The Annals of Probability
1986, Vol. 14, No. 1, 136–172

A CONTROL PROBLEM ARISING IN THE SEQUENTIAL DESIGN OF EXPERIMENTS

BY S. P. LALLEY and G. LORDEN

Columbia University and California Institute of Technology

1. Introduction and summary

The Pelé problem. Starting from an initial point $\mathbf{x}$ not in his playing field, a football player is to dribble onto the field. Due to irregularities in the surface on which the player is dribbling, and perhaps also to small inconsistencies in his kick, the movement of the ball has a "random" component; moreover, a kick with the left foot tends to have a somewhat different effect than a kick with the right foot. The player's objective is to move the ball onto the playing field with as few kicks as possible.

To make the problem more precise, suppose that the "playing field" is the first quadrant $\mathscr{Q}_+$ of $\mathbb{R}^2$. Assume that a kick with the left foot results in a (random) displacement whose distribution is F_A, while a kick with the right foot results in a (random) displacement distributed according to a law F_B (different from F_A). Assume also that F_A and F_B have finite second moments, and mean vectors $\boldsymbol{\mu}_A$ and $\boldsymbol{\mu}_B$ satisfying

$$\mu_A^{(1)} > \mu_B^{(1)} \geq 0, \qquad \mu_B^{(2)} > \mu_A^{(2)} \geq 0. \tag{1.1}$$

Then the Pelé problem may be phrased as an optimal control problem: Among all nonanticipating policies (a "policy" being a rule for deciding, at each step, which foot to kick with), find the policy $\mathscr{P}$, which minimizes $E_{\mathbf{x}}^{\mathscr{P}}T$, where T is the (random) number of kicks made before the first entry into $\mathscr{Q}_+$.

The problem as stated is, of course, a problem of "Markovian decision theory" [cf., for example, Derman (1970)], and may be "solved" by dynamic programming. But given that football players generally have neither the time nor adequate computational facilities for backward inductions, it seems reasonable to ask whether there is some simple policy which, although not optimal, is "nearly" optimal in some sense. One class of rules, which is particularly appealing, is the class of "straight-line switching" (SLS) policies: An SLS policy calls for a kick with the left foot whenever the ball is *above* a certain line (or ray) and a kick with the right whenever the ball is below this line.

We will prove

THEOREM A. *Suppose F_A and F_B have finite second moments and mean vectors satisfying* (1.1). *If $\mathscr{P}$ is any straight-line switching policy whose switch-*

Received February 1984; revised May 1985.
AMS 1980 *subject classifications.* Primary 60G40, 93E20.
Key words and phrases. Controlled random walk, first passage problem, martingale.

ing line (ray) lies in the interior of the negative cone determined by μ_A *and* μ_B, *then there is a constant* $C = C(\mathscr{P})$ *such that*

$$E_{\mathbf{x}}^{\mathscr{P}}T \le \inf_{\mathscr{R}} E_{\mathbf{x}}^{\mathscr{R}}T + C \tag{1.2}$$

for every $\mathbf{x} \in \mathbb{R}^2 \setminus \mathscr{Q}_+$.

The infimum in (1.2) is meant to be taken over *all* (nonanticipating) policies. The point of the theorem is that no matter how far from the field the player starts, he uses at most C more kicks (on the average) when proceeding by straight-line switching than when he uses the optimal strategy. The constant C will depend on the distributions F_A and F_B and also on the switching line. The methods of this paper may be used to obtain upper bounds for C in terms of familiar renewal-theoretic quantities, but we shall defer discussion of this matter to a subsequent paper.

Sequential design of experiments and acquisition of commodities. The Pelé problem is a peculiar manifestation of a more general problem concerning the sequential selection of experiments for discriminating between several competing hypotheses. Imagine an experimenter confronted by several different (simple) hypotheses $\{\theta_1, \theta_2, \dots, \theta_{d+1}\}$ concerning some facet of the "true state" of nature; suppose that the correct hypothesis is θ_{d+1}, and that this is known to the experimenter by unscientific means (e.g., divine revelation, subjective probability, the writings of Velikovsky). Suppose also that the experimenter has at his disposal a finite set of experiments $\{1, 2, \dots, d^*\}$, which he may perform repeatedly, one at a time. Each repetition of experiment i results in a random increment to the vector of log-likelihood ratios whose distribution is F_i, and for each repetition of i the experimenter is assessed a fixed cost δ_i. The experimenter's objective is to accumulate a certain amount a of evidence against *each* of the false hypotheses $\theta_1, \dots, \theta_d$ (where "evidence" is measured by the total change in the log-likelihood ratio during the experimentation).

This problem may be cast in a somewhat different setting. Imagine an investor, consumer, or other variety of capitalist roader interested in acquiring certain amounts a_i of various commodities θ_i $(i = 1, \dots, d)$. Suppose the buyer has several "actions" $i = 1, \dots, d^*$ available to him, each of which will result in the acquisition of (random) amounts of each of the commodities θ_i: Thus the effect of each i may be described by a probability distribution F_i on $\mathbb{R}^d$. (For example, the available actions might be (1) go to a hardware store, (2) go to a junkyard, and (3) go to an auto parts store.) Each action i may be performed repeatedly, with each repetition resulting in a (random) increment to his (vector of) acquisitions; however, each action has associated with it fixed cost δ_i. The buyer's objective is, of course, to attain his goal of acquiring amounts of $\theta_1, \theta_2, \dots, \theta_d$ greater than $a_1, a_2, \dots, a_d$ at minimum cost.

The problem concerning the sequential design of experiments as stated here was apparently introduced by Chernoff (1959), who described a procedure ("Procedure A") having the property that

$$\frac{E^{A}(\text{cost}(a))}{E^{\text{opt}}(\text{cost}(a))} \to 1 \quad \text{as } a \to \infty. \tag{1.3}$$

(Here E^A and E^{opt} denote expectations when Procedure A and the optimal procedure, respectively, are followed; cost (a) denotes the random cost incurred in acquiring a units of evidence.) The problem has subsequently been studied by Box and Hill (1967), Blot and Meeter (1973), and Keener (1980), (1981). Our results concern only the case $d = d^*$, where the number of experiments is the same as the number of competing hypotheses, but provide procedures which improve considerably on Chernoff's Procedure A. (These results also extend to analogous procedures for problems with $d^* \geq d$, but we defer discussion of this case to a subsequent paper.)

Assume that each of the distributions F_i $(i = 1, \dots, d)$ has a finite second moment, and let the mean vectors $\mu(F_i) = \int_{\mathbb{R}^d} \mathbf{x} F_i(d\mathbf{x})$ satisfy the following conditions:

$$\mu(F_i) \geq 0 \quad \forall\, i = 1, \dots, d; \tag{1.4}$$

(1.5) *the vector* $(1, 1, \dots, 1)$ *lies in the interior of the cone*

$$\mathscr{K} = \left\{ \sum_{i=1}^{d} \lambda_i \mu(F_i) \colon \lambda_i > 0,\ i = 1, \dots, d \right\}$$

(1.6) *no convex combination of a subset of* $\{\mu(F_1)/\delta(1), \dots, \mu(F_d)/\delta(d)\}$ *dominates a convex combination of a disjoint subset.*

NOTE. Throughout the paper the notation $\mathbf{x} \geq 0$ will mean $x^{(i)} \geq 0$ for each coordinate $x^{(i)}$ of $\mathbf{x}$; $\mathbf{x} > 0$ will mean that $x^{(i)} > 0$ for each coordinate. The terminology "dominates" used in (1.6) refers to the partial order $\geq$ on $\mathbb{R}^d$: Thus $\nexists$ permutation $\pi(\cdot)$ of $\{1, 2, \dots, d\}$, and probability vectors $(p_1, \dots, p_k)$, $(q_{k+1}, \dots, q_d)$ such that

$$\sum_{j=1}^{k} p_j \mu(F_{\pi(j)}) - \sum_{j=k+1}^{d} q_j \mu(F_{\pi(j)}) \geq 0.$$

At each stage of experimentation we are to choose one of the available distributions F_i, which then provides a random increment to the vector $\mathbf{W}_n$: Thus

$$\mathbf{W}_n = \sum_{j=1}^{n} \mathbf{w}_j,$$

$$\mathscr{L}(\mathbf{w}_{n+1} | \mathscr{A}_n) = F_{\Gamma_{n+1}},$$

where $\Gamma_j \in \{1, \dots, d\}$ $(j = 1, 2, \dots)$ are the successive choices of experiment, and $\{\mathscr{A}_n\}_{n \geq 0}$ is an increasing sequence of σ algebras such that for each $n \geq 0$, $\mathbf{w}_n$, and Γ_{n+1} are measurable with respect to $\mathscr{A}_n$ (cf. the definition of a "policy" at the end of this section).

The SLS policies for the Pelé problem have natural analogues in the sequential design problem, policies which we will call "diagonal-stabilizing" (DS). For a given point $\mathbf{x} \in \mathbb{R}^d$, let $\mathbf{d}(\mathbf{x})$ be the orthogonal projection of $\mathbf{x}$ onto the diagonal [the one-dimensional vector subspace of $\mathbb{R}^d$ spanned by $(1, 1, \dots, 1)$] and let $\mathbf{r}(\mathbf{x}) = \mathbf{d}(\mathbf{x}) - \mathbf{x}$. A policy $\mathscr{P}$ will be called "diagonal-stabilizing" (DS(α; r)) if there exist constants $\alpha > 0$ and $r < \infty$ such that

$$\mu(F_{\Gamma_{n+1}})^t \left(\frac{\mathbf{r}(\mathbf{W}_n)}{|\mathbf{r}(\mathbf{W}_n)|} \right) \geq \alpha \quad \text{on } \{|\mathbf{r}(\mathbf{W}_n)| > r\}, \tag{1.7}$$

i.e., if whenever $\mathbf{W}_n$ is a distance $> r$ from the diagonal, $\mathscr{P}$ is constrained to chose an F_i whose mean pushes inward toward the diagonal [in the direction of $\mathbf{r}(\mathbf{W}_n)$] at least α units. Whenever condition (1.5) is satisfied diagonal-stabilizing policies exist (see Section 4).

Example. Let $d = 3$ and $\mathscr{Y}$ be the subspace of $\mathbb{R}^3$ orthogonal to $(1, 1, 1)$. Suppose the mean vectors of the available controls are

$$\mu_A = (2, 2, 1),$$
$$\mu_B = (2, 1, 2),$$

and

$$\mu_C = (1, 2, 2).$$

A natural diagonal-stabilizing policy is as follows:

$$\Gamma_n = A \quad \text{iff } W_n^{(3)} \geq W_n^{(1)} \vee W_n^{(2)},$$
$$\Gamma_n = B \quad \text{iff } W_n^{(2)} \geq W_n^{(1)} \quad \text{and} \quad W_n^{(2)} > W_n^{(3)},$$
$$\Gamma_n = C \quad \text{iff } W_n^{(1)} > W_n^{(2)} \vee W_n^{(3)}.$$

Notice that Γ_n depends only on the projection of $\mathbf{W}_n$ on the subspace $\mathscr{Y}$. (See Figure 1.)

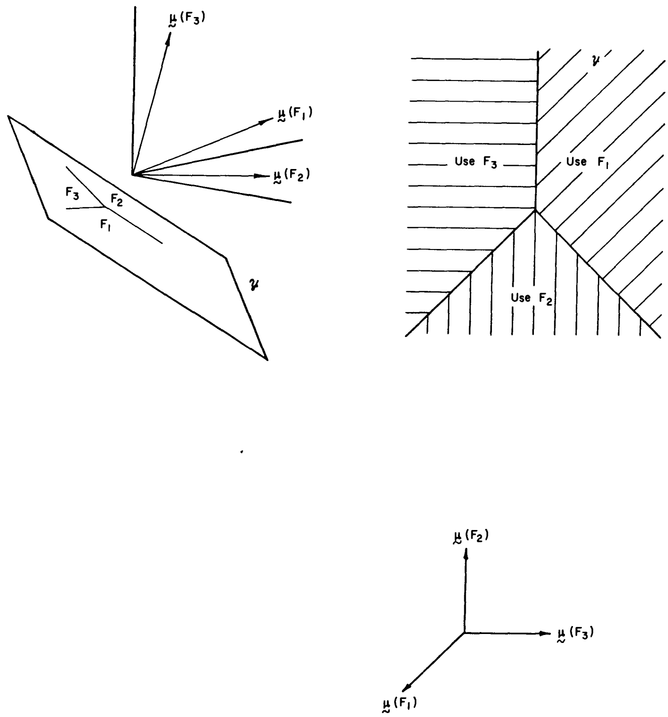
Let $T_a = \inf\{n\colon W_n^{(i)} > a \quad \forall\ i = 1, 2, \dots, d\}$.

Theorem B. *For each* $\alpha > 0$, $r_0 < \infty$ *there exists a constant* $C = C(\alpha, r_0)$ *such that for every* DS(α; r_0) *policy* $\mathscr{P}$ *and every* $a > 0$,

$$E_{\mathbf{0}}^{\mathscr{P}} \sum_{n=1}^{T_a} \delta(\Gamma_n) \leq \inf_{\mathscr{R}} E_{\mathbf{0}}^{\mathscr{R}} \sum_{n=1}^{T_a} \delta(\Gamma_n) + C(\alpha, r_0). \tag{1.8}$$

The constant $C(\alpha, r_0)$ also depends on the underlying distributions $F_1, \dots, F_d$ only through their mean vectors and second moments. The infimum in (1.8) is, again, over *all* nonanticipating policies.

The major difference between Theorems A and B concerns the choice of initial points: Theorem A allows the initial point to be anything, whereas Theorem B stipulates that the initial point should lie on the diagonal. We have had no success in extending the stronger result of Theorem A to dimensions $d \geq 3$ except



Cross Section (Projection on ν)

FIG. 1.

in one special case, the so-called "additive case," which will be discussed in a subsequent paper. However, it is possible to relax the hypothesis that the initial point lies on the diagonal to the somewhat weaker hypothesis that it merely lies "near" the diagonal. This is of some significance in sequential design of experiments, so we will given a complete statement of the result.

THEOREM C. *For each* $\alpha > 0$, $r_0 < \infty$ *there exists a constant* $C = C(\alpha, r_0)$ *such that for every* $DS(\alpha;\, r_0)$ *policy* $\mathscr{P}$ *and every initial point* $\mathbf{x} \in \mathbb{R}^d$ *such that* $|\mathbf{r}(\mathbf{x})| \le r_0$,

$$E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{T} \delta(\Gamma_n) \le \inf_{\mathscr{R}} E_{\mathbf{x}}^{\mathscr{R}} \sum_{n=1}^{T} \delta(\Gamma_n) + C(\alpha, r_0). \tag{1.9}$$

Here $T = T_0 = \inf\{n\colon \mathbf{W}_n \in \mathscr{Q}_+\}$ *is the first hitting time of the positive orthant* $\mathscr{Q}_+ \triangleq \{\mathbf{y} \in \mathbb{R}^d\colon \mathbf{y} > 0\}$.

Definition of a "policy". To allow for auxiliary randomization, but to avoid measure-theoretic complications [see, for example, Dynkin and Yushkevich (1979)], we adopt the following conventions regarding the notion of a policy. We assume that the underlying probability space $(\Omega, \mathscr{F}, P)$ supports independent iid sequences

$$
(1.10)\qquad
\begin{array}{ll}
Y_1^{(1)}, Y_2^{(1)}, \dots & \text{iid } F_1 \\
Y_1^{(2)}, Y_2^{(2)}, \dots & \text{iid } F_2 \\
 & \dots \\
Y_1^{(d)}, Y_2^{(d)}, \dots & \text{iid } F_d \\
U_1, U_2, \dots & \text{iid uniform } (0,1).
\end{array}
$$

By a "policy" $\mathscr{P}$ we will mean a sequence of functions $\psi_1, \psi_2, \dots$ valued in the finite set $\{1, 2, \dots, d\}$, with

$$
(1.11)\qquad
\begin{aligned}
\psi_1 &= \psi_1(\mathbf{x}; u_1), \\
\psi_2 &= \psi_2(\mathbf{x}; \mathbf{y}; u_1, u_2), \\
\psi_3 &= \psi_3(\mathbf{x}; \mathbf{y}_1, \mathbf{y}_2; u_1, u_2, u_3),
\end{aligned}
$$

etc.

The control variables Γ_n (under $P_x^{\mathscr{P}}$) are given by

$$
(1.12)\qquad
\begin{aligned}
\Gamma_1 &= \psi_1(\mathbf{x}; U_1), \\
\Gamma_2 &= \psi_2\big(\mathbf{x}; Y_1^{(\Gamma_1)}; U_1, U_2\big), \\
\Gamma_3 &= \psi_3\big(\mathbf{x}; Y_1^{(\Gamma_1)}, Y_2^{(\Gamma_2)}; U_1, U_2, U_3\big),
\end{aligned}
$$

etc.

The σ algebras $\mathscr{A}_n$ may be specified in a number of ways: The most natural filtration is

$$
(1.13)\qquad
\begin{aligned}
\mathscr{A}_0 &= \sigma(U_1), \\
\mathscr{A}_1 &= \sigma\big(U_1, U_2; \Gamma_1; Y_1^{(\Gamma_1)}\big), \\
\mathscr{A}_2 &= \sigma\big(U_1, U_2, U_3; \Gamma_1, \Gamma_2; Y_1^{(\Gamma_1)}, Y_2^{(\Gamma_2)}\big),
\end{aligned}
$$

etc.

Finally the process $\{\mathbf{W}_n\}$, which we will refer to as a "controlled random walk," is determined by

$$
(1.14)\qquad
\begin{aligned}
\mathbf{W}_0 &= \mathbf{x} \quad \big(\text{under } P_x^{\mathscr{P}}\big), \\
\mathbf{W}_{n+1} &= \mathbf{W}_n + \mathbf{w}_{n+1},
\end{aligned}
$$

where

$$
\mathbf{w}_n = Y_n^{(\Gamma_n)}.
$$

It is evident that $\mathscr{L}_x^{\mathscr{P}}(\mathbf{w}_{n+1} | \mathscr{A}_n) = F_{\Gamma_{n+1}}$.

These conventions will be in force in Sections 3 and 4. In Section 5 the notation will be altered somewhat.

2. First passage problems for a class of submartingales. The controlled random walks described in the introduction are not random walks (in the usual sense of the word); if the control policy is not stationary the march may not even be Markovian. To study first passage problems for such processes effectively a surrogate for the usual potential theory is therefore necessary.

Let $X_1, X_2, \ldots$ be a sequence of random variables adapted to the (nested) sequence of σ fields $\{\mathscr{F}_j\}_{j\geq 0}$: Let

$$\begin{aligned} \mu_j &= E\left[X_j | \mathscr{F}_{j-1}\right], \\ \sigma_j^2 &= \operatorname{var}(X_j | \mathscr{F}_{j-1}), \\ S_n &= X_1 + \cdots + X_n, \\ M_n &= \mu_1 + \cdots + \mu_n, \\ V_n &= \sigma_1^2 + \cdots + \sigma_n^2, \end{aligned} \tag{2.1}$$

and

$$\tau_a = \inf\{n \geq 1;\ S_n > a\}.$$

Assume that there exist constants μ_*, μ^*, and σ^2 such that with probability one, for all $j = 1, 2, \ldots,$

$$\text{(a)}\quad 0 < \mu_* \leq \mu_j \leq \mu^* < \infty, \quad \text{and} \quad \text{(b)}\quad \sigma_j^2 \leq \sigma^2. \tag{2.2}$$

Under these conditions it is natural to expect that $\{S_n\}_{n\geq 0}$ should exhibit some of the crude behavioral features of random walk with positive drift.

PROPOSITION 2.1. *There exist constants $C_1, C_2 < \infty$, depending only on μ_*, μ^*, and σ^2, such that*

$$E\left[\tau_a | \mathscr{F}_0\right] \leq C_1(a+1) \tag{2.3}$$

and

$$E\left[\tau_a^2 | \mathscr{F}_0\right] \leq C_1(a+1)^2. \tag{2.4}$$

PROOF. This is an easy exercise in the use of the first two Wald identities (i.e., the optional stopping theorem applied to the martingales $S_n - M_n$ and $((S_n - M_n)^2 - V_n)$.

Note first that there is a finite constant A such that $E[X_j \wedge A | \mathscr{F}_{j-1}] \geq \mu_*/2$, by (2.2): This is because

$$\begin{aligned} E\left[X_j 1\{X_j \geq A\} | \mathscr{F}_{j-1}\right] &\leq E\left[X_j^2 | \mathscr{F}_{j-1}\right]^{1/2} \cdot P(X_j \geq A | \mathscr{F}_{j-1})^{1/2} \\ &\leq \left((\mu^*)^2 + \sigma^2\right)^{1/2} \cdot \left(\sigma^2/(A - \mu^*)^2\right)^{1/2}. \end{aligned}$$

Since truncation of the increments X_j can only have the effect of increasing the first passage times τ_a, it suffices to prove the proposition under the additional hypothesis that $X_j \leq A$ w.p.l.

According to the first Wald identity, for each $n \in \mathbb{N}$ and $a > 0$, with $\tau = \tau_a$,

$$\begin{aligned} &E\left[S_{\tau\wedge n} | \mathscr{F}_0\right] = E\left[M_{\tau\wedge n} | \mathscr{F}_0\right] \Rightarrow \\ &\quad E\left[\tau \wedge n | \mathscr{F}_0\right] \cdot \mu_* \leq a + A \Rightarrow E\left[\tau | \mathscr{F}_0\right] \leq (a + A)/\mu_*. \end{aligned}$$

According to the second Wald identity,

$$E\left[(S_{\tau\wedge n} - M_{\tau\wedge n})^2|\mathscr{F}_0\right] = E[V_{\tau\wedge n}|\mathscr{F}_0]$$
$$\Rightarrow E\left[M^2_{\tau\wedge n}|\mathscr{F}_0\right] = E[V_{\tau\wedge n}|\mathscr{F}_0] - E\left[S^2_{\tau\wedge n}|\mathscr{F}_0\right] + 2E[M_{\tau\wedge n}S_{\tau\wedge n}|\mathscr{F}_0]$$
$$\leq \sigma^2 E[\tau|\mathscr{F}_0] + 2(a+A)\mu^* E[\tau|\mathscr{F}_0]$$
$$\Rightarrow (\mu_*)^2 E\left[\tau^2|\mathscr{F}_0\right] \leq \sigma^2 E[\tau|\mathscr{F}_0] + 2(a+A)\mu^* E[\tau|\mathscr{F}_0],$$

and thus (2.4) follows from (2.3). □

This very simple proposition will serve us well in the analysis to follow.

Having obtained bounds for the moments of the first passage times, we may now use them to obtain bounds for the moments of the excess over the boundary.

PROPOSITION 2.2. *Suppose in addition to* (2.2) *that there is a function* $h(z)\downarrow (z \geq 0)$ *satisfying*

$$\int_0^\infty h(z)\,dz < \infty \tag{2.5}$$

and

$$E\left[(X_n - z)_+^b|\mathscr{F}_{n-1}\right] \leq h(z) \quad \forall\, z \geq 0, \forall\, n. \tag{2.6}$$

Then there is a constant $C = C(\mu_*, \mu^*, \sigma^2, h) < \infty$ *such that for every* $a > 0$

$$E(S_{\tau_a} - a)^b \leq C. \tag{2.7}$$

Notice that if $E[(X_n)_+^{b+1+\varepsilon}|\mathscr{F}_{n-1}] \leq K$ for some $K < \infty$ and $\varepsilon > 0$, and all n, then (2.5) and (2.6) are satisfied by

$$h(z) = (1+K)(1 \wedge z^{-(1+\varepsilon)})$$

(this is just the Markov inequality). If $X_1, X_2, \dots$ are iid and $E(X_1)_+^{b+1} < \infty$, then (2.5) and (2.6) are satisfied by $h(z) = E(X_1 - z)_+^b$. However, in the general case the assumption $E[(X_n)_+^{b+1}|\mathscr{F}_{n-1}] \leq K < \infty$ is *not* enough to guarantee the conclusion (2.7).

EXAMPLE. The process $\{S_n\}_{n\geq 0}$ will be a Markov chain on the nonnegative integers $\mathbb{Z}^+$. For $x \in \mathbb{Z}^+$ with $4^k \leq x < 4^{k+1}$, let the transition mechanism be

$$\mathbf{P}(x, x+1) = 1 - (4^{k+1} - x)^{-2},$$
$$\mathbf{P}(x, 2^{2k+3} - x) = (4^{k+1} - x)^{-2}.$$

It is clear that (2.1) and (2.2) are satisfied, so the increments $S_n - S_{n-1} = X_n$ are L^2-bounded (conditional on $\mathscr{F}_{n-1}$); however, an easy computation shows that for $\tau = \tau_{4^k}$,

$$E(S_\tau - 4^k) \geq (\text{const})\cdot(k-1)$$

for some positive constant. □

Proof of Proposition 2.2. It obviously suffices to consider only integer values of a. Letting $\tau = \tau_a$, we have

$$\begin{aligned}
E(S_\tau - a)^b &\le E \sum_{w=0}^{a-1} \sum_{n=1+\tau_w}^{\tau_{w+1}} E\left[(X_n + S_{n-1} - a)_+^b | \mathscr{F}_{n-1}\right] \\
&\le E \sum_{w=0}^{a-1} \sum_{n=1+\tau_w}^{\tau_{w+1}} E\left[(X_n + w + 1 - a)_+^b | \mathscr{F}_{n-1}\right] \\
&\le E \sum_{w=0}^{a-1} \sum_{n=1+\tau_w}^{\tau_{w+1}} h(w+1-a) \\
&\le 2C_1 \cdot \sum_{w=0}^{a-1} h(w+1-a) \\
&\le 2C_1 \cdot \left(h(0) + \int_{z=0}^{\infty} h(z)\, dz \right),
\end{aligned}$$

where $C_1 < \infty$ is the constant in (2.3). □

Proposition 2.1 shows that, at least in a weak sense, $\{S_n\}$ drifts to $+\infty$ at a roughly linear rate. It is natural to inquire about the extent to which conditions (2.2) limit excursions below 0: More precisely, how big can the tail of the distribution of $\inf_{n \ge 0} S_n$ be?

Proposition 2.3. *Suppose in addition to* (2.2) *that there is a function* $h(z) \downarrow (z \ge 0)$ *such that*

$$\int_0^\infty z h(z)\, dz < \infty \tag{2.8}$$

and

$$P(X_n \le -z | \mathscr{F}_{n-1}) \le h(z) \quad \forall z \ge 0, \forall n. \tag{2.9}$$

Then there exists a constant $C = C(\mu_*, \mu^*, \sigma^2, h) < \infty$ *such that*

$$-E\left[\min_{n \ge 0} S_n | \mathscr{F}_0\right] \le C. \tag{2.10}$$

Remark. The proposition asserts not only the the expected value is finite, but that there is an upper bound depending only on the parameters μ_*, μ^*, σ^2, and h, and not on any other feature of the distribution of $\{S_n\}$.

In the special case where $X_1, X_2, \ldots$ are iid, this theorem was first noticed (we think) by Kiefer and Wolfowitz (1956) (cf. Theorem 5 there). We will use the Kiefer–Wolfowitz result in proving Proposition 2.3.

Proof. As in the proof of Proposition 2.1, we may assume that $X_n \le A$ almost surely, for all $n \ge 1$, where $A < \infty$ is a fixed constant. Let $H_{n+1}(x) = P(X_{n+1} \le x | \mathscr{F}_n)$ be defined for real x in such a way that $H_{n+1}(x)$ is, with probability one, a distribution function on $\mathbb{R}$. We may assume, without loss of generality, that $H_{n+1}(x)$ is, with probability one, a *continuous* distribution function.

[If $H_{n+1}(x)$ is not continuous, then modify the problem as follows. Let $\xi_1, \xi_2, \dots$ be iid with exponential distribution $P\{\xi_i \geq y\} = e^{-2y/\mu_*}$, and all independent of $\mathscr{F}_\infty \triangleq \sigma(\bigcup_{n=0}^\infty \mathscr{F}_n)$. Let

$$\tilde{X}_n = X_n - \xi_n,$$
$$\tilde{\mathscr{F}}_n = \sigma(\mathscr{F}_n, \xi_1, \dots, \xi_n),$$
$$\tilde{S}_n = \tilde{X}_1 + \cdots + \tilde{X}_n,$$
$$\tilde{H}_{n+1}(x) = P(\tilde{X}_{n+1} \leq x | \tilde{\mathscr{F}}_n),$$

etc. Then $\tilde{H}_{n+1}$ is continuous. Moreover, $\inf_{n\geq 0}\tilde{S}_n \leq \inf_{n\geq 0} S_n$. Thus if we prove (2.10) for $\tilde{X}_n$, $\tilde{S}_n$, etc., it will follow for X_n, S_n, etc.]

The first step of the argument consists of showing that there is a finite collection of probability distributions $\{G_1, G_2, \dots, G_l\}$ and a sequence $\{J_n\}$ of random variables valued in $\{1, 2, \dots, l\}$ such that

(2.11) $\quad J_{n+1}$ is measurable with respect to $\mathscr{F}_n$, $\quad \forall\, n$;

(2.12) $$G_j((-\infty, A]) = 1 \quad \forall\, j = 1, 2, \dots, l;$$

(2.13) $$\mu_*/2 \leq \int_{\mathbb{R}} x G_j(dx) \quad \forall\, j = 1, 2, \dots, l;$$

(2.14) $$\int_{\mathbb{R}} x^2 G_j(dx) < \infty \quad \forall\, j = 1, 2, \dots, l;$$

and

(2.15) $$H_{N+1}(x) \leq G_{J_{N+1}}(x) \quad \forall\, x \in \mathbb{R}, \quad \text{w.p.1.}$$

This is actually quite easy in view of (2.9). Choose a constant $B < \infty$ large enough that

$$\int_B^\infty h(x)\, dx < \mu_*/4,$$
$$h(B) < \varepsilon,$$

and hence

$$P(X_{n+1} \leq -B | \mathscr{F}_n) < \varepsilon \quad \forall\, n = 1, 2, \dots,$$

where $\varepsilon > 0$ is some small constant to be chosen later. Notice that such a $B < \infty$ exists by virtue of (2.8) and (2.9). We set $G_j(-z) = h(z)$ for all $z > B$, $j = 1, 2, \dots, l$ (l will be specified in due course). Clearly (2.12), (2.14), and (2.15) will hold regardless of how $G_1, G_2, \dots, G_l$ are defined on $[-B, A]$.

By Helly's Selection Theorem [cf. Feller (1966), Chapter VIII, Section 6], the family of subprobability measures on the interval $[-B, A]$ is compact in the weak-* topology. Consequently, there exists a finite collection $H_1, H_2, \dots, H_l$ of subprobability measures on $[-B, A]$, each with total mass $\int_{[-B, A]} H_j(dx) = 1 - h(B)$, and such that for every subprobability measure F on $[-B, A]$ with total mass $\geq 1 - h(B)$, there is a $j \in \{1, 2, \ldots, l\}$ with

$$|F(x) - H_j(x)| < 2\varepsilon, \qquad \forall\, x \in [-B, A].$$

Define $G_j(x) = (H_j(x) + 2\varepsilon) \wedge 1$ for $x \in [-B, A]$. Then

$$F(x) \leq G_j(x) \quad \forall\, x \in [-B, A]$$

and

$$|F(x) - G_j(x)| < 4\varepsilon \quad \forall\, x \in [-B, A].$$

It is now clear that if $\varepsilon > 0$ is sufficiently small ($\varepsilon < \mu_*/16$ should do), then (2.11)–(2.15) must be satisfied.

[NOTE. In defining J_n it might be that several different G_j would satisfy (2.15). In such a case J_n should be defined as the *least* j for which (2.15) holds; then there will be no measurability difficulties. Notice also that the selection of $\{G_1, \ldots, G_l\}$ depends only on μ_*, μ^*, σ^2, and h.]

The next part of the proof consists of showing that there exist independent sequences of random variables $\{Y_k^{(j)}\}$ $k = 1, 2, \ldots$ and $j = 1, 2, \ldots, l$ such that

$$\{Y_k^{(j)}\}_{k=1,2,\ldots} \text{ are iid } G_j; \tag{2.16}$$

and

$$Y^{(J_n)}_{N_n(J_n)} \leq X_n \quad \text{w.p.1.} \quad \forall\, n, \tag{2.17}$$

where

$$N_n(j) \triangleq \sum_{i=1}^{n} 1\{J_i = j\}.$$

However, it may be that we will need a somewhat larger probability space to manage this. We will assume, therefore, that the underlying probability space supports independent sequences $\{y_k^{(j)}\}_{k=1,2,\ldots}$, $j = 1, \ldots, l$ which are independent of $\mathscr{F}_\infty \triangleq \sigma(\bigcup_{n=0}^{\infty} \mathscr{F}_n)$, and such that for each $j = 1, 2, \ldots, l$, the random variables $\{y_k^{(j)}\}_{k=1,2,\ldots}$ are iid with distribution G_j.

Recall that the conditional distributions $H_{n+1}(x)$ were assumed to be continuous; thus it is permissible to define Z_{n+1} by

$$Z_{n+1} = G^{-1}_{J_{n+1}}(H_{n+1}(X_{n+1})).$$

By (2.11) and (2.15) it follows that

$$Z_{n+1} \leq X_{n+1} \quad \text{w.p.1} \quad \forall\, n \geq 0,$$

and

$$P(Z_{n+1} \leq x | \mathscr{F}_n) = G_{J_{n+1}}(x) \quad \forall\, x \in \mathbb{R}, \quad \text{w.p.1.}$$

Set

$$\begin{aligned} Y_k^{(j)} = Z_n \quad & \text{on } \{N_n(j) = k > N_{n-1}(j)\} \\ y_k^{(j)} \quad & \text{on } \left\{\sup_{n>0} N_n(j) < k\right\}; \end{aligned}$$

then by construction (2.16) and (2.17) are satisfied, and the sequences $\{Y_k^{(j)}\}$, $j = 1, 2, \dots, l$ are independent.

It is now apparent that

$$\inf_{n\geq 0} S_n \geq \inf_{n\geq 0}\left(Y_1^{(1)} + \cdots + Y_n^{(1)}\right) + \cdots + \inf_{n\geq 0}\left(Y_1^{(l)} + \cdots + Y_n^{(l)}\right).$$

By the Kiefer–Wolfowitz theorem,

$$\sum_{j=1}^{l} -E\left[\inf_{n\geq 0}\left(Y_1^{(j)} + \cdots + Y_n^{(j)}\right)|\mathscr{F}_0\right]$$
$$= C(G_1) + \cdots + C(G_l) < \infty. \qquad \square$$

The final result of this section concerns randomly stopped vector-valued processes. Let $\{X_n\}$, $\{S_n\}$, $\{\mu_n\}$, etc., satisfy conditions (2.1)–(2.2). In addition, suppose $\xi_1, \xi_2, \dots$ is a sequence of nonnegative random variables adapted to $\mathscr{F}_1, \mathscr{F}_2, \dots$, and let $\zeta_n = \xi_1 + \cdots + \xi_n$.

PROPOSITION 2.4. *Suppose that there exist continuous functions $h(z)\downarrow$ and $h^*(z)\downarrow$ such that*

$$\int_0^\infty z(h(z) + h^*(z))\,dz < \infty; \tag{2.18}$$

$$P(X_{n+1} \leq -z|\mathscr{F}_n) \leq h(z) \quad \forall\, z > 0 \quad \text{w.p.l;} \tag{2.19}$$

and

$$P(\xi_{n+1} \geq z|\mathscr{F}_n) \leq h^*(z) \quad \forall\, z > 0 \quad \text{w.p.l.} \tag{2.20}$$

Then there is a function $k(z)\downarrow$ such that $1 \geq k(z) \geq 0$,

$$\int_0^\infty zk(z)\,dz < \infty \tag{2.21}$$

and

$$P\left(\zeta_{\tau_1} \geq z|\mathscr{F}_0\right) \leq k(z) \quad \forall\, z > 0 \quad \text{w.p.l.} \tag{2.22}$$

NOTE. As in (2.1) $\tau_a = \min\{n \geq 0\colon S_n \triangleq X_1 + \cdots + X_n \geq a\}$.

COROLLARY 2.1. *Suppose $\nu \geq 1$ is an integer-valued random variable measurable $(\mathscr{F}_0)$, and suppose that for some σ algebra $\mathscr{F}_{-1} \subset \mathscr{F}_0$ and some sequence $g(n)\downarrow$ with $g(1) = 1$,*

$$\sum_1^\infty ng(n) < \infty, \tag{2.23}$$

and

$$P(\nu \geq n|\mathscr{F}_{-1}) \leq g(n). \tag{2.24}$$

If the hypotheses of Proposition 2.4 are satisfied, then

$$P\left(\zeta_{\tau_\nu} \geq z|\mathscr{F}_{-1}\right) \leq g^*(z) \quad \forall\, z \geq 0, \tag{2.25}$$

where $g^* \downarrow$ *and*

$$\int_0^\infty zg^*(z)\,dz < \infty. \tag{2.26}$$

Observe that the function $k(z)$ may always be chosen in such a way that $K(z) = 1 - k(z)$ is a distribution function on $[0, \infty)$. Iterating the inequality (2.22) gives

$$P\big(\zeta_{\tau_m} \geq z | \mathscr{F}_0\big) \leq 1 - K^{*m}(z)$$

for $m \geq 1$, where $K^{*m} = K*K* \dots *K$ is the usual convolution product. The inequality (2.25) follows from this for

$$g^*(z) = 1 - \sum_{n=1}^{\infty} (g(n) - g(n+1))K^{*n}(z),$$

since ν is $\mathscr{F}_0$-measurable. Notice that

$$\begin{aligned}\int_0^\infty zg^*(z)\,dz &= \sum_{n=1}^{\infty} (g(n) - g(n+1))\int_0^\infty z(1 - K^{*n}(z))\,dz \\ &\leq \sum_{n=1}^{\infty} n^2(g(n) - g(n+1))\int_0^\infty zk(z)\,dz < \infty.\end{aligned}$$

Thus Corollary 2.1 follows from (2.21) and (2.22).

The functions $k(z)$, $g^*(z)$, whose existence is asserted in Proposition 2.4 and Corollary 2.1, may actually be chosen in such a way that they depend only on μ_*, μ^*, σ^2, and the functions $h(z)$ and $h^*(z)$. This should be apparent from the proof.

The proof of Proposition 2.4 will, like the proof of Proposition 2.3, essentially consist of reducing the problem to the case of a (vector-valued) random walk with iid increments by stochastic comparison argument. In this case the reduction is substantially harder, however, since the "quantile transformation" is not available in dimensions larger than one, and since (2.19)–(2.20) give no information about the *joint* distribution of X_{n+1} and ξ_{n+1} conditional on $\mathscr{F}_n$. We know of no approach to Propositions 2.3 or 2.4 which does not make use of a stochastic comparison-coupling argument: In particular, it seems impossible to prove them by using "martingale" methods.

The iid analogues of Proposition 2.4 and Corollary 2.1 are easy consequences of the Wald identities for martingales.

LEMMA 2.1. *Suppose that the random vectors* (X_n, ξ_n) *are iid, and that* (X_{n+1}, ξ_{n+1}) *is independent of* $\mathscr{F}_n$ *for each* $n \geq 0$. *If*

$$\begin{aligned}&E\big(X_{n+1}^2 + \xi_{n+1}^2\big) < \infty, \\ &E(X_{n+1}) \equiv \mu > 0,\end{aligned} \tag{2.27}$$

and

$$\xi_{n+1} \geq 0 \quad \text{w.p.1}$$

for all $n \geq 0$, *then*

$$E\left(\zeta_{\tau_1}^2 | \mathscr{F}_0\right) = E\zeta_{\tau_1}^2 \leq 4\left\{(E\xi_1)^2 E\tau_1^2 + E\tau_1 \operatorname{var} \xi_1\right\} < \infty. \tag{2.28}$$

If in addition $\nu \geq 1$ *is an integer-valued random variable measurable with respect to* $\mathscr{G} \subset \mathscr{F}_1$, *if* (X_1, ξ_1) *is independent of* $\mathscr{G}$, *and if* $E(\nu^2 | \mathscr{F}_0) \leq C < \infty$ *for some constant* C, *then*

$$E\left(\zeta_{\tau_\nu}^2 | \mathscr{F}_0\right) \leq C\left(E\zeta_{\tau_1}\right)^2 + C^{1/2} E\zeta_{\tau_1}^2 < \infty. \tag{2.29}$$

Proof. By Proposition 2.1 $E(\tau_1^2 | \mathscr{F}_0) = E\tau_1^2 < \infty$. Now

$$\begin{aligned} E\left(\zeta_{\tau_1}^2 | \mathscr{F}_0\right) &= E\zeta_{\tau_1}^2 \\ &= E\left(\sum_{n=1}^{\tau_1} \xi_n\right)^2 \\ &= E\left(\sum_{n=1}^{\tau_1} (\xi_n - \mu)\right)^2 + 2\mu E\tau \sum_{n=1}^{\tau_1} (\xi_n - \mu) + \mu^2 E\tau^2. \end{aligned}$$

But the Cauchy–Schwarz inequality gives

$$\mu E\tau \sum_{n=1}^{\tau_1} (\xi_n - \mu) \leq \mu (E\tau^2)^{1/2} \left(E\left(\sum_{n=1}^{\tau_1} (\xi_n - \mu)\right)^2\right)^{1/2},$$

so

$$\begin{aligned} E\zeta_{\tau_1}^2 &\leq \left\{\mu (E\tau^2)^{1/2} + \left(E\left(\sum_{n=1}^{\tau_1} (\xi_n - \mu)\right)^2\right)^{1/2}\right\}^2 \\ &\leq 4\mu^2 E\tau_1^2 + 4E\left(\sum_{n=1}^{\tau_1} (\xi_n - \mu)\right)^2. \end{aligned}$$

Finally,

$$E\left(\sum_{n=1}^{\tau_1} (\xi_n - \mu)\right)^2 = E\tau \operatorname{var} \xi_1$$

by Wald's identity [cf., for example, Chow, Robbins, and Siegmund (1971), Theorem 2.3]. This proves (2.28).

Next notice that $(\zeta_{\tau_{k+1}} - \zeta_{\tau_k})$, $k = 0, 1, \ldots,$ are iid, with $(\zeta_{\tau_{k+1}} - \zeta_{\tau_k})$ independent of $\mathscr{F}_{\tau_k}$ for each $k = 0, 1, \ldots$ (here $\tau_0 = 0$). Also ζ_{τ_1} is independent of $\mathscr{G}$. Thus

$$\begin{aligned} E\left(\zeta_{\tau_\nu}^2 | \mathscr{F}_0\right) &= E\left(E\left(\left(\sum_{k=1}^{\nu} (\zeta_{\tau_k} - \zeta_{\tau_{k-1}})\right)^2 \middle| \mathscr{G}\right) \middle| \mathscr{F}_0\right) \\ &= E\left(\nu \cdot \operatorname{var}(\zeta_{\tau_1}) + \nu^2 (E\zeta_{\tau_1})^2 \middle| \mathscr{F}_0\right) \end{aligned}$$

and (2.29) follows. □

The quantile transformation which was used in proving Proposition 2.3 will be replaced by the transformation provided by the following lemma. Let

$$\mathscr{H} = \big\{\text{probability distributions } H \text{ on } \mathbb{R}^2_+ \text{ which are absolutely continuous, with } H(\{(x, y)\colon x \geq z\}) \leq h(z) \text{ and } H(\{(x, y)\colon y \geq z\}) \leq h^*(z) \quad \forall\, z \geq 0\big\},$$

where $h(\cdot)$ and $h^*(\cdot)$ are the functions from the hypothesis of Proposition 2.4. Let $\mathscr{B}$ be the smallest σ algebra of subsets of $\mathscr{H}$ such that all maps $H \to H(B)$ are measurable, where B is a Borel set of $\mathbb{R}^2_+$.

Lemma 2.2. *For any $\varepsilon > 0$ there exists a finite collection $\{G_1, G_2, \ldots G_l\}$ of probability measures on $\mathbb{R}^2_+$ and measurable* (wrt $\mathscr{B}$) *maps*

$$J\colon \mathscr{H} \to \{1, 2, \ldots, l\}$$

$$T\colon \mathscr{H} \times \mathbb{R}^2_+ \to \mathbb{R}^2_+$$

such that

(2.30) *for each $H \in \mathscr{H}$, $T(H; (x, y)) \geq (x, y)$ (relative to the usual partial order $\mathbb{R}^2_+$);*

(2.31) *for each $H \in \mathscr{H}$ and $f\colon \mathbb{R}^2_+ \to [0, 1]$ measurable,*

$$\int_{\mathbb{R}^2_+} f(T(H; (x, y)))\, dH(x, y) = \int_{\mathbb{R}^2_+} f((x, y))\, dG_{J(H)}(x, y);$$

$$\int x\, dG_{J(H)}(x, y) \leq \int x\, dH(x, y) + \varepsilon \quad \forall\, H \in \mathscr{H}; \tag{2.32}$$

and

$$G_j(\{(x, y)\colon \max(x, y) \geq z\}) \leq 2(h(z) + h^*(z)) \tag{2.33}$$

for all $j = 1, 2, \ldots, l$ and all large $z \geq 0$.

Thus whenever $\mathbf{X}$ is a random vector with law H, $T(H; \mathbf{X}) \geq \mathbf{X}$, and $T(H; \mathbf{X})$ has law $G_{J(H)}$. Furthermore, under (2.18), each of the distributions G_j has finite second moment.

There is a theorem in the literature which is closely related to Lemma 2.2 [cf. Kamae, Krengel, and O'Brien (1977), Theorem 1, who derive it as a special case of Theorem 11 in Strassen (1965)]. This theorem states that if H and G are probability distributions on $\mathbb{R}^d$ satisfying

$$\int f\, dG \leq \int f\, dH \quad \forall \text{ bounded increasing } f \text{ on } \mathbb{R}^d,$$

then on some probability space ([0, 1] with Lebesgue measure will do), there exist random vectors $\mathbf{X}, \mathbf{Y}$ such that $\mathbf{X} \leq \mathbf{Y}$ almost surely, $\mathbf{X}$ has law G, and $\mathbf{Y}$ has law H. Unfortunately, it seems impossible to conclude from this that the construction can be made to depend measurably on H and G (Strassen's proof is nonconstructive, relying on the Hahn–Banach Theorem).

The proof of Lemma 2.2 has been banished to the appendix.

PROOF OF PROPOSITION 2.4. We may assume, as in the proofs of Propositions 2.1 and 2.3, that $X_n \le A$ for some constant $A < \infty$. Define conditional distributions H_n on $\mathbb{R}^2_+$ by

$$H_{n+1}(z_1, z_2) = P(-(X_{n+1} - A) \le z_1;\ \xi_{n+1} \le z_2 | \mathscr{F}_n).$$

As in the proof of Proposition 2.3 we may assume that $H_{n+1}(z_1, z_2)$ is absolutely continuous in (z_1, z_2) (by adding small exponentially distributed random variables to the ξ_n and subtracting others from the X_n).

Again we assume that the underlying probability space is sufficiently accommodating to support sequences $\{\mathbf{y}_k(j)\}_{k=1,2,\dots}$, $(j = 1, \dots, l$ of random vectors which are independent of each other and of $\mathscr{F}_\infty \triangleq \sigma(\bigcup_{n=0}^\infty \mathscr{F}_n)$, and such that

$$P\{\mathbf{y}_k(j) \in d\mathbf{y}\} = G_j(d\mathbf{y}) \quad \forall\, \mathbf{y} \in \mathbb{R}^2_+,\ k \ge 1,\ j = 1, 2, \dots, l.$$

Here $G_1, G_2, \dots, G_l$ are the distributions provided by Lemma 2.3, with $\varepsilon = \mu_*/2$. [For Corollary 2.1 $\mathscr{F}_\infty$ should be defined as $\mathscr{F}_\infty = \sigma((\bigcup_{n=0}^\infty \mathscr{F}_n) \cup \mathscr{G})$].

Define random vectors $\{\mathbf{Z}_{n+1} = (Z^{(1)}_{n+1}, Z^{(2)}_{n+1})\}$ by

$$\mathbf{Z}_{n+1} = T(H_{n+1}; (-(X_{n+1} - A), \xi_{n+1})).$$

Notice that

$$-Z^{(1)}_{n+1} + A \le X_{n+1} \tag{2.34}$$

and

$$Z^{(2)}_{n+1} \ge \xi_{n+1}. \tag{2.35}$$

Notice also that each $\mathbf{Z}_{n+1}$ is measurable with respect to $\mathscr{F}_{n+1}$ since T is measurable and H_{n+1} (as an element of $\mathscr{K}$) is measurable with respect to $\mathscr{F}_n$. By (2.31)

$$P(\mathbf{Z}_{n+1} \in d\mathbf{z} | \mathscr{F}_n) = G_{J(H_{n+1})}(d\mathbf{z}). \tag{2.36}$$

As in the proof of Proposition 2.3, set

$$N_n(j) = \sum_{m=1}^n 1\{J(H_m) = j\}$$

and

$$\mathbf{Y}_k(j) = \mathbf{Z}_n \quad \text{on } \{N_n(j) = k > N_{n-1}(j)\}$$
$$\mathbf{y}_k(j) \quad \text{on } \left\{\sup_n N_n(j) < k\right\}.$$

Then once again, by construction, the sequences $\{\mathbf{Y}_k(j)\}_{k=1,2,\dots}$, $j = 1, 2, \dots, l$ are independent sequences of iid random vectors, with $\mathscr{L}(\mathbf{Y}_k(j)) = G_j$ [cf. (2.36)]. By (2.33)–(2.35)

$$\mathbf{Y}_{N_n(J(H_n))}(J(H_n)) \le (-(X_n - A), \xi_n) \quad \text{w.p.1.}$$

Finally, notice that because we chose $\varepsilon < \mu_*/2$, and by (2.32) and (2.2),

$$\sup_n N_n(j) = 0$$

for all $j \in \{1, 2, \dots, l\}$ such that

$$\int_{\mathbb{R}_+^2} x\, dG_j(x, y) > \mu_*/2. \tag{2.37}$$

The result (2.22) now follows easily from Lemma 2.1, since (2.28) and (2.29) may be applied to each of the independent sequences

$$\{\mathbf{X}_n(j), \tilde{\xi}_n(j)\},$$

where

$$\mathbf{X}_n(j) = A - Y_n^{(1)}(j)$$

and

$$\tilde{\xi}_n(j) = Y_n^{(2)}(j),$$

where $j \in \{1, 2, \dots, l\}$ is such that (2.37) does not hold (thus $E\tilde{X}_n(j) \geq \mu_*/2$). That the second moment condition is (2.27) of Lemma 2.1 holds is guaranteed by (2.18) and (2.33). □

3. A martingale. The purpose of this section is to identify a certain martingale function of $W_n - W_0$ and to verify that the martingale property persists under optional stopping (i.e., that the first Wald identity holds). This will give both an explicit representation and a lower bound for the expected cost $E_{\mathbf{x}}^{\mathscr{P}}(\sum_{n=1}^{T}\delta(\Gamma_n))$.

Recall from (1.6) that no convex combination of a subset of

$$\{\mu(F_1)/\delta(1), \dots, \mu(F_d)/\delta(d)\}$$

may dominate a convex combination of a disjoint subset. In other words, there do not exist a permutation π of $\{1, 2, \dots, d\}$ and probability vectors $(p_1, \dots, p_k), (q_{k+1}, \dots, q_d)$ such that

$$\sum_{j=1}^{k} p_j \mu(F_{\pi(j)}) \leq \sum_{j=k+1}^{d} q_j \mu(F_{\pi(j)}).$$

Recall also that each mean lies in the first orthant $\bar{\mathscr{Q}}_+$: i.e., $\mu(F_j) \geq 0$ for each $j = 1, \dots, d$.

LEMMA 3.1. *Under assumptions* (1.4) *and* (1.6), *the vectors* $\boldsymbol{\mu}(F_1)/\delta(1)$, $\boldsymbol{\mu}(F_2)/\delta(2), \dots, \boldsymbol{\mu}(F_d)/\delta(d)$ *are linearly independent.*

PROOF. Suppose there exist $\beta_1, \beta_2, \dots, \beta_d \in \mathbb{R}$, not all zero, such that

$$\sum_{i=1}^{d} \beta_i \boldsymbol{\mu}(F_i)/\delta(i) = \mathbf{0}.$$

Let

$$B_+ = \{i: \beta_i > 0\} \quad \text{and} \quad B_- = \{i: \beta_i < 0\};$$

then since $\boldsymbol{\mu}(F_i)/\delta(i) \geq \mathbf{0}$ for each $i = 1, \dots, d$, we must have both $B_- \neq \varnothing$ and

$B_+ \neq \varnothing$. [NOTE: We cannot have $\boldsymbol{\mu}(F_i)/\delta(i) = \mathbf{0}$ for any i, since $\boldsymbol{\mu}(F_i)/\delta(i)$ would then trivially be dominated by $\boldsymbol{\mu}(F_j)/\delta(j)$ for any other j, contradicting (1.6).]

Since both $B_- \neq \varnothing$ and $B_+ \neq \varnothing$, one of the following relations obtain [again, since each $\boldsymbol{\mu}(F_i)/\delta(i) \geq 0$]:

$$\sum_{B_+} (\boldsymbol{\mu}(F_i)/\delta(i)) \cdot \Big(\beta_i / \sum_{B_+} \beta_i\Big) \leq \sum_{B_-} (\boldsymbol{\mu}(F_i)/\delta(i)) \cdot \Big(\beta_i / \sum_{B_-} \beta_i\Big),$$

or

$$\sum_{B_-} (\boldsymbol{\mu}(F_i)/\delta(i)) \cdot \Big(\beta_i / \sum_{B_-} \beta_i\Big) \leq \sum_{B_+} (\boldsymbol{\mu}(F_i)/\delta(i)) \cdot \Big(\beta_i / \sum_{B_+} \beta_i\Big).$$

Both contradict assumption (1.6). □

LEMMA 3.2. *If assumptions* (1.4) *and* (1.6) *hold then there is a unit vector* $\mathbf{u} > 0$ *and a real number* $\mu > 0$, *such that for each* $i = 1, 2, \ldots, d$,

$$\mathbf{u}^t \boldsymbol{\mu}(F_i)/\delta(i) = \mu. \tag{3.1}$$

We will reserve the letter $\mathbf{u}$ *for the unit vector satisfying* (3.1) *throughout the paper.*

PROOF. This consists of showing that the hyperplane $\mathscr{Y} = \{\sum_{i=1}^d \beta_i \boldsymbol{\mu}(F_i)/\delta(i)$: $\sum_{i=1}^d \beta_i = 1, \beta_i \in \mathbb{R}\}$ contains points $\lambda_1 \mathbf{e}_1, \lambda_2 \mathbf{e}_2, \ldots, \lambda_d \mathbf{e}_d$, where each $\lambda_i > 0$ and $\mathbf{e}_i$ is the ith unit vector [$e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, 0, \ldots, 0)$, etc.]. For if this is true then the vector

$$\mathbf{u} = \Big(\prod_{i \neq 1} \lambda_i, \prod_{i \neq 2} \lambda_i, \ldots, \prod_{i \neq d} \lambda_i\Big) \Big/ \Big\| \Big(\prod_{i \neq 1} \lambda_i, \ldots, \prod_{i \neq d} \lambda_i\Big) \Big\|$$

will satisfy (3.2) (since $\mathbf{u}^t \mathbf{v} = \mu$ for $\mathbf{v} = \lambda_i \mathbf{e}_i$, $i = 1, \ldots, d$, it must be true for *every* $\mathbf{v} \in \mathscr{Y}$, as $\lambda_1 \mathbf{e}_1, \ldots, \lambda_d \mathbf{e}_d$ span $\mathscr{Y}$).

There are two ways that $\mathscr{Y}$ might fail to contain a positive multiple of $\mathbf{e}_1$: Either (i) $\lambda \mathbf{e}_1 \in \mathscr{Y}$ for some $\lambda \leq 0$, or (ii) $\lambda \mathbf{e}_1 \in \mathscr{Y}$ for *no* $\lambda \in \mathbb{R}$. We will show that each of these possibilities leads to a contradiction of assumption (1.6).

Consider first (i). If $\sum_{i=1}^d \beta_i \boldsymbol{\mu}(F_i)/\delta(i) = \lambda \mathbf{e}_1$ for some $\lambda \leq 0$ and $\beta_i \in \mathbb{R}$, $\sum_{i=1}^d \beta_i = 1$, then let

$$B_+ = \{i: \beta_i > 0\} \quad \text{and} \quad B_- = \{i: \beta_i < 0\}.$$

Recall that each $\boldsymbol{\mu}(F_i)/\delta(i) \geq \mathbf{0}$ and that *no* $\boldsymbol{\mu}(F_i)/\delta(i) = \mathbf{0}$ (Lemma 3.1). Thus in order for

$$\sum \beta_i \boldsymbol{\mu}(F_i)/\delta(i) = \lambda \mathbf{e}_i$$

with $\lambda \leq 0$, it must be the case that $B_- \neq \varnothing$. Moreover, since $\sum \beta_i = 1$,

$$0 < \sum_{B_-} (-\beta_i) < \sum_{B_+} \beta_i.$$

Using again the fact that $\mu(F_i)/\delta(i) \geq 0$ for each i, we conclude that

$$\sum_{B_+} (\mu(F_i)/\delta(i)) \cdot \left(\beta_i / \sum_{B_+} \beta_i \right) + \sum_{B_-} (\mu(F_i)/\delta(i)) \left(\beta_i / \sum_{-B} (-\beta_i) \right) \leq \lambda \mathbf{e}_1 \leq \mathbf{0},$$

which clearly contradicts assumption (1.6).

Consider now the possibility (ii). Notice first that the hyperplane $\mathscr{Y}$ is $(d-1)$-dimensional, since the vectors $\mu(F_1)/\delta(1), \dots, \mu(F_d)/\delta(d)$ are linearly independent. Thus if *no* real multiple of $\mathbf{e}_1$ is contained in $\mathscr{Y}$, then there exist reals $\alpha_1, \dots, \alpha_d$ and $\beta_1, \dots, \beta_d$ such that $\sum_{i=1}^d \alpha_i = \sum_{i=1}^d \beta_i = 1$, and

$$\sum_{i=1}^{d} (\alpha_i - \beta_i) \mu(F_i)/\delta(i) = \mathbf{e}_1.$$

Set

$$B_+ = \{i: (\alpha_i - \beta_i) > 0\} \quad \text{and} \quad B_- = \{i: (\alpha_i - \beta_i) < 0\};$$

by now familiar arguments we have $B_+ \neq \varnothing$ and $B_- \neq \varnothing$. We also have

$$\sum_{B_+} (\alpha_i - \beta_i) = \sum_{B_-} (\beta_i - \alpha_i),$$

so

$$\begin{aligned}\sum_{B_+} (\mu(F_i)/\delta(i)) \cdot & \left[(\alpha_i - \beta_i) / \sum_{B_+} (\alpha_i - \beta_i) \right] \\ &= \mathbf{e}_1 + \sum_{B_-} (\mu(F_i)/\delta(i)) \cdot \left[(\beta_i - \alpha_i) / \sum_{B_-} (\beta_i - \alpha_i) \right] \\ &\geq \sum_{B_-} (\mu(F_i)/\delta(i)) \cdot \left[(\beta_i - \alpha_i) / \sum_{B_-} (\beta_i - \alpha_i) \right],\end{aligned}$$

which again contradicts assumption (1.6). □

COROLLARY 3.1. *For any control policy $\mathscr{P}$ and any initial point $\mathbf{W}_0 = \mathbf{x}$,*

$$E_{\mathbf{x}}^{\mathscr{P}} \left[\mathbf{u}^t \mathbf{w}_{n+1} | \mathscr{A}_n \right] = \mu \delta(\Gamma_{n+1}). \tag{3.2}$$

In other words, the process $\{\mathbf{u}^t(\mathbf{W}_n - W_0) - \mu \sum_{j=1}^n \delta(\Gamma_j)\}_{n \geq 0}$ is a martingale relative to the filtration $\{\mathscr{A}_n\}_{n \geq 0}$.

This is an immediate consequence of Lemma 3.2 and the definition of a control policy (cf. Section 1).

COROLLARY 3.2. *For any control policy $\mathscr{P}$, any initial point $\mathbf{W}_0 = \mathbf{x}$, and any stopping time τ, either of the relations*

$$E_{\mathbf{x}}^{\mathscr{P}} \tau < \infty \tag{3.3}$$

or

$$E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{\tau} \delta(\Gamma_n) < \infty \tag{3.4}$$

implies

$$(3.5)\qquad \mu E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{\tau} \delta(\Gamma_n) = -\mathbf{u}^t\mathbf{x} + E_{\mathbf{x}}^{\mathscr{P}}(\mathbf{u}^t\mathbf{W}_\tau).$$

In particular, for any policy $\mathscr{P}$ *and any initial point* $\mathbf{W}_0 = \mathbf{x}$, *and* $T = \inf\{n\colon \mathbf{W}_n \in \mathscr{Q}_+\}$,

$$(3.6)\qquad E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{T} \delta(\Gamma_n) \geq (-\mathbf{u}^t\mathbf{x})/\mu.$$

With Corollary 3.2 it becomes clear that the problem of minimizing $E_{\mathbf{x}}^{\mathscr{P}}\Sigma_1^T\delta(\Gamma_n)$ is equivalent to minimizing $E_{\mathbf{x}}^{\mathscr{P}}(\mathbf{u}^t\mathbf{W}_T)$, i.e., to hitting the target $\mathscr{Q}_+$ as close to the corner as possible. It is also worth noting that the lower bound $(-\mathbf{u}^t\mathbf{x})/\mu$ is, for certain $\mathbf{x}$ (namely those $\mathbf{x}$ in the negative cone $-\mathscr{K} = \{\Sigma_1^d\lambda_i\mu(F_i)\colon \lambda_i \leq 0\}$) the solution of the linear programming problem

$$\min \sum_{i=1}^{d} s_i\delta(i)$$

$$\text{subject to } s_i \geq 0 \qquad (i = 1, \dots, d) \quad \text{and} \quad \sum_{i=1}^{d} s_i\mu(F_i) + \mathbf{x} \geq 0.$$

Thus the quantity $\inf_{\mathscr{P}} E_{\mathbf{x}}^{\mathscr{P}}(\mathbf{u}^t\mathbf{W}_T)$ may be interpreted as the additional cost for noise and discretization of time.

Notice that the condition $E_{\mathbf{x}}^{\mathscr{P}}\tau < \infty$ and $E_{\mathbf{x}}^{\mathscr{P}}\Sigma_1^\tau\delta(\Gamma_n) < \infty$ are equivalent, since $0 < \delta(i) < \infty$, for each $i = 1, 2, \dots, d$. Also (3.6) follows immediately from (3.5), since $\mathbf{u}^t\mathbf{W}_T \geq 0$ with probability one. Therefore Corollary 3.2 follows from Corollary 3.1 and the Wald Identity (Theorem 2.3 of Chow, Robbins, and Siegmund (1971).

4. Stability. Proof of Theorem C. The defining characteristic of a diagonal-stabilizing policy [cf. (1.7)] is that it applies a "restoring force" to the random walk when it begins to wander away from the diagonal. The optimality theorem (Theorem C) depends to a large extent on the fact that when the available control distributions have finite second moments, the random walk does not wander far away from the diagonal. Thus, in the regime of a diagonal-stabilizing policy, the diagonal is "stable" for the random walk; the main part of the proof of Theorem C consists of quantifying the stability property.

Before embarking on the proof of Theorem C is behooves us to make certain that, under the hypotheses (1.4)–(1.6), there do exist diagonal-stabilizing policies. Let $\mathscr{Y}$ be the linear subspace of $\mathbb{R}^d$ consisting of vectors orthogonal to the diagonal, i.e.,

$$\mathscr{Y} = \left\{\mathbf{v} \in \mathbb{R}^d\colon \sum_{i=1}^{d} v^{(i)} = 0\right\};$$

and let $\mathscr{U}$ be the unit vectors in $\mathscr{Y}$, i.e.,

$$\mathscr{U} = \left\{ \mathbf{v} \in \mathscr{Y} \colon \sum_{i=1}^{d} (v^{(i)})^2 = 1 \right\}.$$

LEMMA 4.1. *If conditions* (1.4)–(1.6) *are satisfied, then*

$$\inf_{\mathbf{v} \in \mathscr{U}} \max_{1 \le i \le d} \mathbf{v}^t \boldsymbol{\mu}(F_i) > 0. \tag{4.1}$$

PROOF. By Lemma 3.1 the vectors $\boldsymbol{\mu}(F_1), \dots, \boldsymbol{\mu}(F_d)$ are linearly independent, so for each $\mathbf{v} \in \mathscr{U}$ there exists $i \in \{1, 2, \dots, d\}$ such that

$$\mathbf{v}^t \boldsymbol{\mu}(F_i) \neq 0.$$

Since $\mathbf{1}$ is in the interior of the cone $\mathscr{H}$ generated by $\boldsymbol{\mu}(F_1), \dots, \boldsymbol{\mu}(F_d)$ moreover, there must exist positive numbers $\lambda_1, \lambda_2, \dots, \lambda_d$ such that

$$\mathbf{1} = \sum_{i=1}^{d} \lambda_i \boldsymbol{\mu}(F_i).$$

Consequently, for each $\mathbf{v} \in \mathscr{U}$,

$$\begin{aligned} 0 &= \mathbf{v}^t \mathbf{1} \\ &= \sum_{i=1}^{d} \lambda_i \big(\mathbf{v}^t \boldsymbol{\mu}(F_i) \big). \end{aligned}$$

It must be the case that at least one of $\mathbf{v}^t \boldsymbol{\mu}(F_1), \dots, \mathbf{v}^t \boldsymbol{\mu}(F_d)$ is strictly positive, since each $\lambda_i > 0$.

Now each of the functions $h_i(\mathbf{v}) = \mathbf{v}^t \boldsymbol{\mu}(F_i)$ is continuous in $\mathbf{v}$, and therefore

$$h(\mathbf{v}) \triangleq \max_{1 \le i \le d} h_i(\mathbf{v})$$

is continuous in $\mathbf{v}$. A continuous function on a compact set must attain its infimum, so by the result of the last paragraph

$$\inf_{\mathbf{v} \in \mathscr{U}} h(\mathbf{v}) > 0. \qquad \square$$

COROLLARY 4.1. *If conditions* (1.4)–(1.6) *are satisfied, then there exist diagonal-stabilizing policies.*

PROOF. As in Section 1 for each $\mathbf{x} \in \mathbb{R}^d$, let $\mathbf{d}(\mathbf{x})$ be the orthogonal projection of $\mathbf{x}$ onto the one-dimensional vector subspace of $\mathbb{R}^d$ containing the vector $\mathbf{1}$ and let $\mathbf{r}(\mathbf{x}) = \mathbf{d}(\mathbf{x}) - \mathbf{x}$. Then $\mathbf{r}(\mathbf{x})/|\mathbf{r}(\mathbf{x})| \in \mathscr{U}$ for each $\mathbf{x} \in \mathbb{R}^d$. By Lemma 4.1

$$\inf_{x \in \mathbb{R}^d} \max_{1 \le i \le d} \boldsymbol{\mu}(F_i)^t \frac{\mathbf{r}(\mathbf{x})}{|\mathbf{r}(\mathbf{x})|} > 0.$$

Let $\{\Gamma_n\}$ be defined by

$$\Gamma_{n+1} = \begin{cases} 1 & \text{if } |\mathbf{r}(\mathbf{W}_n)| = 0; \\ \min\left\{ i \colon \mu(F_i)^t \dfrac{\mathbf{r}(\mathbf{W}_n)}{|\mathbf{r}(\mathbf{W}_n)|} = \max\limits_{1 \le j \le d} \mu(F_j)^t \dfrac{\mathbf{r}(\mathbf{W}_n)}{|\mathbf{r}(\mathbf{W}_n)|} \right\} & \text{otherwise.} \end{cases}$$

Then clearly Γ_{n+1} is a measurable function of $\mathbf{W}_n$ and (1.7) is satisfied, so $\{\Gamma_n\}$ determines a diagonal-stabilizing policy. □

Stability. For a controlled random walk $\{\mathbf{W}_n\}_{n \ge 0}$, define stopping times $\{\eta(n)\}_{n \ge 0}$ by

$$\begin{aligned} \eta(0) &= 0, \\ \eta(n+1) &= \inf\{m > \eta(n) \colon |\mathbf{r}(\mathbf{W}_m)| < r\}, \end{aligned} \tag{4.2}$$

for some constant $r > 0$; and random variables $\{X_n\}$, $\{\mu_n\}$, and $\{\sigma_n^2\}$ by

$$\begin{aligned} X_n &= \mathbf{1}^t(\mathbf{W}_{\eta(n)} - \mathbf{W}_{\eta(n-1)}), \\ S_n &= \sum_{j=1}^{n} X_j, \\ \mu_n &= E_{\mathbf{x}}^{\mathscr{P}}\left[X_n | \mathscr{A}_{\eta(n-1)}\right], \\ \sigma_n^2 &= \operatorname{var}_{\mathbf{x}}^{\mathscr{P}}\left(X_n | \mathscr{A}_{\eta(n-1)}\right). \end{aligned} \tag{4.3}$$

The intervals $\{m \colon \eta(n-1) < m \le \eta(n)\}$ should be thought of as "excursions" away from the diagonal; thus $\eta(n) - \eta(n-1)$ is the length of the excursion and X_n is the displacement of $\{\mathbf{W}_m\}$ in the direction $\mathbf{1}$ over the excursion.

PROPOSITION 4.1. *For each $\alpha > 0$ and $r < \infty$ there exist constants μ_*, μ^*, σ^2, and β and a function $h(z)\downarrow$ with $\int_0^\infty h(z)\,dz < \infty$ (all depending only on α, r, and the underlying distributions $F_1, \dots, F_d$) such that for each DS(α; r) policy $\mathscr{P}$ and each $\mathbf{x} \in \mathbb{R}^d$ with $|\mathbf{r}(\mathbf{x})| < r$,*

$$0 < \mu_* \le \mu_n = E_{\mathbf{x}}^{\mathscr{P}}\left[X_n | \mathscr{A}_{\eta(n-1)}\right] \le \mu^* < \infty \qquad \forall\, n \ge 1, \tag{4.4}$$

$$\sigma_n^2 = E_{\mathbf{x}}^{\mathscr{P}}\left[(X_n - \mu_n)^2 | \mathscr{A}_{\eta(n-1)}\right] \le \sigma^2 < \infty \qquad \forall\, n \ge 1, \tag{4.5}$$

$$E_{\mathbf{x}}^{\mathscr{P}}\left[(X_n - z)_+ | \mathscr{A}_{\eta(n-1)}\right] \le h(z) \quad \forall\, z \ge 0, \quad \forall\, n \ge 1, \tag{4.6}$$

and

$$E_{\mathbf{x}}^{\mathscr{P}}\left[\eta(n) - \eta(n-1) | \mathscr{A}_{\eta(n-1)}\right] \le \beta \qquad \forall\, n \ge 1. \tag{4.7}$$

This result limits, if only in a weak sense, the average "size" of an excursion away from the diagonal, hence it seems natural to call it a stability theorem. The diagonal-stabilizing condition (1.7) is a natural stochastic analogue of the Lyapunov function often used for establishing stability of an ordinary differential equation [cf., for example, Hale (1969), Chapter 20].

Before proving Proposition 4.1 we will show how, in conjunction with Proposition 2.2 and Corollary 3.2, it may be used to prove Theorem C.

PROOF OF THEOREM C. Let

$$\tau = \inf_{n \geq 0} \left\{ \eta(n) \colon \mathbf{1}^t \mathbf{W}_{\eta(n)} > r \right\};$$

clearly $\tau \geq T = \min\{n\colon \mathbf{W}_n \in \mathscr{Q}_+\}$, and, since $\delta(i) > 0$ for each i, $\sum_{n=1}^{\tau} \delta(\Gamma_n) \geq \sum_{n=1}^{T} \delta(\Gamma_n)$. We will show that, for some constant $C(\alpha, r) < \infty$

$$E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{\tau} \delta(\Gamma_n) \leq -\mathbf{u}^t\mathbf{x} + C(\alpha, r) \tag{4.8}$$

for all initial points $\mathbf{x} \in \mathbb{R}^d$ such that $|\mathbf{r}(\mathbf{x})| < r$. (Recall that $\mathbf{u}$ is the unit vector defined in Lemma 3.2.) In view of Corollary 3.2 this will suffice to establish Theorem C.

Now Corollary 3.2 guarantees that if $E_{\mathbf{x}}^{\mathscr{P}}\tau < \infty$, then

$$E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{\tau} \delta(\Gamma_n) = -\mathbf{u}^t\mathbf{x} + E_{\mathbf{x}}^{\mathscr{P}} \mathbf{u}^t \mathbf{W}_\tau.$$

Hence to prove (4.8) we need only show

$$E_{\mathbf{x}}^{\mathscr{P}}\tau < \infty \quad \text{for each } \mathbf{x} \in \mathbb{R}^d, \qquad |\mathbf{r}(\mathbf{x})| < r; \tag{4.9}$$

and

$$E_{\mathbf{x}}^{\mathscr{P}} |W_\tau| \leq C(\alpha, r) \quad \text{for each } \mathbf{x} \in \mathbb{R}^d, \qquad |\mathbf{r}(\mathbf{x})| < r. \tag{4.10}$$

Let $\nu = \inf\{n\colon S_n > \mathbf{1}^t\mathbf{W}_0 + r\}$ then $\tau = \eta(\nu)$, and furthermore, ν is a stopping time of the type considered in Section 2. Consequently, by Proposition 2.1 and (4.4)–(4.5) of Proposition 4.1,

$$E_{\mathbf{x}}^{\mathscr{P}} \nu < \infty \quad \forall\, \mathbf{x} \in \mathbb{R}^d, \qquad |\mathbf{r}(\mathbf{x})| < r. \tag{4.11}$$

Also, by Proposition 2.2 and (4.4)–(4.6) of Proposition 4.1 there exists a constant $\tilde{C}(\alpha, r) < \infty$ such that for every $x \in \mathbb{R}^d$ with $|\mathbf{r}(\mathbf{x})| < r$

$$E_{\mathbf{x}}^{\mathscr{P}}\big(S_\nu - (\mathbf{1}^t\mathbf{x} + r)\big) \leq \tilde{C}(\alpha, r). \tag{4.12}$$

The inequalities (4.9) and (4.10) follow easily from (4.11) and (4.12). First, since $\tau = \eta(\nu)$ and $\mathbf{W}_\tau$ is at a distance no greater than r from the diagonal, $|\mathbf{W}_\tau| \leq 2r + (S_\nu - (\mathbf{1}^t\mathbf{W}_0 + r))$, so (4.10) follows immediately from (4.12). To get (4.9) from (4.11) reason as follows:

$$\begin{aligned} \tau = \eta(\nu) &= \sum_{n=1}^{\nu} (\eta(n) - \eta(n-1)) \\ &= \lim_{m \uparrow \infty} \uparrow \sum_{n=1}^{\nu \wedge m} (\eta(n) - \eta(n-1)), \end{aligned}$$

so by (4.7), (4.11) and the optional stopping theorem for supermartingales

$$\begin{aligned} E_{\mathbf{x}}^{\mathscr{P}}\tau &= \lim_{m} \uparrow E_{\mathbf{x}}^{\mathscr{P}} \sum_{n=1}^{\nu \wedge m} (\eta(n) - \eta(n-1)) \\ &\leq \lim_{m} \uparrow \big(\beta \cdot E_{\mathbf{x}}^{\mathscr{P}}(\nu \wedge m)\big) \\ &\leq \beta \cdot E_{\mathbf{x}}^{\mathscr{P}} \nu < \infty. \end{aligned}$$

□

The remainder of this section will be devoted to the proof of Proposition 4.1. The general results of Section 2, especially Proposition 2.4 and Corollary 2.1, will play a key role.

Let

$$\begin{aligned} R_n &= |\mathbf{r}(\mathbf{W}_n)| \\ &= \operatorname{dist}(\mathbf{W}_n, \text{diagonal}). \end{aligned}$$

LEMMA 4.2. *For each fixed $\alpha > 0$ there is an $r > 0$ so large that for every $DS(\alpha;\, r)$ policy $\mathscr{P}$*

$$E_x^{\mathscr{P}}(R_n - R_{n+1}|\mathscr{A}_n) \geq \alpha/2 \quad \text{on } \{R_n > r\}. \tag{4.13}$$

The upshot of this is that the process $-R_n$ satisfies the conditions (2.2) of Section 2. [The fact that (2.2)(b) holds follows from the fact that each increment $\mathbf{w}_n$ has its law (conditional on $\mathscr{F}_{n-1}$) in the finite collection $\{F_1, \dots, F_d\}$, and $\int_{\mathbb{R}^d} |v|^2 F_i(dv) < \infty$ for each i.]

PROOF. Since each F_i has a finite first moment

$$\int_{\{\mathbf{w}: |\mathbf{w}| > z\}} |\mathbf{w}| F_i(d\mathbf{w}) \to 0 \quad \text{as } x \to +\infty.$$

Thus it is possible to choose $z < \infty$ so large that all of the integrals $(i = 1, 2, \dots, d)$ are less than $\alpha/8$.

Next, the geometry of $\mathbb{R}^d$ guarantees that for any fixed $z > 0$, the level surfaces $\{\mathbf{w}\colon |\mathbf{r}(\mathbf{w})| = \text{const}\}$ differ by increasingly small amounts from hyperplanes over patches of diameter z, as (const) $\to \infty$. Consequently, for r sufficiently large,

$$\begin{aligned} &E_{\mathbf{x}}^{\mathscr{P}}((R_n - R_{n+1}) 1\{|\mathbf{w}_{n+1}| \leq z\}|\mathscr{A}_n) 1\{R_n > r\} \\ &\quad \geq E_{\mathbf{x}}^{\mathscr{P}}\left((3/4)\mathbf{w}_{n+1}^t \left(\frac{\mathbf{r}(\mathbf{w}_n)}{R_n}\right) 1\{|\mathbf{w}_{n+1}| \leq z\} \Big| \mathscr{A}_n\right) 1\{R_n > r\}. \end{aligned}$$

It then follows from the defining property of a diagonal-stabilizing policy [cf. (1.7)] that

$$E_{\mathbf{x}}^{\mathscr{P}}((R_n - R_{n+1})|\mathscr{F}_n) 1\{R_n > r\} \geq (3/4)\alpha - 2 \cdot (\alpha/8). \qquad \square$$

PROOF OF PROPOSITION 4.1.

(1) PROOF OF (4.7). For each integer $m \geq 0$ define a sequence $\{y_n\}_{n \geq 1}$ of random variables as follows:

$$\begin{aligned} y_n &= R_{\eta(m)+n} - R_{\eta(m)+n+1} \quad \text{if } \eta(m) + n \leq \eta(m+1) \\ &= 1 \quad \text{otherwise.} \end{aligned}$$

These random variables satisfy the conditions (2.2) (with respect to the σ

algebras $\{\mathscr{A}_{\eta(m)+n}\}_{n\geq 1}$) because by Lemma 4.2

$$E_{\mathbf{x}}^{\mathscr{P}}\big(y_n|\mathscr{A}_{\eta(m)+n}\big) \geq \alpha/2,$$

and clearly, since $\mathbf{w}_{n+1}$ has its (conditional) distribution in the finite collection $\{F_1,\ldots,F_d\}$,

$$E_{\mathbf{x}}^{\mathscr{P}}\big(y_n^2|\mathscr{A}_{\eta(m)+n}\big) \leq 1 + \sum_{i=1}^{d}\int_{\mathbb{R}^d}|\mathbf{v}|^2 F_i(d\mathbf{v}).$$

Notice also that the constants μ_*, μ^*, and σ^2 in (2.2) can be chosen in such a way that they depend only on α and the distributions $F_1,\ldots,F_d$, not on the initial point $\mathbf{x}$ or the policy $\mathscr{P}$.

As in (2.1), define first passage times

$$\tau(a) = \min\{n \geq 0\colon y_1 + y_2 + \cdots + y_n > a\}.$$

Then it is evident that

$$\begin{aligned}\eta(m+1) &= \eta(m) + \tau\big(R_{\eta(m)+1} - r\big) &&\text{if } R_{\eta(m)+1} \geq r\\ &= \eta(m) + 1 &&\text{if } R_{\eta(m)+1} < r,\end{aligned}$$

since by definition $\eta(m+1)$ is the first $n > \eta(m)$ for which $R_n < r$. Consequently, by Proposition 2.1

$$E_{\mathbf{x}}^{\mathscr{P}}\big(\eta(m+1) - \eta(m)|\mathscr{A}_{\eta(m)+1}\big) \leq C_1\big(\mu_*,\mu^*,\sigma^2\big)\big(R_{\eta(m)+1}\big) + 1$$

implies that

$$E_{\mathbf{x}}^{\mathscr{P}}\big(\eta(m+1) - \eta(m)|\mathscr{A}_{\eta(m)}\big) \leq C_1\big(\mu_*,\mu^*,\sigma^2\big)\left(r + \sum_{i=1}^{d}\int_{\mathbb{R}^d}|\mathbf{v}|F_i(d\mathbf{v})\right) + 1,$$

for each $x \in \mathbb{R}^d$ with $|\mathbf{r}(\mathbf{x})| < r$. This proves (4.7).

Similarly, Proposition 2.1 implies that

$$\text{(4.14)}\quad E_{\mathbf{x}}^{\mathscr{P}}\Big(\big(\eta(m+1) - \eta(m) - 1\big)^2\big|\mathscr{A}_{\eta(m)}\Big) \leq C_2\sum_{i=1}^{d}\int_{\mathbb{R}^d}\big(r + |\mathbf{v}|\big)^2 F_i(d\mathbf{v}),$$

for all $\mathbf{x} \in \mathbb{R}^d$ with $|\mathbf{r}(\mathbf{x})| < r$ and $C_2 = C_2(\mu_*,\mu^*,\sigma^2)$ the constant in (2.4).

(2) Proof of (4.4). Recall that $X_n = \sum_{j=1}^{\eta(n)-\eta(n-1)}\mathbf{1}^t\mathbf{w}_{\eta(n-1)+j}$ where $\mathbf{w}_k = \mathbf{W}_k - \mathbf{W}_{k-1}$ is the kth increment in the (controlled) random walk. Now

$$\begin{aligned}0 &< \min_{i=1,2,\ldots,d}\mathbf{1}^t\boldsymbol{\mu}(F_i)\\ &\leq E_{\mathbf{x}}^{\mathscr{P}}\big(\mathbf{1}^t\mathbf{w}_{\eta(n-1)+j}|\mathscr{A}_{\eta(n-1)+j-1}\big)\\ &\leq \max_{i=1,2,\ldots,d}\mathbf{1}^t\boldsymbol{\mu}(F_i)\\ &< \infty,\end{aligned}$$

and

$$E_{\mathbf{x}}^{\mathscr{P}}\Big(\big(\mathbf{1}^t\mathbf{w}_{\eta(n-1)+j}\big)^2\big|\mathscr{A}_{\eta(n-1)+j-1}\Big) \leq d\cdot\sum_{i=1}^{d}\int_{\mathbb{R}^d}|\mathbf{v}|^2F_i(d\mathbf{v}).$$

Consequently, by (4.7) and the Wald identity

$$
\begin{aligned}
0 &< \beta \cdot \min_{1 \le i \le d} \mathbf{1}^t \boldsymbol{\mu}(F_i) \\
&\le E_{\mathbf{x}}^{\mathscr{P}}(X_n | \mathscr{A}_{\eta(n-1)}) \\
&\le \beta \cdot \max_{1 \le i \le d} \mathbf{1}^t \boldsymbol{\mu}(F_i),
\end{aligned}
$$

for all $\mathbf{x} \in \mathbb{R}^d$ with $|\mathbf{r}(\mathbf{x})| < r$. This proves (4.4).

(3) Proof of (4.5). For notational convenience (in this paragraph only) set

$$\nu(n) = \eta(n) - \eta(n-1)$$

and

$$Y_j = \mathbf{1}^t \mathbf{w}_{\eta(n-1)+j} - E\left(\mathbf{1}^t \mathbf{w}_{\eta(n-1)+j} | \mathscr{A}_{\eta(n-1)+j}\right).$$

By (4.7) $E(\nu(n) | \mathscr{A}_{\eta(n-1)}) \le \beta$ and $\nu(n)$ is a stopping time relative to the filtration $\{\mathscr{A}_{\eta(n-1)+j}\}_{j \ge 0}$, so by the Wald identity

$$
\begin{aligned}
\mathrm{E}_{\mathbf{x}}^{\mathscr{P}}\left((X_n - \mu_n)^2 | \mathscr{A}_{\eta(n-1)}\right) &= E_{\mathbf{x}}^{\mathscr{P}}\left(\left(\sum_{j=1}^{\nu(n)} Y_j\right)^2 \middle| \mathscr{A}_{\eta(n-1)}\right) \\
&= E_{\mathbf{x}}^{\mathscr{P}}\left(\sum_{j=1}^{\nu(n)} E_{\mathbf{x}}^{\mathscr{P}}\left(Y_j^2 | \mathscr{A}_{\eta(n-1)+j}\right) \middle| \mathscr{A}_{\eta(n-1)}\right) \\
&\le E_{\mathbf{x}}^{\mathscr{P}}\left(\nu(n) \cdot \left(\sum_{i=1}^{d} \int |\mathbf{1}^t \mathbf{w}|^2 F_i(d\mathbf{w})\right) \middle| \mathscr{A}_{\eta(n-1)}\right) \\
&\le \beta \cdot \sum_{i=1}^{d} \int |\mathbf{1}^t \mathbf{w}|^2 F_i(d\mathbf{w}).
\end{aligned}
$$

(4) Proof of (4.6). This is really the large intestine of the proof. We actually did most of the work in Section 2—namely, Proposition 2.4—in the hope that the entire project would be more easily digested.

Set

$$
\begin{aligned}
y_n &= R_{\eta(m)+n} - R_{\eta(m)+n+1} \quad \text{if } \eta(m) + n + 1 \le \eta(m+1) \\
&= 1 \quad \text{otherwise,} \\
Y_n &= \left(\mathbf{1}^t \mathbf{w}_{\eta(m)+n+1}\right)_+
\end{aligned}
$$

and

$$\nu = d \cdot |\mathbf{w}_{\eta(m)+1}| + r.$$

Then

$$
\begin{aligned}
\eta(m+1) &\le \eta(m) + 2 + \min\{n : y_1 + \cdots + y_n \ge \nu + r\} \\
&= \eta(m) + 2 + \tau(\nu + r).
\end{aligned}
$$

Notice that

$$P_{\mathbf{x}}^{\mathscr{P}}\left(y_n \leq -z \mid \mathscr{A}_{\eta(m)+n}\right) \leq \sum_{i=1}^{d} \int_{\{\mathbf{v}: |\mathbf{v}| \geq z\}} F_i(d\mathbf{v})$$
$$\triangleq h(z);$$

also

$$P_{\mathbf{x}}^{\mathscr{P}}\left(Y_n \geq z \mid \mathscr{A}_{\eta(m)+n}\right) \leq \sum_{i=1}^{d} \int_{\{\mathbf{v}: \mathbf{1}^t \mathbf{v} \geq z\}} F_i(dv)$$
$$\triangleq h^*(z);$$

and $\int_0^\infty z(h(z) + h^*(z))\, dz < \infty$ because of the fact that each F_i has a finite second moment! In addition,

$$X_{m+1} = \mathbf{1}^t\left(\mathbf{W}_{\eta(m+1)} - \mathbf{W}_{\eta(m)}\right)$$
$$\leq \nu + \left(Y_1 + Y_2 + \cdots + Y_{\tau(\nu+r)}\right).$$

Since

$$P_{\mathbf{x}}^{\mathscr{P}}\left(\nu \geq z \mid \mathscr{A}_{\eta(m)}\right) \leq \sum_{i=1}^{d} \int_{\{\mathbf{v}: d|\mathbf{v}| + R \geq z\}} F_i(d\mathbf{v}),$$

the conditions of Corollary 2.1 are satisfied, and hence

$$P_{\mathbf{x}}^{\mathscr{P}}\left(X_{m+1} \geq z \mid \mathscr{A}_{\eta(m)}\right) \leq g^*(z) \quad \forall\, z > 0, \quad \text{w.p.1},$$

where $g^*(z)$ is the function given in Corollary 2.1. This clearly implies (4.6). □

5. Global optimality in dimension 2. Using Theorem C we now proceed to prove the much stronger Theorem A for straight-line switching policies. In fact an analogous result holds for *all* diagonal-stabilizing policies in dimension 2, but we will refrain from proving this.

The proof of Theorem A essentially consists of a path-by-path comparison of two controlled random walks, one evolving according to the straight-line switching rule $\mathscr{P}$, the other according to a competing policy $\mathscr{R}$. Such a comparison requires that the two random walks be constructed on the same probability space; this in turn will force us to abandon some of the notation and conventions established in the preceding sections.

Let $(\Omega, \mathscr{F}, P)$ be a probability space on which are defined two independent sequences of iid random vectors, and an independent sequence of iid uniform-(0, 1) variables:

$$\mathbf{Y}_1, \mathbf{Y}_2, \ldots \text{ are iid } (F_A),$$
$$\mathbf{Z}_1, \mathbf{Z}_2, \ldots \text{ are iid } (F_B),$$

and

$$U_0, U_1, \ldots \text{ are idd uniform on } (0, 1).$$

Recall that a (randomized) control policy $\mathscr{R}$ consists of a sequence of functions

$\psi_0, \psi_1, \ldots$ which carry instructions for choosing from the two available controls A and B [cf. (1.12)]. For each $\mathbf{x} \in \mathbb{R}^2$, let

$$
\begin{aligned}
&\mathbf{W}_0(\mathbf{x}, \mathscr{R}) = \mathbf{x}, \\
(5.1)\quad &\mathbf{W}_{n+1}(x, \mathscr{R}) = \mathbf{W}_n(\mathbf{X}, \mathscr{R}) + \mathbf{Y}_{n+1} \quad \text{if } \psi_n(\mathbf{W}_0, \ldots, \mathbf{W}_n; U_0, \ldots, U_n) = \mathrm{A} \\
&\qquad\qquad = \mathbf{W}_n(\mathbf{x}, \mathscr{R}) + \mathbf{Z}_{n+1} \quad \text{if } \psi_n(\mathbf{W}_0, \ldots, \mathbf{W}_n; U_0, \ldots, U_n) = \mathrm{B}.
\end{aligned}
$$

Then clearly $\{\mathbf{W}_n(\mathbf{x}, \mathscr{R})\}_{n \ge 0}$ evolves according to the policy $\mathscr{R}$. Similarly, let

$$
\begin{aligned}
&\mathbf{W}_0(\mathbf{x}) = \mathbf{x}, \\
(5.2)\qquad &\mathbf{W}_{n+1}(\mathbf{x}) = \mathbf{W}_n(x) + \mathbf{Y}_{n+1} \quad \text{if } W_n^{(2)}(\mathbf{x}) \ge W_n^{(1)}(\mathbf{x}) \\
&\qquad\qquad = \mathbf{W}_n(x) + \mathbf{Z}_{n+1} \quad \text{if } W_n^{(2)}(\mathbf{x}) < W_n^{(1)}(\mathbf{x}),
\end{aligned}
$$

where, as usual, $W^{(1)}$ and $W^{(2)}$ are the first and second coordinates of $\mathbf{W}$. The process $\{\mathbf{W}_n(\mathbf{x})\}_{n \ge 0}$ evolves according to the straight-line switching rule $\mathscr{P}$.

NOTE. We will continue to use a lower case $\mathbf{w}$ to denote the increments in the controlled random walks W. Thus

$$\mathbf{w}_{n+1}(\mathbf{X}) = \mathbf{W}_{n+1}(\mathbf{x}) - \mathbf{W}_n(\mathbf{x})$$

and

$$\mathbf{w}_{n+1}(\mathbf{x}, \mathscr{R}) = \mathbf{W}_{n+1}(\mathbf{x}, \mathscr{R}) - \mathbf{W}_n(\mathbf{x}, \mathscr{R}).$$

REMARK. Strictly speaking, a straight-line switching policy, as defined in Section 1, may use *any* line (ray) as the switching line, as long as it lies in the negative cone determined by $\boldsymbol{\mu}_A$ and $\boldsymbol{\mu}_B$. However, we may, by a simple linear transformation of $\mathbb{R}^2$ having the form $(x^{(1)}, x^{(2)}) \to (c_1 x^{(1)}, c_2 x^{(2)})$ with $c_1, c_2 > 0$, always reduce the problem to one where the switching line is the diagonal $x^{(1)} = x^{(2)}$, and the mean vectors satisfy

$$
(5.3)\qquad
\begin{aligned}
&\mu_A^{(1)} > \mu_B^{(1)} \ge 0, \\
&\mu_B^{(2)} > \mu_A^{(2)} \ge 0, \\
&\mu_A^{(1)} > \mu_A^{(2)}, \quad \text{and} \\
&\mu_B^{(2)} > \mu_B^{(1)}.
\end{aligned}
$$

More notation will be needed. Let

$$(5.4)\qquad T(\mathbf{x}) = \inf\{n \ge 0 \colon \mathbf{W}_n(\mathbf{x}) \in \mathscr{Q}_+\}$$

and

$$T(\mathbf{x}, \mathscr{R}) = \inf\{n \ge 0 \colon \mathbf{W}_n(\mathbf{x}, \mathscr{R}) \in \mathscr{Q}_+\}.$$

Also let A_+, A_0, and A_- be the sets

$$
(5.5)\qquad
\begin{aligned}
&A_+ = \{(x^{(1)}, x^{(2)}) \colon x^{(2)} \ge x^{(1)} + r\}, \\
&A_- = \{(x^{(1)}, x^{(2)}) \colon x^{(1)} \ge x^{(2)} + r\}, \\
&A_0 = \mathbb{R}^2 \setminus (A_+ \cup A_-).
\end{aligned}
$$

Here $r > 0$ is a large but fixed constant. Define stopping times

$$\tau_F(\mathbf{x}) = \inf\{n \geq 0 \colon \mathbf{W}_n(\mathbf{x}) \in F\}. \tag{5.6}$$

We note at the outset that $\{W_n^{(i)}(\mathbf{x}, \mathscr{R})\}$ ($i = 1$ or 2) is always a submartingale, since $\mu_A \geq 0$ and $\mu_B \geq 0$. It will be convenient to have a sufficient condition for when the submartingale property is preserved by optional stopping.

LEMMA 5.1. *Suppose $S_n = \sum_1^n x_j$ is a submartingale relative to a filtration $\{\mathscr{F}_n\}_{n \geq 0}$ and suppose t_1, t_2 are stopping times. If*

$$E\sum_1^{t_2} E\left[x_j^2 | \mathscr{F}_{j-1}\right] < \infty, \tag{5.7}$$

then

$$E\left(S_{t_2} | \mathscr{F}_{t_1 \wedge t_2}\right) \geq S_{t_1 \wedge t_2} \quad \text{w.p.1.} \tag{5.8}$$

PROOF. This is a simple consequence of the Doob decomposition of S_n into the sum of a martingale and an increasing process, and the Wald identity. □

LEMMA 5.2. *There is a constant $C < \infty$ such that for all $\mathbf{x} \in \mathbb{R}^2 \setminus \mathscr{Q}_+$,*

$$E| \mathbf{W}_{T(\mathbf{x})}(\mathbf{x})| 1\left\{\tau_{A_0}(\mathbf{x}) < T(\mathbf{x})\right\} \leq C. \tag{5.9}$$

PROOF. Straight-line switching is a diagonal-stabilizing policy, and so Proposition 4.1 applies. Let

$$\begin{aligned} &\mathbf{y} \in A_0 \setminus \mathscr{Q}_+, \\ T^*(\mathbf{y}) &= \min\left\{n \geq 0 \colon \mathbf{W}_n(\mathbf{y}) \in A_0 \text{ and } W_n^{(1)}(\mathbf{y}) + W_n^{(2)}(\mathbf{y}) > r\right\} \\ &\geq T(\mathbf{y}). \end{aligned}$$

Proposition 4.1 implies that the conditions of Proposition 2.1 and 2.2 are satisfied, so

$$ET^*(\mathbf{y}) < \infty$$

and

$$E\left[W_{T^*(\mathbf{y})}^{(1)}(\mathbf{y}) + W_{T^*(\mathbf{y})}^{(2)}(\mathbf{y})\right] \leq C'$$

for all $\mathbf{y} \in A \setminus \mathscr{Q}_+$. Here C' is the constant provided by Proposition 2.2.

Now by the submartingale property of $\{W_n^{(i)}(\mathbf{y})\}$, $i = 1$ or 2, and Lemma 2.1, it follows that

$$E\left[W_{T(\mathbf{y})}^{(1)}(\mathbf{y}) + W_{T(\mathbf{y})}^{(2)}(\mathbf{y})\right] \leq C'.$$

The inequality (5.9) follows easily from this and the Markov property of $\{\mathbf{W}(\mathbf{x})\}$ [condition on $\mathbf{W}_{\tau_{A_0}(\mathbf{x})}(\mathbf{x}) = \mathbf{y}$]. □

LEMMA 5.3. *There is a constant $C < \infty$ such that*

$$E| \mathbf{W}_{T(\mathbf{x})}(\mathbf{x})| 1\left\{\tau_{A_-}(\mathbf{x}) < T(\mathbf{x}) \leq \tau_{A_0}(\mathbf{x})\right\} \leq C \tag{5.10}$$

for all $x \in A_+ \setminus \mathscr{Q}_+$, *and*

$$E|\mathbf{W}_{T(\mathbf{x})}(\mathbf{x})|\,1\{\tau_{A_+}(\mathbf{x}) < T(\mathbf{x}) \le \tau_{A_0}(\mathbf{x})\} \le C \tag{5.11}$$

for all $x \in A_- \setminus \mathscr{Q}_+$

PROOF. First notice that for each $\mathbf{y} \in \mathbb{R}^2$,

$$\begin{aligned} &EW^{(i)}_{T(\mathbf{y})}(\mathbf{y})1\{T(\mathbf{y}) \le \tau_{A_0}(\mathbf{y})\} \\ &\quad\le E\Big[W^{(i)}_{\tau_{A_0}(\mathbf{y})}(\mathbf{y})\Big]_+ 1\{T(\mathbf{y}) \le \tau_{A_0}(\mathbf{y})\}, \qquad i = 1,2. \end{aligned} \tag{5.12}$$

This is because $\{W^{(i)}_n(\mathbf{y})\}_{n \ge 0}$ is a submartingale with

$$E\Big[\big(W^{(i)}_{n+1}(\mathbf{y}) - W^{(i)}_n(\mathbf{y})\big)^2 \big| \mathscr{A}_n\Big] \le \int_{\mathbb{R}^2} |\mathbf{v}|^2 (F_A + F_B)(d\mathbf{v}),$$

and $E\tau_{A_0}(\mathbf{y}) < \infty$ by Proposition 2.1 and 4.1. Thus Lemma 5.3 applies.

Next notice that there is a constant $C < \infty$ such that for every $\mathbf{y} \in \mathbb{R}^2 \setminus \mathscr{Q}_+$,

$$E\Big[W^{(i)}_{\tau_{A_0}(\mathbf{y})}(\mathbf{y})\Big]_+ \le C|\mathbf{r}(\mathbf{y})| = C \operatorname{dist}(\mathbf{y}, \text{diagonal}). \tag{5.13}$$

This is a relatively straightforward consequence of Corollary 2.1: Just make the identifications

$$\begin{aligned} X_{n+1} &\leftarrow |\mathbf{r}(\mathbf{W}_n(\mathbf{y})| - |\mathbf{r}(\mathbf{W}_{n+1}(\mathbf{y}))| \quad \text{if } |\mathbf{r}(\mathbf{W}_n(\mathbf{y}))| \ge r, \\ &\leftarrow 1 \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{otherwise} \\ \xi_{n+1} &\leftarrow \big(w^{(i)}_{n+1}(\mathbf{y})\big)_+, \end{aligned}$$

and

$$\nu \leftarrow [\![|r(y)| + 1]\!].$$

The hypotheses of Corollary 2.1 are satisfied because (i) straight-line switching is a diagonal-stabilizing policy, so $E[X_{n+1}|\mathscr{A}_n] \ge \mu_* > 0$ by Lemma 4.2, provided $r > 0$ is sufficiently large; and (ii) the increments $\mathbf{w}_n(\mathbf{y})$ are always chosen from the finite set $\{F_A, F_B\}$, so (2.19) and (2.20) hold.

When one of the random walks $\{\mathbf{W}_n(\mathbf{x})\}$ crosses from one side of the diagonal to the other, the jump cannot be too large. More precisely, there exists a constant $C^* < \infty$ such that for all $\mathbf{x} \in A_+$,

$$E\mathbf{r}\Big(\mathbf{W}_{\tau_{A_0 \cup A_{0-}}(\mathbf{x})}(\mathbf{x})\Big) \le C^*. \tag{5.14}$$

This is a straightward consequence of Proposition 2.1: Just make the identification

$$\begin{aligned} X_n &\leftarrow w^{(1)}_n(\mathbf{x}) - w^{(2)}_n(\mathbf{x}) \quad \text{if } n < \tau_{A_0 \cup A_-}(x) \\ &\leftarrow 1 \quad \text{otherwise.} \end{aligned}$$

Using the Markov property of $\{\mathbf{W}_n(\mathbf{x})\}$ and the inequalities (5.12)–(5.14) one may easily deduce (5.8). Obviously (5.9) follows by symmetry: Just reflect the entire problem through the diagonal. $\square$

LEMMA 5.4. *There exists a constant $C < \infty$ such that for every $\mathbf{x} \in A_+ \setminus \mathscr{Q}_+$,*

$$EW^{(1)}_{T(\mathbf{x})}(\mathbf{x})1\left\{T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\right\} \le C, \tag{5.15}$$

and for every $x \in A_- \setminus \mathscr{Q}_+$,

$$EW^{(2)}_{T(\mathbf{x})}(\mathbf{x})1\left\{T(\mathbf{x}) \le \tau_{A_0 \cup A_+}(\mathbf{x})\right\} \le C. \tag{5.16}$$

PROOF. This is an easy exercise in the use of Proposition 2.2. Make the identification

$$\begin{aligned} X_n &\leftarrow w^{(1)}_n(\mathbf{x}) \quad \text{if } n < T(\mathbf{x}) \wedge \tau_{A_0 \cup A_-}(\mathbf{x}), \\ &\leftarrow 1 \quad \text{otherwise.} \end{aligned}$$

Then

$$\begin{aligned} \mathscr{L}(X_n | \mathscr{F}_{n-1}) &= F_A \quad \text{if } n - 1 \le T(\mathbf{x}) \wedge \tau_{A_0 \cup A_-}(\mathbf{x}), \\ &= \delta_1 \quad \text{otherwise.} \end{aligned}$$

Since $\mu^{(1)}_A > 0$ and since F_A has finite second moment, the conditions of Proposition 2.2 are satisfied. On $\{T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\}$,

$$T(\mathbf{x}) = \min\left\{n\colon X_1 + \cdots + X_n > -x^{(1)}\right\}$$

so (5.15) follows from (2.7). The inequality (5.16) comes about the same way. □

LEMMA 5.5. *There exists a constant $C < \infty$ such that for all $\mathbf{x} \in A_+ \setminus \mathscr{Q}_+$,*

$$E\left(W^{(2)}_{T(\mathbf{x})}(\mathbf{x}) - x^{(2)}\right)1\left\{T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\right\} \le -Cx^{(1)}. \tag{5.17}$$

PROOF. Let

$$\begin{aligned} X_n &= w^{(1)}_n(\mathbf{x}) \quad \text{if } n < T(\mathbf{x}) \wedge \tau_{A_0 \cup A_-}(\mathbf{x}) \\ &= 1 \quad \text{otherwise,} \\ \xi_n &= w^{(2)}_n(\mathbf{x})_+. \end{aligned}$$

Clearly all of the hypotheses of Corollary 2.1 are satisfied, so by integrating (2.25) over $z \in (0, \infty)$, we obtain (5.17). □

PROOF OF THEOREM A. In accordance with Corollary 3.2, it is enough to show that there is a constant $C < \infty$ such that for each policy $\mathscr{R}$ and each $\mathbf{x} \in \mathbb{R}^2 \setminus \mathscr{Q}_+$,

$$E\left[W^{(1)}_{T(\mathbf{x})}(\mathbf{x}) + W^{(2)}_{T(\mathbf{x})}(\mathbf{x})\right] \le E\left[W^{(1)}_{T(\mathbf{x}, \mathscr{R})}(\mathbf{x}, \mathscr{R}) + W^{(2)}_{T(\mathbf{x}, \mathscr{R})}(\mathbf{x}, \mathscr{R})\right] + C.$$

In view of Lemmas 5.2–5.4, it suffices to show that there is a constant $C < \infty$ such that for each policy $\mathscr{R}$ and each $\mathbf{x} \in A_+ \setminus \mathscr{Q}_+$,

$$\begin{aligned} &EW^{(2)}_{T(\mathbf{x})}(\mathbf{x})1\left\{T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\right\} \\ &\quad \le EW^{(2)}_{T(\mathbf{x}, \mathscr{R})}(\mathbf{x}, \mathscr{R})1\left\{T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\right\} + C. \end{aligned} \tag{5.18}$$

[NOTE. $T(\mathbf{x}, \mathscr{R})$ is the first time $\mathbf{W}_n(\mathbf{x}, \mathscr{R})$ enters $\mathscr{Q}_+$ [cf. (5.4)]. However, $T(\mathbf{x})$ and $\tau_{A_0 \cup A_-}(\mathbf{x})$ are first passage times for $\mathbf{W}_n(\mathbf{x})$].

The proof rests on the observation that when $T(x) \le \tau_{A_0 \cup A_-}(\mathbf{x})$, the increments $\mathbf{w}_n(\mathbf{x})$ in the random walk $\mathbf{W}_n(\mathbf{x})$ are all drawn from F_A. The competing policy $\mathscr{R}$ may draw from either F_A or F_B; the resulting increment $\mathbf{w}_n(\mathbf{x}, \mathscr{R})$ will be different from $\mathbf{w}_n(\mathbf{x})$ only when $\mathscr{R}$ draws from F_B. Recall that

$$\mu_A^{(1)} > \mu_B^{(1)} \quad \text{and} \quad \mu_B^{(2)} > \mu_A^{(2)}.$$

This will be crucial since it allows us to use the (martingale version of the) Kiefer–Wolfowitz Theorem (cf. Proposition 2.3) on the components $\{W_n^{(1)}(\mathbf{x}, \mathscr{R}) - W_n^{(1)}(\mathbf{x})\}$ and $\{W_n^{(2)}(\mathbf{x}) - W_n^{(2)}(\mathbf{x}, \mathscr{R})\}$. More precisely, let

$$t(\mathbf{x}) = T(\mathbf{x}) \wedge T(\mathbf{x}, \mathscr{R}) \wedge \tau_{A_0 \cup A_-}(\mathbf{x}),$$

then Proposition 2.3 implies that

$$E \max_{n \le t(\mathbf{x})} \left(W_n^{(1)}(\mathbf{x}, \mathscr{R}) - W_n^{(1)}(\mathbf{x})\right) \le C \tag{5.19}$$

and

$$E \max_{n \le t(\mathbf{x})} \left(W_n^{(2)}(\mathbf{x}) - W_n^{(2)}(\mathbf{x}, \mathscr{R})\right) \le C. \tag{5.20}$$

Here $C < \infty$ is the constant provided by Proposition 2.3: It should be borne in mind that C depends only on F_A and F_B not on $\mathbf{x}$ or $\mathscr{R}$.

To prove (5.18) we partition the event $\{T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\}$ as follows:

$$B_1 = \{T(\mathbf{x}) = T(\mathbf{x}, \mathscr{R}) = t(\mathbf{x})\},$$
$$B_2 = \{T(\mathbf{x}) = t(\mathbf{x}) < T(\mathbf{x}, \mathscr{R})\},$$

and

$$B_3 = \left\{T(\mathbf{x}, \mathscr{R}) = t(\mathbf{x}) < T(\mathbf{x}) \le \tau_{A_0 \cup A_-}(\mathbf{x})\right\}.$$

We will show that for suitable constants $C_1, C_2, C_3 < \infty$,

$$E\left[W_{T(\mathbf{x})}^{(2)}(\mathbf{x}) - W_{T(\mathbf{x}, \mathscr{R})}^{(2)}(\mathbf{x}, \mathscr{R})\right] 1_{B_i} \le C_i \tag{5.21}$$

for $i = 1, 2, 3$. This will prove (5.18) and hence Theorem A.

The easiest of the three cases is $i = 1$. In this case (5.20) directly implies (5.21).

For the case $i = 2$ we must assume that $ET(\mathbf{x}, \mathscr{R}) < \infty$. This is no great loss, however, because if it is *not* satisfied then (1.2) holds trivially. Now recall that $\{W_n^{(2)}(\mathbf{x}, \mathscr{R})\}$ is a submartingale; thus by Lemma 5.1

$$EW_{T(\mathbf{x}, \mathscr{R})}^{(2)}(\mathbf{X}, \mathscr{R}) 1_{B_2} \ge EW_{T(\mathbf{x})}^{(2)}(\mathbf{x}, \mathscr{R}) 1_{B_2}. \tag{5.22}$$

But (5.20) implies that

$$E\left[W_{T(\mathbf{x})}^{(2)}(\mathbf{x}) - W_{T(\mathbf{x})}^{(2)}(\mathbf{x}, \mathscr{R})\right] 1_{B_2} \le C. \tag{5.23}$$

The case $i = 2$ of (5.21) follows from (5.22) and (5.23).

On the event B_3 we must use both (5.19) and (5.20). These imply that

$$EW_{T(\mathbf{x}, \mathscr{R})}^{(2)}(\mathbf{x}) 1_{B_3} \le EW_{T(\mathbf{x}, \mathscr{R})}^{(2)}(\mathbf{x}, \mathscr{R}) 1_{B_3} + C \tag{5.24}$$

and

$$E\left[-W^{(1)}_{T(\mathbf{x},\mathscr{R})}(\mathbf{x})\right]1_{B_3} \le C \tag{5.25}$$

[notice that on B_3 $W^{(1)}_{T(\mathbf{x},\mathscr{R})}(\mathbf{x},\mathscr{R}) > 0$]. Now by Lemma 5.5 and the Markov property,

$$\begin{aligned} E\left[W^{(2)}_{T(\mathbf{x})}(\mathbf{x}) - W^{(2)}_{T(\mathbf{x},\mathscr{R})}(\mathbf{x})\right]1_{B_3} &\le \tilde{C}E\left[-W^{(1)}_{T(\mathbf{x},\mathscr{R})}(\mathbf{x})\right]1_{B_3} \\ &\le \tilde{C}\cdot C, \end{aligned} \tag{5.26}$$

by (5.25); here $\tilde{C}$ is the constant provided by Lemma 5.5, and C is the constant in (5.25). Combining (5.24) and (5.26) yields the case $i = 3$ of (5.21) (with $C_3 = C + C\cdot\tilde{C}$).

This complete the proof of Theorem A. □

APPENDIX

PROOF OF LEMMA 2.3. The distributions G_j are easy to describe. Let m_1, m_2 and n be large integers, $m_1 \ll m_2 \ll n$, and let $\beta > 0$ be a real constant such that

$$2(h(z), h^*(z)) = 2^{-m_2}.$$

The distributions G_j will all be concentrated on

$$\begin{aligned} &\{(k_1 2^{-m_1}\beta, k_2 2^{-m_1}\beta) \colon 1 \le k_1, k_2 \le 2^{m_1}\} \cup \{(z,z)\colon z \in \mathbb{R} \text{ and } z \ge \beta\} \\ &\quad = \Delta_1 \cup \Delta_2. \end{aligned}$$

For $z_2 > z_1 \ge \beta$,

$$G_j\{(z,z)\colon z_2 \ge z \ge z_1\} = 2\cdot(h(z_1) + h^*(z_1) - h(z_2) - h^*(z_2))$$

and for any point $(k_1 2^{-m_1}\beta, k_2 2^{-m_1}\beta) \in \Delta_1$,

$$G_j(\{(k_1 2^{-m_1}\beta, k_2 2^{-m_1}\beta)\}) = \alpha(k_1,k_2)2^{-n}$$

where $\alpha = \alpha(k_1, k_2)$ is an integer, $0 \le \alpha < 2^n$. It is clear that for each triple (m_1, m_2, n) the set of G_j satisfying these specifications is finite.

The mapping $T(H,\cdot)\colon \mathbb{R}^2_+ \to \Delta_1 \cup \Delta_2$ is also relatively simple. The main features are as follows.

(A.1) For $(x, y) \in \mathbb{R}^2_+$ with $x \vee y \ge \beta$, $T(H,\cdot)$ first moves (x, y) to the diagonal point $(x\vee y, x \vee y)$ and then moves it up along the diagonal to a point $(z, z) = T(H(x, y))$ with $z \ge x \vee y$.

(A.2) For $(x, y) \in \mathbb{R}^2_+$ with $x \vee y < \beta$, $T(H,\cdot)$ either moves (x, y) to the nearest point $(k2^{-m_1}\beta, k_2 2^{-m_1}\beta) \ge (x, y)$, to (β,β), or to a point of Δ_2 to be specified later.

[The reason for not simply moving all points (x, y) with $x \vee y < \beta$ to the nearest upper corner is that the distributions G_j must assign dyadic weights $\alpha 2^{-n}$ to these points.]

Let

$$\begin{aligned} \hat{H}(z) &= H(\{(x, y): x \vee y > z\}), \qquad z \geq \beta, \\ \hat{\hat{H}}(z) &= 2(h(z) + h^*(z)), \qquad z \geq \beta, \\ S_H(z) &= \min\{z_1 \geq z: \hat{\hat{H}}(z_1) = \hat{H}(z)\}, \qquad \hat{H}(z) > 0 \\ &\quad +\infty \quad \text{if } \hat{H}(z) = 0. \end{aligned}$$

Notice that $\hat{H}(z) \leq (\frac{1}{2})\hat{\hat{H}}(z)$, according to the definition of the class $\mathscr{K}$. Now define

$$T(H,(x, y)) = (S_H(x \vee y), S_H(x \vee y)), x \vee y \geq \beta.$$

This is in agreement with (A.1); moreover it is clear that for any $H \in \mathscr{K}$, and G_j,

$$\begin{aligned} H\{(x, y): S_H(x \vee y) \geq z\} &= 2(h(z) + h^*(z)) \\ &= G_j\{(z_1, z_2): z_1 \geq z_2\} \quad \text{for } z \geq S_H(\beta). \end{aligned}$$

It is also clear that the component of the mapping $T(\cdot, \cdot)$ specified thus far is measurable.

The next step is to fill in the remainder of Δ_2: This consists of those points

$$\{(z, z): \beta < z \leq S_H(\beta)\} \triangleq \Delta_2^0(H).$$

The only points remaining to be moved by $T(H, \cdot)$ are those in the square $\{(x, y): x \vee y < \beta\}$, which we subdivide into the squares

$$K(k_1, k_2) = \{(x, y): k_1 - 1 \leq 2^{m_1}\beta^{-1}x < k_1 \text{ and } k_2 - 1 \leq 2^{m_1}\beta^{-1}y < k_2\}$$

and columns

$$J(k_1) = \bigcup_{k_2=1}^{2^{m_1}} K(k_1; k_2).$$

Our plan is to move through the columns $J(k_1)$ from left to right until we have accumulated just enough H mass to fill in $\Delta_2^0(H)$: This is indicated schematically in Figure 2. Thus let

$$\begin{aligned} L(k, y) &= H\left(\bigcup_{k_1=1}^{k-1} J(k_1)\right) + H\{(x, y') \in J(k); y' \leq y\}, \\ \hat{k}(H) &= \max\{k \geq 1: L(k-1, \beta) < 2^{-m_2} - \hat{H}(\beta)\}, \\ \hat{y}(H) &= \max\{y \leq \beta: L(\hat{k}(H), y) < 2^{-m_2} - \hat{H}(\beta)\} \\ &= \min\{y \leq \beta: L(\hat{k}(H), y) \geq 2^{-m_2} - \hat{H}(\beta)\}. \end{aligned}$$

The set $(\bigcup_{k_1=1}^{\hat{k}(H)-1} J(k_1)) \cup \{(x, y) \in J(\hat{k}(H)): y \leq \hat{y}(H)\}$ may now be moved by the one-dimensional quantile transformation to fill in the segment $\Delta_2^0(H)$. For $(x, y) \in J(k)$, $k < \hat{k}(H)$, and also $(x, y) \in J(\hat{k}(H))$ with $y \leq \hat{y}(H)$, let

$$S_L((x, y)) = \min\{z > \beta: L(k, y) = 2^{-m_2} - \hat{\hat{H}}(z)\}$$

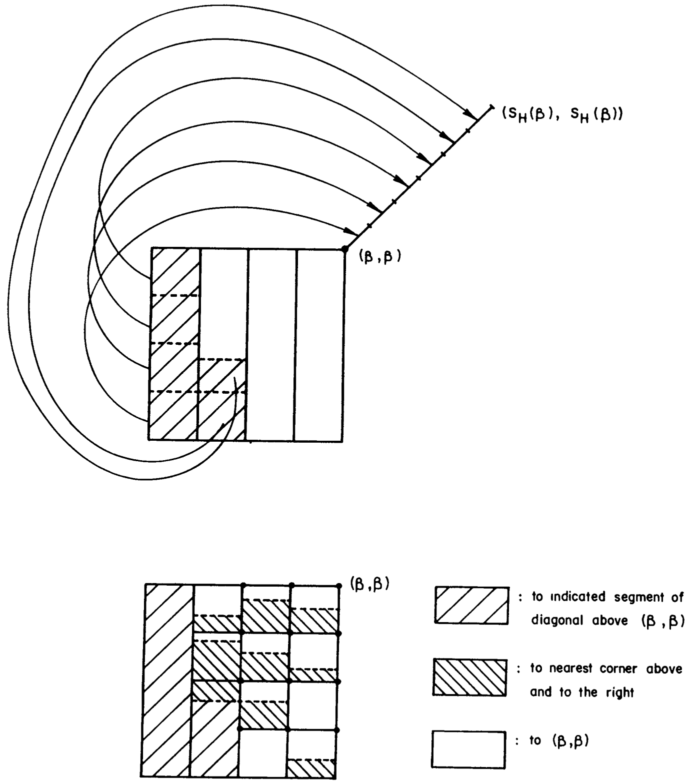


FIG. 2.

and let

$$T(H,(x,y)) = \left(S_L((x,y)), S_L((x,y))\right).$$

Notice that by construction, $\forall\ \beta \le z_1 < z_2 \le S_H(\beta)$,

$$\begin{aligned} H(\{(x,y)\colon z_1 < S_L((x,y)) \le z_2\}) &= \hat{H}(z_1) - \hat{H}(z_2) \\ &= G_j\{(z,z)\colon z_1 < z \le z_2\} \end{aligned}$$

for all of the G_j in our collection. It is once again clear that the component of $T(\cdot,\cdot)$ we have just defined is measurable.

To complete the definition of $T(H,\cdot)$ we need to specify its action on those points (x, y) such that

(A.3) $\quad 2^{m_1}\beta^{-1}x \ge \hat{k}(H) \quad$ or $\quad \hat{k}(H) - 1 \le 2^{m_1}\beta^{-1}x < \hat{k}(H) \quad$ and $\quad y \ge \hat{y}(H).$

The plan is to move most of these points to the upper right-hand corners of their respective squares $K(k_1, k_2)$. However, because the mass assigned to any of these corners must be an integer multiple of 2^{-n}, there will in general be "leftover" points (x, y) in some of the squares $K(k_1, k_2)$: These points will all be moved to

(β, β). It should be noted that when this plan is carried out, the total H mass of those points assigned to (β, β) will be an integer multiple of 2^{-n}. This is because the total H mass of points (x, y) remaining for this third phase of the operation [i.e., those satisfying (A.3)] is exactly $1 - 2^{-m_2}$.

Within any square $K(k_1, k_2)$ containing points satisfying (A.3) there should be a number $y(k_1, k_2)$ such that

$$T(H,(x, y)) = (k_1\beta 2^{-m_1}, k_2\beta 2^{-m_1}) \quad \text{for } (x, y) \in K(k_1, k_2)$$

satisfying (A.3) and with $y < y(k_1, k_2)$

$$= (\beta, \beta) \quad \text{for } (x, y) \in K(k_1, k_2)$$

satisfying (A.3) and with $y \geq y(k_1, k_2)$. The number $y(k_1, k_2)$ is uniquely determined by the requirement that

$$H(\{(x, y) \in K(k_1, k_2) \text{ satisfying (A.3) with } y < y(k_1, k_2)\})$$

is an integer multiple of 2^{-n}; and

$$H(\{(x, y) \in K(k_1, k_2) \text{ satisfying (A.3) with } y \geq y(k_1, k_2)\}) < 2^{-n}.$$

[Such a $y(k_1, k_2)$ always exists, since $H \in \mathscr{K}$ is absolutely continuous.] Notice that with this requirement in force, for any square $K(k_1, k_2)$ the H mass of those points satisfying (A.3) which are *not* moved to the upper corner $(k_1\beta 2^{-m_1}, k_2\beta 2^{-m_1})$ is less than 2^{-n}.

This completes the definition of the mapping $T(\cdot, \cdot)$. It is apparent from the construction that T is measurable, and that (2.32), (2.33), and (2.35) are satisfied. To assure that (2.34) holds, we choose m_1, m_2, and n very large:

(i) If m_2 is chosen very large, then β is large, and $\int_{z \geq \beta} 2 \cdot (h(z) + h^*(z))\, dz \ll \varepsilon$. Consequently, the contributions to the integrals in (2.34) from those points $(z, z) \in \Delta_2$, and those (x, y) such that $T(H,(x, y) \in \Delta_2$, is negligible compared to ε.

(ii) If m_1 is chosen very large, so that $2^{-m_1} \ll \varepsilon$, then those points in the squares $K(k_1, k_2)$ which $T(H, \cdot)$ moves to the upper corner $(k_1\beta 2^{-m_1}, k_2\beta 2^{-m_1})$ have a negligible $(\leq 2^{-m_1})$ effect on the *difference* of the two integrals in (2.34).

(iii) If n is chosen sufficiently large, then no more an H mass of points than $2^{2m_1} \cdot 2^{-n} \ll \varepsilon/\beta$ will be moved to (β, β). Thus the contribution to the integral on RHS (2.34) from those (x, y) which $T(H, \cdot)$ maps to (β, β) will be no larger than $\beta \cdot 2^{2m_1} \cdot 2^{-n} \ll \varepsilon$. Also the contribution to the integral on LHS (2.34) from the images of these points will be negligible.

The mapping $J\colon \mathscr{K} \to \{1, 2, \ldots, l\}$ has not yet been mentioned. It is, however, obviously determined by T. □

REFERENCES

BLOT W. and MEETER, D. (1973). Sequential experimental design procedures, *J. Amer. Statist. Assoc.* **68** 586–593.

BOX, G. and HILL, W. (1967). Discrimination among mechanistic models. *Technometrics* **9** 57–71.

CHERNOFF, H. (1959). Sequential design of experiments. *Ann. Math. Statist.* **30** 755–770.
CHOW, Y. S., ROBBINS, H. and SIEGMUND D. (1977). *Great Expectations: The Theory of Optimal Stopping.* Houghton-Mifflin, New York.
DERMAN, D. (1970). *Finite State Markovian Decision Processes.* Academic Press, New York.
DYNKIN, E. B. and YUSHKEVICH, A. (1979). *Controlled Markov Processes.* Springer, New York.
FELLER, W. (1966). *An Introduction to Probability Theory and its Applications* 2. Wiley, New York.
HALE, J. (1969). *Ordinary Differential Equations.* Wiley, New York.
KAMAE, T., KRENGEL, U. and O'BRIEN, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5** 899–912.
KEENER, R. (1980). Renewal theory and the sequential design of experiments with two states of nature. *Comm. Statist. A—Theory Methods* **9** (16), 1699–1726.
KEENER, R. (1981). A control problem. *Univ. of Michigan Tech. Report.*
KIEFER, J. and WOLFOWITZ, J. (1956). On the characteristics of the general queueing process, with applications to random walk. *Ann. Math. Statist.* **27** 147–161.
STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

DEPARTMENT OF STATISTICS
618 MATHEMATICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027

DEPARTMENT OF MATHEMATICS
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA