The Annals of Probability
1991, Vol. 19, No. 3, 1367–1392

GENERATING A RANDOM LINEAR EXTENSION OF A PARTIAL ORDER[1]

BY PETER MATTHEWS

University of Maryland Baltimore County

Given a partial order of N items, a linear extension that is almost uniformly distributed, in the sense of variation distance, is generated. The algorithm runs in polynomial time. The technique used is a coupling for a random walk on a polygonal subset of the unit sphere in $\mathbb{R}^N$. Included is a discussion of how accurately the steps of the random walk must be computed.

1. Introduction. Consider the following artificial situation. A tennis tournament, in which a stronger player will always beat a weaker player, is in progress. The tournament will continue until the players are completely ranked. At the start of the tournament it was fair to say that all $N!$ rankings of the players were equally likely. Given the results from matches already played, all rankings consistent with these results are equally likely. We may ask questions like: What is the probability that player A will end up ranked 1? What is the probability that players A, B and C will end up among the top 5? What is the probability of a particular final ranking? (Or equivalently: What is the number of possible rankings remaining?) Or any number of other similar questions. These questions are, in general, too difficult to answer analytically. One must resort then to Monte Carlo methods. If one could sample nearly uniformly from the set of possible final rankings consistent with the current information, then one could estimate any of the above quantities accurately. If the sampling was not too slow, then this could be accomplished much faster than a computation based on a complete enumeration of all the possible final rankings. Note that such an enumeration could require an amount of computation growing exponentially in N. With Monte Carlo methods it is possible to reduce the calculations to an amount growing polynomially in N, as explained below.

The subject of this paper is the generation of random linear extensions of a partial order. The tennis example above informally defines a partial order. More formally, a partial order on a set $\{x_1, \dots, x_N\}$ of N items is a set of consistent pairwise restrictions $x_{i_k} < x_{j_k}$ for some collection of pairs $(i_1, j_1), \dots, (i_K, j_K)$. The set of linear extensions of the partial order, or the set of all total orderings satisfying all the restrictions, will be denoted $\mathscr{H}$. Though the example above is frivolous, the structure of $\mathscr{H}$ for general partial orders is a serious question. For examples see the collections Pouzet and Richards

Received July 1989; revised February 1990.
[1]Research supported by the National Security Agency under grant MDA 904-88-H-2014.
AMS 1980 *subject classifications*. Primary 60B10; secondary 60J15, 06A10.
Key words and phrases. Random walk, coupling, uniform generation, volume of a polyhedron.

(1984) and Rival (1982). For an application to learning see Goldman, Rivest and Schapire (1989).

Variation distance will be used as a measure of the uniformity of a probability distribution on $\mathscr{H}$. If $\mathscr{U}$ is the uniform distribution on $\mathscr{H}$, $\mathscr{U}(A) = |A|/|\mathscr{H}|$ for $A \subset \mathscr{H}$, and μ is any other probability distribution on $\mathscr{H}$, then

$$\|\mu - \mathscr{U}\| = \max_{A \subset \mathscr{H}} |\mu(A) - \mathscr{U}(A)|.$$

Given $\delta > 0$, we wish to generate an $\mathscr{H}$-valued random variable L with $\|\mathscr{L}(L) - \mathscr{U}\| \leq \delta$, where $\mathscr{L}(L)$ denotes the distribution of L.

To generate such a random ordering L, we first replace $\mathscr{H}$ by a more manageable space. Let S_{N-1} denote the unit sphere in R^N, $S_{N-1} = \{x \in R^N: \sum_1^N x_i^2 = 1\}$. For convenience, $N \geq 6$ will always be assumed. Except on a set of measure 0, the coordinates of $x \in S_{N-1}$ are distinct. Each such point determines an ordering of $\{1, 2, \dots, N\}$, by setting $j < k$ if $x_j < x_k$. The ordering is simply a list of the coordinate indices in order of increasing coordinate values. Ignoring the set of points with two or more coordinates equal, the set of points in S_{N-1} that correspond to linear extensions of a partial order can be described as follows. If the partial order prescribes that $i < j$, let h_{ij} denote the hyperplane $\{x \in R^N: x_i = x_j\}$. Let H_{ij} denote the half-space $\{x \in R^N: x_i \leq x_j\}$. Then $H' = \bigcap H_{ij}$, the intersection of all such half-spaces determined by the partial order, is a convex cone, and $H = S_{N-1} \cap H'$ is a convex polyhedral subset of S_{N-1} whose faces are the intersections of the hyperplanes h_{ij} with S_{N-1}.

The problem of sampling uniformly from $\mathscr{H}$ can be replaced by the problem of sampling uniformly from H. The hyperplanes h_{ij} for $i < j$ divide H into $|\mathscr{H}|$ pieces, one for each ordering of the coordinates. Each ordering of the coordinates corresponds to a linear extension of $\mathscr{H}$. A symmetry argument verifies that these pieces have equal volumes. Thus to sample from $\mathscr{H}$, one simply chooses a point in H and determines the ordering of its coordinates, which is unique with probability 1 and thus specifies a point in $\mathscr{H}$.

The general problem of sampling nearly uniformly from a convex polyhedron is quite difficult. The problem is in some sense as hard as determining the volume of a polyhedron. If one could approximate the volume of an arbitrary convex polyhedron efficiently, then one could sample nearly uniformly from one. Split the polyhedron of interest into two pieces, find the volume of each, and flip an appropriately biased coin to choose one of the pieces to sample from. Split this piece and iterate. This procedure can be repeated until the remaining piece is as small as desired, leading to a point that is nearly uniformly distributed. Going the other way, random sampling allows one to estimate the relative volumes of two pieces of a polyhedron. This procedure can be bootstrapped to give an estimate of the volume of a convex polyhedron to any desired accuracy. See Dyer, Frieze and Kannan (1989) for details.

Dyer and Frieze (1988) have shown that the general problem of computing the volume of a convex polyhedron is #P complete. So, barring any miraculous developments in computer science, Monte Carlo methods will be useful in approximating volumes of polyhedra. Dyer, Frieze and Kannan (1989) give a general algorithm for sampling nearly uniformly from a convex polyhedron in polynomial time, with the sampling distribution close enough to uniform to obtain accurate volume estimates in polynomial time. The essence of their approach is to place a regular lattice of points in the polyhedron, connect nearest neighbors to create a graph and run a random walk on this graph. They use conductance techniques [see Sinclair and Jerrum (1989)] to bound the second largest eigenvalue of the transition matrix, and hence the rate of convergence to uniformity. An alternative technique is suggested here. At least in the special case of a spherical polyhedron H determined by a partial order, it appears theoretically to be faster.

The method discussed here, running a reflecting random walk, can be motivated by Brownian motion. Consider a Brownian motion on H with normal reflection at the boundaries of H. Matthews (1990) shows that such a process converges to uniformity fairly rapidly in variation distance. Unfortunately, there is no technique known to the author for efficiently simulating a Brownian path on H or even its position at some time t with a guarantee on the accuracy of the simulation. A natural approach to try is to approximate the Brownian path by the path of a random walk taking small steps.

The technique used here is to run a random walk X on H with normal reflection at the boundaries. We study the rate of convergence in distribution to uniformity of $X(i)$, without worrying about how well it approximates a Brownian motion. Coupling [see Section 4E of Diaconis (1988)] is used to study the rate of convergence of $\mathscr{L}(X(i))$ to the uniform distribution $\mathscr{U}$ on H. Let Y denote a random walk on H with the same step distribution as X, but with $Y(0) \sim \mathscr{U}$. The steps of X and Y may be dependent on each other. Then $\|\mathscr{L}(X(i)) - \mathscr{U}\| \leq P(X(i) \neq Y(i))$. The coupling technique, then, is to construct the joint process (X, Y) so that their paths meet and stick together as soon as possible. Of course, the process Y is only used to study rates of convergence; it is not involved in any simulations.

The random walk X used here can be described as follows. It is straightforward to find some partial order in $\mathscr{H}$ and a point $X(0)$ in the corresponding section of H. Given a current position $X(i)$, let $C(X(i))$ denote the cap of points in S_{N-1} that are within a distance ε of $X(i)$. Linear rather than geodesic distance will be used here; $x \in C(X(i))$ if $|X(i) - x| \leq \varepsilon$ and $|x| = 1$. A point $X'(i+1)$ is chosen from the uniform distribution on $C(X(i))$. If $X'(i+1) \in H$, then $X(i+1) = X'(i+1)$. Otherwise, $X'(i+1)$ must be reflected back into H. This is formally described in Section 4, but intuitively the reflection is straightforward. Imagine a ball at $X(i)$, shot in the direction of $X'(i+1)$ with just enough energy to go a distance $|X'(i+1) - X(i)|$. The ball performs reflections in the faces of H as it encounters them. The terminal position of the ball is $X(i+1)$. As this random walk is reversible, it is easy to check that its stationary distribution is uniform on H. All random walks discussed hereafter will take steps uniformly chosen from $C(X(i))$ for some $\varepsilon < \sqrt{2}$.

Consider coupling such a random walk with another random walk Y with the same transition distribution, but with $Y(0) \sim \mathscr{U}$. The natural coupling, used in Lindvall and Rogers (1986) for diffusions on manifolds and Matthews

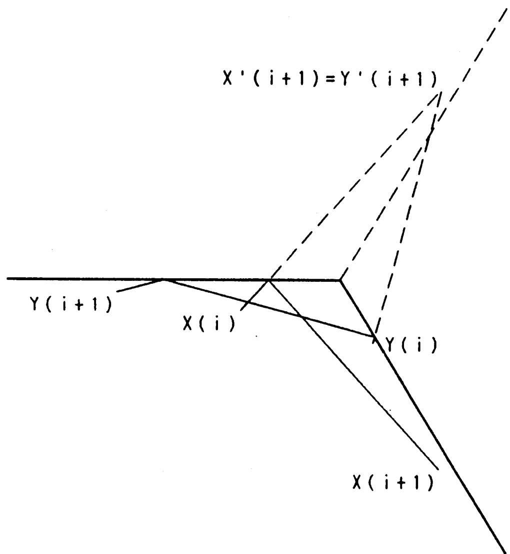

FIG. 1.

(1990) for Brownian motion in convex polyhedra, is as follows. Let h_i denote the hyperplane of points equidistant from $X(i)$ and $Y(i)$. Choose $X'(i+1)$ uniformly from $C(X(i))$. If $X'(i+1) \in C(Y(i))$, let $Y'(i+1) = X'(i+1)$. Otherwise, let $Y'(i+1)$ be the mirror image of $X'(i+1)$ in h_i. This type of joint transition will later be called a reflection step. This coupling is somewhat difficult to study, as noncommuting reflections can force $|X(i+1) - Y(i+1)|$ to be larger than $|X'(i+1) - Y'(i+1)|$. See Figure 1 for an illustration.

Ignoring technical details, to avoid this difficulty the coupling is modified to the following. While X and Y are far apart, if there is any risk of a bad set of reflections pushing them further apart, they move in parallel (later called a rotation step), attempting to preserve their distance. If there is no danger of complicated reflections they take a reflection step. Once they are close together they take only reflection steps. If they couple within a specified number of steps, they are done; otherwise, they begin anew their attempt to couple. The bulk of this paper involves showing that there is a positive probability of coupling in one iteration of this procedure. Theorem 1 in Section 5 is a statement of this result.

From the point of view of a person doing actual simulations, Theorem 1 is not yet satisfactory. Simulation can only be done to a finite number of decimal places, so the question arises: How much accuracy is enough? The notion of a random walk of accuracy γ is discussed. Essentially it is a random walk whose steps are simulated with an error of no more than γ at each step. Section 6 discusses this problem; a sample result is Corollary 2.

COROLLARY 2. *Let $\mathscr{H}$ be the linear extensions of a given partial order of N items. Suppose without loss of generality that $(1,2,3,\dots,N) \in \mathscr{H}$. Let $\mathscr{U}$ be* *the uniform distribution on $\mathscr{H}$. Suppose $\delta > 0$ is given, and a random linear extension L of $\mathscr{H}$ satisfying $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$ is desired. Let $\varepsilon = (6(\log(N-1) + 10)N^4)^{-1}$ and $K = 253[2.17\log(2\delta^{-1})](\log(N-1)+10)^3 N^8$. Let*

$$\gamma = \min\left(\frac{7\delta}{18N^{2.5}(K+1)}, \left(\frac{6\delta}{N^3(K+1)^3}\right)^{1/2}, \frac{\pi\delta}{2\sqrt{2}\,(K+1)^2 N^3 \varepsilon}\right).$$

Let $Z(0)$ be uniformly distributed to accuracy γ on the subset S_I of H, where $S_I = \{x \in S_{N-1}\colon x_1 < x_2 < \cdots < x_N\}$. If a finite precision random walk Z with maximum step size ε is run for K steps, allowing at each step a maximum error of γ in the calculation of $Z(i)$, and L is the order of the coordinates of $Z(K)$, then $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$.

The amount of calculation required is polynomial in N and $\log(\delta^{-1})$. Though only a polynomial amount of work is required, the procedure is probably too slow for practical implementation. However, it is interesting theoretically and has potential to become much faster. To avoid the difficulties of complex reflections, the maximum step size ε was chosen very small here, leading to a large number of steps K. It seems intuitively that complex reflections should not delay coupling very much very often, so that it should be possible to derive a coupling that allows for a much larger maximum step size ε. As the number of steps depends on ε through ε^{-2}, increasing ε by a factor of N could reduce the number of steps required by a factor of N^2. A few clever ideas could reduce the number of required steps K enough to make this a practical algorithm.

The rest of this paper is organized as follows. Sections 2, 3 and 4 give preliminary technical results. Section 2 gives bounds on boundary crossing probabilities for the type of random walk discussed here. Preliminary results concerning geometric probability are in Section 3. They involve probabilities of random points or short arcs on S_{N-1} approaching or crossing one or more hyperplanes of the type bounding H. Reflection and rotation steps are defined in Section 4, and some of their properties are derived. Section 5 gives the coupling discussed above. Finally, Section 6 is a discussion of random walks simulated with finite accuracy.

2. First-passage problems. In this section, the necessary results involving hitting times of Brownian motion W and random walks X on S_{N-1} are given. The time taken to hit a line of latitude $\{(x_1, \dots, x_N)\colon x_1 = -\varepsilon\}$ and the time to move a short distance ε are studied for Brownian motion. The latter gives the time taken to generate a step of an embedded random walk. Together, the two give a bound on the distribution of the numbers of steps required by a random walk to cross two boundaries, which are necessary ingredients of the coupling given later. For brevity, routine computations will generally be omitted.

We first note an elementary fact about projections of a uniformly distributed point on S_{N-1}.

PROPOSITION 2.1. *Let $X = (X_1, X_2, \dots, X_N)$ be uniformly distributed on S_{N-1}. Let h be a k-dimensional subspace of $\mathbb{R}^N$ and let $P_h X$ be the projection of X onto h. Then $|P_h X|^2$ has a Beta$(k/2, (N-k)/2)$ distribution. Further, the squared distance from X to h has a Beta$((N-k)/2, k/2)$ distribution.*

PROOF. Let $Y_1, Y_2, \dots, Y_N$ be iid Normal(0, 1). By spherical symmetry the vector $(\Sigma Y_i^2)^{-1/2}(Y_1, \dots, Y_N)$ is uniformly distributed on S_{N-1}. Also, by spherical symmetry we can take h to be the hyperplane spanned by the first k coordinate vectors without loss of generality. Thus $|P_h X|^2$ has the same distribution as $\Sigma_{i=1}^k Y_i^2 / \Sigma_{i=1}^N Y_i^2$, which is a Beta$(k/2, (N-k)/2)$ distribution.

Since $P(|X| = 1) = 1$, the squared distance from X to h is just the squared length of the projection of X onto the orthogonal complement of h. By the first part of the proposition, this has a Beta$((N-k)/2, k/2)$ distribution. □

For Brownian motion W on S_{N-1}, let W_1 be the first coordinate of W. Let P_x and E_x denote probability and expectation for W_1 started at x. As in Karlin and Taylor (1981), W_1 is a diffusion on $[-1, 1]$ with drift $\mu(x) = -(N-1)x/2$ and infinitesimal variance $\sigma^2(x) = 1 - x^2$. The diffusion W_1 thus has speed measure $M(x)$ with density $m(x) = (1-x^2)^{(N-3)/2}$ and scale function $S(x)$ with derivative $s(x) = (1-x^2)^{-(N-1)/2}$.

Let $T(x) = \inf\{t \geq 0: W_1(t) \leq x\}$. A slight modification of the argument leading to (3.11) and (3.38) of Chapter 15 of Karlin and Taylor (1981) yields

$$E_1 T(x) = 2\int_x^1 s(y) \int_y^1 m(z)\, dz\, dy, \tag{2.1}$$

$$E_1 T^k(x) = 2k\int_x^1 s(y) \int_y^1 m(z) E_z T^{k-1}(x)\, dz\, dy, \tag{2.2}$$

and for $1 > w > x$,

$$E_w T(x) = 2\int_x^w s(y) \int_y^1 m(z)\, dz\, dy. \tag{2.3}$$

PROPOSITION 2.2. *For $\varepsilon > 0$, $N \geq 6$ and $0 < k < \varepsilon^{-1}$,*

$$E_1 T(\varepsilon) \leq \frac{\log(N-1) + 10}{N-1}, \tag{2.4}$$

$$\mathrm{Var}_1 T(\varepsilon) \leq \left(\frac{\log(N-1) + 10}{N-1}\right)^2, \tag{2.5}$$

$$E_{k\varepsilon} T(-\varepsilon) \leq \frac{6\varepsilon}{\sqrt{N-1}}\left(\frac{2}{(1-\varepsilon^2)^{(N-1)/2}} + \frac{k}{(1-k^2\varepsilon^2)^{(N-1)/2}}\right). \tag{2.6}$$

PROOF. These results follow from simple manipulations and bounds on (2.1), (2.2) and (2.3). The calculations are omitted. □

Next consider a random walk X on S_{N-1} taking steps of length less than or equal to ε, where $\varepsilon < \sqrt{2}$. Given $X(i)$, $X(i+1)$ is chosen uniformly from the cap $C(X(i))$. Let X_1 denote the first coordinate of X. Let F_ε denote the cumulative distribution function of $|X(1) - X(0)|$. With $X(0) = (1, 0, \dots, 0)$, $F_\varepsilon(x) = P(|X(1) - X(0)| \le x) = P(X_1(1) \ge 1 - x^2/2)$. If X_1 were uniform on S_{N-1}, $X_1^2(1)$ would have a Beta$(1/2, (N-1)/2)$ distribution by Proposition 2.1. Requiring $X(1)$ to be in $C(X(0))$ is equivalent to conditioning on $X_1(1) \ge 1 - \varepsilon^2/2$. It follows that $F_\varepsilon(x)$ has density function

$$f_\varepsilon(x) = \frac{2x^{N-2}\left(1 - \dfrac{x^2}{4}\right)^{(N-3)/2}}{\int_{(1-\varepsilon^2/2)^2}^{1} y^{-1/2}(1-y)^{(N-3)/2}\,dy} \quad \text{for } x \in [0, \varepsilon]. \tag{2.7}$$

Next embed the random walk X in Brownian motion W. Started at $(1, 0, \dots, 0)$, the process X is distributed like W, also started at $(1, 0, \dots, 0)$, at a sequence of randomized stopping times. Let $R_1, R_2, \dots$ be iid with distribution F_ε. Let $Q_0 = 0$ and

$$Q_i = \inf\{t > Q_{i-1} \colon |W(t) - W(Q_{i-1})| \ge R_i\}.$$

Note that $Q_1, Q_2 - Q_1, Q_3 - Q_2, \dots$ are iid. By the rotational symmetry of X and W, $\{W(Q_i),\ i = 0, 1, \dots\}$ has the same joint distribution as the random walk X. We wish to bound the mean of Q_i from below, and the variance of Q_i from above. Together these will give an upper bound on the amount of time a Brownian motion must be run in order to generate a step of an embedded random walk. Combining these with Proposition 2.2 will give upper bounds on right tail probabilities for hitting times for random walks.

Proposition 2.3. *For Q_i as above*

$$EQ_i \ge \frac{\varepsilon^2(1 - \varepsilon^2/4)^{(N-3)/2}}{N+1}, \tag{2.8}$$

$$EQ_i^2 \le 2\left(\frac{\varepsilon^2}{(N-1)(1-\varepsilon^2/2)}\right)^2. \tag{2.9}$$

Proof. Start W at $W(0) = (1, 0, \dots, 0)$. Then given R_1, Q_1 is the first time W_1 hits $1 - R_1^2/2$.

Using (2.1), simple calculations yield $E(Q_1|R_1) \ge (N-1)^{-1}R_1^2$. Integrating with respect to the density (2.7) of R_1 gives (2.8).

To prove (2.9), note that by (2.2), $EQ_i^2 \le E_1T^2(\varepsilon) \le 2(E_1T(\varepsilon))^2$. A calculation similar to the lower bound on $E(Q_i|R_i)$ gives

$$E_1T(\varepsilon) \le \frac{\varepsilon^2}{(N-1)(1-\varepsilon^2/2)},$$

which gives (2.9). □

Define

$$(2.10)\qquad \rho = \sqrt{\frac{2(N+1)}{(N-1)(1-\varepsilon^2/2)(1-\varepsilon^2/4)^{(N-3)/2}} - 1}.$$

By (2.8) and (2.9) ρ is an upper bound on the coefficient of variation of Q_i.

Now consider two first-passage problems for a random walk. The first is the time taken to come within 1.5ε of the hyperplane $\{x: x_1 = 0\}$ from $(1, 0, \dots, 0)$; the second is the time to cross the hyperplane $\{x: x_1 = 0\}$ from a distance 1.5ε away.

PROPOSITION 2.4. *Let* $X(0) = (1, 0, \dots, 0)$ *and let* $T_x(1.5\varepsilon) = \min\{i: X_1(i) \le 1.5\varepsilon\}$. *Let*

$$\alpha = \frac{(N+1)(\log(N-1)+10)}{(N-1)\varepsilon^2(1-\varepsilon^2/4)^{(N-3)/2}}.$$

For $b > 0$ *let*

$$(2.11)\qquad l = l(b) = \alpha(1+\sqrt{b}) + \rho\sqrt{\alpha}\sqrt{b^{3/2}+b} + b\rho^2,$$

where ρ *is given by* (2.10). *Then* $P(T_X(1.5\varepsilon) > l) \le 2b^{-1}$.

PROOF. Embed X in a Brownian motion W with $W(0) = (1, 0, \dots, 0)$. For any $t > 0$, the two events $\{T_W(0.5\varepsilon) \le t\}$ and $\{Q_l \ge t\}$ imply that $W_1(Q_i) \le 1.5\varepsilon$ for some $i \le l$, and hence $T_X(1.5\varepsilon) \le l$. Thus

$$(2.12)\qquad P(T_X(1.5\varepsilon) > l) \le \inf_{t>0}\left[P_1(T_W(0.5\varepsilon) > t) + P(Q_l < t)\right].$$

For

$$t > \frac{\log(N-1)+10}{N-1},$$

by Proposition 2.2 and Chebychev's inequality,

$$(2.13)\qquad P_1(T_W(0.5\varepsilon) > t) \le \frac{\operatorname{Var}_1 T_W(0.5\varepsilon)}{(t - E_1 T_W(0.5\varepsilon))^2} \le \frac{\left(\dfrac{\log(N-1)+10}{N-1}\right)^2}{\left(t - \dfrac{\log(N-1)+10}{N-1}\right)^2}.$$

This probability will be less than or equal to b^{-1} if

$$(2.14)\qquad t = \frac{(1+\sqrt{b})(\log(N-1)+10)}{N-1}.$$

For

$$t < \frac{l\varepsilon^2(1-\varepsilon^2/4)^{(N-3)/2}}{N+1},$$

by Chebychev's inequality the second term of (2.12) is no larger than

$$\frac{l \operatorname{Var} Q_1}{(lEQ_1 - t)^2}.$$

For t as in (2.14), this term will be less than or equal to b^{-1} if

$$l \geq \frac{2tEQ_1 + b \operatorname{Var} Q_1 + \sqrt{b^2(\operatorname{Var} Q_1)^2 + 4tbEQ_1 \operatorname{Var} Q_1}}{2(EQ_1)^2}.$$

Proposition 2.3 and some simplification yield the value of l quoted in the proposition. □

Next consider the time taken by a random walk X with $X_1(0) = 1.5\varepsilon$ to cross $\{x\colon x_1 = 0\}$.

PROPOSITION 2.5. *Let* $X_1(0) = 1.5\varepsilon$ *and let* $T_X(0) = \min\{i\colon X_1(i) \leq 0\}$. *Let*

$$\beta = \frac{(N+1)}{\varepsilon^2(1-\varepsilon^2/4)^{(N-3)/2}} \frac{6\varepsilon}{\sqrt{N-1}} \left(\frac{2}{(1-\varepsilon^2)^{(N-1)/2}} + \frac{1.5}{\left(1-(1.5)^2\varepsilon^2\right)^{(N-1)/2}} \right).$$

Let

$$m = m(c) = c\left(\beta + \rho\sqrt{\beta} + \rho^2\right). \tag{2.15}$$

Then $P_{1.5\varepsilon}(T_X(0) > m) \leq 2c^{-1}$.

PROOF. The proof is exactly like that of Proposition 2.4 and hence omitted. □

3. Preliminary calculations for a random walk. This section contains some preliminary calculations for random walks on S_{N-1} that will be used in Sections 5 and 6. All the random walks will be started in a distribution similar to the uniform distribution on S_{N-1} and will have no reflections. All the problems involve probabilities of individual positions or paths of a random walk coming close to hyperplanes of the form $h_{ij} = \{x \in \mathbb{R}^N\colon x_i = x_j\}$, or intersections of several hyperplanes of this form. These can be though of as problems in geometric probability. A random arc is chosen on S_{N-1} by choosing an initial point at random and choosing a terminal point uniformly according to a distance requirement. What is the probability that the arc crosses several hyperplanes h_{ij}? What is the probability that the arc comes close to the intersection of several of the hyperplanes h_{ij}? Again, straightforward calculations are omitted.

Let P_{ij} denote projection onto h_{ij}. $I - P_{ij}$ is then projection onto the orthogonal complement $h_{ij}^{\perp}$. For a point $(x_1, x_2, \dots, x_N)$, its order statistics are its coordinates sorted into nondecreasing order, $(x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(N)})$. The uniform order statistics are the order statistics of a point uniformly distributed on S_{N-1}. Let $h_{ijk} = \{x \in R^N\colon x_i = x_j = x_k\}$ for $i < j < k$.

PROPOSITION 3.1. *For Z distributed on S_{N-1} such that its order statistics have the same distribution as the uniform order statistics, $N \geq 6$ and $\delta > 0$,*

$$P\Big(\min_{i<j} |(I - P_{ij})Z| \leq \delta\Big) \leq \frac{3}{7} N^{2.5}\delta$$

and

$$P\Big(\min_{i<j<k} |(I - P_{ijk})Z| \leq \delta\Big) \leq \frac{N^4\delta^2}{12}.$$

PROOF. First note that without loss of generality, Z can be taken to be uniform on S_{N-1}. Let $\sigma \in \Sigma_N$ be a uniformly distributed permutation on N letters, independent of Z. Let $Z_\sigma = (Z_{\sigma(1)}, Z_{\sigma(2)}, \dots, Z_{\sigma(N)})$. Then Z_σ is uniformly distributed on S_{N-1}. The minimum distance to any hyperplane h_{ij} or h_{ijk} is unaffected by this relabeling, so it is sufficient to prove the proposition for Z uniform.

By Proposition 2.1, for any two indices $i < j$, $|(I - P_{ij})Z|^2$ has a Beta$(1/2, (N-1)/2)$ distribution. Elementary bounds and summing over all $\binom{N}{2}$ choices of indices gives the first result. The second result is proved analogously. □

For the next problems define an interaction graph, which describes the set of hyperplanes that a step of a random walk on S_{N-1} encounters. For any two points $y, z \in S_{N-1}$, define a graph G as follows. G has N vertices, corresponding to the N coordinates of S_{N-1}. The edge jk is included in G if the geodesic from y to z intersects h_{ij}. If there is a path from j to k in G, then j and k are said to be connected. A graph is connected if every pair of vertices in it is connected. We say G has a component of order K if there is a set of K vertices such that the subgraph of G consisting of these K vertices and the edges of G between them is connected.

Consider the probability that the interaction graph G of two random points Y and Z contains a component of order 3 or 4.

PROPOSITION 3.2. *Given Y distributed on S_{N-1} such that its order statistics have the same distribution as uniform order statistics, Z chosen uniformly from the cap $C(Y)$ of points within a distance ε of Y, and G the interaction graph of (Y, Z), let C_K be the event that G has a component of order K. Then*

$$P(C_3) \leq \frac{N^3\varepsilon^2}{18(1 - \varepsilon^2/2)^2}$$

and

$$P(C_4) \leq \frac{N^4\varepsilon^3}{17(1 - \varepsilon^2/2)^3}.$$

PROOF. Again Y can be taken to be uniform on S_{N-1}. Consider the first K coordinates of Y, $(Y_1, Y_2, \dots, Y_K)$. Without loss of generality suppose $Y_1 < Y_2 < \cdots < Y_K$. Let C'_K be the event that the subgraph of G made up of vertices $1, 2, \dots, K$ is connected. For example if $K = 3$ this is the event that the segment from Y to Z crosses at least two of the hyperplanes h_{12}, h_{13} and h_{23}. Then $P(C_K) \le \binom{N}{K} P(C'_K)$.

Moving Z further from Y along the geodesic connecting them can only add edges to the interaction graph; hence it can only increase the probability of a component of order K. Thus for simplicity assume $|Y - Z| = \varepsilon$; we will obtain an upper bound on $P(C'_K)$.

First assume $K = 3$. Let h be the hyperplane $\{x \in R^N \colon x_1 = x_2 = x_3\}$. The squared distance from Y to h is

$$|(I - P_{123})Y|^2 = \tfrac{1}{3}\big((Y_1 - Y_2)^2 + (Y_1 - Y_3)^2 + (Y_2 - Y_3)^2\big).$$

By Proposition 2.1 this has a Beta$(1, (N-2)/2)$ distribution. Since $Y_1 < Y_2 < Y_3$ was assumed, $(Y_1 - Y_2)^2 + (Y_2 - Y_3)^2 \le (Y_1 - Y_3)^2$. Thus if the distance from Y to h is d, then $(Y_1 - Y_3)^2 \ge 3d^2/2$. Formally,

$$(3.1) \qquad |(P_{123} - I)Y|^2 = d^2 \Rightarrow |(P_{13} - I)Y|^2 = \frac{(Y_1 - Y_3)^2}{2} \ge \frac{3d^2}{4}.$$

To have the subgraph $(1, 2, 3)$ connected here, $Z_3 \le Z_1$ must hold. Recall how a point Z at a distance ε from Y can be generated. Let T denote the plane tangent to S_{N-1} at Y. Choose a point η in T uniformly from the sphere of points of norm $\varepsilon' = \varepsilon\sqrt{1 - \varepsilon^2/4}\,/(1 - \varepsilon^2/2)$. Then Z is the projection of $Y + \eta$ onto S_{N-1}, or $Z = (Y + \eta)/|Y + \eta|$. Geometrically, Z is across h_{13} from Y if and only if $Y + \eta$ is across h_{13} from Y. This in turn holds if and only if $\eta \cdot (P_{13} - I)Y \ge |(P_{13} - I)Y|^2$. With P_T denoting projection onto T, this in turn implies that $\eta \cdot P_T(P_{13} - I)Y \ge |(P_{13} - I)Y|^2$. Let Θ denote a unit vector in the direction of $P_T(P_{13} - I)Y$, $\Theta = P_T(P_{13} - I)Y / |P_T(P_{13} - I)Y|$. Since η is uniformly distributed on the sphere of radius ε' in T, Proposition 2.1 implies that $\varepsilon'^{-2}|\eta \cdot \Theta|^2$ has a Beta$(1/2, (N-2)/2)$ distribution. Further, it is independent of $d^2 = |(I - P_{123})Y|^2 \sim$ Beta$(1, (N-2)/2)$. Combining all this with (3.1) gives

$$\begin{aligned} P((1,2,3) \text{ is connected}) &\le P(\eta \cdot \Theta \ge |(P_{13} - I)Y|) \\ &\le P\big(|\eta \cdot \Theta|^2 \ge 3d^2/4\big)/2. \end{aligned}$$

This is simply $P(\varepsilon'^2 U \ge 3V/4)/2$, where U is Beta$(1/2, (N-2)/2)$ and V is Beta$(1, (N-2)/2)$ and U and V are independent.

Thus

$$P(C'_K) \le \frac{\Gamma\left(\frac{N}{2}\right)}{2\Gamma\left(\frac{N-2}{2}\right)} \int_{v=0}^{\frac{4}{3}\varepsilon'^2} P\left(U \ge \frac{3v}{4\varepsilon'^2}\right)(1 - v)^{(N-4)/2}\, dv.$$

Changing the variable of integration to $t = 3v/4\varepsilon'^2$ and simplifying gives

$$P(C'_K) \le \frac{N-2}{4}\frac{4\varepsilon'^2}{3}\int_{t=0}^{1} P(U \ge t)\,dt.$$

This is $(N-2)\varepsilon'^2 EU/3 \le \varepsilon'^2/3$, since $EU = (N-1)^{-1}$.

Multiplying by $\binom{N}{3}$ gives the first result.

Next consider $P(C_4)$, the probability that the interaction graph G has a component of order 4. Again we assume $Y_1 < Y_2 < Y_3 < Y_4$ and let h be the hyperplane $\{x \in R^N\colon x_1 = x_2 = x_3 = x_4\}$. The squared distance from Y to h is $|(I - P_{1234})Y|^2$, which has a Beta$(3/2, (N-3)/2)$ distribution by Proposition 2.1. If this squared distance from Y to h is d^2, then $(Y_1 - Y_4)^2 \ge d^2$. Further, $\max((Y_1 - Y_3)^2, (Y_2 - Y_4)^2) \ge d^2/2$.

To have the subgraph of G with vertices $(1,2,3,4)$ connected, either $Z_1 > Z_4$ or both $Z_1 \ge Z_3$ and $Z_2 \ge Z_4$ must hold. Thus

$$P(C'_4) \le P(Z_1 \ge Z_4) + P(Z_1 \ge Z_3 \cap Z_2 \ge Z_4). \tag{3.2}$$

The first term of (3.2) can be bounded as in the case of $K = 3$. It is no larger than $P(\varepsilon'^2 U \ge V)/2$, where U and V are independent Beta$(1/2, (N-2)/2)$ and Beta$(3/2, (N-3)/2)$ random variables, respectively. Calculation shows

$$P(Z_1 \ge Z_4) \le \frac{1.15\varepsilon'^3}{\pi}. \tag{3.3}$$

The second term is bounded in the same fashion. Suppose without loss of generality that $(Y_1 - Y_3)^2 \ge d^2/2$. Proceeding as above,

$$\begin{aligned} P(Z_1 \ge Z_3 \cap Z_2 \ge Z_4) &\le P\left(Z_1 \ge Z_3 | (Y_1 - Y_2)^2 \ge \frac{d^2}{2}\right) \\ &\le \frac{2\sqrt{2}\,(1.15)\varepsilon'^3}{\pi}. \end{aligned} \tag{3.4}$$

Adding (3.3) and (3.4) and multiplying by $\binom{N}{4}$ gives the bound on $P(C_4)$. □

The last problem concerns how close the path from the current position to the next comes to a hyperplane of the form $\{x \in R^N\colon x_i = x_j = x_k\}$ for some $i < j < k$.

PROPOSITION 3.3. *Given Y distributed on S_{N-1} such that its order statistics have the same distribution as the uniform order statistics and Z chosen uniformly $C(Y)$, let*

$$d_3 = \min_{1 \le j < k \le N}\ \min_{0 \le \lambda \le 1} \frac{|(I - P_{ijk})((1-\lambda)Y + \lambda Z)|}{|(1-\lambda)Y + \lambda Z|}.$$

Let $\delta > 0$ be given. Then

$$P(d_3 \le \delta) \le \binom{N}{3}\left(\frac{N\delta^2(1+\varepsilon'^2)}{2} + \frac{2N\sqrt{2(1+\varepsilon'^2)}\,\delta\varepsilon}{\pi}\right).$$

PROOF. Again, without loss of generality, Y can be taken to be uniformly distributed on S_{N-1}. Further

$$P(d_3 \le \delta) \le \binom{N}{3} P\left(\min_{0\le\lambda\le 1} \frac{|(I-P_{123})(\lambda Y + (1-\lambda)Z)|}{|(\lambda Y + (1-\lambda)Z)|} \le \delta\right).$$

Again the squared distance $d^2 = |(I-P_h)Y|$ from Y to $h = \{x \in R^N\colon x_1 = x_2 = x_3\}$ has a Beta$(1, (N-2)/2)$ distribution.

Consider again the tangent plane T to S_{N-1} at Y. Except on a set of probability 0 (which will be ignored hereafter) $T \cap h$ is $N-3$ dimensional; it is spanned by vectors $(x_1, \dots, x_N)$ satisfying the three equations $x_1 = x_2$, $x_2 = x_3$ and $x \cdot Y = 0$. T also contains a vector that is orthogonal to h, namely $(Y_2 - Y_3, Y_3 - Y_1, Y_1 - Y_2, 0, \dots, 0)$. Let n denote a unit vector in this direction; $n = (\sqrt{3}\,d)^{-1}(Y_2 - Y_3, Y_3 - Y_1, Y_1 - Y_2, 0, \dots, 0)$. Let m denote a unit vector that, along with n and a basis for $T \cap h$, gives an orthonormal basis for T. Assume m is oriented so that $m \cdot (I-P_h)Y \ge 0$. Note that $(I-P_h)m$ is orthogonal to n and Y, since both m and $P_h m$ are, and $(I-P_h)m$ is orthogonal to every vector in h, so it must be a multiple of $(Y_1 - \bar{Y}_{123}, Y_2 - \bar{Y}_{123}, Y_3 - \bar{Y}_{123}, 0, \dots, 0) = (I-P_h)Y$.

It suffices to consider Z a distance ε from Y, since increasing the step length can only increase the probability in question. Z can be generated by choosing a random vector η of length $\varepsilon' = \varepsilon\sqrt{1-\varepsilon^2/4}\,/(1-\varepsilon^2/2)$ in T, and taking $Z = (Y+\eta)/|Y+\eta|$. Let the random vector $(U, V) = (\eta \cdot n, \eta \cdot m)$. A simple generalization of Proposition 2.1 shows that $\varepsilon'^{-2}(U^2, V^2)$ has a bivariate Beta$(1/2, 1/2, (N-2)/2)$ distribution. Thus $U^2/(U^2+V^2)$ has a Beta$(1/2, 1/2)$ distribution.

Consider the geodesic arc from Y to Z. Since Y and η are orthogonal, this arc can be written as $(Y+\lambda\eta)/\sqrt{1+\lambda^2\varepsilon'^2}$ for $0 \le \lambda \le 1$. Let $\hat{d}_3^2(\lambda)$ be the squared distance from a point on this arc to h in this parameterization:

$$\hat{d}_3^2(\lambda) = \frac{|(I-P_h)(Y+\lambda\eta)|^2}{1+\lambda^2\varepsilon'^2} = \frac{|(I-P_h)Y + \lambda(Un + V(I-P_h)m)|^2}{1+\lambda^2\varepsilon'^2}.$$

Also let $\hat{d}_3^2 = \min_{\lambda\in[0,1]} \hat{d}_3^2(\lambda)$. Then $P(d_3 \le \delta) \le \binom{N}{3} P(\hat{d}_3 \le \delta)$. Let g denote $|(I-P_h)m|$, and recall $d = |(I-P_h)Y|$. Then

$$\hat{d}_3^2(\lambda) = \frac{(d+\lambda Vg)^2 + \lambda^2 U^2}{1+\lambda^2\varepsilon'^2}.$$

To bound the minimum of $\hat{d}_3^2(\lambda)$ over $(0,1)$ from below, note that

$$\hat{d}_3^2(\lambda) \geq \frac{1}{1+\varepsilon'^2}\left((d+\lambda V g)^2 + \lambda^2 U^2\right).$$

The minimum value of this quadratic is obtained at

$$\lambda^* = \frac{-dVg}{V^2g^2+U^2},$$

with minimizing value

$$\frac{d^2U^2}{(1+\varepsilon'^2)(V^2g^2+U^2)}.$$

If $\lambda^* < 0$, then the minimum value in the interval $[0,1]$ is $\hat{d}_3^2(0)$. Note also that $\hat{d}_3$ cannot be less than or equal to δ if $d > \delta + \varepsilon$.

Thus

$$P\left(\hat{d}_3 \leq \delta\right) \leq P(d^2 \leq \delta^2) + P\left(\left[(\delta+\varepsilon)^2 \geq d^2 > \delta^2\right] \cap [Vgd < 0] \cap \left[\frac{d^2U^2}{(1+\varepsilon'^2)(V^2g^2+U^2)} \leq \delta^2\right]\right).$$

Note that since g is independent of U and V, this probability will not be decreased if g is set to its maximum value 1. Therefore let $g = 1$. Since the sign of V is independent of V^2, d and U, $1/2 = P(V < 0 | g, d, U^2, V^2)$. Using this along with the distributions of d^2 and $U^2/(U^2+V^2)$, we obtain

$$\begin{aligned} P\left(\hat{d}_3 \leq \delta\right) \leq{}& \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-2}{2}\right)} \int_0^{\delta^2} (1-x)^{(N-4)/2}\,dx \\ &+ \frac{1}{2}\frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-2}{2}\right)} \frac{1}{\Gamma^2\left(\frac{1}{2}\right)} \\ &\times \int_{x=\delta^2}^{(\delta+\varepsilon)^2} \int_{z=0}^{\min((1+\varepsilon'^2)\delta^2/x,\,1)} \frac{(1-x)^{(N-4)/2}}{\sqrt{z(1-z)}}\,dz\,dx. \end{aligned}$$

Some manipulation and multiplication by $\binom{N}{3}$ gives the result. □

4. Preliminary results on coupling. In this section some results on two kinds of joint transition distributions for random walks on H are given. The two types of joint transitions will be called reflection and rotation steps. In a reflection step the two processes move as reflections in the hyperplane of points equidistant from them. This is the coupling used in Lindvall and Rogers (1986). In the second type of step, a rotation step, the processes try to preserve their distance. This type of joint transition will be used when the processes are close to a corner of H, so that a reflection step could move them apart.

PROPOSITION 4.1. *Suppose $x, y \in S_{N-1}$ are on opposite sides of an $N-1$-dimensional hyperplane h. Then*

$$|x - (2P_h - I)y| \le |x - y|.$$

Note that $(2P_h - I)y$ is the reflection of y in h.

PROOF. The straightforward proof is omitted. □

Next define some notation involved in generating a step of the random walk. Given $Y(i) \in \operatorname{int} H$, let $Y'(i+1)$ be chosen uniformly from $C(Y(i))$. The next position $Y(i+1)$ of the random walk can be computed as follows. Let $g(t) = Y(i) + t(Y'(i+1) - Y(i))$ for $0 \le t \le 1$. Let $t_1 = \inf\{t > 0\colon g(t) \notin H\}$. Then $g(t_1) \in \partial H$. One can show that with probability 1, $g(t_1)$ lies in a single face of H, say h_{cd}. Define $g_1(t)$ by

$$g_1(t) = \begin{cases} g(t) & \text{for } t \le t_1, \\ (2P_{cd} - I)g(t) & \text{for } t > t_1. \end{cases}$$

Then $g_1(t)$ matches the path from $Y(i)$ to $Y(i+1)$ through its first reflection. Similarly define $t_2, \dots$ and $g_2, \dots$ by $t_j = \inf\{t > t_{j-1}\colon g_{j-1}(t) \notin H\}$ and

$$g_j(t) = \begin{cases} g_{j-1}(t) & \text{for } t \le t_j, \\ (2P_{ef} - I)g_{j-1}(t) & \text{for } t > t_j \end{cases}$$

if $g_{j-1}(t_j) \in h_{ef}$.

Note that each reflection $(2P - I)$ amounts to a transposition of a pair of coordinates. Thus since $g(t)$ does not hit $\bigcup_{a<b<c}\{x\colon x_a = x_b = x_c\}$ with probability 1, neither will any $g_j(t)$. Further each $g_j(t)$ crosses at least one less hyperplane h_{kl} from the entire set of $\binom{N}{2}$ possible hyperplanes than g_{j-1} does, and $g(t)$ crosses each hyperplane at most once, so for some $j \le \binom{N}{2}$, $g_j(t)$ must have its terminal point inside H. This terminal point is then $Y(i+1)$.

Recall the interaction graph G_i of the step from $Y(i)$ to $Y(i+1)$. Vertices j and k are joined by an edge in G_i if the segment from $Y(i)$ to $Y'(i+1)$ crosses h_{jk}.

PROPOSITION 4.2. *If a and b are not connected in G_i, then no reflection of the transformation from $Y'(i+1)$ to $Y(i+1)$ can be in h_{ab}.*

PROOF. Clearly the reflection in the construction of $g_1(t)$ involves connected coordinates. Now consider the interaction graph of the segment

$(g_1(t_1), g_1(1))$. It is the graph of the segment $(g_1(t_1), Y'(i+1))$, except that two coordinates in one connected component have switched labels, and at least one edge has disappeared. Thus a and b cannot be connected in the graph of $(g_1(t_1), g_1(1))$ unless they were connected in G_i. The reflection in the construction of g_2 will involve coordinates that are connected. Continuing inductively, every reflection will involve only connected coordinates. □

Consider now a reflection step. Given $X(i)$ and $Y(i)$ in S_{N-1}, let h be the hyperplane $\{z \in \mathbb{R}^N: |z - X(i)| = |z - Y(i)|\}$ of points equidistant from $X(i)$ and $Y(i)$. Each step $Y(i) \to Y'(i+1)$ is defined in terms of the step $X(i) \to X'(i+1)$ as follows. If $X'(i+1) \in C(X(i)) \cap C(Y(i))$, then $Y'(i+1) = X'(i+1)$. Otherwise, $Y'(i+1) = (2P_h - I)X'(i+1)$, the reflection of $X'(i+1)$ in h. Clearly $Y'(i+1)$ is uniformly distributed in $C(Y(i))$.

PROPOSITION 4.3. *If a reflection step is taken by random walks X and Y on H from $X(i)$ and $Y(i)$ and if the interaction graph G_i of $(X(i), X'(i+1))$ contains no connected component of order* 3 *or more, then*

$$|X(i+1) - Y(i+1)| \le |X'(i+1) - Y'(i+1)|.$$

PROOF. By assumption G_i is composed of isolated singletons and doublets. Since the reflections involved in transforming $X'(i+1)$ to $X(i+1)$ are simply interchanges of coordinates, the reflections in disjoint components of G_i commute. Thus the reflections in transforming $X'(i+1)$ to $X(i+1)$ can be done in any order. Denote this list of reflections by R.

We give a sequential procedure for calculating $X(i+1)$ and $Y(i+1)$ from $X'(i+1)$ and $Y'(i+1)$ with the property that, at each step of the sequence, the distance between the points cannot increase. Let $X_0(i+1) = X'(i+1)$. Begin performing the reflections involved in transforming $Y'(i+1)$ to $Y(i+1)$ in the proper order. Suppose at step j, having performed $j-1$ reflections already, we have $Y_{j-1}(i+1)$ and $X_{j-1}(i+1)$, and suppose the next reflection Y is to perform is in the hyperplane h_{ab}. Define $Y_j(i+1) = (2P_{ab} - I)$ $Y_{j-1}(i+1)$ and

$$X_j(i+1) = X_{j-1}(i+1) \quad \text{if } 2P_{ab} - I \notin R$$

and

$$X_j(i+1) = (2P_{ab} - I)X_{j-1}(i+1) \quad \text{if } 2P_{ab} - I \in R.$$

Delete $2P_{ab} - I$ from R if it is in R, and proceed to the step $j+1$. Eventually $Y_j(i+1)$ will be in H. At this time, perform all the remaining reflections in R.

Proposition 4.1 guarantees that each step of this process only brings the points closer together. If at step j, both $X_{j-1}(i+1)$ and $Y_{j-1}(i+1)$ are reflected in h_{ab}, then certainly $|X_{j-1}(i+1) - Y_{j-1}(i+1)| = |X_j(i+1) - Y_j(i+1)|$. If at step j, only $Y_{j-1}(i+1)$ is reflected in h_{ab}, then $X_{j-1}(i+1)$ must be on the same side of h_{ab} as H, while $Y_{j-1}(i+1)$ must be on the opposite side. Thus by Proposition 4.1,

$$|X_j(i+1) - Y_j(i+1)| \le |X_{j-1}(i+1) - Y_{j-1}(i+1)|.$$

Thus the distance is nondecreasing until $Y(i+1) \in H$ is obtained. Any further reflections involving $X(i+1)$ are in hyperplanes separating it from H. Thus again by Proposition 4.1, the distance cannot be increased by these reflections. □

PROPOSITION 4.4. *Suppose W, X, Y and Z are random walks on H. Suppose $|W(i) - X(i)| \le |Y(i) - Z(i)|$. Suppose each of the pairs (W, X) and (Y, Z) takes a reflection step. Then there is a joint distribution for the four steps such that $|W'(i+1) - X'(i+1)| \le |Y'(i+1) - Z'(i+1)|$.*

PROOF. Let h_W and h_Y denote, respectively, the hyperplanes of points equidistant from $W(i)$ and $X(i)$, and $Y(i)$ and $Z(i)$. Consider the one-dimensional orthogonal subspaces $h_W^{\perp}$ and $h_Y^{\perp}$, and unit vectors e_W and e_Y in each. Assume e_W and e_Y are chosen so that $Y(i) \cdot e_Y \ge W(i) \cdot e_W \ge 0$. Let F_W and F_Y denote, respectively, the cumulative distribution functions of $W'(i+1) \cdot e_W$ and $Y'(i+1) \cdot e_Y$. Elementary geometry shows that F_W is stochastically smaller than F_Y. Further, $P(|W'(i+1) - X'(i+1)| \le 2x) = F_W(x) + F_W(-x)$ for $x > 0$. It follows that $|W'(i+1) - X'(i+1)|$ has a distribution that is stochastically smaller than that of $|Y'(i+1) - Z'(i+1)|$. Thus a joint distribution satisfying $|W'(i+1) - X'(i+1)| \le |Y'(i+1) - Z'(i+1)|$ can easily be constructed. □

Now define a rotation step. We are given $X(i)$, $Y(i)$, and assume $X(i) \ne Y(i)$. Generate $X'(i+1)$ uniformly from $C(X(i))$. $Y'(i+1)$ will be constructed next. With probability 1 there is an $N-2$-dimensional subspace h of points equidistant from $X(i)$, $Y(i)$ and $X'(i+1)$; $h = \{x \in R^N\colon |x - X(i)| = |x - Y(i)| = |x - X'(i+1)|\}$. There is a unique rotation R fixing h satisfying $RX(i) = X'(i+1)$. Let $Y'(i+1) = RY(i)$.

To check that $Y'(i+1)$ is uniformly distributed in $C(Y(i))$, note that the probability distribution of $Y'(i+1)$ is determined by the joint probability distribution of h and R. It is easy to check that picking a point $Y^*(i+1)$ uniformly from $C(Y(i))$, letting $h^* = \{x \in R^N\colon |x - X(i)| = |x - Y(i)| = |x - Y^*(i+1)|\}$, and letting R^* fix h^* and satisfy $R^*Y(i) = Y^*(i+1)$ gives a pair h^*, R^* with the same joint distribution as h, R. Thus $Y'(i+1)$ has the proper distribution.

The following proposition shows that in a rotation step, as long as $X(i)$ and $Y(i)$ are not close together, they cannot both cross from one side of a hyperplane to the other.

PROPOSITION 4.5. *Let h be any $(N-1)$-dimensional hyperplane and suppose $X(i)$ and $Y(i)$ are on the same side of h. Also assume $|X(i) - Y(i)| > \varepsilon$. Suppose $X'(i+1)$ is chosen uniformly from $C(X(i))$ and $Y'(i+1)$ is chosen* *by a rotation step. If $X(i)$ and $X'(i+1)$ are on opposite sides of h, then $Y(i)$ and $Y'(i+1)$ are on the same side of h.*

PROOF. Consider the affine plane Q spanned by $X(i)$, $X'(i+1)$ and $Y(i)$. R as defined above acts as a rotation on Q, fixing the point in it that is equidistant from $X(i)$, $X'(i+1)$ and $Y(i)$. Assume without loss of generality that the angle of rotation of R is no larger than π. R cannot move $Y(i)$ as far as $X(i)$, since their distance is larger than ε. Geometrically, it is clear that for $Y(i)$ to be rotated across $Q \cap h$, it must first pass through $X(i)$, which is impossible. □

PROPOSITION 4.6. *Given random walks X and Y on H, suppose $|X(i) - Y(i)| > \varepsilon$, a rotation step is taken, and the interaction graph G_i^X of $(X(i), X'(i+1))$ has no connected component of order 4 or more. Then $|X(i+1) - Y(i+1)| \le |X(i) - Y(i)|$.*

PROOF. By Proposition 4.5 any edge that is present in G_i^X cannot be present in G_i^Y. By assumption all the connected components of G_i^X are of the four types ($\triangle$, $\wedge$, $|$, $\cdot$). As in Proposition 4.3, in transforming $X'(i+1)$ to $X(i+1)$, the reflections in disjoint components can be done in any order. Again, we give a sequential procedure for performing the reflections transforming $X'(i+1), Y'(i+1)$ to $X(i+1), Y(i+1)$ such that at each step, the distance between the points is not increased. Begin by performing all the X-reflections corresponding to connected components of types $\triangle$ and $|$, that is, complete subgraphs. By Proposition 4.5 none of the edges in which reflections are performed may be present in G_i^Y. By Proposition 4.1 each of these reflections can only decrease the distance between the current values of $X_j(i+1)$ and $Y'(i+1)$. Next for each component of the form $\wedge$, perform the first required reflection, and a second reflection if it is required and does not involve an edge present in G_i^Y. Again each of these reflections can only bring $X_j(i+1)$ closer to $Y'(i+1)$.

After this is complete the remaining edges of G_i^X have no connected component of order 3 or more. An argument exactly like that of the proof of Proposition 4.3 completes the proof. □

Next consider the behavior of two random walks taking rotation steps when their current positions are very close together. The following pair of propositions concludes by showing that, unless both random walks come very close to the intersection of two hyperplanes of the form $\{x \in R^N: x_j = x_k\}$, their final positions will be at least as close as their initial positions.

Suppose $X(i), Z(i) \in S_{N-1}$ satisfy $|X(i) - Z(i)| < \delta$. Suppose R is the rotation fixing vectors orthogonal to the span of $X(i)$, $Z(i)$ and $X'(i+1)$ that satisfies $RX(i) = X'(i+1)$ and $RZ(i) = Z'(i+1)$. From previous discussion R fixes an $N-2$-dimensional subspace of R^N, and rotates the orthogonal complement of this subspace by an angle θ with $|\theta| \le \pi$. For $0 \le t \le 1$ let R_t be the rotation of angle $t\theta$ in the same plane as R. We will say a hyperplane separates two points x and z if they are on opposite sides of it or if at least one of them lies on the hyperplane.

PROPOSITION 4.7. *Suppose* $0 \le t \le 1$ *and* $|X(i) - Z(i)| < \delta$. *Let* h_a, h_b *denote hyperplanes of the form* $\{x \in R^N: x_j = x_k\}$. *Suppose also that*

$$\min_{h_a \neq h_b} |(I - P_{h_a \cap h_b}) R_t X(i)| > \delta.$$

Then at most one hyperplane h_a *separates* $R_t X(i)$ *and* $R_t Z(i)$.

PROOF. Fix $h_a \neq h_b$, and suppose both h_a and h_b separate $R_t X(i)$ and $R_t Z(i)$. Let h be the plane spanned by the normals to h_a and h_b, that is, the plane orthogonal to $h_a \cap h_b$. Then the projections of h_a and h_b onto h are lines, and both lines separate $P_h R_t X(i)$ and $P_h R_t Z(i)$. Also, since $|(I - P_{h_a \cap h_b}) R_t X(i)| > \delta$, $|P_h R_t X(i)| > \delta$. The angle between $P_h h_a$ and $P_h h_b$ is the same as the angle between the normals to h_a and h_b, either $\pi/3$ or $\pi/2$. In either case it is elementary geometry to check that if both lines separate $P_h R_t X(i)$ and $P_h R_t Z(i)$ and $|P_h R_t X(i)| > \delta$, then $|P_h R_t X(i) - P_h R_t Z(i)| \ge \delta$. This implies $|R_t X(i) - R_t Z(i)| \ge \delta$, which is a contradiction, proving the proposition. □

PROPOSITION 4.8. *Suppose* $X(i)$ *and* $Z(i)$ *satisfy* $|X(i) - Z(i)| < \delta$ *and take a rotation step. Let* h_a, h_b *denote hyperplanes of the form* $\{x \in R^N: x_j = x_k\}$. *Suppose also that*

$$\min_{0 \le t \le 1} \min_{h_a \neq h_b} |(I - P_{h_a \cap h_b}) R_t X(i)| > \delta.$$

Then $|X(i+1) - Z(i+1)| \le |X(i) - Z(i)|$.

PROOF. By Proposition 4.7, $X(i)$ and $Z(i)$ are separated by at most one hyperplane h_a. Let $h_1, h_2, \dots, h_n$ denote the sequence of hyperplanes crossed by the path from $X(i)$ to $X'(i+1)$. Then if $h_1 \neq h_a$, by Proposition 4.7 the path from $Z(i)$ to $Z'(i+1)$ must cross $h_a, h_1, \dots, h_{n-1}$ in that order. It may also cross h_n and one additional hyperplane h_{n+1}. If $h_1 = h_a$, then the path from $Z(i)$ to $Z'(i+1)$ must cross $h_2, \dots, h_{n-1}$ in that order, plus possibly h_n and then some h_{n+1}. In any of these situations, the sequences of reflections used to transform $X'(i+1)$ to $X(i+1)$ and $Z'(i+1)$ to $Z(i+1)$ are essentially the same, with the only possibilities for differences being one first crossing a hyperplane inside H, which would not cause a reflection, and one crossing a final hyperplane that the other does not. In any case, the common initial sequence of reflections will preserve their distance, while by Proposition 4.1 a final reflection by one process can only decrease the distance. □

5. The coupling. In this section a coupling for a random walk on $H \subset S_{N-1}$ is given. The random walk starting in the uniform (stationary) distribution will be denoted Y. The random walk whose rate of convergence to uniformity is of interest will be denoted X. $X(0)$ can have any probability distribution on H; often point masses are considered. In the next section, on finite precision calculations, it will be convenient for $X(0)$ to be uniform on a subset of H.

Before giving the final coupling result, we give a proposition that contains the main substance of the result.

PROPOSITION 5.1. *Given $H \subset S_{N-1}$, the points in S_{N-1} satisfying a given partial order $\mathscr{H}$, and x and y arbitrary points in H, let X and Y be random walks on H started at x and y. Let $k_1 = (1 - aN^4\varepsilon^2/12)^{-1}l(b)$ and $k_2 = m(c)$, with $l(b)$ and $m(c)$ given by* (2.11) *and* (2.15). *Let $k = k_1 + k_2$. Suppose ε is small enough to satisfy*

$$\frac{k_1 N^4 \varepsilon^3}{17(1 - \varepsilon^2/2)^3} \le p_1, \tag{5.1}$$

$$\frac{k_2 N^3 \varepsilon^2}{18(1 - \varepsilon^2/2)^2} \le p_2. \tag{5.2}$$

Then there is a coupling for X and Y such that

$$P(X(k) = Y(k)) \ge p = 1 - \left(a^{-1} + 4b^{-1} + 4c^{-1} + p_1 + p_2\right). \tag{5.3}$$

PROOF. Define a third random walk W with the same marginal transitions as X and Y, but with $W(0)$ uniform on H. If both X and Y couple with W in k steps, then they also couple with one another. The steps of X and Y will be defined by the steps of W as follows.

The definition will change at time k_1; first consider the paths up to time k_1. If $i < k_1$ and $W(i)$ is within a distance ε of any hyperplane of the form $\{x \in R^N: x_l = x_m = x_n\}$ for $l < m < n$, then both X and Y take rotation steps based on the step $W(i) \to W'(i + 1)$. If $|X(i) - W(i)| \le 3\varepsilon$, then X takes a rotation step based on W. Similarly if $|Y(i) - W(i)| \le 3\varepsilon$, then Y takes a rotation step based on W. If X (or Y) has not been required to take a rotation step above, then it takes a reflection step based on $W(i) \to W'(i + 1)$. It will be shown that, with probability at least $1 - (a^{-1} + 4b^{-1} + p_1)$, both $X(k_1)$ and $Y(k_1)$ are within 3ε of $W(k_1)$.

First consider the number of reflection steps generated by W in its first k_1 steps. Since $W(i)$ is always uniformly distributed, by Proposition 3.1 at each time point it has probability at most $N^4\varepsilon^2/12$ of being within ε of any hyperplane of the form $\{x \in R^N: x_l = x_m = x_n\}$ for $l < m < n$, so the expected number of rotation steps among the first k_1 is at most $k_1 N^4 \varepsilon^2/12$. By Markov's inequality with probability at least $1 - a^{-1}$, there are at least $k_1(1 - aN^4\varepsilon^2/12)$ reflection steps among the first k_1 steps. This can be restated as

$$P(\text{at least } l(b) \text{ reflection steps among first } k_1) \ge 1 - a^{-1}. \tag{5.4}$$

Let $\mathscr{A}$ denote the event that no component of order 4 arises in the interaction graph of $(W(i-1), W'(i))$ for all $i < k_1$. $W(i)$ is uniformly distributed for each i, and Proposition 3.2 gives the probability of a component of order 4 or more at each step, so by assumption (5.1),

$$P(\mathscr{A}) \geq 1 - \frac{k_1 N^4 \varepsilon^3}{17(1 - \varepsilon^2/2)^3} = 1 - p_1. \tag{5.5}$$

Now we can show that with high probability both $X(k_1)$ and $Y(k_1)$ are within 3ε of $W(k_1)$. First consider $X(k_1)$. Define a counting random variable $N(j)$ for $j > 0$ to be the number of potential reflection steps generated by W in its walk from $W(0)$ to $W(j)$; $N(j+1) = N(j)$ if $W(j)$ is within ε of a hyperplane of the form $\{x \in R^N\colon\ x_l = x_m = x_n\}$ for $l < m < n$, and $N(j+1) = N(j) + 1$ otherwise. For $j \geq 1$ let $N^{-1}(j)$ be the time W generates its jth reflection step; $N^{-1}(1)$ is the first time $i \geq 0$ that $W(i)$ is not within ε of a hyperplane of the form $\{x \in R^N\colon\ x_l = x_m = x_n\}$ for $l < m < n$, and for $j \geq 2$, $N^{-1}(j)$ is the first time i after $N^{-1}(j-1)$ that $W(i)$ is not within ε of a hyperplane of the form $\{x \in R^N\colon\ x_l = x_m = x_n\}$ for $l < m < n$.

Define two new random walks W^* and X^* on S_{N-1} (nonreflecting) as follows. Let $W^*(0) = W(0)$ and $X^*(0) = X(0)$. Now define $W^*(i)$ and $X^*(i)$ for $i \leq N(k_1)$. Given $W^*(j-1)$ and $X^*(j-1)$, consider the pairs $\{W(N^{-1}(j)), W'(N^{-1}(j)+1)\}$ and $\{X(N^{-1}(j)), X'(N^{-1}(j)+1)\}$. These pairs are part of the construction of a reflection step involving W and X. If $|W(N^{-1}(j)) - X(N^{-1}(j))| \leq |W^*(j-1) - X^*(j-1)|$, then by Proposition 4.4 there is a conditional distribution for $W^*(j)$ and $X^*(j)$ given $W(N^{-1}(j)), W'(N^{-1}(j)+1)$, $X(N^{-1}(j))$ and $X'(N^{-1}(j)+1)$ involving a reflection step such that $|X'(N^{-1}(j)+1) - W'(N-1(j)+1)| \leq |X^*(j) - W^*(j)|$. In this case, generate $X^*(j)$ and $W^*(j)$ accordingly. Otherwise, generate $X^*(j)$ and $W^*(j)$ independently with the proper marginal distributions.

Note that on $\mathscr{A}$, for $0 \leq i \leq k_1$, $|X(i) - W(i)| \leq \max(3\varepsilon, |X^*(N(i)) - W^*(N(i))|)$. This follows by induction. At time 0, the two distances are equal. Suppose the inequality holds at time i. For the next step of W and X, one of three things can happen. If $|W(i) - X(i)| \leq 3\varepsilon$, then X takes a rotation step based on W, and by Proposition 4.6, on $\mathscr{A}$, $|X(i+1) - Y(i+1)| \leq 3\varepsilon$. Second, if $|W(i) - X(i)| > 3\varepsilon$, and $W(i)$ is within ε of a hyperplane of the form $\{x \in R^N\colon\ x_l = x_m = x_n\}$ for $l < m < n$, then N is not incremented, so $X^*(N(i+1)) = X^*(N(i))$ and $W^*(N(i+1)) = W^*(N(i))$. W and X take a rotation step, and by Proposition 4.6, on $\mathscr{A}$, their distance is not increased. Thus the assertion holds here as well. In the third case, $|W(i) - X(i)| > 3\varepsilon$ and $W(i)$ generates a reflection step, incrementing N. By induction and the construction of $W^*(N(i+1))$ and $X^*(N(i+1))$ above, $|X'(i+1) - W'(i+1)| \leq |W^*(N(i+1)) - X^*(N(i+1))|$. However, by Proposition 4.3, $|X(i+1) - W(i+1)| \leq |X'(i+1) - W'(i+1)|$, so the claim is proved.

On $\mathscr{A}$, $X(k_1)$ and $W(k_1)$ will be within 3ε of each other if X^* and W^* come within 3ε of each other before the random time $N(k_1)$. By (5.4), $P(N(k_1) \geq$

$l(b)) \geq a^{-1}$. Combining this with Proposition 2.4 and (5.5) shows that

$$P(|X(k_1) - W(k_1)| \leq 3\varepsilon) \geq 1 - p_1 - a^{-1} - 2b^{-1}.$$

The exact same argument can be made for $Y(k_1)$ and $W(k_1)$ being within 3ε of each other except on a set of probability at most $1 - p_1 - a^{-1} - 2b^{-1}$. Since the two sets where $X(k_1)$ and $Y(k_1)$ may not be close to $W(k_1)$ both contain $\mathscr{A}^c$ and the set where W generated less than $l(b)$ potential reflection steps,

$$P(|X(k_1) - W(k_1)| \leq 3\varepsilon \cap |Y(k_1) - W(k_1)| \leq 3\varepsilon) \geq 1 - \left(p_1 + a^{-1} + 4b^{-1}\right). \tag{5.6}$$

Now let X and Y take reflection steps based on W for the next k_2 steps. If both are already close to W, they will both couple to W with high probability.

Let $\mathscr{B}$ denote the set of sample paths where the interaction graphs of $(W(i-1), W'(i))$ contain no components of order 3 or more for $k_1 + 1 \leq i \leq k$. Since each of $W(i)$, $k_1 \leq i < k$, is uniformly distributed on H, by Proposition 3.2 and assumption (5.2),

$$P(\mathscr{B}) \geq 1 - p_2. \tag{5.7}$$

To show X and W couple with high probability by time k, again define two random walks X' and W' on S_{N-1} with $|X'(0) - W'(0)| = 3\varepsilon$. X' and W' always take reflection steps. As above X' and W' can be defined so that on $\mathscr{B} \cap \{|X(k_1) - W(k_1)| \leq 3\varepsilon\}$ for $0 \leq i \leq k_2$, $|X'(i) - W'(i)| \geq |X(k_1 + i) - W(k_1 + i)|$. X' and W' will certainly couple by time k if W' crosses the hyperplane of points equidistant from $W'(0)$ and $X'(0)$ between times 1 and k_2. By Proposition 2.5 this has probability at least $1 - 2c^{-1}$. As in the previous construction involving a pair W^* and X^*, one can show that on $\mathscr{B}$ a coupling for X' and W' by time k_2 implies a coupling for X and W by time k, assuming $X(k_1)$ and $W(k_1)$ were within 3ε of each other. A similar argument can be made for Y coupling with W. Again note that $\mathscr{B}^c$ is part of the two sets of sample points where X and Y may not couple with W.

Tying it all together, X and Y will couple by time k unless one of the following events occurs.

1. The sample path is not in $\mathscr{A}$. This has probability at most p_1.
2. W does not generate $l(b)$ potential reflection steps by time k_1. This has probability at most a^{-1}.
3. $|X^*(l(b)) - W^*(l(b))| > 3\varepsilon$. This has probability at most $2b^{-1}$.
4. $|Y^*(l(b)) - W^*(l(b))| > 3\varepsilon$. This has probability at most $2b^{-1}$.
5. The sample path is not in $\mathscr{B}$. This has probability at most p_1.
6. X' and W' do not couple by time k_2. This has probability at most $2c^{-1}$.
7. Y' and W' do not couple by time k_2. This has probability at most $2c^{-1}$.

Subtracting all these probabilities from one gives the result. □

THEOREM 1. *Let $H \subset S_{N-1}$ be the points in S_{N-1} satisfying a given partial order $\mathscr{H}$. Let $X(0)$ have an arbitrary probability distribution on H. Let $\delta > 0$*

be given. Assume a step size ε and a number of steps k have been chosen in Proposition 5.1, *giving a coupling probability p. Assume K satisfies*

$$K \geq k\left\lceil \frac{\log \delta}{\log(1-p)} \right\rceil.$$

Let X take steps uniformly chosen from a cap of radius ε. Recall that $\mathscr{U}$ denotes the uniform distribution on H. Then

$$\|\mathscr{L}(X(K)) - \mathscr{U}\| \leq \delta.$$

PROOF. By an easy extension of Lemmas 4 and 5 of Chapter 4E of Diaconis (1988), for any coupling of X and Y with $Y(0)$ uniform, $\|\mathscr{L}(X(K)) - \mathscr{U}\| \leq P(X(K) \neq Y(K))$. Therefore it suffices to show that X and Y can be coupled by time K with probability at least $1 - \delta$.

Let $j = \lceil \log \delta / \log(1 - p) \rceil$. Run the coupling used in Proposition 5.1 j times, each time for k steps. Each time, independently of the past, there is probability at least p of coupling. Therefore the probability of not being coupled after j attempts is at most $(1-p)^j$. By the definition of j, this is at most δ. □

To avoid vacuousness, we must exhibit some choice of ε and k for which the probability of coupling (5.3) is positive. This is implicit in the following example of the allocation of error probabilities to obtain a specific step size and number of steps.

COROLLARY 1. *Suppose $N \geq 6$, and a random walk X on $H \subset S_{N-1}$ taking steps of maximum size $\varepsilon = (6N^4(\log(N-1) + 10))^{-1}$ starting with an arbitrary initial distribution is run for $K = 253\lceil 2.17 \log(\delta^{-1})\rceil(\log(N-1) + 10)^3 N^8$ steps. Then*

$$\|\mathscr{L}(X(K)) - \mathscr{U}\| \leq \delta.$$

PROOF. Set $a = 1000$, $b = 16$ and $c = 16$, and apply Theorem 1. Then

$$k_1 \leq 252(\log(N-1) + 10)^3 N^8,$$

so $p_1 \leq 0.07$. Also

$$k_2 \leq 27(\log(N-1) + 10) N^{4.5},$$

so $p_2 \leq 0.059$. Thus $k = k_1 + k_2 \leq 253(\log(N-1) + 10)^3 N^8$. In each attempt at coupling, the failure probability is at most 0.63. Straightforward calculation gives the result. □

A better leading constant can be obtained by a more careful analysis, particularly if the assumption $N \geq 6$ is changed to a larger lower bound. The factor $\log(N-1) + 10$ can also be improved. However, the factor $\log(\delta^{-1})N^8$ appears to be intrinsic to this technique; some new ideas will be required to reduce it to a more practical value.

6. Finite precision calculations. The calculations thus far have assumed that the random walk X on H was generated with perfect accuracy. For actually simulating $X(0), \dots, X(K)$, this is not realistic. Intuitively, if enough digits of X were generated at each step, the final position $X(K)$ would be close to uniform anyhow. The natural question is: How many digits are enough? This section gives a rigorous answer.

Think of an "infinite precision" random walk X and a "finite precision" random walk Z. Z starts off close to X and tries to stay close by taking rotation steps. It may get separated for two reasons. The first is the gradual accumulation of roundoff errors. The second is more subtle. Once X and Z are slightly separated, if they cross several hyperplanes in a different order and if the reflections in these hyperplanes do not commute, then they may be pushed apart a distance on the order of magnitude of the step size. Many of these could push them very far apart. This section discusses the allowable roundoff error at each step such that the roundoff error alone does not separate $X(K)$ and $Z(K)$ too much, and that $X(i)$ and $Z(i)$ are never far enough apart so that a separation from noncommuting reflections is very likely.

Define a finite precision random walk Z of accuracy γ on H as follows. Given any probability distribution $\mathscr{V}$ on H, say Y has distribution $\mathscr{V}$ to accuracy γ if there exists a bivariate random variable (X', Y'), where X' has distribution $\mathscr{V}$, Y' has the same distribution as Y, and $P(|Y' - X'| > \gamma) = 0$. Given a planned initial distribution, this specifies a class of possible distributions for $Z(0)$, all of which match it to accuracy γ. The steps of Z are defined similarly. Given $Z(i)$, $Z'(i+1)$ must be chosen from a distribution which, to accuracy γ, is the uniform distribution on $C(Z(i))$. The choice of the distribution approximating the uniform may depend on $Z(i)$; it need not be the same at every step. Again this does not uniquely specify a random walk; any Markov process satisfying these requirements will do. Simulating a random walk of accuracy $\gamma > 0$ is possible with finite precision normal number random numbers.

Let X be an infinite precision random walk on H, the type discussed in the previous sections. To obtain results easily, it is necessary to make an assumption about the distribution of $X(0)$. Let $S_I \subset H$ be the set of points in S_{N-1} satisfying a particular linear extension in $\mathscr{H}$. Assume $X(0)$ is uniformly distributed on S_I. By relabeling coordinates if necessary, S_I can be taken to be $\{x \in S_{N-1}: x_1 < x_2 < \cdots < x_N\}$. In theory $X(0)$ could be generated by taking $Y_1, Y_2, \dots, Y_N$ iid $N(0,1)$, normalizing them to have sum of squares equal to 1, then sorting the normalized values.

Now define a finite precision approximation Z to X. Let $Z(0) \in H$ be such that $P(|Z(0) - X(0)| \le \gamma) = 1$. Given $Z(i)$ and $X(i)$, define $Z(i+1)$ and $X(i+1)$ as follows. Let $X'(i+1)$ be chosen uniformly from $C(X(i))$. Choose $Z''(i+1)$ by the rotation step method. Let $Z'(i+1)$ be any point satisfying $P(|Z'(i+1) - Z''(i+1)| \le \gamma) = 1$. Now construct $X(i+1)$ and $Z(i+1)$ by performing reflections as usual. By using random number generators and a finite amount of computation, the random walk Z can be simulated. Of course,

Z can and in practice should be simulated without consideration of X; marginally Z is just a random walk of accuracy γ.

THEOREM 2. *Let $\mathscr{H}$ be the linear extensions of a given partial order of N items. Suppose without loss of generality that* $(1, 2, 3, \dots, N) \in \mathscr{H}$. *Let $\mathscr{U}$ be the uniform distribution on $\mathscr{H}$. Suppose $\delta > 0$ is given, and a random linear extension L of $\mathscr{H}$ satisfying $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$ is desired. Suppose ε and K have been chosen so that for an infinite precision random walk, $\|\mathscr{L}(X(K)) - \mathscr{U}\| \le \delta/2$. Let*

$$\gamma = \min\left(\frac{7\delta}{18N^{2.5}(K+1)}, \left(\frac{6\delta}{N^3(K+1)^3}\right)^{1/2}, \frac{\pi\delta}{2\sqrt{2}\,(K+1)^2N^3\varepsilon}\right).$$

Let $Z(0)$ be uniformly distributed to accuracy γ on the subset S_I of H, where $S_I = \{x \in S_{N-1}\colon x_1 < x_2 < \cdots < x_N\}$. If a finite precision random walk Z with maximum step size ε is run for K steps, allowing at each step a maximum error of γ in the calculation of $Z(i)$, and L is the order of the coordinates of $Z(K)$, then $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$.

PROOF. Consider an infinite precision random walk X with $X(0)$ uniform on S_I and $|X(0) - Z(0)| \le \gamma$. Attempt to couple X with an infinite precision random walk Y started with $Y(0)$ uniform on H. By Proposition 1 and the assumptions, $P(X(K) = Y(K)) \ge 1 - \delta/2$. We will consider the probability that the coordinates of $Z(K)$ have the same ordering as those of $Y(K)$.

First consider $P(|Z(K) - X(K)| \ge (K+1)\gamma)$. Let R_t, h_a and h_b be as in Proposition 4.8. Let $\mathscr{A}$ be the set of sample paths where

$$\min_{0\le t\le 1}\ \min_{0\le i\le K}\ \min_{h_a\ne h_b} |(I - P_{h_a\cap h_b})R_t X(i)| \ge (K+1)\gamma.$$

By Proposition 3.3,

$$P(\mathscr{A}) \le K\binom{N}{3}\left(\frac{N(K+1)^2\gamma^2(1+\varepsilon'^2)}{2} + \frac{2N\sqrt{2(1+\varepsilon'^2)}\,(K+1)\gamma\varepsilon}{\pi}\right),$$

where $\varepsilon' = \varepsilon\sqrt{1 - \varepsilon^2/4}\,/(1 - \varepsilon^2/2)$. For γ, K and ε above, this simplifies to $P(\mathscr{A}) \le \delta/3$. Also, by Proposition 3.1, the probability that $X(K)$ is within $(K+1)\gamma$ of any hyperplane h_a is less than or equal to $\delta/6$.

Thus, except on a set of probability at most $\delta(\frac{1}{2} + \frac{1}{3} + \frac{1}{6})$, X is coupled with Y at time K, $|X(K) - Z(K)| \le (K+1)\gamma$ and $\min_{i<j}|(I - P_{ij})X(K)| > (K+1)\gamma$. Thus, except on a set of probability δ, the coordinates of $Z(K)$ and $Y(K)$ have the same order. It follows that if L is the ordering of the coordinates of $Z(K)$, then $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$. □

COROLLARY 2. *Let $\mathscr{H}$ be the linear extensions of a given partial order of N items. Suppose without loss of generality that* $(1, 2, 3, \dots, N) \in \mathscr{H}$. *Let $\mathscr{U}$ be* *the uniform distribution on $\mathscr{H}$. Suppose $\delta > 0$ is given, and a random linear extension L of $\mathscr{H}$ satisfying $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$ is desired. Let $\varepsilon = (6(\log(N-1)+10)N^4)^{-1}$ and $K = 253[2.17\log(2\delta^{-1})](\log(N-1)+10)^3N^8$. Let*

$$\gamma = \min\left(\frac{7\delta}{18N^{2.5}(K+1)}, \left(\frac{6\delta}{N^3(K+1)^3}\right)^{1/2}, \frac{\pi\delta}{2\sqrt{2}\,(K+1)^2N^3\varepsilon}\right).$$

Let $Z(0)$ be uniformly distributed to accuracy γ on the subset S_I of H, where $S_I = \{x \in S_{N-1}\colon x_1 < x_2 < \cdots < x_N\}$. If a finite precision random walk Z with maximum step size ε is run for K steps, allowing at each step a maximum error of γ in the calculation of $Z(i)$, and L is the order of the coordinates of $Z(K)$, then $\|\mathscr{L}(L) - \mathscr{U}\| \le \delta$.

PROOF. This is Theorem 2 applied to the case of Corollary 1. □

The accuracy γ requires simulation of each coordinate of Z to about $\log_{10}(N/\gamma)$ decimal places. This is logarithmic in N and δ^{-1}, and for reasonable N and δ is not much worse than double precision arithmetic.

Acknowledgment. I would like to thank a referee for a careful and helpful reading of this paper.

REFERENCES

DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. IMS, Hayward, Calif.

DYER, M. and FRIEZE, A. (1988). On the complexity of computing the volume of a polyhedron. *SIAM J. Comput.* **17** 967–974.

DYER, M., FRIEZE, A. and KANNAN, R. (1989). A random polynomial time algorithm for approximating the volume of convex bodies. In *Proc. 21st Ann. ACM Symp. Theory of Computing* 375–381. Assoc. Comput. Mach., New York.

GOLDMAN, S., RIVEST, R. and SCHAPIRE, R. (1989). Learning binary relations and total orders. In *Proc. 30th IEEE Symp. Foundations of Computer Science* 46–51. IEEE, New York.

KARLIN, S. and TAYLOR, H. (1981). *A Second Course in Stochastic Processes*. Academic, New York.

LINDVALL, T. and ROGERS, L. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872.

MATTHEWS, P. (1990). Mixing rates for Brownian motion in a convex polyhedron. *J. Appl. Probab.* **27** 259–268.

POUZET, M. and RICHARD, D., eds. (1984). *Orders, Description and Roles*. North-Holland, Amsterdam.

RIVAL, I., ed. (1982). *Ordered Sets*. Reidel, Dordrecht.

SINCLAIR, A. and JERRUM, M. (1989). Approximate counting, uniform generation, and rapidly mixing Markov chains. *Inform. and Comput.* **82** 93–133.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MARYLAND BALTIMORE COUNTY
CATONSVILLE, MARYLAND 21228