JOURNAL OF MULTIVARIATE ANALYSIS 25, 111–118 (1988)

Random Walks on Z_2^n

PAUL ERDÖS

Hungarian Academy of Sciences, Budapest, Hungary

AND

ROBERT W. CHEN

Department of Mathematics and Computer Science, University of Miami, Coral Gables, Florida 33124

Communicated by P. R. Krishnaiah

For each positive integer $n \geq 1$, let Z_2^n be the direct product of n copies of Z_2, i.e., $Z_2^n = \{(a_1, a_2, ..., a_n) | a_i = 0 \text{ or } 1 \text{ for all } i = 1, 2, ..., n\}$ and let $\{W_j^n\}_{j \geq 0}$ be a random walk on Z_2^n such that $P\{W_0^n = A\} = 2^{-n}$ for all A's in Z_2^n and $P\{W_{j+1}^n = (a_2, a_3, ..., a_n, 0) | W_j^n = (a_1, a_2, ..., a_n)\} = P\{W_{j+1}^n = (a_2, a_3, ..., a_n, 1) | W_j^n = (a_1, a_2, ..., a_n)\} = \frac{1}{2}$ for all $j = 0, 1, 2, ...,$ and all $(a_1, a_2, ..., a_n)$'s in Z_2^n. For each positive integer $n \geq 1$, let C_n denote the covering time taken by the random walk W_j^n on Z_2^n to cover Z_2^n, i.e., to visit every element of Z_2^n. In this paper, we prove that, among other results, $P\{\text{except finitely many } n,\ c2^n \ln(2^n) < C_n < d2^n \ln(2^n)\} = 1$ if $c < 1 < d$. © 1988 Academic Press, Inc.

For each positive integer $n \geq 1$, let Z_2^n be the direct product of n copies of Z_2, i.e., $Z_2^n = \{(a_1, a_2, ..., a_n) | a_i = 0 \text{ or } 1 \text{ for all } i = 1, 2, ..., n\}$ and let $\{W_j^n\}_{j \geq 0}$ be a random walk on Z_2^n such that $P\{W_0^n = A\} = 2^{-n}$ for all A's in Z_2^n and $P\{W_{j+1}^n = (a_2, a_3, ..., a_n, 0) | W_j^n = (a_1, a_2, ..., a_n)\} = P\{W_{j+1}^n = (a_2, a_3, ..., a_n, 1) | W_j^n = (a_1, a_2, ..., a_n)\} = \frac{1}{2}$ for all $j = 0, 1, 2, ...$ and all $(a_1, a_2, ..., a_n)$'s in Z_2^n. For each positive integer $n \geq 1$, let C_n denote the covering time taken by the random walk W_j^n on Z_2^n to cover Z_2^n, i.e., to visit every element of Z_2^n. In this paper, we prove that, among other results, $P\{\text{except finitely many } n,\ c2^n \ln(2^n) < C_n < d2^n \ln(2^n)\} = 1$ if $c < 1 < d$.

In [2], Matthews studied a different random walk on Z_2^n. His random walk can be described as follows: Let μ_n be a probability measure on Z_2^n, for each positive integer $n \geq 1$, that puts mass p_n on $(0, 0, ..., 0)$ and mass $(1 - p_n)/n$ on each of $(1, 0, ..., 0)$, $(0, 1, 0, ..., 0)$, ..., $(0, 0, ..., 0, 1, 0)$, and

Received September 17, 1986.

AMS 1980 subject classifications: primary 60F; secondary 60G.

Key words and phrases: random walks, Borel–Cantelli lemma, covering time.

0047-259X/88 $3.00

683/25/1-8

(0, 0, ..., 0, 1). For each step the random walk on Z_2^n corresponding to μ_n does not move with probability p_n, otherwise it changes exactly one coordinate, with each coordinate equally likely to be changed. He proved that $P\{(C_n - 2^n ln(2^{n+1}))\ 2^{-n} \leqslant x\} \to \exp(-e^{-x})$ for all x if $\sup_n p_n < 1$. Our result is similar to his. However, his technique does not seem applicable to the random walk w_γ^n in this paper. A completely different method is used to obtain our results.

For ease of presentation, we introduce the following fair coin tossing process $\{X_m\}_{m \geqslant 1}$ as follows: $\{X_m\}_{m \geqslant 1}$ is a sequence of independent and identically distributed random variables such that $P(X_1 = 0) = P(X_1 = 1) = \frac{1}{2}$. For each positive integer $n \geqslant 1$, let T_n denote the first occurrence time such that $(X_1, X_2, ..., X_{T_n})$ contains all A's in Z_2^n, i.e., $T_n = \inf\{k \mid \text{each } A \text{ in } Z_2^n \text{ appears in } (X_1, X_2, ..., X_k) \text{ at least once}\}$, $= \infty$ if no such k exists. It is easy to see that $C_n = T_n - n$ for all $n \geqslant 1$. Now we start with the following notation and definitions.

For each element $A = (a_1, a_2, ..., a_n)$ in Z_2^n, the positive integer i $(1 \leqslant i \leqslant n)$ is called a period of A if $(a_1, a_2, ..., a_{n-i}) = (a_{i+1}, a_{i+2}, ..., a_n)$. Let τ_A denote the minimal period of A which is defined by $\tau_A = \min\{i \mid 1 \leqslant i \leqslant n \text{ and } i \text{ is a period of } A\}$.

LEMMA 1. *For any two elements A and B in Z_2^n and any positive integer m, $P\{(X_1, X_2, ..., X_m)$ contains $A\} \leqslant P\{(X_1, X_2, ..., X_m)$ contains $B\}$ if $\tau_A < \tau_B$.*

Proof. See page 186 of [1].

LEMMA 2. *For any element A in Z_2^n and $\tau_A \geqslant k$, then $\{1 - n2^{-k}\}(n+1) \times 2^{-n} \leqslant P\{(X_1, X_2, ..., X_n)$ contains $A\} \leqslant (n+1)\, 2^{-n}$.*

Proof. For each integer $i = 1, 2, ..., n+1$, let $E_i = \{(X_i, X_{i+1}, ..., X_{i+n-1}) = A\}$. Then $P\{(X_1, X_2, ..., X_{2n})$ contains $A\} = P\{\bigcup_{i=1}^{n+1} E_i\}$. By Lemma 1, we only have to consider the case when $\tau_A = k$. Now if $\tau_A = k$, then it is easy to see that E_i and E_j are disjoint if $|i-j| < k$. Hence $\sum_{i=1}^{n+1} P(E_i) \geqslant P(\bigcup_{i=1}^{n+1} E_i) \geqslant \sum_{i=1}^{n+1} P(E_i) - \sum_{1 \leqslant i < j \leqslant n+1} P(E_i \cap E_j)$. Therefore, $\{1 - n2^{-k}\}(n+1)\, 2^{-n} \leqslant P\{\bigcup_{i=1}^{n+1} E_i\} \leqslant (n+1)\, 2^{-n}$, since $P(E_1) = 2^{-n}$ and $P(E_1 \cap E_j) \leqslant 2^{-n-k}$ for all $k+1 \leqslant j \leqslant n+1$.

LEMMA 3. *For any element A in Z_2^n, $((n+1)/2)\ 2^{-n} \leqslant P\{(X_1, X_2, ..., X_{2n})$ contains $A\} \leqslant (n+1)\, 2^{-n}$.*

Proof. Let $A_0 = (0, 0, ..., 0)$ be the unit element of Z_2^n. Then, by Lemma 1, $P\{(X_1, X_2, ..., X_{2n})$ contains $A\} \geqslant P\{(X_1, X_2, ..., X_{2n})$ contains $A_0\}$. Now it is easy to see that $P\{(X_1, X_2, ..., X_{2n})$ contains $A_0\} =$

$((n+1)/2)\,2^{-n}$. Therefore, $((n+1)/2)\,2^{-n} \leqslant P\{(X_1, X_2, ..., X_{2n})$ contains $A\} \leqslant (n+1)\,2^{-n}$ for any element A in Z_2^n.

LEMMA 4. *For any positive integer m and any element A in Z_2^n such that $\tau_A \geqslant k$. Then $P\{(X_1, X_2, ..., X_{(m+1)n})$ contains $A\} \geqslant m(n+1)\,2^{-n}\{1 - n2^{-k} - ((n+1)\,2^{-n})^{1/2} - \frac{1}{2}m(n+1)\,2^{-n}\}$.*

Proof. For each positive integer $i = 1, 2, ..., m$, let B_i be the event that B_i occurs if $(X_{(i-1)n+1}, X_{(i-1)n+2}, ..., X_{(i+1)n})$ contains A. It is easy to see that $P\{(X_1, X_2, ..., X_{(m+1)n})$ contains $A\} = P\{\bigcup_{i=1}^m B_i\} \geqslant \sum_{i=1}^m P(B_i) - \sum_{1 \leqslant i < j \leqslant m} P(B_i \cap B_j) = mP(B_1) - (m-1)\,P(B_1 \cap B_2) - \frac{1}{2}(m-1)(m-2) \times P^2(B_1)$, since $B_1, B_2, ..., B_m$ are exchangeable and B_i, B_j are mutually independent if $|i-j| > 1$. Now by the lemma of [5, p. 278] and Lemma 2, we have Lemma 4.

LEMMA 5. *For any positive integer m and any element A in Z_2^n. Then $P\{(X_1, X_2, ..., X_{(m+1)n})$ contains $A\} \geqslant \frac{1}{2}m(n+1)\,2^{-n}\{1 - 2((n+1)\,2^{-n})^{1/2} - m(n+1)\,2^{-n}\}$.*

Proof. Similar to the proof of Lemma 4; use Lemma 3 in the final substitution.

For each positive integer $k = 1, 2, ..., n$, let $n_k = \mathrm{card}\{A \mid A \in Z_2^n$ and $\tau_A = k\}$. It is easy to see that $n_k \leqslant 2^k$ for all $k = 1, 2, ..., n$.

LEMMA 6. $\sum_{n=1}^\infty P\{T_n > d2^n \ln(2^n)\} < \infty$ *if* $d > 1$.

Proof. $\sum_{n=1}^\infty P\{T_n > d2^n \ln(2^n)\} \leqslant \sum_{n=1}^\infty \sum_{k=1}^n P\{(X_1, X_2, ..., X_{d2^n\ln(2^n)})$ does not contain

$$A \mid \tau_A = k\} \leqslant \sum_{n=1}^\infty 2^k \Big\{1 - \frac{m}{2}(n+1)\,2^{-n}(1 - 2((n+1)\,2^{-n})^{1/2} - m(n+1)\,2^{-n}\Big\}^{[d2^n \ln(2)/(m+1)]}$$

$$+ \sum_{n=1}^\infty 2^n \Big\{1 - m(n+1)\,2^{-n}\Big(1 - n2^{-k} - ((n+1)\,2^{-n})^{1/2} - \frac{1}{2}m(n+1)\,2^{-n}\Big)\Big\}^{[d2^n \ln(2)/(m+1)]}.$$

It is easy to see that if $k \leqslant 2\ln(n)$, then

$$\sum_{n=1}^{\infty} 2^k \left\{1 - \frac{m}{2}(n+1)\,2^{-n}(1 - 2((n+1)\,2^{-n})^{1/2} - m(n+1)\,2^{-n})\right\}^{[d2^n \ln(2)/(m+1)]} < \infty$$

if $md > m+1$; it is possible since $d > 1$. Now since $n2^{-k} \to 0$ as $n \to \infty$ if $k \geqslant 2\ln(n)$, there exists an n_0 such that if $n \geqslant n_0$ and $m \leqslant n$, $n2^{-k} + ((n+1)\,2^{-n})^{1/2} + \frac{1}{2}m(n+1)\,2^{-n} < \varepsilon$, where $(1-\varepsilon)\,d > 1$. Hence

$$\sum_{n=1}^{\infty} 2^n \left\{1 - m(n+1)\,2^{-n}\left(1 - n2^{-k} - ((n+1)\,2^{-n})^{1/2} - \frac{1}{2}(n+1)\,2^{-n}\right)\right\}^{[d2^n \ln(2)/(m+1)]}$$

$$\leqslant 2^{n_0+1} + \sum_{n > n_0} 2^n \{1 - (1-\varepsilon)\,mn2^{-n}\}^{[d2^n \ln(2)/(m+1)]}$$

$$\approx 2^{n_0+1} + \sum_{n > n_0} 2^n e^{-[d(1-\varepsilon)\,mn \ln(2)/(m+1)]} < \infty$$

$$\text{if} \quad d(1-\varepsilon)\,m > m+1;$$

it is possible since $d(1-\varepsilon) > 1$. The proof of Lemma 6 now is complete.

Now we are in a position to state and prove our upper bound for the covering time C_n.

THEOREM 1. *$P\{C_n > d2^n \ln(2^n)$ only finitely often$\} = 1$ for any constant $d > 1$.*

Proof. Since $C_n = T_n - n$ for all $n = 1, 2, \ldots$, $\sum_{n=1}^{\infty} P\{C_n > d2^n \ln(2^n)\} \leqslant \sum_{n=1}^{\infty} P\{T_n > d2^n \ln(2^n)\} < \infty$ if $d > 1$. By the Borel–Cantelli lemma, we have $P\{C_n > d2^n \ln(2^n)$ only finitely often$\} = 1$ for any constant $d > 1$.

With respect to the fair coin tossing process $\{X_m\}_{m \geqslant 1}$, we define a new sequence $\{Y_m\}_{m \geqslant 1}$ of random variables as follows: For each positive integer $m \geqslant 1$, $Y_m = 0$ or 1 according to $(X_1, X_2, \ldots, X_{m+n-2})$ contains $(X_m, X_m + 1, \ldots, X_{m+n-1})$ or not. For each positive integer $n \geqslant 1$, let $S_{2^n} = \sum_{i=1}^{2^n} Y_i$. It is easy to see that $S_2n = \text{card}\{W_0^n, W_1^n, \ldots, W_{2^n-1}^n\}$ is the number of distinct states which the random walk W_i^n visited before the 2^nth step.

LEMMA 7. $\lim_{n \to \infty} E(S_2n)\,2^{-n} \geqslant (e-1)/e$.

Proof. To show that $\lim_{n\to\infty} E(S_2 n)\, 2^{-n} \geqslant (e-1)/e$, it suffices to show that $\lim_{n\to\infty} E(S_2 n)\, 2^{-n} \geqslant (e-1)/e-\varepsilon$ for any $\varepsilon>0$.

Let m be a fixed positive integer and $c=[2^n/(mn)]$ be the largest integer $\leqslant 2^n/(mn)$. Since $0\leqslant E(Y_i)\leqslant 1$ and is non-increasing in i, $mn\sum_{j=1}^{c} E(Y_{jmn+1})\leqslant E(S_2 n)\leqslant mn\sum_{j=0}^{c} E(Y_{jmn+1})$. Since $mn\{\sum_{j=0}^{c} E(Y_{jmn+1})-\sum_{j=1}^{c} E(Y_{jmn+1})\} = mnE(Y_1) = mn$, $\lim_{n\to\infty} 2^{-n}mn\sum_{j=1}^{c} E(Y_{jmn+1}) = \lim_{n\to\infty} 2^{-n}E(S_2 n)=\lim_{n\to\infty} 2^{-n}mn\sum_{j=0}^{c} E(Y_{jmn+1})$. Hence it is sufficient to show that $\lim_{n\to\infty} 2^{-n}mn\sum_{j=0}^{c} E(Y_{jmn+1})\geqslant (e-1)/e-\varepsilon$ for any $\varepsilon>0$.

By the definition of Y_j's, it is easy to see that $E(Y_{jmn+1})= P(Y_{jmn+1}=1)=\sum_{A\in Z_2^n} P\{(X_1, X_2, ..., X_{jmn+n-1})$ does not contain A and $(X_{jmn+1}, X_{jmn+2}, ..., X_{jmn+n})=A\} \geqslant \sum_{A\in Z_2^n} P\{(X_1, X_2, ..., X_{jmn})$ does not contain A and $(X_{jmn+1}, X_{jmn+2}, ..., X_{jmn+n})=A\}-n2^{-n}\geqslant \sum_{A\in Z_2^n} 2^{-n}\times P\{\bigcap_{i=1}^{j}[(X_{(i-1)mn+1}, X_{(i-1)mn+2}, ..., X_{imn})$ does not contain $A]\}-jn2^{-2}\geqslant (1-mn2^{-n})^j-jn2^{-n}$ for all $j=0, 1, 2, ..., c$. Hence $\sum_{j=0}^{c} E(Y_{jmn+1})\geqslant \sum_{j=0}^{c}\{(1-mn2^{-n})^j-jn2^{-n}\} = 2^n(mn)^{-1}\{1-(1-mn2^{-n})^{c+1}\}-n2^{-n}\{c(c+1)/2\}$. Therefore, $\lim_{n\to\infty} 2^{-n}mn\sum_{j=0}^{c} E(Y_{jmn+1})\geqslant \lim_{n\to\infty}\{\{1-(1-mn2^{-n})^{c+1}\}-(n/2)\,2^{-n}(2^n/mn+1)\}=(e-1)/e-1/2m$. Since m can be as large as possible, $\lim_{n\to\infty} 2^{-n}mn\sum_{j=0}^{c} E(Y_{jmn+1})\geqslant (e-1)/e-\varepsilon$ for any $\varepsilon>0$ and it completes the proof of Lemma 7.

LEMMA 8. $\lim_{n\to\infty} E(S_2 n)\, 2^{-n}\leqslant (e-1)/e$.

Proof. By a similar argument used in the proof of Lemma 7, it is sufficient to show that $\lim_{n\to\infty} 2^{-n}mn\sum_{j=0}^{c} E(Y_{jmn+1})\leqslant (e-1)/e+\varepsilon$ for any $\varepsilon>0$. Now $E(Y_{jmn+1})=P(Y_{jmn+1}=1)\leqslant \sum_{A\in Z_2^n} P\{(X_1, X_2, ..., X_{jmn})$ does not contain A and $(X_{jmn+1}, X_{jmn+1}, ..., X_{jmn+n})=A\}\leqslant \sum_{k=1}^{n} n_k 2^{-n} P\{\bigcap_{i=1}^{j}[(X_{(i-1)mn+1}, X_{(i-1)mn+2}, ..., X_{imn})$ does not contain $A\,|\,\tau_A=k\}=\sum_{k=1}^{n} 2^{-n} n_k\{P\{(X_1, X_2, ..., X_{mn})$ does not contain $A\}\}^j$. Now for sufficiently large n and $k\geqslant 2\ln(n)$, $P\{(X_1, X_2, ..., X_{mn})$ does not contain $A\,|\,\tau_A=k\}\leqslant (1-mn2^{-n}(1-\varepsilon))$. Since $n_i\leqslant 2^i$, $\sum_{i=1}^{k} n_i 2^{-n}\to 0$ as $n\to\infty$ if $k<2\ln(n)$. Hence, for sufficiently large n, $E(Y_{jmn+1})\leqslant (1-mn2^{-n}(1-\varepsilon))^j+\varepsilon$. Therefore, $2^{-n}mn\sum_{j=0}^{c} E(Y_{jmn+1})\leqslant 2^{-n}mn\sum_{j=0}^{c}(1-mn2^{-n}(1-\varepsilon))^j+\varepsilon=1-(1-mn2^{-n}(1-\varepsilon))^{c+1}+\varepsilon\to 1-e^{-1/(1-\varepsilon)}+\varepsilon$ as $n\to\infty$ and it completes the proof of Lemma 8.

For each positive integer $k=1, 2, ...,$ let $\mathscr{A}_k=\{W_t^n\,|\,(k-1)\,2^n\leqslant t<k2^n\}$, $\mathscr{B}_k=\bigcup_{j=1}^{k} A_j$, $\mathscr{D}_k=Z_2^n-\mathscr{B}_k$, and $E_k=\mathscr{A}_k-\mathscr{B}_{k-1}$.

THEOREM 2. *For all* $k=1, 2, ...,$

(i) $\lim_{n\to\infty} 2^{-n}E\{\text{card}(\mathscr{A}_k)\}=1-e^{-1}$,

(ii) $\lim_{n\to\infty} 2^{-n}E\{\text{card}(\mathscr{B}_k)\}=1-e^{-k}$,

(iii) $\lim_{n\to\infty} 2^{-n}E\{\text{card}(\mathscr{D}_k)\}=e^{-k}$.

Proof. By the fact that card($\mathscr{A}_k$) has the same distribution as of $S_2 n$ for all $k=1, 2, \ldots$. Now, by Lemmas 7 and 8, we have (i).

By the fact that W_t^n and $W_{t'}^n$ are independent if $|t-t'| \geqslant 2$ and (i), we have (ii).

By the fact that $\mathscr{D}_k \cap \mathscr{B}_k = \varnothing$, $Z_2^n = \mathscr{D}_k \cup \mathscr{B}_k$, and (ii), we have (iii).

In order to obtain the lower bound for the covering time C_n, we have to estimate the asymptotic upper bound for the variance of card($\mathscr{B}_k$) for all $k=1, 2, \ldots$. We start with the following lemmas.

For each pair (i, j) of positive integers, let $\varepsilon_{ij}=0$ or 1 according to $(X_i, X_{i+1}, \ldots, X_{i+n-1}) \neq (X_j, X_{j+1}, \ldots, X_{j+n-1})$ or $(X_i, X_{i+1}, \ldots, X_{i+n-1}) = (X_j, X_{j+1}, \ldots, X_{j+n-1})$. For each positive integer $N \geqslant n$, let $\xi(n, N) = \sum_{1 \leqslant i < j \leqslant N} \varepsilon_{ij}$ and for each positive integer n, let $t_n = \sup\{N \mid N \geqslant n$ and $\xi(n, N) = 0\}$. It is easy to see that $\xi(n, N)$ is the number of recurrences in $N+n-1$ trials and t_n is the number of trials before the first recurrence. The next lemma is a special case of Theorems 1 and 2 of [3].

LEMMA 9. *If $N \to \infty$ and n varies so that* (i) $\binom{N}{2} 2^{-n-1} \to \lambda > 0$ *and* (ii) $n^t N 2^{-n} \to 0$ *for all $t < \infty$. Then*

$$(1) \quad E\{Z^{\xi(n, N)}\} \to \exp\{\lambda(Z-1)/(1-\tfrac{1}{2}Z)\},$$

$$(2) \quad P\{t_n > x 2^{n/2}\} \to e^{-x^2}.$$

Proof. See pages 172–179 of [3].

For each positive integer $k=1, 2, \ldots$, we define a finite sequence $\{\tau_i^k \mid 1 \leqslant i \leqslant \text{card}(\mathscr{D}_k)\}$ (probably empty) of hitting times of $\mathscr{D}_k$ as follows: $\tau_1^k = \min\{t \mid W_t^n \in \mathscr{D}_k,\ k2^n \leqslant t < (k+1)\, 2^n\}, = \infty$ if no such t exists, and for each $j = 2, 3, \ldots, \text{card}(\mathscr{D}_k)$, $\tau_j^k = \min\{t \mid W_t^n \in \mathscr{D}_k,\ \tau_{j-1}^k < t < (k+1)\, 2^n\}, = \infty$ if no such t exists. Let $V_k = \{\tau_i^k \mid i = 1, 2, \ldots,$ and $\tau_i^k < \infty\}$. It is easy to see that $E_{k+1} = \{W_{\tau_i^k}^n \mid \tau_i^k \in V_k\}$.

If $E_{k+1} \neq \varnothing$, we define a finite sequence $\{Z_i^k \mid 1 \leqslant i \leqslant \text{card}(E_{k+1})\}$ of random variables as follows: $Z_1^k = 1$ and for each $i = 2, 3, \ldots, \text{card}(E_{k+1})$, $Z_i^k = 0$ or 1 according as $W_{\tau_i^k}^n \in \{W_{\tau_j^k}^n \mid 1 \leqslant j < i\}$ or $W_{\tau_i^k}^n \notin \{W_{\tau_j^k}^n \mid 1 \leqslant j < i\}$. It is easy to check that $S(E_{k+1}) = \sum_{i=1}^{\text{card}(E_{k+1})} Z_i^k = \sum_{i=k2^n+1}^{(k+1)\, 2^n} Y_i$ is the number of new states which the random walk W_t^n visited between the $(k2^n)$th step and the $((k+1)\, 2^n - 1)$th step.

LEMMA 10. $\text{Var}(S(E_{k+1})) \leqslant a n e^{-1}\, \text{card}(E_{k+1})$ *for some constant $a > 0$.*

Proof.

$$\begin{aligned}\operatorname{Var}(S(E_{k+1})) &= \operatorname{Var}\left(\sum_{i=1}^{\operatorname{card}(E_{k+1})} Z_i^k\right)\\ &= \sum_{i=1}^{\operatorname{card}(E_{k+1})} \operatorname{Var}(Z_i^k) + \sum_{i\neq j} \operatorname{Cov}(Z_i^k, Z_j^k)\\ &= \sum_{i=1}^{\operatorname{card}(E_{k+1})} \{P(Z_i^k=1) - P^2(Z_i^k=1)\}\\ &\quad + \sum_{i\neq j} \{P\{(Z_i^k=1)\cap(Z_j^k=1)\} - P\{Z_i^k=1\}\,P\{Z_j^k=1\}\}.\end{aligned}$$

Since $Z_1^k, Z_2^k, \ldots,$ are 0–1 random variables, $\operatorname{Var}(Z_j^k)\leqslant\frac{1}{4}$. Since the distribution $W_{\tau_i}^n$ is independent of $W_{\tau_j}^n$ if $|i-j|\geqslant n$, $P\{Z_j^k=1|Z_i^k=1\}\leqslant P\{Z_j^k=1|Z_i^k=0\}+n2^{-n}$ (by Lemma 9) as $n\to\infty$ and $j\geqslant i+n$. Hence $\sum_{i\neq j}\operatorname{Cov}(Z_i^k, Z_j^k)=\sum_{|i-j|<n}\operatorname{Cov}(Z_i^k, Z_j^k)+\sum_{|i-j|\geqslant n}\operatorname{Cov}(Z_i^k, Z_j^k)\leqslant (n/4)\operatorname{card}(E_{k+1}) + (n/n2^{-n})\operatorname{card}^2(E_{k+1})$. Since $\operatorname{card}(E_{k+1})\leqslant 2^n$, $\operatorname{Var}(S(E_{k+1}))\leqslant an\operatorname{card}(E_{k+1})$, for some constant $a>0$ and it completes the proof of Lemma 10.

Lemma 11. *$\lim_{n\to\infty} n^{-1}2^{-n}\operatorname{Var}\{\operatorname{card}(\mathscr{B}_k)\}\leqslant ae^{-k}$ for some constant $a>0$.*

Proof. We will prove Lemma 11 by induction on k. By Lemma 10, Lemma 11 holds when $k=1$. Now we assume that Lemma 11 holds for all $k=1, 2, \ldots, M$, and we will show that $\lim_{n\to\infty} 2^{-1}2^{-n}\operatorname{Var}\{\operatorname{card}(\mathscr{B}_{M+1})\}\leqslant ae^{-M-1}$. Since $\mathscr{B}_{M+1}=\mathscr{B}_{M+1}=\mathscr{B}_M\cup E_{M+1}$ and $\mathscr{B}_M\cap E_{m+1}=\varnothing$,

$$\begin{aligned}\operatorname{Var}(\operatorname{card}(\mathscr{B}_{M+1}))\} &= E\{(\operatorname{card}(\mathscr{B}_{M+1}) - E(\operatorname{card}(\mathscr{B}_{M+1})))^2\}\\ &= E\{[\operatorname{card}(\mathscr{B}_{M+1}) - E\{\operatorname{card}(\mathscr{B}_{M+1})|\operatorname{card}(\mathscr{B}_M)\}]^2\}\\ &\quad + E\{[E\{\operatorname{card}(\mathscr{B}_{M+1})|\operatorname{card}(\mathscr{B}_M)\} - E\{\operatorname{card}(\mathscr{B}_{M+1})\}]^2\}\\ &\approx e^{-2}\operatorname{Var}(\mathscr{B}_M) + E\{2^n - \operatorname{card}(\mathscr{B}_M)\}\cdot ane^{-1}\\ &\approx e^{-2}ane^{-M2^n} + 2^ne^{-M}ane^{-1} = an2^ne^{-M-1}(1+e^{-1}).\end{aligned}$$

Since $\sum_{i=0}^{\infty} e^{-i}=e/(e-1)$, by induction, we have $\lim_{n\to\infty} n^{-1}2^{-n}\operatorname{Var}\{\operatorname{card}(\mathscr{B}_k)\}\leqslant ae^{-k}$ for some constant $a>0$ and for all $k\geqslant 1$.

Lemma 12. *$\sum_{n=1}^{\infty} P\{T_n<c2^n\ln(2^n)\}<\infty$ for any $c<1$.*

Proof. $P\{T_n<c2^n\ln(2^n)\} = P\{\sum_{i=1}^{c2^n\ln(2^n)} Y_i=2^n\} = P\{\operatorname{card}(\mathscr{B}_{c\ln(2^n)})=2^n\}\approx P\{\operatorname{card}(\mathscr{B}_{c\ln(2^n)}) - E\{\operatorname{card}(\mathscr{B}_{c\ln(2^n)})\geqslant 2^n - 2^n(1-2^{-nc})\}\leqslant \operatorname{Var}\{\operatorname{card}(\mathscr{B}_{c\ln(2^n)})\}/2^{2n(1-c)}\approx an2^n2^{-2n(1-c)}e^{-c\ln(2^n)} = an2^{-n(1-c)}$. Hence $\sum_{n=1}^{\infty} P\{T_n<c2^n\ln(2^n)\}\approx\sum_{n=1}^{\infty} an2^{-n(1-c)}<\infty$ since $c<1$.

Now we are in the position to state and prove our lower bound for the covering time C_n.

THEOREM 3. $P\{C_n > c2^n \ln(2^n)$ *except finitely many* $n\} = 1$, *if* $c < 1$.

Proof. By Lemma 12 and the fact that $C_n = T_n - n$ for all $n \geq 1$, $\sum_{n=1}^{\infty} P\{C_n < c2^n \ln(2^n)\} < \infty$. By the Borel–Cantelli lemma, $P\{C_n < c2^n \ln(2^n)$ infinitely often$\} = 0$. Hence $P\{C_n > c2^n \ln(2^n)$ except finitely many $n\} = 1$.

Combining Theorem 1 and Theorem 3, we have the following theorems.

THEOREM 4. $P\{\lim_{n \to \infty} C_n/(2^n \ln(2^n)) = 1\} = 1$.

THEOREM 5. $\lim_{n \to \infty} E(C_n)/(2^n \ln(2^n)) = 1$.

THEOREM 6. $P\{\sum_{n=1}^{\infty} 2^n(1 - 2^{-n})^{C_n} = \infty\} = 1$.

ACKNOWLEDGMENTS

Our thanks to Michelle Wachs and Larry A. Shepp for their valuable comments.

REFERENCES

[1] GUIBAS, L. J., AND ODLYZKO, A. M. (1981). String overlaps, pattern matching, and nontransitive games, *J. Combin. Theory Ser. A* **30** 183–208.

[2] MATTHEWS, P. C., (1984). *Covering Problems for Random Walks on Spheres and Finite Groups*, Tech. Report 234, Department of Statistics, Stanford University.

[3] Mikhailov, V. G., and Zubkov, A. M. (1973). Limit distributions of random variables associated with long duplications in a sequence of independent trials, *Theory Probab. Appl.* **18** 172–179.

[4] SHEPP, L. A. (1972) Covering the circle with random arcs, *Israel J. Math.* **11** 328–345.

[5] SOLOVEV, A. D. (1966). A combinatorial identity and its application to the problem concerning the first occurrence of a rare event, *Theory Probab. Appl.* **11** 276–282.