MATHEMATICS OF OPERATIONS RESEARCH
Vol. 18, No. 2, May 1993
Printed in U.S.A.

HIT-AND-RUN ALGORITHMS FOR GENERATING MULTIVARIATE DISTRIBUTIONS

CLAUDE J. P. BÉLISLE, H. EDWIN ROMEIJN AND ROBERT L. SMITH

We introduce a general class of Hit-and-Run algorithms for generating essentially arbitrary absolutely continuous distributions on $\mathbf{R}^d$. They include the Hypersphere Directions algorithm and the Coordinate Directions algorithm that have been proposed for identifying nonredundant linear constraints and for generating uniform distributions over subsets of $\mathbf{R}^d$. Given a bounded open set S in $\mathbf{R}^d$, an absolutely continuous probability distribution π on S (the target distribution) and an arbitrary probability distribution ν on the boundary of the d-dimensional unit sphere centered at the origin (the direction distribution), the (ν, π)-Hit-and-Run algorithm produces a sequence of iteration points as follows. Given the nth iteration point x, choose a direction θ according to the distribution ν and then choose the $(n+1)$st iteration point according to the conditionalization of the distribution π along the line defined by x and $x + \theta$. Under mild conditions, we show that this sequence of points is a Harris recurrent reversible Markov chain converging in total variation to the target distribution π.

1. Introduction. Let S be a bounded open subset of $\mathbf{R}^d$ and let π be an absolutely continuous probability measure on S. Let $f(x)$ be a probability density function for π and assume that it is bounded, almost everywhere continuous (with respect to Lebesgue measure on S), and strictly positive. Let D denote the d-dimensional unit sphere centered at the origin and let ∂D denote its topological boundary. Thus $D = \{x \in \mathbf{R}^d\colon \|x\| < 1\}$ and $\partial D = \{x \in \mathbf{R}^d\colon \|x\| = 1\}$. Finally, let ν be an arbitrary probability measure on ∂D. The *Hit-and-Run algorithm* with *direction distribution* ν and with *target distribution* π (in short, the (ν, π)-Hit-and-Run algorithm on S) can be described as follows:

Step 0. Choose a starting point $x_0 \in S$ and set $k = 0$.

Step 1. Choose a direction θ_{k+1} on ∂D, with distribution ν.

Step 2. Choose $\lambda_k \in \Lambda_k = \{\lambda \in \mathbf{R}\colon x_k + \lambda\theta_{k+1} \in S\}$, from the distribution with density

$$f_k(\lambda) = \frac{f(x_k + \lambda\theta_{k+1})}{\int_{\Lambda_k} dr f(x_k + r\theta_{k+1})}, \qquad \lambda \in \Lambda_k.$$

Step 3. Set $x_{k+1} = x_k + \lambda_k\theta_{k+1}$ and set $k = k + 1$.

Step 4. Go to 1.

The function $f_k(\lambda)$ introduced in Step 2 is the (reparametrized) density function of the conditionalization of the distribution π along the line Λ_k. Geometrically, f_k is a

Received June 1, 1990; revised June 5, 1991.
AMS 1980 subject classification. Primary: 60J20, 62E20; Secondary: 62E25.
IAOR 1973 subject classification. Main: Simulation. Cross references: Markov processes.
OR/MS Index 1978 subject classification. Primary: 571 Probability/regenerative processes. Secondary: 761 Simulation.
Key words. Hit-and-run algorithms, Monte-Carlo methods, Markov chains, reversibility, Harris recurrence.

0364-765X/93/1802/0255/$01.25

cross-section of the multivariate density function f, normalized to be a density function.

If ν is the uniform distribution on ∂D and if π is the uniform distribution on S the algorithm is known as the Hypersphere Directions algorithm (HD). This special case was first suggested by Boneh and Golan (1979) in the context of nonredundant constraint identification and later independently by Smith (1980) for generating points uniformly distributed over S. Later Smith (1984) showed that, for S open and bounded, the sequence of iteration points of the Hypersphere Directions algorithm converges to the uniform distribution on S. If ν is the discrete distribution that assigns mass $1/2d$ to each of the $2d$ coordinate directions and if π is the uniform distribution on S, the algorithm is known as the Coordinate Directions algorithm (CD). This special case was suggested by Telgen (1980). Later Berbee et al. (1987) showed that if S is a convex polyhedron then the sequence of iteration points of the Coordinate Directions algorithm converges to the uniform distribution on S. Recently Boender et al. (1991) have introduced a related class of algorithms, known as Shake-and-Bake algorithms, for generating points that are (asymptotically) uniformly distributed on the boundary ∂S of a full-dimensional convex polyhedron S. These algorithms are similar to the Hit-and-Run algorithms in the sense that each iteration involves the choice of a new direction θ on ∂D followed by the choice of a new iteration point along the line defined by the current iteration point and the new direction θ.

Clearly the (ν, π)-Hit-and-Run algorithm defines a discrete time Markov chain on S with stationary transition probabilities. The main purpose of this paper is to investigate the convergence properties of this Markov chain and to generalize the results of Smith (1984) and Berbee et al. (1987) in the following three directions: (a) the region S will be an arbitrary bounded open subset of $\mathbf{R}^d$, (b) the direction distribution ν will be completely arbitrary, (c) the target distribution π will be an arbitrary absolutely continuous distribution on S with a bounded, almost everywhere continuous, and strictly positive density.

2. Statement of the Main Results. Let $(X_n;\ n = 0, 1, 2, \ldots)$ denote the (ν, π)-Hit-and-Run Markov chain and let $P = (P(x, A);\ x \in S,\ A \in \mathscr{B}_S)$ denote its one-step transition probabilities. Thus P is a Markov kernel, that is (i) $\forall x \in S$, $P(x, \cdot\,)$ is a probability measure on $\mathscr{B}_S$, and (ii) $\forall A \in \mathscr{B}_S$, $P(\cdot\,, A)$ is a measurable function on S. See, e.g., Nummelin (1984, §1), Orey (1971, chapter 1, §0), Revuz (1975, chapter 1 §1). This Markov kernel will be called the *(ν, π)-Hit-and-Run Markov kernel* with direction distribution ν and target distribution π.

The first main result of this paper is that P is reversible with respect to π, i.e.,

$$(1)\qquad \int_A \pi(dx) P(x, B) = \int_B \pi(dy) P(y, A) \qquad \forall A, B \in \mathscr{B}_S.$$

(For a general discussion of reversibility, see, e.g., Kelly (1979) or Serfozo (1990).) If the direction distribution ν is absolutely continuous with respect to the uniform distribution on ∂D, then the probability measures $P(x, \cdot\,)$ are absolutely continuous with respect to the Lebesgue measure on S and our proof of reversibility is an elementary direct verification of (1). If ν is not absolutely continuous with respect to the uniform distribution on ∂D, then the probability measures $P(x, \cdot\,)$ are singular and proving reversibility is a more delicate task. Our proof involves expressing the singular Markov kernel as a weak limit of reversible absolutely continuous Markov kernels and then showing that reversibility is preserved in the limit.

After reversibility is established, stationarity of π follows at once since (1), with $B = S$, yields

$$\pi(A) = \int_S \pi(dy) P(y, A) \qquad \forall A \in \mathscr{B}_S.$$

However, the stationary distribution π is not necessarily unique. In order to have uniqueness, some restrictions have to be imposed. Recall that the *support* of a probability measure μ on $\mathbf{R}^d$, to be denoted $\mathrm{Supp}(\mu)$, is defined as $\mathrm{Supp}(\mu) = \{x \in \mathbf{R}^d: \mu(B(x, \epsilon)) > 0, \forall \epsilon > 0\}$, where $B(x, \epsilon)$ is the open ball of radius ϵ centered at x. See, e.g., Chung (1974, p. 31). Recall also that every open subset of $\mathbf{R}^d$ is a union of at most countably many disjoint open connected sets called the *connected components*. See, e.g., Franz (1965, §2). Finally, let P^n denote the nth iterate of P.

DEFINITION 1. The direction distribution ν is said to be full dimensional if $\mathrm{Supp}(\nu)$ contains a set of vectors that span $\mathbf{R}^d$ (in other words if the set $\{0\} \cup \mathrm{Supp}(\nu)$ is not contained in any linear subspace of dimension $d - 1$).

DEFINITION 2. The connected components of S are said to be ν-communicating if for every pair of connected components A and B there exists an $x \in A$ and an $n \geqslant 1$ such that $P^n(x, B) > 0$.

For instance, if $d = 2$, $S = B((0,0),1) \cup B((10,10),1)$ and $\mathrm{Supp}(\nu) = \{(1,0),(0,1)\}$ then the connected components of S are not ν-communicating. If $d = 2$, $S = B((0,0),1) \cup B((10,1),1) \cup B((10,10),1)$ and $\mathrm{Supp}(\nu) = \{(1,0),(0,1)\}$ then the connected components of S are ν-communicating. In §3 we show that the stationary distribution π is unique if and only if the direction distribution ν is full dimensional and the connected components of S are ν-communicating. In particular, the uniform distribution is the unique stationary distribution for CD on a connected bounded open set or for HD on an arbitrary bounded open set.

Now let us turn our attention to the asymptotic behaviour of the (ν, π)-Hit-and-Run Markov chain. Throughout this paper, φ will denote the Lebesgue measure on S. The second main result of this paper says that $(X_n;\ n = 0, 1, 2, \ldots)$ is *Harris recurrent* with respect to φ if and only if, again, ν is full dimensional and the connected components of S are ν-communicating. Harris recurrence with respect to φ, also known as φ-recurrence, means that

$$\mathbf{P}_x[X_n \in B \text{ for some } n \geqslant 1] = 1 \qquad \forall x \in S \tag{2}$$

whenever $\varphi(B) > 0$. See, e.g., Orey (1971), Revuz (1975). Here $\mathbf{P}_x$ denotes conditional probability given that $X_0 = x$. If the direction distribution ν has a density, with respect to the uniform distribution on ∂D, which is bounded away from 0 and if the density f of the target distribution π is also bounded away from 0, then the probability measures $P(x, \cdot\,)$ have densities, with respect to Lebesgue measure on S, which are bounded away from 0, uniformly in x. In this case a standard geometric trial argument yields Harris recurrence. The general case is more complicated. Our proof proceeds as follows. First we prove φ-irreducibility (i.e., we show that (2) holds with "$= 1$" replaced by "> 0") and then we deduce Harris recurrence using the ergodic theorem for stationary sequences (Breiman 1968, Theorem 6.28).

Once Harris recurrence is established, we verify aperiodicity and we obtain convergence of the (ν, π)-Hit-and-Run Markov chain to the stationary distribution π via a theorem of Orey (1971) which says that if an aperiodic Harris recurrent Markov chain possesses a stationary probability distribution π, then for every initial distribution μ, the distribution of X_n converges in total variation to the stationary distribution π.

Our main results can then be summarized as follows:

THEOREM 1. (a) *The* (ν, π)*-Hit-and-Run Markov kernel* P *is reversible with respect to* π *and therefore* π *is a stationary distribution for it.*

(b) *If* ν *is full dimensional and if the connected components of* S *are* ν*-communicating, then* P *is Harris recurrent with respect to Lebesgue measure on* S *and for every initial distribution the* (ν, π)*-Hit-and-Run Markov chain converges in total variation to* π.

Recall that if $\pi, \pi_1, \pi_2, \pi_3, \ldots$ are probability measures on S then, by definition, π_n is said to converge in total variation to π if

$$\lim_{n \to \infty} \pi_n(B) = \pi(B) \qquad \forall B \in \mathscr{B}_S,$$

uniformly in B. See, e.g., Dudley (1989). (In comparison, π_n is said to converge in distribution to π, or converge weakly to π, if

$$\lim_{n \to \infty} \pi_n(B) = \pi(B) \qquad \forall B \in \mathscr{B}_S \text{ with } \pi(\partial B) = 0,$$

where ∂B denotes the topological boundary of B; see Billingsley (1968, Theorem 2.1).) Thus Theorem 1 asserts that if $(X_n;\ n \geqslant 0)$ is the (ν, π)-Hit-and-Run Markov chain on S and if ν is full dimensional and the connected components of S are ν-communicating, then for every initial distribution μ we have

$$\lim_{n \to \infty} \mathbf{P}_\mu[X_n \in B] = \pi(B) \qquad \forall B \in \mathscr{B}_S,$$

uniformly in B. Here $\mathbf{P}_\mu[\cdot] = \int_S \mu(dx) \mathbf{P}_x[\cdot]$. This result has significant practical implications for Monte-Carlo methods. It says that for any absolutely continuous distribution π possessing a bounded, strictly positive and almost everywhere continuous density, one can use the Hit-and-Run algorithm to generate points that are asymptotically π-distributed.

3. Proof of the Main Results.

3.1. *Reversibility of* P *and stationarity of* π. In this section we prove part (a) of Theorem 1. Let Θ and U be random variables, defined on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$, such that Θ has distribution ν, such that U is uniformly distributed over the interval $(0, 1)$, and such that Θ and U are independent. For $x, y \in S$, with $x \neq y$, let

$$\Lambda(x, y) = \left\{ \lambda \in \mathbf{R}\colon x + \lambda \frac{y - x}{\|y - x\|} \in S \right\}$$

and let $f_{(x,y)}$ be the probability density function on $\mathbf{R}$ defined by

$$f_{(x,y)}(\lambda) = \frac{f(x + \lambda(y - x)/\|y - x\|)}{\int_{\Lambda(x,y)} dr f(x + r(y - x)/\|y - x\|)} \quad \text{if } \lambda \in \Lambda(x, y)$$

and $f_{(x,y)}(\lambda) = 0$ if $\lambda \notin \Lambda(x, y)$. Let $F_{(x,y)}$ denote the cumulative distribution function of $f_{(x,y)}$. Following the description given in §1, the (ν, π)-Hit-and-Run

Markov kernel can be written as

$$(3)\qquad P(x,A) = \mathbf{P}\left[x + \left(F^{-1}_{(x,x+\Theta)}(U)\right)\Theta \in A\right]$$

$$= \int_{\partial D} \nu(d\theta)\mathbf{P}\left[x + \left(F^{-1}_{(x,x+\theta)}(U)\right)\theta \in A\right]$$

$$= \int_{\partial D} \nu(d\theta) \int_{\Lambda(x,x+\theta)} d\lambda\, 1_A(x+\lambda\theta) \frac{f(x+\lambda\theta)}{\int_{\Lambda(x,x+\theta)} dr f(x+r\theta)}$$

for all $x \in S$ and all $A \in \mathscr{B}_S$. Here 1_A is the indicator function of the set A and $F^{-1}_{(x,x+\theta)}$ is the usual left-continuous inverse of $F_{(x,x+\theta)}$, i.e., for $0 < u < 1$, $F^{-1}_{(x,x+\theta)}(u) = \inf\{v \in \mathbf{R}\colon F_{(x,x+\theta)}(v) > u\}$. First, consider the case where the direction distribution ν is absolutely continuous with respect to the uniform distribution ν_0 on ∂D, i.e., the case where ν is of the form $\nu(G) = \int_G \nu_0(d\theta)h(\theta)$, $\forall G \in \mathscr{B}_{\partial D}$. In this case (3) becomes $P(x,A) = \int_A \pi(dx) g(x,y)$ with

$$g(x,y) = \frac{h((y-x)/\|y-x\|) + h((x-y)/\|x-y\|)}{\|y-x\|^{d-1} C_d \int_{\Lambda(x,y)} dr f(x + r(y-x)/\|y-x\|)}$$

and with $C_d = 2\pi^{d/2}/\Gamma(d/2)$ (i.e., C_d is the surface area of ∂D). The left-hand side and right-hand side of (1) then become $\int_A \pi(dx)\int_B \pi(dy) g(x,y)$ and $\int_B \pi(dy) \int_A \pi(dx) g(y,x)$, respectively. Reversibility then follows from Fubini's theorem and from the fact that $g(x,y)$ is symmetric in x and y. Now consider the case where ν is not absolutely continuous with respect to the uniform distribution on ∂D. Let $\nu_1, \nu_2, \nu_3, \dots$ be a sequence of absolutely continuous direction distribution on ∂D such that ν_n converges weakly to ν. (For a definition of weak convergence, see, e.g., Billingsley (1968); the fact that such a sequence exists can be seen as follows. Let Θ be a ∂D-valued random variable with distribution ν, let U_n be an $\mathbf{R}^d$-valued random variable with uniform distribution over the open ball of radius $1/n$ centered at the origin, and assume that Θ and U_n are independent. Let ν_n be the distribution of $(\Theta + U_n)/\|\Theta + U_n\|$. Then ν_n has the desired properties.) The above argument shows that if P_n denotes the (ν_n, π)-Hit-and-Run Markov kernel, then

$$(4)\qquad \int_A \pi(dx) P_n(x,B) = \int_B \pi(dy) P_n(y,A) \qquad \forall A, B \in \mathscr{B}_S.$$

To complete the proof, it suffices to show that reversibility is preserved in the limit, i.e., it suffices to show that letting $n \to \infty$ in (4) will yield

$$(5)\qquad \int_A \pi(dx) P(x,B) = \int_B \pi(dy) P(y,A) \qquad \forall A, B \in \mathscr{B}_S.$$

This can be done as follows. Since ν_n converges weakly to ν, the measure $P_n(x,\cdot\,)$ converges weakly to the measure $P(x,\cdot\,)$, for each x. Consequently if R is a rectangle in S then we have $\lim_{n\to\infty} P_n(x,R) = P(x,R)$ for almost all $x \in S$, because $P(x,\partial R) = 0$ for almost all x in S. See, e.g., Billingsley (1968, Theorem 2.1). Thus, letting $n \to \infty$ in (4) and using the Lebesgue dominated convergence theorem (see, e.g., Billingsley (1986, Theorem 16.4), we obtain (5) whenever A and B are rectangles in S. A standard measure theoretic extension argument (e.g., Breiman 1968, Proposition 2.23 applied twice) shows that (5) must, therefore, hold for every A and B in $\mathscr{B}_S$. Thus P is reversible with respect to π.

As explained in §2, reversibility of P with respect to π implies that π is stationary for P. This completes the proof of part (a) of Theorem 1.

3.2. *Irreducibility of P and uniqueness of the stationary distribution.* We begin with some preliminary results. The nth iterate of the Markov kernel P, say P^n, has a unique decomposition

$$P^n(x,B) = P^n_{\text{sing}}(x,B) + P^n_{\text{abs}}(x,B)$$

$$= P^n_{\text{sing}}(x,B) + \int_B dy\, p_n(x,y)$$

where $P^n_{\text{sing}}(x,B)$ is a singular substochastic kernel, where $P^n_{\text{abs}}(x,B)$ is an absolutely continuous substochastic kernel and where $p_n(x,y)$ is jointly measurable. (see Orey 1971, §1.1). If $P^n_{\text{abs}}(x,S) > 0$ then we say that the probability measure $P^n(x,\cdot)$ has an absolutely continuous part. If the density $p_n(x,y)$ is positive for almost all y's in some Borel set B, then we say that the probability measure $P^n(x,\cdot)$ *spreads over* B.

PROPOSITION 1. *Suppose that ν is full dimensional. Then*

(a) *$\forall n \geqslant d$ and $\forall x \in S$, the probability measure $P^n(x,\cdot)$ has an absolutely continuous part.*

(b) $\forall x \in S$, $\lim_{n\to\infty} P^n_{\text{abs}}(x,S) = 1$.

(c) *There exists an $r > 0$, depending only on ν, such that for all $x \in S$ and for all $n \geqslant d$, the probability measure $P^n_{\text{abs}}(x,\cdot)$ has a density which is strictly positive on $B(x, r\gamma_x)$, where $\gamma_x = \inf_{y\in S^c}\|x-y\|$, and where S^c denotes the complement of S.*

(d) *If A is a connected component of S and if G is a measurable subset of A with positive Lebesgue measure, then for all $x \in A$, there exists an integer n such that* $P^n(x,G) > 0$.

(e) *If A is a connected component of S, if G is a measurable subset of A with positive Lebesgue measure and if μ is a measure on $(S, \mathscr{B}_S)$ such that $\mu(A) > 0$, then there exists an integer n such that*

$$\int_A \mu(dx) P^n(x,G) > 0.$$

PROOF. Let L_n denote the subset of $(\partial D)^n$ consisting of those points $(\theta_1, \theta_2, \ldots, \theta_n)$ for which there exist integers $1 \leqslant i_1 < i_2 < \cdots < i_d \leqslant n$ such that $\theta_{i_1}, \theta_{i_2}, \ldots, \theta_{i_d}$ are linearly independent in $\mathbf{R}^d$. Let $\Theta_1, \Theta_2, \Theta_3, \ldots$ be the successive directions taken by the (ν,π)-Hit-and-Run Markov chain. Since $\Theta_1, \Theta_2, \Theta_3, \ldots$ are independent random variables with distribution ν and since ν is full dimensional,

$$\mathbf{P}[(\Theta_1, \Theta_2, \ldots, \Theta_n) \in L_n] > 0 \qquad \forall n \geqslant d$$

and

$$\mathbf{P}[(\Theta_1, \Theta_2, \ldots, \Theta_n) \in L_n] \to 1 \qquad \text{as } n \to \infty.$$

One of the important features of the Hit-and-Run Markov chain is that if $(\theta_1, \theta_2, \ldots, \theta_n) \in L_n$ then for every x in S the conditional distribution

$$\mathbf{P}_x[X_n \in G | (\Theta_1, \ldots, \Theta_n) = (\theta_1, \ldots, \theta_n)]$$

is absolutely continuous with respect to Lebesgue measure on S. Thus parts (a) and (b) follow from the fact that

$$P^n(x,G) = \int_{(\partial D)^n} \nu(d\theta_1)\nu(d\theta_2)\cdots\nu(d\theta_n)\mathbf{P}_x[X_n \in G|(\Theta_1,\Theta_2,\dots,\Theta_n) = (\theta_1,\theta_2,\dots,\theta_n)].$$

Now consider part (c). Fix $n \geqslant 1$ and $x \in S$, and suppose that $v \in \operatorname{Supp}(P^n_{\text{abs}}(x,\cdot\,))$. Fix $\epsilon > 0$ small enough so that $B = B(v,\epsilon) \subset S$. Then

$$P^{n+1}(x,B) = \int_S P^n(x,dy)P(y,B) = \int_S P^n_{\text{sing}}(x,dy)P(y,B) + \int_S P^n_{\text{abs}}(x,dy)P(y,B).$$

Since B is an open ball, we have $P(y,B) > 0$ for every y in B and since $v \in \operatorname{Supp}(P^n_{\text{abs}}(x,\cdot\,))$, we obtain $P^n_{\text{abs}}(x,B) > 0$. Thus the second integral on the right-hand side of the last equality is strictly positive. This implies that $P^{n+1}_{\text{abs}}(x,B) > 0$. Since this holds for every $\epsilon > 0$, we conclude that $v \in \operatorname{Supp}(P^{n+1}_{\text{abs}}(x,\cdot\,))$. Thus

$$\text{(6)} \qquad \operatorname{Supp}(P^n_{\text{abs}}(x,\cdot)) \subset \operatorname{Supp}(P^{n+1}_{\text{abs}}(x,\cdot)) \qquad \forall x \in S, \forall n \geqslant 1.$$

Now consider linearly independent $\theta_1,\theta_2,\dots,\theta_d$ in $\operatorname{Supp}(\nu)$. By continuity, we can choose $r > 0$ and $\epsilon > 0$ small enough so that the sets $A_k = B(\theta_k,\epsilon) \cap \partial D$, $k = 1,2,\dots,d$ are linearly independent (in the sense that $u_1,u_2,\dots,u_d$ are linearly independent whenever $u_k \in B(\theta_k,\epsilon) \cap \partial D$, $k = 1,2,\dots,d$) and

$$B(0,r) \subset \left\{ \sum_{k=1}^{d} \alpha_k u_k \colon -1 < \alpha_k < 1,\, k = 1,2,\dots,d \right\}$$

for every $(u_1,u_2,\dots,u_d) \in A_1 \times A_2 \times \cdots \times A_d$. Then for every $(u_1,u_2,\dots,u_d) \in A_1 \times A_2 \times \cdots \times A_d$, the conditional distribution $\mathbf{P}_x[X_d \in G|(\Theta_1,\Theta_2,\dots,\Theta_d) = (u_1,u_2,\dots,u_d)]$ is absolutely continuous and it has a density which is strictly positive on $B(x,r\gamma_x)$. Since $\Theta_1,\Theta_2,\dots,\Theta_d$ are independent random vectors with common distribution ν and since $\theta_1,\theta_2,\dots,\theta_d$ are in the support of ν,

$$\mathbf{P}[(\Theta_1,\Theta_2,\dots,\Theta_d) \in A_1 \times A_2 \times \cdots \times A_d] = \prod_{i=1}^{d} \mathbf{P}[\Theta_i \in A_i] = \prod_{i=1}^{d} \nu(A_i) > 0.$$

Thus $P^d_{\text{abs}}(x,\cdot\,)$ has a density which is strictly positive on $B(x,r\gamma_x)$. Combined with (6), this proves part (c).

Finally, consider parts (d) and (e). Let A be a connected component of S. Fix $x \in A$. Let G be a measurable subset of A with positive Lebesgue measure. Choose $y \in A$ such that for every $\epsilon > 0$ the set $B(y,\epsilon) \cap G$ has positive Lebesgue measure. Let $g\colon [0,1] \to S$ be such that $g(0) = x$, $g(1) = y$, and g is continuous on $[0,1]$. Let

$$\gamma = \inf\{\|u - v\|;\, u \in g([0,1]),\, v \in A^c\}.$$

Observe that γ is strictly positive since it is the distance between the compact set $g([0,1])$ and the closed set A^c. Let r be as in part (c) and let $\epsilon_* = r\gamma$. Now consider open balls $B_i = B(x_i, \epsilon_*)$, $i = 1, 2, \dots, k$, where $x_1 = x$, $x_k = y$, and $x_i \in B_{i-1} \cap g([0,1])$ for $i = 2, 3, \dots, k$. Such a construction is possible since the set $g([0,1])$ is compact. It follows from part (c) that $P^{id}(x, \cdot)$ spreads over B_i. In particular, $P^{kd}(x, \cdot)$ spreads over B_k. This implies that $P^{kd}(x, G) > 0$. This proves part (d). The argument actually shows that $P^{kd}(x', G) > 0$, $\forall x' \in B_1$. Thus if in the proof of part (d) we take $x \in \mathrm{Supp}(\mu)$, then we obtain

$$\int_A \mu(dx') P^{kd}(x', G) \geqslant \int_{B_1} \mu(dx') P^{kd}(x', G) > 0.$$

This proves part (e). ∎

Using Theorems 5.2 and 5.5 of Billingsley (1968), one concludes that the Hit-and-Run Markov kernels and their iterates are continuous in the sense that if $x_k \to x$ then, for $n \geqslant 1$, the probability measure $P^n(x_k, \cdot)$ converges weakly, as $k \to \infty$, to the probability measure $P^n(x, \cdot)$. This implies that if B is an open set and if $x_k \to x$ as $k \to \infty$ then $\liminf_{k \to \infty} P^n(x_k, B) \geqslant P^n(x, B)$. (See, e.g., Billingsley 1968, Theorem 2.1.) A consequence of this observation is that the condition of Definition 2 (that there exists an $x \in A$ and an $n \geqslant 1$ such that $P^n(x, B) > 0$) is equivalent to the condition that for some $n \geqslant 1$

$$\varphi(\{x \in A \colon P^n(x, B) > 0\}) > 0,$$

where, as before, φ denotes Lebesgue measure on $\mathbf{R}^d$. The next result confirms the relevance of Definitions 1 and 2. Recall that a Markov kernel P is called *φ-irreducible* (or *irreducible* with respect to φ) if for every $x \in S$ and for every $B \in \mathscr{B}_S$ with $\varphi(B) > 0$ there exists an $n \geqslant 1$ such that $P^n(x, B) > 0$. In terms of the associated Markov chain $(X_n; n \geqslant 0)$, φ-irreducibility is equivalent to the statement that

$$\mathbf{P}_x[X_n \in B \text{ for some } n \geqslant 1] > 0 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$.

THEOREM 2. *The (ν, π)-Hit-and-Run Markov kernel P is φ-irreducible if and only if ν is full dimensional and the connected components of S are ν-communicating.*

PROOF. The fact that φ-irreducibility implies full dimensionality of ν and ν-communicability of the connected components of S follows almost immediately from the definitions. Now suppose that ν is full dimensional and the connected components of S are ν-communicating. Fix $x \in S$. Fix $G \in \mathscr{B}_S$ with $\varphi(G) > 0$. We need to show that

$$P^n(x, G) > 0 \qquad \text{for some } n \geqslant 1. \tag{7}$$

Let A be the open connected component of S to which x belongs. Let B be an open connected component of S such that $\varphi(G \cap B) > 0$. If $A = B$, then (7) follows from part (d) of Proposition 1. If $A \neq B$, then let $n_2 \geqslant 1$ and $y \in A$ be such that $P^{n_2}(y, B) > 0$. By the remark preceding the statement of Theorem 2, the set $C = \{z \in A \colon P^{n_2}(z, B) > 0\}$ has positive Lebesgue measure. By part (d) of Proposition 1, there exists an integer $n_1 \geqslant 1$ such that $P^{n_1}(x, C) > 0$. We then obtain

$$\begin{aligned} P^{n_1 + n_2}(x, B) &= \int_S P^{n_1}(x, dz) P^{n_2}(z, B) \\ &\geqslant \int_C P^{n_1}(x, dz) P^{n_2}(z, B) > 0 \end{aligned}$$

and therefore by part (e) of Proposition 1 there exists an integer $n_3 \geqslant 1$ such that

$$\int_B P^{n_1+n_2}(x, dz) P^{n_3}(z, G \cap B) > 0.$$

Now if we take $n = n_1 + n_2 + n_3$, then

$$P^n(x, G) \geqslant P^n(x, G \cap B)$$

$$\geqslant \int_B P^{n_1+n_2}(x, dz) P^{n_3}(z, G \cap B) > 0,$$

i.e., (7) holds. ■

Recall that a Markov kernel P is said to be *indecomposible* if there are no disjoint nonempty sets A and B such that $P(x, A) = 1\ \forall x \in A$ and $P(x, B) = 1\ \forall x \in B$ (see Breiman 1968). It is easy to verify that φ-irreducibility implies indecomposability. Thus if ν is full dimensional and if the connected components of S are ν-communicating, then the (ν, π)-Hit-and-Run Markov kernel is indecomposible. Now Theorem 7.16 of Breiman (1968) says that an indecomposible Markov chain possesses at most one stationary probability distribution. The result is stated for the case where the state space is a Borel subset of $\mathbf{R}$ but the proof that Breiman presents is also valid for the case where the state space is a Borel subset of $\mathbf{R}^d$. Thus Theorem 2, combined with the first part of Theorem 1, implies that if ν is full dimensional and if the connected components of S are ν-communicating then π is the unique stationary distribution for the (ν, π)-Hit-and-Run Markov kernel. It is easy to see that these two conditions are actually necessary for uniqueness of the stationary distribution. If the connected components of S do not ν-communicate then S is a union of two connected open sets, say A and B, that do not ν-communicate, and the probability measures defined by $\pi_A(G) = \pi(G \cap A)/\pi(A)$ and $\pi_B(G) = \pi(G \cap B)/\pi(B)$ are both stationary. If ν is not full dimensional then the linear subspace of $\mathbf{R}^d$ spanned by $\operatorname{Supp}(\nu)$, say H, is less than d dimensional and any conditionalization of the probability measure π to a set of the form $x + H$, with $x \in S$, will be stationary. Thus we have proved the following:

THEOREM 3. *The target distribution π is the unique stationary distribution for the (ν, π)-Hit-and-Run Markov kernel if and only if ν is full dimensional and the connected components of S are ν-communicating.* ■

3.3. *Harris recurrence and convergence.* Orey (1971) discusses periodicity for general φ-irreducible Markov kernels. The next proposition states that Hit-and-Run Markov kernels are aperiodic in Orey's sense.

PROPOSITION 2. *Suppose that ν is full dimensional and that the connected components of S are ν-communicating. Then the (ν, π)-Hit-and-Run Markov kernel is aperiodic.*

PROOF. By Theorem 2 we have φ-irreducibility. Thus Theorem 3.1 of Orey (1971) applies. By part (c) of Proposition 1, we cannot have a cycle of length $k > 1$ (see Definition 3.2 of Orey 1971). Thus the (ν, π)-Hit-and-Run Markov kernel is aperiodic. ■

Finally, we recall that the Markov kernel P is said to be *φ-recurrent* (or *Harris recurrent with respect to φ*) if the corresponding Markov chain $(X_n; n \geqslant 0)$ satisfies

$$\mathbf{P}_x[X_n \in B \text{ for some } n \geqslant 1] = 1 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$. This is equivalent to the statement that

$$\mathbf{P}_x\left[\sum_{n=1}^{\infty} 1_B(X_n) = \infty\right] = 1 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$. Note that φ-recurrence implies φ-irreducibility.

THEOREM 4. *The (ν, π)-Hit-and-Run Markov kernel P is Harris recurrent with respect to Lebesgue measure on S if and only if ν is full dimensional and the connected components of S are ν-communicating.*

PROOF. Suppose that P is Harris recurrent with respect to φ. Then P is φ-irreducible and Theorem 2 implies that ν is full dimensional and the connected components of S are ν-communicating. Conversely, suppose that ν is full dimensional and that the connected components of S are ν-communicating. By the first part of Theorem 1 π is invariant for P and by Theorem 2 and the remark following the proof of Theorem 2, P is indecomposible. Thus, under P_π, the Hit-and-Run Markov chain $(X_n; n \geqslant 0)$ is a stationary ergodic sequence (see Breiman 1968, Theorem 7.16, p. 136). Now fix $B \in \mathscr{B}_S$. Then under P_π the sequence $(1_B(X_n); n \geqslant 0)$ is also a stationary ergodic sequence (see Breiman 1968, Proposition 6.31, p. 119; here again Breiman states the result for the case where the state space is a Borel subset of $\mathbf{R}$ but his proof is also valid for the case where the state space is a Borel subset of $\mathbf{R}^d$). Thus by the ergodic theorem (see, e.g., Breiman 1968, Theorem 6.28, p. 118) we obtain

$$\int_S dx f(x) \mathbf{P}_x\left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B)\right] = 1.$$

Since $f(x) > 0\ \forall x \in S$, this implies that

$$\mathbf{P}_x\left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B)\right] = 1 \tag{8}$$

for almost all x in S. We will now show that (8) holds for all x in S. Let H denote the event $\{\lim_{n\to\infty}(1/n)\sum_{i=1}^{n} 1_B(X_i) = \pi(B)\}$. Fix x in S and fix $\epsilon > 0$. By part (b) of Proposition 1 we can choose k so that $P_{\text{abs}}^k(x, S) > 1 - \epsilon$. Now let

$$S_0 = \{y \in S: \mathbf{P}_y[H] = 1\}.$$

Then, using the fact that $\mathbf{P}_x[H | X_k = y] = \mathbf{P}_y[H]$, we obtain

$$\begin{aligned}
\mathbf{P}_x[H] &= \int_S P^k(x, dy) \mathbf{P}_x[H | X_k = y] \\
&= \int_S P^k(x, dy) \mathbf{P}_y[H] \\
&\geqslant \int_{S_0} P_{\text{abs}}^k(x, dy) \mathbf{P}_y[H] \\
&= P_{\text{abs}}^k(x, S_0) = P_{\text{abs}}^k(x, S) > 1 - \epsilon.
\end{aligned}$$

The last equality follows from the fact that $\varphi(S_0) = \varphi(S)$. This holds for every $\epsilon > 0$.

Thus (8) holds for all $x \in S$. Now if $B \in \mathscr{B}_S$ and $\varphi(B) > 0$, then $\pi(B) > 0$ and therefore (8) yields

$$\mathbf{P}_x\left[\sum_{i=1}^{\infty} 1_B(X_i) = \infty\right] = 1.$$

Thus the Markov chain is Harris recurrent with respect to Lebesgue measure on S. ■

As mentioned in §2, convergence in total variation to the stationary distribution now follows from Orey's theorem (Orey 1971, Theorem 7.1 p. 30). The proof of Theorem 1 is now complete.

4. Conclusion. During the last decade, the Coordinate Directions and the Hypersphere Directions Hit-and-Run algorithms have been used for generating uniform distributions over bounded open subsets of $\mathbf{R}^d$. In that context they are known to be efficient when compared to standard methods such as the acceptance-rejection method, especially when d is large. See, for example, Berbee et al. (1987). In this paper we have generalized these traditional Hit-and-Run algorithms to allow for an arbitrary direction distribution ν and an essentially arbitrary absolutely continuous target distribution π. Although not addressed here, the boundedness conditions on the open set S and on the density function f can both be removed. Moreover, the condition that f be strictly positive on S and the condition that S be open both can be substantially relaxed.

As a final comment, we mention that the general (ν, π)-Hit-and-Run algorithms introduced in this paper show considerable promise for solving certain global optimization problems. For example, by setting the density f to be a suitable increasing function of the objective function g, the target distribution π will concentrate near the global maximum of g and therefore the Hit-and-Run Markov chain will, with high probability, find itself near that global maximum. This idea is currently being investigated.

Acknowledgements. We want to thank the editors and the referee for suggestions that improved the exposition of this paper.

This work was supported in part by NATO Collaborative Research Grant No. 0119/89.

References

[1] Berbee, H. C. P., Boender, C. G. E., Rinnooy Kan, A. H. G., Scheffer, C. L., Smith, R. L. and Telgen, J. (1987). Hit-and-Run algorithms for the identification of nonredundant linear inequalities. *Math. Programming*, **37** 184–207.

[2] Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

[3] ______ (1986). *Probability and Measure*, Second Edition, Wiley, New York.

[4] Boender, C. G. E., Caron, R. J., Rinnooy Kan, A. H. G., McDonald, J. F., Romeijn, H. E., Smith, R. L., Telgen, J. and Vorst, A. C. F. (1991). Shake-and-Bake Algorithms for Generating Uniform Points on the Boundary of Bounded Polyhedra. *Oper. Res.*, **39** 935–953.

[5] Boneh, A., and Golan, A. (1979). *Constraints' Redundancy and Feasible Region Boundedness by Random Feasible Point Generator* (*RFPG*). Third European Congress on Operations Research, EURO III, Amsterdam.

[6] Breiman, L. (1968). *Probability*. Addison-Wesley, Reading, MA.

[7] Chung, K. L. (1974). *A Course in Probability Theory*, Second Edition. Academic Press, London.

[8] Dudley, R. M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Pacific Grove, CA.

[9] Franz, W. (1965). *General Topology*. Frederick Ungar Publishing Company, New York.

[10] Kelly, F. P. (1979). *Reversibility and Stochastic Networks*. Wiley, New York.

[11] Nummelin, E. (1984). *General Irreducible Markov Chains and Non-negative Operators*. Cambridge University Press, Cambridge.
[12] Orey, S. (1971). *Limit Theorems for Markov Chain Transition Probabilities*. Van Nostrand Reinhold, New York.
[13] Revuz, D. (1975), *Markov Chains*. North-Holland Publishing Company, Amsterdam.
[14] Serfozo, R. F. (1990). Reversibility and Compound Birth-Death and Migration Processes. Submitted for publication.
[15] Smith, R. L. (1980). Monte Carlo Procedures for Generating Random Feasible Solutions to Mathematical Programs. ORSA/TIMS Conference, Washington, D.C.
[16] ______ (1984). Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper. Res.*, **32** 1296–1308.
[17] Telgen, J. (1980). Private communication with A. Boneh.

C. J. P. Belisle: Department of Statistics, The University of Michigan, Ann Arbor, Michigan 48109-1027

H. E. Romeijn: Department of Operations Research and Tinbergen Institute, Erasmus University Rotterdam, 3000 dr Rotterdam, The Netherlands

R. L. Smith: Department of Industrial and Operations Engineering. The University of Michigan, Ann Arbor, Michigan 48109-2117