Probability in the Engineering and Informational Sciences, **18**, 2004, 411–422. Printed in the U.S.A.

SOME STRONG LIMIT THEOREMS FOR MARKOV CHAIN FIELDS ON TREES

Wen Liu
Department of Mathematics
Hebei University of Technology
Tianjin 300130, China

Weiguo Yang
Faculty of Science
Jiangsu University
Zhenjiang 212013, China
E-mail: wgyang@ujs.edu.cn

In this article, we introduce the notion of the Markov chain fields on the generalized Bethe trees or generalized Cayley trees, and some strong limit theorems on the frequencies of states and ordered couples of states, including the Shannon–McMillan theorem on Bethe tree $T_{B,N}$ and Cayley tree $T_{C,N}$, are obtained. In the proof, a new technique in the study of the strong limit theorem in probability theory is applied.

1. INTRODUCTION

We begin with notations and definitions, which mainly follow from Spitzer [6] and Berger and Ye [2].

A tree is a graph $G = \{T, E\}$ which is connected and contains no circuits. Thus, G is a tree if and only if, given any two vertices $x \neq y \in T$, there exists an unique path $x = z_1, z_2, \ldots, z_m = y$ from x to y with $z_1, \ldots, z_m$ distinct. The distance between x and y is defined to be $m - 1$, the number of edges in the path connecting x and y.

To index the vertices on T, we first assign a vertex as the "root" and label it $\{0\}$. A vertex is said to be on the nth level if the path linking it to the root has n edges. The root $\{0\}$ is also said to be on the 0th level.

Definition 1: *Let T be a tree with root $\{0\}$, and let $\{N_n, h \geq 1\}$ be a sequence of positive integers. T is said to be a generalized Bethe tree or a generalized Cayley tree if each vertex on the nth level has N_{n+1} branches to the $(n+1)$st level.*

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

For example, when $N_1 = N + 1 \geq 2$ and $N_n = N (n \geq 2)$, T is a rooted Bethe tree $T_{B,N}$ on which each vertex has $N + 1$ neighboring vertices ($T_{B,2}$ drawn in Fig. 1), and when $N_n = N \geq 1 (n \geq 1)$, T is a rooted Cayley tree $T_{C,N}$ on which each vertex has N branches to the next level.

In the following, we always assume that T is a generalized Cayley tree and denote by $T^{(n)}$ the subgraph of T containing the vertices from level 0 (the root) to level n. We use $(n, j)(1 \leq j \leq N_1, \ldots, N_n, n \geq 1)$ to denote the jth vertex at the nth level and denote by $|B|$ the number of vertices in the subgraph B. It is easy to see that for $n \geq 1$,

$$|T^{(n)}| = 1 + \sum_{k=1}^{n} N_1, \ldots, N_k. \tag{1}$$

Let b be a positive integer, $S = \{1, 2, \ldots, b\}$, $\Omega = S^T$, $\omega = \omega(\cdot) \in \Omega$, where $\omega(\cdot)$ is a function defined on T and taking values in S, and $\mathbf{F}$ be the smallest Borel field containing all cylinder sets in Ω. Let $\mathbf{X} = \{X_t, t \in T\}$ be the coordinate stochastic process defined on the measurable space $(\Omega, \mathbf{F})$; that is, for any $\omega = \omega(\cdot) \in \Omega$, define

$$X_t(\omega) = \omega(t), \qquad t \in T. \tag{2}$$

Let μ be a probability measure on $(\Omega, \mathbf{F})$. Denote

$$X^{T^{(n)}} = \{X_t, t \in T^{(n)}\}, \qquad \mu(X^{T^{(n)}} = x^{T^{(n)}}) = \mu(x^{T^{(n)}}).$$

Now, we give a definition of Markov chain fields on the tree T by using the cylinder distribution directly, which is a natural extension of the classical definition of Markov chains (see Feller [3, p. 372]).

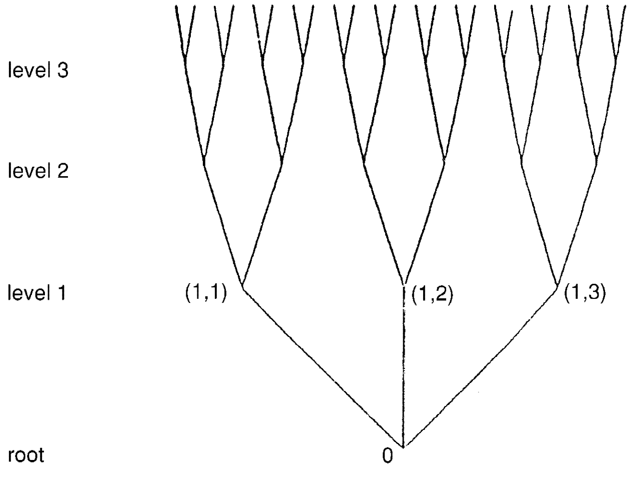

Figure 1. Bethe tree $T_{B,2}$.

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

DEFINITION 2: *Let $P = P(j|i)$ be a strictly positive stochastic matrix on S, $q = (q(1),\ldots,q(b))$ be a strictly positive distribution on S, and μ_P be a measure on $(\Omega,\mathbf{F})$. If*

$$\mu_P(x_0) = q(x_0), \qquad \mu_P(x^{T^{(1)}}) = q(x_0)\prod_{j=1}^{N_1} P(x_{1,j}|x_0), \tag{3}$$

$$\mu_P(x^{T^{(n)}}) = q(x_0)\prod_{j=1}^{N_1} P(x_{1,j}|x_0)\prod_{m=1}^{n-1}\prod_{i=1}^{N_1\ldots N_m}\prod_{j=N_{m+1}(i-1)+1}^{N_{m+1}i} P(x_{m+1,j}|x_{m,i}), \quad n \geq 2, \tag{4}$$

then μ_P will be called a Markov chain field on tree T determined by the stochastic matrix P and the distribution q.

Remark 1: μ_P given by (3) and (4) also depends on q. Hence, Definition 2 slightly extends the one given by Spitzer [6] and Berger and Ye [2], where q is taken to be the stationary distribution $\pi = (\pi(1),\ldots,\pi(b))$ determined by P.

Remark 2: If for all $n \geq 1, N_n = 1$ and $x_{n,1}$ is denoted by x_n, then we have by (3) and (4),

$$\mu_P(x^{T^{(n)}}) = \mu_P(X_0 = x_0,\ldots,X_n = x_n) = q(x_0)\prod_{m=1}^{n-1} P(x_{m+1}|x_m).$$

This is the cylinder distribution of Markov chains.

The tree model has drawn increasing interest from specialists in physics, probability, and information theory. Berger and Ye [2] have studied the existence of entropy rate for G-invariant random fields on trees. Recently, Ye and Berger [7] have also studied the ergodic property and the Shannon–McMillan theorem for PPG-invariant fields on trees. However, their main work is restricted to Bethe tree $T_{B,2}$ or Cayley tree $T_{C,2}$, and the convergence of the results is only the convergence in probability. Benjamini and Peres [1] have introduced the notions of the tree-indexed Markov chains and the tree-indexed random walk and have studied the recurrence and ray-recurrence for them.

In Section 3, we first prove a strong limit theorem on the frequencies of the ordered couples of states for the Markov chain fields on the generalized Cayley trees (or generalized Bethe trees), from which some strong limit theorems, including the Shannon–McMillan theorem with a.s. convergence, for the Markov chain fields on the Cayley tree $T_{C,N}$ or Bethe tree $T_{B,N}$ follow.

In the proof, a new technique in the study of a.s. convergence proposed in our previous works (see Liu and Yang [4,5]) is applied.

2. SOME LEMMAS

LEMMA 1: *Let μ_1 and μ_2 be two probability measures on the measurable space $(\Omega, \mathbf{F})$, and let $\{\tau_n, n \geq 1\}$ be a sequence of positive random variables such that*

$$\liminf_n \frac{\tau_n}{|T^{(n)}|} > 0, \qquad \mu_1 - a.s. \tag{5}$$

Then,

$$\limsup_n \frac{1}{\tau_n} \ln\left(\frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})}\right) \leq 0, \qquad \mu_1 - a.s. \tag{6}$$

PROOF: Let $Z_n = \mu_2(X^{T^{(n)}})/\mu_1(X^{T^{(n)}})$. It is easy to see that $E_{\mu_1}(Z_n) \leq 1$, where E_{μ_1} denotes the expectation under μ_1. Hence, for all $\varepsilon > 0$, we have by the Markov's inequality,

$$\sum_{n=1}^{\infty} \mu_1(|T^{(n)}|^{-1} \ln Z_n \geq \varepsilon) \leq \sum_{n=1}^{\infty} \exp(-|T^{(n)}|\varepsilon) < \infty. \tag{7}$$

Since $\varepsilon > 0$ is arbitrary, by the Borel–Cantelli lemma, it follows from (7) that

$$\limsup_n \frac{1}{|T^{(n)}|} \ln \frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})} \leq 0, \qquad \mu_1 - a.s. \tag{8}$$

Obviously, (5) and (8) imply (6). ■

Let $k, l \in S, S_n(k, \omega)$ be the number of k in $X^{T^{(n)}} = \{X_t, t \in T^{(n)}\}$, and $S_n(k, l, \omega)$ be the number of couple (k, l) in the couples of random variables

$$\{(X_0, X_{1,j}), \quad 1 \leq j \leq N_j, \quad (X_{m,i}, X_{m+1,j}), \quad 1 \leq m \leq n-1,$$
$$1 \leq i \leq N_1 \cdots N_m, \quad N_{m+1}(i-1) + 1 \leq j \leq N_{m+1} i, \, n \geq 2\};$$

that is,

$$S_n(k, \omega) = I_k(X_0) + \sum_{m=1}^{n} \sum_{j=1}^{N_1, \ldots, N_m} I_k(X_{m,j}); \tag{9}$$

$$S_n(k, l, \omega) = \sum_{j=1}^{N_1} I_k(X_0) I_l(X_{1,j}) + \sum_{m=1}^{n-1} \sum_{i=1}^{N_1, \ldots, N_m} \sum_{j=N_{m+1}(i-1)+1}^{N_{m+1} i} I_k(X_{m,i}) I_l(X_{m+1,j}), \tag{10}$$

where

$$I_k(x) = \begin{cases} 1, & x = k \\ 0, & x \neq k \end{cases} \qquad x, k \in S.$$

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

Let

$$\sigma_n(k,\omega) = \sum_{j=1}^{b} S_n(k,j,\omega). \tag{11}$$

It is easy to see that

$$\sum_{k=1}^{b} S_n(k,\omega) = |T^{(n)}|, \tag{12}$$

$$\sum_{i=1}^{b} S_n(i,k,\omega) = S_n(k,\omega) - I_k(X_0), \tag{13}$$

$$\sigma_n(k,\omega) = N_1 I_k(X_0) + \sum_{m=1}^{n-1} \sum_{j=1}^{N_1,\dots,N_m} N_{m+1} I_k(X_{m,j}), \tag{14}$$

$$\sum_{k=1}^{b} \sigma_n(k,\omega) = |T^{(n)}| - 1. \tag{15}$$

In the following, we always assume that μ_P is the Markov chain field on tree T determined by the stochastic matrix $P = (P(j|i))$ and the distribution q.

LEMMA 2: *For all $k \in S$, we have*

$$\liminf_n \frac{S_n(k,\omega)}{|T^{(n)}|} > 0, \qquad \mu_P - a.s. \tag{16}$$

PROOF: Let $0 < \lambda < 1$ be a constant, and $Q = (Q(j|i))$, $i,j \in S$, be another stochastic matrix, where for all $i \in S$,

$$Q(k|i) = \lambda, \qquad Q(j|i) = \frac{(1-\lambda)P(j|i)}{1-P(k|i)}, \qquad j \neq k. \tag{17}$$

Denote by μ_Q the Markov chain field on the tree T determined by Q and distribution q. Then,

$$\mu_Q(x^{T^{(n)}}) = q(x_0) \prod_{j=1}^{N_1} Q(x_{1,j}|x_0) \prod_{m=1}^{n-1} \prod_{i=1}^{N_1,\dots,N_m} \prod_{j=N_{m+1}(i-1)+1}^{N_{m+1}\,i} Q(x_{m+1,j}|x_{m,i}), \qquad n \geq 2. \tag{18}$$

Let

$$a_k = \min\{P(k|i), i \in S\}, \qquad b_k = \max\{P(k|i), i \in S\}. \tag{19}$$

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

By (4), (11), (15), and (17)–(19), we have

$$\begin{aligned}\frac{\mu_Q(X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})} &= \prod_{i=1}^{b}\prod_{j=1}^{b}\left[\frac{Q(j|i)}{P(j|i)}\right]^{S_n(i,j,\omega)}\\ &= \prod_{i=1}^{b}\left[\frac{\lambda}{P(k|i)}\right]^{S_n(i,k,\omega)}\left[\frac{1-\lambda}{1-P(k|i)}\right]^{\sigma_n(i,\omega)-S_n(i,k,\omega)}\\ &\geq \prod_{i=1}^{b}\left[\frac{\lambda}{b_k}\right]^{S_n(i,k,\omega)}\left[\frac{1-\lambda}{1-a_k}\right]^{\sigma_n(i,\omega)-S_n(i,k,\omega)}\\ &= \left(\frac{\lambda}{b_k}\right)^{S_n(k,\omega)-I_k(X_0)}\left(\frac{1-\lambda}{1-a_k}\right)^{|T^{(n)}|-1-S_n(k,\omega)+I_k(X_0)}. \qquad \textbf{(20)}\end{aligned}$$

Hence, by using Lemma 1, it follows from (20) that there exists $A(\lambda) \in \mathbf{F}$, $\mu_P(A(\lambda)) = 1$, such that

$$\limsup_n \frac{1}{|T^{(n)}|} S_n(k,\omega)\ln\left(\frac{\lambda(1-a_k)}{b_k(1-\lambda)}\right) \leq \ln\left(\frac{(1-a_k)}{(1-\lambda)}\right), \qquad \omega \in A(\lambda). \qquad \textbf{(21)}$$

Taking $\lambda \in (0, a_k)$ and noting that

$$0 < \frac{\lambda(1-a_k)}{b_k(1-\lambda)} < 1, \qquad 0 < \frac{1-a_k}{1-\lambda} < 1,$$

we have by (21),

$$\liminf_n \frac{1}{|T^{(n)}|} S_n(k,\omega) \geq \left[\ln\left(\frac{1-a_k}{1-\lambda}\right)\right]\left[\ln\left(\frac{\lambda(1-a_k)}{b_k(1-\lambda)}\right)\right]^{-1} > 0,$$
$$\omega \in A(\lambda). \qquad \textbf{(22)}$$

Hence, (16) holds. ■

LEMMA 3: *If there exist positive integers N_*, N, and d such that $N_* \leq N_n \leq N^*$ when $n \geq d$, then*

$$\liminf_n \frac{\sigma_n(k,\omega)}{|T^{(n)}|} \geq \frac{N_*}{N^*}\liminf_n\left(\frac{S_n(k,\omega)}{|T^{(n)}|}\right) > 0, \qquad \mu_P - a.s., \qquad \textbf{(23)}$$

$$\liminf_n \frac{S_n(k,\omega)}{|T^{(n)}|} \geq \frac{N_*}{N^*}\liminf_n \frac{\sigma_n(k,\omega)}{|T^{(n)}|}, \qquad \textbf{(24)}$$

$$\limsup_n \frac{\sigma_n(k,\omega)}{|T^{(n)}|} \leq \frac{N^*}{N_*}\limsup_n \frac{S_n(k,\omega)}{|T^{(n)}|}, \qquad \textbf{(25)}$$

$$\limsup_n \frac{S_n(k,\omega)}{|T^{(n)}|} \leq \frac{N^*}{N_*}\limsup_n \frac{\sigma_n(k,\omega)}{|T^{(n)}|}. \qquad \textbf{(26)}$$

PROOF: It is easy to see that there exist finite numbers a and b and finite random variables $\alpha(\omega)$ and $\beta(\omega)$ such that

$$a + N_* |T^{(n-1)}| \le |T^{(n)}| \le b + N^* |T^{(n-1)}|, \tag{27}$$

$$\alpha(\omega) + N_* S_{n-1}(k,\omega) \le \sigma_n(k,\omega) \le N^* S_{n-1}(k,\omega) + \beta(\omega). \tag{28}$$

Hence,

$$\frac{\sigma_n(k,\omega)}{|T^{(n)}|} \ge \frac{\alpha(\omega) + N_* S_{n-1}(k,\omega)}{b + N^* |T^{(n-1)}|}.$$

This together with Lemma 2 implies (23) evidently. In a similar way, we can verify (24)–(26) by using inequalities (27) and (28). ■

LEMMA 4: *Let* $0 < p < 1$ *and* $\{c_n, n \ge 1\}$ *be a sequence of nonnegative real numbers. If there exists a sequence of real numbers* $\{\alpha_k, k \ge 1\}$ *such that* $0 < \alpha_k < p, \alpha_k \to p$, *and*

$$\limsup_n \left(\frac{\alpha_k}{p}\right)^{c_n} \left(\frac{1-\alpha_k}{1-p}\right)^{1-c_n} \le 1, \tag{29}$$

then

$$\liminf_n c_n \ge p; \tag{30}$$

if there exists a sequence of real numbers $\{\beta_k, k \ge 1\}$ *such that* $p < \beta_k < 1, \beta_k \to p$, *and*

$$\limsup_n \left(\frac{\beta_k}{p}\right)^{c_n} \left(\frac{1-\beta_k}{1-p}\right)^{1-c_n} \le 1,$$

then

$$\limsup_n c_n \le p.$$

PROOF: By (29),

$$\liminf_n c_n \ln\left(\frac{\alpha_k(1-p)}{p(1-\alpha_k)}\right) \le \ln\left(\frac{1-p}{1-\alpha_k}\right).$$

Hence,

$$\liminf_n c_n \ge \left[\ln\left(\frac{1-p}{1-\alpha_k}\right)\right]\left[\ln\left(\frac{\alpha_k(1-p)}{p(1-\alpha_k)}\right)\right]^{-1} \tag{31}$$

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

since

$$0 < \frac{\alpha_k(1-p)}{p(1-\alpha_k)} < 1, \qquad 0 < \frac{1-p}{1-\alpha_k} < 1.$$

It is easy to see that

$$\lim_k \left[\ln\left(\frac{1-p}{1-\alpha_k}\right)\right]\left[\ln\left(\frac{\alpha_k(1-p)}{p(1-\alpha_k)}\right)\right]^{-1} = p. \tag{32}$$

Inequality (30) follows from (31) and (32) directly. In a similar way, we can verify the second part of the lemma. ■

3. MAIN RESULTS

THEOREM 1: *If there exist positive integers N_*, N^*, and d such that $N_* \le N_n \le N^*$ when $n \ge d$, then for all $k, l \in S$,*

$$\lim_n \frac{S_n(k,l,\omega)}{\sigma_n(k,\omega)} = P(l|k), \qquad \mu_P - a.s. \tag{33}$$

PROOF: Let $0 < \lambda < 1$ be a constant and $D = (D(j|i))$, $i, j \in S$, be another stochastic matrix, where

$$D(l|k) = \lambda, \qquad D(j|k) = \frac{(1-\lambda)P(j|k)}{1-P(l|k)}, \qquad j \ne l,$$
$$D(j|i) = P(j|i), \qquad i \ne k,\ j \in S.$$

Denote by μ_D the Markov chain field on tree T determined by D and the distribution q. Then,

$$\begin{aligned}\frac{\mu_D(X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})} &= \prod_{i=1}^{b}\prod_{j=1}^{b}\left[\frac{D(j|i)}{P(j|i)}\right]^{S_n(i,j,\omega)} \\ &= \prod_{j=1}^{b}\left[\frac{D(j|k)}{P(j|k)}\right]^{S_n(k,j,\omega)} \\ &= \left[\frac{\lambda}{P(l|k)}\right]^{S_n(k,l,\omega)}\left[\frac{1-\lambda}{1-P(l|k)}\right]^{\sigma_n(k,\omega)-S_n(k,l,\omega)}.\end{aligned} \tag{34}$$

By (23) and Lemma 1, there exist $A(k,l,\lambda) \in \mathbf{F}$ and $\mu_P(A(k,l,\lambda)) = 1$, such that

$$\limsup_n \left[\frac{\mu_D(X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})}\right]^{1/\sigma_n(k,\omega)} \le 1, \qquad \omega \in A(k,l,\lambda). \tag{35}$$

Take $\alpha_i \in (0, P(l|k))$ and $\beta_i \in (P(l|k), 1)$, $i = 1,2,\ldots$, such that $\alpha_i \to P(l/k)$, $\beta_i \to P(l/k)(i \to \infty)$. Let $A_*(k,l) = \cap_{i=1}^{\infty} A(k,l,\alpha_i)$. Then, by (34) and (35), we have for all $i \ge 1$,

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

$$\limsup_n \left[\frac{\alpha_i}{P(l|k)}\right]^{S_n(k,l,\omega)/\sigma_n(k,\omega)} \left[\frac{1-\alpha_i}{1-P(l|k)}\right]^{1-S_n(k,l,\omega)/\sigma_n(k,\omega)} \leq 1,$$
$$\omega \in A_*(k,l). \qquad \textbf{(36)}$$

By Lemma 4, it follows from (36) that

$$\liminf_n \frac{S_n(k,l,\omega)}{\sigma_n(k,\omega)} \geq P(l|k), \qquad \omega \in A_*(k,l). \qquad \textbf{(37)}$$

Let $A^*(k,l) = \cap_{i=1}^{\infty} A(k,l,\beta_i)$. In a similar way, it can be shown that

$$\limsup_n \frac{S_n(k,l,\omega)}{\sigma_n(k,\omega)} \leq P(l|k), \qquad \omega \in A^*(k,l). \qquad \textbf{(38)}$$

The theorem follows because $\mu_P(A_*(k,l) \cap \mu_P(A^*(k,l)) = 1$. ■

COROLLARY 1: *Under the conditions of Theorem 1, we have*

$$\liminf_n \frac{S_n(k,l,\omega)}{S_{n-1}(k,\omega)} \geq N_* P(l|k), \qquad \mu_P - a.s.,$$

$$\limsup_n \frac{S_n(k,l,\omega)}{S_{n-1}(k,\omega)} \leq N^* P(l|k), \qquad \mu_P - a.s.$$

In particular, if T is a Bethe tree $T_{B,N}$ or a Cayley tree $T_{C,N}$, then

$$\lim_n \frac{S_n(k,l,\omega)}{S_{n-1}(k,\omega)} = NP(l|k), \qquad \mu_P - a.s. \qquad \textbf{(39)}$$

PROOF: The corollary follows from Theorem 1 and (28) directly. ■

THEOREM 2: *If T is a Bethe tree $T_{B,N}$ or a Cayley tree $T_{C,N}$, then for all $k \in S$,*

$$\lim_n \frac{S_n(k,\omega)}{|T^{(n)}|} = \pi(k), \qquad \mu_P - a.s., \qquad \textbf{(40)}$$

$$\lim_n \frac{\sigma_n(k,\omega)}{|T^{(n)}|} = \pi(k), \qquad \mu_P - a.s., \qquad \textbf{(41)}$$

where $\pi = (\pi(1),\ldots,\pi(b))$ is the stationary distribution determined by P.

PROOF: Let

$$H(i,j) = \left\{\omega : \lim_n \frac{S_n(i,j,\omega)}{S_{n-1}(i,\omega)} = NP(j|i)\right\},$$

$$H = \bigcap_{i,j=1}^{b} H(i,j).$$

By (39), $\mu_P(H) = 1$. Let $\omega \in H$. Then,

$$S_n(j,k,\omega) - S_{n-1}(j,\omega)P(k|j) = \alpha_n(j,k,\omega)S_{n-1}(j,\omega),$$

where $\alpha_n(j,k,\omega) \to 0 (n \to \infty)$. Adding the b equalities for $j = 1,2,\ldots,b$ and using (13), we have

$$S_n(k,\omega) - \sum_{j=1}^{b} NS_{n-1}(j,\omega)P(k|j) = \sum_{j=1}^{b} \alpha_n(j,k,\omega)S_{n-1}(j,\omega) + I_k(X_0). \tag{42}$$

It follows from (42) and (12) that

$$\lim_n \left[\frac{S_n(k,\omega)}{|T^{(n)}|} - \frac{1}{|T^{(n-1)}|} \sum_{j=1}^{b} S_{n-1}(j,\omega)P(k|j) \right] = 0, \qquad \omega \in H. \tag{43}$$

Multiplying the kth equality of (43) by $P(i|k)(k = 1,2,\ldots,b)$, adding them together, and using (43) once again, we obtain

$$\begin{aligned}
\lim_n \Bigg\{ & \frac{1}{|T^{(n)}|} \sum_{k=1}^{b} S_n(k,\omega)P(i|k) - \frac{S_{n+1}(i,\omega)}{|T^{(n+1)}|} \\
& + \left[\frac{S_{n+1}(i,\omega)}{|T^{(n+1)}|} - \frac{1}{|T^{(n-1)}|} \sum_{k=1}^{b}\sum_{j=1}^{b} S_{n-1}(j,\omega)P(k|j)P(i|k) \right] \Bigg\} \\
= \lim_n & \left[\frac{S_{n+1}(i,\omega)}{|T^{(n+1)}|} - \frac{1}{|T^{(n-1)}|} \sum_{j=1}^{b} S_{n-1}(j,\omega)P^{(2)}(i|j) \right] = 0, \qquad \omega \in H,
\end{aligned}$$

where $P^{(h)}(i|j)$ (h is a positive integer) is the hth-order transition probability determined by stochastic matrix $(P(j|i))$. By induction, we have

$$\lim_n \left\{ \frac{1}{|T^{(n+h)}|} S_{n+h}(i,\omega) - \frac{1}{|T^{(n-1)}|} \sum_{j=1}^{b} S_{n-1}(j,\omega)P^{(h+1)}(i|j) \right\} = 0, \qquad \omega \in H. \tag{44}$$

Let

$$\alpha_h(i) = \min\{P^{(h+1)}(i|j), j \in S\}, \qquad \beta_h(i) = \max\{P^{(h+1)}(i|j), j \in S\}.$$

By (44) and (12), we have

$$\limsup_n \frac{1}{|T^{(n+h)}|} S_{n+h}(i,\omega) \le \beta_h(i), \qquad \omega \in H, \tag{45}$$

$$\liminf_n \frac{1}{|T^{(n+h)}|} S_{n+h}(i,\omega) \ge \alpha_h(i), \qquad \omega \in H, \tag{46}$$

Since $\lim_h P^{(h+1)}(i|j) = \pi(i)$,

$$\lim_h \alpha_h(i) = \lim_h \beta_h(i) = \pi(i). \tag{47}$$

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

Equation (40) follows from (45)–(47), and (41) follows from (40) and (23)–(26) directly. ■

COROLLARY 2: *Under the conditions of Theorem 2, we have*

$$\lim_n \frac{S_n(k,l,\omega)}{|T^{(n)}|} = \pi(k)P(l|k), \qquad \mu_P - a.s., \tag{48}$$

$$\lim_n \frac{S_n(k,l,\omega)}{S_n(k,\omega)} = P(l|k), \qquad \mu_P - a.s. \tag{49}$$

PROOF: Equation (48) follows from (33) and (41), and (49) follows from (48) and (40). ■

THEOREM 3: *Let T be a Bethe tree $T_{B,N}$ or a Cayley tree $T_{C,N}$, and $f(x,y)$ be a function defined on S^2. Set*

$$Y_n(\omega) = \sum_{j=1}^{N_1} f(X_0, X_{1,j}) + \sum_{m=1}^{n-1} \sum_{i=1}^{N_1,\ldots,N_m} \sum_{j=N_{m+1}(i-1)+1}^{N_{m+1}i} f(X_{m,i}, X_{m+1,j}). \tag{50}$$

Then,

$$\lim_n \frac{Y_n(\omega)}{|T^{(n)}|} = \sum_{k=1}^{b} \sum_{l=1}^{b} \pi(k)P(l|k)f(k,l), \qquad \mu_P - a.s. \tag{51}$$

PROOF: By (50) and (10), we have

$$\begin{aligned} Y_n(\omega) &= \sum_{j=1}^{N_1} \sum_{k=1}^{b} \sum_{l=1}^{b} f(k,l) I_k(X_0) I_l(X_{1,j}) \\ &\quad + \sum_{m=1}^{n-1} \sum_{i=1}^{N_1,\ldots,N_m} \sum_{j=N_{m+1}(i-1)+1}^{N_{m+1}i} \sum_{k=1}^{b} \sum_{l=1}^{b} f(k,l) I_k(X_{m,i}) I_l(X_{m+1,j}) \\ &= \sum_{k=1}^{b} \sum_{l=1}^{b} f(k,l) S_n(k,l,\omega). \end{aligned} \tag{52}$$

The theorem follows from (52) and (48) directly. ■

Let μ be a probability measure on $(\Omega, \mathbf{F})$ and let

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \ln \mu(X^{T^{(n)}}).$$

https://doi.org/10.1017/S0269964804183083 Published online by Cambridge University Press

$f_n(\omega)$ is called the entropy density on subgraph $T^{(n)}$ with respect to μ. If $\mu = \mu_P$, then by (4) we have

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}\Bigg[\ln q(X_0) + \sum_{j=1}^{N_1} \ln P(X_{1,j}|X_0) + \sum_{m=1}^{n-1} \sum_{i=1}^{N_1,\dots,N_m} \sum_{j=N_{m+1}(i-1)+1}^{N_{m+1}i} \ln P(X_{m+1,j}|X_{m,i})\Bigg]. \quad \textbf{(53)}$$

The convergence of $f_n(\omega)$ to a constant in a sense (L_1 convergence, convergence in probability, or a.s. convergence) is called the Shannon–McMillan theorem or the asymptotic equipartition property (AEP) in formation theory. By using Theorem 3, we can easily obtain the Shannon–McMillan theorem for Markov chain fields on the Bethe tree $T_{B,N}$ and the Cayley tree $T_{C,N}$ with a.s. convergence.

THEOREM 4: *Let μ_P be a Markov chain field on Bethe tree $T_{B,N}$ or the Cayley tree $T_{C,N}$, and $f_n(\omega)$ be defined by (53). Then,*

$$\lim_n f_n(\omega) = -\sum_{k=1}^{b} \sum_{l=1}^{b} \pi(k)P(l|k)\ln P(l|k), \qquad \mu_P - a.s. \quad \textbf{(54)}$$

PROOF: Letting $f(x,y) = -\ln P(y|x)$ in Theorem 3, the proof follows from (51) directly. ■

Remark: As we have mentioned in Section 1, Ye and Berger have studied the Shannon–McMillan theorem for PPG-invariant random field on trees, but the convergence in their results is only the convergence in probability. They conjectured that these results also hold with a.s. convergence. Since the Markov chain field is a particular case of the PPG-invariant random fields, Theorem 4 partly solved the conjecture of Ye and Berger.

References

1. Benjamini, I. & Peres, Y. (1994). Markov chains indexed by trees. *Annals of Probability* 22: 219–243.
2. Berger, T. & Ye, Z. (1990). Entropic aspects of random fields on trees. *IEEE Transactions on Information Theory* 36(5): 1006–1018.
3. Feller, W. (1968). An introduction to probability theory and its applications, Vol. 1, 3rd ed. New York: Wiley.
4. Liu, W. & Yang, W. G. (1995). A limit theorem for the entropy density of nonhomogeneous Markov information source. *Statistics and Probability Letters* 22: 295–301.
5. Liu, W. & Yang, W. G. (1996). An extension of Shannon–McMillan theorem and some limit properties for nonhomogeneous Markov chains. *Stochastic Processes and Their Applications* 61: 129–146.
6. Spitzer, F. (1975). Markov random fields on an infinite tree. *Annals of Probability* 3: 387–398.
7. Ye, Z. & Berger, T. (1996). Ergodic, regularity and asymptotic equipartition property of random fields on trees. *Journal of Combinatorics, Information and System Sciences* 21(2): 157–184.