be parameterized by

$$\phi^-(x) = \begin{pmatrix} \phi^-(x, 0) \\ \phi^-(x, 1) \end{pmatrix} = \gamma \begin{pmatrix} \lambda_1 \\ -\psi(s_3) \end{pmatrix} e^{s_3 x} + \delta \begin{pmatrix} -\lambda_1 \\ \psi(s_4) \end{pmatrix} e^{s_4 x}, \quad (6.11)$$

where $\gamma, \delta \in \mathbb{R}$, $\psi(s) = s^2 + ds - \lambda_0$ and s_3, s_4 are the positive roots of the polynomial

$$s^3 - (r-2d)s^2 - [(r-d)d + \lambda_0 + \lambda_1]s + [r\lambda_0 - d(\lambda_1 + \lambda_0)] \quad (6.12)$$

ordered by $0 < s_3 < s_4$. It can be verified that $\psi(s_3) < 0$, $\psi(s_4) > 0$. We need to satisfy

$$\phi^+(0) = \phi^-(0) \quad (6.13)$$

and

$$\int_0^\infty \phi^+(x) dx + \int_{-\infty}^0 \phi^-(x) dx = \frac{1}{\lambda_0 + \lambda_1} \begin{pmatrix} \lambda_1 \\ \lambda_0 \end{pmatrix}. \quad (6.14)$$

(6.13) is simply the continuity requirement, whereas (6.14) should hold since $S(t)$ has a unique invariant probability measure π which satisfies

$$\pi(\{0\}) = \frac{\lambda_1}{\lambda_0 + \lambda_1}, \quad \pi(\{1\}) = \frac{\lambda_0}{\lambda_0 + \lambda_1}.$$

The conditions (6.13), (6.14) are equivalent to the set of linear equations

$$\alpha - \beta = \gamma - \delta \quad (6.15a)$$

$$\lambda_0\alpha + \lambda_1\beta = -\gamma\psi(s_3) + \delta\psi(s_4) \quad (6.15b)$$

$$\lambda_1 \left(\frac{\alpha}{s_1} - \frac{\beta}{s_2} + \frac{\gamma}{s_3} - \frac{\delta}{s_4} \right) = \frac{\lambda_1}{\lambda_0 + \lambda_1} \quad (6.15c)$$

$$\frac{\alpha\lambda_0}{s_1} + \frac{\beta\lambda_1}{s_2} - \frac{\gamma\psi(s_3)}{s_3} + \frac{\delta\psi(s_4)}{s_4} = \frac{\lambda_0}{\lambda_0 + \lambda_1}. \quad (6.15d)$$

Solving (6.15a)–(6.15d), we obtain a unique solution $(\alpha, \beta, \gamma, \delta)$ satisfying

$$0 < \beta < \alpha, \quad 0 < \delta < \gamma. \quad (6.16)$$

It follows from (6.16) and the foregoing that

$$\phi^+(x) > 0, \quad \text{for } x \geq 0 \quad \text{and} \quad \phi^-(x) > 0, \quad \text{for } x < 0.$$

Note that if $r\lambda_0 \leq d(\lambda_1 + \lambda_0)$, then $s_3 = 0$ and there exists no positive solution $\phi^-(x)$ of (6.10) in $L^1(-\infty, 0]$.

In [3], Bielecki and Kumar have studied the mean square stability of the piecewise deterministic system, i.e., (6.1) with $\sigma = 0$. They have shown that for

$$\frac{r-d}{\lambda_1} \geq \frac{d}{\lambda_0}$$

the policy (6.5) is mean square stable. Our analysis shows that additive noise in (6.1) retains the stability of the zero-inventory policy as long as strict inequality holds in the above.

Acknowledgement. The authors wish to thank Prof. S. R. S. Varadhan for explaining to us the work of Krylov and Safonov on Harnack's inequality.

REFERENCES

1. R. Akella and P. R. Kumar, *Optimal control of production rate in a failure prone manufacturing system*, IEEE Trans. Automat. Control **AC-31** (1986), 116–126.
2. R. N. Bhattacharya, *Criteria for recurrence and existence of invariant measures for multidimensional diffusions*, Annals of Probability **6** (1978), 541–553.
3. T. Bielecki and P. R. Kumar, *Optimality of zero-inventory policies for unreliable manufacturing systems*, Oper. Res. **36** (1988), 532–546.
4. V. S. Borkar, *Optimal Control of Diffusion Processes*, Pitman Research Notes in Math. Series, No. 203, Longman, Harlow, UK, 1989.
5. V. S. Borkar and M. K. Ghosh, *Ergodic control of multidimensional diffusions, I: The existence results*, SIAM J. Control Optim. **26** (1988), 112–126.
6. ———, *Ergodic control of multidimensional diffusions, II: Adaptive control*, Appl. Math. Optim. **21** (1990), 191–220.
7. ———, *Controlled diffusions with constraints*, J. Math. Anal. Appl. **152** (1990), 88–108.
8. E. B. Dynkin, *Markov Processes Vols. I and II*, Springer-Verlag, New York, 1965.
9. M. K. Ghosh, A. Arapostathis and S. I. Marcus, *An optimal control problem arising in flexible manufacturing systems*, Proc. 30th IEEE Conf. on Decision and Control, Brighton, England (1991), 1844–1849.
10. ———, *Optimal control of switching diffusions with application to flexible manufacturing systems*, SIAM J. Control Optim. (to appear).