Chapter 2

Uniform CLT for Markov chains with a countable state space

2.1 Introduction

Let $(S, \mathcal{G}, P)$ be a probability space and let $\mathcal{F}$ be a set of measurable functions on S with an envelope function F finite everywhere. Let $X_1, X_2, ...$ be a strictly stationary sequence of random variables with distribution P, and define the empirical measures P_n, based on $\{X_i\}$, as $P_n = n^{-1} \sum_{i=1}^{n} \delta_{X_i}$. We say the uniform CLT holds over $\mathcal{F}$, if $n^{\frac{1}{2}}(P_n - P)$ converges in law, in the space $l^\infty(\mathcal{F})$ to a Gaussian process. Of course, $l^\infty(\mathcal{F})$ is not separable unless $\mathcal{F}$ is a finite set, but Giné and Zinn [8, p56] includes a suitable definition of weak convergence in non-separable spaces.

Let V be a subspace of the space of measurable functions on S such that $\mathcal{F} \subseteq V$, and let $\|\cdot\|$ be a norm on V. Define the bracketing number of $\mathcal{F}$ with respect to the norm $\|\cdot\|$ and V by letting, for $\varepsilon > 0$, $N_{[\,]}(\varepsilon, \mathcal{F}, \|\cdot\|)$ be the minimal number of brackets $[g_1, h_1], ..., [g_n, h_n]$, with all $g_i, h_i \in V$, such that for all $f \in \mathcal{F}$ there exists $[g_i, h_i]$, for some $i, 1 \leq i \leq n$ with $g_i \leq f \leq h_i$ and $\|h_i - g_i\| < \varepsilon$. Ossiander [14] proved that if $\{X_i\}$ is i.i.d. and $\mathcal{F} \subseteq L^2(S, P)$ satisfies

$$\int_0^1 (\ln N_{[\,]}(\varepsilon, \mathcal{F}, \|\cdot\|_2))^{\frac{1}{2}} d\varepsilon < \infty, \tag{2.1}$$

then the uniform CLT holds over $\mathcal{F}$.

In the discrete space case, the Borisov-Durst theorem [5, p47] says that for $\{X_i\}$ i.i.d. on $S = \{1, 2, 3, \cdots\}$ with distribution π, let $\mathcal{F} = \{1_A : A \subseteq S\}$, then the following are equivalent:
a) the uniform CLT holds over $\mathcal{F}$,
b) $\sum_{k=1}^{\infty} \pi^{\frac{1}{2}}(k) < \infty$,
c) $\int_0^1 (\ln N_{[\,]}(\varepsilon, \mathcal{F}, \|\cdot\|_2))^{\frac{1}{2}} d\varepsilon < \infty$.

Let $\{X_i\}_{i\geq 0}$ be a positive recurrent irreducible Markov chain taking values in $S = \{1, 2, 3, \cdots\}$ with the unique invariant probability measure π, N_i be the i-th hitting time of state 1,

$$m_{i,j} = E\left(\min\{n : n \geq 1, X_n = j\} \mid X_0 = i\right).$$

Levental (1990) [12] generalized Durst and Dudley's result (1981) [7] that (a) and (b) above in the i.i.d. case are equivalent by showing that for Markov chains when $E(N_2 - N_1)^2 < \infty$, then the uniform CLT holds over $\{1_A : A \subseteq S\}$ if and only if

$$\sum_{k=1}^{\infty} \pi(k) m_{1,k}^{\frac{1}{2}} < \infty. \tag{2.2}$$

Of course, in the i.i.d. case $m_{1,k} = (\pi(k))^{-1}$, so (2.2) coincides with (b).

We will prove a uniform CLT for regenerative processes and then apply the theorem to Markov chains to generalize Levental's theorem from the set of indicator functions to the set of possibly unbounded functions. Let $\{X_i\}_{i\geq 0}$ be a positive recurrent irreducible Markov chain. Let F be a non-negative function on S and $\mathcal{F} = \{f : |f| \leq F\}$. We have that the uniform CLT holds over $\mathcal{F}$ if and only if $E(N_2 - N_1)^2 < \infty$,

$$E\Big(\sum_{N_1 < j \leq N_2} F(X_j)\Big)^2 < \infty$$

and

$$\sum_{k=1}^{\infty} F(k)\pi(k) m_{1,k}^{\frac{1}{2}} < \infty.$$

To compare our results to those obtained from weakly dependent observations, we let $\{X_i\}$ be a strictly stationary sequence of random variables and recall the definitions of some classical mixing coefficients.
Strong mixing coefficient:

$$\alpha_k = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \sigma_1^l, B \in \sigma_{l+k}^{\infty}, l \geq 1\}.$$

Absolutely regular mixing coefficient:

$$\beta_k = \frac{1}{2}\sup\{\sum_{i=1}^{I}\sum_{j=1}^{J}|P(A_i\cap B_j) - P(A_i)P(B_j)| : \{A_i\}_{i=1}^{I}\text{is a partition in } \sigma_1^l$$
$$\text{and } \{B_j\}_{j=1}^{J}\text{is a partition in } \sigma_{l+k}^{\infty}, l\geq 1\}.$$

Obviously $\alpha_k \leq \frac{1}{2}\beta_k$.

The best known results for the uniform CLT based on strong mixing empirical processes require at least $\alpha_k = O(k^{-a})$ for some $a > 3$, [3], [13]. For absolutely regular empirical processes, Doukhan, Massart and Rio (1995) [4] obtained a uniform CLT over classes of functions which satisfy a bracketing condition with respect to a norm $\|\cdot\|_{2,\beta}$. This norm depends on P and on the mixing structure of the sequence, and coincides with the usual $L^2(P)$-norm in the independent case. Their result generalizes Ossiander's theorem for independent observations.

In the last section we will prove that when the β_k decay at a polynomial rate, the bracketing condition in Doukhan, Massart and Rio with respect to the class of indicator functions implies that $\sum_{x\in S}(\pi(x))^{\frac{1}{2+\delta}} < \infty$ for some $\delta > 0$. Our uniform CLT over the family of functions dominated by a non-negative function is the best possible result for positive recurrent irreducible Markov chains.We also present an example to illustrate that, in the Markov chain case, applying the known mixing empirical process results will not get the results obtained by our approach. The mixing approach requires that $\alpha_k \ll k^{-3}$ or the bracketing conditions hold. In our example the rates of decay of α_k and β_k can be taken such that $\alpha_k \gg k^{-3}$ and also $\sum_{x\in S}(\pi(x))^{\frac{1}{2+\delta}}$ diverges. Hence the bracketing condition of the mixing approach fails.

2.2 Main Results

A regenerative process, informally speaking, is a stochastic process that can be divided into blocks which are identically distributed and independent. To state the results, we need a formal definition and some notation. The following is a simplification of Levental's [11] general space notation.

(i) $S = \{1, 2, 3, ...\}$ is a discrete space.

(ii) Ω stands for the set of all sequences $\{y_i\}_{1\leq i<\infty}$ such that $y_i = (x_i, \phi_i)$ where $x_i \in S$ and $\phi_i \in \{0, 1\}$.

(iii) P is a probability measure on Ω.

(iv) The coordinate maps $X_n : \Omega \to S$ are defined by $X_n(\{y_i\}) = x_n$ and $\Phi_n : \Omega \to \{0,1\}$ are defined by $\Phi_n(\{y_i\}) = \phi_n$.

(v) $N_i = \min\{j \geq 1 : \sum_{1\leq k\leq j} \Phi_k = i\}, i = 1, ...$ or $N_i = \infty$ if the set that we minimize over is empty. $\{N_i\}$ are called renewal times. For every $i \geq 1$ N_i is a stopping time relative to the increasing sequence of σ-algebras $(\sigma\{W_1, ..., W_n\})_{1\leq n}$ where by W_n we denote the coordinate maps $W_n(\{y_i\}) = y_n$. $\mathcal{G}_{N_i}$ is the σ-algebra associated with the stopping time N_i, i.e. : $\mathcal{G}_{N_i} = \sigma\{W_{k\wedge N_i} : k = 1, 2, ...\}$. θ_k is a shift operator: $\theta_k : \{y_i\}_{i\geq 1} \to \{y_{i+k}\}_{i\geq 1}$ for every $k \geq 1$.

Definition 2.1. $\{X_i\}$ will be called a regenerative process if $N_i < \infty$ almost surely for every $i \geq 1$ and if for every $f : \Omega \to \mathbf{R}$ which is bounded $E[f(\theta_{N_i}) \mid \mathcal{G}_{N_i}] = E[f(\theta_{N_1})]$.

The following two properties of the process $\{W_i\}$ are equivalent to the above definition:
(i) The post $N_i{+}1$ process is independent of the occurence up to and including N_i, and
(ii) $$\mathcal{L}((W_{N_1+1}, ...)) = \mathcal{L}((W_{N_i+1}, ...)) \quad \text{for all } i = 1, 2, ...$$

We assume that $E(N_2 - N_1) < \infty$ and denote $\mu = E(N_2 - N_1)$ throughout the paper. Define

$$\pi(A) = \frac{1}{\mu} E(\sum_{N_1<j\leq N_2} 1_A(X_j)) \quad \text{for all } A \subseteq S.$$

Then π is a probability measure on S (called a steady state distribution). For Markov chains, let N_i be the i-th hitting time of state 1, then π is the invariant propability measure [6, p262]. We also assume that $\pi(k) > 0$ for all $k \in S$ and write $\pi(f) = \sum_{k\in S} f(k)\pi(\{k\})$ for all $f \in L^1(S, \pi)$. Then

$$n^{-1} \sum_{k=1}^{n} f(X_k) \to \pi(f) \quad \text{a.s. for all } f \in L^1(S, \pi).$$

For the proof of the above property see [1, p92] where the statement is formulated for Markov chains but the same proof works for regenerative processes.

Define for every $f \in L^1(S, \pi)$

$$S_n(f) = \sum_{j=1}^{n} \left(f(X_j) - \pi(f) \right).$$

Our generalization of Levental's theorem is the following result.

Theorem 2.2. Suppose $E(N_2 - N_1)^2 < \infty$. Let F be a non-negative function on S and $\mathcal{F} = \{f : |f| \leq F\}$. Then

$$E\Big(\sum_{N_1 < j \leq N_2} F(X_j)\Big)^2 < \infty \tag{2.3}$$

and

$$\sum_{k=1}^{\infty} F(k)\Big(E\Big(\sum_{N_1 < j \leq N_2} 1_{\{k\}}(X_j)\Big)^2\Big)^{\frac{1}{2}} < \infty, \tag{2.4}$$

if and only if the uniform CLT holds over $\mathcal{F}$, namely $\{n^{-\frac{1}{2}} S_n(f)\}_{f \in \mathcal{F}}$, converges in law, as random elements of $l^\infty(\mathcal{F})$, to a Gaussian process $\{G(f)\}_{f \in \mathcal{F}}$ whose covariance is

$$E\left(G(f)G(g)\right) = \frac{1}{\mu} E\left(Z_1(f) Z_1(g)\right) \tag{2.5}$$

where $Z_1(f) = \sum_{N_1 < j \leq N_2} (f(X_j) - \pi(f))$. Furthermore, $\{G(f)\}_{f \in \mathcal{F}}$ is uniformly continuous with respect to the L^2-norm metric of $L^2(S, \pi)$ restricted to $\mathcal{F}$.

Remark. The condition (2.3) implies $F \in L^2(S, \pi)$, since

$$\left(\|F\|_{L^2(S,\pi)}\right)^2 = \frac{1}{\mu} E\Big(\sum_{N_1 < j \leq N_2} F^2(X_j)\Big) \leq \frac{1}{\mu} E\Big(\sum_{N_1 < j \leq N_2} F(X_j)\Big)^2 < \infty.$$

The equality above readily follows from the definition of π.

Proof of Theorem 2.2. First we show (2.3) and (2.4) imply the uniform CLT over $\mathcal{F}$. Define

$$Z_k(f) = \sum_{N_k < j \leq N_{k+1}} \left(f(X_j) - \pi(f) \right)$$

for all $k \geq 1$. By the properties of regenerative processes the $Z_k(f)$ are i.i.d.. Let

$$R_n(f) = \sum_{1 \leq j \leq N_1 \text{ or } N_{l(n)} < j \leq n} (f(X_i) - \pi(f)),$$

and $l(n) = \max\{k : N_k \leq n\}$. Then

$$S_n(f) = \sum_{i=1}^{l(n)-1} Z_i(f) + R_n(f).$$

We have the following lemma.

Lemma 2.3. Let F be a non-negative function in $L^1(S, \pi)$ and $\mathcal{F} = \{f : |f| \leq F\}$. Then

$$n^{-\frac{1}{2}} \sup_{f \in \mathcal{F}} |R_n(f)| \to 0 \text{ in probability as } n \to \infty.$$

Proof. This follows from Chung's proof [1, p99], since

$$\sup_{f \in \mathcal{F}} |R_n(f)| \leq \sum_{1 \leq j \leq N_1 \text{ or } N_{l(n)} < j \leq n} F(X_j) + (N_1 + n - N_{l(n)})\pi(F).$$

Chung's proof is for Markov chains but the same proof will work for regenerative processes.

Lemma 2.4. Suppose $E(N_2 - N_1)^2 < \infty$ and (2.4) holds. Then for all $\varepsilon > 0$,

$$\lim_{\delta \to 0} \operatorname*{limsup}_{n \to \infty} P(\sup_{[\delta]} n^{-\frac{1}{2}} \mid \sum_{i=1}^{n} (Z_i(f) - Z_i(g)) \mid > \varepsilon) = 0.$$

where $[\delta] = \{f, g \in \mathcal{F}, \|f - g\|_2 < \delta\}$.

Proof. For $f, g \in \mathcal{F}$, $|f(k) - g(k)| \leq 2F(k)$ for all $k \in S$, and $\|f - g\|_2 < \delta$ implies $|f(k) - g(k)| \leq \delta(\pi(k))^{-\frac{1}{2}}$ for all $k \in S$. Let $\delta(k) = \min\{2F(k), \frac{\delta}{\sqrt{\pi(k)}}\}$. Then for $(f, g) \in [\delta]$, $|f(k) - g(k)| \leq \delta(k)$ for all $k \in S$. Let $m(\delta)$ be the largest integer such that $\min\{\pi(1), \pi(2), \cdots, \pi(m(\delta))\} \geq \delta^{\frac{2}{3}}$ or $m(\delta) = 1$ if $\pi(1) < \delta^{\frac{2}{3}}$. Clearly, $m(\delta) \to \infty$ as $\delta \to 0$. Since

$$\begin{aligned}
\sup_{[\delta]} \mid \sum_{i=1}^{n} (Z_i(f) - Z_i(g)) \mid &\leq \sup_{[\delta]} \sum_{k=1}^{\infty} |f(k) - g(k)| \mid \sum_{i=1}^{n} Z_i(1_{\{k\}}) \mid \\
&\leq \sum_{k=1}^{\infty} \delta(k) \mid \sum_{i=1}^{n} Z_i(1_{\{k\}}) \mid,
\end{aligned}$$

we have

$$
\begin{aligned}
P(\sup_{[\delta]} n^{-\frac{1}{2}} \mid \sum_{i=1}^{n} (Z_i(f) - Z_i(g)) \mid > \varepsilon) &\leq P(n^{-\frac{1}{2}} \sum_{k=1}^{\infty} \delta(k) \mid \sum_{i=1}^{n} Z_i(1_{\{k\}}) \mid > \varepsilon) \\
&\leq \frac{1}{\varepsilon} \sum_{k=1}^{\infty} \delta(k)\ E(n^{-\frac{1}{2}} \mid \sum_{i=1}^{n} Z_i(1_{\{k\}}) \mid).
\end{aligned}
$$

Since $Z_i(\cdot)$ are i.i.d. and centered,

$$
E(n^{-\frac{1}{2}} \mid \sum_{i=1}^{n} Z_i(1_{\{k\}}) \mid) \leq (n^{-1} E \mid \sum_{i=1}^{n} Z_i(1_{\{k\}}) \mid^2)^{\frac{1}{2}} = \left(E\left(Z_1^2(1_{\{k\}})\right)\right)^{\frac{1}{2}}. \tag{2.6}
$$

Denote $\omega(k) = (E(\sum_{N_1 < j \leq N_2} 1_{\{k\}}(X_j))^2)^{\frac{1}{2}}$, and then by definition,

$$
\begin{aligned}
\left(E\left(Z_1^2(1_{\{k\}})\right)\right)^{\frac{1}{2}} &= (E(\sum_{N_1 < j \leq N_2} 1_{\{k\}}(X_j) - (N_2 - N_1)\pi(k))^2)^{\frac{1}{2}} \\
&\leq \omega(k) + \pi(k) \left(E(N_2 - N_1)^2\right)^{\frac{1}{2}}.
\end{aligned} \tag{2.7}
$$

Thus

$$
\begin{aligned}
& P(\sup_{[\delta]} n^{-\frac{1}{2}} \mid \sum_{i=1}^{n} (Z_i(f) - Z_i(g)) \mid > \varepsilon) \\
\leq\ & \frac{1}{\varepsilon} \sum_{k=1}^{\infty} \delta(k)(\omega(k) + \pi(k) \left(E(N_2 - N_1)^2\right)^{\frac{1}{2}}).
\end{aligned}
$$

Since $\omega(k) \geq \mu\pi(k)$, it is enough to show that $\sum_{k=1}^{\infty} \delta(k)\omega(k) \to 0$ as $\delta \to 0$. If $\delta^{\frac{2}{3}} \leq \pi(1)$ then $(\pi(k))^{-\frac{1}{2}} \leq \delta^{-\frac{1}{3}}$ for $1 \leq k \leq m(\delta)$, and hence we have

$$
\begin{aligned}
\sum_{k=1}^{\infty} \delta(k)\omega(k) &\leq \sum_{k=1}^{m(\delta)} \frac{\delta}{\sqrt{\pi(k)}} \omega(k) + \sum_{k=m(\delta)+1}^{\infty} 2F(k)\omega(k) \\
&\leq \delta^{\frac{2}{3}} \sum_{k=1}^{m(\delta)} \omega(k) + 2 \sum_{k=m(\delta)+1}^{\infty} F(k)\omega(k).
\end{aligned}
$$

The right term converges to zero since $m(\delta) \to \infty$ as $\delta \to 0$ and (2.4).The left term satisfies

$$
\delta^{\frac{2}{3}} \sum_{k=1}^{m(\delta)} \omega(k) \leq \delta^{\frac{2}{3}} m(\delta) \left[E\left(N_2 - N_1\right)^2\right]^{\frac{1}{2}}.
$$

Hence for $m > 0$ fixed,

$$\limsup_{\delta\to 0}\delta^{\frac{2}{3}}m(\delta) = \limsup_{\delta\to 0}\delta^{\frac{2}{3}}\left(m(\delta)-m\right) \le \limsup_{\delta\to 0}\sum_{k=m+1}^{m(\delta)}\pi(k) = \sum_{k=m+1}^{\infty}\pi(k),$$

and the right hand side converges to zero as $m\to\infty$.

Lemma 2.5. Suppose that $n^{-\frac{1}{2}}\left(l(n)-\frac{n}{\mu}\right)$ converges in law to a normal distribution and $\pi(F)<\infty$. Then

$$\sup_{f\in\mathcal{F}} n^{-\frac{1}{2}} \mid \sum_{i=1}^{l(n)-1} Z_i(f) - \sum_{i=1}^{\left[\frac{n}{\mu}\right]} Z_i(f) \mid \to 0 \qquad \text{in probability.}$$

Proof. Put $a=\min\{l(n)-1,[n/\mu]\}$, $b=\max\{l(n)-1,[n/\mu]\}$ and fix $\varepsilon>0$. Then

$$P(\sup_{f\in\mathcal{F}} n^{-\frac{1}{2}} \mid \sum_{i=1}^{l(n)-1} Z_i(f) - \sum_{i=1}^{\left[\frac{n}{\mu}\right]} Z_i(f) \mid > \varepsilon) = P(\sup_{f\in\mathcal{F}} n^{-\frac{1}{2}} \mid \sum_{a<i\le b} Z_i(f) \mid > \varepsilon).$$

Fix $\delta>0$, then there exits a constant $c>0$ such that $P(|l(n)-[n/\mu]|\le c\sqrt{n})>1-\delta$ for n large enough. Thus the right hand side of the last equation is equal to or less than

$$\begin{aligned}
& \delta + P(\{\sup_{f\in\mathcal{F}} n^{-\frac{1}{2}} \mid \sum_{a<i\le b} Z_i(f) \mid > \varepsilon\} \cap \{\mid l(n)-[\tfrac{n}{\mu}] \mid \le c\sqrt{n}\}) \\
\le\ & \delta + P(\sup_{f\in\mathcal{F}} \max_{\left|s-\left[\frac{n}{\mu}\right]\right|\le c\sqrt{n}} 2n^{-\frac{1}{2}} \mid \sum_{\left[\frac{n}{\mu}\right]-c\sqrt{n}\le i\le s} Z_i(f) \mid > \varepsilon) \\
=\ & \delta + P(\sup_{f\in\mathcal{F}} \max_{1\le s\le 2c\sqrt{n}} 2n^{-\frac{1}{2}} \mid \sum_{1\le i\le s} Z_i(f) \mid > \varepsilon).
\end{aligned}$$

Now we consider Z_i as random vetors in $\mathcal{C}(\mathcal{F},\mathbf{R})$, the separable Banach space of all continuous functions from $\mathcal{F}$ to the real number equipped with supremun norm. Since

$$\begin{aligned}
E(\sup_{f\in\mathcal{F}}|Z_1(f)|) &= E(\sup_{f\in\mathcal{F}} \mid \sum_{N_1<j\le N_2} f(X_i) - \pi(f)\,(N_2-N_1) \mid) \\
&\le E(\sum_{N_1<j\le N_2} F(X_i)) + \pi(F)\,E\,(N_2-N_1) < \infty,
\end{aligned}$$

thus $E\|Z_1\| < \infty$ and hence by the uniform SLLN [10, Corollary 7.10],

$$n^{-1} \parallel \sum_{i=1}^{n} Z_i \parallel = n^{-1} \sup_{f\in\mathcal{F}} \mid \sum_{i=1}^{n} Z_i(f) \mid \to 0 \quad \text{a.s.},$$

Thus $n^{-\frac{1}{2}} \sup_{f\in\mathcal{F}} \mid \sum_{i=1}^{\sqrt{n}} Z_i(f) \mid \to 0$ a.s.. This implies

$$\max_{1\leq s\leq 2c\sqrt{n}} n^{-\frac{1}{2}} \sup_{f\in\mathcal{F}} \mid \sum_{i=1}^{s} Z_i(f) \mid \to 0 \quad \text{a.s.}$$

(since $n^{-1}a_n \to 0$ implies $\max_{1\leq s\leq n} n^{-1}a_s \to 0$), and hence

$$P(\sup_{f\in\mathcal{F}} \max_{1\leq s\leq 2c\sqrt{n}} 2n^{-\frac{1}{2}} \mid \sum_{1\leq i\leq s} Z_i\,(f) \mid > \varepsilon) \to 0 \quad \text{as } n\to\infty.$$

Since δ is arbitrary, we have

$$P(\sup_{f\in\mathcal{F}} n^{-\frac{1}{2}} \mid \sum_{i=1}^{l(n)-1} Z_i(f) - \sum_{i=1}^{\left[\frac{n}{\mu}\right]} Z_i(f) \mid > \varepsilon) \to 0 \quad \text{as } n\to\infty,$$

which completes the proof.

We know that $E(N_2 - N_1)^2 < \infty$ implies that $n^{-\frac{1}{2}}\left(l(n) - \frac{n}{\mu}\right)$ converges in law to a normal distribution. So we have the following lemma.

Lemma 2.6. Suppose $E(N_2 - N_1)^2 < \infty$ and (2.4) holds. Then for all $\varepsilon > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{[\delta]} n^{-\frac{1}{2}} |S_n(f) - S_n(g)| > \varepsilon) = 0.$$

Proof. By Lemma 2.3, we only have to show that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{[\delta]} n^{-\frac{1}{2}} \mid \sum_{i=1}^{l(n)-1} (Z_i(f) - Z_i(g)) \mid > \varepsilon) = 0.$$

Since

$$\begin{aligned} & \sup_{[\delta]} \mid \sum_{i=1}^{l(n)-1} (Z_i(f) - Z_i(g)) \mid \\ \leq\ & \sup_{[\delta]} \mid \sum_{i=1}^{\left[\frac{n}{\mu}\right]} (Z_i(f) - Z_i(g)) \mid + 2\sup_{f\in\mathcal{F}} \mid \sum_{i=1}^{l(n)-1} Z_i(f) - \sum_{i=1}^{\left[\frac{n}{\mu}\right]} Z_i(f) \mid, \end{aligned}$$

that (2.3) and (2.4) imply the uniform CLT can be completed from Lemma 2.4 and Lemma 2.5.

That is, we now have:

(i) finite dimensional convergence by (2.3) [1, p99], (there the convergence is formulated for Markov chains but the same proof works for regenerative processes)

(ii) $\mathcal{F}$ is compact in $L^2(S,\pi)$, thus totally bounded, since $F \in L^2(S,\pi)$,

(iii) asymptotically stochastic equicontinuity i.e. for each $\varepsilon > 0$ and $\eta > 0$ there exists a $\delta > 0$ such that $\limsup_{n\to\infty} P(\sup_{[\delta]} n^{-\frac{1}{2}} |S_n(f-g)| > \varepsilon) < \eta$.

Applying Theorem 21 in Pollard [15, p157], shows that the conditions are sufficient for the uniform CLT.

For the converse part, (adapted from the proof of [12, Theorem 4]). Suppose the uniform CLT holds over $\mathcal{F}$, and let $\{G(f)\}_{f\in\mathcal{F}}$ be the limiting Gaussian process. If support$(f) = A$ is a finite set, then $E(G(f) - \sum_{k\in A} f(k)G(1_{\{k\}}))^2 = 0$ by using covariance (2.5) and hence $G(f) = \sum_{k\in A} f(k)G(1_{\{k\}})$ a.s.. On the other hand, $\sup_{f\in\mathcal{F}} \mid G(f) \mid = \|G\|_\infty < \infty$ a.s., thus

$$\sum_{k=1}^{\infty} F(k) \mid G(1_{\{k\}}) \mid < \infty \quad \text{a.s.},$$

since $F \in \mathcal{F}$. This means that the law of the process $\{F(k)G(1_{\{k\}})\}_{k\in S}$ defines a Gaussian measure on the separable Banach space $l^1(S)$. The first moment of the norm of a random vector, which has a Gaussian distribution in a separable Banach space, is finite, and hence

$$\infty > E \parallel \{F(k)G(1_{\{k\}})\}_{k\in S} \parallel_1 = \sum_{k=1}^{\infty} F(k)E \mid G(1_{\{k\}}) \mid .$$

Since $G(1_{\{k\}})$ are centered normal random variables,

$$E\left|G(1_{\{k\}})\right| = \sqrt{\frac{2}{\pi}} \left(E\left(G^2\left(1_{\{k\}}\right)\right)\right)^{\frac{1}{2}},$$

and hence
$$\sum_{k=1}^{\infty} F(k)\left(E\left(G^2\left(1_{\{k\}}\right)\right)\right)^{\frac{1}{2}} < \infty. \qquad (2.8)$$
From Chung [1, p99],
$$\begin{aligned}\left(E\left(G^2\left(1_{\{k\}}\right)\right)\right)^{\frac{1}{2}} &= \left(\mu^{-1}E\left(Z_1\left(1_{\{k\}}\right)\right)\right)^{\frac{1}{2}} \\ &= \mu^{-\frac{1}{2}} \parallel \sum_{N_1<j\leq N_2} 1_{\{k\}}(X_j) - \pi(k)(N_2-N_1) \parallel_2 .\end{aligned}$$

Thus
$$\begin{aligned}&\sum_{k=1}^{\infty} F(k)(E(\sum_{N_1<j\leq N_2} 1_{\{k\}}(X_j))^2)^{\frac{1}{2}} \\ \leq\ &\sum_{k=1}^{\infty} F(k)(\parallel \sum_{N_1<j\leq N_2} 1_{\{k\}}(X_j) - \pi(k)(N_2-N_1) \parallel_2 + \|\pi(k)(N_2-N_1)\|_2) \\ =\ &\sum_{k=1}^{\infty} F(k)\sqrt{\mu}\left(E\left(G^2\left(1_{\{k\}}\right)\right)\right)^{\frac{1}{2}} + \sum_{k=1}^{\infty} F(k)\pi(k)\left\|N_2-N_1\right\|_2 .\end{aligned}$$

Using (2.8), we have
$$\sum_{k=1}^{\infty} F(k)(E(\sum_{N_1<j\leq N_2} 1_{\{k\}}(X_j))^2)^{\frac{1}{2}} < \infty. \qquad (2.9)$$

We also need to show (2.3) is necessary. Since $n^{-\frac{1}{2}}S_n(F)$ converges in law to a normal distribution $G(F)$ and $n^{-\frac{1}{2}}R_n(F) \to 0$ in probability, we have
$$n^{-\frac{1}{2}} \sum_{k=1}^{l(n)-1} Z_k(F)$$
converges in law to $G(F)$. By Lemma 2.5 $n^{-\frac{1}{2}} \sum_{k=1}^{[n/\mu]} Z_k(F)$ converges in law to $G(F)$ since (2.9) impies $\pi(F) < \infty$. This is equivalent to $EZ_1^2(F) < \infty$ [9, p181], thus
$$\begin{aligned}&(E(\sum_{N_1<j\leq N_2} F(X_j))^2)^{\frac{1}{2}} \\ \leq\ &\parallel \sum_{N_1<j\leq N_2} F(X_j) - \pi(F)(N_2-N_1) \parallel_2 + \|\pi(F)(N_2-N_1)\|_2 \\ =\ &\left(E\left(Z_1^2(F)\right)\right)^{\frac{1}{2}} + \pi(F)\left\|N_2-N_1\right\|_2 < \infty.\end{aligned}$$

Now let $\{X_j\}_{j\geq 0}$ be a positive recurrent irreducible Markov chain taking values in $S=\{1,2,3,\cdots\}$ with an invariant probability measure π. Let N_i be the i-th hitting time of state 1, and consider $\{X_i\}$ as a regenerative process with renewal times $\{N_i\}$. Then apply Theorem 2.2 to the Markov chain $\{X_i\}$ to obtain the following theorem.

Theorem 2.7. Let $\{X_i\}$ be a positive recurrent irreducible Markov chain. Let F be a non-negative function on S and $\mathcal{F}=\{f : |f| \leq F\}$. Then $E(N_2-N_1)^2<\infty$,

$$E(\sum_{N_1<j\leq N_2} F(X_j))^2 < \infty \tag{2.10}$$

and

$$\sum_{k=1}^{\infty} F(k)\pi(k)m_{1,k}^{\frac{1}{2}} < \infty \tag{2.11}$$

if and only if the uniform CLT holds over $\mathcal{F}$, i.e. $\{n^{-\frac{1}{2}}\sum_{j=1}^{n}(f(X_j)-\pi(f))\}_{f\in\mathcal{F}}$ converges in law to a Gaussian process in the sense of Theorem 2.2.

Remark. The condition $E(N_2-N_1)^2<\infty$ and (2.10) does not depend on the state which we choose to use to define the hitting times N_i, [1, p84] and (2.11) is equivalent to $\sum_{k=1}^{\infty} F(k)\pi(k)m_{j,k}^{\frac{1}{2}}<\infty$ for any $j\in S$. Since $m_{i,k}\leq m_{i,j}+m_{j,k}$, we have that for j fixed

$$m_{1,k}^{\frac{1}{2}} \leq m_{1,j}^{\frac{1}{2}} + m_{j,k}^{\frac{1}{2}} \quad \text{and} \quad m_{j,k}^{\frac{1}{2}} \leq m_{j,1}^{\frac{1}{2}} + m_{1,k}^{\frac{1}{2}}.$$

Thus (2.11) is equivalent to $\sum_{k=1}^{\infty} F(k)\pi(k)m_{j,k}^{\frac{1}{2}}<\infty$.

Proof of Theorem 2.7. First we show that $E(N_2-N_1)^2<\infty$ is a necessary condition for the uniform CLT holds over $\mathcal{F}$. Without loss of generality we assume that $F(1)>0$. From the fact that $\{1_{\{1\}}(X_j)\}$ satisfy the CLT, that is $n^{-\frac{1}{2}}(\sum_{j=1}^{n} 1_{\{1\}}(X_j)-n\pi(1))$ converges in law to a normal distribution $G(1_{\{1\}})$. Then we have $n^{-\frac{1}{2}}(l(n)-\frac{n}{m_{1,1}})$ converges in law to $G(1_{\{1\}})$, since $\sum_{j=1}^{n} 1_{\{1\}}(X_j)=l(n)$. On the other hand,

$$n^{-\frac{1}{2}}\sum_{j=1}^{l(n)-1} Z_j(1_{\{1\}})$$

converges in law to $G(1_{\{1\}})$ since $S_n(f) = \sum_{j=1}^{l(n)-1} Z_j(f) + R_n(f)$ and $n^{-\frac{1}{2}} R_n(1_{\{1\}}) \to 0$ in probability. Using Lemma 2.5, we thus have

$$n^{-\frac{1}{2}} \sum_{j=1}^{[n/m_{1,1}]} Z_j(1_{\{1\}})$$

converges in law to $G(1_{\{1\}})$, and hence $E(Z_1^2(1_{\{1\}})) < \infty$ [9, p181]. But

$$\begin{aligned} E\left(Z_1^2(1_{\{1\}})\right) &= E\left(\sum_{N_1 < j \le N_2} 1_{\{1\}}(X_j) - \pi(1)(N_2 - N_1)^2\right)^2 \\ &= E\left(1 - \pi(1)(N_2 - N_1)\right)^2. \end{aligned}$$

Therefore $E(N_2 - N_1)^2 < \infty$.

By virtue of Theorem 2.2, we only need to show that (2.4) and (2.11) are equivalent. First suppose (2.11) holds. From Chung [1, p88],

$$E(\sum_{N_1 < j \le N_2} 1_{\{k\}}(X_j))^2 = 2m_{1,1}\pi^2(k)\left(m_{1,k} + m_{k,1}\right) - m_{1,1}\pi(k) \quad \text{for } k \ge 1 \tag{2.12}$$

and

$$E\left(N_2 - N_1\right)^2 = 2m_{1,1}\sum_{k=1}^{\infty} \pi(k) m_{k,1} - m_{1,1}. \tag{2.13}$$

Since $E\left(N_2 - N_1\right)^2 < \infty$ and $m_{1,1} < \infty$, we thus have $\sum_{k=1}^{\infty} \pi(k) m_{k,1} < \infty$, and hence

$$\sum_{k=1}^{\infty} F(k)\pi(k) m_{k,1}^{\frac{1}{2}} \le (\sum_{k=1}^{\infty} F^2(k)\pi(k))^{\frac{1}{2}} (\sum_{k=1}^{\infty} \pi(k) m_{k,1})^{\frac{1}{2}} < \infty, \tag{2.14}$$

since $F \in L^2(S, \pi)$. Thus $\sum_{k=1}^{\infty} F(k)\pi(k) m_{k,1}^{\frac{1}{2}} < \infty$, and hence

$$\begin{aligned} & \sum_{k=1}^{\infty} F(k)(E(\sum_{N_1 < j \le N_2} 1_{\{k\}}(X_j))^2)^{\frac{1}{2}} \\ = & \sum_{k=1}^{\infty} F(k)\left(2m_{1,1}\pi^2(k)\left(m_{1,k} + m_{k,1}\right) - m_{1,1}\pi(k)\right)^{\frac{1}{2}} \\ \le & \sqrt{2m_{1,1}} \sum_{k=1}^{\infty} F(k)\pi(k)(m_{1,k}^{\frac{1}{2}} + m_{k,1}^{\frac{1}{2}}) < \infty. \end{aligned}$$

Conversely, suppose (2.4) holds. Then

$$\sum_{k=1}^{\infty} F(k)\pi(k)m_{1,k}^{\frac{1}{2}} \leq m_{1,1}^{-\frac{1}{2}} \sum_{k=1}^{\infty} F(k)\left(m_{1,1}\pi^2(k)\left(m_{1,k}+m_{k,1}\right)\right)^{\frac{1}{2}}. \tag{2.15}$$

Since

$$m_{1,1}\pi^2(k)\left(m_{1,k}+m_{k,1}\right) - m_{1,1}\pi(k) = m_{1,1}\pi(k)(\frac{m_{1,k}+m_{k,1}}{m_{k,k}} - 1) \geq 0,$$

the right hand side of (2.15) is equal to or less than

$$m_{1,1}^{-\frac{1}{2}} \sum_{k=1}^{\infty} F(k)\left(2m_{1,1}\pi^2(k)\left(m_{1,k}+m_{k,1}\right) - m_{1,1}\pi(k)\right)^{\frac{1}{2}}.$$

Combining with (2.4) and (2.12), we obtain (2.11).

2.3 Comparison with mixing results

It is known that a positive recurrent irreducible Markov chain has convergent absolutely regular mixing coefficients [2]. Using empirical process CLT's for stationary sequences satisfying absolutely regular mixing conditions one can also obtain results similar to those above. However, our conditions are less restrictive than those required for a mixing process application to these problems, and an example is given to show these differences.

Let $X_1, X_2, ...$ be a strictly stationary sequence of random variables with distribution P, and assume that the absolutely regular mixing coefficient sequence $\{\beta_k\}$ satisfies the summability condition $\sum_{k\geq 1}\beta_k < \infty$. Define the mixing rate function $\beta(\cdot)$ by $\beta(t) = \beta_{[t]}$ if $t \geq 1$, and $\beta(t) = 1$ otherwise. For any numerical function f, we denote by Q_f the quantile function of $|f(X_1)|$, that is

$$Q_f(u) = \inf\{t : P(|f(X_1)| > t) \leq u\}.$$

Let $\mathcal{F}$ be a class of functions in the function space $L_{2,\beta}(P)$, here the norm is defined by

$$\|f\|_{2,\beta} = (\int_0^1 \beta^{-1}(u)\left(Q_f(u)\right)^2 du)^{\frac{1}{2}},$$

where $\beta^{-1}(u) = \inf\{t : \beta(t) \leq u\}$. Doukhan, Massart and Rio (1995) [4] proved that a sufficient condition for the uniform CLT holding over $\mathcal{F}$ is that

$$\int_0^1 (\log N_{[\ ]}(\varepsilon, \mathcal{F}, \|\cdot\|_{2,\beta}))^{\frac{1}{2}} d\varepsilon < \infty \tag{2.16}$$

where $N_{[\ ]}(\varepsilon, \mathcal{F}, \|\cdot\|_{2,\beta})$ is the bracketing number of $\mathcal{F}$ with respect to the norm $\|\cdot\|_{2,\beta}$ and $L_{2,\beta}(P)$.

¿From now on, we assume $S = \{1, 2, 3, ...\}$ with distribution $\{p_k\}_{k\geq 1}$ and let $\mathcal{F} = \{1_A : A \subseteq S\}$. Then $N_{[\ ]}(\varepsilon, \mathcal{F}, \|\cdot\|)$ does not depend on the space V from which we choose brackets. In fact, we can choose brackets just from the set of all indicator functions $\mathcal{F}$. Since for any bracket $[g, h]$, we define $[\overline{g}, \overline{h}]$ by setting

$$\begin{array}{ll} \overline{h}(x) = 1,\ \overline{g}(x) = 0 & \text{when } h(x) \geq 1,\ g(x) \leq 0, \\ \overline{h}(x) = 1,\ \overline{g}(x) = 1 & \text{when } h(x) \geq 1,\ g(x) > 0, \\ \overline{h}(x) = 0,\ \overline{g}(x) = 0 & \text{when } h(x) < 1. \end{array}$$

Then $[g, h] \subseteq [\overline{g}, \overline{h}]$ and $\left|\overline{h} - \overline{g}\right| \leq |h - g|$, thus we can replace $[g, h]$ by $[\overline{g}, \overline{h}]$.

Proposition 2.8. Suppose $\beta_k \gg k^{-a}$ for some $a > 1$, then (2.16) implies

$$\sum_{k=1}^{\infty} p_k^{(2+\frac{2}{2a-1})^{-1}} < \infty.$$

Proof. From assumption, $\beta_k \geq ck^{-a}$ for some constant $c > 0$, thus $\beta^{-1}(u) \geq c^{\frac{1}{a}} u^{-\frac{1}{a}}$, and we also have $Q_{1_A} = 1_{[0, \sum_{k\in A} p_k)}$ for all $A \subseteq S$. Then

$$\begin{array}{rcl} \|1_A\|_{2,\beta} & \geq & (\int_0^1 c^{\frac{1}{a}} u^{-\frac{1}{a}} 1_{[0,\sum_{k\in A} p_k)} du)^{\frac{1}{2}} \\ & = & c^{\frac{1}{2a}} ((1 - a^{-1})^{-1} u^{1-\frac{1}{a}} |_0^{\sum_{k\in A} p_k})^{\frac{1}{2}} \\ & = & (\frac{c^{\frac{1}{a}} a}{a-1})^{\frac{1}{2}} (\sum\limits_{k\in A} p_k)^{\frac{a-1}{2a}}. \end{array}$$

Since $\|\cdot\|_{2,\beta} \geq \|\cdot\|_2$, then $N_{[\ ]}(\varepsilon, \mathcal{F}, \|\cdot\|_{2,\beta}) \geq N_{[\ ]}(\varepsilon, \mathcal{F}, \|\cdot\|_2)$ and hence (2.16) implies $\sum_{k=1}^{\infty} p_k^{\frac{1}{2}} \leq M < \infty$ by the Borisov-Durst theorem [5, p47]. Then for all $A \subseteq S$,

$$\sum_{k\in A} p_k^{\frac{a-1}{a}+\frac{1}{2a}} \leq (\sum_{k\in A} p_k)^{\frac{a-1}{a}} (\sum_{k\in A} p_k^{\frac{1}{2}})^{\frac{1}{a}},$$

and hence

$$(\sum_{k\in A} p_k)^{\frac{a-1}{a}} \geq M^{-\frac{1}{a}} \sum_{k\in A} p_k^{\frac{2a-1}{2a}}.$$

Define a new norm by

$$\|f\|_{2,a} = (\sum_{k=1}^{\infty} f^2(k) p_k^{\frac{2a-1}{2a}})^{\frac{1}{2}}.$$

Then for all $A \subseteq S$,

$$\|1_A\|_{2,\beta} \geq (\frac{c^{\frac{1}{a}} a}{a-1})^{\frac{1}{2}} M^{-\frac{1}{2a}} \|1_A\|_{2,a}. \tag{2.17}$$

We observe that (2.16) and (2.17) assert that

$$\int_0^1 (\log N_{[\,]}(\varepsilon, \mathcal{F}, \|.\|_{2,a}))^{\frac{1}{2}} d\varepsilon < \infty. \tag{2.18}$$

By the Borisov-Durst theorem (2.18) is equivalent to $\sum_{k=1}^{\infty} p_k^{\frac{1}{2}(\frac{2a-1}{2a})} < \infty$.

Under the condition $F \equiv 1$ and $\mathcal{F} = \{1_A : A \subseteq S\}$ our conditions for the uniform CLT are $E(N_2 - N_1)^2 < \infty$ and

$$\sum_{k=1}^{\infty} (E(\sum_{N_1 < j \leq N_2} 1_{\{k\}}(X_j))^2)^{\frac{1}{2}} < \infty. \tag{2.19}$$

The mixing results in Doukhan, Massart and Rio require at least the bracketing condition (2.16). We present an example in which our conditions hold but the bracketing condition fails.

Example 2.9. Let $\{X_i\}$ be a stationary Markov chain with transition probability

$$p_{n,n+1} = (\frac{n}{n+1})^s, p_{n,1} = 1 - p_{n,n+1} \text{ for all } n \geq 1 \text{ and some } s > 1.$$

Then the invariant probability measure is $p_k = ck^{-s}$ where $c = (\sum_{k=1}^{\infty} k^{-s})^{-1}$.
(i) Since $P(N_2 - N_1 = n) = \left(\frac{1}{2}\right)^s \left(\frac{2}{3}\right)^s \cdots \left(\frac{n-1}{n}\right)^s (1 - (\frac{n}{n+1})^s) = n^{-s} \frac{(n+1)^s - n^s}{(n+1)^s}$,

$$E(N_2 - N_1)^2 = \sum_{n=1}^{\infty} n^2 P(N_2 - N_1 = n) \approx \sum_{n=1}^{\infty} n^{1-s}.$$

Thus $E\left(N_2-N_1\right)^2<\infty$ if $s>2$.
(ii) Examine the condition (2.19). Note that

$$E(\sum_{N_1<j\leq N_2} 1_{\{k\}}(X_j))^2=\sum_{n=k}^{\infty} P\left(N_2-N_1=n\right) \approx \sum_{n=k}^{\infty} n^{-s-1} \approx k^{-s},$$

thus (2.19) holds if $s>2$.
(iii) $\sum_{k=1}^{\infty} p_k^{\frac{1}{2+\delta}}<\infty$, only if $s>2+\delta$.
(iv) We claim that $\alpha_k \gg k^{-s}$ and $\beta_k \gg k^{1-s}$. Recall that mixing coefficients α_k and β_k are defined by

$$\alpha_k=\sup\{|P(A\cap B)-P(A)P(B)|: A\in\sigma_1^l, B\in\sigma_{l+k}^{\infty}, l\geq 1\}$$

and

$$\beta_k=\tfrac{1}{2}\sup\{\sum_{i=1}^{I}\sum_{j=1}^{J}\left|P\left(A_i\cap B_j\right)-P(A_i)P(B_j)\right|;\{A_i\}_{i=1}^{I}\text{is a partition of the sample space in } \sigma_1^l, \{B_j\}_{j=1}^{J} \text{ is a partition of the sample space in } \sigma_{l+k}^{\infty}, l\geq 1\}.$$

We take $A_k=\{X_1=1\}$ and $B_k=\{X_{k+1}=k+2\}$ then $P(A_k\cap B_k)=0$ and hence

$$\alpha_k\geq P(A_k)P(B_k)\approx(k+2)^{-s}\approx k^{-s}.$$

For the absolutely regular mixing coefficient, we take

$$l=1,\quad A_1=\{X_1=1\},\quad A_2=A_1^c,$$

$$B_j=\{X_{k+1}=k+1+j\}\quad\text{for } j=1,2,\cdots,J\quad\text{and}\quad B_{J+1}=(\bigcup_{j=1}^{J} B_j)^c.$$

Note that $P(A_1\cap B_j)=0$ for $j=1,2,\cdots,J$. Thus

$$\beta_k\geq\frac{1}{2}\sup_J\sum_{j=1}^{J}P(A_1)P(B_j)\approx\sum_{j=1}^{\infty}(\frac{1}{k+1+j})^s\approx k^{1-s}.$$

Now if we restrict $2<s<5/2$, obviously our conditions $E(N_2-N_1)^2<\infty$ and (2.19) hold. Suppose that the bracketing condition (2.16) holds, then by Proposition 2.8, we have

$$\textstyle\sum_{k=1}^{\infty} p_k^{\frac{1}{2+\delta}}<\infty$$

where $\delta=2/(2\left(s-1\right)-1)$. Thus $2/(2\left(s-1\right)-1)<s-2$ by (iii), but this contradicts to $2<s<5/2$. Therefore the bracketing condition fails.

Bibliography

[1] Chung, K. L. Markov chain with stationary probability, Springer, 1967.

[2] Davydov, Yu. A. Mixing conditions for Markov chains. Theory of probability and its Application 28,1973, 313-318.

[3] Dhompongsa, S. A note on the almost sure approximation of empirical prosess of weakly dependent random vectors. Yokohama math. J. 32, 1984, 113-121.

[4] Doukhan, P., Massart, P. and Rio, E. Invariance principles for absolutely regular empirical processes. Ann. Inst. H. Poincaré Probab. Statist. 31, no. 2, 1995, 393-427.

[5] Dudley, R. M. A course on empirical processes. Lecture Notes in Math, Vol. 1097. Spring, New York, 1984.

[6] Durrett, R. Probability: theory and examples. Wadsworth & Brooks/Cole Statistics/Probability Series, 1991.

[7] Durst, M. and Dudley, R. M. Empirical processes, Vapnik-Cervonenkis classes and Poisson processes. Probab. Math. Statist. 1, 1981, 109-115.

[8] Giné, E., Zinn, J. Lectures on the central limit theorem for empirical processes. Lect. Notes Math. Val. 1221, 1986, 50-113.

[9] Gnedenko, B. V. and Kolmogorov, A. N. Limit distributions for sums of independent random variables, 1967.

[10] Ledoux, M. and Talagrad, M. Probability in Banach spaces, Springer-Verlag, 1991.

[11] Levental, S. Uniform limit theorems for Harris recurrent Markov chains. Probability theory and related fields 80, 1988, 101-118.

[12] Levental, S. Uniform CLT for Markov chain with a countable state space. Stochastic processes and their Application 34, 1990, 245-253.

[13] Massart, P. Invariance principles for empirical processes: the weakly case, Quelques problèmes de vitesse de convergence pour des measures empiriques. Thèse d'Etat, Université de Paris-Sud, 1987.

[14] Ossiander, M. A central limit theorem under metric entropy with L_2-bracketing. Annals of probability 15, 1987, 897-919.

[15] Pollard, D. Convergence of stochastic processes, Springer, New York, 1984.