Journal of Inequalities and Applications
a SpringerOpen Journal

A class of small deviation theorems for random fields on a uniformly bounded tree

Shuai Wang and Weiguo Yang[*]

[*]Correspondence:
wgyang@ujs.edu.cn
Faculty of Science, Jiangsu
University, Zhenjiang, Jiangsu
212013, China

Abstract

By introducing the asymptotic logarithmic likelihood ratio as a measure of the Markov approximation of arbitrary random fields on a uniformly bounded tree, by constructing a non-negative martingale on a uniformly bounded tree, a class of small deviation theorems of functionals and a class of small deviation theorems of the frequencies of occurrence of states for random fields on a uniformly bounded tree are established. Some known results are generalized in this paper.

MSC: 60J10; 60F15

Keywords: deviation theorems; uniformly bounded tree; random field

1 Introduction

A tree is a graph $G = \{T, E\}$ which is connected and contains no circuits. Given any two vertices $\alpha \neq \beta \in T$, let $\overline{\alpha\beta}$ be the unique path connecting α and β. Define the graph distance $d(\alpha, \beta)$ to be the number of edges contained in the path $\overline{\alpha\beta}$.

Let T be an infinite tree which is locally finite and has no leaves. We choose a vertex as the root. We call the number of neighbors of a vertex of the tree the degree of this vertex. When the degrees of any vertices of T are uniformly bounded, we say T is a uniformly bounded tree. A tree is said to be a Bethe tree if each vertex has $N + 1$ neighboring vertices, which is denoted by $T_{B,N}$. A tree is said to be a Cayley tree if the root has only N neighbors and the other vertices have $N + 1$ neighbors, which is denoted by $T_{C,N}$. Both kinds of trees are common homogeneous trees. Obviously, the two kinds of homogeneous trees are the special cases of the uniformly bounded tree. When the context permits, uniformly bounded trees are all denoted simply by T.

Let T be an infinite tree with root o. The set of all vertices with distance n from the root is called the nth generation of T, which is denoted by L_n. We denote by $T^{(n)}$ the subtree comprised of level 0 (the root o) through level n. For each vertex t, there is a unique path from o to t, and $|t|$ for the number of edges on this path. We denote the first predecessor of t by 1_t, the second predecessor of t by 2_t, and denote by n_t the nth predecessor of t. For any two vertices s and t of tree T, write $s \leq t$ if s is on the unique path from the root o to t. We denote by $s \wedge t$ the vertex farthest from o satisfying $s \wedge t \leq s$ and $s \wedge t \leq t$.

Let $(\Omega, \mathcal{F})$ be a measurable space, $\{X_t, t \in T\}$ be a collection of random variables defined on $(\Omega, \mathcal{F})$ and taking values in $S = \{0, 1, \ldots, b-1\}$, where b is a positive integer. Let A be a subgraph of T, $X^A = \{X_t, t \in A\}$, and we denote by $|A|$ the number of vertices of A. Let the realization of $X^{T^{(n)}}$ be $x^{T^{(n)}}$. Let μ be a probability measure on the measurable space $(\Omega, \mathcal{F})$.

We will call μ the random field on tree T. Let the distribution of $\{X_t, t \in T\}$ under probability measure μ be $\mu(x^{T^{(n)}}) = \mu(X^{T^{(n)}} = x^{T^{(n)}})$. $\mu(x^{T^{(n)}})$ is actually the marginal distribution of μ.

Definition 1 (see [1]) Let T be an infinite tree which is locally finite and has no leaves, S be a finite state space, $\{X_t, t \in T\}$ be a collection of S-valued random variables defined on the measurable space $(\Omega, \mathcal{F})$, and let

$$p = \{p(x), x \in S\} \tag{1}$$

be a distribution on S,

$$P = (P(y|x)), \quad x, y \in S, \tag{2}$$

be a stochastic matrix on S^2. Let μ_P be a probability measure on the measurable space $(\Omega, \mathcal{F})$. If for any vertex $t \in T$,

$$\mu_P(X_t = y | X_{1_t} = x \text{ and } X_s \text{ for } s \wedge t \leq 1_t) = \mu_P(X_t = y | X_{1_t} = x) = P(y|x), \quad x, y \in S, \tag{3}$$

and

$$\mu_P(X_0 = x) = p(x), \quad x \in S,$$

$\{X_t, t \in T\}$ will be called S-valued Markov chains indexed by an infinite tree T with initial distribution (1) and transition matrix (2) under probability measure μ_P.

Let $\{X_t, t \in T\}$ be Markov chains indexed by tree T under probability measure μ_P defined above. It is easy to see that

$$\mu_P(x^{T^{(n)}}) = \mu_P(X^{T^{(n)}} = x^{T^{(n)}}) = p(x_o) \prod_{m=1}^{n} \prod_{t \in L_m} P(x_t | x_{1_t}). \tag{4}$$

In order to avoid technical problems, we always assume that $\mu(x^{T^{(n)}})$, $P(y|x)$ and $p(x)$ are positive.

Definition 2 Let T be a uniformly bounded tree which has no leaves and $\{X_t, t \in T\}$ be a collection of S-valued random variables defined on $(\Omega, \mathcal{F})$, $P = (P(y|x))$, $x, y \in S$ be a positive stochastic matrix, μ, μ_P be two probability measures on $(\Omega, \mathcal{F})$, and $\{X_t, t \in T\}$ be Markov chains indexed by tree T under probability measure μ_P determined by P. Assume that $\mu(x^{T^{(n)}})$ is always strictly positive. Let

$$\varphi_n(\omega) = \frac{\mu(X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})}, \tag{5}$$

$$\varphi(\omega) = \limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \ln \varphi_n(\omega), \tag{6}$$

$\varphi(\omega)$ will be called the asymptotic logarithmic likelihood ratio.

Remark 1 If $\mu = \mu_P$, then $\varphi(\omega) \equiv 0$. In Lemma 1 we will show that in a general case $\varphi(\omega) \geq 0$ μ-a.e., hence $\varphi(\omega)$ can be regarded as a measure of the Markov approximation of an arbitrary random field on T.

The tree model has recently drawn increasing interest from specialists in physics, probability and information theory. Benjamini and Peres [1] have given the notion of the tree-indexed homogeneous Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye [2] have studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Ye and Berger [3] have studied the asymptotic equipartition property (AEP) in the sense of convergence in probability for a PPG-invariant and ergodic random field on a homogeneous tree. Recently, Yang [4] has studied some strong limit theorems for countable homogeneous Markov chains indexed by a homogeneous tree and the strong law of large numbers and the asymptotic equipartition property (AEP) for finite homogeneous Markov chains indexed by a homogeneous tree. Huang and Yang [5] have studied the strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree. Liu and Wang [6] have studied the small deviation theorems between the arbitrary random fields and the Markov chain fields on the Cayley tree. Peng, Yang, and Wang [7] have further studied a class of small deviation theorems for functionals of random fields on a homogeneous tree which partially extend the result of [6].

In this paper, by introducing the asymptotic logarithmic likelihood ratio as a measure of the Markov approximation of the arbitrary random field on a uniformly bounded tree, and by constructing a non-negative martingale, we obtain the following two results: a class of small deviation theorems of functionals and a class of small deviation theorems of the frequencies of occurrence of states for random fields on a uniformly bounded tree. In fact, our present outcomes can imply the case in [5] and [7].

Lemma 1 *Let T be a uniformly bounded tree which has no leaves. Let μ_1, μ_2 be two probability measures on $(\Omega, \mathcal{F})$, $D \in \mathcal{F}$, $\{\tau_n, n \geq 1\}$ be a sequence of positive random variables such that*

$$\liminf_{n \to \infty} \frac{\tau_n}{|T^{(n)}|} > 0 \quad \mu_1\text{-a.e. on } D. \tag{7}$$

Then

$$\limsup_{n \to \infty} \frac{1}{\tau_n} \ln \frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})} \leq 0 \quad \mu_1\text{-a.e. on } D. \tag{8}$$

Proof The proof of this lemma is similar to that of Lemma 1 of [6], so we omit it. □

Remark 2 Let $\mu_1 = \mu$, $\mu_2 = \mu_P$ and $\tau_n = |T^{(n)}|$ in Lemma 1, by (8) there exists $A \in \mathcal{F}$, $\mu(A) = 1$ such that

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \ln \frac{\mu_P(X^{T^{(n)}})}{\mu(X^{T^{(n)}})} \leq 0, \quad \omega \in A, \tag{9}$$

hence we have $\varphi(\omega) \geq 0$, $\omega \in A$.

From this remark, we know that $D \in \mathcal{F}$ and the sequence of $\{\tau_n, n \geq 1\}$ are existent. Let T be a uniformly bounded tree, $k, l \in S$. Let

$$S_n(k) = \left|\left\{t \in T^{(n)} : X_t = k\right\}\right|, \tag{10}$$

$$S_n(k, l) = \left|\left\{t \in \left(T^{(n)} \setminus \{0\}\right) : (X_{1_t}, X_t) = (k, l)\right\}\right|, \tag{11}$$

that is,

$$S_n(k) = \sum_{t \in T^{(n)}} \delta_k(X_t), \qquad S_n(k, l) = \sum_{m=1}^{n} \sum_{t \in L_m} \delta_k(X_{1_t}) \delta_l(X_t),$$

where $\delta_k(\cdot)$ $(k \in S)$ is the Kronecker δ-function

$$\delta_k(x) = \begin{cases} 1, & \text{if } x = k, \\ 0, & \text{if } x \neq k. \end{cases}$$

Let the degree of each vertex σ $(\sigma \neq o)$ on the tree T be $d(\sigma)$. Since T is a uniformly bounded tree which has no leaves, we know that there are two positive numbers m and M such that $2 \leq m \leq d(\sigma) \leq M$.

Lemma 2 *Let T be a uniformly bounded tree which has no leaves, $P = (P(y|x))$, $x, y \in S$ be a positive stochastic matrix, μ, μ_P be two probability measures on $(\Omega, \mathcal{F})$, $\{X_t, t \in T\}$ be Markov chains indexed by T under probability measure μ_P determined by P, $\varphi(\omega)$ be denoted by (6), M be defined as above, $0 \leq c < \ln(1 - a_k)^{-1}$ be a constant, and*

$$D(c) = \left\{\omega : \varphi(\omega) \leq c\right\}, \tag{12}$$

$$M_k = \max\left\{\left[\ln \frac{1 - a_k}{1 - \lambda} + c\right] \bigg/ \ln \frac{\lambda(1 - a_k)}{b_k(1 - \lambda)}, 0 < \lambda \leq 1 + (a_k - 1)e^c\right\}, \tag{13}$$

where $a_k = \min\{P(k|i), i \in S\}$, $b_k = \max\{P(k|i), i \in S\}$. Then

$$\liminf_{n \to \infty} \frac{S_{n-1}(k)}{|T^{(n)}|} \geq \frac{M_k}{M - 1} \quad \mu\text{-a.e. on } D(c). \tag{14}$$

Proof By using a similar proof as that of Lemma of [6], we can obtain

$$\liminf_{n \to \infty} \frac{S_n(k)}{|T^{(n)}|} \geq M_k \quad \mu\text{-a.e. on } D(c). \tag{15}$$

Since $\frac{1}{|T^{(n+1)}|} \geq \frac{1}{M-1} \frac{1}{|T^{(n)}|}$, this corollary follows from (15). □

In the following, we always let $N \geq 0$, $k \in S$, $d^0(t) = 1$, and denote by $d^N(t) = |\tau \in T : N_\tau = t|$, where N_τ is defined as above. Let

$$S_k^N(A) = \sum_{t \in A} \delta_k(X_t) d^N(t), \tag{16}$$

$$S_{k,l}^N(A) = \sum_{t \in A} \delta_k(X_{1_t}) \delta_l(X_t) d^N(t). \tag{17}$$

Corollary 1 *Let m, M be defined as above. Under the assumption of Lemma 2, we have*

$$\liminf_{n \to \infty} \frac{S_k^{N+1}(T^{(n-1)})}{|T^{(n)}|} \geq \frac{M_k(m-1)^{N+1}}{M-1} \quad \mu\text{-a.e. on } D(c). \tag{18}$$

Proof Because T is a uniformly bounded tree and has no leaves, then $(m-1)^N \leq d^N(t) \leq (M-1)^N$. By Lemma 2 we have

$$
\begin{aligned}
\liminf_{n \to \infty} \frac{S_k^{N+1}(T^{(n-1)})}{|T^{(n)}|} &= \liminf_{n \to \infty} \frac{\sum_{t \in T^{(n-1)}} \delta_k(X_t) d^{N+1}(t)}{|T^{(n)}|} \\
&\geq (m-1)^{N+1} \liminf_{n \to \infty} \frac{\sum_{t \in T^{(n-1)}} \delta_k(X_t)}{|T^{(n)}|} \\
&= (m-1)^{N+1} \liminf_{n \to \infty} \frac{S_{n-1}(k)}{|T^{(n)}|} \\
&\geq \frac{M_k(m-1)^{N+1}}{M-1} \quad \mu\text{-a.e. on } D(c).
\end{aligned}
\tag{19}
$$

The proof is finished. $\qquad\square$

Lemma 3 *Let T be a uniformly bounded tree which has no leaves, $P = (P(y|x))$, $x, y \in S$ be a positive stochastic matrix, μ, μ_P be two probability measures on $(\Omega, \mathcal{F})$, $\{X_t, t \in T\}$ be Markov chains indexed by T under probability measure μ_P determined by P, $\{g_t(x, y), t \in T\}$ be functions defined on S^2, $L_0 = \{o\}$ (where o is the root of the tree T), $\mathcal{F}_n = \sigma(X^{T^{(n)}})$, λ be a real number. Let*

$$t_n(\lambda, \omega) = \frac{e^{\lambda \sum_{t \in T^{(n)} \setminus \{o\}} g_t(X_{1_t}, X_t)}}{\prod_{t \in T^{(n)} \setminus \{o\}} E_{\mu_P}[e^{\lambda g_t(X_{1_t}, X_t)} | X_{1_t}]} \frac{\mu_P(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}, \tag{20}$$

where E_{μ_P} is the expectation under probability measure μ_P. Then $(t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1)$ is a non-negative martingale under probability measure μ.

Proof The proof of this lemma is similar to that of Lemma 3 of [7], so we omit it. $\qquad\square$

2 Small deviation theorem

Small deviation theorems are a class of strong limit theorems expressed by inequalities. They are the extensions of strong limit theorems expressed by equalities. It is a new research topic proposed by Liu (see [8]).

In this section, we will establish a class of small deviation theorems of functionals and a class of small deviation theorems of the frequencies of occurrence of states for random fields on a uniformly bounded tree.

Theorem 1 *Let T be a uniformly bounded tree which has no leaves, $P = (P(y|x))$, $x, y \in S$ be a positive stochastic matrix, μ, μ_P be two probability measures on $(\Omega, \mathcal{F})$, $\{X_t, t \in T\}$ be Markov chains indexed by T under probability measure μ_P determined by P. Let $\{g_t(x, y), t \in T\}$ be a collection of uniformly bounded functions defined on S^2. Let $\{|g_t(x, y)| \leq K, x, y \in S, \forall t \in T\}$, and let $\varphi(\omega)$ be denoted by (6). Let $c \geq 0$, $D(c)$ be defined*

by (12),

$$F_n(\omega) = \sum_{t \in T^{(n)} \setminus \{o\}} g_t(X_{1_t}, X_t), \tag{21}$$

$$G_n(\omega) = \sum_{t \in T^{(n)} \setminus \{o\}} E_{\mu_P}\big[g_t(X_{1_t}, X_t)|X_{1_t}\big]. \tag{22}$$

Then

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \big[F_n(\omega) - G_n(\omega)\big] \leq K(c + \sqrt{2c}) \quad \mu\text{-}a.e. \text{ on } D(c), \tag{23}$$

$$\liminf_{n \to \infty} \frac{1}{|T^{(n)}|} \big[F_n(\omega) - G_n(\omega)\big] \geq -K(c + \sqrt{2c}) \quad \mu\text{-}a.e. \text{ on } D(c). \tag{24}$$

Proof Let $t_n(\lambda, \omega)$ be defined by (20). By Lemma 3, $(t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1)$ is a non-negative martingale under probability measure μ with $E_\mu(t_n(\lambda, \omega)) = 1$. By Doob's martingale convergence theorem, we have

$$\lim_{n \to \infty} t_n(\lambda, \omega) = t(\lambda, \omega) < \infty \quad \mu\text{-a.e.} \tag{25}$$

Hence

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \ln t_n(\lambda, \omega) \leq 0 \quad \mu\text{-a.e.} \tag{26}$$

We have by (20) and (26)

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \Bigg[\sum_{t \in T^{(n)} \setminus \{o\}} \big(g_t(X_{1_t}, X_t)\lambda - \ln E_{\mu_P}\big(e^{\lambda g_t(X_{1_t}, X_t)}|X_{1_t}\big)\big) - \ln \frac{\mu(X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})} \Bigg]$$
$$\leq 0 \quad \mu\text{-a.e.} \tag{27}$$

By (5), (6), (12) and (27), we have

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \Bigg[\sum_{t \in T^{(n)} \setminus \{o\}} \big(g_t(X_{1_t}, X_t)\lambda - \ln E_{\mu_P}\big(e^{\lambda g_t(X_{1_t}, X_t)}|X_{1_t}\big)\big) \Bigg]$$
$$\leq c \quad \mu\text{-a.e. on } D(c). \tag{28}$$

Taking $\lambda > 0$, we arrive at

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \Bigg[\sum_{t \in T^{(n)} \setminus \{o\}} \big(g_t(X_{1_t}, X_t) - E_{\mu_P}\big(g_t(X_{1_t}, X_t)|X_{1_t}\big)\big) \Bigg]$$

$$\overset{(a)}{\leq} \limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \Bigg[\sum_{t \in T^{(n)} \setminus \{o\}} \bigg(\frac{\ln E_{\mu_P}(e^{\lambda g_t(X_{1_t}, X_t)}|X_{1_t})}{\lambda} - E_{\mu_P}\big(g_t(X_{1_t}, X_t)|X_{1_t}\big) \bigg) \Bigg] + \frac{c}{\lambda}$$

$$\overset{(b)}{\leq} \limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \Bigg[\sum_{t \in T^{(n)} \setminus \{o\}} \bigg(\frac{E_{\mu_P}(e^{\lambda g_t(X_{1_t}, X_t)}|X_{1_t}) - 1}{\lambda} - E_{\mu_P}\big(g_t(X_{1_t}, X_t)|X_{1_t}\big) \bigg) \Bigg] + \frac{c}{\lambda}$$

$$\overset{(c)}{\leq} \frac{\lambda}{2} \limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \left[\sum_{t \in T^{(n)} \setminus \{o\}} E_{\mu_P}\left(g_t^2(X_{1_t}, X_t) e^{\lambda |g_t(X_{1_t}, X_t)|} \big| X_{1_t}\right) \right] + \frac{c}{\lambda}$$

$$\overset{(d)}{\leq} \frac{\lambda}{2} K^2 e^{\lambda K} + \frac{c}{\lambda}$$

$$\overset{(e)}{\leq} \frac{\lambda}{2} K^2 e^{\lambda K} + \frac{c}{\lambda e^{\lambda K}} + Kc \quad \mu\text{-a.e. on } D(c), \tag{29}$$

where (a) follows by (28), (b) follows by the inequality $\ln x \leq x - 1$ ($x > 0$), (c) follows by the inequality $0 \leq e^x - x - 1 \leq \frac{1}{2} x^2 e^{|x|}$, (d) follows by $|g_t(x, y)| \leq K$, $\forall t \in T$, and (e) follows by the inequality $e^{-x} \geq 1 - x$. In the case $c > 0$, noticing that $\frac{\lambda}{2} K^2 e^{\lambda K} + \frac{c}{\lambda e^{\lambda K}}$ attains its smallest value $\sqrt{2c}K$ when $\lambda e^{\lambda K} = \sqrt{\frac{2c}{K^2}}$, by (29) we have

$$\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \left[F_n(\omega) - G_n(\omega)\right] \leq K(c + \sqrt{2c}) \quad \mu\text{-a.e. on } D(c).$$

Hence (23) holds. In the case $c = 0$, (23) also holds by choosing $\lambda_i \to 0^+$ ($i \to \infty$) in (29). Taking $\lambda < 0$ and using a similar approach, we can prove (24). This completes the proof of Theorem 1. □

In the following, we will provide an example showing that $D(c)$ maybe has a positive probability, even has probability 1.

Example Let T be a Cayley tree $T_{C,2}$, μ_P and μ be two probability measures on the measurable space $(\Omega, \mathcal{F})$, and $\{X_t, t \in T\}$ be a collection of random variables taking values in the state space $\{0, 1\}$ defined on the measurable space $(\Omega, \mathcal{F})$. Let $\{X_t, t \in T\}$ be i.i.d. process indexed by tree T under the probability measure μ_P with the common distribution $\mu_P(X_t = 1) = p$, $\mu_P(X_t = 0) = 1 - p$, $0 < p < 1$, and $\{X_t, t \in T\}$ be also i.i.d. process indexed by tree T under the probability measure μ with the common distribution $\mu(X_t = 1) = q$, $\mu(X_t = 0) = 1 - q$, $0 < q < 1$. It is easy to see that $\{X_t, t \in T\}$ are Markov chains indexed by tree T with the transitions matrices

$$\begin{pmatrix} 1 - p & p \\ 1 - p & p \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 - q & q \\ 1 - q & q \end{pmatrix}$$

and stationary distributions $(1 - p, p)$ and $(1 - q, q)$ under the probability measures μ_P and μ, respectively. It is also easy to see that

$$\mu_P\left(X^{T^{(n)}} = x^{T^{(n)}}\right) = p^{s_1(T^{(n)})} (1 - p)^{|T^{(n)}| - s_1(T^{(n)})},$$

$$\mu\left(X^{T^{(n)}} = x^{T^{(n)}}\right) = q^{s_1(T^{(n)})} (1 - q)^{|T^{(n)}| - s_1(T^{(n)})},$$

where $S_1(T^{(n)}) = S_1^0(T^{(n)})$, and $s_1(T^{(n)})$ is its realization of $S_1(T^{(n)})$. In this case

$$\varphi_n(\omega) = \frac{\mu(X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})} = \left(\frac{q}{p}\right)^{S_1(T^{(n)})} \left(\frac{1 - q}{1 - p}\right)^{|T^{(n)}| - S_1(T^{(n)})}.$$

By the strong law of large numbers for Markov chains indexed by tree (see [4]), we have

$$\lim_{n \to \infty} \frac{S_1(T^{(n)})}{|T^{(n)}|} = q \quad \mu\text{-a.e.}$$

Hence we have

$$\varphi(\omega) = \limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \ln \varphi_n(\omega)$$

$$= \lim_{n \to \infty} \frac{S_1(T^{(n)})}{|T^{(n)}|} \ln \frac{q(1-p)}{p(1-q)} + \ln \frac{1-q}{1-p}$$

$$= q \ln \frac{q(1-p)}{p(1-q)} + \ln \frac{1-q}{1-p} \quad \mu\text{-a.e.}$$

Let

$$f(q) = q \ln \frac{q(1-p)}{p(1-q)} + \ln \frac{1-q}{1-p}, \quad 0 < q < 1.$$

It is easy to see that $f(p) = 0$ and $\lim_{q \to 1^-} f(q) = \infty$. Since $f(q)$ is a continuous function, for any $0 \le c < \infty$, there exists q such that $f(q) = c$. Thus $\mu(D(c)) = 1$.

Theorem 2 *Let T be a uniformly bounded tree which has no leaves. M_k, m, M, $D(c)$ are defined as above. Let $0 \le c < \ln(1 - a_k)^{-1}$. Under the assumption of Theorem* 1, *we have*

$$\limsup_{n \to \infty} \left[\frac{S_{k,l}^N(T^{(n)} \setminus \{o\})}{S_k^{N+1}(T^{(n-1)})} - P(l|k) \right]$$

$$\le \frac{(M-1)^N}{(m-1)^N} \left(\sqrt{\frac{2c(M-1)P(l|k)}{M_k(m-1)}} + \frac{c(M-1)}{M_k(m-1)} \right) \quad \mu\text{-a.e. on } D(c), \tag{30}$$

$$\liminf_{n \to \infty} \left[\frac{S_{k,l}^N(T^{(n)} \setminus \{o\})}{S_k^{N+1}(T^{(n-1)})} - P(l|k) \right]$$

$$\ge -\frac{(M-1)^N}{(m-1)^N} \left(\sqrt{\frac{2c(M-1)P(l|k)}{M_k(m-1)}} + \frac{c(M-1)}{M_k(m-1)} \right) \quad \mu\text{-a.e. on } D(c). \tag{31}$$

Proof Letting $g_t(x, y) = \delta_k(x)\delta_l(y)d^N(t)$ in Theorem 1, by (21) and (22), we have

$$F_n(\omega) = \sum_{t \in T^{(n)} \setminus \{o\}} g_t(X_{1_t}, X_t) = \sum_{t \in T^{(n)} \setminus \{o\}} \delta_k(X_{1_t})\delta_l(X_t)d^N(t) = S_{k,l}^N\big(T^{(n)} \setminus \{o\}\big), \tag{32}$$

$$G_n(\omega) = \sum_{t \in T^{(n)} \setminus \{o\}} E_{\mu_P}\big[g_t(X_{1_t}, X_t)|X_{1_t}\big]$$

$$= \sum_{t \in T^{(n)} \setminus \{o\}} \sum_{x_t \in S} \delta_k(X_{1_t})\delta_l(x_t)d^N(t)P(x_t|X_{1_t})$$

$$= \sum_{t \in T^{(n)} \setminus \{o\}} \delta_k(X_{1_t})d^N(t)P(l|k)$$

$$= \sum_{t \in T^{(n-1)}} \delta_k(X_t)d^{N+1}(t)P(l|k)$$

$$= S_k^{N+1}\big(T^{(n-1)}\big)P(l|k). \tag{33}$$

By (28), when $c \geq 0$, we have

$$
\limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \left[\sum_{t \in T^{(n)} \setminus \{o\}} \left(\lambda \delta_k(X_{1_t}) \delta_l(X_t) d^N(t) - \ln E_{\mu_P} \left(e^{\lambda \delta_k(X_{1_t}) \delta_l(X_t) d^N(t)} | X_{1_t} \right) \right) \right]
$$
$$
\leq c \quad \mu\text{-a.e. on } D(c). \tag{34}
$$

By Corollary 1 and (34), when $0 \leq c < \ln(1 - a_k)^{-1}$, we arrive at

$$
\limsup_{n \to \infty} \frac{\sum_{t \in T^{(n)} \setminus \{o\}} (\lambda \delta_k(X_{1_t}) \delta_l(X_t) d^N(t) - \ln E_{\mu_P}(e^{\lambda \delta_k(X_{1_t}) \delta_l(X_t) d^N(t)} | X_{1_t}))}{S_k^{N+1}(T^{(n-1)})}
$$
$$
\leq \frac{c(M-1)}{M_k(m-1)^{N+1}} \quad \mu\text{-a.e. on } D(c). \tag{35}
$$

Taking $\lambda > 0$, we have

$$
\limsup_{n \to \infty} \left[\frac{S_{k,l}^N(T^{(n)} \setminus \{o\})}{S_k^{N+1}(T^{(n-1)})} - P(l|k) \right]
$$
$$
= \limsup_{n \to \infty} \frac{\sum_{t \in T^{(n)} \setminus \{o\}} (\delta_k(X_{1_t}) \delta_l(X_t) d^N(t) - E_{\mu_P}[\delta_k(X_{1_t}) \delta_l(X_t) d^N(t) | X_{1_t}])}{S_k^{N+1}(T^{(n-1)})}
$$
$$
\overset{(f)}{\leq} \limsup_{n \to \infty} \frac{1}{S_k^{N+1}(T^{(n-1)})} \sum_{t \in T^{(n)} \setminus \{o\}} \left(\frac{\ln E_{\mu_P}[e^{\lambda \delta_k(X_{1_t}) \delta_l(X_t) d^N(t)} | X_{1_t}]}{\lambda} \right.
$$
$$
\left. - E_{\mu_P}[\delta_k(X_{1_t}) \delta_l(X_t) d^N(t) | X_{1_t}] \right) + \frac{c(M-1)}{\lambda M_k(m-1)^{N+1}}
$$
$$
\overset{(g)}{\leq} \frac{\lambda}{2} \limsup_{n \to \infty} \frac{\sum_{t \in T^{(n)} \setminus \{o\}} E_{\mu_P}[(\delta_k(X_{1_t}) \delta_l(X_t) d^N(t))^2 e^{\lambda | \delta_k(X_{1_t}) \delta_l(X_t) d^N(t) |} | X_{1_t}]}{S_k^{N+1}(T^{(n-1)})}
$$
$$
+ \frac{c(M-1)}{\lambda M_k(m-1)^{N+1}}
$$
$$
= \frac{\lambda}{2} \limsup_{n \to \infty} \frac{\sum_{t \in T^{(n)} \setminus \{o\}} \sum_{x_t} (\delta_k(X_{1_t}) \delta_l(x_t) d^N(t))^2 e^{\lambda | \delta_k(X_{1_t}) \delta_l(x_t) d^N(t) |} P(x_t | X_{1_t})}{S_k^{N+1}(T^{(n-1)})}
$$
$$
+ \frac{c(M-1)}{\lambda M_k(m-1)^{N+1}}
$$
$$
= \frac{\lambda}{2} \limsup_{n \to \infty} \frac{\sum_{t \in T^{(n)} \setminus \{o\}} (\delta_k(X_{1_t}) d^N(t))^2 e^{\lambda | \delta_k(X_{1_t}) d^N(t) |} P(l | X_{1_t})}{S_k^{N+1}(T^{(n-1)})} + \frac{c(M-1)}{\lambda M_k(m-1)^{N+1}}
$$
$$
\overset{(h)}{\leq} \frac{\lambda (M-1)^{2N} e^{\lambda(M-1)^N} P(l|k)}{2} \limsup_{n \to \infty} \frac{\sum_{t \in T^{(n)} \setminus \{o\}} \delta_k(X_{1_t})}{S_k^{N+1}(T^{(n-1)})} + \frac{c(M-1)}{\lambda M_k(m-1)^{N+1}}
$$
$$
\overset{(i)}{\leq} \frac{\lambda (M-1)^{2N} e^{\lambda(M-1)^N} P(l|k)}{2(m-1)^N} + \frac{c(M-1)}{\lambda M_k(m-1)^{N+1}}
$$
$$
\overset{(j)}{\leq} \frac{1}{(m-1)^N} \left(\frac{\lambda (M-1)^{2N} e^{\lambda(M-1)^N} P(l|k)}{2} + \frac{c(M-1)}{\lambda e^{\lambda(M-1)^N} M_k(m-1)} + \frac{c(M-1)^{N+1}}{M_k(m-1)} \right)
$$
$$
\mu\text{-a.e. on } D(c), \tag{36}
$$

where (f) follows by (35), (g), similarly to (b) and (c) of (29), follows by the inequalities $\ln x \leq x - 1$ $(x > 0)$ and $0 \leq e^x - x - 1 \leq \frac{x^2}{2} e^{|x|}$, (h) follows by the inequality $(m-1)^N \leq d^N(t) \leq (M-1)^N$, (i) follows by the inequality

$$\frac{\sum_{t \in T^{(n)} \setminus \{o\}} \delta_k(X_{1_t})}{S_k^{N+1}(T^{(n-1)})} = \frac{\sum_{t \in T^{(n-1)}} \delta_k(X_t) d^1(t)}{\sum_{t \in T^{(n-1)}} \delta_k(X_t) d^{N+1}(t)} \leq \frac{1}{(m-1)^N},$$

and (j) follows by the inequality $e^{-x} \geq 1 - x$. In the case $c > 0$, notice that

$$\frac{\lambda(M-1)^{2N} e^{\lambda(M-1)^N} P(l|k)}{2} + \frac{c(M-1)}{\lambda e^{\lambda(M-1)^N} M_k(m-1)}$$

attains its smallest value $2\sqrt{\frac{c(M-1)^{2N+1} P(l|k)}{2 M_k(m-1)}}$ when

$$\lambda(M-1)^N e^{\lambda(M-1)^N} = \sqrt{\frac{2c(M-1)}{M_k P(l|k)(m-1)}}.$$

By (36), we have

$$\limsup_{n \to \infty} \left[\frac{S_{k,l}^N(T^{(n)} \setminus \{o\})}{S_k^{N+1}(T^{(n-1)})} - P(l|k) \right]$$

$$\leq \frac{(M-1)^N}{(m-1)^N} \left(\sqrt{\frac{2c(M-1)P(l|k)}{M_k(m-1)}} + \frac{c(M-1)}{M_k(m-1)} \right) \quad \mu\text{-a.e. on } D(c).$$

Hence (30) holds. In the case $c = 0$, (30) also holds by choosing $\lambda_i \to 0^+$ $(i \to \infty)$ in (36). Taking $\lambda < 0$ and using a similar approach, we can prove (31). This completes the proof of Theorem 2. □

Corollary 2 *Under the assumption of Theorem* 2, *we have*

$$\limsup_{n \to \infty} \left[\frac{S_{k,l}^N(T^{(n)} \setminus \{o\})}{S_k^{N+1}(T^{(n-1)})} - P(l|k) \right] = 0 \quad \mu\text{-a.e. on } D(0), \tag{37}$$

Proof Letting $c = 0$, (37) follows from (30) and (31). □

Corollary 3 (see [5]) *Under the assumption of Theorem* 2, *we have*

$$\limsup_{n \to \infty} \left[\frac{S_{k,l}^N(T^{(n)} \setminus \{o\})}{S_k^{N+1}(T^{(n-1)})} - P(l|k) \right] = 0 \quad \mu_P\text{-a.e.} \tag{38}$$

Proof Let $\mu = \mu_P$. Then $\varphi_n(\omega) \equiv 0$, $D(0) = \Omega$. Hence (38) follows (37) directly. □

Corollary 4 (see [7]) *Under the assumptions of Theorem* 2, *if* T *is a* T_{C,N_1}, *we have*

$$\limsup_{n \to \infty} \left[\frac{S_n(k,l)}{N_1 S_{n-1}(k)} - P(l|k) \right] \leq \sqrt{\frac{2cP(l|k)}{M_k}} + \frac{c}{M_k} \quad \mu\text{-a.e. on } D(c), \tag{39}$$

$$\liminf_{n \to \infty} \left[\frac{S_n(k,l)}{N_1 S_{n-1}(k)} - P(l|k) \right] \geq -\sqrt{\frac{2cP(l|k)}{M_k}} - \frac{c}{M_k} \quad \mu\text{-a.e. on } D(c). \tag{40}$$

Proof Since T is a T_{C,N_1}, we can take $M - 1 = N_1$, $m - 1 = N_1$. Since

$$S_{k,l}^0\big(T^{(n)} \setminus \{o\}\big) = \sum_{t \in T^{(n)} \setminus \{o\}} \delta_k(X_{1_t})\delta_l(X_t) = S_n(k,l), \tag{41}$$

$$S_k^{0+1}\big(T^{(n-1)}\big) = \sum_{t \in T^{(n-1)}} \delta_k(X_t)d^1(t) = N_1 S_{n-1}(k). \tag{42}$$

By (30), (31), (41) and (42), we have

$$\limsup_{n \to \infty}\left[\frac{S_n(k,l)}{N_1 S_{n-1}(k)} - P(l|k)\right] \le \sqrt{\frac{2cN_1 P(l|k)}{M_k N_1}} + \frac{cN_1}{M_k N_1}$$

$$= \sqrt{\frac{2cP(l|k)}{M_k}} + \frac{c}{M_k} \quad \mu\text{-a.e. on } D(c),$$

$$\liminf_{n \to \infty}\left[\frac{S_n(k,l)}{N_1 S_{n-1}(k)} - P(l|k)\right] \ge -\sqrt{\frac{2cN_1 P(l|k)}{M_k N_1}} - \frac{cN_1}{M_k N_1}$$

$$= -\sqrt{\frac{2cP(l|k)}{M_k}} - \frac{c}{M_k} \quad \mu\text{-a.e. on } D(c).$$

This completes the proof of Corollary 4.   □

Remark 3 Unfortunately, the upper and lower bounds obtained in Theorem 1 and Theorem 2 are not tight. For example, if we let $|g_t(x,y)| \le 1$ and $c = 1$ in Theorem 1, then $\limsup_{n \to \infty} \frac{1}{|T^{(n)}|}|F_n(\omega) - G_n(\omega)| \le 2$, but $c + \sqrt{2c} > 2$.

References
1. Benjamini, I, Peres, Y: Markov chains indexed by trees. Ann. Probab. **22**, 219-243 (1994)
2. Berger, T, Ye, Z: Entropic aspects of random fields on trees. IEEE Trans. Inf. Theory **36**(5), 1006-1018 (1990)
3. Ye, Z, Berger, T: Entropic, regularity and asymptotic equipartition property of random fields on trees. J. Comb. Inf. Syst. Sci. **21**(2), 157-184 (1996)
4. Yang, WG: Some limit properties for Markov chains indexed by a homogeneous tree. Stat. Probab. Lett. **65**, 241-250 (2003)
5. Huang, HL, Yang, WG: Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree. Sci. China Ser. A **51**(2), 195-202 (2008)
6. Liu, W, Wang, LY: The Markov approximation of the random field on Cayley tree and a class of small deviation theorems. Stat. Probab. Lett. **63**, 113-121 (2003)
7. Peng, WC, Yang, WG, Wang, B: A class of small deviation theorems for functionals of random fields on a homogeneous tree. J. Math. Anal. Appl. **361**, 293-301 (2010)
8. Liu, W: Relative entropy densities and a class of limit theorems of the sequence of *m*-valued random variables. Ann. Probab. **18**, 829-839 (1990)