Hindawi Publishing Corporation
International Journal of Stochastic Analysis
Volume 2014, Article ID 628321, 6 pages
http://dx.doi.org/10.1155/2014/628321

Research Article

Strong Law of Large Numbers for Hidden Markov Chains Indexed by an Infinite Tree with Uniformly Bounded Degrees

Huilin Huang

College of Mathematics and Information Science, Wenzhou University, Zhejiang 325035, China

Correspondence should be addressed to Huilin Huang; huilin_huang@sjtu.org

Received 29 August 2014; Accepted 24 November 2014; Published 9 December 2014

Academic Editor: Lukasz Stettner

We study strong limit theorems for hidden Markov chains fields indexed by an infinite tree with uniformly bounded degrees. We mainly establish the strong law of large numbers for hidden Markov chains fields indexed by an infinite tree with uniformly bounded degrees and give the strong limit law of the conditional sample entropy rate.

1. Introduction

The impact of the theory of hidden Markov models (HMMs) has become widespread for modeling sequences of dependent random variables during the last two decades. Hidden Markov models have many applications in a wide range of areas, such as speech recognition (Rabiner (1998)) [1], image processing [2], DNA sequence analysis (see, e.g., [3, 4]), DNA microarray time course analysis [5], and econometrics [6, 7]. For a good review of statistical and information-theoretic aspects of hidden Markov processes (HMPs), please see [8]. In recent years, the work of Baum and Petrie [9] on finite-state finite-alphabet HMMs has been extended to HMM with finite as well as continuous state spaces and a general alphabet. In particular, statistical properties and ergodic theorems for relative entropy densities of HMMs were developed, and consistency and asymptotic normality of the maximum-likelihood (ML) parameter estimator were proved under some mild conditions [9–12].

In this paper we want to expand tree-indexed homogeneous Markov chain fields to hidden Markov chain fields indexed by an infinite tree.

A tree T is a graph which is connected and contains no loops. 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 with root 0. 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 union of the first n generations of T. For each vertex t, there is a unique path from 0 to t and $|t|$ for the number of edges on this path. We denote the first predecessor of t by ${}^1t$, the second predecessor of t by ${}^2t$, and the nth predecessor of t by ${}^nt$. The degree of a vertex is defined to be the number of neighbors of it. In this paper, we mainly consider an infinite tree which has uniformly bounded degree; that is, the numbers of neighbors of any vertices in this tree are uniformly bounded. For any two vertices s and t of tree T, write $s \leq t$ if s is on the unique path from the root 0 to t. We denote by $s \wedge t$ the vertex farthest from 0 satisfying $s \wedge t \leq s$ and $s \wedge t \leq t$. $X^A = \{X_t,\ t \in A\}$ and denote by $|A|$ the number of vertices of A.

When the context permits, this type of trees is all denoted simply by T.

Definition 1 (homogeneous Markov chains indexed by tree T (see [13, 14])). Let T be an infinite tree with uniformly bounded degrees, let $\mathcal{X}$ be a finite state space, and let $\{X_t,\ t \in T\}$ be a stochastic process defined on probability space $(\Omega, \mathcal{F}, \mathbf{P})$, taking values in the finite set $\mathcal{X}$. Let

$$p = \{p(i),\ i \in \mathcal{X}\} \tag{1}$$

be a distribution on $\mathcal{X}$, and let

$$A = (a(j \mid i)), \quad i, j \in \mathcal{X}, \tag{2}$$

be a transition probability matrix on $\mathcal{X}^2$. If, for any vertex t,

$$P\left(X_t = j \mid X_{1_t} = i,\ X_s = x_s \text{ for } t \wedge s \leq {}^1t\right), \quad \forall x_s \in \mathcal{X}$$
$$= P\left(X_t = j \mid X_{1_t} = i\right) = a\left(j \mid i\right), \quad \forall i, j \in \mathcal{X}, \tag{3}$$

$$P\left(X_0 = i\right) = p\left(i\right), \quad \forall i \in \mathcal{X}. \tag{4}$$

Thus $\{X_t,\ t \in T\}$ will be called $\mathcal{X}$-valued homogeneous Markov chains indexed by infinite tree with the initial distribution (1) and transition probability matrix A whose elements are determined by (3).

Definition 2. Let T be an infinite tree with uniformly bounded degrees and let $\{X_t,\ t \in T\}$ and $\{Y_t,\ t \in T\}$ be two stochastic processes on a probability space $(\Omega, \mathcal{F}, P)$ with finite state spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. Let $A = (a(j \mid i))_{i,j\in\mathcal{X}}$ and $B = (b(y \mid i))_{i\in\mathcal{X}, y\in\mathcal{Y}}$ be two stochastic matrices on $\mathcal{X}^2$ and $\mathcal{X} \times \mathcal{Y}$, respectively. Suppose

$$P\left(X_0 = i\right) = p\left(i\right), \quad \forall i \in \mathcal{X},$$
$$P\left(Y_0 = y_0 \mid X_0 = x_0\right) = b\left(y_0 \mid x_0\right). \tag{5}$$

If for any vertex $t \in T$,

$$P\left(Y_t = y_t, X_t = x_t \mid Y_{1_t} = y_{1_t}, X_{1_t} = x_{1_t},\right.$$
$$\left. Y_s = y_s, X_s = x_s,\ \text{for } t \wedge s \leq {}^1t\right)$$
$$\forall x_t, x_{1_t}, x_s \in \mathcal{X},\ y_t, y_{1_t}, y_s \in \mathcal{Y}$$
$$= P\left(Y_t = y_t, X_t = x_t \mid X_{1_t} = x_{1_t}\right). \tag{6}$$

Moreover, we suppose

$$P\left(Y_t = y_t, X_t = x_t \mid X_{1_t} = x_{1_t}\right)$$
$$= P\left(Y_t = y_t \mid X_t = x_t\right) P\left(X_t = x_t \mid X_{1_t} = x_{1_t}\right) \tag{7}$$
$$:= b\left(y_t \mid x_t\right) a\left(x_t \mid x_{1_t}\right),$$

where

$$a\left(x_t \mid x_{1_t}\right) = \mathbf{P}\left(X_t = x_t \mid X_{1_t} = x_{1_t}\right),$$
$$b\left(y_t \mid x_t\right) = \mathbf{P}\left(Y_t = y_t \mid X_t = x_t\right). \tag{8}$$

Then $\{X_t, Y_t,\ t \in T\}$ will be called $\mathcal{X} \times \mathcal{Y}$-valued *hidden Markov chain indexed by an infinite tree* T, or called *tree-indexed hidden Markov chain* taking values in the finite set $\mathcal{X} \times \mathcal{Y}$.

Remark 3. If we sum over y_t in (6) and take conditional expectations with respect to $\{X_{1_t}, X_s,\ t \wedge s \leq {}^1t\}$ on both sides of the result equation, we can easily arrive at (3). In Definition 2, We can also call the processes $\{X_t,\ t \in T\}$ and $\{Y_t,\ t \in T\}$, respectively, to be *state process* and the *observed process* indexed by an infinite tree.

2. Two Useful Lemmas

Let T be an infinite tree with uniformly bounded degrees and let $\{X_t, Y_t,\ t \in T\}$ be $\mathcal{X} \times \mathcal{Y}$-valued hidden Markov chains indexed by T. Let $g_t(i, j, y)$ be functions defined on $\mathcal{X} \times \mathcal{X} \times \mathcal{Y}$. Let λ be a real number, $L_0 = \{0\}$, $\mathcal{F}_n = \sigma(X^{T^{(n)}}, Y^{T^{(n)}})$, and now we define a stochastic sequence as follows:

$$t_n\left(\lambda, \omega\right) = \frac{e^{\lambda \sum_{t\in T^{(n)}\setminus\{0\}} g_t(X_{1_t}, X_t, Y_t)}}{\prod_{t\in T^{(n)}\setminus\{0\}} E\left[e^{\lambda g_t(X_{1_t}, X_t, Y_t)} \mid X_{1_t}\right]}. \tag{9}$$

At first we come to prove the following fact.

Lemma 4. *$\{t_n(\lambda, \omega), \mathcal{F}_n,\ n \geq 1\}$ is a nonnegative martingale.*

Proof of Lemma 4. Obviously, by (6), the process $Z_t = (X_t, Y_t)$ is a *tree-indexed Markov chain* with state space $\mathcal{X}\times\mathcal{Y}$, so that we have

$$\mathbf{P}\left(Y^{T^{(n)}} = y^{T^{(n)}}, X^{T^{(n)}} = x^{T^{(n)}}\right)$$
$$= p\left(x_0\right) \mathbf{P}\left(Y_0 = y_0 \mid X_0 = x_0\right)$$
$$\times \prod_{t\in T^{(n)}\setminus\{0\}} P\left(Y_t = y_t, X_t = x_t \mid X_{1_t} = x_{1_t}\right) \tag{10}$$
$$= p\left(x_0\right) \mathbf{P}\left(Y_0 = y_0 \mid X_0 = x_0\right)$$
$$\times \prod_{t\in T^{(n)}\setminus\{0\}} \mathbf{P}\left(X_t = x_t \mid X_{1_t} = x_{1_t}\right) \mathbf{P}\left(Y_t = y_t \mid X_t = x_t\right)$$
$$= p\left(x_0\right) b\left(y_0 \mid x_0\right) \prod_{t\in T^{(n)}\setminus\{0\}} b\left(y_t \mid x_t\right) a\left(x_t \mid x_{1_t}\right). \tag{11}$$

Then we have

$$\mathbf{P}\left(X^{L_n} = x^{L_n}, Y^{L_n} = y^{L_n} \mid X^{T^{(n-1)}} = x^{T^{(n-1)}}, Y^{T^{(n-1)}} = y^{T^{(n-1)}}\right)$$
$$= \frac{\mathbf{P}\left(X^{T^{(n)}} = x^{T^{(n)}}, Y^{T^{(n)}} = y^{T^{(n)}}\right)}{\mathbf{P}\left(X^{T^{(n-1)}} = x^{T^{(n-1)}}, Y^{T^{(n-1)}} = y^{T^{(n-1)}}\right)}$$
$$= \prod_{t\in L_n} \mathbf{P}\left(X_t = x_t, Y_t = y_t \mid X_{1_t} = x_{1_t}\right); \tag{12}$$

here the second equation holds because of (10). Furthermore, we have

$$E\left[e^{\lambda \sum_{t\in L_n} g_t(X_{1_t}, X_t, Y_t)} \mid \mathcal{F}_{n-1}\right]$$
$$= \sum_{x^{L_n}, y^{L_n}} e^{\lambda \sum_{t\in L_n} g_t(X_{1_t}, x_t, y_t)}$$
$$\times \mathbf{P}\left(X^{L_n} = x^{L_n}, Y^{L_n} = y^{L_n} \mid X^{T^{(n-1)}}, Y^{T^{(n-1)}}\right)$$
$$= \sum_{x^{L_n}, y^{L_n}} \prod_{t\in L_n} e^{\lambda g_t(X_{1_t}, x_t, y_t)} \mathbf{P}\left(X_t = x_t, Y_t = y_t \mid X_{1_t}\right)$$

$$= \prod_{t\in L_n} \sum_{(x_t,y_t)\in\mathcal{X}\times\mathcal{Y}} e^{\lambda g_t(X_{1_t},x_t,y_t)} \mathbf{P}\left(X_t = x_t, Y_t = y_t \mid X_{1_t}\right)$$
$$= \prod_{t\in L_n} E\left[e^{\lambda g_t(X_{1_t},X_t,Y_t)} \mid X_{1_t}\right] \quad \text{a.e.} \tag{13}$$

On the other hand, we also have

$$t_n(\lambda,\omega) = t_{n-1}(\lambda,\omega) \frac{e^{\lambda \sum_{t\in L_n} g_t(X_{1_t},X_t,Y_t)}}{\prod_{t\in L_n} E\left[e^{\lambda g_t(X_{1_t},X_t,Y_t)} \mid X_{1_t}\right]}. \tag{14}$$

Combining (13) and (14), we get

$$E\left[t_n(\lambda,\omega) \mid \mathcal{F}_{n-1}\right] = t_{n-1}(\lambda,\omega) \quad \text{a.e.} \tag{15}$$

Thus we complete the proof of Lemma 4. □

Lemma 5. *Let $\{X_t, Y_t,\ t \in T\}$ be $\mathcal{X}\times\mathcal{Y}$-valued hidden Markov chains indexed by an infinite tree T with uniformly bounded degrees. $\{g_t(i,j,y),\ t\in T\}$ are functions defined as above; denote*

$$R_n(\omega) = \sum_{t\in T^{(n)}\setminus\{0\}} E\left[g_t\left(X_{1_t},X_t,Y_t\right) \mid X_{1_t}\right]. \tag{16}$$

Let $\alpha > 0$, $\{a_n,\ n\geq 1\}$ be a sequence of nonnegative random variables. Denote

$$D(\alpha) = \Bigg\{\omega : \lim_{n\to\infty} a_n = \infty, \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}\setminus\{0\}} E\left[g_t^2\left(X_{1_t},X_t,Y_t\right) \times e^{\alpha|g_t(X_{1_t},X_t,Y_t)|} \mid X_{1_t}\right] = M(\omega) < \infty\Bigg\}, \tag{17}$$

$$H_n(\omega) = \sum_{t\in T^{(n)}\setminus\{0\}} g_t\left(X_{1_t},X_t,Y_t\right). \tag{18}$$

Then

$$\lim_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{a_n} = 0 \quad \textit{a.e. on } D(\alpha). \tag{19}$$

Proof. By Lemma 4, we have known that $\{t_n(\lambda,\omega), \mathcal{F}_n,\ n\geq 1\}$ is a nonnegative martingale. According to Doob martingale convergence theorem, we have

$$\lim_n t_n(\lambda,\omega) = t(\lambda,\omega) < \infty \quad \text{a.e.,} \tag{20}$$

which implies that

$$\limsup_{n\to\infty} \frac{\ln t_n(\lambda,\omega)}{a_n} \leq 0 \quad \text{a.e.} \tag{21}$$

Combining (9), (18), and (21), we arrive at

$$\limsup_{n\to\infty} \frac{1}{a_n} \left\{\lambda H_n(\omega) - \sum_{t\in T^{(n)}\setminus\{0\}} \ln\left[E\left[e^{\lambda g_t(X_{1_t},X_t,Y_t)} \mid X_{1_t}\right]\right]\right\} \leq 0 \quad \text{a.e.} \tag{22}$$

Let $\lambda > 0$. Dividing two sides of the above equation by λ, we get

$$\limsup_{n\to\infty} \frac{1}{a_n} \left\{H_n(\omega) - \sum_{t\in T^{(n)}\setminus\{0\}} \frac{\ln\left[E\left[e^{\lambda g_t(X_{1_t},X_t,Y_t)} \mid X_{1_t}\right]\right]}{\lambda}\right\} \leq 0 \quad \text{a.e.} \tag{23}$$

For case $0 < \lambda \leq \alpha$, by using inequalities $\ln x \leq x - 1$ $(x > 0)$, $0 \leq e^x - 1 - x \leq (1/2)x^2 e^{|x|}$, and (23)

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{a_n} \left[H_n(\omega) - \sum_{t\in T^{(n)}\setminus\{0\}} E\left[g_t\left(X_{1_t},X_t,Y_t\right) \mid X_{1_t}\right]\right] \\
&\leq \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}\setminus\{0\}} \left\{\frac{\ln\left[E\left[e^{\lambda g_t(X_{1_t},X_t,Y_t)} \mid X_{1_t}\right]\right]}{\lambda} - E\left[g_t\left(X_{1_t},X_t,Y_t\right) \mid X_{1_t}\right]\right\} \\
&\leq \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}\setminus\{0\}} \left\{\frac{E\left[e^{\lambda g_t(X_{1_t},X_t,Y_t)} \mid X_{1_t}\right] - 1}{\lambda} - E\left[g_t\left(X_{1_t},X_t,Y_t\right) \mid X_{1_t}\right]\right\} \\
&\leq \frac{\lambda}{2} \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}\setminus\{0\}} E\left[g_t^2\left(X_{1_t},X_t,Y_t\right) e^{\lambda|g_t(X_{1_t},X_t,Y_t)|} \mid X_{1_t}\right] \\
&\leq \frac{\lambda}{2} \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}\setminus\{0\}} E\left[g_t^2\left(X_{1_t},X_t,Y_t\right) e^{\alpha|g_t(X_{1_t},X_t,Y_t)|} \mid X_{1_t}\right] \\
&\leq \frac{\lambda}{2} M(\omega) \quad \text{a.e. } \omega \in D(\alpha).
\end{aligned} \tag{24}$$

Letting $\lambda \to 0^+$ in (24), combining with (16), we have

$$\limsup_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{a_n} \leq 0 \quad \text{a.e. } \omega \in D(\alpha). \tag{25}$$

Let $-\alpha \leq \lambda < 0$. Similarly to the analysis of the case $0 < \lambda \leq \alpha$, it follows from (22) that

$$\liminf_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{a_n} \geq \frac{\lambda}{2} M(\omega) \quad \text{a.e. } \omega \in D(\alpha). \tag{26}$$

Letting $\lambda \to 0^-$, we can arrive at

$$\liminf_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{a_n} \geq 0 \quad \text{a.e. } \omega \in D(\alpha). \tag{27}$$

Combining (25) and (27), we obtain (19) directly. □

Corollary 6. *Let $\{X_t, Y_t,\ t \in T\}$ be $\mathcal{X} \times \mathcal{Y}$-valued hidden Markov chains indexed by an infinite tree T with uniformly bounded degrees. If $\{g_t(i, j, y),\ t \in T\}$ are the uniformly bounded functions defined on $\mathcal{X}^2 \times \mathcal{Y}$, let N be nonnegative integer. Then*

$$\lim_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{\left|T^{(n+N)}\right|} = 0 \quad a.e. \tag{28}$$

Proof. Letting $a_n = |T^{(n+N)}|$ in Lemma 5, noticing that $\{g_t(i, j, y),\ t \in T\}$ are uniformly bounded, then $D(\alpha) = \Omega$ for all $\alpha > 0$. This theorem follows from Lemma 5 directly. □

3. Strong Law of Large Numbers

In the following, we always let $N \geq 0$, $k \in S$, $d^0(t) := 1$, and denote

$$d^N(t) := \left|\tau \in T : {}^N\tau = t\right|, \tag{29}$$

$$S_n^N(x) = \frac{\sum_{t\in T^{(n)}} d^N(t)\,\delta_x(X_t)}{\left|T^{(n+N)}\right|}, \quad \forall x \in \mathcal{X}; \tag{30}$$

here and thereafter $\delta_x(\cdot)$ denotes the Kronecker function. The following lemma is very useful for proving our main result.

Lemma 7 (see [14]). *Let T be an infinite tree T with uniformly bounded degrees and let $\{X_t,\ t \in T\}$ be a tree-indexed Markov chain with finite state space $\mathcal{X}$, which is determined by initial distribution (1) and finite transition probability matrix A. Suppose that the stochastic matrix A is ergodic, whose unique stationary distribution is π; that is, $\pi A = \pi$ and $\sum_{x\in\mathcal{X}} \pi(x) = 1$. Let $S_n^N(x)$ be defined as (30). Thus one has*

$$\lim_{n\to\infty} S_n^N(x) = \pi(x) \quad a.e. \tag{31}$$

Theorem 8. *Let T be an infinite tree with uniformly bounded degrees and let $\{X_t, Y_t,\ t \in T\}$ be $\mathcal{X}\times\mathcal{Y}$-valued hidden Markov chains indexed by T. For all nonnegative integer N, define the following weighted empirical measure of triples (X_{1_t}, X_t, Y_t):*

$$M_n^N(\tilde{x}, x, y) = \frac{\sum_{t\in T^{(n)}\setminus\{0\}} d^N(t)\,\delta_{\tilde{x}}(X_{1_t})\,\delta_x(X_t)\,\delta_y(Y_t)}{\left|T^{(n+N)}\right|} \quad \forall (\tilde{x}, x, y) \in \mathcal{X}\times\mathcal{X}\times\mathcal{Y}. \tag{32}$$

If the transition probability matrix A of $\{X_t,\ t \in T\}$ is ergodic, one has

$$\lim_{n\to\infty} \left[M_n^N(\tilde{x}, x, y) - \pi(\tilde{x})\,b(y \mid x)\,a(x \mid \tilde{x})\right] = 0 \quad a.e., \tag{33}$$

where π is the stationary distribution of the ergodic matrix A.

Proof. For any $t \in T$, let

$$g_t(i, j, k) = d^N(t)\,\delta_{\tilde{x}}(i)\,\delta_x(j)\,\delta_y(k). \tag{34}$$

Then we have

$$\begin{aligned}
R_n(\omega) &= \sum_{t\in T^{(n)}\setminus\{0\}} E\left[g_t\left(X_{1_t}, X_t, Y_t\right) \mid X_{1_t}\right] \\
&= \sum_{t\in T^{(n)}\setminus\{0\}} \sum_{(x_t,y_t)\in\mathcal{X}\times\mathcal{Y}} g_t\left(X_{1_t}, x_t, y_t\right) \\
&\qquad \times \mathbf{P}\left(X_t = x_t, Y_t = y_t \mid X_{1_t}\right) \\
&= \sum_{t\in T^{(n)}\setminus\{0\}} \sum_{(x_t,y_t)\in\mathcal{X}\times\mathcal{Y}} d^N(t)\,\delta_{\tilde{x}}\left(X_{1_t}\right)\delta_x(x_t)\,\delta_y(y_t) \\
&\qquad \times \mathbf{P}\left(X_t = x_t \mid X_{1_t}\right) \\
&\qquad \times \mathbf{P}\left(Y_t = y_t \mid X_t = x_t\right) \\
&= \sum_{t\in T^{(n)}\setminus\{0\}} d^N(t)\,\delta_{\tilde{x}}\left(X_{1_t}\right) b(y \mid x)\,a(x \mid \tilde{x}) \\
&= \sum_{t\in T^{(n-1)}} d^{N+1}(t)\,\delta_{\tilde{x}}(X_t)\,b(y \mid x)\,a(x \mid \tilde{x}) \\
&= \left|T^{(n+N)}\right| \cdot S_{n-1}^{N+1}(\tilde{x})\,b(y \mid x)\,a(x \mid \tilde{x}), \\
H_n(\omega) &= \sum_{t\in T^{(n)}\setminus\{0\}} g_t\left(X_{1_t}, X_t, Y_t\right) \\
&= \sum_{t\in T^{(n)}\setminus\{0\}} d^N(t)\,\delta_{\tilde{x}}\left(X_{1_t}\right)\delta_x(X_t)\,\delta_y(Y_t) \\
&= \left|T^{(n+N)}\right| \cdot M_n^N(\tilde{x}, x, y).
\end{aligned} \tag{35}$$

Combining (35) and Corollary 6, we obtain

$$\lim_{n\to\infty} \left[M_n^N(\tilde{x}, x, y) - S_{n-1}^{N+1}(\tilde{x})\,b(y \mid x)\,a(x \mid \tilde{x})\right] = 0. \tag{36}$$

By using Lemma 7, our conclusion (33) holds. □

Corollary 9. *Under the conditions of Theorem 8, denote $M_n(\tilde{x}, x, y) = M_n^0(\tilde{x}, x, y)$; then one has*

$$\lim_{n\to\infty} \left[M_n(\tilde{x}, x, y) - \pi(\tilde{x})\,b(y \mid x)\,a(x \mid \tilde{x})\right] = 0 \quad a.e., \tag{37}$$

where π is the stationary distribution of the ergodic matrix A.

For every finite $n \in \mathbf{N}$, let $\{X_t, Y_t,\ t \in T\}$ be $\mathcal{X} \times \mathcal{Y}$-valued hidden Markov chains indexed by an infinite tree T with uniformly bounded degrees. We define the offspring empirical measure as follows:

$$S_n(x, y) = \frac{\sum_{t\in T^{(n)}} \delta_x(X_t)\,\delta_y(Y_t)}{\left|T^{(n)}\right|}, \quad \forall (x, y) \in \mathcal{X}\times\mathcal{Y}. \tag{38}$$

In the following, we consider the limit law of the random sequence of $S_n(x, y)$ which are defined as above.

Theorem 10. *Let T be an infinite tree T with uniformly bounded degrees, and let $\{X_t, Y_t,\ t \in T\}$ be $\mathcal{X} \times \mathcal{Y}$-valued hidden Markov chains indexed by T. If the transition probability matrix A of $\{X_t,\ t \in T\}$ is ergodic, then*

$$\lim_{n\to\infty} S_n(x, y) = \pi(x)\, b(y \mid x) \quad a.e., \tag{39}$$

where π is the stationary distribution of the ergodic matrix A.

Proof. Letting $N = 0$, we have

$$\begin{aligned} M_n(\tilde{x}, x, y) &= M_n^0(\tilde{x}, x, y) \\ &= \frac{\sum_{t\in T^{(n)}\setminus\{0\}} \delta_{\tilde{x}}(X_{1_t})\, \delta_x(X_t)\, \delta_y(Y_t)}{|T^{(n)}|} \\ &\qquad \forall (\tilde{x}, x, y) \in \mathcal{X} \times \mathcal{X} \times \mathcal{Y}. \end{aligned} \tag{40}$$

Comparing (38) with (40), it is easy to see

$$\begin{aligned} S_n(x, y) &= \frac{\sum_{t\in T^{(n)}} \delta_x(X_t)\, \delta_y(Y_t)}{|T^{(n)}|} \\ &= \sum_{\tilde{x}\in\mathcal{X}} M_n(\tilde{x}, x, y) + \frac{\delta_x(X_0)\, \delta_y(Y_0)}{|T^{(n)}|}. \end{aligned} \tag{41}$$

Taking limit on both sides of above equation as n tends to infinity, it follows from Corollary 9 that

$$\begin{aligned} \lim_{n\to\infty} S_n(x, y) &= \sum_{\tilde{x}\in\mathcal{X}} \pi(\tilde{x})\, b(y \mid x)\, a(x \mid \tilde{x}) \\ &= \pi(x)\, b(y \mid x) \quad \text{a.e.}, \end{aligned} \tag{42}$$

where the last equation holds because π is unique stationary distribution of the ergodic stochastic matrix A; that is, $\pi A = \pi$. Thus we complete the proof of Theorem 10. □

From the expression of (38) we can easily obtain the empirical measure of the observed chain $\{Y_t,\ t \in T\}$ which is denoted by $M_n(\cdot)$,

$$M_n(y) = \sum_{x\in\mathcal{X}} S_n(x, y) = \frac{\sum_{t\in T^{(n)}} \delta_y(Y_t)}{|T^{(n)}|}, \quad \forall y \in \mathcal{Y}. \tag{43}$$

Thus we can obtain Corollary 11.

Corollary 11. *Under the same conditions of Theorem 10, one has*

$$\lim_{n\to\infty} M_n(y) = \sum_{x\in\mathcal{X}} \pi(x)\, b(y \mid x) \quad a.e. \tag{44}$$

Let $f(x, y)$ be any function defined on $\mathcal{X} \times \mathcal{Y}$. Denote

$$G_n(\omega) = \sum_{t\in T^{(n)}} f(X_t, Y_t). \tag{45}$$

By simpe computation, we arrive at Corollary 12.

Corollary 12. *Under the same conditions of Theorem 10, one also has*

$$\lim_{n\to\infty} \frac{G_n(\omega)}{|T^{(n)}|} = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi(x)\, b(y \mid x)\, f(x, y) \quad a.e. \tag{46}$$

Now we define conditional entropy rate of $Y^{T^{(n)}}$ given $X^{T^{(n)}}$ by

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \ln \mathbf{P}\left(Y^{T^{(n)}} \mid X^{T^{(n)}}\right). \tag{47}$$

From (6), we obtain that

$$f_n(\omega) = -\frac{\sum_{t\in T^{(n)}} \ln P(Y_t \mid X_t)}{|T^{(n)}|}. \tag{48}$$

The convergence of $f_n(\omega)$ to a constant in a sense (L_1 convergence, convergence in probability, a.e. convergence) is called the conditional version of *Shannon-McMillan theorem* or the entropy theorem or the AEP (asymptotic equipartition property) in information theory. Here from Corollary 12, if we let

$$f(X_t, Y_t) = -\ln P(Y_t \mid X_t), \tag{49}$$

we can easily obtain the Shannon-McMillan theorem with a.e. convergence for conditional entropy theorem of hidden Markov chain fields on tree T.

Corollary 13. *Under the same conditions of Theorem 10, one has*

$$\lim_{n\to\infty} f_n(\omega) = -\sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} \pi(x)\, b(y \mid x) \ln b(y \mid x) \quad a.e. \tag{50}$$

Here one also specifies $0 \ln 0$ *as zero by convention.*

4. Conclusion

This paper gives some strong limit theorems for hidden Markov chains fields indexed by an infinite tree with uniformly bounded degrees. We study the strong law of large numbers for hidden Markov chains fields indexed by an infinite tree with uniformly bounded degrees and give the strong law of the conditional sample entropy rate.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by National Natural Science Foundation of China (Grant no. 11201344).

References

[1] L. R. Rabiner, "Tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[2] J. Li and R. M. Gray, *Image Segmentation and Compression Using Hidden Markov Models*, Kluwer Academic Publishers, 2000.

[3] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison, *Biological Sequence Analysis*, Cambridge University Press, Cambridge, UK, 1998.

[4] T. Koski, *Hidden Markov Models for Bioinformatics*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.

[5] M. Yuan and C. Kendziorski, "Hidden Markov models for microarray time course data in multiple biological conditions," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1323–1332, 2006.

[6] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, vol. 57, no. 2, pp. 357–384, 1989.

[7] C. A. Sims and T. Zha, "Were there regime switches in U.S. monetary policy?" *The American Economic Review*, vol. 96, no. 1, pp. 54–81, 2006.

[8] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1518–1569, 2002.

[9] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *Annals of Mathematical Statistics*, vol. 37, pp. 1554–1563, 1966.

[10] T. Petrie, "Probabilistic functions of finite state Markov chains," *Annals of Mathematical Statistics*, vol. 40, pp. 97–115, 1969.

[11] B. G. Leroux, "Maximum-likelihood estimation for hidden Markov models," *Stochastic Processes and their Applications*, vol. 40, no. 1, pp. 127–143, 1992.

[12] P. J. Bickel and Y. Ritov, "Inference in hidden Markov models. I. Local asymptotic normality in the stationary case," *Bernoulli*, vol. 2, no. 3, pp. 199–228, 1996.

[13] I. Benjamini and Y. Peres, "Markov chains indexed by trees," *The Annals of Probability*, vol. 22, no. 1, pp. 219–243, 1994.

[14] H. Huang and W. Yang, "Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree," *Science in China Series A: Mathematics*, vol. 51, no. 2, pp. 195–202, 2008.