A simple variance inequality for U-statistics of a Markov chain with applications ☆

G. Fort[a,*], E. Moulines[a], P. Priouret[b], P. Vandekerkhove[c]

[a] *LTCI-CNRS/TELECOM ParisTech, 46 rue Barrault, 75634 Paris Cedex 13, France. Fax: +(33) 1 45 81 71 44*
[b] *LPMA, Université Pierre et Marie Curie, Boîte Courrier 188, 75252 Paris Cedex 5, France*
[c] *LAMA, Université de Marne-la-Vallée 5, boulevard Descartes, 77454 Marne-la-Vallée cedex 2, France*

Abstract

We establish a simple variance inequality for U-statistics whose underlying sequence of random variables is an ergodic Markov Chain. The constants in this inequality are explicit and depend on computable bounds on the mixing rate of the Markov Chain. We apply this result to derive the strong law of large number for U-statistics of a Markov Chain under conditions which are close from being optimal.

Keywords: U-statistics, Markov chains, Inequalities, Limit theorems, Law of large numbers

1. Introduction

Let $\{Y_n\}_{n=0}^{\infty}$ be a sequence of random variables with values in a measurable space $(\mathsf{Y}, \mathcal{Y})$. Let m be an integer and $h : \mathsf{Y}^m \to \mathbb{R}$ be a symmetric function. For $n \geq m$, the U-statistic associated to h is defined by

$$U_{n,m}(h) \stackrel{\text{def}}{=} \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} h(Y_{i_1}, \ \ldots, \ Y_{i_m}) \,. \tag{1}$$

The function h is often referred to as the kernel of the U-statistics and m is called the degree of h. We refer to Serfling (1980), Lee (1990), and Koroljuk and Borovskich (1994) for U-statistics whose underlying sequence is an i.i.d. sequence of random variables.

☆This work is partially supported by the French National Research Agency, under the program ANR-08-BLAN-0218 BigMC

*Corresponding Author

Email addresses: gersende.fort@telecom-paristech.fr (G. Fort), eric.moulines@telecom-paristech.fr (E. Moulines), p.priouret@upmc.jussieu.fr (P. Priouret), Pierre.Vandekerkhove@univ-mlv.fr (P. Vandekerkhove)

Preprint submitted to Statistics and Probability Letters *July 13, 2011*

Several authors have studied U-statistics for *stationary* sequences of dependent random variables under different dependence conditions: see Arcones (1998), Borovkova et al. (2001), Dehling (2006) and the references therein. Much less efforts have been spent on the behavior of U-statistics for non-stationary and asymptotically stationary processes; see Harel and Puri (1990) and Elharfaoui and Harel (2008). In this letter, we establish a variance inequality for U-statistics whose underlying sequence is an ergodic Markov Chain (which is not assumed to be stationary). This inequality is valid for U-statistics of any order and the constants appearing in the bound can be explicitly computed (for example, using Foster-Lyapunov drift and minorization conditions if the chain is geometrically ergodic). This inequality can be used to derive, with minimal effort, limit theorems for U-statistics of a non-stationary Markov chain. In this paper, for the purpose of illustration, we derive the strong law of large numbers (SLLN) under weak conditions.

Notations

Let $(\mathsf{Y}, \mathcal{Y})$ be a general state space (see e.g. (Meyn and Tweedie, 2009, Chapter 3)) and P be a Markov transition kernel. P acts on bounded measurable functions f on Y and on measures μ on $\mathcal{Y}$ via

$$Pf(x) \stackrel{\text{def}}{=} \int P(x, \mathrm{d}y) f(y) , \qquad \mu P(A) \stackrel{\text{def}}{=} \int \mu(\mathrm{d}x) P(x, A) .$$

We will denote by P^n the n-iterated transition kernel defined by induction

$$P^n(x, A) \stackrel{\text{def}}{=} \int P^{n-1}(x, \mathrm{d}y) P(y, A) = \int P(x, \mathrm{d}y) P^{n-1}(y, A) ;$$

where P^0 coincides with the identity kernel. For a function $V : \mathsf{Y} \to [1, +\infty)$, define the V-norm of a function $f : \mathsf{Y} \to \mathbb{R}$ by

$$|f|_V \stackrel{\text{def}}{=} \sup_{\mathsf{Y}} |f|/V .$$

When $V = 1$, the V-norm is the supremum norm and will be denoted by $|f|_\infty$. Let $\mathcal{L}_V$ be the set of measurable functions such that $|f|_V < +\infty$. For two probability measures μ_1, μ_2 on $(\mathsf{Y}, \mathcal{Y})$, $\|\mu_1 - \mu_2\|_{\mathrm{TV}}$ denotes the total variation distance.

For μ a probability distribution on $(\mathsf{Y}, \mathcal{Y})$ and P a Markov transition kernel on $(\mathsf{Y}, \mathcal{Y})$, denote by $\mathbb{P}_\mu$ the distribution of the Markov chain $(Y_n)_{n \in \mathbb{N}}$ with initial distribution μ and transition kernel P; let $\mathbb{E}_\mu$ be the associated expectation. For $p > 0$ and Z a random variable measurable with respect to the σ-algebra $\sigma\left((Y_n)_{n \in \mathbb{N}}\right)$, set $\|Z\|_{\mu,p} \stackrel{\text{def}}{=} \left(\mathbb{E}_\mu\left[|Z|^p\right]\right)^{1/p}$.

2. Main Results

Let P be a Markov transition kernel on $(\mathsf{Y}, \mathcal{Y})$. We assume that the transition kernel P satisfies the following assumption:

A1 The kernel P is positive Harris recurrent and has a unique stationary distribution π. In addition, there exist a measurable function $V : \mathsf{Y} \to [1, +\infty)$ and a nonnegative non-increasing sequence $(\rho(k))_{k\in\mathbb{N}}$ such that $\lim_n \rho(n) = 0$ and for any probability distributions μ and μ' on $(\mathsf{Y}, \mathcal{Y})$, and any integer k,

$$\left\|\mu P^k - \mu' P^k\right\|_{\mathrm{TV}} \leq \rho(k)\left[\mu(V) + \mu'(V)\right] \ , \tag{2}$$

and

$$\pi(V) < \infty \ . \tag{3}$$

A2 The function h is symmetric and π-canonical, *i.e,* for all $(y_1, \ldots, y_{m-1}) \in \mathsf{Y}^{m-1}$, $y \mapsto h(y_1, \ldots, y_{m-1}, y)$ is π-integrable and

$$\int \pi(\mathrm{d}y) h(y_1, \ldots, y_{m-1}, y) = 0\,. \tag{4}$$

For μ a probability measure on $(\mathsf{Y}, \mathcal{Y})$, we denote

$$M(\mu, V) \stackrel{\mathrm{def}}{=} \sup_{k\geq 0} \mu P^k(V) \ . \tag{5}$$

Note that, under **A 1**, for any probability measure μ on $(\mathsf{X}, \mathcal{X})$, $\pi(V) \leq M(\mu, V)$. We can now state the main result of this paper, which is an explicit bound for the variance of bounded π-canonical U-statistics. The proof of Theorem 2.1 is given in Section 3.

Theorem 2.1. *Assume **A1-A2**. If $|h|_\infty < \infty$ then, for any initial probability measure μ on $(\mathsf{Y}, \mathcal{Y})$,*

$$\|U_{n,m}(h)\|_{\mu,2} \leq C_{n,m}\sqrt{M(\mu, V)}|h|_\infty\, n^{-m/2} \tag{6}$$

with

$$C_{n,m} \stackrel{\mathrm{def}}{=} 2^{m/2+1}\sqrt{(2m)!}\left(\sum_{k=0}^{n}(k+1)^m \rho(k)\right)^{1/2} \frac{n^m}{\binom{n}{m}} \ . \tag{7}$$

Remark 1. In the case where $\rho(k) = \varrho^k$ for some $\varrho \in (0, 1)$, for all $(m, n) \in \mathbb{N}$,

$$\sum_{k=0}^{n}(k+1)^m \rho(k) \leq \frac{1}{\varrho\,(-\ln(\varrho))^{m+1}} \frac{m^{m+1} - (-\ln(\varrho))^{m+1}}{m + \ln(\varrho)} \ .$$

We may extend Theorem 2.1 to symmetric functions h which are not canonical. For any integer p and any $\mu_1, \ldots, \mu_p$, p (signed) finite measures on $(\mathsf{Y}, \mathcal{Y})$, denote by $\mu_1 \otimes \cdots \otimes \mu_p \stackrel{\mathrm{def}}{=} \bigotimes_{i=1}^{p} \mu_i$, the product measure on $(\mathsf{Y}^p, \mathcal{Y}^{\otimes p})$. For μ a (signed) finite measure on $(\mathsf{Y}, \mathcal{Y})$, define $\mu^{\otimes p} \stackrel{\mathrm{def}}{=} \mu \otimes \cdots \otimes \mu$.

Let $h : \mathsf{Y}^m \to \mathbb{R}$ be a measurable and symmetric function such that $\pi^{\otimes m}(|h|) < \infty$. Define for any $c \in \{1, \ldots, m-1\}$ the measurable function $\pi_{c,m}h : \mathsf{Y}^c \to \mathbb{R}$ given by

$$\pi_{c,m}h(y_1, \ldots, y_c) \stackrel{\text{def}}{=} (\delta_{y_1} - \pi) \otimes \cdots \otimes (\delta_{y_c} - \pi) \otimes \pi^{\otimes(m-c)}[h] \ , \tag{8}$$

where for any $y \in \mathsf{Y}$, δ_y denotes the Dirac mass at y. Set

$$\pi_{0,m}h \stackrel{\text{def}}{=} \pi^{\otimes m}h \quad \text{and} \quad \pi_{m,m}h(y_1, \ldots, y_m) \stackrel{\text{def}}{=} \bigotimes_{i=1}^{m}(\delta_{y_i} - \pi)[h] \ . \tag{9}$$

Note that for any $c \in \{1, \ldots, m\}$ $\pi_{c,m}h$ is a π-canonical function. The Hoeffding decomposition allows to write any U-statistics associated to a symmetric function h as the following sum of canonical U-statistics (see e.g. (Serfling, 1980, p. 178, Lemma A)):

$$U_{n,m}(h) = \sum_{c=0}^{m} \binom{m}{c} U_{n,c}(\pi_{c,m}h) \ , \tag{10}$$

where $U_{n,c}$ is defined in (1) when $c \geq 1$ and $U_{n,0}(f) \stackrel{\text{def}}{=} f$. The symmetric function h is said to be d-degenerated (for $d \in \{0, \ldots, m\}$) if $\pi_{d,m}h \not\equiv 0$ and $\pi_{c,m}h \equiv 0$ for $c \in \{0, \ldots, d-1\}$. By construction, a π-canonical function h is m-degenerated (it is also said "completely degenerated").

Corollary 2.2. *Assume **A1**. Let h be a bounded symmetric $d(h)$-degenerated function. Then*

$$\left\| U_{n,m}(h) - \pi^{\otimes m}h \right\|_{\mu,2} \leq \sqrt{M(\mu, V)} |h|_\infty \sum_{c=d(h)\vee 1}^{m} \binom{m}{c} 2^c \, C_{n,c} n^{-c/2} \ ,$$

where $C_{n,c}$ is defined in (7).

It is possible to extend the previous result to unbounded canonical functions. Define, for any $q \geq 1$,

$$B_q(h) \stackrel{\text{def}}{=} \sup_{(y_1, \ldots, y_m) \in \mathsf{Y}^m} \frac{|h(y_1, \cdots, y_m)|}{\sum_{j=1}^{m} V^{1/q}(y_j)} \ , \tag{11}$$

where V is defined in **A1**. The proof of Corollary 2.3 is given in Section 3.

Corollary 2.3. *Assume **A1**-**A2** and that, for some $p \in [0, \infty)$, $B_{2(p+1)}(h) < \infty$ holds. Then, for any initial probability measure μ on $(\mathsf{Y}, \mathcal{Y})$,*

$$\|U_{n,m}(h)\|_{\mu,2} \leq 2^{m/2} m \sqrt{(2m)!} \; D(p, \mu, V, h) \; \left(\sum_{k=0}^{n} (k+1)^m \left(\rho(k)\right)^{\frac{p}{(p+1)}} \right)^{1/2} \frac{n^{m/2}}{\binom{n}{m}} \ ,$$

where the constant $D(p, \mu, V, h)$ is given by

$$D(p, \mu, V, h) \; \stackrel{\text{def}}{=} \; 2^{\frac{2p+1}{2(p+1)}} \; \left[p^{\frac{1}{p+1}} + p^{-\frac{p}{p+1}} \right]^{1/2} \; \sqrt{M(\mu, V)} \; B_{2(p+1)}(h) \ . \tag{12}$$

Using again the Hoeffding decomposition (10), Corollary 2.3 can be extended to the case when h is d-degenerated for $d \in \{0, \cdots, m-1\}$. Details are left to the reader. When used in combination with explicit ergodicity bounds for Markov chains, Theorem 2.1 and the corollaries can be used to obtain non-asymptotic computable bounds for the variance of U- and V-statistics. As a simple illustration, assume that the transition kernel P is phi-irreducible, aperiodic and that

1. *(Drift condition)* there exist a drift function $V : \mathsf{Y} \to [1, +\infty)$ and constants $1 < b < \infty$, and $\lambda \in (0, 1)$ such that

$$PV \leq \lambda V + b \, .$$

2. *(Minorization condition)* for any $d \geq 1$, the level sets $\{V \leq d\}$ are petite for P.

Then, there exists a probability distribution π such that $\pi P = \pi$ and $\pi(V) \leq b(1-\lambda)^{-1}$. In addition, there exist computable constants $C < \infty$ and $\rho \in (0, 1)$ such that for any probability measures μ, μ' on $(\mathsf{Y}, \mathcal{Y})$ and any $n \geq 0$,

$$\|\mu P^n - \mu' P^n\|_{\mathrm{TV}} \leq C \, \rho^n \left[\mu(V) + \mu'(V)\right] \; ;$$

(see for example Roberts and Rosenthal (2004), Douc et al. (2004) or Baxendale (2005)). Assumption **A1** is thus satisfied with $\rho(k) = C\rho^k$ and we may thus apply Theorem 2.1 to obtain a non-asymptotic bound.

It can also be used to derive limiting theorems for U-statistics of Markov chains. In what follows, as an illustration of our result, we derive a law of large numbers which holds true under conditions which are, to the best of our knowledge, the weakest known so far and more likely pretty close from being optimal.

Theorem 2.4. *Assume **A1** with $\rho(n) = O\left(n^{-r}\right)$ for some $r > 1$. Let $m \geq 1$ and $h : \mathsf{Y}^m \to \mathbb{R}$ be a symmetric function such that for some $\delta > 0$,*

$$\sup_{(y_1, \ldots, y_m) \in \mathsf{Y}^m} \frac{|h(y_1, \ldots, y_m)| (\log^+ |h(y_1, \ldots, y_m)|)^{1+\delta}}{\sum_{i=1}^m V(y_i)} < \infty \, . \tag{13}$$

Then, for any probability measure μ on $(\mathsf{Y}, \mathcal{Y})$ such that $M(\mu, V) < \infty$,

$$\binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} \{h(Y_{i_1}, \ldots, \; Y_{i_m}) - \mathbb{E}_\mu[h(Y_{i_1}, \ldots, \; Y_{i_m})]\} \to 0 \, , \quad \mathbb{P}_\mu\text{-}a.s. \tag{14}$$

when $n \to +\infty$ and

$$\lim_{n \to +\infty} \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} \mathbb{E}_\mu[h(Y_{i_1}, \ldots, \; Y_{i_m})] = \int \pi(\mathrm{d}y_1) \cdots \pi(\mathrm{d}y_m) h(y_1, \cdots, y_m) \, . \tag{15}$$

3. Proof of Theorem 2.1 and Corollary 2.3

For any probability measure μ on $(\mathsf{Y},\mathcal{Y})$, for any positive integer ℓ and any ordered ℓ-uplet $k_0 = 0 \leq k_1 \leq \cdots \leq k_\ell$, consider the probability measure $\mathbb{P}_\mu^{k_1,k_2,\ldots,k_\ell}$ defined for any nonnegative measurable function $f : \mathsf{Y}^\ell \to \mathbb{R}_+$, by

$$\mathbb{P}_\mu^{k_1,k_2,\ldots,k_\ell}(f) \stackrel{\mathrm{def}}{=} \int \cdots \int \mu(\mathrm{d}y_0) \prod_{i=1}^{\ell} P^{k_i - k_{i-1}}(y_{i-1}, \mathrm{d}y_i) f(y_{1:\ell}) , \tag{16}$$

where $y_{1:\ell} \stackrel{\mathrm{def}}{=} (y_1, \ldots, y_\ell)$. Note that, by construction,

$$\mathbb{E}_\mu \left[f(Y_{k_1}, \ldots, Y_{k_\ell}) \right] = \mathbb{P}_\mu^{k_1,k_2,\ldots,k_\ell}(f) .$$

For any positive integer m and any ordered $2m$-uplet $\mathcal{I} = (1 \leq i_1 \leq i_2 \leq \cdots \leq i_{2m})$, we denote for $\ell \in \{1, \ldots, m\}$,

$$j_\ell(\mathcal{I}) \stackrel{\mathrm{def}}{=} \min(i_{2\ell-1} - i_{2\ell-2}, i_{2\ell} - i_{2\ell-1}) , \tag{17}$$

$$j_\star(\mathcal{I}) \stackrel{\mathrm{def}}{=} \max \left[j_1(\mathcal{I}), j_2(\mathcal{I}), \ldots, j_m(\mathcal{I}) \right] , \tag{18}$$

where, by convention, we set $i_0 = 1$. Denote by $\mathcal{B}_+(\mathsf{Y}^{2m})$ the set of nonnegative measurable function $f : \mathsf{Y}^{2m} \to \mathbb{R}_+$. For any probability measure μ on $(\mathsf{Y},\mathcal{Y})$ and any ordered $2m$-uplet $\mathcal{I}$, denote $\mathbb{P}_\mu^{\mathcal{I}} \stackrel{\mathrm{def}}{=} P_\mu^{i_1,\ldots,i_{2m}}$. We consider the probability measure $\tilde{\mathbb{P}}_\mu^{\mathcal{I}} = \tilde{\mathbb{P}}_\mu^{i_1,\ldots,i_{2m}}$ on $(\mathsf{Y}^{2m}, \mathcal{Y}^{\otimes 2m})$ given for $f \in \mathcal{B}_+(\mathsf{Y}^{2m})$ by

$$\tilde{\mathbb{P}}_\mu^{\mathcal{I}}(f) \stackrel{\mathrm{def}}{=} \int \pi(\mathrm{d}y_1) \mathbb{P}_\mu^{i_2,\ldots,i_{2m}}(\mathrm{d}y_{2:2m}) f(y_{1:2m}) , \tag{19}$$

if $\inf \{k \in \{1, \ldots, m\}, j_\star(\mathcal{I}) = j_k(\mathcal{I})\} = 1$ and

$$\tilde{\mathbb{P}}_\mu^{\mathcal{I}}(f) \stackrel{\mathrm{def}}{=} \int \mathbb{P}_\mu^{i_1,\ldots,i_{2\ell-2}}(\mathrm{d}y_{1:2\ell-2}) \pi(\mathrm{d}y_{2\ell-1}) \mathbb{P}_\mu^{i_{2\ell},\ldots,i_{2m}}(\mathrm{d}y_{2\ell:2m}) f(y_{1:2m}) , \tag{20}$$

if $\ell = \inf \{k \in \{1, \ldots, m\}, j_\star(\mathcal{I}) = j_k(\mathcal{I})\} \in \{2, \ldots, m\}$. For any permutation σ on $\{1, \ldots, 2m\}$, define $f_\sigma : \mathsf{Y}^{2m} \to \mathbb{R}$ the function

$$f_\sigma(y_1, \ldots, y_{2m}) \stackrel{\mathrm{def}}{=} h\left(y_{\sigma(1)}, \ldots, y_{\sigma(m)}\right) h\left(y_{\sigma(m+1)}, \ldots, y_{\sigma(2m)}\right) . \tag{21}$$

Since the function h is π-canonical, it follows from the definition of $\tilde{\mathbb{P}}_\mu^{\mathcal{I}}$ that, for any ordered $2m$-uplet $\mathcal{I}$ and any permutation σ,

$$\tilde{\mathbb{P}}_\mu^{\mathcal{I}}(f_\sigma) = 0 . \tag{22}$$

This relation plays a key role in all what follows and is the main motivation for considering the probability measures $\tilde{\mathbb{P}}_\mu^{\mathcal{I}}$.

Proposition 3.1. *Assume **A1**-**A2**. Then, for any probability measure μ, any positive integer n and any ordered $2m$-uplet $\mathcal{I}$ in $\{1, \cdots, n\}$,*

$$\left\| \mathbb{P}_\mu^{\mathcal{I}} - \tilde{\mathbb{P}}_\mu^{\mathcal{I}} \right\|_{\mathrm{TV}} \leq 4\,\rho\left(j_\star(\mathcal{I})\right)\, M(\mu, V)\;, \tag{23}$$

where the sequence $(\rho(n))_{n\in\mathbb{N}}$, $M(\mu, V)$ and $j_\star(\mathcal{I})$ are defined respectively in (2), (5), *and* (18).

Proof. Let $\mathcal{I} = (1 \leq i_1 \leq i_2 \leq \cdots \leq i_{2m} \leq n)$. To simplify the notation, in what follows, the dependence in $\mathcal{I}$ of $j_1, \ldots, j_m$ - defined in (17) - is implicit. Assume first that $j_\star = j_1$.

Let $f \in \mathcal{B}_+(\mathsf{Y}^{2m})$. The definition of (16) implies that

$$\mathbb{P}_\mu^{\mathcal{I}}(f) \stackrel{\text{def}}{=} P_\mu^{i_1,\ldots,i_{2m}}(f) = \int \mu P^{i_1}(\mathrm{d}y_1)\mathbb{P}_{y_1}^{i_2-i_1,\ldots,i_{2m}-i_1}(\mathrm{d}y_{2:m})f(y_{1:2m})\,.$$

Combining this expression with the definition (19) of $\tilde{\mathbb{P}}_\mu^{\mathcal{I}}$ yields

$$\left|\mathbb{P}_\mu^{\mathcal{I}}(f) - \tilde{\mathbb{P}}_\mu^{\mathcal{I}}(f)\right| \leq T_1 + T_2\;, \tag{24}$$

with

$$T_1 \stackrel{\text{def}}{=} \left| \int \left[\mu P^{i_1}(\mathrm{d}y_1) - \pi(\mathrm{d}y_1)\right] \mathbb{P}_\mu^{i_2,\ldots,i_{2m}}(\mathrm{d}y_{2:2m})f(y_{1:2m}) \right|\;,$$
$$T_2 \stackrel{\text{def}}{=} \left| \int \mu P^{i_1}(\mathrm{d}y_1) \left[\mathbb{P}_{y_1}^{i_2-i_1,\ldots,i_{2m}-i_1}(\mathrm{d}y_{2:2m}) - \mathbb{P}_\mu^{i_2,\ldots,i_{2m}}(\mathrm{d}y_{2:2m})\right] f(y_{1:2m}) \right|\,.$$

Consider first T_1. Since $\left|\int \mathbb{P}_\mu^{i_2,\ldots,i_{2m}}(\mathrm{d}y_{2:2m})f(y_{1:2m})\right| \leq |f|_\infty$, **A1** and (5) imply that

$$T_1 \leq \left\|\mu P^{i_1} - \pi\right\|_{\mathrm{TV}} |f|_\infty \leq \rho(i_1)\ \left[\mu(V) + \pi(V)\right]\ |f|_\infty \leq 2\rho(i_1)M(\mu, V)\;,$$

where we have used that $\mu(V) \leq M(\mu, V)$ and $\pi(V) \leq M(\mu, V)$. On the other hand, for any bounded measurable function $g : \mathsf{Y}^{2m-1} \to \mathbb{R}$, and $y \in \mathsf{Y}$,

$$\mathbb{P}_y^{i_2-i_1,\ldots,i_{2m}-i_1}(g) = \int \delta_y(\mathrm{d}y_1)P^{i_2-i_1}(y_1, \mathrm{d}y_2)\mathbb{P}_{y_2}^{i_3-i_2,\ldots,i_{2m}-i_2}(\mathrm{d}y_{3:2m})g(y_{2:2m})$$

and

$$\mathbb{P}_\mu^{i_2,\ldots,i_{2m}}(g) = \int \mu P^{i_1}(\mathrm{d}y_1)P^{i_2-i_1}(y_1, \mathrm{d}y_2)\mathbb{P}_{y_2}^{i_3-i_2,\ldots,i_{2m}-i_2}(\mathrm{d}y_{3:2m})g(y_{2:2m})\,.$$

Therefore, under **A1**,

$$\left|\mathbb{P}_y^{i_2-i_1,\ldots,i_{2m}-i_1}(g) - \mathbb{P}_\mu^{i_2,\ldots,i_{2m}}(g)\right| \leq \rho(i_2 - i_1)\left[V(y) + \mu P^{i_1}(V)\right] |g|_\infty\,.$$

Integrating this bound shows that $T_2 \leq 2\rho(i_2 - i_1)M(\mu, V)|f|_\infty$, where $M(\mu, V)$ is defined in (5). In conclusion we get

$$\left|\mathbb{P}_\mu^{\mathcal{I}}(f) - \tilde{\mathbb{P}}_\mu^{\mathcal{I}}(f)\right| \leq 2\left[\rho(i_2 - i_1) + \rho(i_1)\right]\, M(\mu, V)\, |f|_\infty\,. \tag{25}$$

Assume now that, for some $\ell \in \{2, \ldots, m\}$, $j_\star = j_\ell$. With these notations, for any nonnegative function $f : \mathsf{Y}^{2m} \to \mathbb{R}$,

$$\mathbb{P}^{\mathcal{I}}_{\mu}(f) = \int \mathbb{P}^{i_1,\ldots,i_{2\ell-1}}_{\mu}(\mathrm{d}y_{1:2\ell-1})\mathbb{P}^{i_{2\ell}-i_{2\ell-1},\ldots,i_{2m}-i_{2\ell-1}}_{y_{2\ell-1}}(\mathrm{d}y_{2\ell:2m})\, f\left(y_{1:2m}\right) \;.$$

Combining this expression with the definition (20) of $\tilde{\mathbb{P}}^{\mathcal{I}}_{\mu}$, we get

$$\left|\mathbb{P}^{\mathcal{I}}_{\mu}(f) - \tilde{\mathbb{P}}^{\mathcal{I}}_{\mu}(f)\right| \leq T_1 + T_2 \;, \tag{26}$$

with

$$T_1 = \left|\int \mathbb{P}^{i_1,\ldots,i_{2\ell-2}}_{\mu}(\mathrm{d}y_{1:2\ell-2})[P^{i_{2\ell-1}-i_{2\ell-2}}(y_{2\ell-2}, \mathrm{d}y_{2\ell-1}) - \pi(\mathrm{d}y_{2\ell-1})] \right.$$
$$\left. \times \quad \mathbb{P}^{i_{2\ell},\ldots,i_{2m}}_{\mu}(\mathrm{d}y_{2\ell:2m})f(y_{1:2m})\right| \;,$$

and

$$T_2 = \left|\int \mathbb{P}^{i_1,\ldots,i_{2\ell-1}}_{\mu}(\mathrm{d}y_{1:2\ell-1})\right.$$
$$\left. \times \left[\mathbb{P}^{i_{2\ell}-i_{2\ell-1},\ldots,i_{2m}-i_{2\ell-1}}_{y_{2\ell-1}}(\mathrm{d}y_{2\ell:2m}) - \mathbb{P}^{i_{2\ell},\ldots,i_{2m}}_{\mu}(\mathrm{d}y_{2\ell:2m})\right] f(y_{1:2m})\right| \;.$$

Consider first T_1. Under **A1**, (2), for any $y_{2\ell-2} \in \mathsf{Y}$, and any bounded measurable function $g : \mathsf{Y} \mapsto \mathbb{R}$,

$$\int [P^{i_{2\ell-1}-i_{2\ell-2}}(y_{2\ell-2}, \mathrm{d}y_{2\ell-1}) - \pi(\mathrm{d}y_{2\ell-1})]g(y_{2\ell-1})$$
$$\leq \rho\left(i_{2\ell-1} - i_{2\ell-2}\right)\, \left[V(y_{2\ell-2}) + \pi(V)\right] |g|_\infty \;.$$

Applying this relation with

$$g_{y_{1:2\ell-2}}(y_{2\ell-1}) = \int \cdots \int \mathbb{P}^{i_{2\ell},\ldots,i_{2m}}_{\mu}(\mathrm{d}y_{2\ell:2m})f(y_{1:2\ell-2}, y_{2\ell-1}, y_{2\ell:2m}) \;,$$

and using that, for any $y_{1:2\ell-2} \in \mathsf{Y}^{2\ell-2}$, $|g_{y_{1:2\ell-2}}|_\infty \leq |f|_\infty$, yields to

$$T_1 \leq \rho(i_{2\ell-1} - i_{2\ell-2})\left[\mu P^{i_{2\ell-2}}(V) + \pi(V)\right] |f|_\infty$$
$$\leq 2\rho(i_{2\ell-1} - i_{2\ell-2})M(\mu, V)|f|_\infty \;.$$

Consider now T_2. Note that, for any bounded measurable function $g : \mathsf{Y}^{2m-2\ell+1} \to \mathbb{R}$ that

$$\mathbb{P}^{i_{2\ell}-i_{2\ell-1},\ldots,i_{2m}-i_{2\ell-1}}_{y_{2\ell-1}}(g) = \int P^{i_{2\ell}-i_{2\ell-1}}(y_{2\ell-1}, \mathrm{d}y_{2\ell})$$
$$\times\, \mathbb{P}^{i_{2\ell+1}-i_{2\ell},\ldots,i_{2m}-i_{2\ell}}_{y_{2\ell}}(\mathrm{d}y_{2\ell+1:2m})g(y_{2\ell:2m}),$$

and

$$\mathbb{P}_{\mu}^{i_{2\ell},\dots,i_{2m}}(g) = \int \mu P^{i_{2\ell-1}}(\mathrm{d}y_{2\ell-1}) P^{i_{2\ell}-i_{2\ell-1}}(y_{2\ell-1}, \mathrm{d}y_{2\ell}) \\ \times \mathbb{P}_{y_{2\ell}}^{i_{2\ell+1}-i_{2\ell},\dots,i_{2m}-i_{2\ell}}(\mathrm{d}y_{2\ell+1:2m}) g(y_{2\ell:2m}) \;.$$

Therefore, under **A1**, for any bounded measurable function $g : \mathsf{Y}^{2m-2\ell+1} \to \mathbb{R}$ and $y_{2\ell-1} \in \mathsf{Y}$,

$$\left| \mathbb{P}_{y_{2\ell-1}}^{i_{2\ell}-i_{2\ell-1},\dots,i_{2m}-i_{2\ell-1}}(g) - \mathbb{P}_{\mu}^{i_{2\ell},\dots,i_{2m}}(g) \right| \\ \leq \rho(i_{2\ell} - i_{2\ell-1}) \left[V(y_{2\ell-1}) + M(\mu, V) \right] |g|_\infty \;.$$

Therefore, by integrating this bound with respect to $\mathbb{P}_{\mu}^{i_1,\dots,i_{2\ell-1}}$ yields to the bound

$$T_2 \leq 2\rho(i_{2\ell} - i_{2\ell-1}) M(\mu, V) |f|_\infty \;,$$

which concludes the proof. □

Lemma 3.2. *Let $(\mathsf{X}, \mathcal{X})$ be a measurable space. Let ξ and ξ' be two probability measures on $(\mathsf{X}, \mathcal{X})$ and $p \in [0, +\infty)$. Then, for any measurable function f satisfying $\xi(|f|^{1+p}) + \xi'(|f|^{1+p}) < \infty$,*

$$|\xi(f) - \xi'(f)| \leq C(p) \left[\xi(|f|^{1+p}) + \xi'(|f|^{1+p}) \right]^{1/(p+1)} \|\xi - \xi'\|_{\mathrm{TV}}^{p/(p+1)} \;,$$

where $C(p) \stackrel{\mathrm{def}}{=} \left[p^{1/(p+1)} + p^{-p/(p+1)} \right]$.

Proof. For any $M > 0$,

$$|\xi(f) - \xi'(f)| \leq M \|\xi - \xi'\|_{\mathrm{TV}} |f|_\infty + \xi \left[|f| \mathbb{1}\{|f| \geq M\} \right] + \xi' \left[|f| \mathbb{1}\{|f| \geq M\} \right] \\ \leq M \|\xi - \xi'\|_{\mathrm{TV}} |f|_\infty + M^{-p} \left[\xi(|f|^{1+p}) + \xi'(|f|^{1+p}) \right] \;.$$

The proof follows by optimizing in M. □

Proposition 3.3. *Assume **A1-A2**. Then, for any ordered $2m$-uplet $\mathcal{I} = (1 \leq i_1 \leq \cdots \leq i_{2m} \leq n)$, any permutation σ on $\{1, \dots, 2m\}$, and any initial distribution μ on $(\mathsf{Y}, \mathcal{Y})$,*

$$|\mathbb{E}_\mu \left[f_\sigma(Y_{i_1}, \dots, Y_{i_{2m}}) \right]| \leq 4M(\mu, V) \, \rho\left(j_\star(\mathcal{I}) \right) \, |h|_\infty^2 \;, \tag{27}$$

where the sequence $(\rho(n))_{n \in \mathbb{N}}$, the index $j_\star(\mathcal{I})$ and the function f_σ are defined in (2), (18), and (21), respectively. If, for some $p \in [0, \infty)$, the constant $B_{2(p+1)}(h)$, defined in (11) is finite, then

$$|\mathbb{E}_\mu \left[f_\sigma(Y_{i_1}, \dots, Y_{i_{2m}}) \right]| \leq m^2 \; D(p, \mu, V, h)^2 \left(\rho\left(j_\star(\mathcal{I}) \right) \right)^{\frac{p}{(p+1)}} \tag{28}$$

where the constant $D(p, \mu, V, h)$ is defined in (12).

Proof. The proof of (27) follows immediately from (22) and Proposition 3.1.

By applying the inequality $ab \leq 1/2(a^2 + b^2)$ and the Jensen inequality, it follows from **A2** that

$$|f_\sigma(y_1, \ldots, y_{2m})|^{p+1} \leq (1/2)\, B_{2(p+1)}^{2(p+1)}(h)\, m^{2p+1} \sum_{i=1}^{2m} V(y_i) \ ,$$

where f_σ is defined in (21). Therefore, for any ordered $2m$-uplet $\mathcal{I} = (1 \leq i_1 \leq \cdots \leq i_{2m} \leq n)$,

$$\mathbb{P}_\mu^{\mathcal{I}}\left[|f_\sigma|^{p+1}\right] \leq M(\mu, V) B_{2(p+1)}^{2(p+1)}(h)\, m^{2(p+1)} \ , \tag{29}$$

$$\tilde{\mathbb{P}}_\mu^{\mathcal{I}}\left[|f_\sigma|^{p+1}\right] \leq M(\mu, V) B_{2(p+1)}^{2(p+1)}(h)\, m^{2(p+1)} \ . \tag{30}$$

The proof then follows by using (22) and by applying Proposition 3.1 and Lemma 3.2. □

Proof of Theorem 2.1 and Corollary 2.3. Denote by $\Gamma(2m)$ the collection of all permutations of $2m$ elements. We have

$$\mathbb{E}_\mu\left[\left(\sum_{1 \leq i_1 < \cdots < i_m \leq n} h(Y_{i_1}, \ldots, Y_{i_m})\right)^2\right] \leq$$
$$\sum_{\sigma \in \Gamma(2m)} \sum_{1 \leq i_1 \leq \cdots \leq i_{2m} \leq n} \left|\mathbb{E}_\mu\left(h(Y_{i_{\sigma(1)}}, \ldots, Y_{i_{\sigma(m)}}) h(Y_{i_{\sigma(m+1)}}, \ldots, Y_{i_{\sigma(2m)}})\right)\right| \ .$$

Let $k \geq 0$. Denote by $\mathsf{I}_{m,n}^k$ the set of all ordered $2m$-uplet $\mathcal{I} = (1 \leq i_1 \leq \cdots \leq i_{2m} \leq n)$ such that $j_\star(\mathcal{I}) = k$, where $j_\star(\mathcal{I})$ is defined in (18). By definition, for $\mathcal{I} \in \mathsf{I}_{m,n}^k$, and $\ell \in \{1, \ldots, m\}$, $j_\ell(\mathcal{I}) \leq k$. It is easily seen that the cardinal of $\mathsf{I}_{m,n}^k$ is at most $2^m n^m (k+1)^m$. The proof of Theorem 2.1 follows from Proposition 3.3, (27).

The proof of Corollary 2.3 follows from Proposition 3.3, (28). □

4. Proof of Theorem 2.4

We will use the following elementary Lemma.

Lemma 4.1. *Let $(s_n)_{n \in \mathbb{N}}$ be a non-decreasing sequence of real numbers. Let $(u_n)_{n \in \mathbb{N}}$ be a non-decreasing sequence of positive numbers. Assume that*

- *the sequence $(\ln(u_n)/\ln(n))_{n \in \mathbb{N}}$ converges to a positive limit δ.*
- *for any $\alpha > 1$, the sequence $\left(u_{\lfloor \alpha^n \rfloor}^{-1} s_{\lfloor \alpha^n \rfloor}\right)_{n \in \mathbb{N}}$ converges to L.*

Then, the sequence $\left(u_n^{-1} s_n\right)_{n \in \mathbb{N}}$ converges to L.

Proof. Let $\alpha > 1$. For any $n \in \mathbb{N}$, denote by $k_n \stackrel{\text{def}}{=} \sup\{k \in \mathbb{N}, \lfloor \alpha^k \rfloor \leq n\}$. Since the sequences $(s_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ are non decreasing and $u_n > 0$ for any $n \in \mathbb{N}$,

$$\frac{u_{\lfloor \alpha^{k_n} \rfloor}}{u_{\lfloor \alpha^{k_n+1} \rfloor}} \frac{s_{\lfloor \alpha^{k_n} \rfloor}}{u_{\lfloor \alpha^{k_n} \rfloor}} \leq \frac{s_n}{u_n} \leq \frac{u_{\lfloor \alpha^{k_n+1} \rfloor}}{u_{\lfloor \alpha^{k_n} \rfloor}} \frac{s_{\lfloor \alpha^{k_n+1} \rfloor}}{u_{\lfloor \alpha^{k_n+1} \rfloor}} .$$

Since $\lim_{n\to\infty} u_{\lfloor \alpha^{n+1} \rfloor}/u_{\lfloor \alpha^n \rfloor} = \alpha^\delta$,

$$\frac{1}{\alpha^\delta} L \leq \liminf_n \frac{s_n}{u_n} \leq \limsup_n \frac{s_n}{u_n} \leq \alpha^\delta L$$

□

Proof of Theorem 2.4. Note that the positive and negative parts of h satisfy the conditions of Theorem 2.4 so that we can assume without loss of generality that h is non negative.

Proof of (15). For any $\tau > 0$, denote

$$h_\tau(y_1, \ldots, y_m) \stackrel{\text{def}}{=} h(y_1, \ldots, y_m) \mathbb{1}_{\{|h(y_1,\ldots,y_m)| \leq \tau\}} .$$

By A1, we have, for any $1 \leq i_1 < \cdots < i_m \leq n$,

$$\left| \mathbb{E}_\mu \left[h_\tau(Y_{i_1}, \ldots, Y_{i_m}) \right] - \pi^{\otimes m}[h_\tau] \right| \leq 2M(\mu, V) |h_\tau|_\infty \sum_{j=1}^m \rho(i_j - i_{j-1}) ,$$

where by convention, $i_0 = 0$. Note that $\sum_{1 \leq i_1 < i_2 \leq n} \rho(i_2 - i_1) = \sum_{k=1}^{n-1} (n - k)\rho(k) \leq n \sum_{k=1}^{n-1} \rho(k)$. Therefore,

$$\begin{aligned}
&\left| \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} \mathbb{E}_\mu \left[h_\tau(Y_{i_1}, \ldots, Y_{i_m}) \right] - \pi^{\otimes m}[h_\tau] \right| \\
&\leq 2M(\mu, V)\tau \sum_{j=1}^m \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} \rho(i_j - i_{j-1}) \\
&\leq 2M(\mu, V)\tau \sum_{j=1}^m \binom{n}{m}^{-1} n^{m-2} \sum_{1 \leq i_{j-1} < i_j \leq n} \rho(i_j - i_{j-1}) \\
&\leq 2M(\mu, V)\tau \sum_{j=1}^m \binom{n}{m}^{-1} n^m n^{-1} \sum_{k=1}^n \rho(k) ,
\end{aligned}$$

which goes to zero since $n^{-1} \sum_{k=1}^n \rho(k) \to 0$. Under the stated assumptions, there exists a constant C such that

$$\mathbb{E}_\mu \left[|h(Y_{i_{1:m}})| \mathbb{1}_{\{|h(Y_{i_{1:m}})| \geq \tau\}} \right] \leq C \left(\log^+ \tau \right)^{-(1+\delta)} .$$

Since $\lim_{\tau\to\infty} \pi^{\otimes m}[h_\tau] = \pi^{\otimes m}[h]$, the proof follows. □

Proof of (14). Let $m \geq 1$ be fixed. We prove that

$$\lim_n \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} h(Y_{i_{1:m}}) = \pi^{\otimes m}[h] \,, \quad \mathbb{P}_\mu - \text{a.s.} \tag{31}$$

Using Lemma 4.1, we have to prove that (31) holds if for any $\alpha > 1$,

$$\lim_{k \to +\infty} \binom{\phi_k}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq \phi_k} h(Y_{i_{1:m}}) = \pi^{\otimes m}[h] \quad \mathbb{P}_\mu - \text{a.s.} \tag{32}$$

where $\phi_k \stackrel{\text{def}}{=} \lfloor \alpha^k \rfloor$. By the Hoeffding decomposition (10), it suffices to prove that for any $c \in \{1, \cdots, m\}$,

$$\lim_{k \to +\infty} \binom{\phi_k}{c}^{-1} \sum_{1 \leq i_1 < \cdots < i_c \leq \phi_k} \pi_{c,m} h(Y_{i_{1:c}}) \xrightarrow{\text{a.s.}} 0$$

where $\pi_{c,m}h$ is the symmetric π-canonical function defined in (8); note that under (13),

$$\sup_{(y_1, \cdots, y_c) \in \mathsf{Y}^c} \frac{|\pi_{c,m}h(y_1, \cdots, y_c)| \ \log^+(|\pi_{c,m}h(y_1, \cdots, y_c)|)^{1+\delta}}{\sum_{i=1}^c V(y_i)} < +\infty \,. \tag{33}$$

The case $c = 1$ is the ergodic theorem for Markov Chain (see for example (Meyn and Tweedie, 2009, Theorem 17.1.7)).

We consider now the case $c \in \{2, \cdots, m\}$. In all what follows, the index $c \in \{2, \ldots, m\}$ is given and for ease of notations, we denote by g an arbitrary π-canonical symmetric function of c variables . Take $s > 0$ such that

$$2s < r - 1 \,. \tag{34}$$

By A1 and (33), there exists a constant C depending upon s and $M(\mu, V)$, such that

$$\mathbb{E}_\mu \left[\sum_{k=1}^{\infty} \phi_k^{-c} \sum_{1 \leq i_1 < \cdots < i_c \leq \phi_k} |g(Y_{i_{1:c}})| \mathbb{1}_{\{|g(Y_{i_{1:c}})| \geq \phi_k^s\}} \right] \leq C \sum_{k=1}^{\infty} (\log \phi_k)^{-\delta - 1} \,, \tag{35}$$

and the RHS is finite since $\alpha > 1$ and $\delta > 0$. Therefore,

$$\phi_k^{-c} \sum_{1 \leq i_1 < \cdots < i_c \leq \phi_k} g(Y_{i_{1:c}}) \mathbb{1}_{\{|g(Y_{i_{1:c}})| \geq \phi_k^s\}} \to 0 \quad \mathbb{P}_\mu - \text{a.s.} \,. \tag{36}$$

We must now prove that

$$\lim_k \phi_k^{-c} \sum_{1 \leq i_1 < \cdots < i_c \leq \phi_k} g_{\phi_k^s}(Y_{i_{1:c}}) = 0 \,, \quad \mathbb{P}_\mu - \text{a.s.} \,, \tag{37}$$

where for $\tau > 0$, $g_\tau(y_{1:c}) \stackrel{\text{def}}{=} g(y_{1:c})\mathbb{1}_{\{|g(y_{1:c})|<\tau\}}$. We apply again the Hoeffding decomposition (10) to the function $g_{\phi_k^s}$. Observe that since g is π-canonical, satisfies (33) and $\pi(V) < +\infty$, the dominated convergence theorem implies that $\lim_k \pi^{\otimes c}(g_{\phi_k^s}) = \pi^{\otimes c}(g) = 0$. Hence, by (10), the limit (36) holds provided for any $\ell \in \{1, \cdots, c\}$,

$$\lim_{k\to\infty} \phi_k^{-\ell} \sum_{1\le i_1<\cdots<i_\ell\le\phi_k} \pi_{\ell,c}[g_{\phi_k^s}](Y_{i_{1:\ell}}) = 0\,, \quad \mathbb{P}_\mu - \text{a.s.}\ . \tag{38}$$

Since g is π canonical, for $\ell \in \{1, \cdots, c-1\}$, we have $\pi_{\ell,c} g = 0$ which implies

$$\pi_{\ell,c}[g_{\phi_k^s}] = \pi_{\ell,c}\left[g - g\mathbb{1}_{\{|g|\ge\phi_k^s\}}\right] = -\pi_{\ell,c}\left[g\mathbb{1}_{\{|g|\ge\phi_k^s\}}\right]\ .$$

Therefore, (38) is equivalent to

$$\lim_{k\to\infty} \phi_k^{-\ell} \sum_{1\le i_1<\cdots<i_\ell\le\phi_k} \pi_{\ell,c}[g\mathbb{1}_{\{|g|\ge\phi_k^s\}}] = 0\,, \quad \mathbb{P}_\mu - \text{a.s.}\ ,$$

which holds true by using an argument similar to (35); details are omitted.

When $\ell = c$, by definition of $\pi_{c,c}$ (see (9)) we have by applying Theorem 2.1

$$\mathbb{E}_\mu\left[\left(\phi_k^{-c} \sum_{1\le i_1<\cdots<i_c\le\phi_k} \pi_{c,c}[g_{\phi_k^s}](Y_{i_{1:c}})\right)^2\right] \\ \le C\ \phi_k^{-c}\left(\sum_{j=0}^{\phi_k}(j+1)^c\rho(j)\right)\phi_k^{2s} \le C'\ \phi_k^{1-r+2s},$$

which by (34) implies (38) when $\ell = c$. This concludes the proof.

□

□

References

Arcones, M. A. (1998). The law of large numbers for U-statistics under absolute regularity. *Electron. Comm. Probab. 3*, 13–19 (electronic).

Baxendale, P. H. (2005). Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab. 15*(1A), 700–738.

Borovkova, S., R. Burton, and H. Dehling (2001). Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Trans. Amer. Math. Soc. 353*(11), 4261–4318.

Dehling, H. (2006). Limit theorems for dependent U-statistics. In *Dependence in probability and statistics*, Volume 187 of *Lecture Notes in Statist.*, pp. 65–86. New York: Springer.

Douc, R., E. Moulines, and J. Rosenthal (2004). Quantitative bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab. 14*(4), 1643–1665.

Elharfaoui, E. and M. Harel (2008). Central limit theorem of the smoothed empirical distribution functions for asymptotically stationary absolutely regular stochastic processes. *J. Appl. Math. Stoch. Anal.*, Art. ID 735436, 18.

Harel, M. and M. L. Puri (1990). Weak invariance of generalized U-statistics for nonstationary absolutely regular processes. *Stochastic Process. Appl. 34*(2), 341–360.

Koroljuk, V. S. and Y. V. Borovskich (1994). *Theory of U-statistics*, Volume 273 of *Mathematics and its Applications*. Dordrecht: Kluwer Academic Publishers Group. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors.

Lee, A. J. (1990). *U-statistics*, Volume 110 of *Statistics: Textbooks and Monographs*. New York: Marcel Dekker Inc. Theory and practice.

Meyn, S. P. and R. L. Tweedie (2009). *Markov Chains and Stochastic Stability*. London: Cambridge University Press.

Roberts, G. O. and J. S. Rosenthal (2004). General state space Markov chains and MCMC algorithms. *Probab. Surv. 1*, 20–71.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley.