JULY 1986 LIDS-P-1583

CONTROL OF MARKOV CHAINS WITH LONG-RUN AVERAGE COST CRITERION

by

Vivek S. Borkar

Tata Inst. of Fundamental Research
Bangalore Centre
P.O. Box 1234, I.I.Sc. Campus
Bangalore 560012, INDIA

Abstract

The long-run average cost control problem for discrete time Markov chains is studied in an extremely general framework. Existence of stable stationary strategies which are optimal in the appropriate sense is established and these are characterized via the dynamic programming equations. The approach here differs from the conventional approach via the discounted cost problem and covers situations not covered by the latter.

Key Words: Markov chains, optimal control, long-run average cost, stationary strategy, stable strategy, dynamic programming

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

I. INTRODUCTION

The aim of this paper is to provide a new framework for the long-run average cost control problem for discrete time Markov chains with a countable state space. The tranditional approach to this problem has been to treat it as a limiting case of the discounted cost control problem as the discount factor approaches unity. (See [13] for a succinct account and the bibliographies of [1], [13], [14] for further references. [1] contains some major recent extensions of this approach.) However, this limiting argument needs a strong stability condition, various forms of which are used in the literature [8]. This condition fails in many important applications such as control of queueing networks. A concrete example of the failure of the classical argument was provided in [12]. Motivated by these problems, [3], [4] used an alternative approach to tackle a special class of Markov chains, viz., those exhibiting a 'nearest-neighbour motion'. More precisely, the hypotheses were that the chain moves from any state to at most finitely many neighbouring states and the length of the shortest path from a state i to a prescribed state is unbounded as a function of i. Two cases were considered: the first being the case when the cost function penalizes unstable behaviour and the second the case when there is no such restriction on cost, but the stationary strategies satisfy a stability condition. The aim in both cases was to establish the existence of a stable optimal stationary strategy and to characterize it in terms of the dynamic programming equations. Reasonably complete results were established in the first case and comparatively weaker results in the second. This paper considers a very general framework that subsumes both the paradigm of [3],

[4] and that of [13].

The paper relies heavily on parts of [4]. In order to avoid excessive overlap with [4] and to keep the present paper from getting too unwieldy, we refer to [4] for a great many details. In view of this, we use the same notation as [4]. This notation is recalled in section II. Section III gives a necessary and sufficient condition for the tightness of invariant probabilities under stationary strategies. Section IV gives a dynamic programming characterization of a stationary strategy which gives less cost than any other stationary strategy. Section V and VI study the stability and statistical behaviour of a chain governed by an arbitrary strategy under the conditions spelt out in section III. Section VII establishes the existence of a stable stationary strategy which is optimal under various definitions of optimality. Section VIII considers the case when at least one stationary strategy is not stable, but the cost function penalizes unstable behaviour. Section IX gives simple, more easily verifiable sufficient conditions for the stability condition of section III to hold. Finally, section X indicates how to extend all these results to control-dependent cost functions and concludes with a few plausible conjectures.

II. NOTATION AND PRELIMINARIES

Let X_n, $n=1,2,\ldots$, be a controlled Markov chain on state space $S = \{1,2,\ldots\}$ with transition matrix $P_u = [[p(i,j,u_i)]]$, $i, j \varepsilon S$, where $u = [u_1,u_2,\ldots]$ is the control vector satisfying $u_i \varepsilon D(i)$ for some prescribed compact metric spaces $D(i)$. The functions $p(i,j,.)$ are assumed to be continuous. Let $L = \Pi D(i)$ with the product topology. A control strategy $C(S)$ is a sequence of L-valued random variables $\{\xi_n\}$, $\xi_n = [\xi_n(1), \xi_n(2),\ldots]$, such that for all $i \varepsilon S$, $n \geq 1$,

$$P(X_{n+1} = i/X_m, \xi_m, m \leq n) = p(X_n, i, \xi_n(X_n)).$$

As noted in [4], there is no loss of generality if we assume that (a) the $D(i)$'s are identical or (b) the law of ξ_n is the product of the individual laws of $\xi_n(i)$, $i \varepsilon S$, for each n. If $\{\xi_n\}$ are i.i.d. with a common law Φ, call it a stationary randomized strategy (SRS), denoted by $\gamma[\Phi]$. If Φ is a Dirac measure at some $\xi \varepsilon L$, call it a stationary strategy (SS), denoted by $\gamma\{\xi\}$. The corresponding transition matrices are denoted by $P[\Phi] = [[\int p(i,j,u)\Phi_i(du)]]$ (Φ_i being the image of Φ under the projection $L \to D(i)$) and $P\{\xi\} = P_\xi$. The expectations under the corresponding laws of $\{X_n\}$ are denoted by $E_\Phi[\cdot]$, $E_\xi[\cdot]$ respectively (with the initial law either arbitrary or inferred from the context).

We shall assume that the chain has a single communicating class under all SRS. If in addition it is positive recurrent under some $\gamma[\Phi]$ or $\gamma\{\xi\}$, call the latter a stable SRS (SSRS) or a stable SS (SSS) respectively and denote by $\pi[\Phi] = [\pi[\Phi](1), \pi[\Phi](2),\ldots]$ or $\pi\{\xi\} = [\pi\{\xi\}(1), \pi\{\xi\}(2),\ldots]$ the corresponding unique invariant probability measures. For $f:S \to R$ bounded, let $C_f[\Phi] = \sum \pi[\Phi](i)f(i)$ and $C_f\{\xi\} = \sum \pi\{\xi\}(i)f(i)$.

Let $k:S \to R$ be a bounded cost function. Define

$$\varphi_n = \frac{1}{n} \sum_{m=1}^{n} k(X_m)$$

$$\varphi_\infty = \liminf_{r \to \infty} \varphi_n$$

Under an SSRS $\gamma[\Phi]$ or an SSS $\gamma\{\xi\}$, we have $\varphi_n \to C_k[\Phi]$ (or $C_k\{\xi\}$ resp.) a.s. Let $\alpha = \inf C_k\{\xi\}$, $\beta = \inf C_k[\Phi]$, the infima being over all SSS (SSRS resp.). Clearly, $\beta \leq \alpha$. We say that an SSS $\gamma\{\xi_0\}$ is

(i) optimal in the mean if $C_k\{\xi_0\} = \alpha = \beta$ and under any CS,
$\liminf_{n \to \infty} E[\varphi_n] \geq \alpha$,

(ii) optimal in probability if $C_k\{\xi_0\} = \alpha = \beta$ and under any CS,
$\lim_{n \to \infty} P(\varphi_n \geq \alpha - \varepsilon) = 0$ for all $\varepsilon > 0$,

(iii) a.s. optimal if $C_k\{\xi_0\} = \alpha = \beta$ and under any CS,
$\varphi_\infty \geq \alpha$ a.s.

Next, we summarize the relevant results of [3], [4], [13]. Call k almost monotone if $\liminf k(i) > \beta$. Let $\tau = \min\{n>1 | X_n = 1\}$ and define $V\{\xi\} = [V\{\xi\}(1), V\{\xi\}(2), \ldots]$ by

$$V\{\xi\}(i) = E_\xi[\sum_{m=2}^{\tau} (k(X_m) - \alpha)/X_j = i]$$

for each SSS $\gamma\{\xi\}$. Define $V[\Phi] = [V[\Phi](1), V[\Phi](2),\ldots]$ analogously for an SSRS $\gamma[\Phi]$. Since $E[\tau/X_1 = 1] \lesssim \infty$ under positive recurrence, these are well-defined. The relevant results of [3], [4], [13] are as follows:

(1) Suppose all SS are SSS and

$$\sup_{i,\xi} E_\xi[\tau/X_1 = i] < \infty. \tag{2.1}$$

(see [8] for equivalent conditions.) Then there exists an a.s. optimal SSS $\gamma\{\xi_0\}$. Furthermore, $\gamma\{\xi_0\}$ is a.s. optimal if and only if the following holds termwise:

$$C_k\{\xi_0\}1_c = (P\{\xi_0\} - U)V\{\xi_0\} + Q = \min_u(P_u - U)V(\xi_0\} + Q \tag{2.2}$$

where $1_c = [1,2,\ldots]^T$, $Q = [k(1), k(2), \ldots]^T$ and U is the infinite identity matrix [13].

(2) Suppose that at lest one SSS exists, k is almost monotone and (*) below holds:

(*) For each i, $p(i,j,.) \equiv 0$ for all but finitely many j and for any finite subset A of S and $M \gtrsim 1$, there exists an $N \gtrsim 1$ such that whenever $i \gtrsim N$, the length of the shortest path from i to any state in A exceeds M under any SRS.

Then there exists an a.s. optimal SSS. Also, $\gamma\{\xi\}$ is an a.s. optimal SSS if any only if (2.2) holds. [3], [4].

(3) Suppose all SS are SSS, the set $\{\pi\{\xi\} | \xi \in L\}$ is tight and (*) holds. Then there exists an SSS which is optimal in probability. Also, $\gamma\{\xi_0\}$ is optimal in probability if and only if (2.2) holds. [4]

In the sections to follow, we generalize all three cases above. In the last section, we also indicate how to extend our results to control-dependent cost functions.

The main results of this paper are contained in Lemma 4.4, Theorems 7.1, 7.2, 7.3, 8.1 and 8.2. The key hypothesis is condition C of section III for Lemma 4.4, Theorem 7.1 and 7.3, condition D of section VII for Theorem 7.2 and condition E of section VIII for Theorem 8.2.

III. TIGHTNESS OF THE INVARIANT PROBABILITIES.

Throughout sections III–VII, we shall assume that all SS are SSS. Consider the condition:

Condition C: The family $\{\tau(\xi)\}$ of random variables is uniformly integrable, where $\tau(\xi) = \tau$ corresponding to the chain governed by $\gamma\{\xi\}$ with initial condition $X_1 = 1$.

Our aim is to prove that this condition is necessary and sufficient for the tightness of $\{\pi\{\xi\} | \xi \varepsilon L\}$. We shall proceed via several lemmas.

Lemma 3.1. A sequence X_n, n=1,2,..., of S-valued random variables is a Markov chain with transition matrix P_ξ if and only if for all bounded $f:S \to R$, the sequence Y_n, n = 1,2,..., defined by

$$Y_n = \sum_{m=1}^{n} (f(X_m) - \sum_{i \varepsilon S} p(X_{m-1}, i, \xi(X_{m-1}))f(i)) \tag{3.1}$$

is a martingale with respect to the natural filtration of $\{X_n\}$. (We use the convention $p(X_0, i, \xi(X_0)) = P(X_1=i)$.)

The proof is elementary. Let $p^n(i,j,\xi)$ denote the probability of going from i to j in n transitions under $\gamma\{\xi\}$.

Lemma 3.2. For each i and n, the family of probability measures $\{p^n(i,.,\xi), \xi \varepsilon L\}$ on S is tight.

Proof. Let $\xi_m \to \xi$ in L. Then $\liminf p^n(i,j,\xi_m) \geq p^n(i,j,\xi)$ for each i,j,n by Fatou's lemma. Since $1 = \sum p^n(i,j,\xi_m) = \sum p^n(i,j,\xi)$, it is easily seen that $\lim p^n(i,j,\xi_m) = p^n(i,j,\xi)$. By Scheffe's theorem [2], $p^n(i,.,\xi_m) \to p^n(i,.,\xi)$ in total variation and hence weakly. The claim follows.

QED

Let $\xi_n \to \xi_\infty$ in L and for $m = 1,2,\ldots,\infty$, let X_n^m, $n = 1,2,\ldots$, be a Markov chain governed by $\gamma\{\xi_m\}$ with $X_1^m = 1$ and Q^m the law of $\{X_n^m, n = 1,2,\ldots\}$ viewed as a probability measure on the canonical space $S^\infty = SxSx\ldots$

Lemma 3.3. $Q^m \to Q^\infty$ weakly.

Proof. A family of probability measures on a product space is tight if and only if its image measures under each coordinate projections are. Thus Lemma 3.2 implies the tightness of $\{Q^1, Q^2,\ldots\}$. Let Q be any limit point of this set. Let $[X_1, X_2,\ldots]$ be an S^∞-valued random variable with law Q. Clearly, $X_1 = 1$. Pick $N \geq 1$. Let $f:S \to R$ and $g:SxSx\ldots xS$ (N times) $\to R$ be bounded maps. Let Y_n^m, $\bar{Y}_n$ be defined as in (3.1) with $(\{X_n\}, \xi)$ replaced by $(\{X_n^m\}, \xi_m)$, $(\{X_n\}, \xi_\infty)$ respectively. By Lemma 3.1, for $n \geq N$,

$$E[(Y_n^m - Y_N^m)g(X_1^m,\ldots,X_N^m)] = 0.$$

Passing to the limit along an appropriate subsequence, we get

$$E[(\bar{Y}_n - \bar{Y}_N)g(X_1,\ldots,X_n)] = 0.$$

A standard argument based on a monotone class theorem shows that $\{\bar{Y}_n\}$ is a martingale with respect to the natural filtration of $\{X_n\}$. The claim now follows from Lemma 3.1.

QED

By Skorohod's theorem, we can construct on a common probability space (Ω,F,P) random variables $\tilde{X}_n^m$, $n = 1,2,\ldots,m=1,2,\ldots,\infty$, such that the law of $[\tilde{X}_1^m, \tilde{X}_2^m,\ldots]$ agrees with the law of $[X_1^m, X_2^m,\ldots]$ for each m and $\tilde{X}_n^m \to \tilde{X}_n^\infty$ for each n, a.s. We shall assume this done and by abuse of notation, write X_n^m for $\tilde{X}_n^m$. Let $\tau^m = \min\{n>1 | X_n^m = 1\}$ and

$$u^m(i) = E[\sum_{j=2}^{\tau^m} I\{X_j = i\}]$$

for $i\varepsilon S$, $m = 1,2,,\ldots,\infty$. (Here and elsewhere, $I\{\ldots\}$ stands for the indicator function.)

<u>Corollary 3.1</u>. Under condition C, $E[\tau^m] \to E[\tau^\infty]$ and $u^m(i) \to u^\infty(i)$ for each $i\varepsilon S$.

<u>Proof</u>. Outside a set of zero probability, $X_n^m \to X_n^\infty$ for each n. Since these are discrete valued, $X_n^m = X_n^\infty$ from some n onwards. It is easy to see from this that $\tau^m \to \tau^\infty$ a.s. The first claim follows. The second is proved similarly.

QED

Theorem 3.1. The following are equivalent:

(i) Condition C holds.

(ii) $\{\pi\{\xi\}|\xi\varepsilon L\}$ is tight.

(iii) $\{\pi\{\xi\}|\xi\varepsilon L\}$ is compact in the topology of weak convergence.

Proof. It is clear that (iii) implies (ii). Suppose (i) holds. For $\{\xi_n\}$ as above, we have

$$\pi\{\xi_n\}(i) = u^n(i)/E[\tau^m] \to u^\infty(i)/E[\tau^\infty] = \pi\{\xi_\infty\}(i)$$

for each i. By Scheffe's theorem, $\pi\{\xi_n\} \to \pi\{\xi_\infty\}$ weakly and in total variation. Thus the map $\xi \to \pi\{\xi\}$ is continuous. Since L is compact, (iii) follows. We shall now prove that (ii) implies (i). Suppose (ii) holds and (i) is false. Then $\{\xi_n\}$ above can be picked so that lim inf $E[\tau^m] > E[\tau^\infty]$. (The $\geq$ inequality always holds by Fatou's lemma.) Since $\pi\{\xi^m\}(1) = (E[\tau^m])^{-1}$ for each m, we have lim sup $\pi\{\xi_m\}(1) < \pi\{\xi_\infty\}(1)$. Now for each $N\geq 1$,

$$\sum_{i=1}^{N} \pi\{\xi_m\}(i)p(i,j,\xi_m(i)) \leq \pi\{\xi_m\}(j).$$

Let $\pi = [\pi(1), \pi(2),\ldots]$ be any weak limit point of $\{\pi\{\xi_n\}\}$. Then passing to the limit along an appropriate subsequence in the above inequality, we have

$$\sum_{i=1}^{N} \pi(i)p(i,j,\xi_\infty(i)) \leq \pi(j) \ .$$

Letting $N\uparrow\infty$,

$$\sum_{i\varepsilon S} \pi(i)p(i,j,\xi_\infty(i)) \leq \pi(j).$$

Since both sides of the inequality add up to one when summed over j, equality must hold. Hence π must be of the form $\pi(i) = \alpha\pi\{\xi_\infty\}(i)$, $i\varepsilon S$, for some $\alpha \ \varepsilon \ [0,1)$. This contradicts the tightness of $\{\pi\{\xi_n\}, n=1,2,\ldots\}$. Thus (ii) implies (i).

QED

<u>Corollary 3.2</u>. Under condition C, there exists a $\xi_0 \ \varepsilon \ L$ such that $C_k\{\xi_0\} = \alpha$.

Later on in section IX, we shall give simpler sufficient conditions that ensure condition C. See [6] for some related results.

IV. THE DYNAMIC PROGRAMMING EQUATIONS

The results of this section are essentially the same as the corresponding results of [4] except for the much more general set-up being used here. The proofs of [4] apply with a little extra work and thus we omit many details, referring the reader to [4]. Assume condition C. In particular, this implies that all SS are SSS.

Lemma 4.1. For any $i \varepsilon S$, $u \varepsilon D(i)$,

$$\sum_j p(i,j,u)\ E_\xi[\tau/X_1 = j] < \infty$$

Proof. Let $\xi' \varepsilon L$ be such that $\xi'(j) = \xi(j)$ for $j \neq i$, $\xi'(i) = u$. Let $\tau_j = \min\{n>1 \,|\, X_n = j\}$. Then

$$\sum_j p(i,j,u)E_\xi[\tau/X_1=j] = \sum_j p(i,j,u)E_\xi[\tau I\{\tau_i<\tau\}/X_1=j]$$

$$+ \sum_j p(i,j,u)E_\xi[\tau I\{\tau_i>\tau\}/X_1=j]$$

$$\leq (E_{\xi'}[\tau_i/X_i = i] + E_\xi[\tau/X_1 = i]) + E_{\xi'}[\tau/X_1 = i] < \infty.$$

QED

In particular, it follows that $\sum p(i,j,u)V\{\xi\}(j)$ is well-defined for all i,u,ξ.

Lemma 4.2. For an SSRS $\gamma[\Phi]$, $v[\Phi](1) = 0$ and

$$C_k[\Phi]1_c = (P[\Phi] - U)V[\Phi] + Q \tag{4.1}$$

termwise. In particular, for an SSS $\gamma\{\xi\}$,

$$C_k\{\xi\}1_c = (P\{\xi\} - U)V[\Phi] + Q. \tag{4.2}$$

Furthermore, any $W = [W(1),W(2),\ldots]^T$ satisfying

$$C_k[\Phi]1_c = (P[\Phi] - U)W + Q$$

with $\sup (W(i) - V[\Phi](i)) < \infty$ must differ from $V[\Phi]$ at most by a constant multiple of 1_c.

For a proof, see Lemmas 3.1 and 3.2 of [4]. Let $\{X_n\}$ be governed by an SSS $\gamma\{\xi\}$ with $X_1 = 1$. Consider

$$Y = \sum_{m=2}^{\tau} (V\{\xi\}(X_m) - E[V\{\xi\}(X_m)/X_{m-1}]).$$

Since $V\{\xi\}(X_\tau) = V\{\xi\}(X_1) = 0$ by the above lemma, this equals

$$\sum_{m=1}^{\tau-1} (V\{\xi\}(X_m) - E[V\{\xi\}(X_{m+1})/X_m])$$

$$= \sum_{m=1}^{\tau-1} (k(X_m) - C_k\{\xi\}),$$

by (4.2). Thus $E[|Y|] < \infty$.

<u>Lemma 4.3</u>. $E[Y] = 0$.

For a proof, see Lemma 5.2, [4].

<u>Lemma 4.4</u>. $C_k\{\xi\} = \alpha$ if and only if the following holds termwise:

$$C_k\{\xi\}1_c = \min_u (P_u - U)V\{\xi\} + Q \tag{4.3}$$

The proof is as in Lemma 5.3, [4], and the remarks that immediately follow it. (4.3) are called the dynamic programming equations. It should be remarked that in this paper we do not address the important issue of studying the set of general solutions (c,W) of the system of equations

$$c1_c = \min(P_u - U)W + Q$$

and characterizing the particular class $(C_k\{\xi\}, V\{\xi\} + \text{constant x } 1_c)$ with ξ as above, from among this set.

V. STABILITY UNDER ARBITRARY STRATEGIES

In this section, we assume condition C and show that the mean return time to state 1 remains bounded under arbitrary control strategies. Define τ as before. As observed in section III, the map $\xi \to E_\xi[\tau/X_1 = 1]$ is continuous. Hence there exists a $\bar{\xi} \varepsilon L$ such that $E_{\bar{\xi}}[\tau/X_1=1] = \max E_\xi[\tau/X_1=1]$.

Let $A_1, A_2 \ldots$ be finite subsets of S containing 1 such that $\cup A_n = S$. Define $\tau_n = \min\{m\geq 1 | X_m \notin A_n \text{ or } X_m = 1\}$, $n = 1,2,\ldots$ Define $v_n: S \to R$ by

$$v_n(i) = \begin{cases} \inf E[\tau_n/X_1 = i] & \text{if } i \varepsilon A_n,\ i\neq 1, \\ 0 & \text{if } i \notin A_n, \\ \inf E[v(X_2)I\{X_2\neq 1\}/X_1=1] & \text{if } i=1, \end{cases}$$

where the infima are over all CS. Standard dynamic programming arguments [10] show that v_n satisfies

$$v_n(i) = \max_\xi [1 + \sum_j p(i,j,\xi(i))v_n(j)] \tag{5.1}$$

for $i \varepsilon A_n$, $i\neq 1$, and

$$v_n(1) = \max_\xi [1 + \sum_{j\neq 1} p(1,j,\xi(1))v_n(j)] \tag{5.2}$$

Note that the summations on the right are finite in both cases because $v_n(j) \neq 0$ for at most finitely many j. Hence by continuity of p(i,j,.) and compactness of the D(i)'s, the maximum is attained in each case. Let ξ_n be such that $\xi_n(i)$, $\xi_n(1)$ attain the maxima in (5.1), (5.2).

Lemma 5.1. $E_{\xi_n}[\tau_n/X_1 = 1] = \max E_{\xi}[\tau_n/X_1 = 1]$, where the maximum is over all CS.

The proof follows by standard dynamic programming arguments [10] and is omitted.

Corollary 5.1. $\sup E[\tau/X_1=1] \leq E_{\bar{\xi}}[\tau/X_1=1]$, where the supremum is over all CS.

Proof. Under any CS,

$$E[\tau_n/X_1 = 1] \leq E_{\xi_n}[\tau/X_1 = 1]$$

$$\leq E_{\xi_n}[\tau/X_1 = 1] \leq E_{\bar{\xi}}[\tau/X_1 = 1]$$

Let $n \to \infty$.

QED

Corollary 5.2. All SRS are SSRS.

Lemma 5.2. $\{\pi[\Phi] \mid \Phi$ a product probability measure on L$\}$ is a compact set in the topology of weak convergence.

Proof. The arguments leading to Lemma 5.1 and Corollary 5.1 can also be used to prove the following: For any SRS $\gamma[\Phi]$,

$$E_{\Phi}[\sum_{m=2}^{\tau} I\{X_m \geq N\}/X_1=1] \leq \max_{\xi} E_{\xi}[\sum_{m=2}^{\tau} I\{X_m \geq N\}/X_1=1],$$

$$N = 1,2,\ldots,$$

$$E_{\Phi}[\tau/X_1=1] \geq \min_{\xi} E_{\xi}[\tau/X_1 = 1],$$

where the maximum (resp. minimum) are attained for some ξ. In particular, the right hand side of the second inequality is strictly positive. Thus

$$\sum_{i \geq N} \pi[\Phi](i) = E_{\Phi}[\sum_{m=2}^{\tau} I\{X_m \geq N\}/X_1=1]/E_{\Phi}[\tau/X_1=1]$$

$$\leq \text{constant} \times \max_{\xi}(\sum_{i \geq N} \pi\{\xi\}(i)).$$

The tightness of $\{\pi[\Phi]\}$ now follows from the tightness of $\{\pi\{\xi\}\}$. Let $\Phi_n \to \Phi_\infty$ in the topology of weak convergence of probability measures on L. The space M(L) of probability measures on L with this topology is compact by Prohorov's theorem and the set $\bar{M}(L)$ of probability measures in M(L) of the product form is a closed and hence compact subset of this space. Let π be any weak limit point of $\{\pi[\Phi_n]\}$, i.e., $\pi[\Phi_{n_j}] \to \pi$ for some subsequence $\{n_j\}$. By Scheffe's theorem, $\pi[\Phi_{n_j}] \to \pi$ in total variation. Hence letting $j \to \infty$ in

$$\pi[\Phi_{n_j}]P[\Phi_{n_j}] = \pi[\Phi_{n_j}],$$

we obtain,

$$\pi P[\Phi_\infty] = \pi,$$

i.e., $\pi = \pi[\Phi_\infty]$. Thus the map $\Phi \rightarrow \pi[\Phi]$ is continuous. The claim follows.

QED

Corollary 5.3. There exists an SSRS $\gamma[\Phi]$ such that $C_k[\Phi] = \beta$.

Using arguments identical to those leading to Lemma 4.4, one can prove the following.

Lemma 5.3. $C_k[\Phi_0] = \beta$ if and only if the following holds termwise:

$$C_k[\Phi_0]1_c = \min_{\Phi}(P[\Phi] - U)V[\Phi_o] + Q \tag{5.3}$$

Corollary 5.4. $\beta = \alpha$.

We omit the proof here. It follows exactly along the lines of the proof of Lemma 8.2 in section VIII, which in fact treats a slightly more complicated situation.

As in section III, we have

Corollary 5.5. The set $\{\tau | \{X_n\}$ is governed by some SSRS and $X_1 = 1\}$ is uniformly integrable.

VI. STATISTICAL BEHAVIOUR UNDER ARBITRARY STRATEGIES

In preparation for proving the optimality of the $\gamma\{\xi_0\}$ of Corollary 3.2 with respect to arbitrary CS, we prove further properties of a chain governed by an arbitrary CS.

Consider a fixed CS $\{\xi_n\}$ and let $\{X_n\}$ be the Markov chain governed by $\{\xi_n\}$ with $X_1 = 1$. As before, $\hat{\Phi}_i$ will denote the image of $\Phi \varepsilon \bar{M}(L)$ under the i-th coordinate projection. Let τ be as before. Note that we continue to assume condition C.

Lemma 6.1. For each $a \varepsilon (0,1)$, there exists a $\Phi(a) \varepsilon \bar{M}(L)$ such that for any bounded $f:S \to R$,

$$E[\sum_{m=1}^{\tau-1} a^{m-1} \sum_{j \varepsilon S} p(X_m, j, \xi_m(X_m)) f(j)]$$

$$E[\sum_{m=1}^{\tau-1} a^{m-1} \sum_{j \varepsilon S} (\int p(X_m, j, u) \hat{\Phi}(a)_{X_m}(du)) f(j)].$$

Proof. Fix a and construct $\Phi(a)$ as follows: Define a probability measure u on SxL by

$$\int g du = E[\sum_{m=1}^{\tau-1} a^{m-1} g(X_m, \xi_m)]/(E[(1-a)^{-1}(1-a^{\tau})]$$

for all bounded continuous $g:SxL \to R$. Disintegrate u as

$$u(dx,dy) = u_1(dx)u_2(x)(dy)$$

where u_1 is the image of u under the projection $SxL \to S$ and $u_2:S \to M(L)$ is the regular conditional distribution. Let $u_{2j}(i)$ denote the image of $u_2(i)$ under the projection $L \to D(j)$. Define $\Phi(a)$ as the product measure $u_{2i}(i)$. The claim now follows by direct verification.

QED

Let $f:S \to R$ be a bounded map. Define

$$h_a(i) = E_{\Phi(a)}[\sum_{m=1}^{\tau-1} a^{m-1}f(x_m)/X_1 = 1], \ i\varepsilon S.$$

$$Z_1 = h_a(1)$$

$$Z_n = \sum_{m=1}^{n-1} a^{m-1}f(X_m) + a^{n-1}h_a(X_n)I\{\tau>n\}, \ n=2,3,4,\ldots$$

$$W_n = Z_{n+1} - Z_n$$

$$= a^{n-1}f(X_n) + a^n h_a(X_{n+1})I\{\tau>n+1\} - a^{n-1}h_a(X_n)I\{\tau>n\}, \ n=1,2,\ldots$$

Clearly, $h_a(\cdot)$ is bounded.

Lemma 6.2.

$$E\Big[\sum_{m=1}^{\tau-1} a^{m-1}f(X_m)\Big] = E_{\Phi(a)}\Big[\sum_{m=1}^{\tau-1} a^{m-1}f(X_m)/X_1 = 1\Big].$$

<u>Proof</u>. The right hand side is $h_a(1)$. Note that

$$\sum_{m=1}^{\tau-1} a^{m-1}f(X_m) - h_a(1) = \sum_{m=1}^{\tau-1} W_n$$

Letting $\mathbf{F}_n = \sigma(X_m, \xi_m, m \le n)$, $n=1,2,\ldots,$ the sequence

$$\sum_{m=1}^{n-1} (W_n - E[W_m/\mathbf{F}_m])$$

becomes an $\{\mathbf{F}_{n+1}\}$-martingale. By the optional sampling theorem,

$$E\Big[\sum_{m=1}^{(\tau-1)\wedge n} W_m\Big] = E\Big[\sum_{m=1}^{(\tau-1)\wedge n} E[W_m/\mathbf{F}_m]\Big].$$

Since the expressions inside both expectations are bounded by a constant times τ, we can let $n \to \infty$ and apply the dominated convergence theorem to claim

$$E\Big[\sum_{m=1}^{\tau-1} a^{m-1}f(X_m)\Big] - h_a(1) = E\Big[\sum_{m=1}^{\tau-1} W_m\Big]$$

$$= E[\sum_{m=1}^{\tau-1} E[W_m/F_m]]$$

$$= E[\sum_{m=1}^{\tau-1} a^{m-1}(f(X_m) - h_a(X_m) +$$

$$a \sum_{1 \neq j \varepsilon S} p(X_m, j, \xi_m(X_m))h_a(j))]$$

By Lemma 6.1, the right hand side equals

$$E[\sum_{m=1}^{\tau-1} a^{m-1}(f(X_m) - h_a(X_m) + a \sum_{1 \neq j \varepsilon S} (\int p(i,j,u)\hat{\Phi}(a)_i(du))h_a(j))].$$

Hence the expression in (6.1) is zero. The claim follows.

QED

Note that for a given a, the claim above holds for all bounded f: S $\rightarrow$ R. Let $a_n \rightarrow 1$ in (0,1) such that $\Phi(a_n) \rightarrow \Phi$ in $\bar{M}(L)$ for some Φ. Then a trivial limiting argument (as for Corollary 3.1) shows that for all bounded f:S $\rightarrow$R.

$$E[\sum_{m=1}^{\tau-1} f(X_m)] = E_{\Phi}[\sum_{m=1}^{\tau-1} f(X_m)].$$

Summarizing the results,

Lemma 6.3. For any CS$\{\xi_n\}$, there exists an SSRS $\gamma[\Phi]$ such that for all bounded $f:S \to R$,

$$E[\sum_{m=1}^{\tau-1} f(X_m)/X_1=1] = E_{\Phi}[\sum_{m=1}^{\tau-1} f(X_m)/X_1=1],$$

where the expectation on the left is with respect to the law under $\{\xi_n\}$.

Let $\{\xi_n\}$ be a fixed CS as before and $\{X_n\}$ the chain governed by $\{\xi_n\}$ with $X_1=1$. Let $\sigma_0=1$, $\sigma_n=\min\{m>\sigma_{n-1}|X_m=1\}$ ($<\infty$ a.s. by Corollary 5.1), $F_n = \sigma(X_m, \xi_m, m\leq n)$ and F_{σ_n} = the stopped σ-field for the stopping time σ_n. We say that an S-valued sequence $\{Y_n\}$ of random variables is an acceptable controlled Markov chain if $Y_1=1$ and there exists an L-valued sequence of random of random variables $\{\xi_n'\}$ such that $P(Y_{n+1} = j/Y_m, \xi_m, m\leq n) = p(Y_n, j, \xi_n'(Y_n))$.

Lemma 6.4. For each n, the regular conditional law of the process X_{σ_n+m}, m=1,2,..., conditioned on F_{σ_n} is a.s. the law of an acceptable controlled Markov chain.

<u>Proof</u>. Fix $n,m\geq 1$ and let $\omega \rightarrow P_\omega(\cdot)$ denote a version of the regular conditional law of X_{σ_n+i}, $i=1,2,\ldots,$ given $\mathbf{F}_{\sigma_n}$. Recall the definition of $P[\Phi]$. Let $p_\Phi(i,j)$ denote the (i,j)-th element of $P[\Phi]$. We need to show that there exists an $\bar{M}(L)$-valued $\mathbf{F}_{\sigma_n}$-measurable random variable $\Phi(\omega)$ such that for any bounded real random variable Y which is measurable with respect to $\mathbf{F}_{\sigma_n+m}$,

$$E[E[I\{X_{\sigma_n+m+1} = j\}/\mathbf{F}_{\sigma_n+m}]Y/\mathbf{F}_{\sigma_n}] = E[p_{\Phi(\omega)}(X_{\sigma_n+m},j)Y/\mathbf{F}_{\sigma_n}]$$

a.s. But the left hand side equals

$$E[I\{X_{\sigma_n+m+1} = j\}Y/\mathbf{F}_{\sigma_n}] = \sum_{i\varepsilon S} E[I\{X_{\sigma_n+m+1}=j\}I\{X_{\sigma_n+m}=i\}Y/\mathbf{F}_{\sigma_n}]$$

$$= \sum_{i\varepsilon S} E[p(i,j,\xi_{\sigma_n+m}(i))YI\{X_{\sigma_n+m}=i\}/\mathbf{F}_{\sigma_n}] \tag{6.2}$$

Define new probability measures $\bar{P}_i$ on the underlying probability space by

$$\frac{d\bar{P}_i}{dP} = YI\{X_{\sigma_n+m} = i\}$$

and let $\tilde{E}_i[\cdot]$ denote the expectation with respect to $\bar{P}_i$. A standard change of measure argument shows that (6.2) equals

$$\sum_{i\varepsilon S} \tilde{E}_i[p(i,j,\xi_{\sigma_n+m}(i))/F_{\sigma_n}]E[YI\{X_{\sigma_n+m} = i\}/F_{\sigma_n}].$$

Define the $\bar{M}(L)$-valued random variable $\Phi(\omega)$ as follows: Let $\Phi(\omega) = \Pi\Phi_i(\omega)$ where $\Phi_i(\omega)$ is the random measure of $D(i)$ defined by

$$\int f d\Phi_i(\omega) = E[f(\xi_{\sigma_n+m}(X_{\sigma_n+m}))/F_{\sigma_n}]$$

for all bounded continuous $f:D(i) \to R$. (The definition of $\Phi_i(\omega)$ and hence that of $\Phi(\omega)$ is specified only a.s.) Then (6.2) equals

$$\sum_{i\varepsilon S} p_{\Phi(\omega)}(i,j)E[YI\{X_{\sigma_n+m} = i\}/F_{\sigma_n}]$$

$$= E[(\sum_{i\varepsilon S} p_{\Phi(\omega)}(i,j)I\{X_{\sigma_n+m} = i\})Y/F_{\sigma_n}]$$

$$= E[p_{\Phi(\omega)}(X_{\sigma_n+m}, j)Y/F_{\sigma_n}]$$

QED

Corollary 6.1. For each $a \varepsilon (0,1]$, there exists a sequence $\{\phi_a(n)\}$ of $\bar{M}(L)$-valued random variables such that $\phi_a(n)$ is F_{σ_n}-measurable for each n and for all bounded $f:S \to R$

$$E\Big[\sum_{m=\sigma_n}^{\sigma_{n+1}-1} a^{m-1} f(X_m)/F_{\sigma_n}\Big] = E_{\phi_a(n)}\Big[\sum_{m=1}^{\tau-1} a^{m-1} f(X_m)/X_1 = 1\Big] \quad \text{a.s.}$$

where $E_{\Phi}[\cdot]$ is the expectation with respect to the law under $\gamma[\Phi]$.

<u>Proof.</u> Follows from Lemmas 6.2, 6.3 and 6.4. QED

In particular, this implies that

$$E\Big[\sum_{m=\sigma_n}^{\sigma_{n+1}-1} a^{m-1} f(X_m)/f_{\sigma_n}\Big] \leq \sup_{\Phi} E_{\Phi}\Big[\sum_{m=1}^{\tau-1} a^{m-1} f(X_m)/X_1=1\Big] \quad \text{a.s.}$$

for a, f as above. This weaker conclusion is all we need for what follows and can be proved also by combining Lemma 6.4 with a dynamic programming argument similar to the one used to prove Corollary 5.1. However, we prefer the above approach because the content of Corollary 6.1 is of independent interest and is not captured by the alternative argument suggested above.

VII. OPTIMALITY OF $\gamma\{\xi_0\}$

In this section, we establish the optimality of $\gamma\{\xi_0\}$ of Corollary 3.2 under condition C. Let $\{\xi_n\}$, $\{X_n\}$ be as in Section VI.

Let $\bar{S} = S\cup\{\infty\}$ denote the one point compactification of S and M(S), M($\bar{S}$) respectively the spaces of probability measures on S, $\bar{S}$ with the topology of weak convergence. Define a sequence $\{\nu_n\}$ of M($\bar{S}$)-valued random variables and another sequence $\{u_n\}$ in M($\bar{S}$) as follows: For $A\subset\bar{S}$, n=1,2,...,

$$\nu_n(A) = \frac{1}{n}\sum_{m=1}^{n} I\{X_m \varepsilon A\}$$

$$u_n(A) = E[\nu_n(A)].$$

Lemma 7.1. Outside a set of zero probability, every limit point ν^* of $\{\nu_n\}$ in M($\bar{S}$) (which is compact) has a decomposition

$$\nu^* = (1-c)\pi[\Phi] + c\delta \tag{7.1}$$

for some c ε [0,1], Φ ε $\bar{M}(L)$, δ being the Dirac measure at ∞.

The proof is as in Lemma 3.6, [4] except that now we need to consider $f\varepsilon\bar{G}$ (in the notation of [4]) with finite supports. A similar argument proves the following:

Lemma 7.2. Each limit point u^* of $\{u_n\}$ has a decomposition

$$u^* = (1-c)\pi[\Phi] + c\delta \tag{7.2}$$

for some $c \in [0,1]$ and $\Phi \in \bar{M}(L)$.

<u>Lemma 7.3</u>. Under condition C, c=0 in (7.2).

<u>Proof</u>. Let $\bar{u}_n$ denote the restriction to u_n to S for each n, viewed as an element of M(S). It suffices to prove that $\{\bar{u}_n\}$ is tight. Let $A_n = \{n+1, n+2, \ldots\} \subset S$, $n = 1,2,\ldots$ We have

$$\bar{u}_n(A_N) = E[\frac{1}{n} \sum_{m=1}^{n} I\{X_m \in A_N\}]$$

$$\leq E[\frac{1}{n} \sum_{m=1}^{\sigma_{n+1}} I\{X_m \in A_N\}]$$

$$= E[\frac{1}{n} \sum_{m=0}^{n} E[\sum_{j=\sigma_m}^{\sigma_{m+1}-1} I\{X_j \in A_N\}/\mathcal{F}_{\sigma_m}]]$$

$$= E[\frac{1}{n} \sum_{m=0}^{n} E_{\phi_1(n)}[\sum_{m=1}^{\tau-1} I\{X_m \in A_N\}/X_1 = 1]]$$

$$\leq 2 \sup_{\Phi} E_{\Phi}[\sum_{m=1}^{\tau-1} I\{X_m \in A_N\}/X_1 = 1]$$

$$\leq 2 \sup_{\Phi} \left(\sum_{i=N+1}^{\infty} \pi[\bar{\Phi}](i) \right) \sup_{\Phi} E_{\Phi}[\tau/X_1 = 1]$$

where $\{\phi_1(n)\}$ are as in Corollary 6.1. By Lemma 5.2, the right hand side can be made arbitrarily small by making N large. Thus $\{\bar{u}_n\}$ are tight and the claim follows.

QED

Theorem 7.1. $\gamma\{\xi_0\}$ is optimal in the mean.

Proof. From (7.2) and the above lemma,

$$\liminf_{n \to \infty} E[\varphi_n] \geq \min_{\Phi} C_k[\Phi] = \alpha. \qquad \text{QED}$$

To ensure a.s. optimality, we need the following stronger condition:

Condition D: $\sup E[\tau^2/X_1 = 1] < \infty$, where the supremum is over all CS.

It is easy to see that this implies condition C.

Theorem 7.2. Under condition D, $\gamma\{\xi_0\}$ is a.s. optimal.

Proof. Let $\bar{\nu}_n$ denote the restriction to S of ν_n, n=1,2,..., viewed as an element of M(S). By Corollary 5.4 and Lemma 7.1, it suffices to prove that for each sample point outside a set of zero probability, the sequence $\{\bar{\nu}_n\}$ is tight. Let A_n, n=1,2,..., be as in the proof of Lemma 7.3. Then

$$\nu_n(A_N) = \frac{1}{n} \sum_{m=1}^{n} I\{X_m \varepsilon A_N\} \leq \frac{1}{n} \sum_{m=0}^{n} F_m$$

where

$$F_m = \sum_{j=\sigma_m}^{\sigma_{m+1}-1} I\{X_j \varepsilon A_N\}, \ m=0,1,2,\ldots$$

For any $\varepsilon>0$ and $m=1,2,\ldots$.

$$E[F_m/\mathcal{F}_{\sigma_m}] = E_{\phi_1(m)}[\sum_{m=1}^{\tau-1} I\{X_m \varepsilon A_N\}/X_1 = 1]$$

$$\leq \sup_{\Phi}(\sum_{i=N+1}^{\infty} \pi[\Phi](i)) \sup_{\Phi} E_{\Phi}[\tau/X_1=1] < \varepsilon$$

for N sufficiently large. Under condition D, we can use the strong law of large number for martingales ([11], pp. 53) to conclude that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n} (F_m - E[F_m/\mathcal{F}_{\sigma_m}]) = 0 \quad \text{a.s.}$$

Hence

$$\limsup_{n\to\infty} \nu_n(A_n) < \varepsilon \quad \text{a.s.}$$

for N sufficiently large. The claim follows. QED

There is one important case for which we have a.s. optimality even without condition D.

<u>Theorem 7.3</u>. If k is almost monotone, $\gamma\{\xi\}$ is a.s. optimal even if condition D does not hold.

<u>Proof</u>. From Lemma 7.1 and the definition of almost monotonicity, it follows that $\varphi_\infty \geq \beta = \alpha$ a.s. QED

<u>Corollary 7.1</u>. $\gamma\{\xi_0\}$ is optimal in probability.

<u>Remark</u>. We do not assume condition D here.

<u>Proof</u>. Using Theorem 7.1, 7.3 above, imitate the arguments of Lemmas 5.4, 5.5, Corollary 5.2 and Theorem 5.1 of [4] to conclude. QED

Theorem 7.2 is important in adaptive control situations involving self-tuning as in [5]. Taking k to be minus the indicator of successive A_N's defined as above, one verifies easily that Theorem 7.2 implies 'condition T' of [5]. Thus all the results of [5] can be rederived in the much more general set-up here if we impose, as in [5], the additional restriction that p(i,j,u) is either = 0 for all u or >0 for all u, for each pair i, j$\in$S.

(One can work around this extra restriction. However, we won't digress into these matters here.)

VIII. THE GENERAL CASE

In this section, we assume that there exists at least one SS that is not an SSS. Clearly, conditions C and D fail. In general, one cannot expect to find an optimal SSS. For example, if $k(i) > 0$ for all i and $\lim k(i) = 0$, then $C_k\{\xi\} > 0$ for all SSS $\gamma\{\xi\}$, but $\lim \varphi_n = 0$ a.s. under any SS that is not an SSS. This suggests that we should put a restriction on k that will penalize the unstable behaviour. Almost monotonicity does precisely that. In what follows, we assume that k is almost monotone and at least one SSRS exists.

Lemma 8.1. There exists an SSRS $\gamma[\Phi]$ such that $C_k[\Phi] = \beta$.

This is proved exactly as in Lemma 4.1, [4].

Lemma 8.2. $\beta=\alpha$

Proof. Let Φ be as above. Writing $P[\Phi] = [[p_\Phi(i,j)]]$, we have

$$\sum_j p_\Phi(i,j)E_\Phi[\tau/X_1=j] = E_\Phi[\tau/X_1=i] < \infty \tag{8.1}$$

for all $i \varepsilon S$ and

$$\beta 1_c = (P[\Phi] - U)V[\Phi] + Q \tag{8.2}$$

Recall that $\hat{\Phi}_i$ is the image of Φ under the projection $L \rightarrow D(i)$. For each i, (8.1) implies that

$$\overset{\bullet}{\Phi}_i(\{u\varepsilon D(i) \mid \sum_j p(i,j,u)E_{\Phi}[\tau/X_1=j] < \infty\}) = 1$$

Suppose that for some $i_0 \varepsilon S$, $\sum p(i_0,j,.)\ V[\Phi](j)$ is not $\overset{\bullet}{\Phi}_i$ - a.s. constant. Relabelling S if necessary, assume that $i_0=1$. Then we can pick a $u\varepsilon D(1)$ such that

$$\sum_j p(1,j,u)V[\Phi](j) < \sum_j p_{\Phi}(1,j)V[\Phi](j) \tag{8.3}$$

and

$$\sum_j p(1,j,u)E_{\Phi}[\tau/X_1=j] < \infty \tag{8.4}$$

Let $\Phi' \ \varepsilon \ \bar{M}(L)$ be such that $\overset{\bullet}{\Phi}{}'_i = \overset{\bullet}{\Phi}_i$ for $i\neq 1$ and $\overset{\bullet}{\Phi}{}'_1$ = the Dirac measure at u. By (8.4), $E_{\Phi'}[\tau/X_1=1] < \infty$ and hence $\gamma[\Phi']$ is an SSRS. It is also easy to see that $V[\Phi] = V[\Phi']$. Thus

$$\beta 1_c \geq (P[\Phi'] - U)V[\Phi'] + Q \tag{8.5}$$

with a strict inequality in the first row and an equality in the rest. Imitate the proof of Lemma 5.3, [4], to conclude that $C_k[\Phi'] < \beta$, a contradiction. Hence for each $i\varepsilon S$, the quantity $\sum p(i,j,.)V[\Phi](j)$ is $\overset{\bullet}{\Phi}_i$- a.s. a constant. Hence we can find a $u_i \ \varepsilon \ D(i)$ for each i such that

$$\sum_j p(i,j,u_i)E_{\Phi}[\tau/X_1=j] < \infty$$

and

$$\sum_j p(i,j,u_i)V[\Phi](j) = \sum_j p_{\Phi}(i,j)V[\Phi](j)$$

Construct SRS $\gamma[\Phi(1)]$, $\gamma[\Phi(2)]$,... such that $\hat{\Phi}_i(n) = \hat{\Phi}_i$ for $i>n$ and $\hat{\Phi}_i(n)$ = the Dirac measure at u_i for $i\leq n$. Let $\xi = [u_1,u_2,\ldots] \in L$. The arguments leading to (8.5) show that $\gamma[\Phi(1)]$ is an SSRS and

$$\beta 1_c = (P[\Phi(1)] - U)V[\Phi(1)] + Q,$$

implying $C_k[\Phi(1)] = \beta$. The same argument can be repeated to inductively claim that $C_k[\Phi(n)] = \beta$, n=1,2,... (In this connection, recall from the second half of the proof of Lemma 5.3, [4], that relabeling S can change $V[\Phi]$ at most by a constant multiple of 1_c.) Since k is almost monotone, it follows that $\{\pi[\Phi(n)], n=1,2,\ldots\}$ is tight. Let π be any weak limit point of this set. By Scheffe's theorem, $\pi[\Phi(n)] \to \pi$ in total variation along appropriate subsequence. Letting $n \to \infty$ along this subsequence in the equation $\pi[\Phi(n)]P[\Phi(n)] = \pi[\Phi(n)]$, we have $\pi P\{\xi\} = \pi$. Thus $\pi = \pi\{\xi\}$ and $\gamma\{\xi\}$ is an SSS. Moreover, $C_k\{\xi\} = \lim C_k[\Phi(n)] = \beta$. The claim follows.

QED

<u>Theorem 8.1</u>. An a.s. optimal SSS exists.

Proof. For ξ as above, $C_k\{\xi\} = \alpha$. From Lemma 7.1 and the almost monotonicity assumption on k, it follows that

$$\varphi_\infty \geq \beta = \alpha \quad \text{a.s.} \qquad \text{QED}$$

The dynamic programming equations are harder to come by. Say that an SSS $\gamma\{\xi\}$ is stable under local perturbation if for any $\xi' \in L$ such that $\xi'(i) = \xi(i)$ for all but one i, $\gamma\{\xi'\}$ is an SSS.

Condition E. All SSS are stable under local perturbation.

The proof of Lemma 5.3, [4] can then be repeated to prove the following:

Theorem 8.2. Under condition E, an SSS $\gamma\{\xi\}$ satisfies $C_k\{\xi\} = \alpha$ if and only if (4.3) holds.

A sufficient condition for condition E to hold is that for each $i \in S$, $p(i,j,.) \equiv 0$ for all j not belonging to some finite subset A_i of S. To see this, note that if $\xi'(i) = \xi(i)$ except for $i=i_0$, relabelling S so that $i_0 = 1$, we have

$$E_{\xi'}[\tau/X_1=1] = \sum_j p(1,j,\xi'(1))E_\xi[\tau/X_1=j] \leq \max_{j \in A_1} E_\xi[\tau/X_1=j] < \infty.$$

Thus the above subsumes the case studied in [4]. In particular, we did not need the second half of the assumption (*) recalled in section II.

Without condition E, the same proof can be used to show the following much weaker statement:

<u>Theorem 8.3</u>. If $\gamma\{\xi\}$ is an SSS satisfying $C_k\{\xi\} = \alpha$, then for any $i \varepsilon S$, $u \varepsilon D(i)$ such that $\gamma\{\xi'\}$ defined by $\xi'(j) = \xi(j)$ for $j \neq i$ and $\xi'(i) = u$ is an SSS, the following holds

$$\sum_j p(i,j,u) V\{\xi\}(j) - V\{\xi\}(i) + k(i) \geq \alpha.$$

IX. SUFFICIENT CONDITIONS FOR UNIFORM INTEGRABILITY

It seems unlikely that there is a simple characterization of conditions C or D applicable to all cases. Instead, one looks for simple sufficient conditions that may be used either singly or in combination depending on the problem at hand. In this section, we briefly outline a few such conditions.

Suppose we need a bound of the type sup $E[\tau^m/X_1=1] < \infty$ for some $m \geq 1$, the supremum being over all CS belonging to a prescribed class A. One obvious way of ensuring this would be to prove that for some integer $N \geq 1$ and reals $K>0$, $\varepsilon \in (0,1)$,

$$P(\tau \geq nN/X_1=1) \leq K\ \varepsilon^n,\ n=1,2,...$$

for all CS in A. Most methods below are based on this idea.

(1) Suppose there exists a map $V:S \rightarrow R$ such that $V(1) = 0$, $V(i) > 0$ for $i \neq 1$ and the following hold:

(A1) For some $\varepsilon_0 > 0$,

$$E[(V(X_{n+1}) - V(X_n))I\{V(X_n) > 0\}/F_n] \leq -\varepsilon_0 \tag{9.2}$$

under all CS.

(A2) There exists a random variable Z and a $\lambda > 0$ such that

$$E[\exp(\lambda Z)] < \infty$$

and for all $c\varepsilon R$,

$$P(|V(X_{n+1}) - V(X_n)| > c/F_n) \leq P(Z>c) \qquad (9.3)$$

under any CS.

Then a bound of the type (9.1) under all CS follows from Proposition 2.4 of [9], ensuring condition D. If we require (9.2), (9.3) to hold only for all SS, we still have condition C. See [9] also for potential applications to queueing and for a list of references that give other closely related criteria. This method is similar to the "Liapunov function" approach of [6], [10].

(2) Suppose that there exists a map $V:S \to R$ such that $V(i) \geq 1$ for all i and the following hold:

(A3) $\sup_{\xi} \sup_{1\leq i\leq n} E_{\xi}[V(X_2)/X_1=1] < \infty$ for all n=1,2,...

(A4) $\bigcup_{\xi\varepsilon L} \{i\varepsilon S|(V(i)/E_{\xi}[V(X_2)/X_1=i]) \leq \ell\}$ is a finite set of all ℓ=1,2,...

It is observed in [7], pp.415, that this ensures that all SS are SSS. Assume furthermore that

$$c = \inf_{\xi} \pi\{\xi\}(1) > 0.$$

Let $C = \{\nu \varepsilon M(S) | \nu(1) \leq c/2\}$ and pick an integer $N > 4/c$. Then

$$P(\tau \geq N/X_1 = 1) = P(\frac{1}{N} \sum_{m=1}^{N} I\{X_m = 1\} \leq \frac{2}{N}/X_1 = 1)$$

$$\leq P(\frac{1}{N} \sum_{m=1}^{N} I\{X_m = 1\} \leq \frac{c}{2}/X_1 = 1)$$

$$= P(\nu_N \varepsilon C/X_1 = 1). \qquad (9.3)$$

where $\{\nu_n\}$ are defined as in section VII. Under our assumptions, the methods of [7] yield an exponential bound of the type (9.1) for all SSS. This can be verified by checking that the estimates of [7], pp. 415-421, for the logarithm of the right hand side of (9.3) hold uniformly for all SSS under our hypotheses. Thus we have condition C. It is not clear whether this method can be adapted to ensure condition D as well.

(3) Suppose that there exists an $N \geq 1$ such that the following holds:

(A5) $$d = \sup \sup_{i \varepsilon S} P(\tau \geq N/X_1 = i) < 1 \qquad (9.4)$$

where the first supremum is over all CS. Then for any acceptable chain $\{X_n\}$,

$$P(\tau \geq 2N) = E[\prod_{m=1}^{2N-1} I\{X_m \neq 1\}]$$

$$= E[\prod_{m=2}^{N} I\{X_m \neq 1\} E[\prod_{m=N}^{2N-1} I\{X_m \neq 1\}/F_N]]$$

$$\leq d\ E[\prod_{m=2}^{N} I\{X_m \neq 1\}]$$

$$\leq d^2$$

Repeating this argument, we get an estimate of the type (9.1), ensuring condition D. If we take the first supremum in (9.4) over all SS only, we still have condition C.

(4) Suppose that

$$V(i) = \sup_{\xi} E_{\xi}[\tau/X_1=i] < \infty \tag{9.5}$$

for all i and furthermore,

$$\sup E_{\xi}[\sum_{m=1}^{\tau-1} V(X_m)/X_1] < \infty. \tag{9.6}$$

These imply condition D as we verify below: Under any CS,

$$E[\tau^2/X_1 = 1] = 2E[\sum_{m=1}^{\tau-1} m/X_1 = 1] + E[\tau/X_1=1]$$

By (9.5) with i=1 and Corollary 5.1, we have

$$\sup E[\tau/X_1=1] < \infty$$

where the supremum is over all CS. Also, for any acceptable $\{X_n\}$,

$$E[\sum_{m=1}^{\tau-1} m] = E[\sum_{m=1}^{\infty} (\tau-m)I\{\tau>m\}]$$

$$= E[\sum_{m=1}^{\infty} E[(\tau-m)/F_m]I\{\tau>m\}]$$

$$\leq E[\sum_{m=1}^{\infty} V(X_m)I\{\tau>m\}]$$

$$= E[\sum_{m=1}^{\tau-1} V(X_m)]$$

$$\leq \sup_{\xi} E_{\xi}[\sum_{m=1}^{\tau-1} V(X_m)/X_1 = 1] \qquad (9.7)$$

where the last inequality is obtained by dynamic programming arguments identical to those used to prove Lemma 5.1 and Corollary 5.1. The only difference here is that the supremum on the right hand side of (9.7) may not be attained for some ξ. However, it is finite by (9.6) and we are done.

Note that if we assume

$$\sup_{\xi \varepsilon L} \sup_{i \varepsilon S} E_{\xi}[\tau / X_1 = i] < \infty. \tag{9.8}$$

then sup V(i) < ∞ and (9.5), (9.6) hold automatically. Recall from section II that (9.8) is a typical assumption in the classical threatment of the problem, which is thus subsumed by our results.

X. CONTROL-DEPENDENT COST FUNCTIONS

The foregoing results are easily extended to control-dependent cost functions. For simplicity, let D(i) = a fixed compact metric space D for all i. This causes no lack of generality since we can always replace all D(i) by L and each p(i,j,.) by its composition with the projection map $L \to D(i)$. Let $k: S\mathrm{x}D \to [0,\infty)$ be a bounded continuous cost function and suppose that we seek to a.s. minimize

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} k(X_m, \xi_m(X_m))$$

(or "in mean" or "in probability" versions thereof.)

Let $M(\bar{S}\mathrm{x}D)$, $M(S\mathrm{x}D)$ denote the spaces of probability measures on $\bar{S}\mathrm{x}D$, SxD resp. with the topology of weak convergence. For each SSRS $\gamma[\Phi]$, define

$$\hat{\pi}[\Phi](A\mathrm{x}B) = \sum_{i \varepsilon A} \pi[\Phi](i)\hat{\Phi}_i(B),$$

for all $A \subset S$, B Borel in D. Write $\hat{\pi}\{\xi\} = \hat{\pi}[\Phi]$ when Φ = the Dirac measure at $\xi \varepsilon L$.

<u>Lemma 10.1</u>. A collection of $\hat{\pi}[\Phi]$'s is tight if and only if the corresponding collection of $\pi[\Phi]$'s is.

This is immediate from the compactness of D.

<u>Lemma 10.2</u>: (a) If $\hat{\pi}[\Phi(n)] \to \hat{\pi}$ in M(SxD) for a sequence of SSRS $\{\gamma[\Phi(n)]\}$,

then $\hat{\pi} = \hat{\pi}[\Phi]$ for some SSRS $\gamma[\Phi]$.

(b) If $\hat{\pi}\{\xi_n\} \to \hat{\pi}$ in M(SxD) for a sequence of SSS $\{\gamma\{\xi_n\}\}$, then $\hat{\pi} = \hat{\pi}\{\xi\}$ for some SSS $\gamma\{\xi\}$.

<u>Proof</u>. Let π denote the image of $\hat{\pi}$ under the projection $SxD \to S$. Then $\pi[\Phi(n)] \to \pi$ weakly and hence in total variation (by Scheffe's theorem). Let Φ be a limit point of $\{\Phi(n)\}$ in $\bar{M}(L)$. For each n, we have

$$\pi[\Phi(n)]P[\Phi(n)] = \pi[\Phi(n)].$$

Letting $n \to \infty$ along an appropriate subsequence, it follows that

$$\pi P[\Phi] = \pi,$$

i.e., $\pi = \pi[\Phi]$. Next, let $f{:}SxD \to R$ be bounded continuous with compact support. Let $\hat{f}(i,u) = \sum p(i,j,u)f(j)$. Then for each n, $\sum \pi[\Phi(n)](i)f(i) = \sum \pi[\Phi(n)](i)\int\hat{f}(i,.)d\hat{\Phi}(n) = \int\hat{f}d\hat{\pi}[\Phi(n)]$. Letting $n \to \infty$ along an appropriate subsequence, we get, $\sum \pi(i)f(i) = \sum \pi(i)\int\hat{f}(i,.)d\hat{\Phi}_i = \int\hat{f}d\hat{\pi}$. Hence $\hat{\pi} = \hat{\pi}[\Phi]$. This proves (a). The proof of (b) is similar. QED

Define a sequence $\{u_n'\}$ of $M(\bar{S}xD)$-valued random variables by:

$$u_n'(AxB) = \frac{1}{n}\sum_{m=1}^{n} I\{X_m \in A,\ \xi_m(X_m) \in B\}$$

for A,B Borel in $\bar{S}$, D respectively.

<u>Lemma 10.3</u>. For each sample point outside a set of zero probability, every limit point u^* of $\{u_n'\}$ in $M(\bar{S}xD)$ is of the form

$$u^* = (1-c)\hat{\pi}[\Phi] + c\eta$$

for some $c \varepsilon [0,1]$, some SSRS $\gamma[\Phi]$ and some probability measure η on $\{\infty\}xD$.

For a proof, see Lemma 3.6 of [4]. Using the above lemmas, it is a routine exercise to recover the results of preceding sections for this more general cost function, as long as we make the following obvious modifications:

(1) Redefine $C_k[\Phi]$ as $C_k[\Phi] = \int kd\hat{\pi}[\Phi]$.

(2) Let $Q_u = [k(1,u), k(2,u),\ldots]^T$. Replace (4.3) by

$$C_k\{\xi\}1_c = \min_u((P_u-U)V\{\xi\} + Q_u)$$

and (5.3) by

$$C_k[\Phi_0]1_c = \min_{\Phi}(P[\Phi] - U)V[\Phi_0] + Q_u)$$

(3) Redefine "almost monotonicity" as

$$\liminf_{i\to\infty} \inf_u k(i,u) > \beta.$$

The author has not attempted the case of unbounded cost functions, but is seems reasonable to expect similar results under suitable growth conditions. See [1] to get a feeling of what these could be.

Finally, it is tempting to conjecture the following:

(1) If all SS are SSS, condition C automatically holds.

(2) Theorem 7.2 holds even without condition D.

(3) In Theorem 8.2, condition E cannot be relaxed.

ACKNOWLEDGEMENTS

This work was done while the author was visiting the Institute for Mathematics and Its Applications, University of Minnesota. The author would like to thank Professors Bruce Hajek and Rabi Bhattacharya for useful conversations.

REFERENCES

[1] R.N. Bhattacharya, M. Majumdar, preprint.

[2] P. Billingsley, Convergence of Probability Measures, John Wiley, New York, 1968.

[3] V. Borkar, Controlled Markov Chains and Stochastic Networks, SIAM J. Control and Opt. 21 (1983), 652-666.

[4] V. Borkar, On Minimum Cost Per Unit Time Control of Markov Chains, SIAM J. Control and Opt. 22 (1984), 965-978.

[5] V. Borkar, P. varaiya, Identification and Adaptive Control of Markov Chains, SIAM J. Control and Opt. 20 (1982), 470-489.

[6] H. Deppe, Continuity of Mean Recurrence Times in Denumerable Semi-Markov Processes, Zeit. Wahr. 69 (1985), 581-592.

[7] M. Donsker, S.R.S. varadhan, Asymptotic Evaluation of Certain Markov Process Expectations for Large Time III, Comm. in Pure and Appl. Math. XXIX 4(1976), 389-461.

[8] A. Federgruen, A. Hordijk, H.C. Tijms, A Note on Simultaneous Recurrence Conditions on a Set of Denumerable Stochastic Matrices, J. Appl. Prob. 15 (1978), 842-847.

[9] B. Hajek, Hitting-Time and Occupation-Time Bounds Implied by Drift Analysis with Applications, Adv. Appl. Prob. 14 (1982), 502-525.

[10] H.J. Kushner, Introduction to Stochastic Control, Holt, Rinehart and WInston Inc., New York, 1971.

[11] M. Loeve, Probability Theory II, 4th Ed., Springer Verlag, New York, 1978.

[12] Z. Rosberg, P. Varaiya, J. Walrand, Optimal Control of Service in Tandem Queues, IEEE Trans. on Automatic Control, AC-27 (1982), 600-609.

[13] S. Ross, Applied Probability Models with Optimization Applications, Holden Day, san Francisco, 1970.

[14] J. Wessels, J. Van De Wal, Markov Decision Processes, Statistica Neerlandica 39(1985) 219-233.