Limit Theorems in Change-Point Analysis for Dependent Data

INAUGURAL-DISSERTATION
zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Alexander Schmitz

aus Lindlar

Köln 2011

Erster Berichterstatter:	Prof. Dr. Josef G. Steinebach
Zweiter Berichterstatter:	Prof. Dr. Wolfgang Wefelmeyer
Dritter Berichterstatter:	Prof. Dr. Lajos Horváth

Tag der mündlichen Prüfung:	18. Mai 2011

Abstract. This thesis concerns dependence issues arising from nonparametric change-point analysis based on weighted approximations. We will establish new approximation results under strong mixing conditions. Based on coupling methods, approximations for weighted tied-down partial sum processes by standardized Brownian bridge processes will be derived. Moreover, we will present some new "backward" strong invariance principles for linear processes with strongly mixing errors. As a consequence, we are able to establish Darling-Erdős type limit theorems for weighted tied-down partial sum processes within a financial time series framework.

Zusammenfassung. Die vorliegende Arbeit behandelt gewisse Probleme, die in der nicht-parametrischen Strukturbruch Analyse, basierend auf gewichteten Approximationen, unter stochastischen Abhängigkeiten auftreten. Es werden unter starken Mischungsbedingungen neue Approximationen entwickelt. Unter Anwendung von Kopplungsmethoden, werden Approximationen für gewichtete "Tied-down" Partialsummenprozesse durch standardisierte Brown'sche Brücken-Prozesse hergeleitet. Zusätzlich werden neue "rückwärts gerichtete" starke Invarianzprinzipien für lineare Prozesse mit stark mischenden Fehlern präsentiert. Unter Annahmen, wie sie in der Analyse von Finanzzeitreihen üblich sind, erhalten wir Darling-Erdős Grenzwertsätze für gewichtete "Tied-down" Partialsummenprozesse.

Contents

Introduction vii

Chapter 1. Strong Mixing Conditions and Time Series 1
1.1. Strongly Mixing Linear Processes 1
1.2. Absolute Regularity and Time Series 6

Chapter 2. Limit Theorems for Weighted Partial Sums 13
2.1. Darling-Erdős Limit Theorems 13
2.2. Limit Theorems via Coupling Methods 31

Chapter 3. Limit Theorems in Change-Point Analysis 53
3.1. Quasi-Likelihood Approach 53
3.2. Asymptotics for Rejection Regions 70

Chapter 4. Strong Approximations for Partial Sums 73
4.1. Approximations of Sums of Independent R.V. 73
4.2. Strong Approximations for Linear Processes 94

Chapter 5. Time-Reversibility and Invariance 115
5.1. Reversed Approximations 115
5.2. Applications in Change-Point Analysis 123

Appendix A. Maximal Inequalities for Mixingales 133

Bibliography 137

Introduction

Suppose that $X_1, \ldots, X_n$ are independent random variables with mean $\mu_i \quad (i = 1, \ldots, n)$ and we are interested to test the "no change in the mean" null hypothesis

$$H_0 \ : \ \mu_1 = \cdots = \mu_n$$

against the "at most one change-point" alternative

$$H_{1n}(k) \ : \ \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_n,$$

where $\mu_1, \ldots, \mu_k$ and $\mu_{k+1}, \ldots, \mu_n$ are unknown. Assuming in a first stage that the possible change-point k is known, the first requirement for an ad hoc decision procedure is the *invariance* under a common shift in location of all observation. Since the chronologically ordered observations split up into two subsamples, such an invariant decision procedure can be based on the standardized difference between the sample mean of the first k observations and the sample mean of the last $n-k$ observations. However, the change-point k is unknown in many applications. Then the maximally selected version of this standardized ad hoc statistic is a reasonable choice for testing H_0 against the alternative $H_A = \cup_{k \in [1,n)} H_{1n}(k)$. One would reject the "no change" hypothesis if the standardized max-type statistic

$$\sigma^{-1} \max_{1 \leq k < n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| S(k) - \frac{k}{n} S(n) \right|$$

is large in a certain sense, where $S(k)$ denotes the k-th partial sum and σ^2 is the common variance. Assuming in a second stage independent and identically distributed random variables, under H_0, it turns out that the asymptotic behavior of this test statistic is also invariant under changes in the underlying distribution. This invariance property is a consequence of Donsker's invariance principle applied to the constrained version of the max-type statistic, i.e., for each fixed value $0 < \varepsilon < 1/2$, we have

$$\sigma^{-1} \max_{\varepsilon n \leq k \leq (1-\varepsilon)n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| S(k) - \frac{k}{n} S(n) \right| \xrightarrow{\mathcal{D}} \sup_{\varepsilon \leq t \leq 1-\varepsilon} \frac{|B(t)|}{\sqrt{t(1-t)}},$$

as $n \to \infty$, where $\{B(t),\ 0 \leq t \leq 1\}$ denotes a Brownian bridge process. Observe that $\varepsilon = 0$ is excluded above due to a local limit theorem for the

Brownian bridge process. This constrained version seems reasonable under the assumption that no early or late change occurs within the sample of chronologically ordered observations. Given a fixed level of significance, the critical value can be derived asymptotically from tail approximations of the limit distribution due to Vostrikova [106] and Miller and Siegmund [83].

But on the other hand the constrained version neglects an increasing amount of early and late observations as the sample size increases. The drawbacks of the constrained statistics are obvious and these to overcome is an area of application for the weighted approximation theory.

Let Z_n be the tied-down partial sum process defined by

$$Z_n(t) = \begin{cases} S\left(\lfloor (n+1)t \rfloor\right) - \frac{\lfloor (n+1)t \rfloor}{n} S(n) & , 0 \leq t < 1; \\ 0 & , t = 1. \end{cases}$$

In order to detect early and also late changes, one would choose a nonnegative weight function $q(t)$ on $(0,1)$ that increases in a neighborhood of zero and decreases in a neighborhood of one. Szyszkowicz [102] established under a finite second moment assumption weighted sup-norm versions of Donsker's theorem for the tied-down partial sum process. Using strong approximation results from Major [78], she also obtained convergence-in-probability versions, that is, a construction of Brownian bridge processes, such that

$$\sup_{0<t<1} \left| \frac{Z_n(t)}{\sqrt{n}} - \sigma B_n(t) \right| \Big/ q(t) = o_P(1) \quad (n \to \infty)$$

holds if and only if

$$\int_0^1 \frac{1}{t(1-t)} \exp\left(-\frac{cq^2(t)}{t(1-t)} \right) dt < \infty \quad \text{for all} \quad c > 0.$$

With a view towards our further dependent data studies in this thesis, it is important to notice that the construction above can be viewed as a *coupling method*: The sequence $\{X_n,\ n \geq 1\}$ is redefined, without changing its distribution, together with a sequence of Brownian bridge processes $\{B_n(t),\ 0 \leq t \leq 1\}$ on a *common* probability space $(\Omega, \mathcal{F}, P)$,

The weight function $q(t) = (t(1-t))^{1/2}$, which corresponds to the standardized tied-down partial sum statistic, is still excluded above. Csörgő and Horváth [24] established approximations by standardized Brownian bridge processes of the form

$$\left| \sup_{\frac{1}{n} \leq t \leq 1 - \frac{1}{n}} \frac{|B_n(t)|}{(t(1-t))^{1/2}} - \sigma^{-1} \sup_{\frac{1}{n} \leq t \leq 1 - \frac{1}{n}} \left(\frac{n}{nt(n-nt)} \right)^{1/2} |Z_n(t)| \right|$$

and derived, via using Strassen's invariance principle for the law of the iterated logarithm, the approximation rate $o_P((\log \log n)^{1/2})$, as $n \to \infty$. Moreover, assuming only slightly stronger moment conditions, they strengthened the approximation rate considerably. From these approximation results they derived further asymptotics for critical rejection regions which are designed for small sample sizes. Furthermore, in light of the classical invariance-principle-based results by Darling and Erdős [27], the following extreme value asymptotic holds: Let E and E' be independent random variables satisfying $P[E \leq y] = P[E' \leq y] = \exp\{-\exp(-y)\}$ for each real y and consider the functions

$$A(x) = (2\log x)^{1/2} \text{ and } D(x) = 2\log x + \frac{1}{2}\log\log x - \frac{1}{2}\log\pi,$$

then, as $n \to \infty$,

$$A(\log n) \sup_{\frac{1}{n}\leq t\leq 1-\frac{1}{n}} \left(\frac{n}{\sigma^2 nt(n-nt)}\right)^{1/2} |Z_n(t)| - D(\log n) \xrightarrow{\mathcal{D}} E \vee E'.$$

The Darling-Erdős type limit theorems for the standardized tied-down partial sum process opened the scope for a variety of applications in change-point analysis. Especially in the dependent data context, researchers focus on the further development of these methods towards tests for structural breaks in time series models. In Chapter 1 we will present an overview of the related dependence concepts.

Since the Darling-Erdős type limit theorems are mainly derived via invariance-principle-based techniques, the extensions to strong mixing conditions are numerous and find applications in change-analysis of temperature data, river flow data and financial time series. Although the standardized Brownian bridge type approximations have promising features with respect to small sample sizes, these kind of approximations are regarded less in the literature. The main reasons are the involved constructions which rely heavily on independence assumptions. Using *coupling methods*, we will develop throughout Chapter 2 and Chapter 3 these approximation results within a strong mixing framework.

Considering dependence conditions beyond strong mixing conditions, it turns out that even the standard strong-invariance-principle-based approach may fail to establish Darling-Erdős type limit theorems for weighted tied-down partial sum processes. These problems typically arise when dealing with linear processes with dependent errors. Using *representation* and *coupling methods*, we will establish throughout Chapter 4 and Chapter 5 new "backward" invariance principles to cope the difficulties. Finally, we will discuss further applications within a financial time series context.

Acknowledgements. For help and advice, I thank my teacher Professor Josef G. Steinebach. This work is dedicated to my parents, Rosita and Josef Schmitz, for their support through all the years.

CHAPTER 1

Strong Mixing Conditions and Time Series

In the first section we will discuss the strong mixing property of linear processes. The topics of the second section concern the absolute regularity condition and geometric ergodicity of time series models allowing for conditional heteroscedasticity.

1.1. Strongly Mixing Linear Processes

Let $\mathcal{B}$ denote the Borel σ-field on the real line $\mathbb{R}$. Throughout all chapters let $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$. Let $\{X_k,\ k \geq 1\}$ be a one-sided sequence of real-valued random variables. Without changing its probability distribution the one-sided sequence can be redefined on $\left(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, P\right)$, where the measure P is constructed via Kolmogorov's existence theorem so that for each $\omega = \{\omega_k,\ k \geq 1\} \in \mathbb{R}^{\mathbb{N}}$ the projections $X_k(\omega) = \omega_k$ have the right (joint) distribution. Suppose the one-sided sequence is stationary. Then the unilateral shift $\{\omega_k,\ k \geq 1\} \mapsto \{\omega_{k+1},\ k \geq 1\}$ is measure preserving and the sequence is called ergodic if the unilateral shift is ergodic. Similarly, a one-sided *stationary* sequence can be embedded in a two-sided stationary sequence on the probability space $\left(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}}, P\right)$ and the bilateral shift

$$T : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}} \quad \{\omega_k,\ k \in \mathbb{Z}\} \mapsto \{\omega_{k+1},\ k \in \mathbb{Z}\}$$

is a measure preserving Borel isomorphism. The two-sided extension is called ergodic if the bilateral shift T is ergodic. The mapping T is ergodic if and only if for all A and $B \in \mathcal{B}^{\mathbb{Z}}$

$$n^{-1} \sum_{k=1}^{n} P\left(A \cap T^{-k} B\right) \to P(A)\, P(B) \quad (n \to \infty),$$

cf. e.g. Rosenblatt [91, p. 95, Corollary 4]. Moreover, according to Billingsley [10, Problem 24.2], it suffices to consider only sets A and B from the generating π-system of $\mathcal{B}^{\mathbb{Z}}$. That is to say ergodicity depends only on the finite dimensional distributions and the bilateral shift is ergodic if the unilateral shift is ergodic. The stationary two-sided extension is called mixing (in the ergodic sense) if

$$P\left(A \cap T^{-n} B\right) \to P(A)\, P(B) \quad (n \to \infty)$$

for all A and $B \in \mathcal{B}^{\mathbb{Z}}$. Mixing (in the ergodic sense) clearly implies ergodicity.

We now introduce the strong mixing condition due to Rosenblatt [90]. Suppose a probability space $(\Omega, \mathcal{F}, P)$. Let the measure of dependence between two σ-fields $\mathcal{A}$ and $\mathcal{B} \subset \mathcal{F}$ be

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A) P(B)|.$$

Let $\{X_k,\ k \in \mathbb{Z}\}$ be a two-sided sequence of random variables on $(\Omega, \mathcal{A}, P)$. For $-\infty \leq J < L \leq \infty$ define $\mathcal{F}_J^L = \sigma(X_k,\ J \leq k \leq L)$, i.e. the σ-field generated by the family $\{X_k,\ J \leq k \leq L\}$. For each $n \in \mathbb{N}$ define the dependence (mixing) coefficient $\alpha(n)$ by

$$\alpha(n) = \sup_{-\infty < J < \infty} \alpha\left(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^{\infty}\right).$$

The sequence $\{X_k,\ k \in \mathbb{Z}\}$ is said to be strongly mixing (α-mixing) if

$$\lim_{n \to \infty} \alpha(n) = 0.$$

For a one-sided sequence $\{X_k,\ k \geq 1\}$ one can define $\alpha(n)$ by

$$\alpha(n) = \sup_{1 \leq J < \infty} \alpha\left(\mathcal{F}_1^J, \mathcal{F}_{J+n}^{\infty}\right).$$

Suppose a strictly stationary two-sided sequence $\{X_k,\ k \in \mathbb{Z}\}$. The strong mixing condition is satisfied if $\lim_{n \to \infty} \alpha(n) = 0$, where

$$\alpha(n) = \alpha\left(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^{\infty}\right)$$

and the sequence is ergodic. To prove ergodicity, redefine the sequence without changing its distribution on the space $\left(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}}, P\right)$. Using Billingsley [10, Problem 24.2], it suffices to consider only sets A and B from the generating π-system of $\mathcal{B}^{\mathbb{Z}}$ with $A \in \mathcal{F}_{-\infty}^i$ and $B \in \mathcal{F}_j^{\infty}$ for some i and $j \in \mathbb{Z}$. This implies for $n \in \mathbb{N}$ large enough:

$$\left|P\left(A \cap T^{-n} B\right) - P(A) P(B)\right| \leq \alpha(j - i + n).$$

For a converse statement involving the notion of uniform ergodicity we refer to the recent contribution of Bradley [16]. We conclude with a remark claimed in Bradley [13, p. 170], which will be useful in the proof of Lemma 2.1.3.

Remark. *If a one-sided strictly stationary sequence is strongly mixing with mixing coefficients* $\alpha(n)$, *so is the two-sided strictly stationary extension with the same mixing coefficients.*

The proof of the remark is based on the following approximation of measure argument: For a fixed $\epsilon > 0$ and a set $A \in \mathcal{B}^{\mathbb{Z}}$ we can find an approximating set A_ϵ from the generating π-system such that $P(A \Delta A_\epsilon) < \epsilon$.

The shift is measure preserving. Thus we can find an appropriate translation of the approximating set, i.e. there exist k and $\ell \in \mathbb{N}$ such that $T^{-k}A_\epsilon \in \mathcal{F}_0^\ell$ and the mixing property of the one-sided sequence applies.

For more properties of *strong mixing conditions* (plural) we refer to Bradley [14] and the references therein.

The rest of this section is devoted to linear processes. Let $\{\xi_k,\ k \in \mathbb{Z}\}$ be a sequence of independent and identically distributed random variables with $E\xi_1 = 0$ and $0 < E\xi_1^2 = \sigma^2 < \infty$. We define the two-sided linear process by

$$X_n = \sum_{k=-\infty}^{\infty} a_k \xi_{n-k} \quad (n = 0, \pm 1, \pm 2, \dots),$$

where the sequence of real weights $\{a_k,\ k \in \mathbb{Z}\}$ satisfies

$$\sum_{k\in\mathbb{Z}} a_k^2 < \infty.$$

The series converges with probability one by making use of Kolmogorov's series theorem, see [103, Theorem 3.11]. Hence, $\{X_n,\ n \in \mathbb{Z}\}$ is a stationary and ergodic process, see [10, Theorem 36.4].

Suppose that the two-sided linear process is generated by centered normal random variables ξ_k. Then $\{X_n,\ n \in \mathbb{Z}\}$ is a (discrete) Gaussian process, i.e. a stationary process such that the finite-dimensional joint distributions are centered and normal. Since the Gaussian distribution is determined by mean and covariance matrix, the covariance function $\rho(k) = EX_nX_{n+k}$, defined on the integers k, describes the whole process. Moreover, from $\rho(k) = \rho(-k)$ and due to positive semi-definiteness of the covariance function in the Gaussian case, there exists a uniquely determined spectral distribution function F on the circle satisfying

$$\rho(k) = \int_{[-\pi,\pi]} \exp\{ik\nu\} F(d\nu),$$

which follows from Herglotz's theorem, cf. e.g. Brockwell and Davis [18].

The connection between Gaussian processes and linear processes admits a converse statement.

EXAMPLE 1.1. *Consider a discrete Gaussian process. If the spectral distribution function is absolute continuous with spectral density f, then there is a representation of the process as linear process in terms of independent normal random variables.*

The representation follows via using the Fourier coefficients of the square integrable function $\sqrt{f}$ as weights, cf. Varadhan [103, p. 148].

A stationary process with covariance function $\rho(k)$ is called a "short-memory" process if

$$\sum_{k=1}^{\infty} |\rho(k)| < \infty$$

holds. In this case, cf. e.g. [45, Theorem 2.11], there exists a spectral density function and it is given by

$$f(\nu) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \rho(k) \exp\{-ik\nu\}.$$

We refer to [45, Section 2.5] for an introduction to so-called "long-memory" processes. Consider the partial sums $S(n) = \sum_{i=1}^{n} X_i \quad (n = 1, 2, , \dots)$ of a stationary "short-memory" process. Since

$$\operatorname{Var} S(n) = n\rho(0) + 2\sum_{k=2}^{n} (n - k + 1)\rho(k),$$

the so-called "long-run" variance $\lim_{n\to\infty} n^{-1} \operatorname{Var} S(n)$ exists. Moreover, see [72, Chapter 2], if the spectral density is continuous at $\nu = 0$ then

$$\operatorname{Var} S(n) = 2\pi f(0)n + o(n) \quad (n \to \infty).$$

Towards this end, let us consider autoregressive moving average (ARMA) time series. The backshift operator B has the property $BX_n = X_{n-1}$ and $B\xi_n = \xi_{n-1}$. The ARMA(p, q) sequence $\{X_n,\ n \in \mathbb{Z}\}$ is defined as the stationary solution of

$$\phi\,(B)\, X_n = \theta\,(B)\, \xi_n,$$

where $\phi(x)$ and $\theta(x)$ are polynomials of degree p and $q \in \{0, 1, \dots\}$ and the constant term of both polynomials is assumed to be one.

Example 1.2. *If $E \log_+ (|\xi_1|) < \infty$ and $\phi(x)$ has no zeros of absolute value one , then there is a stationary* ARMA *solution. This solution has a representation as linear process and is ergodic.*

Proof. Consider the Laurent expansion of $\frac{\theta(z)}{\phi(z)}$, i.e. there is some $0 < \epsilon < 1$ such that

$$\frac{\theta(z)}{\phi(z)} = \sum_{k=-\infty}^{\infty} a_k z^k, \quad 1 - \epsilon < |z| < 1 + \epsilon.$$

Thus the weights are of geometric order, i.e. $|a_k| = O\left(\rho_1^k\right)$ $(k \to +\infty)$ for some $0 < \rho_1 < 1$. Similarly, $|a_k| = O\left(\rho_2^k\right)$ $(k \to -\infty)$ for some $\rho_2 > 1$. Using Berkes et al. [7, Lemma 2.2], the series $\sum_{k=0}^{\infty} |\xi_{n-k}|\rho_1^k$ and $\sum_{k=0}^{\infty} |\xi_{n+k}|\rho_2^{-k}$ converge with probability one. $\square$

In classical time series analysis, mainly the weak type stationarity conditions are of interest. In particular, given a second order stationary white noise $\{\xi_n,\ n \in \mathbb{Z}\}$, the ARMA equation admits a causal (one-sided) representation in terms of the white noise variables if $\phi(z) \neq 0$ for all complex z satisfying $|z| \leq 1$, see [18, Theorem 3.1.1]. This is called a causal moving average representation. However, under the condition of the last example, as a consequence of the (two-sided) moving average representation, an application of Brockwell and Davis [18, Theorem 4.4.2] yields the spectral density of the ARMA process:

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(\exp\{-i\lambda\})|^2}{|\phi(\exp\{-i\lambda\})|^2} \quad (-\pi \leq \lambda \leq \pi).$$

Compared with the moment condition in the last example to assure a stationary ARMA solution on an independent and identically distributed noise sequence, we need slightly higher moments to establish the strong mixing property for ARMA time series and linear processes, respectively. Consider the linear process

$$X_n = \sum_{k=-\infty}^{\infty} a_k \xi_{n-k} \quad (n = 0, \pm 1, \pm 2, \dots)$$

and assume that there is a constant $\delta > 0$, such that $E|\xi_1|^\delta < \infty$ and if $\delta \geq 1$ that $E\xi_1 = 0$. Let $\rho = \min\{1, \delta\}$ and suppose

$$\sum_{k \in \mathbb{Z}} |a_k|^\rho < \infty.$$

Let the operator A on the space of bounded two-sided sequences (equipped with the sup-norm) defined by

$$(Ax)_i = \sum_{k \in \mathbb{Z}} a_{i-k} x_k.$$

Let us assume that there is a bounded linear operator K such that

$$AK = I.$$

Suppose further that the density of ξ_1 satisfies

$$\int_{\mathbb{R}} |f(y+x) - f(y)|\, dy \leq C_0 |x|,$$

where $C_0 > 0$ a constant. Then the two-sided process is strongly mixing. In particular, in terms of the even mixing coefficients and for k sufficiently large, the following holds:

$$\alpha_X(2k) \leq C_1 \left\{ \sum_{m=k}^{\infty} d_{m,\delta}^{\frac{1}{1+\delta}} \right\} \vee \left\{ \sum_{m=k}^{\infty} L(d_{m,2}) \right\},$$

where $L(u) = \{u[1 \vee |\log u|]\}^{1/2}$ and $d_{m,\mu} = \sum_{|j|>m} |a_j|^\mu$ $(\mu > 0)$ and $C_1 > 0$ is a constant. This result is due to Doukhan [33, Chapter 2.3.1]. Therein, Doukhan generalizes the results of Gorodetskii [52] from one-sided processes to linear random fields on $\mathbb{Z}^d$. The two-sided case stated above follows from the case $d = 1$, see [33, Chapter 2.3.1, Theorem 1]. Rosenblatt [92, Theorem 4.4.1] proved the case $d = 1$ separately. Moreover, Rosenblatt [92, p. 51] pointed out that if the transfer function $\nu \mapsto A(\exp(-i\nu))$ of the linear filter A has no zeros, then there exists a bounded operator K and the invertibility condition, in terms of $AK = I$, is fulfilled. This follows via Wiener's theorem, cf. Zygmund [112, Theorem VI.5.2]. More sufficient conditions are discussed in [33, pp. 76-77].

1.2. Absolute Regularity and Time Series

Let $\{X_k,\ k \in \mathbb{Z}\}$ be a sequence of random variables on $(\Omega, \mathcal{A}, P)$ with values in a polish space $(S, \mathcal{S})$. For $-\infty \le J < L \le \infty$ define $\mathcal{F}_J^L = \sigma(X_k,\ J \le k \le L)$, i.e. the σ-field generated by the family $\{X_k,\ J \le k \le L\}$. The sequence $\{X_k,\ k \in \mathbb{Z}\}$ obeys the *absolute regularity* condition if

$$\beta(n) = \sup_{k \in \mathbb{Z}} E\left[\sup_{A \in \mathcal{F}_{-\infty}^k} \left|P\left(A|\mathcal{F}_{k+n}^\infty\right) - P(A)\right|\right] \to 0$$

as $n \to \infty$. This condition was proposed by A.N. Kolmogorov under the name *strong regularity*, see Volkonskii and Rozanov [104, p. 179]. Philipp [87, p. 231] remarked explicitely that the supremum is measurable since in a Polish space it is sufficient to extend the supremum only over countably many sets A. The absolute regularity condition can be expressed in terms of the total variation distance, cf. Volkonskii and Rozanov [105, Section 4], i.e. let $P^{(k,n)}$ denote the product measure on $\mathcal{F}_{-\infty}^k \otimes \mathcal{F}_{k+n}^\infty$ defined by $P^{(k,n)}(A \times B) = P(A)\,P(B)$ for $A \in \mathcal{F}_{-\infty}^k$, $B \in \mathcal{F}_{k+n}^\infty$ then

$$\beta(n) = \sup_{k \in \mathbb{Z}} \sup_{C \in \mathcal{F}_{-\infty}^0 \otimes \mathcal{F}_n^\infty} \left|P(C) - P^{(k,n)}(C)\right|.$$

We refer to Pollard [88, pp. 59-60] for a definition of the total variation distance and its connection with the total variation norm, see also Elstrodt [44, Satz VII.1.9]. With a view towards coupling methods, Berbee [4, Chapter 4.4] introduced the equivalent total variation norm condition as a measure of dependence for a stationary sequence under the name *weak Bernoulli*. The stationary sequence $\{X_k,\ k \in \mathbb{Z}\}$ is said to be weak Bernoulli if in terms of the total variation norm

$$\beta(n) = \frac{1}{2}\left\|P_{X^0,Y^n} - P_{X^0} \otimes P_{Y^n}\right\|_{TV} \to 0$$

as $n \to \infty$, where $X^0 = (\dots, X_{-1}, X_0)$ and $Y^n = (X_n, X_{n+1}, \dots)$. He also introduced a *Cesaro weak Bernoulli* version. The equivalence of *weak Bernoulli* and *absolute regularity* is proved in Berbee [4, Proposition 4.1.1]. There are several other equivalent characterization , cf. e.g. Doukhan [33] and Bradley [15, Theorem 3.29]. We conclude with two statements of Bulinski and Shashkin [19, Lemma A.17] baring the idea behind the equivalences mentioned so far. Consider two polish spaces $(S_i, \mathcal{S}_i)$ $(i = 1, 2)$ and let ξ and η be random elements on a common probability space, taking their values in S_1 and S_2, respectively. Then the function $\beta : S_2 \to \mathbb{R}$ defined by

$$\beta(y) = \frac{1}{2}\|P_{\xi|\eta=y} - P_\xi\|_{TV}, \quad y \in S_2,$$

is measurable in y. Moreover,

$$\int_{S_2} \beta(y) P_\eta(dy) = \frac{1}{2}\|P_{\xi,\eta} - P_\xi \otimes P_\eta\|_{TV}.$$

Similar to the α-mixing condition, Bradley [13] introduced the β-mixing condition via a measure of dependence between two σ-fields without any a priori separability conditions. Consider the probability space $(\Omega, \mathcal{F}, P)$. Let the measure of dependence between two sub-σ-fields $\mathcal{A}$ and $\mathcal{B} \subset \mathcal{F}$ be

$$\beta(\mathcal{A}, \mathcal{B}) = \sup \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} \left| P(A_i \cap B_j) - P(A_i) P(B_j) \right|,$$

where the supremum ranges over all pairs of partitions $\{A_1, \dots, A_I\}$ and $\{B_1, \dots, B_J\}$ of Ω into finitely measurable sets, such that $A_i \in \mathcal{A}$ for each i and $B_j \in \mathcal{B}$ for each j.

Let $\{X_k,\ k \in \mathbb{Z}\}$ be a two-sided sequence of random variables on $(\Omega, \mathcal{F}, P)$. For $-\infty \le J < L \le \infty$ define $\mathcal{F}_J^L = \sigma(X_k,\ J \le k \le L)$, i.e. the σ-field generated by the family $\{X_k,\ J \le k \le L\}$. For each $n \in \mathbb{N}$ define the dependence (mixing) coefficient $\beta(n)$ by

$$\beta(n) = \sup_{-\infty < J < \infty} \beta\left(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^\infty\right).$$

The sequence $\{X_k,\ k \in \mathbb{Z}\}$ is said to be absolutely regular (β-mixing) if

$$\lim_{n\to\infty} \beta(n) = 0.$$

For a one-sided sequence $\{X_k,\ k \ge 1\}$ one can define $\beta(n)$ by

$$\beta(n) = \sup_{1 \le J < \infty} \beta\left(\mathcal{F}_1^J, \mathcal{F}_{J+n}^\infty\right).$$

Suppose a strictly stationary two-sided sequence $\{X_k,\ k \in \mathbb{Z}\}$. The β-mixing condition is satisfied if $\lim_{n\to\infty} \beta(n) = 0$, where

$$\beta(n) = \beta\left(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty\right)$$

In the case of random variables with values in a polish space the β-mixing condition coincides with the absolute regularity condition, see Bradley [15, Proposition 3.22]. By the definition it follows that β-mixing implies α-mixing. Finally, the following is true.

REMARK. *If a one-sided strictly stationary sequence is β-mixing with mixing coefficients $\beta(n)$, so is the two-sided strictly stationary extension with the same mixing coefficients.*

The next part of this section concerns Markov chains. Consider a time-homogeneous Markov process with state space $\mathcal{X}$ and with stationary (one-step) transition probability function $\pi(\cdot,\cdot)$. For a fixed initial probability measure μ on $(\mathcal{X},\mathcal{F})$ let

$$P_\mu\left(A_0\times\cdots\times A_n\right) = \int_{A_0}\mu\left(dx_0\right)\int_{A_1}\pi\left(x_0,dx_1\right)\cdots\int_{A_{n-1}}\pi\left(x_{n-2},dx_{n-1}\right)\pi\left(x_{n-1},A_n\right)$$

Then the chain can be represented on $\left(\mathcal{X}^{\mathbb{N}},\mathcal{F}^{\mathbb{N}},P_\mu\right)$ with the theorem of Ionescu-Tulcea, cf. e.g. Rosenblatt [91, Appendix 3]. Higher step transition probabilities can be generated from the one-step's recursively via

$$\pi^{(n+1)}\left(x,A\right) = \int \pi^{(n)}\left(x,dy\right)\pi\left(y,A\right).$$

Let us restrict on the real line as state space. The measure μ is called invariant with respect to the transition probability $\pi(\cdot,\cdot)$ if

$$\int \mu(dx)\pi(x,A) = \mu\left(A\right).$$

Then we get a consistent family on finite product sets and via Kolmogorov's existence theorem we can set up the chain on $\left(\mathbb{R}^{\mathbb{Z}},\mathcal{B}^{\mathbb{Z}},P_\mu\right)$. Then the shift operator T is a one-to-one mapping of the product space onto itself and T, T^{-1} are measurable and preserve the measure P_μ. This (two-sided) stationary process is called stationary Markov chain. It is called ergodic if the shift operator is ergodic with respect to P_μ.

As an example consider the autoregressive model

$$X_{n+1} = f\left(X_n\right)+\xi_{n+1},$$

where $\{\xi_k,\ k\in\mathbb{Z}\}$ is a sequence of independent, identically distributed random variables. Then the transition probability distribution is given by the regular conditional distribution, i.e.

$$\pi\left(x_n,(\infty,y]\right) = P\left(X_{n+1}\leq y|X_n=x_n\right)$$

and if G denotes the distribution function of ξ_1 we have

$$P\left(X_{n+1} \leq y | X_n = x_n\right) = G\left(y - f\left(x_n\right)\right).$$

As a special case consider the discrete Ornstein-Uhlenbeck process, that is an autoregression with centered normal errors.

Example 1.3. *Let* $|\rho| < 1$ *and consider*

$$X_{n+1} = \rho X_n + \xi_{n+1}, \quad n = 1, 2, \ldots,$$

where each ξ_n *is a centered normal random variable with variance* σ^2*. The invariant measure is normal with mean zero and variance* $\sigma^2 / \left(1 - \rho^2\right)$*, cf. Varadhan [103, Example 6.1].*

A strictly stationary Markov chain is said to satisfy "geometric ergodicity" if there exist measurable functions $a : \mathbb{R} \to (0, \infty)$ and $c : \mathbb{R} \to (0, \infty)$ such that the following holds for $\mu - a.s.$ $x \in \mathbb{R}$:

$$\left|\pi^{(n)}\left(x, B\right) - \mu\left(B\right)\right| \leq a(x) \exp\left\{-c(x) n\right\}$$

for all $n \in \mathbb{N}$ and $B \in \mathcal{B}$.

A strictly stationary Markov chain satisfies "geometric ergodicity" if and only if the Markov chain is β-mixing with $\beta(n) \to 0$ at least exponentially fast as $n \to \infty$, cf. Bradley [14, Theorem 3.7]. Further references on the connection between ergodicity conditions for Markov chains and mixing conditions can be found in Bradley [15, Chapter 7.19].

Example 1.4. *The discrete Ornstein-Uhlenbeck is β-mixing.*

Proof. By Varadhan [103, Example 6.1] the n-step transition probability $\pi^n(x, \cdot)$ is a normal distribution with mean $\rho^n x$ and variance $\sigma^2 \sum_{k=0}^{n-1} \rho^{2k}$. Since the invariant measure is normal with mean zero and variance $\sigma^2 / \left(1 - \rho^2\right)$, we apply the inequality in [36, p. 83] and derive an absolute constant C such that for every Borel set B

$$\left|P\left(X_n \in B | X_0 = x\right) - \mu(B)\right| \leq C(2 + x) \exp\left\{-|\ln \rho| n\right\}.$$

$\square$

In the last example we are able to compute the invariant measure and check the definition. Observe that in time series models defined by recurrence equations an approach based on one-step transitions would be more convenient. Among others, Meyn and Tweedie [82] developed within an operator-theoretic framework so-called drift criteria for geometric ergodicity. These criteria are defined by the one-step transition function and in terms of so-called test functions. They introduced the drift operator Δ for any non-negative measurable (test) function V by

$$\Delta V(x) := \int \pi(x, dy)V(y) - V(x), \quad x \in \mathcal{X}.$$

Typically, see [82, p. 376], the geometric drift conditions are of the form: There exists a test function $V : \mathcal{X} \to [1, \infty]$, a set $C \subset \mathcal{X}$ and constants $\beta > 0, b < \infty$, such that

$$\Delta V(x) \leq -\beta V(x) + bI_{\{C\}}(x), \quad x \in \mathcal{X}.$$

In applications, one chooses $V(x) = 1 + |x|^s$, where s is determined by the moment assumptions on the process. These criteria are a common approach to check mixing properties for d-dimensional time series given by stochastic recurrence equations, i.e. $X_k = A_{k-1}X_k + B_k$ with random coefficient $d \times d$-matrices A_{k-1}. It turns out that a variety of time-series models can be "embedded" in these kind of stochastic recurrence equation above. We refer to Basrak et al. [3] and Carrasco and Chen [20] for an exposition of stochastic recurrence equation and their (mixing) properties and the embedding of so-called GARCH time series which find applications in modeling of financial time series.

Let $\{\eta_k,\ k \in \mathbb{Z}\}$ be a sequence of independent and identically distributed random variables with mean zero and variance one. The probability distribution of η_1 has a continuous density and the density is positive on the whole real line. Let $\mathcal{F}_{k-1}$ denote the sigma field generated by the family $\{\dots, \eta_{k-2}, \eta_{k-1}\}$. Carrasco and Chen [20] considered the augmented GARCH$(1,1)$ model, introduced by Duan [34] for modeling stochastic volatilities, defined by

$$\begin{cases} \varepsilon_k = \sigma_k \eta_k, \\ \Lambda\left(\sigma_k^2\right) = c\left(e_k\right)\Lambda\left(\sigma_{k-1}^2\right) + g\left(e_k\right), \end{cases}$$

where σ_k is measurable with respect to $\mathcal{F}_{k-1}$ for every $k \in \mathbb{Z}$; $\Lambda(\cdot), c(\cdot)$ and $g(\cdot)$ are continuous real-valued function and each e_k is a measurable function of η_{k-1}, such that

$$|c(0)| < 1, \quad E\left|c\left(e_k\right)\right|^s < 1 \quad \text{and} \quad E\left|g\left(e_k\right)\right|^s < \infty$$

is satisfied for some integer $s \geq 1$. Carrasco and Chen [20] established geometric ergodicity for the Markov process $\{\Lambda\left(\sigma_k\right),\ k \geq 1\}$. Precisely, if $\Lambda(\cdot)$ is increasing and continuous with domain $[0, \infty)$ and σ_0^2 is initialized from the invariant measure, then $\{\sigma_k,\ k \geq 1\}$ and $\{\varepsilon_k,\ k \geq 1\}$ are strictly stationary and β-mixing with exponential decay, cf. [20, Proposition 5].

Aue et al. [1] proved that certain logarithmic moment conditions are necessary and sufficient for the existence of a strictly stationary solution.

Moreover, Francq and Zakoïan [47] established geometric ergodicity under low moment assumptions allowing even for distributions which are a mixture of an absolutely continuous component and finitely many Dirac measures.

CHAPTER 2

Limit Theorems for Weighted Partial Sums

In the first section we will establish further Darling-Erdős type limit theorems for weighted tied-down sums of α-mixing random variables. In the second section we will derive a convergence-in-probability limit theorem for the running maximum of weighted tied-down sums. The proof is based on coupling methods for β-mixing random variables.

2.1. Darling-Erdős Limit Theorems

Darling and Erdős [27, Theorem 1] derived an extreme value distribution for $\max_{1\le k\le n} k^{-1/2}S(k)$, as $n\to\infty$, where $S(k)$ denotes the k-th partial sum of independent and identically distributed random variables with mean zero and variance one.

Following their method of proof, we first consider a sequence of standard normal random variables $\{X_k,\ k\ge 1\}$ and we let $\{V(t),\ t\ge 0\}$ be an Ornstein-Uhlenbeck process, i.e. a stationary Gaussian process with $EV(t)=0$ and $EV(t)V(s)=\exp\left(-\left|t-s\right|/2\right)$. Comparing the covariance functions yields $\left\{k^{-1/2}S(k),\ k\ge 1\right\} \overset{\mathcal{D}}{=} \left\{V\left(\log k\right),\ k\ge 1\right\}$. Darling and Erdős [27, Section 3] established an asymptotic relation for the maximum of the Ornstein-Uhlenbeck process over a discrete set, i.e. $\max_{1\le k\le n} V\left(k\right)$. They also derived a subsequent result for the maximum of the absolute value, i.e. $\max_{1\le k\le n}\left|V\left(k\right)\right|$. Throughout this chapter let

$$A(x)=\left(2\log x\right)^{1/2} \tag{2.1.1}$$

and

$$D^*(x)=2\log x+\frac{1}{2}\log\log x-\frac{1}{2}\log\left(4\pi\right). \tag{2.1.2}$$

Moreover, let E be a random variable with *Gumbel* type extreme value distribution function, i.e.

$$P\left[E\le y\right]=\exp\left\{-\exp\left(-y\right)\right\},\quad -\infty<y<\infty. \tag{2.1.3}$$

The asymptotic relation for the maximum over a discrete set has the following continuous time variant.

$$A\left(T\right)\max_{0\le t\le T} V\left(t\right)-D^*\left(T\right)\overset{\mathcal{D}}{\to} E\quad (T\to\infty). \tag{2.1.4}$$

Csörgő and Révész [26, Remark 1.9.1] pointed out that this limit theorem is implicitly contained in [27]. Nevertheless, it can also be viewed as a particular case of Leadbetter et al. [66, Theorem 12.3.5], where the extreme value asymptotic is established for the class of stationary normal processes with zero mean and covariance function $r(\tau)$ satisfying

$$r(\tau) = 1 - C\,|\tau|^{\alpha} + o\left(|\tau|^{\alpha}\right) \quad (\tau \to 0),$$

for some constants $C > 0$ and $0 < \alpha \leq 2$. Within this class of stationary normal processes, Bickel and Rosenblatt [9, Theorem A1] established the asymptotical independence of

$$A(T) \max_{0 \leq t \leq T} V(t) - D^*(T) \quad \text{and} \quad -A(T) \min_{0 \leq t \leq T} V(t) + D^*(T)$$

as $T \to \infty$. Moreover, they proved

$$-A(T) \min_{0 \leq t \leq T} V(t) + D^*(T) \xrightarrow{\mathcal{D}} E \quad (T \to \infty). \tag{2.1.5}$$

This implies the corresponding limit theorem for the running supremum of the absolute value, cf. Bickel and Rosenblatt [9, Corollary A1]. For another approach, we refer to Horvath [57, Lemma 2.1] for a continuous time variant for the running supremum of the absolute value. Therein a Darling-Erdős type limit theorem for the maximum of the norm of a d-dimensional Ornstein-Uhlenbeck process is established. This result reduces in the one-dimensional case, i.e. $d = 1$, to

$$A(T) \max_{0 \leq t \leq T} |V(t)| - D^*(T) \xrightarrow{\mathcal{D}} E \vee E' \quad (T \to \infty), \tag{2.1.6}$$

where E' and E are independent and identically distributed random variables with *Gumbel* type extreme value distribution.

Having first established the extreme value asymptotics in the Gaussian case, Darling and Erdős [27] proved in a second step, under the assumption of a finite third moment, the following classical limit theorems:

$$A(\log n) \max_{1 \leq k \leq n} \frac{S(k)}{\sqrt{k}} - D^*(\log n) \xrightarrow{\mathcal{D}} E \tag{2.1.7}$$

and

$$A(\log n) \max_{1 \leq k \leq n} \frac{|S(k)|}{\sqrt{k}} - D^*(\log n) \xrightarrow{\mathcal{D}} E \vee E'. \tag{2.1.8}$$

Moreover, they claimed that the corresponding limit theorems for the partial sums can be established assuming less than a finite third moment condition. The optimal moment condition was derived by Einmahl [41] in the case of independent random variables. For an exposition of the invariance-principle-based methods established in [27] and its further developments to self-normalized Darling-Erdős type limit theorem, we refer to Csörgő [22, pp. 535-538].

With a view towards dependent sequences, Einmahl and Mason [43, p. 438] pointed out that Darling-Erdős limit theorems can be derived, whenever the partial sum process $\{S(k),\ k \geq 1\}$ can be redefined on a suitable probability space, together with a standard Wiener process $\{W(t),\ t \geq 0\}$, such that

$$S(n) - W(n) = o\left(n^{1/2} (\log\log n)^{-1/2}\right) \quad a.s. \quad (n \to \infty).$$

On the one hand, in the independent and identically distributed case this rate of approximation is connected with a moment condition which is slightly stronger than a finite second moment. For instance, Breiman [17] established the rate under $E(|X_1| \log_+ \log_+(|X_1|))^2 < \infty$, using the Skorohod representation. Einmahl [40] provided vector-valued refinements based on a different method.

On the other hand, in the dependent case the rate $o(n^{1/2}(\log\log n)^{-1/2})$ allows for only mild restrictions on the strong mixing coefficient. This is our motivation to study the interplay between decay of the mixing coefficients and Darling-Erdős type limit theorem for weighted tied-down partial sums more precisely. We will first employ the following strong approximation result for strongly mixing random variables due to Bradley [12] which imposes only a logarithmic decay of the mixing coefficients.

Assumption B. *Let* $\{X_k,\ k \geq 1\}$ *be a strictly stationary sequence of centered real-valued random variables with*

$$EX_1^2 < \infty \quad \textit{and} \quad \mathit{Var}\, S(n) \to \infty \quad (n \to \infty). \tag{2.1.9}$$

Suppose $\delta > 0$ *and* $\lambda > 1 + 3/\delta$ *are real numbers such that*

$$\sup_{n \in \mathbb{N}} (\mathit{Var}\, S(n))^{-(2+\delta)/2} E\, |S(n)|^{2+\delta} < \infty \tag{2.1.10}$$

and

$$\alpha(n) = o\left((\log n)^{-\lambda}\right) \quad (n \to \infty). \tag{2.1.11}$$

Theorem (see Bradley (1983, Theorem 4)). *If Assumption B holds, then the sequence* $\{S(k),\ k \geq 1\}$ *can be redefined on another probability space, together with a Wiener process* $\{W(t),\ t \geq 0\}$*, such that*

$$S(n) - \sigma W(n) = o\left(n^{1/2} (\log\log n)^{-1/2}\right) \quad a.s. \quad (n \to \infty), \tag{2.1.12}$$

where $0 < \lim_{n\to\infty} n^{-1}\, \mathit{Var}\, S(n) = \sigma^2 < \infty$.

We will follow the approach of Einmahl and Mason [43] and we will establish Darling-Erdős limit theorems for the standardized tied-down partial sum process $\{G_n(k),\ 1 \le k < n\}$ via the strong approximation results in [12]. We introduce the following notation. Let

$$G_n(k) = \left(\frac{n}{k(n-k)} \right)^{1/2} T_n(k), \quad 1 \le k < n, \tag{2.1.13}$$

and

$$T_n(k) = S(k) - \frac{k}{n} S(n), \quad 1 \le k \le n. \tag{2.1.14}$$

From heuristic reasonings it seems clear that the Darling-Erdős limit theorem applies if the maximum of $G_n(k)$ ranges only over the "small" or over the "large" indices k. In the following theorem we derive a general asymptotic to separate the "small" from the "large" area more precisely.

THEOREM 2.1.1. *Suppose that Assumption B holds. Let* $\{c_k,\ k \ge 1\}$ *be a sequence for which* $c_n \le n-1$ *and* $c_n \uparrow \infty \quad (n \to \infty)$. *If*

$$A(\log n) \sqrt{c_n \log\log n} = o\left(\sqrt{n}\right) \quad (n \to \infty) \tag{2.1.15}$$

and

$$\frac{\log\log c_n}{\log\log n} \to 1 \quad (n \to \infty) \tag{2.1.16}$$

are satisfied, then, as $n \to \infty$,

$$A(\log c_n) \frac{1}{\sigma} \max_{1 \le k \le c_n} |G_n(k)| - D^*(\log c_n) \xrightarrow{\mathcal{D}} E \vee E' \tag{2.1.17}$$

and

$$A(\log c_n) \frac{1}{\sigma} \max_{n - c_n \le k \le n-1} |G_n(k)| - D^*(\log c_n) \xrightarrow{\mathcal{D}} E \vee E', \tag{2.1.18}$$

where E' *and* E *are independent and identically Gumbel distributed random variables and* $\lim_{n\to\infty} n^{-1}\, Var\, S(n) = \sigma^2$.

As a consequence of the main result above, we are able to derive the following Darling-Erdős limit theorems.

THEOREM 2.1.2. *If Assumption B holds, then we have, as* $n \to \infty$,

$$A(\log n) \frac{1}{\sigma} \max_{1 \le k \le \frac{n}{\log n}} |G_n(k)| - D^*(\log n) \xrightarrow{\mathcal{D}} E \vee E' \tag{2.1.19}$$

and

$$A(\log n) \frac{1}{\sigma} \max_{n - \frac{n}{\log n} \le k \le n-1} |G_n(k)| - D^*(\log n) \xrightarrow{\mathcal{D}} E \vee E', \tag{2.1.20}$$

where E' *and* E *are independent and identically Gumbel distributed random variables and* $\lim_{n\to\infty} n^{-1}\, Var\, S(n) = \sigma^2$.

As a consequence of the extreme value asymptotic above, we are able to derive the following convergence-in-probability limit theorems.

THEOREM 2.1.3. *Suppose that Assumption B holds. Let* $0 < \epsilon < 1$ *and put* $c_k = \exp(\log k)^\epsilon \quad (k = 1, 2, \dots)$. *Then we have, as* $n \to \infty$,

$$(\log\log n)^{-1/2} \max_{1 \le k \le c_n} |G_n(k)| \xrightarrow{P} \sigma\epsilon^{1/2} \tag{2.1.21}$$

and

$$(\log\log n)^{-1/2} \max_{n-c_n \le k \le n-1} |G_n(k)| \xrightarrow{P} \sigma\epsilon^{1/2}, \tag{2.1.22}$$

where $\lim_{n\to\infty} n^{-1}\, Var\, S(n) = \sigma^2$.

We established certain precise asymptotics if the maximum of $G_n(k)$ ranges only over the "small" or over the "large" indices k. In light of Theorem 2.1.2, it seems natural to study also the asymptotics for the "middle" section. Even at the cost of a polynomial decay of the mixing coefficients, we need considerably stronger rates than in (2.1.12). We will employ the following one-dimensional version of a general result due to Kuelbs and Philipp [65].

ASSUMPTION K. *Let* $\{X_k,\ k \ge 1\}$ *be a strictly stationary sequence of centered real-valued random variables with*

$$EX_1^2 < \infty \quad \text{and} \quad Var\, S(n) \to \infty \quad (n \to \infty). \tag{2.1.23}$$

Suppose $0 < \delta \le 1$ *and* $0 < \epsilon \le 1/4$ *are real numbers such that*

$$E\,|X_1|^{2+\delta} < \infty \tag{2.1.24}$$

and

$$\alpha(n) = O\left(n^{-(1+\epsilon)(1+2/\delta)}\right) \quad (n \to \infty). \tag{2.1.25}$$

THEOREM (see Kuelbs and Philipp (1980, Theorem 4)). *If Assumption K holds, then the sequence* $\{S(k),\ k \ge 1\}$ *can be redefined on another probability space, together with a Wiener process* $\{W(t),\ t \ge 0\}$, *such that*

$$S(n) - \sigma W(n) = O\left(n^{\frac{1}{2}-\lambda}\right) \quad a.s. \quad (n \to \infty) \tag{2.1.26}$$

for some $\lambda > 0$ *and* $0 \le \sigma^2 < \infty$ *together with* $\lim_{n\to\infty} n^{-1}\, Var\, S(n) = \sigma^2$.

In the independent case it was claimed by Horvath [57, display (3.22)] that the maximum of $G_n(k)$, ranging only over the "middle" section as below, is of order $O_P(\log\log\log n)$. Here we state the exact extreme value asymptotic in the strongly mixing case.

THEOREM 2.1.4. *If Assumption K holds, then we have, as* $n \to \infty$,

$$A\left(2\log\left((\log n)-1\right)\right)\frac{1}{\sigma}\max_{\frac{n}{\log n}\le k\le n-\frac{n}{\log n}}|G_n(k)| - D^*\left(2\log\left((\log n)-1\right)\right)\xrightarrow{\mathcal{D}} E\vee E', \quad (2.1.27)$$

where E' *and* E *are independent and identically Gumbel distributed random variables and* $\lim_{n\to\infty} n^{-1}\,Var\,S(n)=\sigma^2$.

As a consequence of Theorem 2.1.4, we are able to derive the extreme value asymptotic for the maximum of $G_n(k)$ ranging over $\{1,\ldots,n-1\}$. The following result is implicitly contained in Horvath [57] and Davis et. al. [32].

THEOREM 2.1.5. *Suppose that Assumption K holds. Let*

$$D(x)=2\log x+\frac{1}{2}\log\log x-\frac{1}{2}\log\pi. \quad (2.1.28)$$

Then we have, as $n\to\infty$,

$$A(\log n)\frac{1}{\sigma}\max_{1\le k\le n-1}|G_n(k)|-D(\log n)\xrightarrow{\mathcal{D}} E\vee E', \quad (2.1.29)$$

where E' *and* E *are independent and identically Gumbel distributed random variables and* $\lim_{n\to\infty} n^{-1}\,Var\,S(n)=\sigma^2$.

The rest of this section concerns the proof of the theorems. The proofs of the main results are based on a series of lemmas. We first derive some uniform approximation results and certain uniform "backward" invariance principles which are of independent interest.

LEMMA 2.1.1. *Suppose that Assumption B holds. Let* $\{c_k,\ k\ge 1\}$ *and* $\{d_k,\ k\ge 1\}$ *be two sequences with* $c_n\le d_n$ *and* $c_n\uparrow\infty$ $(n\to\infty)$. *Then the sequence* $\{S(k),\ k\ge 1\}$ *can be redefined on another probability space, together with a Wiener process* $\{W(t),\ t\ge 0\}$, *such that, as* $n\to\infty$,

$$\max_{c_n\le k\le d_n}\frac{|S(k)-\sigma W(k)|}{\sqrt{k}}=o_P\left((\log\log c_n)^{-1/2}\right), \quad (2.1.30)$$

where $\lim_{n\to\infty} n^{-1}\,Var\,S(n)=\sigma^2$.

PROOF. Since $c_n\le d_n$ and $c_n\uparrow\infty$ $(n\to\infty)$, an application of Bradley [12, Theorem 4] implies

$$\max_{c_n\le k\le d_n}\left(\frac{\log\log k}{k}\right)^{1/2}|S(k)-\sigma W(k)|=o(1) \quad a.s. \quad (n\to\infty).$$

For each $\epsilon>0$, we derive

$$\lim_{n\to\infty}P\left[(\log\log c_n)^{1/2}\max_{c_n\le k\le d_n}\frac{|S(k)-\sigma W(k)|}{\sqrt{k}}>\epsilon\right]=0.$$

□

LEMMA 2.1.2. *Suppose that Assumption B holds. Let* $\{a_k,\ k \geq 1\}$ *and* $\{d_k,\ k \geq 1\}$ *be two sequences with* $0 < a_n \downarrow 0$ *and* $d_n \uparrow \infty$ $(n \to \infty)$. *Then the sequence* $\{S(k),\ k \geq 1\}$ *can be redefined on another probability space, together with a Wiener process* $\{W(t),\ t \geq 0\}$, *such that, as* $n \to \infty$,

$$a_n \max_{1 \leq k \leq d_n} \frac{|S(k) - \sigma W(k)|}{\sqrt{k}} = o_P(1)\,, \tag{2.1.31}$$

where $\lim_{n\to\infty} n^{-1}\, \mathit{Var}\, S(n) = \sigma^2$.

PROOF. Consider the sequence $\{c_k,\ k \geq 1\}$ defined by

$$c_k = \log\left(\min\{d_k, 1/a_k\}\right) \quad (k = 1, 2, \dots)\,.$$

Obviously, we have

$$c_n \leq d_n, \quad c_n \uparrow \infty \quad \text{and} \quad a_n\sqrt{c_n} \to 0 \quad (n \to \infty).$$

For each $\epsilon > 0$, we have

$$P\left[a_n \max_{1 \leq k \leq c_n} |W(k)| > \epsilon\right] \leq \frac{a_n^2 c_n}{\epsilon^2},$$

where Kolmogorov's inequality, cf. e.g. Révész [89, Theorem 2.1.1], was applied. Whence

$$\lim_{n\to\infty} P\left[a_n \max_{1 \leq k \leq c_n} \frac{|W(k)|}{\sqrt{k}} > \epsilon\right] = 0. \tag{2.1.32}$$

Moreover, using the maximal inequality in Bradley [12, display (4.4)], we obtain a nonnegative constant $C(\epsilon)$, such that

$$P\left[a_n \max_{1 \leq k \leq c_n} |S(k)| > \epsilon\right] \leq C(\epsilon)\left(\sqrt{c_n}a_n\right)^{2+\delta}.$$

Thus

$$\lim_{n\to\infty} P\left[a_n \max_{1 \leq k \leq c_n} \frac{|S(k)|}{\sqrt{k}} > \epsilon\right] = 0 \tag{2.1.33}$$

and the assertion flows from (2.1.32), (2.1.33) together with Lemma 2.1.1.

□

LEMMA 2.1.3. *Suppose that Assumption B holds. Let* $\{c_k,\ k \geq 1\}$ *and* $\{d_k,\ k \geq 1\}$ *be two sequences with* $1 \leq c_n \leq d_n \leq n-1$ *and* $n - d_n \uparrow \infty$ $(n \to \infty)$. *Then the sequence* $\{S(k),\ k \geq 1\}$ *can be redefined on another probability space together with a sequence of Wiener process*

$\{W_n(t),\ t \geq 0\}$, *such that, as* $n \to \infty$,

$$\max_{c_n \leq k \leq d_n} \left(\frac{\log\log(n-k)}{n-k} \right)^{1/2} |S(n) - S(k) - \sigma W_n(n-k)| = o_P(1) \tag{2.1.34}$$

where $\lim_{n\to\infty} n^{-1}\,\mathit{Var}\, S(n) = \sigma^2$.

Proof. Extend the strictly stationary process $\{X_k,\ k \geq 1\}$ as two-sided process $\{X'_k,\ -\infty < k < \infty\}$ (say), which is again strictly stationary and strongly mixing with the same mixing coefficients, see Chapter 1.1 above. For each fixed $n \in \mathbb{N}$, using Lemma 2.1.1, we can redefine the one-sided process $\left\{X'_{n+1-\ell},\ \ell \geq 1\right\}$ on another probability space $(\Omega_1, \mathcal{A}_1, P_1)$ as $\left\{X'_{-\ell},\ \ell \geq 1\right\}$ (say), together with one Wiener process, such that for each $\epsilon > 0$,

$$\lim_{n\to\infty} P_1 \left[\max_{n-d_n \leq k \leq n-c_n} \left(\frac{\log\log k}{k} \right)^{1/2} \left| \sum_{\ell=1}^{k} X'_{-\ell} - W(k) \right| > \epsilon \right] = 0.$$

Consider the following law on the Borel sets of the polish space $\mathbb{R}^n \times D[0,n]$

$$\mathcal{L}\left(\left\{ \sum_{\ell=1}^{k} X'_{-\ell},\ 1 \leq k \leq n \right\},\ \{W(t),\ 0 \leq t \leq n\} \right).$$

Since

$$\left\{ \sum_{\ell=1}^{k} X'_{-\ell},\ 1 \leq k \leq n \right\} \stackrel{\mathcal{D}}{=} \left\{ \sum_{\ell=1}^{k} X_{n+1-\ell},\ 1 \leq k \leq n \right\},$$

an application of Lemma 1 in Billingsley [11, Section 21] yields, for each $n \in \mathbb{N}$, a Wiener process $\{W_n(t),\ 0 \leq t \leq n\}$ on the initial probability space (suitably enlarged), such that

$$\left\{ \sum_{\ell=1}^{k} X_{n+1-\ell} - W_n(k),\ 1 \leq k \leq n \right\} \stackrel{\mathcal{D}}{=} \left\{ \sum_{\ell=1}^{k} X'_{-\ell} - W(k),\ 1 \leq k \leq n \right\}.$$

This implies

$$\lim_{n\to\infty} P \left[\max_{n-d_n \leq k \leq n-c_n} \left(\frac{\log\log k}{k} \right)^{1/2} \left| \sum_{\ell=1}^{k} X_{n+1-\ell} - W_n(k) \right| > \epsilon \right] = 0.$$

The assertion follows from the observation

$$\max_{n-d_n \leq k \leq n-c_n} \left(\frac{\log\log k}{k} \right)^{1/2} \left| \sum_{\ell=1}^{k} X_{n+1-\ell} - W_n(k) \right|$$
$$= \max_{c_n \leq k \leq d_n} \left(\frac{\log\log(n-k)}{n-k} \right)^{1/2} \left| \sum_{\ell=1}^{n-k} X_{n+1-\ell} - W_n(n-k) \right|.$$

□

LEMMA 2.1.4. *Suppose that Assumption B holds. Let* $\{a_k,\ k \geq 1\}$ *and* $\{c_k,\ k \geq 1\}$ *be two sequences with* $a_n \downarrow 0$ *and* $n - c_n \uparrow \infty$ $(n \to \infty)$. *Then the sequence* $\{S(k),\ k \geq 1\}$ *can be redefined on another probability space, together with a sequence of Wiener process* $\{W_n(t),\ t \geq 0\}$, *such that, as* $n \to \infty$,

$$a_n \max_{c_n \leq k \leq n-1} \frac{|S(n) - S(k) - \sigma W_n(n-k)|}{\sqrt{n-k}} = o_P(1), \tag{2.1.35}$$

where $\lim_{n\to\infty} n^{-1}\, Var\, S(n) = \sigma^2$.

PROOF. Consider the sequence $\{d_k,\ k \geq 1\}$ defined by

$$d_k = k - \log\left(\min\{k - c_k, 1/a_k\}\right) \quad (k = 1, 2, \dots).$$

Obviously, we have

$$c_n \leq d_n, \quad n - d_n \uparrow \infty \quad \text{and} \quad a_n\sqrt{n - d_n} \to 0 \quad (n \to \infty)$$

By Lemma 2.1.3

$$a_n \max_{c_n \leq k \leq d_n} \frac{|S(n) - S(k) - \sigma W_n(n-k)|}{\sqrt{n-k}} = o_P(1) \quad (n \to \infty). \tag{2.1.36}$$

Using the maximal inequality in Bradley [12, display (4.4)], we have

$$\begin{aligned} &\lim_{n\to\infty} P\left[a_n \max_{d_n \leq k \leq n-1} \frac{|S(n) - S(k)|}{\sqrt{n-k}} > \epsilon\right] \\ &= \lim_{n\to\infty} P\left[a_n \max_{1 \leq k \leq n-d_n} \frac{\left|\sum_{l=1}^{k} X_{-\ell}\right|}{\sqrt{k}} > \epsilon\right] = 0. \end{aligned} \tag{2.1.37}$$

Moreover, via using Kolmogorov's inequality, cf. e.g. Révész [89, Theorem 2.1.1], we have

$$\begin{aligned} &\lim_{n\to\infty} P\left[a_n \max_{d_n \leq k \leq n-1} \frac{|W_n(n-k)|}{\sqrt{n-k}} > \epsilon\right] \\ &= \lim_{n\to\infty} P\left[a_n \max_{1 \leq k \leq n-d_n} \frac{|W(k)|}{\sqrt{k}} > \epsilon\right] = 0. \end{aligned} \tag{2.1.38}$$

The assertion flows from (2.1.36), (2.1.37) and (2.1.38). □

LEMMA 2.1.5. *Suppose that Assumption B holds. Let* $\{a_k,\ k \geq 1\}$ *and* $\{c_k,\ k \geq 1\}$ *be two sequences with* $a_n \uparrow \infty$ *and* $c_n \uparrow \infty$ $(n \to \infty)$. *If*

$$a_n\sqrt{c_n \log\log n} = o\left(\sqrt{n}\right) \quad (n \to \infty), \tag{2.1.39}$$

then

$$a_n \max_{1 \leq k \leq c_n} \left| \frac{\frac{1}{k}S(k) - \frac{1}{n-k}\left(S(n) - S(k)\right)}{\sqrt{\frac{1}{k} + \frac{1}{n-k}}} - \frac{S(k)}{\sqrt{k}} \right| = o_P(1) \tag{2.1.40}$$

as $n \to \infty$.

PROOF. By Lemma 2.1.2, we have

$$\frac{a_n c_n}{n} \max_{1 \le k \le c_n} \frac{|S(k) - W(k)|}{\sqrt{k}} = o_P(1),$$

and via Darling and Erdős [27, Theorem 2], i.e. (2.1.8), we derive, similarly as in Durrett [36, Exercise 2.3], that

$$(2 \log \log c_n)^{-1/2} \max_{1 \le k \le c_n} \frac{|W(k)|}{\sqrt{k}} \xrightarrow{P} 1 \tag{2.1.41}$$

holds, as $n \to \infty$. Since (2.1.39) implies

$$a_n c_n \sqrt{\log \log c_n} = o(n) \quad (n \to \infty),$$

we have

$$\frac{a_n c_n}{n} \max_{1 \le k \le c_n} \frac{|W(k)|}{\sqrt{k}} = o_P(1),$$

which in turn implies

$$\frac{a_n c_n}{n} \max_{1 \le k \le c_n} \frac{|S(k)|}{\sqrt{k}} = o_P(1).$$

Moreover, consider $f(t) = \sqrt{1-t}$ $(0 < t < 1)$. Since $f(0) = 1$, the mean value theorem yields

$$\begin{aligned} \max_{1 \le k \le c_n} \left| \frac{1}{\sqrt{k}\sqrt{\frac{1}{k} + \frac{1}{n-k}}} - 1 \right| &= \max_{1 \le k \le c_n} \left| f\left(\frac{k}{n}\right) - f(0) \right| \\ &\le \frac{c_n}{n} \sup_{0 < s < \frac{c_n}{n}} \frac{1}{2\sqrt{1-s}}. \end{aligned}$$

We arrive at

$$a_n \max_{1 \le k \le c_n} \left| \frac{\frac{1}{k} S(k)}{\sqrt{\frac{1}{k} + \frac{1}{n-k}}} - \frac{1}{\sqrt{k}} S(k) \right| = o_P(1) \quad (n \to \infty). \tag{2.1.42}$$

Towards this end, consider the inequality

$$a_n \max_{1 \le k \le c_n} \frac{1}{\sqrt{\frac{1}{k} + \frac{1}{n-k}}} \frac{|S(n) - S(k)|}{n-k} \le \frac{a_n \sqrt{c_n}}{\sqrt{n}} \max_{1 \le k \le c_n} \frac{|S(n) - S(k)|}{\sqrt{n-k}}$$

together with

$$\frac{a_n \sqrt{c_n}}{\sqrt{n}} \max_{1 \le k \le c_n} \frac{|S(n) - S(k) - W_n(n-k)|}{\sqrt{n-k}} = o_P(1),$$

where Lemma 2.1.4 was applied. Since

$$P\left[\frac{a_n\sqrt{c_n}}{\sqrt{n}}\max_{1\le k\le c_n}\frac{|W_n(n-k)|}{\sqrt{n-k}}>\epsilon\right]$$
$$\le P\left[\max_{n-c_n\le k\le n-1}\frac{|W(k)|}{\sqrt{k\log\log k}}>\frac{\epsilon\sqrt{n}}{a_n\sqrt{c_n\log\log n}}\right],$$

an application of the iterated logarithm via using (2.1.39), yields

$$a_n\max_{1\le k\le c_n}\frac{1}{\sqrt{\frac{1}{k}+\frac{1}{n-k}}}\frac{|S(n)-S(k)|}{n-k}=o_P(1) \tag{2.1.43}$$

as $n\to\infty$. Now (2.1.42) and (2.1.43) imply (2.1.40). □

Lemma 2.1.6. *Suppose that Assumption B holds. Let* $\{a_k,\ k\ge 1\}$ *and* $\{c_k,\ k\ge 1\}$ *be two sequences with* $a_n\uparrow\infty$ *and* $c_n\uparrow\infty$ $(n\to\infty)$. *If*

$$a_n\sqrt{c_n\log\log n}=o\left(\sqrt{n}\right)\quad(n\to\infty), \tag{2.1.44}$$

then

$$a_n\max_{n-c_n\le k\le n-1}\left|\frac{\left|\frac{1}{k}S(k)-\frac{1}{n-k}\left(S(n)-S(k)\right)\right|}{\sqrt{\frac{1}{k}+\frac{1}{n-k}}}-\frac{|S(n)-S(k)|}{\sqrt{n-k}}\right|=o_P\left(1\right) \tag{2.1.45}$$

as $n\to\infty$.

Proof. Consider $f(t)=\sqrt{1-t}$ $(0<t<1)$. Since $f(0)=1$, the mean value theorem yields

$$\max_{n-c_n\le k\le n-1}\left|\frac{1}{\sqrt{n-k}\sqrt{\frac{1}{k}+\frac{1}{n-k}}}-1\right|=\max_{n-c_n\le k\le n-1}\left|\sqrt{\frac{k}{n}}-1\right|$$
$$\le\max_{\frac{1}{n}\le s\le\frac{c_n}{n}}|f(s)-f(0)|$$
$$\le\frac{c_n}{n}\max_{0<s\le\frac{c_n}{n}}\frac{1}{2\sqrt{1-s}}.$$

Therefore

$$a_n\max_{n-c_n\le k\le n-1}\left|\frac{\left|\frac{1}{n-k}\left(S(n)-S(k)\right)\right|}{\sqrt{\frac{1}{k}+\frac{1}{n-k}}}-\frac{|S(n)-S(k)|}{\sqrt{n-k}}\right|$$
$$=O\left(\frac{a_nc_n}{n}\right)\max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}}.$$

By Lemma 2.1.4, we have

$$\frac{a_nc_n}{n}\max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)-W_n(n-k)|}{\sqrt{n-k}}=o_P\left(1\right)$$

and, via using (2.1.41) and (2.1.44), the following is true

$$\lim_{n\to\infty} P\left[\frac{a_n c_n}{n} \max_{n-c_n\le k\le n-1} \frac{|W_n(n-k)|}{\sqrt{n-k}} > \epsilon\right]$$
$$= \lim_{n\to\infty} P\left[\max_{1\le k\le c_n} \frac{|W(k)|}{\sqrt{k}} > \frac{\epsilon n}{a_n c_n}\right] = 0.$$

This implies

$$a_n \max_{n-c_n\le k\le n-1} \left| \frac{\frac{1}{n-k}\,|S(n)-S(k)|}{\sqrt{\frac{1}{k}+\frac{1}{n-k}}} - \frac{|S(n)-S(k)|}{\sqrt{n-k}} \right| = o_P(1) \qquad (2.1.46)$$

as $n\to\infty$. Next, consider the inequality

$$a_n \max_{n-c_n\le k\le n-1} \frac{1}{\sqrt{\frac{1}{k}+\frac{1}{n-k}}} \frac{|S(k)|}{k}$$
$$= a_n \max_{n-c_n\le k\le n-1} \left(1-\frac{k}{n}\right)^{1/2} \frac{|S(k)|}{\sqrt{k}}$$
$$\le \frac{a_n\sqrt{c_n}}{\sqrt{n}} \max_{n-c_n\le k\le n-1} \frac{|S(k)|}{\sqrt{k}}.$$

Observe

$$\frac{a_n\sqrt{c_n}}{\sqrt{n}} \max_{n-c_n\le k\le n-1} \frac{|S(k)-W(k)|}{\sqrt{k}} = o_P(1),$$

where Lemma 2.1.2 was applied. Moreover, for each $\epsilon>0$, the following is true

$$P\left[\frac{a_n\sqrt{c_n}}{\sqrt{n}} \max_{n-c_n\le k\le n-1} \frac{|W(k)|}{\sqrt{k}} > \epsilon\right]$$
$$\le P\left[\frac{a_n\sqrt{c_n}\log\log n}{\sqrt{n}} \max_{n-c_n\le k\le n-1} \frac{|W(k)|}{\sqrt{k\log\log k}} > \epsilon\right],$$

which in turn, via using the law of the iterated logarithm and (2.1.44), yields

$$\frac{a_n\sqrt{c_n}}{\sqrt{n}} \max_{n-c_n\le k\le n-1} \frac{|W(k)|}{\sqrt{k}} = o_P(1).$$

Hence

$$a_n \max_{n-c_n\le k\le n-1} \frac{1}{\sqrt{\frac{1}{k}+\frac{1}{n-k}}} \frac{|S(k)|}{k} = o_P(1) \qquad (2.1.47)$$

as $n\to\infty$. Finally, (2.1.45) flows from (2.1.46) and (2.1.47). $\square$

LEMMA 2.1.7. *Suppose that Assumption B holds. Let* $\{c_k,\ k\ge 1\}$ *be a sequence with* $c_n \uparrow \infty$ $(n\to\infty)$. *If*

$$\frac{\log\log c_n}{\log\log n} \to 1 \quad (n\to\infty), \qquad (2.1.48)$$

then, as $n \to \infty$,

$$A(\log c_n)\,\frac{1}{\sigma}\max_{1\le k\le c_n}\frac{|S(k)|}{\sqrt{k}} - D^*(\log c_n) \xrightarrow{\mathcal{D}} E \vee E' \tag{2.1.49}$$

where E' *and* E *are independent and identically Gumbel distributed random variables and* $\lim_{n\to\infty} n^{-1}\,Var\,S(n) = \sigma^2$.

Proof. By Lemma 2.1.2, we have

$$\max_{1\le k\le c_n}\frac{|S(k)-\sigma W(k)|}{\sqrt{k}} = o\left((\log\log c_n)^{1/2}\right).$$

Therefore, via using (2.1.41) and (2.1.48),

$$(2\log\log n)^{-1/2}\max_{1\le k\le c_n}\frac{|S(k)|}{\sqrt{k}} \xrightarrow{P} \sigma \tag{2.1.50}$$

Let $0<\epsilon<1$ and put $u_n = \exp(\log n)^{\epsilon}$. Obviously,

$$\frac{\log\log u_n}{\log\log n} = \epsilon. \tag{2.1.51}$$

Whence

$$(2\log\log n)^{-1/2}\max_{1\le k\le u_n}\frac{|S(k)|}{\sqrt{k}} \xrightarrow{P} \sigma\sqrt{\epsilon} \tag{2.1.52}$$

as $n\to\infty$. Therefore, from (2.1.50) and (2.1.52),

$$\lim_{n\to\infty} P\left[\max_{1\le k\le u_n}\frac{|S(k)|}{\sqrt{k}} \ge \max_{1\le k\le c_n}\frac{|S(k)|}{\sqrt{k}}\right] = 0 \tag{2.1.53}$$

and similarly, we have

$$\lim_{n\to\infty} P\left[\max_{1\le k\le u_n}\frac{|W(k)|}{\sqrt{k}} \ge \max_{1\le k\le c_n}\frac{|W(k)|}{\sqrt{k}}\right] = 0. \tag{2.1.54}$$

In light of (2.1.51), we can assume $u_n < c_n$. Therefore, via using (2.1.53), (2.1.54) and Lemma 2.1.1, we have

$$\left|\max_{1\le k\le c_n}\frac{|S(k)|}{\sqrt{k}} - \max_{1\le k\le c_n}\frac{\sigma\,|W(k)|}{\sqrt{k}}\right| = o_P\left((\log\log n)^{-1/2}\right)$$

and the assertion flows from Darling and Erdős [27, Theorem 2], i.e. (2.1.8), that is

$$A(\log c_n)\max_{1\le k\le c_n}\frac{|W(k)|}{\sqrt{k}} - D^*(\log c_n) \xrightarrow{\mathcal{D}} E \vee E'$$

as $n\to\infty$. □

Lemma 2.1.8. *Suppose that Assumption B holds. Let* $\{c_k,\ k\ge 1\}$ *be a sequence with* $c_n \le n-1$ *and* $c_n \uparrow \infty$ $(n\to\infty)$. *If*

$$\frac{\log\log c_n}{\log\log n} \to 1 \quad (n\to\infty), \tag{2.1.55}$$

then, as $n \to \infty$,

$$A(\log c_n)\frac{1}{\sigma}\max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}} - D^*(\log c_n) \xrightarrow{\mathcal{D}} E \vee E' \tag{2.1.56}$$

where E' *and* E *are independent and identically Gumbel distributed random variables and* $\lim_{n\to\infty} n^{-1}\,Var\,S(n) = \sigma^2$.

Proof. By Lemma 2.1.4, we have

$$\max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)-\sigma W_n(n-k)|}{\sqrt{n-k}} = o_P\left((\log\log c_n)^{1/2}\right).$$

Since

$$\max_{n-c_n\le k\le n-1}\frac{|W_n(n-k)|}{\sqrt{n-k}} \stackrel{\mathcal{D}}{=} \max_{1\le k\le c_n}\frac{|W(k)|}{\sqrt{k}}, \tag{2.1.57}$$

we have, via using (2.1.41) and (2.1.55),

$$(\log\log n)^{-1/2}\max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}} \xrightarrow{P} \sigma \tag{2.1.58}$$

Let $0<\epsilon<1$ and put $u_n = \exp(\log n)^{\epsilon}$. Whence

$$(\log\log n)^{-1/2}\max_{n-u_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}} \xrightarrow{P} \sigma\sqrt{\epsilon}. \tag{2.1.59}$$

Therefore, via (2.1.58) and (2.1.59), we have

$$\lim_{n\to\infty} P\left[\max_{n-u_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}} \ge \max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}}\right] = 0 \tag{2.1.60}$$

and similarly

$$\lim_{n\to\infty} P\left[\max_{n-u_n\le k\le n-1}\frac{|W_n(n-k)|}{\sqrt{n-k}} \ge \max_{n-c_n\le k\le n-1}\frac{|W_n(n-k)|}{\sqrt{n-k}}\right] = 0. \tag{2.1.61}$$

In light of (2.1.51), we can assume $n-c_n < n-u_n$. Therefore, via using (2.1.60), (2.1.61) and Lemma 2.1.3, we have

$$\begin{aligned}&\left|\max_{n-c_n\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}} - \max_{n-c_n\le k\le n-1}\frac{\sigma\,|W_n(n-k)|}{\sqrt{n-k}}\right|\\ &= o_P\left((\log\log n)^{-1/2}\right)\end{aligned}$$

and the assertion flows from (2.1.57) and Darling and Erdős [27, Theorem 2], i.e. (2.1.8), that is,

$$A(\log c_n)\max_{1\le k\le c_n}\frac{|W(k)|}{\sqrt{k}} - D^*(\log c_n) \xrightarrow{\mathcal{D}} E \vee E'$$

as $n\to\infty$. □

PROOF OF THEOREM 2.1.1. In light of (2.1.13) and (2.1.14), we immediately derive the decomposition

$$G_n(k) = \frac{\frac{1}{k}S(k) - \frac{1}{n-k}\left(S(n) - S(k)\right)}{\sqrt{\frac{1}{k} + \frac{1}{n-k}}} \quad (k = 1, \ldots, n-1). \tag{2.1.62}$$

Therefore, putting together Lemma 2.1.5, Lemma 2.1.6, Lemma 2.1.7 and Lemma 2.1.8 we obtain (2.1.17) and (2.1.18). □

PROOF OF THEOREM 2.1.2. As a consequence of Theorem 2.1.1, we have, as $n \to \infty$,

$$A\left(\log\left(n/\log n\right)\right)\frac{1}{\sigma}\max_{1\le k\le c_n}|G_n(k)| - D^*\left(\log\left(n/\log n\right)\right) \xrightarrow{\mathcal{D}} E \vee E' \tag{2.1.63}$$

and

$$A\left(\log\left(n/\log n\right)\right)\frac{1}{\sigma}\max_{n-c_n\le k\le n-1}|G_n(k)| - D^*\left(\log\left(n/\log n\right)\right) \xrightarrow{\mathcal{D}} E \vee E'. \tag{2.1.64}$$

Moreover, similarly as in (2.1.41), we derive, as $n \to \infty$,

$$\left(2\log\log\left(n/\log n\right)\right)^{-1/2}\max_{1\le k\le \frac{n}{\log n}}|G_n(k)| \xrightarrow{P} \sigma \tag{2.1.65}$$

and

$$\left(2\log\log\left(n/\log n\right)\right)^{-1/2}\max_{n-\frac{n}{\log n}\le k\le n-1}|G_n(k)| \xrightarrow{P} \sigma. \tag{2.1.66}$$

Letting $n \to \infty$, we obtain

$$\left(A\left(\log\left(n/\log n\right)\right) - A\left(\log n\right)\right)\left(2\log\log\left(n/\log n\right)\right)^{1/2} = o(1) \tag{2.1.67}$$

and

$$D^*\left(\log n\right) - D^*\left(\log\left(n/\log n\right)\right) = o(1). \tag{2.1.68}$$

Concerning the former statement (2.1.67), consider the expression

$$\log\log\left(n/\log n\right) - \log\left(1 - \left(\log\log n/\log n\right)\right) = \log\log n.$$

Whence, from (2.1.1) we have

$$\begin{aligned}&\left(A\left(\log\left(n/\log n\right)\right) - A\left(\log n\right)\right)A\left(\log\left(n/\log n\right)\right)\\ &= \left(-\sqrt{1-\epsilon_n} + 1\right)\left(2\log\log\left(n/\log n\right)\right),\end{aligned}$$

where

$$\epsilon_n = \log\left(1 - \left(\log\log n/\log n\right)\right)/\log\log\left(n/\log n\right).$$

Therefore, for some $0 < \xi_n < \epsilon_n$, the mean value theorem implies

$$\begin{aligned}&\left(A\left(\log\left(n/\log n\right)\right) - A\left(\log n\right)\right)A\left(\log\left(n/\log n\right)\right)\\ &= \left(1-\xi_n\right)^{-1/2}\epsilon_n\left(\log\log\left(n/\log n\right)\right).\end{aligned}$$

This yields (2.1.67). Moreover, (2.1.68) is an immediate consequence of (2.1.2). Towards this end, the assertions (2.1.19) and (2.1.20) follow from (2.1.63) - (2.1.68), via using Slutsky's theorem, cf. e.g. Pollard [88, p. 175]. □

PROOF OF THEOREM 2.1.3. In light of (2.1.62), we derive the first assertion from Lemma 2.1.40 and (2.1.52). Similarly, the second assertion flows from Lemma 2.1.45 and (2.1.59). □

PROOF OF THEOREM 2.1.4. Let $a_n = n/\log n$. If $n \geq \exp\{1\}$, then

$$\sup_{a_n \leq t \leq n-a_n} \left|W(\lfloor t \rfloor) - W(t)\right| \leq \sup_{a_n \leq t \leq n-a_n} \sup_{0 \leq s \leq \log n} \left|W(t) - W(t+s)\right|.$$

Therefore, via using Hanson and Russo [56, display (3.10b)], we have

$$\sup_{a_n \leq t \leq n-a_n} \left|W(\lfloor t \rfloor) - W(t)\right| = O(\log n) \quad a.s. \quad (n \to \infty),$$

which implies, via Kuelbs and Philipp [65, Theorem 4],

$$\sup_{a_n \leq t \leq n-a_n} \frac{|S(t) - \sigma W(t)|}{\lfloor t \rfloor^{\frac{1}{2}-\lambda}} = o(1) \quad a.s. \quad (n \to \infty),$$

for some $0 < \lambda < 1/2$. Therefore

$$n^{-1/2} \sup_{a_n \leq t \leq n-a_n} \frac{|S(t) - \sigma W(t)|}{\left(\frac{\lfloor t \rfloor}{n}\left(1 - \frac{\lfloor t \rfloor}{n}\right)\right)^{1/2}} = o\left(n^{-\lambda}(\log n)^{1/2}\right) \tag{2.1.69}$$

and

$$n^{-1/2} \sup_{a_n \leq t \leq n-a_n} \frac{\left(\frac{\lfloor t \rfloor}{n}\right)^{1/2} |S(n) - \sigma W(n)|}{\left(1 - \frac{\lfloor t \rfloor}{n}\right)^{1/2}} = o\left(n^{-\lambda}(\log n)^{1/2}\right). \tag{2.1.70}$$

hold almost surely as $n \to \infty$. Observe that scaling properties of the Wiener process imply

$$\begin{aligned} &n^{-1/2} \sup_{a_n \leq t \leq n-a_n} \frac{\left|W(t) - \frac{\lfloor t \rfloor}{n} W(n)\right|}{\left(\frac{\lfloor t \rfloor}{n}\left(1 - \frac{\lfloor t \rfloor}{n}\right)\right)^{1/2}} \\ &\stackrel{\mathcal{D}}{=} \sup_{\frac{1}{\log n} \leq t \leq 1 - \frac{1}{\log n}} \frac{\left|W(t) - \frac{\lfloor nt \rfloor}{n} W(1)\right|}{\left(\frac{\lfloor nt \rfloor}{n}\left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{1/2}}. \end{aligned}$$

Next, consider $f(t) = (t(1-t))^{-1/2}$ $(0 < t < 1)$ with derivative

$$f'(t) = \frac{2t-1}{2\,(t(1-t))^{3/2}} \quad (0 < t < 1).$$

Let

$$r_n(y) = (ny - \lfloor nt \rfloor)/n.$$

The law of the iterated logarithm at zero, cf. e.g. Csörgő and Révész [26, Theorem 1.3.3], and the mean value theorem imply that

$$(\log\log\log n)^{-1/2} \sup_{\frac{1}{\log n} \le t \le \frac{1}{2}} |W(t)| \, |f(y) - f(y - r_n(y))| = O\left(\frac{\log n}{n}\right)$$

and

$$\sup_{\frac{1}{2} \le t \le 1 - \frac{1}{\log n}} |W(t)| \, |f(y) - f(y - r_n(y))| = O\left(\frac{(\log n)^{3/2}}{n}\right)$$

hold almost surely as $n \to \infty$. Next, consider $g(t) = (t/(1-t))^{1/2}$ $(0 < t < 1)$ with derivative

$$g'(t) = \frac{t^{1/2}}{2(1-t)^{5/2}} \quad (0 < t < 1).$$

Since

$$\sup_{\frac{1}{2} \le t \le 1 - \frac{1}{\log n}} |g(y) - g(y + r_n(y))| = O\left(\frac{(\log n)^{5/2}}{n}\right) \quad (n \to \infty),$$

we finally derive that

$$\left| \sup_{\frac{1}{\log n} \le t \le 1 - \frac{1}{\log n}} \frac{\left|W(t) - \frac{[nt]}{n} W(1)\right|}{\left(\frac{[nt]}{n}\left(1 - \frac{[nt]}{n}\right)\right)^{1/2}} - \sup_{\frac{1}{\log n} \le t \le 1 - \frac{1}{\log n}} \frac{|W(t) - tW(1)|}{(t(1-t))^{1/2}} \right| = O\left(\frac{(\log n)^{5/2}}{n}\right) \quad (2.1.71)$$

holds almost surely as $n \to \infty$. Towards this end, from Csörgő and Révész [26, display (1.9.7)]

$$\left\{ \frac{W(t) - tW(1)}{(t(1-t))^{1/2}}, \; 0 < t < 1 \right\} \stackrel{\mathcal{D}}{=} \left\{ V\left(\log \frac{t}{1-t}\right), \; 0 < t < 1 \right\}, \quad (2.1.72)$$

where $\{V(t), \; -\infty < t < \infty\}$ denotes an Ornstein-Uhlenbeck process with covariance $EV(t)V(s) = \exp(-|t-s|/2)$. Since the Ornstein-Uhlenbeck process is strictly stationary, we derive

$$\sup_{\frac{1}{\log n} \le y \le 1 - \frac{1}{\log n}} \frac{|W(y) - yW(1)|}{\sqrt{y(1-y)}} \stackrel{\mathcal{D}}{=} \sup_{0 < s < 2\log((\log n) - 1)} |V(s)|. \quad (2.1.73)$$

In light of (2.1.69)-(2.1.73), an application of (2.1.6) yields (2.1.27). □

PROOF OF THEOREM 2.1.5. Since Theorem 2.1.4 implies

$$\max_{\frac{n}{\log n}\le k\le n-\frac{n}{\log n}} |G_n(k)| = O_P\left(\log\log\log n\right) \quad (n\to\infty),$$

it suffices to prove

$$\lim_{n\to\infty} P\Bigg[\max\left\{\max_{1\le k\le n/\log n}|G_n(k)|\,,\max_{n-(n/\log n)\le k\le n-1}|G_n(k)|\right\} \le \left(y+D\left(\log n\right)\right)/A\left(\log n\right)\Bigg] = \exp\left\{-2\exp\left\{-y\right\}\right\}.$$

We claim that Assumption K is stronger than Assumption B. Suppose Assumption K holds. Letting $\delta_0 = \epsilon\delta/8$, Kuelbs and Philipp [65, Lemma 2.4] implies

$$n^{-1}\mathrm{Var}\, S(n) \to \sigma^2 \quad (n\to\infty),$$

and from Kuelbs and Philipp [65, Lemma 2.5], which is attributed to Sotres and Ghosh [98], together with Serfling [94, Theorem B] we have

$$\sup_{n\in\mathbb{N}} n^{-(1+\delta_0/2)} E\left|S(n)\right|^{2+\delta_0} < \infty$$

which implies assumption (2.1.10). Therefore, the assumptions of [12, Theorem 4] are fulfilled, hence Assumption B. As a consequence, we can apply Lemma 2.1.5 and Lemma 2.1.6 and it suffices to prove

$$\lim_{n\to\infty} P\Bigg[\max\left\{\max_{1\le k\le n/\log n}\frac{|S(k)|}{\sqrt{k}},\max_{n-(n/\log n)\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}}\right\} \le \left(y+D\left(\log n\right)\right)/A\left(\log n\right)\Bigg] = \exp\left\{-2\exp\left\{-y\right\}\right\}.$$

Observe that the distance between $S(k)$ and $S(n)-S(k)$ in terms of indices is $n-2(n/\log n)$. Hence, in light of the α-mixing property it suffices to prove

$$\lim_{n\to\infty} P\left[\max_{1\le k\le n/\log n}\frac{|S(k)|}{\sqrt{k}} \le \left(y+D\left(\log n\right)\right)/A\left(\log n\right)\right] \times P\left[\max_{n-(n/\log n)\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}} \le \left(y+D\left(\log n\right)\right)/A\left(\log n\right)\right] = \exp\left\{-2\exp\left\{-y\right\}\right\}.$$

By Theorem 2.1.1

$$P\left[A(\log n)\frac{1}{\sigma}\max_{1\le k\le n/\log n}\frac{|S(k)|}{\sqrt{k}}-D^*(\log n)\le y+\log 2\right]$$
$$\to\exp\{-\exp\{-y\}\}\quad(n\to\infty)$$

and

$$P\left[A(\log n)\frac{1}{\sigma}\max_{n-(n/\log n)\le k\le n-1}\frac{|S(n)-S(k)|}{\sqrt{n-k}}-D^*(\log n)\le y+\log 2\right]$$
$$\to\exp\{-\exp\{-y\}\}\quad(n\to\infty).$$

Noticing (2.1.2) and (2.1.28), the assertion flows from the observation

$$D(\log n)=D^*(\log n)-\log 2.$$

□

2.2. Limit Theorems via Coupling Methods

For further applications in the next chapter we are interested in the growth rate of the running maximum of weighted tied-down partial sum processes. We immediately derive from Theorem 2.1.5 the following limit theorem.

$$(\log\log n)^{-1/2}\max_{1\le k\le n-1}\left(\frac{n}{k(n-k)}\right)^{1/2}\left|S(k)-\frac{k}{n}S(n)\right|\xrightarrow{P}\sigma,\qquad(2.2.1)$$

as $n\to\infty$, where $\sigma^2=\lim_{n\to\infty}n^{-1}\mathrm{Var}\,S(n)$. This convergence-in-probability result holds under a polynomial decay of the strong mixing coefficient, which is Assumption K. Therefore it is natural to ask, whether it holds also under less restrictive mixing assumptions. In particular, we are interested to establish the limit theorem under Assumption B, i.e. assuming only a logarithmic decay of the strong mixing coefficient. The proof of the result will be based on so-called "coupling methods".

Let S and T be uncountable, complete and separable metric spaces. Given a probability space $(\Omega,\mathcal{A},P)$, let X be a measurable mapping from Ω to S. Then, in the sense of Lindvall [73], the quadruple $(\Omega,\mathcal{A},P,X)$ is called *random element* in the space $(S,\mathcal{S})$, where $\mathcal{S}$ denotes the Borel σ-field. The phrase "coupling" of two probability measures refers to constructions where both measures are represented as two random variables on a joint probability space. With a view towards invariance principles, Pollard [88, Chapter 10] provides a discussion of several different coupling results, including "Tusnády's Lemma" and the "Strassen-Dudley Theorem" among others.

DEFINITION (see Lindvall (1992)). *A coupling of two random elements* $(\Omega, \mathcal{A}, P, X)$ *and* $(\Omega', \mathcal{A}', P', X')$ *in the space* $(S, \mathcal{S})$ *and* $(T, \mathcal{T})$*, respectively, is a random element* $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{P}, (\hat{X}, \hat{X}'))$ *in* $(S \times T, \mathcal{S} \otimes \mathcal{T})$ *such that*

$$\hat{X} \overset{\mathcal{D}}{=} X \quad \textit{and} \quad \hat{X}' \overset{\mathcal{D}}{=} X'.$$

A probability space $(\Omega, \mathcal{A}, P)$ can be *extended* with a uniformly distributed random variable U. That is a new product space $(\Omega, \mathcal{A}, P) \otimes \big([0,1], \mathcal{B}, \lambda^1\big)$, where λ^1 denotes the Lebesgue measure on the unit interval and the random variable U is the projection on the second coordinate. Then every random element X (say) on the original space is stochastically independent of U with respect to $P \otimes \lambda^1$. We will keep the notation $(\Omega, \mathcal{A}, P)$ for the extended probability space.

The following construction was already used in the proof of Lemma 2.1.3 to establish a "backward" approximation result.

LEMMA (see Billingsley (1999, Section 21, Lemma 1)). *Consider a random element* $(\Omega, \mathcal{A}, P, X)$ *in* $(S, \mathcal{S})$*. Suppose* ν *is a probability measure on* $S \times T$ *with marginal measure* μ *on* S*, i.e.* $\mu(\cdot) = \nu(\cdot \times T)$*, and* $\mathcal{L}(X) = \mu$*. Then the probability space* $(\Omega, \mathcal{A}, P)$ *can be extended with a uniformly distributed random variable* U *and there is a random element* $(\Omega, \mathcal{A}, P, Y)$*, such that* Y *is a function of* (X, U) *and* $\mathcal{L}(X, Y) = \nu$*.*

Since ν determines the marginal measures, this construction can be viewed as coupling. A similar representation appeared already in Skorokhod [97]. The lemma stated by Billingsley can also be viewed as a part of Berbee's extension lemma.

LEMMA (see Berbee (1980, Extension Lemma 4.2.4)). *Consider a random element* $(\Omega, \mathcal{A}, P, X)$ *in* $(S, \mathcal{S})$*. Suppose* ν *is a probability measure on* $S \times T$ *with marginal measure* μ *on* S*, i.e.* $\mu(\cdot) = \nu(\cdot \times T)$*, and* $\mathcal{L}(X) = \mu$*. Then the probability space* $(\Omega, \mathcal{A}, P)$ *can be extended with a uniformly distributed random variable* U *and there is a random element* $(\Omega, \mathcal{A}, P, Y)$*, such that* Y *is a function of* (X, U) *and* $\mathcal{L}(X, Y) = \nu$*. The construction of* Y *does not affect the dependence structure, in the sense that*

$$P_{Z|X,Y} = P_{Z|X}$$

for any random element $(\Omega, \mathcal{A}, P, Z)$ *on the original space with values in a polish space.*

Using his extension lemma and coupling results due to Schwarz [93], Berbee derived the following result.

LEMMA (see Berbee (1979, Corollary 4.2.5)). *Consider a random element* $(\Omega, \mathcal{A}, P, (X,Y))$ *in* $(S \times T, \mathcal{S} \otimes \mathcal{T})$. *Then the probability space* $(\Omega, \mathcal{A}, P)$ *can be extended with a uniformly distributed random variable* U *and there is a random element* $(\Omega, \mathcal{A}, P, Y')$, *such that* Y' *is a function of* (X, Y, U) *and is independent of* X *and with distribution* $\mathcal{L}(Y') = \mathcal{L}(Y)$, *and such that*

$$P\left[Y \neq Y'\right] = \beta\left(\sigma(X), \sigma(Y)\right)$$

where β *denotes the measure of dependence with respect to* β*-mixing.*

We refer to Lindvall [73] for an account of coupling methods appearing in the different branches of probability theory. Related references and recent contributions concerning coupling methods with a view towards mixing conditions are stated in Merlevède and Peligrad [81]. Among the coupling results concerning strong mixing conditions, Berbee's lemma is the appropriate tool for our efforts. As a consequence we can not abstain from an additional β-mixing condition in the next theorem. But no rate of convergence is required for $\beta(n)$. Let us restate Assumption B, i.e. the set of assumptions of Bradley [12, Theorem 4].

ASSUMPTION B. *Let* $\{X_k,\ k \geq 1\}$ *be a strictly stationary sequence of centered real-valued random variables with*

$$EX_1^2 < \infty \quad \text{and} \quad Var\, S(n) \to \infty \quad (n \to \infty). \tag{2.2.2}$$

Suppose $\delta > 0$ *and* $\lambda > 1 + 3/\delta$ *are real numbers such that*

$$\sup_{n \in \mathbb{N}} (\, Var\, S(n))^{-(2+\delta)/2}\, E\, |S(n)|^{2+\delta} < \infty \tag{2.2.3}$$

and

$$\alpha(n) = o\left((\log n)^{-\lambda}\right) \quad (n \to \infty). \tag{2.2.4}$$

Let us state the result on the growth rate of the running maximum of weighted tied-down partial sum processes.

THEOREM 2.2.1. *Suppose that Assumption* B *holds. Suppose further that* $\{X_k,\ k \geq 1\}$ *is* β*-mixing, i.e.* $\beta(n) \downarrow 0$ $(n \to \infty)$. *Then the sequence* $\{X_k,\ k \geq 1\}$ *can be redefined without changing its distribution on an extended version of the initial probability space together with a sequence of Brownian bridge process* $\{B_n(t),\ 0 \leq t \leq 1\}$, *such that*

$$\begin{aligned} &\left| \max_{1 \leq k \leq n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left| S(k) - \frac{k}{n} S(n) \right| - \sigma \sup_{\frac{1}{n} \leq t \leq 1 - \frac{1}{n}} \frac{|B_n(t)|}{(t(1-t))^{1/2}} \right| \\ &= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty). \end{aligned} \tag{2.2.5}$$

and

$$(\log\log n)^{-1/2} \max_{1\le k\le n-1} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|S(k) - \frac{k}{n}S(n)\right| \xrightarrow{P} \sigma, \tag{2.2.6}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

In the proof we will use the following notation. As defined in (2.1.13) and (2.1.14), we let

$$G_n(k) = (n/(k(n-k)))^{1/2}\,(S(k) - (k/n)S(n)).$$

In a first step of the proof we need to replace the partial sum process $S(k)$ with a modified version $\tilde{S}_n(k)$ which allows for a coupling argument. Let $\{b_k,\ k\ge 1\}$ be a sequence with $1 \le b_n \le n/2$. For each $k \in \{1,\dots,n\}$ let

$$\tilde{S}_n(k) = S\left(\lfloor (n/2) - b_n\rfloor \wedge k\right) + S(k) - S\left((\lfloor (n/2) + b_n\rfloor - 1) \wedge k\right). \tag{2.2.7}$$

Note that $k \in \{1,\dots,\lfloor (n/2) - b_n\rfloor\}$ implies $\tilde{S}_n(k) = S(k)$.

Secondly, using strong approximations from Bradley [12], the newly defined partial sum process will be approximated with suitably discrete time processes $\{\tilde{B}_n\,(k)\,,\ 1\le k\le n\}$, defined as follows: For each n, we will construct two Wiener processes $\{W_{1n}(t),\ t\ge 0\}$ and $\{W_{2n}(t),\ t\ge 0\}$. Then for each $k\in\{1,\dots,\lfloor n/2\rfloor\}$ we let

$$\tilde{B}_n\,(k) = W_{1n}(k) - (k/n)\,(W_{1n}\,(\lfloor n/2\rfloor) + W_{2n}\,(\lfloor n/2\rfloor)) \tag{2.2.8}$$

and for each $k \in \{\lfloor n/2\rfloor + 1,\dots,n\}$

$$\begin{aligned}&\tilde{B}_n\,(k)\\ &= -W_{2n}(n-k) + (1-(k/n))\,(W_{1n}\,(\lfloor n/2\rfloor) + W_{2n}\,(\lfloor n/2\rfloor)).\end{aligned} \tag{2.2.9}$$

In a third step, similar to the Brownian bridge approximations established by Csörgő and Horváth [24], we will switch from $\tilde{B}_n(k)$ to the continuous versions $B_n(t)$, where for each $n\ge 2$ and $1/n \le t \le 1/2$

$$B_n\,(t) = W_{1n}(nt) - t\,(W_{1n}\,(n/2) + W_{2n}\,(n/2)) \tag{2.2.10}$$

and for $1/2 \le t \le 1-(1/n)$

$$B_n\,(t) = -W_{2n}(n-nt) + (1-t)\,(W_{1n}\,(n/2) + W_{2n}\,(n/2)). \tag{2.2.11}$$

Finally, in the proof of Theorem 2.2.1 we will show that the coupling argument applied on (2.2.7) carries over through (2.2.8) -(2.2.11) and yields the supremum of a standardized Brownian bridge process. The proof is based on a series of lemmas.

LEMMA 2.2.1. *Suppose that Assumption B holds. Let* $\{b_k,\ k \geq 1\}$ *be a sequence for which* $1 \leq b_n \leq \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0 \quad (n \to \infty)$. *Then*

$$n^{1/2} \left| \max_{1 \leq k \leq \lfloor \frac{n}{2} \rfloor} |G_n(k)| - \max_{1 \leq k \leq \lfloor \frac{n}{2} - b_n \rfloor} |G_n(k)| \right| = O_P\left((b_n \log\log b_n)^{1/2} \right) \quad (n \to \infty) \tag{2.2.12}$$

and

$$n^{1/2} \left| \max_{\lfloor \frac{n}{2} \rfloor + 1 \leq k \leq n-1} |G_n(k)| - \max_{\lfloor \frac{n}{2} + b_n \rfloor \leq k \leq n-1} |G_n(k)| \right| = O_P\left((b_n \log\log b_n)^{1/2} \right) \quad (n \to \infty). \tag{2.2.13}$$

PROOF. Let γ_n be the integer-valued random variable defined by

$$\gamma_n = \min \left\{ \ell \in \{1, \ldots, n-1\} \,\middle|\, |G_n(\ell)| = \max_{1 \leq k \leq n-1} |G_n(k)| \right\}. \tag{2.2.14}$$

Put

$$J_n = \left\{ \lfloor \frac{n}{2} - b_n \rfloor + 1, \ldots, \lfloor \frac{n}{2} + b_n \rfloor - 1 \right\} \quad \text{and} \tag{2.2.15}$$

$$I_n = \{1, \ldots, n-1\} \setminus J_n. \tag{2.2.16}$$

Then the equality

$$|G_n(\gamma_n)| - \max_{k \in I_n} |G_n(k)| = 0.$$

is true on the event $\{\gamma_n \in I(n)\}$. Now we put

$$k_0 = \lfloor \frac{n}{2} - b_n \rfloor \tag{2.2.17}$$

and observe that the inequality

$$\begin{aligned} \left| |G_n(\gamma_n)| - \max_{k \in I_n} |G_n(k)| \right| &= |G_n(\gamma_n)| - \max_{k \in I_n} |G_n(k)| \\ &\leq |G_n(\gamma_n)| - |G_n(k_0)| \end{aligned}$$

is true on the event $\{\gamma_n \in J_n\}$. To prove assertions (2.2.12) and (2.2.13), it suffices to prove that

$$n^{1/2} \max_{k \in J_n} \left| \left(\frac{n}{k(n-k)} \right)^{1/2} \left(S(k_0) + (S(k) - S(k_0)) - \frac{k}{n} S(n) \right) - \left(\frac{n}{k_0(n-k_0)} \right)^{1/2} \left(S(k_0) - \frac{k_0}{n} S(n) \right) \right| \tag{2.2.18}$$

is of order $O_P((b_n \log\log b_n)^{1/2}) \quad (n \to \infty)$.

Consider the function $f(t) = (t\,(1-t))^{-1/2}$ $(0 < t < 1)$. Since $f'(1/2) = 0$, the mean value theorem implies

$$n \max_{k\in J_n} \left| \left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2} - \left(\frac{k_0}{n}\left(1-\frac{k_0}{n}\right)\right)^{-1/2} \right| = o\left(b_n\right). \qquad (2.2.19)$$

Similar arguments yield

$$n \max_{k\in J_n} \left| \frac{k}{n}\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2} - \frac{k_0}{n}\left(\frac{k_0}{n}\left(1-\frac{k_0}{n}\right)\right)^{-1/2} \right| = O\left(b_n\right). \qquad (2.2.20)$$

By stationarity

$$\max_{k\in J_n} \left|S(k) - S\left(k_0\right)\right| \stackrel{\mathcal{D}}{=} \max_{1\le k\le 2b_n-1} \left|S(k)\right|.$$

Therefore, via

$$\begin{aligned} &\max_{1\le k\le 2b_n-1} |S(k)| \\ &\le (2b_n)^{1/2} \max_{1\le k\le 2b_n-1} \frac{|S(k)-\sigma W(k)|}{\sqrt{k}} \\ &\quad + (2b_n)^{1/2} \max_{1\le k\le 2b_n-1} \frac{|\sigma W(k)|}{\sqrt{k}}, \end{aligned}$$

we arrive at

$$\max_{k\in J_n} \left|S(k) - S\left(k_0\right)\right| = O_P\left(\left(b_n \log\log b_n\right)^{1/2}\right),$$

where Darling and Erdős [27, Theorem 1] was applied, i.e.

$$(2\log\log b_n)^{-1/2} \max_{1\le k\le 2b_n-1} \frac{|W(k)|}{\sqrt{k}} \xrightarrow{P} 1.$$

Moreover, from

$$\lim_{n\to\infty} \max_{k\in J_n} \left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2} = 2,$$

we arrive at

$$n^{1/2} \max_{k\in J_n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|S(k) - S\left(k_0\right)\right| = O_P\left(\left(b_n \log\log b_n\right)^{1/2}\right). \qquad (2.2.21)$$

Towards this end, from (2.2.19), (2.2.20) and (2.2.21), noticing $S\left(k_0\right)/\sqrt{n} = O_P\left(1\right)$ $(n\to\infty)$, expression (2.2.18) is of order $O_P((b_n \log\log b_n)^{1/2})$, as $n\to\infty$. This implies (2.2.12) and (2.2.13). □

LEMMA 2.2.2. *Suppose that Assumption B holds. Let* $\{b_k,\ k \geq 1\}$ *be a sequence for which* $1 \leq b_n \leq \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0 \quad (n \to \infty)$. *Then*

$$\max_{1 \leq k \leq \lfloor \frac{n}{2} - b_n \rfloor} \left| |G_n(k)| - \left(\frac{n}{k(n-k)} \right)^{1/2} \left| S(k) - \frac{k}{n} \tilde{S}(n) \right| \right|$$
$$= O_P\left(b_n^{1/2}\right) \quad (n \to \infty) \tag{2.2.22}$$

and

$$\max_{\lfloor \frac{n}{2} + b_n \rfloor \leq k \leq n-1} \left| |G_n(k)| - \left(\frac{n}{k(n-k)} \right)^{1/2} \left| \tilde{S}(k) - \frac{k}{n} \tilde{S}(n) \right| \right|$$
$$= O_P\left(b_n^{1/2}\right) \quad (n \to \infty). \tag{2.2.23}$$

PROOF. Put

$$J_n = \left\{ \lfloor \frac{n}{2} - b_n \rfloor + 1, \ldots, \lfloor \frac{n}{2} + b_n \rfloor - 1 \right\} \quad \text{and} \tag{2.2.24}$$

$$I_n = \{1, \ldots, n-1\} \setminus J_n. \tag{2.2.25}$$

Then

$$\lim_{n \to \infty} \max_{k \in I_n} \frac{k}{n} \left(\frac{n}{k(n-k)} \right)^{1/2} = 1. \tag{2.2.26}$$

By (2.2.7) and stationarity

$$S(n) - \tilde{S}_n(n) \stackrel{\mathcal{D}}{=} S\left(\lfloor \frac{n}{2} + b_n \rfloor - \lfloor \frac{n}{2} - b_n \rfloor - 1 \right). \tag{2.2.27}$$

Since $\left(\lfloor \frac{n}{2} + b_n \rfloor - \lfloor \frac{n}{2} - b_n \rfloor - 1 \right) \sim 2b_n \quad (n \to \infty)$, via using Bradley [12, Theorem 4], i.e. (2.1.30), we have

$$S\left(\lfloor \frac{n}{2} + b_n \rfloor - \lfloor \frac{n}{2} - b_n \rfloor - 1 \right) = \sigma W(2b_n) + o_P\left(b_n^{1/2}\right). \tag{2.2.28}$$

By (2.2.7), for each $k \in \left\{1, \ldots, \lfloor \frac{n}{2} - b_n \rfloor \right\}$, we have $\tilde{S}_n(k) = S(k)$. Therefore (2.2.22) flows from (2.2.26)- (2.2.28). Moreover, by (2.2.27), (2.2.28) and (2.2.7) we also have

$$\max_{\lfloor \frac{n}{2} + b_n \rfloor \leq k \leq n-1} \left| S(k) - \tilde{S}(k) \right| = O_P\left(b_n^{1/2}\right). \tag{2.2.29}$$

Observe

$$\lim_{n \to \infty} \max_{\lfloor \frac{n}{2} + b_n \rfloor \leq k \leq n-1} \left(\frac{n}{k(n-k)} \right)^{1/2} = 1. \tag{2.2.30}$$

Assertion (2.2.23) follows from (2.2.29) and (2.2.30). □

LEMMA 2.2.3. *Suppose that Assumption B holds. Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$. *Then, for each* $n \in \mathbb{N}$, *we can find two Wiener processes* $\left\{W_{1n}(t),\ 0 \le t \le \lfloor \frac{n}{2} - b_n \rfloor\right\}$ *and* $\left\{W_{2n}(t),\ 0 \le t \le \lfloor \frac{n}{2} - b_n \rfloor\right\}$ *on a possibly different probability space which supports the following construction*

$$n^{-1/2} \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \frac{|S(k) - \sigma W_{1n}(k)|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} = o_P\left((\log\log n)^{1/2}\right) \tag{2.2.31}$$

and

$$n^{-1/2} \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \frac{\left|\tilde{S}_n(n) - \sigma W_{1n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - \sigma W_{2n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right)\right|}{\left(\frac{n}{k}\left(1 - \frac{k}{n}\right)\right)^{1/2}}$$
$$= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty), \tag{2.2.32}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

PROOF. Observe

$$\lim_{n\to\infty} \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(1 - \frac{k}{n}\right)^{-1/2} = \sqrt{2}.$$

Therefore, from (2.1.31) we have

$$\max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(\frac{n}{n-k}\right)^{1/2} \frac{|S(k) - \sigma W_{1n}(k)|}{\sqrt{k}} = o_P\left((\log\log n)^{1/2}\right). \tag{2.2.33}$$

This implies (2.2.31). Similarly, from

$$\lim_{n\to\infty} \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(\frac{k}{n-k}\right)^{1/2} = 1$$

we have

$$\max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(\frac{k}{n-k}\right)^{1/2} \frac{\left|\tilde{S}_n\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - \sigma W_{1n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right)\right|}{\sqrt{n}}$$
$$= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty). \tag{2.2.34}$$

Moreover, from (2.1.35)

$$\max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \frac{|S(n) - S(k) - \sigma W_{2n}(n-k)|}{\sqrt{n-k}} = o_P\left((\log\log n)^{1/2}\right).$$

Whence

$$\max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(\frac{k}{n-k}\right)^{1/2} \frac{\left|\tilde{S}_n(n) - S\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - \sigma W_{2n}\left(n - \lfloor \frac{n}{2} + b_n \rfloor + 1\right)\right|}{\sqrt{n}}$$
$$= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty). \tag{2.2.35}$$

Choose a sequence $1 \le \ell_n \le \lfloor \frac{n}{2} - b_n \rfloor$ such that $\frac{n}{\ell_n} \uparrow \infty \quad (n \to \infty)$. Then, via Csörgő and Révész [26, display (1.2.3)], we have

$$n^{-1/2} \left| W\left(n - \lfloor \frac{n}{2} + b_n \rfloor + 1\right) - W\left(\lfloor \frac{n}{2} - b_n \rfloor + 1\right) \right| = O_P\left(\ell_n / \sqrt{n}\right). \tag{2.2.36}$$

If $\ell_n = o\left(\sqrt{n}\,(n \log\log n)^{-1/2}\right) \quad (n \to \infty)$, assertion (2.2.32) flows from (2.2.34) - (2.2.36). □

LEMMA 2.2.4. *Suppose that Assumption B holds. Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0 \quad (n \to \infty)$. *Then, for each* $n \in \mathbb{N}$, *we can find two Wiener processes* $\left\{W_{1n}(t),\ 0 \le t \le \lfloor \frac{n}{2} - b_n \rfloor\right\}$ *and* $\left\{W_{2n}(t),\ 0 \le t \le \lfloor \frac{n}{2} - b_n \rfloor\right\}$ *on a possibly different probability space which supports the following construction*

$$\begin{aligned} & n^{-1/2} \max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \frac{|S(n) - S(k) - \sigma W_{2n}(n-k)|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} \\ & = o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty) \end{aligned} \tag{2.2.37}$$

and

$$\begin{aligned} & n^{-1/2} \max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \frac{\left|\tilde{S}_n(n) - W_{1n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - W_{2n}\left(\lfloor \frac{n}{2} - b_n \rfloor + 1\right)\right|}{\left(\frac{k}{n}\right)^{1/2}\left(1 - \frac{k}{n}\right)^{-1/2}} \\ & = o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty), \end{aligned} \tag{2.2.38}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

REMARK. *The expressions* (2.2.37) *and* (2.2.38) *are motivated by the decomposition*

$$\tilde{S}_n(k) - \frac{k}{n}\tilde{S}_n(n) = -\left(S(n) - S(k)\right) + \left(1 - \frac{k}{n}\right)\tilde{S}_n(n),$$

where $\lfloor \frac{n}{2} + b_n \rfloor \le k \le n - 1$.

PROOF. Since

$$\lim_{n\to\infty} \max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \left(\frac{k}{n}\right)^{-1/2} = \sqrt{2},$$

we have, via using (2.1.35),

$$\max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \left(\frac{n}{k}\right)^{1/2} \frac{|S(n) - S(k) - \sigma W_{2n}(n-k)|}{\sqrt{n-k}} = o_P\left((\log\log n)^{1/2}\right).$$

This implies (2.2.37). Similarly, we have

$$\max_{\lfloor \frac{n}{2}+b_n \rfloor \le k \le n-1} \left(\frac{n-k}{k}\right)^{1/2} \frac{\left|S\left(\lfloor \frac{n}{2}-b_n \rfloor\right) - \sigma W_{1n}\left(\lfloor \frac{n}{2}-b_n \rfloor\right)\right|}{\sqrt{n}} = o_P(1),$$

where

$$\lim_{n\to\infty} \max_{\lfloor \frac{n}{2}+b_n \rfloor \le k \le n-1} \left(\frac{k}{n-k}\right)^{-1/2} = 1$$

was applied. Moreover, similarly to the proof of (2.2.32),

$$\max_{\lfloor \frac{n}{2}+b_n \rfloor \le k < n} \left(\frac{n-k}{k}\right)^{1/2} \frac{\left|\tilde{S}(n) - S\left(\lfloor \frac{n}{2}-b_n \rfloor\right) - \sigma W_{2n}\left(\lfloor \frac{n}{2}-b_n \rfloor\right)\right|}{\sqrt{n}}$$

is of order $o_P(1)$ $(n \to \infty)$, whence (2.2.38). □

LEMMA 2.2.5. *Suppose that Assumption B holds. Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$. *Let* $\{W(t),\ t \ge 0\}$ *be a standard Wiener process. Then*

$$\max_{1 \le k \le \lfloor \frac{n}{2}-b_n \rfloor} \frac{\left|W\left(\lfloor \frac{n}{2}-b_n \rfloor\right) - W\left(\lfloor \frac{n}{2} \rfloor\right)\right|}{\left(\frac{k}{n}\right)^{-1/2}\left(1-\frac{k}{n}\right)^{1/2}} = O\left((b_n \log(n/b_n))^{1/2}\right) \tag{2.2.39}$$

and

$$\max_{\lfloor \frac{n}{2}+b_n \rfloor \le k < n} \frac{\left|W\left(\lfloor \frac{n}{2}-b_n \rfloor\right) - W\left(\lfloor \frac{n}{2} \rfloor\right)\right|}{\left(\frac{k}{n}\right)^{1/2}\left(1-\frac{k}{n}\right)^{-1/2}} = O\left((b_n \log(n/b_n))^{1/2}\right) \tag{2.2.40}$$

almost surely as $n \to \infty$.

PROOF. We observe that

$$\lim_{n\to\infty} \max_{1 \le k \le \lfloor \frac{n}{2}-b_n \rfloor} \frac{k}{n-k} = \lim_{n\to\infty} \max_{\lfloor \frac{n}{2}+b_n \rfloor \le k \le n-1} \frac{n-k}{k} = 1.$$

By Csörgő and Révész [26, display (1.2.4)]

$$\sup_{1 \le t \le n-b_n} \sup_{0 \le s \le b_n} |W(t) - W(t+s)| = O\left((b_n \log(n/b_n))^{1/2}\right)$$

almost surely as $n \to \infty$. Both assertions follow immediately. □

LEMMA 2.2.6. *Suppose that Assumption B holds. Let the process* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *be as defined in* (2.2.8) *and* (2.2.9). *Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$. *Then*

$$\left| \max_{1 \le k \le \lfloor \frac{n}{2}-b_n \rfloor} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} - \max_{1 \le k \le \lfloor \frac{n}{2} \rfloor} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} \right|$$
$$= O_P\left((b_n \log\log(n/b_n))^{1/2}\right) \quad (n \to \infty). \tag{2.2.41}$$

PROOF. Let $\gamma_n \in \{1, \ldots, \lfloor n/2 \rfloor\}$ be the integer-valued random variable defined by

$$\gamma_n = \min \left\{ \ell : \frac{\left|\tilde{B}_n(\ell)\right|}{\left(\frac{\ell}{n}\left(1-\frac{\ell}{n}\right)\right)^{1/2}} = \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} \right\}.$$

Then, on the event $\left\{\gamma_n \in \left\{1, \ldots, \lfloor \frac{n}{2} - b_n \rfloor\right\}\right\}$ the equality

$$\left(\frac{n}{\gamma_n(n-\gamma_n)}\right)^{1/2} \left|\tilde{B}_n(\gamma_n)\right| - \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right| = 0$$

is true. Now we put

$$k_0 = \lfloor \frac{n}{2} - b_n \rfloor$$

and observe that the inequality

$$\begin{aligned} 0 &\le \left(\frac{n}{\gamma_n(n-\gamma_n)}\right)^{1/2} \left|\tilde{B}_n(\gamma_n)\right| - \max_{1 \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right| \\ &\le \left(\frac{n}{\gamma_n(n-\gamma_n)}\right)^{1/2} \left|\tilde{B}_n(\ell)\right| - \left(\frac{n}{k_0(n-k_0)}\right)^{1/2} \left|\tilde{B}_n(k_0)\right| \end{aligned}$$

is true on the event $\left\{\gamma_n \in \left\{\lfloor \frac{n}{2} - b_n \rfloor + 1, \ldots, \lfloor \frac{n}{2} \rfloor\right\}\right\}$. In order to establish (2.2.41) it suffices to prove

$$\begin{aligned} &\max_{\lfloor \frac{n}{2} - b_n \rfloor + 1 \le k \le \lfloor \frac{n}{2} \rfloor} \left| \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} - \frac{\left|\tilde{B}_n(k_0)\right|}{\left(\frac{k_0}{n}\left(1-\frac{k_0}{n}\right)\right)^{1/2}} \right| \\ &= O_P\left(\left(b_n \log\log(n/b_n)\right)^{1/2}\right). \end{aligned} \tag{2.2.42}$$

Analogously to (2.2.19), we have

$$\begin{aligned} &|W(k_0)| \max_{\lfloor \frac{n}{2} - b_n \rfloor + 1 \le k \le \lfloor \frac{n}{2} \rfloor} \left| \left(\frac{n}{k(n-k)}\right)^{1/2} - \left(\frac{n}{k_0\,(n-k_0)}\right)^{1/2} \right| \\ &= O_P\left(\frac{b_n}{n}\right), \end{aligned} \tag{2.2.43}$$

where $W(k_0)/\sqrt{n} = O_P(1) \quad (n \to \infty)$ was applied. Similarly to (2.2.20), we have

$$\begin{aligned} &\left|W\left(\lfloor \frac{n}{2} \rfloor\right)\right| \max_{\lfloor \frac{n}{2} - b_n \rfloor + 1 \le k \le \lfloor \frac{n}{2} \rfloor} \left| \frac{k}{n}\left(\frac{n}{k(n-k)}\right)^{1/2} - \frac{k_0}{n}\left(\frac{n}{k_0\,(n-k_0)}\right)^{1/2} \right| \\ &= O_P\left(\frac{b_n}{n}\right). \end{aligned} \tag{2.2.44}$$

By scaling properties of the Wiener process

$$n^{-1/2} \max_{\lfloor \frac{n}{2} - b_n \rfloor + 1 \le k \le \lfloor \frac{n}{2} \rfloor} |W(k) - W(k_0)| \stackrel{\mathcal{D}}{=} \max_{1 \le k \le \lfloor \frac{n}{2} \rfloor - \lfloor \frac{n}{2} - b_n \rfloor} \left| W\left(\frac{k}{n}\right)\right|.$$

Moreover, by the law of the iterated logarithm at zero, cf. e.g. Csörgő and Révész [26, display (1.3.10)], we have

$$n^{1/2} \sup_{\frac{1}{n} \le t \le \frac{b_n}{n}} |W(t)| = O_P\left((b_n \log\log(n/b_n))^{1/2} \right).$$

Thus, via

$$\lim_{n\to\infty} \max_{\lfloor \frac{n}{2} - b_n \rfloor + 1 \le k \le \lfloor \frac{n}{2} \rfloor} \left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{-1/2} = 2,$$

we arrive at

$$n^{1/2} \max_{\lfloor \frac{n}{2} - b_n \rfloor + 1 \le k \le \lfloor \frac{n}{2} \rfloor} \left(\frac{n}{k(n-k)} \right)^{1/2} |W(k) - W(k_0)|$$
$$= O_P\left((b_n \log\log(n/b_n))^{1/2} \right). \tag{2.2.45}$$

Assertion (2.2.42) follows from (2.2.43) - (2.2.45). □

Lemma 2.2.7. *Suppose that Assumption B holds. Let the process* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *be as defined in* (2.2.8) *and* (2.2.9). *Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$. *Then*

$$\left| \max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} - \max_{\lfloor \frac{n}{2} \rfloor + 1 \le k \le n-1} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} \right|$$
$$= O_P\left((b_n \log\log(n/b_n))^{1/2} \right) \quad (n \to \infty). \tag{2.2.46}$$

Proof. Let $\gamma_n \in \left\{ \lfloor \frac{n}{2} \rfloor + 1, \ldots, n-1 \right\}$ be the integer-valued random variable defined by

$$\gamma_n = \min\left\{ \ell : \frac{\left|\tilde{B}_n(\ell)\right|}{\left(\frac{\ell}{n}\left(1 - \frac{\ell}{n}\right)\right)^{1/2}} = \max_{\lfloor \frac{n}{2} \rfloor + 1 \le k \le n-1} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} \right\}.$$

Then, on the event $\left\{ \gamma_n \in \left\{ \lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1 \right\} \right\}$ the equality

$$\left(\frac{n}{\gamma_n(n - \gamma_n)} \right)^{1/2} \left|\tilde{B}_n(\gamma_n)\right|$$
$$- \max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n-1} \left(\frac{n}{k(n-k)} \right)^{1/2} \left|\tilde{B}_n(k)\right| = 0$$

is true. Now we put

$$k_0 = \lfloor \frac{n}{2} + b_n \rfloor$$

and observe that the inequality

$$
\begin{aligned}
0 &\le \left(\frac{n}{\gamma_n(n-\gamma_n)}\right)^{1/2}\left|\tilde{B}_n\left(\gamma_n\right)\right| \\
&\quad - \max_{\lfloor\frac{n}{2}+b_n\rfloor\le k\le n-1}\left(\frac{n}{k(n-k)}\right)^{1/2}\left|\tilde{B}_n\left(k\right)\right| \\
&\le \left(\frac{n}{\gamma_n(n-\gamma_n)}\right)^{1/2}\left|\tilde{B}_n\left(\gamma_n\right)\right| - \left(\frac{n}{k_0(n-k_0)}\right)^{1/2}\left|\tilde{B}_n\left(k_0\right)\right|.
\end{aligned}
$$

is true on the event $\left\{\gamma_n \in \left\{\lfloor\frac{n}{2}\rfloor+1,\ldots,\lfloor\frac{n}{2}+b_n\rfloor-1\right\}\right\}$. In order to establish (2.2.46) it suffices to prove

$$
\begin{aligned}
&\max_{\lfloor\frac{n}{2}\rfloor+1\le k\le\lfloor\frac{n}{2}+b_n\rfloor-1}\left|\frac{\left|\tilde{B}_n\left(k\right)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} - \frac{\left|\tilde{B}_n\left(k_0\right)\right|}{\left(\frac{k_0}{n}\left(1-\frac{k_0}{n}\right)\right)^{1/2}}\right| \\
&= O_P\left(\left(b_n\log\log\left(n/b_n\right)\right)^{1/2}\right). \qquad (2.2.47)
\end{aligned}
$$

Analogously to (2.2.19), we have

$$
\begin{aligned}
&\left|W\left(n-k_0\right)\right|\max_{\lfloor\frac{n}{2}\rfloor+1\le k\le\lfloor\frac{n}{2}+b_n\rfloor-1}\left|\left(\frac{n}{k(n-k)}\right)^{1/2}-\left(\frac{n}{k_0\left(n-k_0\right)}\right)^{1/2}\right| \\
&= O_P\left(\frac{b}{n}\right), \qquad (2.2.48)
\end{aligned}
$$

where $W\left(n-k_0\right)/\sqrt{n} = O_P\left(1\right) \quad \left(n\to\infty\right)$ was applied. Similarly to (2.2.20), we have

$$
\begin{aligned}
&\left|W\left(\lfloor\frac{n}{2}\rfloor\right)\right|\max_{\lfloor\frac{n}{2}\rfloor+1\le k\le\lfloor\frac{n}{2}+b_n\rfloor-1}\left|\left(1-\frac{k}{n}\right)\left(\frac{n}{k(n-k)}\right)^{1/2}\right. \\
&\qquad \left. - \left(1-\frac{k_0}{n}\right)\left(\frac{n}{k_0\left(n-k_0\right)}\right)^{1/2}\right| = O_P\left(\frac{b}{n}\right). \qquad (2.2.49)
\end{aligned}
$$

By scaling properties of the Wiener process

$$
\begin{aligned}
&n^{-1/2}\max_{\lfloor\frac{n}{2}\rfloor+1\le k\le\lfloor\frac{n}{2}+b_n\rfloor-1}\left|W\left(n-k_0\right)-W\left(n-k\right)\right| \\
&\overset{\mathcal{D}}{=}\max_{1\le k\le\lfloor\frac{n}{2}+b_n\rfloor-\lfloor\frac{n}{2}\rfloor-1}\left|W\left(\frac{k}{n}\right)\right|.
\end{aligned}
$$

Moreover, by the law of the iterated logarithm at zero, cf. e.g. Csörgő and Révész [26, display (1.3.10)], we have

$$
n^{1/2}\sup_{\frac{1}{n}\le t\le\frac{b}{n}}\left|W(t)\right| = O_P\left(\left(b_n\log\log\left(n/b_n\right)\right)^{1/2}\right).
$$

Thus, via

$$\lim_{n\to\infty} \max_{\lfloor\frac{n}{2}\rfloor+1\le k\le \lfloor\frac{n}{2}+b_n\rfloor-1} \left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2} = 2,$$

we arrive at

$$\begin{aligned} & n^{1/2} \max_{\lfloor\frac{n}{2}\rfloor+1\le k\le \lfloor\frac{n}{2}+b_n\rfloor-1} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|W(k) - W\left(\lfloor\frac{n}{2}+b_n\rfloor\right)\right| \\ & = O_P\left((b_n \log\log(n/b_n))^{1/2}\right) \qquad (2.2.50) \end{aligned}$$

Therefore (2.2.47) follows from (2.2.48) - (2.2.50). □

LEMMA 2.2.8. *Suppose that Assumption B holds. Let the process* $\{\tilde{B}_n(k),\ 1\le k\le n\}$ *be as defined in* (2.2.8) - (2.2.9) *and let the process* $\{B_n(t),\ \frac{1}{n}\le t\le \frac{n-1}{n}\}$ *be as defined in* (2.2.10) - (2.2.11). *Then*

$$\begin{aligned} & \left| \max_{1\le k\le\lfloor\frac{n}{2}\rfloor} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right| - \sup_{\frac{1}{n}\le t\le\frac{1}{2}} \left(\frac{n}{nt(n-nt)}\right)^{1/2} |B_n(t)| \right| \\ & = o_P\left((\log\log n)^{1/2}\right) \quad (n\to\infty). \qquad (2.2.51) \end{aligned}$$

PROOF. Observe

$$\max_{1\le k\le\lfloor\frac{n}{2}\rfloor} |W_{1n}(k)| = \sup_{\frac{1}{n}\le t\le\frac{1}{2}} |W_{1n}(\lfloor nt\rfloor)|$$

and

$$\sup_{\frac{1}{n}\le t\le\frac{1}{2}} \left(\frac{\lfloor nt\rfloor}{n}\left(1-\frac{\lfloor nt\rfloor}{n}\right)\right)^{-1/2} = n(n-1)^{-1/2}.$$

Moreover,

$$\sup_{\frac{1}{n}\le t\le\frac{1}{2}} |W_{1n}(\lfloor nt\rfloor) - W_{1n}(nt)| = \sup_{\frac{1}{n}\le t\le\frac{1}{2}} |W_{1n}(\lfloor nt\rfloor) - W_{1n}(\lfloor nt\rfloor + r_n(t))|.$$

where $0\le r_n(t)<1$. Thus

$$\sup_{\frac{1}{n}\le t\le\frac{1}{2}} |W_{1n}(\lfloor nt\rfloor) - W_{1n}(nt)| \le \sup_{1\le t\le n-\ell_n}\ \sup_{0\le s\le\ell_n} |W_{1n}(t+s) - W_{1n}(t)|.$$

Let $\{\ell_n,\ n\ge 1\}$ be a sequence for which $\ell_n\uparrow\infty$. Then via Hanson and Russo [56, Display (3.12b)] we arrive at

$$\sup_{\frac{1}{n}\le t\le\frac{1}{2}} \left(\frac{n}{\lfloor nt\rfloor(n-\lfloor nt\rfloor)}\right)^{1/2} |W_{1n}(\lfloor nt\rfloor) - W_{1n}(nt)| = O_P(\ell_n) \qquad (2.2.52)$$

as $n\to\infty$. Given some Wiener process, consider the random variable

$$\Delta_n(t) = \left| \left(\frac{n}{\lfloor nt\rfloor(n-\lfloor nt\rfloor)}\right)^{1/2} W(nt) - \left(\frac{n}{nt(n-nt)}\right)^{1/2} W(nt) \right|.$$

There exists some event A satisfying $P(A)=1$, so that for every $\omega \in A$ we can find $\frac{1}{n} \le t_n(\omega) \le \frac{1}{2}$, such that

$$\left| \Delta_n\left(t_n\right)(\omega) - \sup_{\frac{1}{n}\le t\le\frac{1}{2}} \Delta_n(t)(\omega) \right| < \epsilon,$$

where $t_n(\omega)$ depends on $\epsilon > 0$. Therefore, in order to establish that $\sup_{1/n\le t\le 1/2} \Delta_n(t) = o((\log\log n)^{1/2})$ holds almost surely as $n \to \infty$, it suffices to prove the following claim.

CLAIM. *For each sequence* $\{t_n,\ n \ge 1\}$, *satisfying* $\frac{1}{n} \le t_n \le \frac{1}{2}$,

$$\Delta_n\left(t_n\right) = o\left((\log\log n)^{1/2}\right) \quad a.s. \tag{2.2.53}$$

holds as $n \to \infty$.

PROOF OF CLAIM. Consider first the case that $\{t_n,\ n\ge 1\}$ is bounded away from zero, i.e. $\liminf_{n\to\infty} t_n > \epsilon_0$ for some $0<\epsilon_0<1/2$. Then

$$\Delta_n\left(t_n\right) \le \sup_{\epsilon_0\le t\le\frac{1}{2}} \frac{\left|W\left(nt\right)\right|}{\sqrt{n}} \left| \left(\frac{\lfloor nt\rfloor}{n}\left(1-\frac{\lfloor nt\rfloor}{n}\right)\right)^{-1/2} - \left(t(1-t)\right)^{-1/2} \right|.$$

Since $\frac{\lfloor nt\rfloor}{n} = t - \frac{r_n(t)}{n}$, where $0 \le r_n(t) < 1$, we have, as $n\to\infty$,

$$\begin{aligned} &\sup_{\epsilon_0\le t\le\frac{1}{2}} \left| \left(\frac{\lfloor nt\rfloor}{n}\left(1-\frac{\lfloor nt\rfloor}{n}\right)\right)^{-1/2} - \left(t(1-t)\right)^{-1/2} \right| \\ &\le \sup_{\epsilon_0\le t\le\frac{1}{2}} \left| \left(t-\frac{r_n(t)}{n}\left(1-t-\frac{r_n(t)}{n}\right)\right)^{-1/2} - \left(t(1-t)\right)^{-1/2} \right| = o(1). \end{aligned}$$

The functional law of the iterated logarithm, cf. e.g. Csörgő and Révész [26, Display (1.3.2)], implies $\Delta_n\left(t_n\right) = o((\log\log n)^{1/2})$ almost surely as $n\to\infty$.

Secondly, since $\{t_n,\ n\ge1\}$ is bounded by hypothesis, it remains to consider $\liminf_{n\to\infty} t_n = 0$. In light of the first case, we can assume $\lim_{n\to\infty} t_n = 0$ and $t_n = \frac{s_n}{n}$.

Let us consider the case $1 \le s_n \le s < \infty$. Then

$$\begin{aligned} \Delta_n\left(t_n\right) &= \frac{\left|W\left(s_n\right)\right|}{\sqrt{n}} \left| \left(\frac{\lfloor s_n\rfloor}{n}\left(1-\frac{\lfloor s_n\rfloor}{n}\right)\right)^{-1/2} - \left(\frac{s_n}{n}\left(1-\frac{s_n}{n}\right)\right)^{-1/2} \right| \\ &= \frac{\left|W\left(s_n\right)\right|}{\sqrt{n}} \left| g\left(\frac{s_n}{n}-\frac{r_n}{n}\right) - g\left(\frac{s_n}{n}\right) \right|, \end{aligned}$$

where $g(t) = (t(1-t))^{-1/2}$ and $0\le r_n<1$. Since $g'(t) = -\frac{1}{2}(t(1-t))^{-3/2}(1-2t)$, an application of the mean value theorem implies

$$\Delta_n\left(t_n\right) \le \left|W\left(s_n\right)\right|\left(s_n - r_n\right)^{-3/2} = O(1) \quad a.s. \quad (n\to\infty).$$

Finally, suppose $t_n = \frac{s_n}{n}$ and $\limsup_{n\to\infty} s_n = \infty$. In light of the preceding case we can assume $s_n \uparrow \infty$ and $\frac{s_n}{n} \downarrow 0$ $(n \to \infty)$. Similarly, we have

$$\begin{aligned}\Delta_n(t_n) &\le \frac{|W(s_n)|}{n^{1/2}} \frac{\left|g'\left(\frac{s_n}{n} - \frac{r_n}{n}\right)\right|}{n} \\ &\le \frac{|W(s_n)|}{(s_n \log\log s_n)^{1/2}} \frac{(\log\log s_n)^{1/2}}{s_n}.\end{aligned}$$

This completes the proof of the claim (2.2.53). □

Towards this end, similarly to (2.2.52)

$$\begin{aligned}&\sup_{\frac{1}{n}\le t\le \frac{1}{2}} \left(\frac{\lfloor nt\rfloor}{n}\right)\left(\frac{n}{\lfloor nt\rfloor(n-\lfloor nt\rfloor)}\right)^{1/2} |W_{1n}(\lfloor nt\rfloor) - W_{1n}(nt)| \\ &= O_P\left(\frac{\ell_n}{\sqrt{n}}\right) \quad (n\to\infty). \qquad (2.2.54)\end{aligned}$$

Consider

$$\begin{aligned}&\sup_{\frac{1}{n}\le t\le \frac{1}{2}} \left| \left(\frac{\lfloor nt\rfloor}{n}\right)^{1/2}\left(1-\frac{\lfloor nt\rfloor}{n}\right)^{-1/2} - \left(\frac{t}{1-t}\right)^{1/2}\right| \\ &= \sup_{\frac{1}{n}\le t\le \frac{1}{2}} \left| h\left(t - \frac{r_n(t)}{n}\right) - h(t)\right|,\end{aligned}$$

where $h(t) = (t/(1-t))^{1/2}$ and $0 \le r_n < 1$. Since $h'(t) = (1-t)^{-2}$, an application of the mean value theorem yields

$$\begin{aligned}&n^{-1/2} \sup_{\frac{1}{n}\le t\le \frac{1}{2}} \left|W_{1n}\left(\frac{n}{2}\right)\right| \left| \left(\frac{\lfloor nt\rfloor}{n}\right)^{1/2}\left(1-\frac{\lfloor nt\rfloor}{n}\right)^{-1/2} - \left(\frac{t}{1-t}\right)^{1/2}\right| \\ &\le 4n^{-3/2}\left|W_{1n}\left(\frac{n}{2}\right)\right|. \qquad (2.2.55)\end{aligned}$$

The assertions follows from (2.2.52) - (2.2.55). □

Lemma 2.2.9. *Suppose that Assumption B holds. Let the process* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *be as defined in* (2.2.8) - (2.2.9) *and let the process* $\{B_n(t),\ \frac{1}{n} \le t \le \frac{n-1}{n}\}$ *be as defined in* (2.2.10) - (2.2.11). *Then*

$$\begin{aligned}&\left| \max_{\lfloor \frac{n}{2}\rfloor \le k\le n-1} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right| - \sup_{\frac{1}{2}\le t\le 1-\frac{1}{n}} \left(\frac{n}{nt(n-nt)}\right)^{1/2} |B_n(t)| \right| \\ &= o_P\left((\log\log n)^{1/2}\right) \quad (n\to\infty). \qquad (2.2.56)\end{aligned}$$

PROOF. Let $\{\ell_k,\ k \geq 1\}$ be a sequence for which $\ell_n \uparrow \infty$ and $\ell_n = o((\log\log n)^{1/2})$ $(n \to \infty)$. Then, similarly to (2.2.52), we have

$$\sup_{\frac{1}{2} \leq t \leq 1-\frac{1}{n}} \left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}\right)^{1/2} |W_{2n}(n - \lfloor nt \rfloor) - W_{2n}(n - nt)| = O_P(\ell_n) \tag{2.2.57}$$

as $n \to \infty$. Given some Wiener process, consider the random variable

$$\Delta_n(t) = \left| \left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}\right)^{1/2} W(n - nt) - \left(\frac{n}{nt(n - nt)}\right)^{1/2} W(n - nt) \right|$$

There exists some event A satisfying $P(A) = 1$, so that for every $\omega \in A$ we can find $\frac{1}{2} \leq t_n(\omega) \leq 1 - \frac{1}{n}$, such that

$$\left| \Delta_n(t_n)(\omega) - \sup_{\frac{1}{2} \leq t \leq 1-\frac{1}{n}} \Delta_n(t)(\omega) \right| < \epsilon,$$

where $t_n(\omega)$ depends on $\epsilon > 0$. Therefore, in order to establish that $\sup_{1/2 \leq t \leq (n-1)/n} \Delta_n(t) = o((\log\log n)^{1/2})$ holds almost surely as $n \to \infty$, it suffices to prove the following claim.

CLAIM. *For each sequence* $\{t_n,\ n \geq 1\}$, *satisfying* $\frac{1}{2} \leq t_n \leq 1 - \frac{1}{n}$,

$$\Delta_n(t_n) = o\left((\log\log n)^{1/2}\right) \quad a.s. \tag{2.2.58}$$

holds as $n \to \infty$.

PROOF OF CLAIM. Consider first the case that $\{t_n,\ n \geq 1\}$ is bounded away from one, i.e. $\frac{1}{2} \leq \limsup_{n\to\infty} t_n \leq 1 - \epsilon_0$, for some $0 < \epsilon_0 < 1/2$. Then

$$\begin{aligned}
&\Delta_n(t_n) \\
&\leq \sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \frac{|W(n - nt)|}{\sqrt{n}} \left| \left(\frac{\lfloor nt \rfloor}{n}\left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2} \right|.
\end{aligned}$$

Since $\frac{\lfloor nt \rfloor}{n} = t - \frac{r_n(t)}{n}$, where $0 \leq r_n(t) < 1$, we have, as $n \to \infty$,

$$\begin{aligned}
&\sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \left| \left(\frac{\lfloor nt \rfloor}{n}\left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2} \right| \\
&\leq \sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \left| \left(t - \frac{r_n(t)}{n}\left(1 - t - \frac{r_n(t)}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2} \right| = o(1).
\end{aligned}$$

The functional law of the iterated logarithm, cf. e.g. Csörgő and Révész [26, Display (1.3.2)], implies $\Delta_n(t_n) = o((\log\log n)^{1/2})$ almost surely as $n \to \infty$.

Secondly, it remains to consider $\limsup_{n\to\infty} t_n = 1$. In light of the first case, we can assume $\lim_{n\to\infty} t_n = 1$ and $t_n = 1 - \frac{s_n}{n}$.

Let us consider the case $1 \le s_n \le s < \infty$. Then

$$\begin{aligned}
&\Delta_n(t_n) \\
&= \frac{|W(s_n)|}{\sqrt{n}} \left| \left(\frac{\lfloor n - s_n \rfloor}{n} \left(1 - \frac{\lfloor n - s_n \rfloor}{n} \right) \right)^{-1/2} - \left(\left(1 - \frac{s_n}{n} \right) \frac{s_n}{n} \right)^{-1/2} \right| \\
&= \frac{|W(s_n)|}{\sqrt{n}} \left| g\left(\frac{s_n}{n} - \frac{r_n}{n} \right) - g\left(\frac{s_n}{n} \right) \right|,
\end{aligned}$$

where $g(t) = (t(1-t))^{-1/2}$ and $0 \le r_n < 1$. Since $g'(t) = -\frac{1}{2}(t(1-t))^{-3/2}(1-2t)$, an application of the mean value theorem implies

$$\Delta_n(t_n) \le |W(s_n)| (s_n - r_n)^{-3/2} = O(1) \quad a.s. \quad (n \to \infty).$$

Finally, suppose $t_n = 1 - \frac{s_n}{n}$ and $\limsup_{n\to\infty} s_n = \infty$. In light of the preceding case we can assume $s_n \uparrow \infty$ and $\frac{s_n}{n} \downarrow 0$ $(n \to \infty)$. Similarly, we have

$$\begin{aligned}
\Delta_n(t_n) &\le \frac{|W(s_n)|}{n^{1/2}} \frac{\left| g'\left(\frac{s_n}{n} - \frac{r_n}{n} \right) \right|}{n} \\
&\le \frac{|W(s_n)|}{(s_n \log\log s_n)^{1/2}} \frac{(\log\log s_n)^{1/2}}{s_n}.
\end{aligned}$$

This completes the proof of the claim (2.2.58). □

Similar to (2.2.54) and (2.2.55) we have

$$n^{-1/2} \sup_{\frac{1}{2} \le t \le 1 - \frac{1}{n}} \left| \frac{W_{2n}\left(\lfloor \frac{n}{2} \rfloor \right) \left(1 - \frac{\lfloor nt \rfloor}{n} \right)}{\left(\frac{\lfloor nt \rfloor}{n} \left(1 - \frac{\lfloor nt \rfloor}{n} \right) \right)^{1/2}} - W_{2n}\left(\frac{n}{2} \right) \left(\frac{1-t}{t} \right)^{1/2} \right| = O_P(\ell_n) \tag{2.2.59}$$

as $n \to \infty$. The assertion follows from (2.2.57) - (2.2.59). □

PROOF OF THEOREM 2.2.1. For each $n \in \mathbb{N}$, let

$$J_n = \left\{ \lfloor \frac{n}{2} - b_n \rfloor + 1, \ldots, \lfloor \frac{n}{2} + b_n \rfloor - 1 \right\} \quad \text{and} \tag{2.2.60}$$

$$I_n = \{1, \ldots, n-1\} \setminus J_n. \tag{2.2.61}$$

Consider the random vectors

$$V_{1n} = \left(X_1, \ldots, X_{\lfloor \frac{n}{2} - b_n \rfloor} \right) \quad \text{and} \quad V_{2n} = \left(X_{\lfloor \frac{n}{2} + b_n \rfloor}, \ldots, X_n \right). \tag{2.2.62}$$

We extend the initial probability space with three uniformly distributed random variables U^*, U_i $(i = 1, 2)$, with two standard normal random variables N_i $(i = 1, 2)$ and with two Brownian bridge processes $\{B_i(t),\ 0 \le t \le 1\}$ $(i = 1, 2)$.

CLAIM 1. *We can construct two Wiener processes*

$$\left\{W_{1n}(t),\ 0 \le t \le \frac{n}{2}\right\} \quad \textit{and} \quad \left\{W_{2n}(t),\ 0 \le t \le \frac{n}{2}\right\} \tag{2.2.63}$$

on the extended probability space, such that W_{in} *is a measurable function of* (U_i, V_{in}, N_i, B_i) $(i = 1, 2)$ *and*

$$\begin{aligned}&\left|\max_{k\in I_n} |G_n(k)| - \sigma \sup_{\frac{1}{n}\le t\le 1-\frac{1}{n}} \left(\frac{n}{nt(n-nt)}\right)^{1/2} |B_n\,(t)|\right| \\ &= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty).\end{aligned} \tag{2.2.64}$$

PROOF OF CLAIM 1. In light of Lemma 2.2.3 - Lemma 2.2.5, an application of Billingsley [11, Lemma 21.1] yields two Wiener processes

$$\left\{W_{1n}(t),\ 0 \le t \le \lfloor\frac{n}{2} - b_n\rfloor\right\} \quad \text{and} \quad \left\{W_{2n}(t),\ 0 \le t \le \lfloor\frac{n}{2} - b_n\rfloor\right\}$$

on the extended probability space, such that W_{in} is a measurable function of (U_i, V_{in}) $(i = 1, 2)$ and

$$\begin{aligned}&\left|\max_{k\in I_n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{S}_n(k) - \frac{k}{n}\tilde{S}_n(n)\right| - \sigma \max_{k\in I_n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right|\right| \\ &= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty).\end{aligned} \tag{2.2.65}$$

For each $s \in [\lfloor\frac{n}{2} - b_n\rfloor, \frac{n}{2}]$ let

$$\begin{aligned}W_{in}(s) =& W_{in}\left(\lfloor\frac{n}{2} - b_n\rfloor\right) + \frac{s - \lfloor\frac{n}{2} - b_n\rfloor}{\sqrt{\frac{n}{2} - \lfloor\frac{n}{2} - b_n\rfloor}} N_i \\ &+ \sqrt{\frac{n}{2} - \lfloor\frac{n}{2} - b_n\rfloor}\, B_i\left(\frac{s - \lfloor\frac{n}{2} - b_n\rfloor}{\frac{n}{2} - \lfloor\frac{n}{2} - b_n\rfloor}\right) \quad (i = 1, 2).\end{aligned} \tag{2.2.66}$$

By this definition, cf. e.g. Csörgő and Révész [26, Proposition 1.4.1], we constitute two Wiener processes

$$\left\{W_{1n}(t),\ 0 \le t \le \frac{n}{2}\right\} \quad \text{and} \quad \left\{W_{2n}(t),\ 0 \le t \le \frac{n}{2}\right\}$$

on the initial probability space such that W_{in} is a measurable function of $\{U_i, V_{in}, N_i, B_i\}$ $(i = 1, 2)$. Therefore

$$\begin{aligned}&\left|\max_{k\in I_n} |G_n(k)| - \sigma \max_{1\le k\le n-1} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right|\right| \\ &= o_P\left((\log\log n)^{1/2}\right) \quad (n \to \infty),\end{aligned} \tag{2.2.67}$$

where Lemma 2.2.2, Lemma 2.2.6 and Lemma 2.2.7 were applied. The claim (2.2.64) follows via Lemma 2.2.8 and Lemma 2.2.9. □

CLAIM 2. *We can construct a random vector* V_{2n}^*, *such that*

$$V_{2n}^* \text{ is independent of } V_{1n} \quad \text{and} \quad V_{2n}^* \overset{\mathcal{D}}{=} V_{2n}. \tag{2.2.68}$$

Moreover, we can construct a Wiener process $\left\{W_{2n}^*(t),\ 0 \le t \le \frac{n}{2}\right\}$, *such that* W_{2n}^* *is a measurable function of* $(U_2, V_{2n}^*, N_2, B_2)$ *and*

$$\begin{aligned}&\left|\max_{k\in I_n} |G_n(k)| - \sigma \sup_{\frac{1}{n}\le t\le 1-\frac{1}{n}} \left(\frac{n}{nt(n-nt)}\right)^{1/2} |B_n^*(t)|\right|\\ &= o_P\left((\log\log n)^{1/2}\right) \quad (n\to\infty),\end{aligned} \tag{2.2.69}$$

where B_n^* *is defined as function of* W_{1n} *and* W_{2n}^* *as in* (2.2.10) *and* (2.2.11).

PROOF OF CLAIM 2. By Berbee [4, Corollary 4.2.5] there exists a random vector V_{2n}^*, such that V_{2n}^* is a measurable function of (V_{1n}, V_{2n}, U^*) and (2.2.68) is satisfied. Moreover

$$P\left[V_{2n} \ne V_{2n}^*\right] = \beta\left(\sigma\left(V_{1n}\right), \sigma\left(V_{2n}\right)\right). \tag{2.2.70}$$

Towards this end, we introduce the coupled version of $\tilde{S}_n(n)$. Let $\tilde{S}_n^*(k) = S(k)$, $k \in \left\{1, \dots, \lfloor\frac{n}{2} - b_n\rfloor\right\}$, and for each $k \in \left\{\lfloor\frac{n}{2} + b_n\rfloor, \dots, n\right\}$ we put

$$\tilde{S}_n^*(k) = S\left(\lfloor\frac{n}{2} - b_n\rfloor\right) + \sum_{\ell=1}^{k-\lfloor\frac{n}{2}+b_n\rfloor+1} \pi_\ell V_n^{*(2)}, \tag{2.2.71}$$

where π_ℓ denotes the ℓ-th projection. Hence, as in the proof of Lemma (2.2.2),

$$\max_{k\in I_n}\left| |G_n(k)| - \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{S}_n^*(k) - \frac{k}{n}\tilde{S}_n^*(n)\right|\right| = O_P\left(b_n^{1/2}\right) \tag{2.2.72}$$

is true on the event $\left\{V_n^{(2)} = V_n^{*(2)}\right\}$. Since $\beta\left(\sigma\left(V_{1n}\right), \sigma\left(V_{2n}\right)\right) \le 2b_n$, the β-mixing condition applies and we have

$$\lim_{n\to\infty} P\left[V_{2n} \ne V_{2n}^*\right] = 0. \tag{2.2.73}$$

Whence

$$\max_{k\in I_n}\left| |G_n(k)| - \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{S}_n^*(k) - \frac{k}{n}\tilde{S}_n^*(n)\right|\right| = O_P\left(b_n^{1/2}\right). \tag{2.2.74}$$

The claim (2.2.69) follows along the lines in the proof of CLAIM 1 with V_{2n} replaced by V_{2n}^*. □

Let $b_n = \log\log\log n$, then both claims together with Lemma 2.2.1 yield

$$\left| \max_{1\le k\le n} |G_n(k)| - \sigma \sup_{\frac{1}{n}\le t\le 1-\frac{1}{n}} \left(\frac{n}{nt(n-nt)} \right)^{1/2} |B_n^*(t)| \right|$$
$$= o_P\left((\log\log n)^{1/2} \right) \quad (n \to \infty). \tag{2.2.75}$$

Since W_{1n} is a measurable function of $\{U_1, V_{1n}, N_1, B_1\}$ and W_{2n}^* is a measurable function of $\{U_2, V_{2n}^*, N_2, B_2\}$, both Wiener processes are independent. Consequently, the process $\left\{n^{-1/2}B_n^*(t),\ 0\le t\le 1\right\}$ is a Brownian bridge. Similarly as in Csörgő and Horváth [24, Theorem A.4.2], using the representation as strictly stationary Ornstein-Uhlenbeck process, cf. e.g. (2.1.72), the assertion follows from (2.1.6). □

CHAPTER 3

Limit Theorems in Change-Point Analysis

In the first section we will establish the limit distribution for a max-type test statistic via proving a Darling-Erdős type limit theorem. In particular, we will present a new approximation for weighted tied-down sums of mixing random variables by a sequence of standardized Brownian bridge processes. In the second section we will discuss asymptotical results for possible rejection regions.

3.1. Quasi-Likelihood Approach

Suppose that $X_1, \ldots, X_n$ are independent and identically distributed random variables with common distribution P_θ, where $\theta \in \Theta$. Furthermore, we will assume that each P_θ is absolutely continuous with respect to a common σ-finite measure λ, so that p_θ denotes the density of P_θ with respect to λ. The likelihood function is defined by

$$L_n(\theta) = \prod_{i=1}^{n} p\left(x_i, \theta\right) \tag{3.1.1}$$

and is the joint probability density of the observations $x_1, \ldots, x_n$ as a function of θ, cf. Lehmann and Romano [69, p. 503].

Within this general setup, someone wishes to test for a parameter change, i.e. the null hypothesis

$$H_0 \ : \ \theta_1 = \cdots = \theta_n \in \Theta_0 \subset \Theta \tag{3.1.2}$$

versus the alternative

$$H_{1n}(k) \ : \ \theta_1 = \cdots = \theta_k \neq \theta_{k+1} = \cdots = \theta_n, \tag{3.1.3}$$

where $\theta_1, \ldots, \theta_k \in \Theta_0$ and $\theta_{k+1}, \ldots, \theta_n \in \Theta_1 \subset \Theta \setminus \Theta_0$ are unknown. If k is known, one can consider the likelihood ratio statistic $-2 \log \Lambda_k$, where

$$\Lambda_k = \sup_{\theta \in \Theta_0} L_n\left(\theta\right) \Big/ \sup_{(\theta, \tau) \in \Theta_0 \times \Theta_1} \prod_{i=1}^{k} p\left(x_i, \theta\right) \prod_{i=k+1}^{n} p\left(x_i, \tau\right). \tag{3.1.4}$$

If $\hat{\theta}_k$ and $\hat{\tau}_k$ are maximum likelihood estimators for θ and τ, then

$$\Lambda_k = L_n\left(\hat{\theta}_n\right) \Big/ \prod_{i=1}^{k} p\left(x_i, \hat{\theta}_k\right) \prod_{i=k+1}^{n} p\left(x_i, \hat{\tau}_k\right), \tag{3.1.5}$$

cf. e.g. Lehmann and Romano [69, p. 513]. Nevertheless, in the so-called at most one change-point (AMOC) model the change-point k is usually assumed to be unknown. Therefore it is natural to use the maximally selected likelihood ratio and reject H_0, if

$$\max_{1\le k\le n} \left(-2\log\Lambda_k\right) \tag{3.1.6}$$

is large. Csörgő and Horváth [24, Chapter 1] established weighted approximations for $\max_{1\le k\le n}\left(-2\log\Lambda_k\right)$ under rather general regularity assumptions.

Let us consider the AMOC model for a possible change in the mean μ of real-valued *normal* observations. If the variance σ^2 is assumed to be known, then

$$\max_{1\le k<n} \left(-2\log\Lambda_k\right) = \max_{1\le k<n} \frac{1}{\sigma^2}\left(k\hat{X}_k + (n-k)\check{X}_k - n\hat{X}_n\right), \tag{3.1.7}$$

where $\hat{X}_k$ and $\check{X}_k$ are the maximum likelihood estimators for the mean based on $\{X_1, \ldots, X_k\}$ and $\{X_{k+1}, \ldots, X_n\}$, respectively. Therefore

$$\max_{1\le k<n} \left(-2\log\Lambda_k\right) = \max_{1\le k<n} \frac{n}{\sigma^2 k(n-k)}\left(S(k) - \frac{k}{n}S(n)\right)^2, \tag{3.1.8}$$

where $S(k) = \sum_{i=1}^{k} X_i$.

Davis et. al. [31] derived the limit distribution of (3.1.8) via proving a Darling-Erdős type limit theorem, i.e.

$$A\left(\log n\right) \max_{1\le k<n} \left(-2\log\Lambda_k\right)^{1/2} - D^*\left(\log n\right) \xrightarrow{\mathcal{D}} E \vee E', \tag{3.1.9}$$

where E' and E are independent and identically distributed random variables with *Gumbel* type extreme value distribution, see (2.1.3), and $A(x)$, $D^*(x)$ are defined in (2.1.1) and (2.1.2), respectively. Moreover, if the variance is unknown and considered as nuisance parameter, then Csörgő and Horváth [24, p. 31] proved that the maximally selected likelihood ratio is large if and only if

$$\max_{1\le k<n} \frac{1}{\hat{\sigma}_k}\left(\frac{n}{k(n-k)}\right)^{1/2} \left|S(k) - \frac{k}{n}S(n)\right| \tag{3.1.10}$$

is large, where

$$\hat{\sigma}_k^2 = n^{-1}\left(\sum_{i=1}^{k}\left(X_i - \hat{X}_k\right)^2 + \sum_{i=k+1}^{n}\left(X_i - \check{X}_k\right)^2\right). \tag{3.1.11}$$

Horvath [57] considered changes in mean and variance, that is, in the parameter vector (μ, σ^2) of normal observations. From an asymptotic viewpoint the maximally selected likelihood ratio is then similar to (3.1.8), see [57, display (3.27)]. As a consequence of his invariance-principle-based approach, Horvath [57] pointed out that the assumption of normal observations can be dropped. In particular, he applied multivariate strong approximation results due to Einmahl [42] and established a Darling-Erdős type limit theorem for the maximum of the norm of a d-dimensional Ornstein-Uhlenbeck process.

Motivated by the parametric AMOC model of Worsley [108] and with a view towards a non-parametric approach, Gombay and Horváth [51] considered the statistic

$$\max_{1\le k<n} kg\left(\hat{X}_k\right) + (n-k)g\left(\check{X}_k\right) - ng\left(\hat{X}_n\right), \tag{3.1.12}$$

where the function g satisfies a mild smoothness condition. It turns out that (3.1.8) is a special case of (3.1.12). For related results we refer to Csörgő and Horváth [24, Theorem 1.4.1].

Recently, Ling [74] contributes a general approach for testing parameter stability in time series y_t, generated by the model

$$y_t = f\left(\theta, Y_{t-1}, \epsilon_t\right) \quad (t = 0, \pm 1, \pm 2, \dots), \tag{3.1.13}$$

where f is a known function, θ is a d-dimensional parameter vector, $\{\epsilon_t,\ t = 0, \pm1, \pm2, \dots\}$ is a sequence of independent and identically distributed errors and Y_t is the infinite-dimensional vector $(\dots, y_{t-1}, y_t)$. In a first step, Ling [74] considered a known change-point $k \in [1, n)$ and introduced so-called objective function L_n and L_{1n} to estimate the time-series parameters θ_0 and θ_1 from the subsamples $\{y_1, \dots, y_k\}$ and $\{y_{k+1}, \dots, y_n\}$, respectively, i.e.

$$L_n\left(k, \theta_0\right) = \sum_{t=1}^{k} l\left(\theta_0, Y_t\right) \quad \text{and} \quad L_{1n}\left(k, \theta_1\right) = \sum_{t=k+1}^{n} l\left(\theta_1, Y_t\right), \tag{3.1.14}$$

where $l\left(\theta, Y_t\right)$ is almost surely three times differentiable with respect to θ. Assuming minor regularity assumptions, these estimating functions allow for standard estimation methods, see e.g. [28, Chapter 16 & 17], and quasi-likelihood estimates for the unknown time-series parameters θ_0 and θ_1. Let $D_t(\theta) = \partial l_t(\theta)/\partial\theta$, where $l_t(\theta) = l(\theta, Y_t)$. Furthermore let $P_t(\theta) = -\partial^2 l_t(\theta)/\partial\theta\partial\theta'$. Denote $\Sigma = E\left[P_t(\theta_0)\right]$ and $\Omega = E\left[D_t(\theta_0) D_t'(\theta_0)\right]$. Suppose $\hat{\theta}_n(k)$ and $\hat{\theta}_{1n}(k)$ are the maximizers of the two objective functions on Θ. It seems reasonable to choose a Wald type test statistic based on $\hat{\theta}_n(k) - \hat{\theta}_{1n}(k)$ for testing the no change null hypothesis H_0 against the change-point alternative $H_{1n}(k)$, i.e. (3.1.3). We refer to [28, Definition 21.1] for a definition of the Wald test. Under the unknown change-point assumption, Ling [74] proposed a maximally selected Wald type test statistic, i.e. $\max_{1\le k<n} W_n(k)$ for testing H_0 against $\cup_{k\in[1,n)} H_{1n}(k)$. In order to establish the limit distribution of the max-type test statistic, it was shown that $\max_{1\le k<n} W_n(k)$ is asymptotically equivalent, in a certain sense, to the following expression (see [74, display (6.9)]):

$$\max_{\log n\le k\le n-\log n}\left|\frac{k(n-k)}{n}\left[\hat{\theta}_n(k)-\hat{\theta}_{1n}(k)\right]'\Sigma\Omega^{-1}\Sigma\left[\hat{\theta}_n(k)-\hat{\theta}_{1n}(k)\right]\right|.$$

Moreover, let

$$\xi_n(k)=\left(\frac{k(n-k)}{n}\right)^{1/2}\left[\frac{1}{k}\sum_{t=1}^{k}D_t(\theta_0)-\frac{1}{n-k}\sum_{t=k+1}^{n}D_t(\theta_0)\right]. \qquad (3.1.15)$$

Since $\partial L_{1n}\left(k,\hat{\theta}_{1n}(k)\right)/\partial\theta=0$, Ling derived via using Taylor expansions

$$\max_{\log n\le k\le n-\log n}\left|W_n(k)-\xi_n'(k)\Omega^{-1}\xi_n'(k)\right|=o_P(1). \qquad (3.1.16)$$

In light of (3.1.15), the expression $\max_{\log n\le k\le n-\log n}\xi_n'(k)\Omega^{-1}\xi_n(k)$ can be viewed as a vector-valued version of (3.1.8). This suggests that the maximally selected Wald type statistics obeys an extreme value asymptotic. Assuming that the quasi score functions $D_t(\theta_0)$ form a vector-valued martingale difference sequence, Ling proved a Darling-Erdős type limit theorem via using multidimensional strong invariance principles due to Eberlein [37], i.e., for some $\delta>0$,

$$\sum_{t=1}^{k}D_t(\theta_0)-\sum_{t=1}^{k}G_{1t}=O\left(k^{1/2-\delta}\right) \qquad (3.1.17)$$

holds almost surely as $k\to\infty$, where $\{G_{1t},\ t=1,2,\dots\}$ are independent identically d-dimensional normal vectors with covariance matrix Ω. Moreover, approximations for the second partial sum in (3.1.15) were derived with new “backward” invariance principles, i.e.

$$\sum_{t=-k}^{-1}D_t(\theta_0)-\sum_{t=1}^{k}G_{2t}=O\left(k^{1/2-\delta}\right)\quad a.s.\quad(k\to\infty). \qquad (3.1.18)$$

In order to establish (3.1.18), it was additionally assumed that $D_t(\theta_0)$ obeys the so-called near-epoch dependence (NED) condition, in the sense of McLeish [79]. We refer to Chapter 5 for related results involving the NED condition.

Davis et. al. [32] studied the AMOC model for parameter changes in autoregressive time series models, satisfying the α-mixing condition, with a quasi likelihood approach. Based on results in [57], they proved a Darling-Erdős limit theorem for a vector-valued version of the maximally selected (quasi) likelihood ratio statistic (3.1.10). In particular, they employed strong approximation results, similar to (3.1.17), for strongly mixing random vectors due to Kuelbs and Philipp [65, Theorem 4]. Since in the strongly mixing case the time reversed process is still strongly mixing, the backward version as in (3.1.18) follows immediately from the "forward" invariance principle. The crucial point is that Kuelbs and Philipp [65, Theorem 4] assumed a polynomial decay of the strong mixing coefficients.

In conclusion we see that maximally selected (quasi) likelihood ratio tests and max-type Wald tests for the AMOC model result into expressions as in (3.1.10). This is our motivation to study further asymptotic properties of statistics based on weighted tied-down partial sums of dependent random variables. In the same context, Csörgő [22, p. 535] pointed out: "Studying the asymptotic behaviour of these statistics is clearly of interest[...]".

With the aim to meet the requirements of an interesting study, we will drop the main assumption in Davis et. al. [32], that is, the polynomial decay of the strong mixing coefficients. We point out that our Theorem 2.1.5 is closely related to the main result of Davis et. al. [32]. In both results the invariance-principle-based constructions rely heavily on the polynomial decay of the strong mixing coefficients because it guarantees a sharp approximation rate in the invariance principle due to Kuelbs and Philipp [65, Theorem 4].

Under the logarithmic decay of the strong mixing coefficients we will establish a different invariance-principle-based construction below. Moreover, we will show that the weighted tied-down partial sum process can be approximated by a sequence of standardized Brownian bridge processes which in turn implies a Darling-Erdős type limit theorem for the test statistic. We point out that in the independent case these kind of approximations are originally due to Csörgő and Horváth [24].

The construction here employs also coupling arguments as in the proof of our Theorem 2.2.1. This is somewhat different from the constructions in [24], established under independence assumptions. As a consequence we can not abstain from an additional β-mixing condition in the next theorem. However, no rate of convergence is required for $\beta(n)$. Let us restate Assumption B, i.e. the set of assumptions of Bradley [12, Theorem 4].

Assumption B. *Let* $\{X_k,\ k \geq 1\}$ *be a strictly stationary sequence of centered real-valued random variables with*

$$EX_1^2 < \infty \quad \textit{and} \quad Var\, S(n) \to \infty \quad (n \to \infty).$$

Suppose $\delta > 0$ *and* $\lambda > 1 + 3/\delta$ *are real numbers such that*

$$\sup_{n\in\mathbb{N}} (\, Var\, S(n))^{-(2+\delta)/2}\, E\, |S(n)|^{2+\delta} < \infty$$

and

$$\alpha(n) = o\left((\log n)^{-\lambda}\right) \quad (n \to \infty).$$

Let us state the weighted approximation result.

THEOREM 3.1.1. *Suppose that Assumption B holds. Suppose further that* $\{X_k,\ k \geq 1\}$ *is* β*-mixing, i.e.* $\beta(n) \downarrow 0$ $(n \to \infty)$. *Let* $a_n = \lfloor \exp{(\log n)^{\epsilon}} \rfloor$ $(n = 1, 2, \dots)$ *for some* $0 < \epsilon < 1$. *Then the sequence* $\{X_k,\ k \geq 1\}$ *can be redefined, without changing its distribution, on an extended version of the initial probability space together with a sequence of Brownian bridge process* $\{B_n(t),\ 0 \leq t \leq 1\}$, *such that*

$$\begin{aligned} &\left| \max_{a_n \leq k \leq n - a_n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| S(k) - \frac{k}{n} S(n) \right| - \sigma \sup_{\frac{a_n}{n} \leq t \leq 1 - \frac{a_n}{n}} \frac{|B_n\,(t)|}{(t(1-t))^{1/2}} \right| \\ &= o_P\left((\log\log n)^{-1/2} \right) \quad (n \to \infty), \end{aligned} \tag{3.1.19}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

In light of the "tail behavior" (Theorem 2.1.3) and the "overall behavior" (Theorem 2.2.1) of the running maximum, we are able to prove via Theorem 3.1.1 the following result.

THEOREM 3.1.2. *Suppose that Assumption B holds. Suppose further that* $\{X_k,\ k \geq 1\}$ *is* β*-mixing, i.e.* $\beta(n) \downarrow 0$ $(n \to \infty)$. *Then the sequence* $\{X_k,\ k \geq 1\}$ *can be redefined, without changing its distribution, on an extended version of the initial probability space together with a sequence of Brownian bridge process* $\{B_n(t),\ 0 \leq t \leq 1\}$, *such that*

$$\begin{aligned} &\left| \max_{1 \leq k \leq n-1} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| S(k) - \frac{k}{n} S(n) \right| - \sigma \sup_{\frac{1}{n} \leq t \leq 1 - \frac{1}{n}} \frac{|B_n\,(t)|}{(t(1-t))^{1/2}} \right| \\ &= o_P\left((\log\log n)^{-1/2} \right) \quad (n \to \infty), \end{aligned} \tag{3.1.20}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

As a consequence we are able to prove the extreme value asymptotic.

THEOREM 3.1.3. *Suppose that Assumption B holds. Suppose further that* $\{X_k,\ k \geq 1\}$ *is* β*-mixing, i.e.* $\beta(n) \downarrow 0$ $(n \to \infty)$. *Let*

$$A(x) = (2\log n)^{1/2} \quad \textit{and}$$

$$D(x) = 2\log x + \frac{1}{2}\log\log x - \frac{1}{2}\log \pi.$$

Then we have, as $n \to \infty$,

$$A(\log n)\frac{1}{\sigma}\max_{1\le k\le n-1}\left(\frac{n}{k(n-k)}\right)^{1/2}\left|S(k)-\frac{k}{n}S(n)\right| - D(\log n) \xrightarrow{\mathcal{D}} E\vee E', \tag{3.1.21}$$

where E *and* E' *are independent identically Gumbel distributed random variables and* $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

The proof of Theorem 3.1.1 derives the rate $o_P((\log\log n)^{-1/2})$ mainly along the pattern in the proof of Theorem 2.2.1 and its preparatory series of lemmas. See the remarks following Theorem 2.2.1 for abbreviations and notations. We will employ a similar coupling method. The proofs of Theorem 3.1.2 and Theorem 3.1.3 will be given at the very end of this section.

LEMMA 3.1.1. *Suppose that Assumption B holds. Let* $\{b_k,\ k\ge 1\}$ *be a sequence for which* $1\le b_n\le \frac{n}{2}$, $b_n\uparrow\infty$ *and* $\frac{b_n}{n}\downarrow 0$ $(n\to\infty)$. *Then*

$$\begin{aligned} n^{1/2}&\left|\max_{a_n\le k\le\lfloor\frac{n}{2}\rfloor}|G_n(k)| - \max_{a_n\le k\le\lfloor\frac{n}{2}-b_n\rfloor}|G_n(k)|\right| \\ &= O_P\left((b_n\log\log b_n)^{1/2}\right) \quad (n\to\infty) \end{aligned} \tag{3.1.22}$$

and

$$\begin{aligned} n^{1/2}&\left|\max_{\lfloor\frac{n}{2}\rfloor+1\le k\le n-a_n}|G_n(k)| - \max_{\lfloor\frac{n}{2}+b_n\rfloor\le k\le n-a_n}|G_n(k)|\right| \\ &= O_P\left((b_n\log\log b_n)^{1/2}\right) \quad (n\to\infty). \end{aligned} \tag{3.1.23}$$

PROOF. Follow the pattern in the proof of Lemma 2.2.1. □

LEMMA 3.1.2. *Suppose that Assumption B holds. Let* $\{b_k,\ k\ge 1\}$ *be a sequence for which* $1\le b_n\le \frac{n}{2}$, $b_n\uparrow\infty$ *and* $\frac{b_n}{n}\downarrow 0$ $(n\to\infty)$. *Then*

$$\begin{aligned} \max_{a_n\le k\le\lfloor\frac{n}{2}-b_n\rfloor}&\left||G_n(k)| - \left(\frac{n}{k(n-k)}\right)^{1/2}\left|S(k)-\frac{k}{n}\tilde{S}(n)\right|\right| \\ &= O_P\left(\left(\frac{b_n}{a_n}\right)^{1/2}\right) \quad (n\to\infty) \end{aligned} \tag{3.1.24}$$

and

$$\begin{aligned} \max_{\lfloor\frac{n}{2}+b_n\rfloor\le k\le n-a_n}&\left||G_n(k)| - \left(\frac{n}{k(n-k)}\right)^{1/2}\left|\tilde{S}(k)-\frac{k}{n}\tilde{S}(n)\right|\right| \\ &= O_P\left(\left(\frac{b_n}{a_n}\right)^{1/2}\right) \quad (n\to\infty). \end{aligned} \tag{3.1.25}$$

PROOF. Put

$$J_n = \left\{ \lfloor \frac{n}{2} - b_n \rfloor + 1, \ldots, \lfloor \frac{n}{2} + b_n \rfloor - 1 \right\}$$

and let

$$K_n = \{a_n, \ldots, n - a_n\} \setminus J_n.$$

Since

$$\max_{k \in K_n} \frac{k}{n} \left(\frac{n}{k(n-k)} \right)^{1/2} = O\left(a_n^{-1/2}\right) \quad (n \to \infty),$$

both assertions follow along the pattern in the proof of Lemma 2.2.2. □

LEMMA 3.1.3. *Suppose that Assumption B holds. Let* $\{b_k,\ k \geq 1\}$ *be a sequence for which* $1 \leq b_n \leq \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0 \quad (n \to \infty)$. *Then, for each* $n \in \mathbb{N}$, *we can find two Wiener processes* $\left\{W_{1n}(t),\ 0 \leq t \leq \lfloor \frac{n}{2} - b_n \rfloor \right\}$ *and* $\left\{W_{2n}(t),\ 0 \leq t \leq \lfloor \frac{n}{2} - b_n \rfloor \right\}$ *on a possibly different probability space which supports the following construction*

$$n^{-1/2} \max_{a_n \leq k \leq \lfloor \frac{n}{2} - b_n \rfloor} \frac{|S(k) - \sigma W_{1n}(k)|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} = o_P\left((\log\log n)^{-1/2}\right) \tag{3.1.26}$$

and

$$\begin{aligned} n^{-1/2} &\max_{a_n \leq k \leq \lfloor \frac{n}{2} - b_n \rfloor} \frac{\left| \tilde{S}(n) - \sigma W_{1n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - \sigma W_{2n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right) \right|}{\left(\frac{n}{k}\left(1 - \frac{k}{n}\right)\right)^{1/2}} \\ &= o_P\left((\log\log n)^{-1/2}\right) \quad (n \to \infty), \end{aligned} \tag{3.1.27}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

PROOF. Both assertions follow along the lines of the proof of Lemma 2.2.3 together with the approximation results Lemma 2.1.1 and Lemma 2.1.3. □

LEMMA 3.1.4. *Suppose that Assumption B holds. Let* $\{b_k,\ k \geq 1\}$ *be a sequence for which* $1 \leq b_n \leq \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0 \quad (n \to \infty)$. *Then, for each* $n \in \mathbb{N}$, *we can find two Wiener processes* $\left\{W_{1n}(t),\ 0 \leq t \leq \lfloor \frac{n}{2} - b_n \rfloor \right\}$ *and* $\left\{W_{2n}(t),\ 0 \leq t \leq \lfloor \frac{n}{2} - b_n \rfloor \right\}$ *on a possibly different probability space which supports the following construction*

$$\begin{aligned} n^{-1/2} &\max_{\lfloor \frac{n}{2} + b_n \rfloor \leq k \leq n - a_n} \frac{|S(n) - S(k) - \sigma W_{2n}(n-k)|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} \\ &= o_P\left((\log\log n)^{-1/2}\right) \quad (n \to \infty) \end{aligned} \tag{3.1.28}$$

and

$$n^{-1/2} \max_{\lfloor \frac{n}{2}+b_n \rfloor \le k \le n-a_n} \frac{\left|\tilde{S}(n) - W_{1n}\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - W_{2n}\left(\lfloor \frac{n}{2} - b_n \rfloor + 1\right)\right|}{\left(\frac{k}{n}\right)^{1/2}\left(1-\frac{k}{n}\right)^{-1/2}}$$

$$= o_P\left((\log\log n)^{-1/2}\right) \quad (n \to \infty), \tag{3.1.29}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}\, Var\, S(n)$.

Proof. Both assertions follow along the lines of the proof of Lemma 2.2.4 together with the approximation results Lemma 2.1.1 and Lemma 2.1.3. □

Lemma 3.1.5. *Suppose that Assumption B holds. Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$. *Let* $\{W(t),\ t \ge 0\}$ *be a standard Wiener process. Then*

$$\max_{a_n \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \frac{\left|W\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - W\left(\lfloor \frac{n}{2} \rfloor\right)\right|}{\left(\frac{k}{n}\right)^{-1/2}\left(1-\frac{k}{n}\right)^{1/2}} = O\left((b_n \log(n/b_n))^{1/2}\right) \tag{3.1.30}$$

and

$$\max_{\lfloor \frac{n}{2}+b_n \rfloor \le k \le n-a_n} \frac{\left|W\left(\lfloor \frac{n}{2} - b_n \rfloor\right) - W\left(\lfloor \frac{n}{2} \rfloor\right)\right|}{\left(\frac{k}{n}\right)^{1/2}\left(1-\frac{k}{n}\right)^{-1/2}} = O\left((b_n \log(n/b_n))^{1/2}\right) \tag{3.1.31}$$

almost surely as $n \to \infty$.

Proof. The proof follows the lines in the proof of Lemma 2.2.5. □

Lemma 3.1.6. *Suppose that Assumption B holds and let the process* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *be as defined in* (2.2.8) *and* (2.2.9). *Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$. *Then*

$$\left| \max_{a_n \le k \le \lfloor \frac{n}{2} - b_n \rfloor} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} - \max_{a_n \le k \le \lfloor \frac{n}{2} \rfloor} \frac{\left|\tilde{B}_n(k)\right|}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}} \right|$$

$$= O_P\left((b_n \log\log(n/b_n))^{1/2}\right) \quad (n \to \infty). \tag{3.1.32}$$

Proof. The proof follows the lines in the proof of Lemma 2.2.6. □

Lemma 3.1.7. *Suppose that Assumption B holds and let the process* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *be as defined in* (2.2.8) *and* (2.2.9). *Let* $\{b_k,\ k \ge 1\}$ *be a sequence for which* $1 \le b_n \le \frac{n}{2}$, $b_n \uparrow \infty$ *and* $\frac{b_n}{n} \downarrow 0$ $(n \to \infty)$.

Then

$$\left| \max_{\lfloor \frac{n}{2} + b_n \rfloor \le k \le n - a_n} \frac{\left| \tilde{B}_n(k) \right|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} - \max_{\lfloor \frac{n}{2} \rfloor + 1 \le k \le n - a_n} \frac{\left| \tilde{B}_n(k) \right|}{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}} \right|$$

$$= O_P\left(\left(b_n \log\log(n/b_n)\right)^{1/2} \right) \quad (n \to \infty). \tag{3.1.33}$$

PROOF. The proof follows the lines in the proof of Lemma 2.2.7. □

LEMMA 3.1.8. *Suppose that Assumption B holds and let the processes* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *and* $\{B_n(t),\ \frac{1}{n} \le t \le \frac{n-1}{n}\}$ *as in* (2.2.8) - (2.2.9) *and* (2.2.10) - (2.2.11), *respectively. Then*

$$\left| \max_{a_n \le k \le \lfloor \frac{n}{2} \rfloor} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| \tilde{B}_n(k) \right| - \sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left(\frac{n}{nt(n-nt)} \right)^{1/2} \left| B_n(t) \right| \right|$$

$$= o_P\left((\log\log n)^{1/2} \right) \quad (n \to \infty). \tag{3.1.34}$$

PROOF. Observe

$$\max_{a_n \le k \le \lfloor \frac{n}{2} \rfloor} |W_{1n}(k)| = \sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} |W_{1n}(\lfloor nt \rfloor)|$$

and

$$\sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left(\frac{\lfloor nt \rfloor}{n} \left(1 - \frac{\lfloor nt \rfloor}{n} \right) \right)^{-1/2} = O\left(a_n^{-1/2} \right) \quad (n \to \infty).$$

Moreover,

$$\sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} |W_{1n}(\lfloor nt \rfloor) - W_{1n}(nt)| = \sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} |W_{1n}(\lfloor nt \rfloor) - W_{1n}(\lfloor nt \rfloor + r_n(t))| .$$

where $0 \le r_n(t) < 1$. Thus

$$\sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} |W_{1n}(\lfloor nt \rfloor) - W_{1n}(nt)| \le \sup_{1 \le t \le n - \ell_n} \sup_{0 \le s \le \ell_n} |W_{1n}(t+s) - W_{1n}(t)| .$$

Let $\{\ell_n,\ n \ge 1\}$ be a sequence for which $\ell_n \uparrow \infty$. Then via Hanson and Russo [56, Display (3.12b)] we arrive at

$$\sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)} \right)^{1/2} |W_{1n}(\lfloor nt \rfloor) - W_{1n}(nt)| = O_P\left(\frac{\ell_n}{a_n} \right) \tag{3.1.35}$$

as $n \to \infty$. Given some Wiener process, consider the random variable

$$\Delta_n(t) = \left| \left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)} \right)^{1/2} W(nt) - \left(\frac{n}{nt(n-nt)} \right)^{1/2} W(nt) \right| .$$

There exists some event A satisfying $P(A) = 1$, so that for every $\omega \in A$ we can find $\frac{a_n}{n} \leq t_n(\omega) \leq \frac{1}{2}$, such that

$$(\log\log n)^{1/2} \left| \Delta_n(t_n)(\omega) - \sup_{\frac{a_n}{n} \leq t \leq \frac{1}{2}} \Delta_n(t)(\omega) \right| < \epsilon,$$

where $t_n(\omega)$ depends on $\epsilon > 0$. Therefore, in order to establish that $\sup_{a_n/n \leq t \leq 1/2} \Delta_n(t) = o((\log\log n)^{-1/2})$ holds almost surely as $n \to \infty$, it suffices to prove the following claim.

CLAIM. *For each sequence* $\{t_n,\ n \geq 1\}$, *satisfying* $\frac{a_n}{n} \leq t_n \leq \frac{1}{2}$,

$$\Delta_n(t_n) = o\left((\log\log n)^{-1/2}\right) \quad a.s. \tag{3.1.36}$$

holds as $n \to \infty$.

PROOF OF CLAIM. Consider first the case that $\{t_n,\ n \geq 1\}$ is bounded away from zero, i.e. $\liminf_{n\to\infty} t_n > \epsilon_0$ for some $0 < \epsilon_0 < 1/2$. Then

$$\Delta_n(t_n) \leq \sup_{\epsilon_0 \leq t \leq \frac{1}{2}} \frac{|W(nt)|}{\sqrt{n}} \left| \left(\frac{\lfloor nt \rfloor}{n} \left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2} \right|.$$

Since $\frac{\lfloor nt \rfloor}{n} = t - \frac{r_n(t)}{n}$, where $0 \leq r_n(t) < 1$, we have, as $n \to \infty$,

$$\begin{aligned}
&\sup_{\epsilon_0 \leq t \leq \frac{1}{2}} \left| \left(\frac{\lfloor nt \rfloor}{n} \left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2} \right| \\
&\leq \sup_{\epsilon_0 \leq t \leq \frac{1}{2}} \left| \left(t - \frac{r_n(t)}{n} \left(1 - t - \frac{r_n(t)}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2} \right| \\
&\leq \sup_{\epsilon_0 \leq t \leq \frac{1}{2}} \left| g\left(t - \frac{r_n(t)}{n}\right) - g(t) \right|,
\end{aligned}$$

where $g(t) = (t(1-t))^{-1/2}$ and $0 \leq r_n < 1$. Since $g'(t) = -\frac{1}{2}(t(1-t))^{-3/2}(1-2t)$, an application of the mean value theorem implies

$$\Delta_n(t_n) \leq \sup_{\epsilon_0 \leq t \leq \frac{1}{2}} \frac{|W(nt)|}{\sqrt{n}} \sup_{\epsilon_0 \leq t \leq \frac{1}{2}} \frac{g'(t)}{n}.$$

The functional law of the iterated logarithm, cf. e.g. Csörgő and Révész [26, display (1.3.2)], implies $\Delta_n(t_n) = o((\log\log n)^{-1/2})$ almost surely as $n \to \infty$.

Secondly, since $\{t_n,\ n \geq 1\}$ is bounded by hypothesis, it remains to consider $\liminf_{n\to\infty} t_n = 0$. In light of the first case, we can assume

$\lim_{n\to\infty} t_n = 0$ and $t_n = \frac{s_n}{n}$, where $a_n \le s_n \le n/2$. Similarly, we have

$$\begin{aligned}\Delta_n(t_n) &\le \frac{|W(s_n)|}{n^{1/2}} \frac{\left|g'\left(\frac{s_n}{n} - \frac{r_n}{n}\right)\right|}{n} \\ &\le \frac{|W(s_n)|}{(s_n \log\log s_n)^{1/2}} \frac{(\log\log s_n)^{1/2}}{s_n} \\ &\le \frac{|W(s_n)|}{(s_n \log\log s_n)^{1/2}} \frac{(\log\log s_n)^{1/2}}{a_n}.\end{aligned}$$

This completes the proof of the claim (3.1.36). □

Towards this end, similarly as in (3.1.35)

$$\begin{aligned}&\sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left(\frac{\lfloor nt \rfloor}{n}\right) \left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}\right)^{1/2} |W_{1n}(\lfloor nt \rfloor) - W_{1n}(nt)| \\ &= O_P\left(\frac{\ell_n}{\sqrt{n}}\right) \quad (n \to \infty). \qquad (3.1.37)\end{aligned}$$

Consider

$$\begin{aligned}&\sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left| \left(\frac{\lfloor nt \rfloor}{n}\right)^{1/2} \left(1 - \frac{\lfloor nt \rfloor}{n}\right)^{-1/2} - \left(\frac{t}{1-t}\right)^{1/2} \right| \\ &= \sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left| h\left(t - \frac{r_n(t)}{n}\right) - h(t) \right|,\end{aligned}$$

where $h(t) = (t/(1-t))^{1/2}$ and $0 \le r_n < 1$. Since $h'(t) = (1-t)^{-2}$, an application of the mean value theorem yields

$$\begin{aligned}&n^{-1/2} \sup_{\frac{a_n}{n} \le t \le \frac{1}{2}} \left| W_{1n}\left(\frac{n}{2}\right) \right| \left| \left(\frac{\lfloor nt \rfloor}{n}\right)^{1/2} \left(1 - \frac{\lfloor nt \rfloor}{n}\right)^{-1/2} - \left(\frac{t}{1-t}\right)^{1/2} \right| \\ &\le 4n^{-3/2} \left| W_{1n}\left(\frac{n}{2}\right) \right|. \qquad (3.1.38)\end{aligned}$$

The assertions follows from (3.1.35) - (3.1.38). □

LEMMA 3.1.9. *Suppose that Assumption B holds and let the processes* $\{\tilde{B}_n(k),\ 1 \le k \le n\}$ *and* $\{B_n(t),\ \frac{1}{n} \le t \le \frac{n-1}{n}\}$ *as in* (2.2.8) - (2.2.9) *and* (2.2.10) - (2.2.11) *respectively. Then*

$$\begin{aligned}&\left| \max_{\lfloor \frac{n}{2} \rfloor \le k \le n - a_n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right| - \sup_{\frac{1}{2} \le t \le 1 - \frac{a_n}{n}} \left(\frac{n}{nt(n-nt)}\right)^{1/2} |B_n(t)| \right| \\ &= o_P\left((\log\log n)^{-1/2}\right) \quad (n \to \infty). \qquad (3.1.39)\end{aligned}$$

PROOF. Let $\{\ell_k,\ k \geq 1\}$ be a sequence for which $\ell_n \uparrow \infty$. Similarly as in (3.1.35), we have

$$\sup_{\frac{1}{2} \leq t \leq 1-\frac{a_n}{n}} \left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}\right)^{1/2} |W_{2n}(n - \lfloor nt \rfloor) - W_{2n}(n - nt)|$$

$$= O_P\left(\frac{\ell_n}{\sqrt{a_n}}\right) \tag{3.1.40}$$

as $n \to \infty$. Given some Wiener process, consider the random variable

$$\Delta_n(t) = \left|\left(\frac{n}{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}\right)^{1/2} W(n - nt) - \left(\frac{n}{nt(n-nt)}\right)^{1/2} W(n - nt)\right|$$

There exists some event A satisfying $P(A) = 1$, so that for every $\omega \in A$ we can find $\frac{1}{2} \leq t_n(\omega) \leq 1 - \frac{1}{n}$, such that

$$(\log\log n)^{1/2} \left|\Delta_n(t_n)(\omega) - \sup_{\frac{1}{2} \leq t \leq 1-\frac{a_n}{n}} \Delta_n(t)(\omega)\right| < \epsilon,$$

where $t_n(\omega)$ depends on $\epsilon > 0$. Therefore, in order to establish that $\sup_{1/2 \leq t \leq (n-a_n)/n} \Delta_n(t) = o((\log\log n)^{-1/2})$ holds almost surely as $n \to \infty$, it suffices to prove the following claim.

CLAIM. *For each sequence* $\{t_n,\ n \geq 1\}$, *satisfying* $\frac{1}{2} \leq t_n \leq 1 - \frac{a_n}{n}$,

$$\Delta_n(t_n) = o\left((\log\log n)^{-1/2}\right) \quad a.s. \tag{3.1.41}$$

holds as $n \to \infty$.

PROOF OF CLAIM. Consider first the case that $\{t_n,\ n \geq 1\}$ is bounded away from one, i.e. $\frac{1}{2} \leq \limsup_{n\to\infty} t_n \leq 1 - \epsilon_0$, for some $0 < \epsilon_0 < 1/2$. Then

$$\Delta_n(t_n) \leq \sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \frac{|W(n-nt)|}{\sqrt{n}} \left|\left(\frac{\lfloor nt \rfloor}{n}\left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2}\right|$$

Since $\frac{\lfloor nt \rfloor}{n} = t - \frac{r_n(t)}{n}$, where $0 \leq r_n(t) < 1$, we have, as $n \to \infty$,

$$\sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \left|\left(\frac{\lfloor nt \rfloor}{n}\left(1 - \frac{\lfloor nt \rfloor}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2}\right|$$

$$\leq \sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \left|\left(t - \frac{r_n(t)}{n}\left(1 - t - \frac{r_n(t)}{n}\right)\right)^{-1/2} - (t(1-t))^{-1/2}\right|$$

$$\leq \sup_{\frac{1}{2} \leq t \leq 1-\epsilon_0} \left|g\left(t - \frac{r_n(t)}{n}\right) - g(t)\right|,$$

where $g(t) = (t(1-t))^{-1/2}$ and $0 \le r_n < 1$. Since $g'(t) = -\frac{1}{2}(t(1-t))^{-3/2}(1-2t)$, an application of the mean value theorem implies

$$\Delta_n(t_n) \le \sup_{\frac{1}{2}\le t\le 1-\epsilon_0} \frac{|W(n-nt)|}{\sqrt{n}} \sup_{\frac{1}{2}\le t\le 1-\epsilon_0} \frac{g'(t)}{n}.$$

The functional law of the iterated logarithm, cf. e.g. Csörgő and Révész [26, display (1.3.2)], implies $\Delta_n(t_n) = o((\log\log n)^{-1/2})$ almost surely as $n \to \infty$.

It remains to consider $\limsup_{n\to\infty} t_n = 1$. In light of the first case, we can assume $\lim_{n\to\infty} t_n = 1$ and $t_n = 1 - \frac{s_n}{n}$, where $\frac{1}{2} \le 1 - \frac{s_n}{n} \le 1 - \frac{a_n}{n}$. Similarly, we have

$$\begin{aligned}\Delta_n(t_n) &\le \frac{|W(s_n)|}{n^{1/2}} \frac{\left|g'\left(\frac{s_n}{n} - \frac{r_n}{n}\right)\right|}{n} \\ &\le \frac{|W(s_n)|}{(s_n \log\log s_n)^{1/2}} \frac{(\log\log s_n)^{1/2}}{s_n} \\ &\le \frac{|W(s_n)|}{(s_n \log\log s_n)^{1/2}} \frac{(\log\log s_n)^{1/2}}{a_n}.\end{aligned}$$

This completes the proof of the claim (3.1.41). □

Towards this end, similarly as in (3.1.40)

$$\begin{aligned}&n^{-1/2} \sup_{\frac{1}{2}\le t\le 1-\frac{a_n}{n}} \frac{\left|W_{2n}\left(\lfloor\frac{n}{2}\rfloor\right) - W_{2n}\left(\frac{n}{2}\right)\right|}{\left(\frac{\lfloor nt\rfloor}{n}\right)^{1/2}\left(1 - \frac{\lfloor nt\rfloor}{n}\right)^{-1/2}} \\ &= O_P\left(\frac{\ell_n}{\sqrt{n}}\right) \quad (n\to\infty).\end{aligned} \tag{3.1.42}$$

Observe

$$\begin{aligned}&\sup_{\frac{1}{2}\le t\le 1-\frac{a_n}{n}} \left|\left(\frac{\lfloor nt\rfloor}{n}\right)^{1/2}\left(1 - \frac{\lfloor nt\rfloor}{n}\right)^{-1/2} - \left(\frac{1-t}{t}\right)^{1/2}\right| \\ &= \sup_{\frac{1}{2}\le t\le 1-\frac{a_n}{n}} \left|h\left(t - \frac{r_n(t)}{n}\right) - h(t)\right|,\end{aligned}$$

where $h(t) = ((1-t)/t))^{1/2}$ and $0 \le r_n < 1$. Since $h'(t) = -t^{-2}$, an application of the mean value theorem yields

$$\begin{aligned}&n^{-1/2} \sup_{\frac{1}{2}\le t\le 1-\frac{a_n}{n}} \left|W_{2n}\left(\frac{n}{2}\right)\right| \left|\left(\frac{\lfloor nt\rfloor}{n}\right)^{1/2}\left(1 - \frac{\lfloor nt\rfloor}{n}\right)^{-1/2} - \left(\frac{1-t}{t}\right)^{1/2}\right| \\ &\le \frac{\left|W_{2n}\left(\frac{n}{2}\right)\right|}{n^{3/2}} \sup_{\frac{1}{2}\le t\le 1-\frac{a_n}{n}} \left|h'(t)\right|.\end{aligned} \tag{3.1.43}$$

The assertions follows from (3.1.40) - (3.1.43). □

PROOF OF THEOREM 3.1.1. Let

$$J_n = \left\{\lfloor \frac{n}{2} - b_n \rfloor + 1, \ldots, \lfloor \frac{n}{2} + b_n \rfloor - 1\right\} \quad \text{and} \tag{3.1.44}$$

$$K_n = \{a_n, \ldots, n - a_n\} \setminus J_n. \tag{3.1.45}$$

Consider the random vectors

$$V_{1n} = \left(X_{a_n}, \ldots, X_{\lfloor \frac{n}{2} - b_n \rfloor}\right) \quad \text{and} \quad V_{2n} = \left(X_{\lfloor \frac{n}{2} + b_n \rfloor}, \ldots, X_{n-a_n}\right). \tag{3.1.46}$$

We extend the initial probability space with three uniformly distributed random variables U^*, U_i $(i = 1, 2)$, with two standard normal random variables N_i $(i = 1, 2)$ and with two Brownian bridge processes $\{B_i(t),\ 0 \le t \le 1\}$ $(i = 1, 2)$.

CLAIM 1. *We can construct two Wiener processes*

$$\left\{W_{1n}(t),\ 0 \le t \le \frac{n}{2}\right\} \quad \textit{and} \quad \left\{W_{2n}(t),\ 0 \le t \le \frac{n}{2}\right\} \tag{3.1.47}$$

on the extended probability space, such that W_{in} *is a measurable function of* (U_i, V_{in}, N_i, B_i) $(i = 1, 2)$ *and*

$$\begin{aligned}&\left| \max_{k \in K_n} |G_n(k)| - \sigma \sup_{\frac{a_n}{n} \le t \le 1 - \frac{a_n}{n}} \left(\frac{n}{nt(n - nt)}\right)^{1/2} |B_n(t)| \right| \\ &= o_P\left((\log\log n)^{-1/2}\right) \quad (n \to \infty).\end{aligned} \tag{3.1.48}$$

PROOF OF CLAIM 1. In light of Lemma 3.1.3 - Lemma 3.1.5, an application of Billingsley [11, Lemma 21.1] yields two Wiener processes

$$\left\{W_{1n}(t),\ 0 \le t \le \lfloor \frac{n}{2} - b_n \rfloor\right\} \quad \text{and} \quad \left\{W_{2n}(t),\ 0 \le t \le \lfloor \frac{n}{2} - b_n \rfloor\right\}$$

on the extended probability space, such that W_{in} is a measurable function of (U_i, V_{in}) $(i = 1, 2)$ and

$$\begin{aligned}&\left| \max_{k \in K_n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{S}_n(k) - \frac{k}{n}\tilde{S}_n(n)\right| - \sigma \max_{k \in K_n} \left(\frac{n}{k(n-k)}\right)^{1/2} \left|\tilde{B}_n(k)\right| \right| \\ &= o_P\left((\log\log n)^{-1/2}\right) \quad (n \to \infty).\end{aligned} \tag{3.1.49}$$

For each $s \in [\lfloor \frac{n}{2} - b_n \rfloor, \frac{n}{2}]$ let

$$\begin{aligned}W_{in}(s) = &W_{in}\left(\lfloor \frac{n}{2} - b_n \rfloor\right) + \frac{s - \lfloor \frac{n}{2} - b_n \rfloor}{\sqrt{\frac{n}{2} - \lfloor \frac{n}{2} - b_n \rfloor}} N_i \\ &+ \sqrt{\frac{n}{2} - \lfloor \frac{n}{2} - b_n \rfloor}\, B_i\left(\frac{s - \lfloor \frac{n}{2} - b_n \rfloor}{\frac{n}{2} - \lfloor \frac{n}{2} - b_n \rfloor}\right) \quad (i = 1, 2).\end{aligned} \tag{3.1.50}$$

By this definition, cf. e.g. Csörgő and Révész [26, Proposition 1.4.1], we constitute two Wiener processes

$$\left\{W_{1n}(t),\ 0 \le t \le \frac{n}{2}\right\} \quad \text{and} \quad \left\{W_{2n}(t),\ 0 \le t \le \frac{n}{2}\right\}$$

on the initial probability space such that W_{in} is a measurable function of $\{U_i, V_{in}, N_i, B_i\}$ $(i = 1, 2)$. Therefore

$$\begin{aligned} &\left| \max_{k \in K_n} |G_n(k)| - \sigma \max_{a_n \le k \le n-a_n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| \tilde{B}_n(k) \right| \right| \\ &= o_P\left((\log\log n)^{-1/2} \right) \quad (n \to \infty), \end{aligned} \tag{3.1.51}$$

where Lemma 3.1.2, Lemma 3.1.6 and Lemma 3.1.7 were applied. The claim (3.1.48) follows via Lemma 3.1.8 and Lemma 3.1.9. □

CLAIM 2. *We can construct a random vector* V_{2n}^*, *such that*

$$V_{2n}^* \textit{ is independent of } V_{1n} \quad \textit{and} \quad V_{2n}^* \stackrel{\mathcal{D}}{=} V_{2n}. \tag{3.1.52}$$

Moreover, we can construct a Wiener process $\left\{W_{2n}^*(t),\ 0 \le t \le \frac{n}{2}\right\}$, *such that* W_{2n}^* *is a measurable function of* $(U_2, V_{2n}^*, N_2, B_2)$ *and*

$$\begin{aligned} &\left| \max_{k \in K_n} |G_n(k)| - \sigma \sup_{\frac{a_n}{n} \le t \le 1 - \frac{a_n}{n}} \left(\frac{n}{nt(n-nt)} \right)^{1/2} |B_n^*(t)| \right| \\ &= o_P\left((\log\log n)^{-1/2} \right) \quad (n \to \infty), \end{aligned} \tag{3.1.53}$$

where B_n^* *is defined as function of* W_{1n} *and* W_{2n}^* *as in* (2.2.10) *and* (2.2.11).

PROOF OF CLAIM 2. By Berbee [4, Corollary 4.2.5] there exists a random vector V_{2n}^*, such that V_{2n}^* is a measurable function of (V_{1n}, V_{2n}, U^*) and (3.1.52) is satisfied. Moreover

$$P\left[V_{2n} \ne V_{2n}^*\right] = \beta\left(\sigma\left(V_{1n}\right), \sigma\left(V_{2n}\right)\right). \tag{3.1.54}$$

Towards this end, we introduce the coupled version of $\tilde{S}_n(k)$. Let $\tilde{S}_n^*(k) = S(k)$, $k \in \left\{a_n, \dots, \lfloor \frac{n}{2} - b_n \rfloor\right\}$, and for each $k \in \left\{\lfloor \frac{n}{2} + b_n \rfloor, \dots, n - a_n\right\}$ we put

$$\tilde{S}_n^*(k) = S\left(\lfloor \frac{n}{2} - b_n \rfloor\right) + \sum_{\ell=1}^{k - \lfloor \frac{n}{2} + b_n \rfloor + 1} \pi_\ell V_n^{*(2)}, \tag{3.1.55}$$

where π_ℓ denotes the ℓ-th projection. Hence, as in the proof of Lemma 3.1.2,

$$\max_{k\in K_n}\left||G_n(k)|-\left(\frac{n}{k(n-k)}\right)^{1/2}\left|\tilde{S}_n^*(k)-\frac{k}{n}\tilde{S}_n^*(n)\right|\right|=O_P\left(\left(\frac{b_n}{a_n}\right)^{1/2}\right) \tag{3.1.56}$$

is true on the event $\left\{V_n^{(2)}=V_n^{*(2)}\right\}$. Since $\beta\left(\sigma\left(V_{1n}\right),\sigma\left(V_{2n}\right)\right)\le 2b_n$, the β-mixing condition applies and we have

$$\lim_{n\to\infty}P\left[V_{2n}\ne V_{2n}^*\right]=0. \tag{3.1.57}$$

Whence

$$\max_{k\in K_n}\left||G_n(k)|-\left(\frac{n}{k(n-k)}\right)^{1/2}\left|\tilde{S}_n^*(k)-\frac{k}{n}\tilde{S}_n^*(n)\right|\right|=O_P\left(\left(\frac{b_n}{a_n}\right)^{1/2}\right). \tag{3.1.58}$$

The claim (3.1.53) follows along the lines in the proof of CLAIM 1 with V_{2n} replaced by V_{2n}^*. □

Let $b_n=\log\log\log n$, then both claims together with Lemma 3.1.1 yield

$$\begin{aligned}&\left|\max_{a_n\le k\le n-a_n}|G_n(k)|-\sigma\sup_{\frac{a_n}{n}\le t\le 1-\frac{a_n}{n}}\left(\frac{n}{nt(n-nt)}\right)^{1/2}\left|B_n^*\left(t\right)\right|\right|\\&=o_P\left((\log\log n)^{-1/2}\right)\qquad(n\to\infty).\end{aligned} \tag{3.1.59}$$

Since W_{1n} is a measurable function of $\{U_1,V_{1n},N_1,B_1\}$ and W_{2n}^* is a measurable function of $\{U_2,V_{2n}^*,N_2,B_2\}$, both Wiener processes are independent. Consequently, the process $\left\{n^{-1/2}B_n^*(t),\ 0\le t\le 1\right\}$ is a Brownian bridge. Similarly as in Csörgő and Horváth [24, Theorem A.4.2], using the representation as strictly stationary Ornstein-Uhlenbeck process, cf. e.g. (2.1.72), the assertion follows from (2.1.6). □

PROOF OF THEOREM 3.1.2. Using Theorem 2.1.3, Theorem 2.2.1 and Theorem 3.1.1, we can follow the pattern in the proof of Csörgő and Horváth [24, Theorem A.4.2], that is, the proof of [24, display (A.4.37)]. □

PROOF OF THEOREM 3.1.3. Using Theorem 3.1.2 together with expression (2.1.72), the assertion follows along the lines in the proof of Csörgő and Horváth [24, Theorem A.3.1]. □

3.2. Asymptotics for Rejection Regions

Let Z_n be the tied-down partial sum process in $D[0,1]$, defined by

$$Z_n(t) = \begin{cases} n^{-1/2}\left(S\left(\lfloor (n+1)t \rfloor\right) - \frac{\lfloor (n+1)t \rfloor}{n} S(n)\right) & ,0 \le t < 1; \\ 0 & ,t = 1, \end{cases} \tag{3.2.1}$$

where $S(k)$ denotes the k-th partial sum and $S(0) = 0$. Under Assumption B, an application of the invariance principle of Bradley [12, Theorem 4] implies the weak convergence of Z_n, that is

$$Z_n \xrightarrow{D[0,1]} B_0 \quad (n \to \infty), \tag{3.2.2}$$

where $\{B_0(t),\ 0 \le t \le 1\}$ denotes a Brownian bridge process. Moreover, similarly as in the proofs of Lemma 2.1.1 - Lemma 2.1.4, we can extend the probability space with two sequences of Wiener processes $\{W_{in}(t),\ t \ge 0\}$ $(i = 1, 2)$, such that, as $n \to \infty$,

$$\max_{1 \le k \le n} \frac{|S(k) - \sigma W_{1n}(k)|}{\sqrt{k}} = O_P(1) \tag{3.2.3}$$

and

$$\max_{1 \le k \le n-1} \frac{|S(n) - S(k) - \sigma W_{2n}(n-k)|}{\sqrt{n-k}} = O_P(1), \tag{3.2.4}$$

where $\lim_{n\to\infty} n^{-1}\mathrm{Var}\, S(n) = \sigma^2$. Consider strictly positive functions $q(t)$ on $(0,1)$ that increase in a neighborhood of zero and decrease in a neighborhood of one, such that the following condition is fulfilled:

$$\int_0^1 \frac{1}{t(1-t)} \exp\left(-\frac{cq^2(t)}{t(1-t)}\right) dt < \infty \quad \text{for all} \quad c > 0. \tag{3.2.5}$$

In light of (3.2.1) - (3.2.5), we can follow the construction method given in Horvath [58, Theorem A.1] and we derive a sequence of Brownian bridges $\{B_n(t),\ 0 \le t \le 1\}$, such that

$$\sup_{0<t<1} \frac{|Z_n(t) - \sigma B_n(t)|}{q(t)} = o_P(1) \quad (n \to \infty). \tag{3.2.6}$$

We point out that both Wiener processes in (3.2.3) and (3.2.4) are not assumed to be independent of one another.

Horvath [58] employed his basic result, i.e. [58, Theorem A.1], for testing changes in the mean of linear processes satisfying the strong mixing condition. From a converse statement in [58, Theorem A.1] it follows that a Brownian bridge approximation as in (3.2.6) is not possible for $q(t) = (t(1-t))^{1/2}$. Nevertheless, in the standardized case, that is $q(t) = (t(1-t))^{1/2}$, Horváth established truncation arguments to find two independent Wiener processes and derived a Darling-Erdős limit theorem.

Kirch [63, Remark 3.3.1] pointed out that the latter extreme value asymptotic for linear processes still holds even without any mixing conditions. [63, Theorem 3.3.2] rests upon the so-called "Beveridge-Nelson" decomposition, i.e. the partial sums of linear processes can be decomposed as partial sums of the independent noise variables and negligible remainder terms.

Recently, under strong mixing conditions, Hušková et al. [60] derived related limit theorems for testing changes in the parameters of autoregressive time series based on partial sums of weighted residuals. In particular, they considered weight functions $q(t) = (t(1-t))^{\alpha}$, where $\alpha \in [0, 1/2)$. Moreover, in the case $\alpha = 1/2$, they derived a Darling-Erdős limit theorem using arguments as in Davis et. al. [32], where a kind of asymptotic independence holds due to the mixing assumption.

Similarly as in [60, Theorem 2.2], the following holds under Assumption B, for every $0 < \epsilon < 1$, as $n \to \infty$,

$$\sup_{\epsilon < t < 1-\epsilon} \frac{|Z_n(t)|}{(\sigma^2 t(1-t))^{1/2}} \xrightarrow{\mathcal{D}} \sup_{\epsilon < t < 1-\epsilon} \frac{|B_n(t)|}{(t(1-t))^{1/2}}. \tag{3.2.7}$$

On the one hand the truncated version above seems reasonable under the assumption that no early or late change occurs within the sample of chronologically ordered observations. But on the other hand an increasing amount of early and late observations is neglected as the sample size increases.

In light of the related references and results mentioned in [60] and other sources, we believe that our approximation in Theorem 3.1.2 by standardized Brownian bridge processes based on coupling methods under mixing assumptions is novel. We point out that in the independent case these kind of approximations are originally due to Csörgő and Horváth [24], see also Csörgő [22, Remark 25].

Since we derived Theorem 3.1.2 under a logarithmic decay of the mixing coefficients, the approximation rate in (3.1.20) can not attain the rates in the corresponding result for independent random variables, see Csörgő and Horváth [24, Theorem 1.3.2]. Nevertheless, our rate is strong enough to produce asymptotical results for rejection regions. The following statements generalize Csörgő and Horváth [24, Corollary 1.3.1] to the mixing case. Given a fixed level $0 < \alpha < 1$, consider the quantiles

$$\begin{aligned} z_n &= z_n(\alpha) \\ &= \sup \left\{ x \;:\; P\left[\max_{1 \le k < n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| S(k) - \frac{k}{n} S(n) \right| \le x \right] \le 1 - \alpha \right\} \end{aligned} \tag{3.2.8}$$

and

$$u_n = u_n(\alpha) = \sup\left\{x \ : \ P\left[\sup_{\frac{1}{n}\le t\le 1-\frac{1}{n}} \frac{|B(t)|}{\sqrt{t(1-t)}} \le x\right] = 1-\alpha\right\}. \quad (3.2.9)$$

Similarly as in [24, Corollary 1.3.1], Theorem 3.1.2 and Theorem 3.1.3 imply:

$$\lim_{n\to\infty} P\left[\max_{1\le k<n}\left(\frac{n}{k(n-k)}\right)^{1/2}\left|S(k)-\frac{k}{n}S(n)\right| > u_n\right] = \alpha \quad (3.2.10)$$

and

$$z_n(\alpha) - u_n(\alpha) = o\left((\log\log n)^{-1/2}\right) \quad (n\to\infty). \quad (3.2.11)$$

Recently, Gombay [50] presented a change-point test for changes in simple linear regression models with weakly dependent errors and derived less conservative critical values via using (3.2.9) instead of using the critical values obtained from the extreme value distribution. Assuming that the errors are described by linear processes on an independent noise, her approach rests upon results due to Berkes et al. [6]. Therein, the approximations are established with the so-called "Beveridge-Nelson" decomposition for partial sums of linear processes. No mixing assumptions are required. However, the rates of the standardized Brownian bridge type approximation given in [50, p. 69] can not attain the rates in Csörgő and Horváth [24, Theorem 1.3.2] and are slightly slower than ours in Theorem 3.1.2. Critical values are obtained from (3.2.9) via tail approximations due to Vostrikova [106]. We refer also to Miller and Siegmund [83, Appendix] for related tail approximations.

CHAPTER 4

Strong Approximations for Partial Sums

In the first section we will discuss certain invariance principles for independent random variables. We will derive a variant of a strong approximation result of Shao [95] for independent, not necessarily identically distributed random variables. In the second section we will establish an invariance principle for the law of the iterated logarithm for linear processes with dependent errors. This result will be based on our variant of Shao's embedding result. It will be shown that the approximation can be improved via using a strong approximation result due to Einmahl [39].

4.1. Approximations of Sums of Independent R.V.

Let $\{X_n,\ n \geq 1\}$ be a sequence of independent and identically distributed random variables on a probability space $(\Omega, \mathcal{A}, P)$ with zero means and $\sigma^2 = EX_1^2$ for some constant $0 < \sigma^2 < \infty$. Let Φ be the standard normal distribution function. The central limit theorem says that the standardized partial sums $S_n = \sum_{k=1}^n X_k$ converge in distribution, i.e. for all $t \in \mathbb{R}$ we have

$$\lim_{n\to\infty} P\left[n^{-1/2} S_n \leq t\sigma\right] = \Phi(t).$$

Our first example is a convergence-in-probability version.

Example 4.1. *Let $\{X_n,\ n \geq 1\}$ be a sequence of independent, identically distributed random variables on a probability space $(\Omega, \mathcal{A}, P)$ with zero mean and $EX_1^2 = \sigma^2$ for some constant $0 < \sigma^2 < \infty$. Suppose there exists on the same $(\Omega, \mathcal{A}, P)$ another random variable U, uniformly distributed over $[0,1]$, such that U and $\{X_n,\ n \geq 1\}$ are independent. Then there exists an double array of rowwise independent standard normal random variables $\{Y_{nk},\ 1 \leq k \leq n\}$ on the same $(\Omega, \mathcal{A}, P)$, such that*

$$\sum_{k=1}^n X_k - \sigma \sum_{k=1}^n Y_{nk} = o_P\left(n^{1/2}\right) \quad (n \to \infty).$$

Proof. Let

$$\bar{\sigma}_n^2 = EX_1^2 I\left\{|X_1| < \sqrt{n}\right\} - E^2 X_1 I\left\{|X_1| < \sqrt{n}\right\}$$

and

$$F_n(t) = P\left[n^{-1/2}S_n \le t\bar{\sigma}_n\right], \quad t \in \mathbb{R}.$$

The quantile function F_n^{-1} is defined by

$$F_n^{-1}(u) = \inf\{x : u \le F_n(x)\}, \quad 0 < u < 1.$$

Suppose another probability space $(\Omega_1, \mathcal{A}_1, P_1)$ together with two independent, uniformly distributed random variables U_0 and U_1. For every $\delta > 0$ we have

$$\begin{aligned} & P_1\left[\left|\bar{\sigma}_n F_n^{-1}(U_0) - \sigma\Phi^{-1}(U_0)\right| > \delta\right] \\ & \quad \le P_1\left[\left|\bar{\sigma}_n \Phi^{-1}(U_0) - \sigma\Phi^{-1}(U_0)\right| > \delta/2\right] \\ & \quad + P_1\left[\bar{\sigma}_n \left|F_n^{-1}(U_0) - \Phi^{-1}(U_0)\right| > \delta/2\right] \\ & \quad = I_1 + I_2. \end{aligned}$$

By dominated convergence $\bar{\sigma}_n^2 \to \sigma^2 \quad (n \to \infty)$ and

$$I_1 = P_1\left[\left|\Phi^{-1}(U_0)\right| > \delta/\left(2\left|\bar{\sigma}_n - \sigma\right|\right)\right] \to 0 \quad (n \to \infty).$$

Let $Z = \Phi^{-1}(U_0)$ and observe

$$\begin{aligned} I_2 &\le P_1\left[\Phi(Z) > F_n\left(\delta/(2\bar{\sigma}_n) + Z\right)\right] + P_1\left[F_n\left(Z - \delta/(2\bar{\sigma}_n)\right) > \Phi(Z)\right] \\ &= J_1 + J_2. \end{aligned}$$

Let $0 < \epsilon < 1$. Moreover, we set

$$\Delta_n = \inf\left\{\Phi\left(\delta/(2\bar{\sigma}_n) + z\right) - \Phi(z),\ z \in \left[\Phi^{-1}(\epsilon/2), -\Phi^{-1}(\epsilon/2)\right]\right\}$$

Using the mean value theorem we get $\Delta_n > 0$. Since $\bar{\sigma}_n^2 \to \sigma^2 > 0$, we have $\Delta = \inf_{n\in\mathbb{N}} \Delta_n > 0$. Via

$$J_1 = P_1\left[\Phi\left(\delta/(2\bar{\sigma}_n) + Z\right) - \Phi(Z) < \Phi\left(\delta/(2\bar{\sigma}_n) + Z\right) - F_n\left(\delta/(2\bar{\sigma}_n) + Z\right)\right],$$

we arrive at

$$\begin{aligned} J_1 \le P_1\big[&\Delta < \Phi\left(\delta/(2\bar{\sigma}_n) + Z\right) - F_n\left(\delta/(2\bar{\sigma}_n) + Z\right), \\ & Z \in \left[\Phi^{-1}(\epsilon/2), -\Phi^{-1}(\epsilon/2)\right]\big] + \epsilon. \end{aligned}$$

By the central limit theorem and Slutsky's lemma $F_n(t) \to \Phi(t) \quad (n \to \infty)$ for all $t \in \mathbb{R}$. This implies, cf. e.g. Durrett [36, Exercise 2.6], $\sup_{t\in\mathbb{R}} |F_n(t) - \Phi(t)| \to 0$. Since $\epsilon > 0$ can be as small as we wish, we have, as $n \to \infty$,

$$J_1 \to 0$$

and similarly $J_2 \to 0$. Therefore

$$\lim_{n\to\infty} P_1\left[\left|\sqrt{n}\bar{\sigma}_n F_n^{-1}(U_0) - \sqrt{n}\sigma\Phi^{-1}(U_0)\right| > \delta\sqrt{n}\right] = 0.$$

Finally, we construct a deconvolution of $\sqrt{n}\Phi^{-1}(U_0)$ on the initial probability space. Suppose for each $n \in \mathbb{N}$ another probability space supporting a sequence $\zeta_1, \ldots, \zeta_n$ of independent standard normal random variables. Let

$$\mathcal{L}\left(\{\zeta_k,\ 1 \le k \le n\},\ \sum_{k=1}^{n} \zeta_k\right)$$

be the corresponding law on the Borel sets of $\mathbb{R}^n \times \mathbb{R}$. Obviously

$$\mathcal{L}\left(\sum_{k=1}^{n} \zeta_k\right) = \mathcal{L}\left(\sqrt{n}\Phi^{-1}(U_0)\right).$$

Therefore, via using Billingsley [11, Lemma 21.1], there exists on $(\Omega_1, \mathcal{A}_1, P_1)$ a random element η_n, taking values in $\mathbb{R}^n$ and which is a measurable function of $\sqrt{n}\Phi^{-1}(U_0)$ and U_1, such that

$$\mathcal{L}\left(\{\zeta_k,\ 1 \le k \le n\},\ \sum_{k=1}^{n} \zeta_k\right) = \mathcal{L}\left(\eta,\ \sqrt{n}\Phi^{-1}(U_0)\right).$$

This implies that $\eta_n = (\eta_{n1}, \ldots, \eta_{nn})$ is an n-dimensional standard normal vector satisfying

$$\sqrt{n}\Phi^{-1}(U_0) = \sum_{k=1}^{n} \eta_{nk} \quad P_1 - a.s.$$

With the same construction as above, there exists on $(\Omega, \mathcal{A}, P)$ a sequence $Y_{n1}, \ldots, Y_{nn}$ of independent standard normal random variables, being measurable functions of U and $(X_1, \ldots, X_n)$, such that

$$\mathcal{L}\left(\sqrt{n}\bar{\sigma}_n F_n^{-1}(U_0),\ \eta_n\right) = \mathcal{L}\left(\sum_{k=1}^{n} X_k,\ \{Y_{1n}, \ldots, Y_{nn}\}\right).$$

Observe that $\sum_{k=1}^{n} X_k$ and $\sum_{k=1}^{n} Y_{nk}$ have the same joint distribution as $\sqrt{n}\bar{\sigma}_n F_n^{-1}(U_0)$ and $\sqrt{n}\sigma\Phi^{-1}(U_0)$. $\square$

The construction of the approximating normal sequences on the initial probability space in the proof above is motivated by the method in Billingsley [11, p. 215]. The quantile function techniques used in the proof above can be found in Gänssler and Stute [49, Kapitel 10]. Therein quantile function techniques are used to prove an invariance principle for the law of the iterated logarithm.

Major [78] pointed out that the functional central limit theorem and Strassen's invariance principle for the law of the iterated logarithm are easy consequences of the following strong approximation result.

Theorem (Major (1979)). *Let $\{X_n,\ n \ge 1\}$ be a sequence of independent, identically distributed random variables with zero mean and $EX_1^2 = \sigma^2$ for some constant $0 < \sigma^2 < \infty$. There exists a probability space supporting $\{X_n,\ n \ge 1\}$ together with a sequence of independent, identically distributed* *standard normal random variables* $\{Y_n,\ n \geq 1\}$ *and a numerical sequence* $\left\{\sigma_n^2,\ n \geq 1\right\}$ *satisfying*

$$\sigma_n^2 \to \sigma^2 \quad (n \to \infty),$$

such that

$$\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \sigma_k Y_k = o\left(n^{1/2}\right) \quad a.s. \quad (n \to \infty).$$

As an example we will use Csörgő-Révész estimates for increments of a Wiener process to derive the following invariance principle for the law of the iterated logarithm. Strassen [100, Theorem 2] originally proved a functional version of the invariance principle for the law of the iterated logarithm.

Example 4.2. *Let* $\{X_n,\ n \geq 1\}$ *be a sequence of independent identically distributed random variables with zero mean and* $EX_1^2 = \sigma^2$ *for some constant* $0 < \sigma^2 < \infty$. *There exists a probability space supporting* $\{X_n,\ n \geq 1\}$ *together with a Wiener process* $\{W(t),\ t \geq 0\}$, *such that*

$$\frac{S(n) - W(\sigma n)}{\sqrt{n \log\log n}} = o(1) \quad a.s. \quad (n \to \infty).$$

Proof. The construction of Major [78] yields a sequence of approximating normal random variables $\{Y_n,\ n \geq 1\}$ and a numerical sequence $\left\{\sigma_n^2,\ n \geq 1\right\}$. Let

$$\mathcal{L}\left(\{X_k,\ k \geq 1\},\ \{\sigma_k Y_k,\ k \geq 1\}\right)$$

be the corresponding law on the Borel sets of the polish space $\mathbb{R}^\infty \times \mathbb{R}^\infty$. We can assume $\sigma_n^2 \to 1$. Furthermore, we put $s_0 = 0$ and $s_k = \sum_{i=1}^{k} \sigma_i$ $(k = 1, 2, \dots)$. Let $\{W(t),\ t \geq 0\}$ be a standard Wiener process on another (irrelevant) probability space and let

$$\mathcal{L}\left(\{W(s_k) - W(s_{k-1}),\ k \geq 1\},\ \{W(t),\ t \geq 0\}\right)$$

be the corresponding law on the Borel sets of the polish space $\mathbb{R}^\infty \times C[0, \infty)$. Since disjunct increments of a Wiener process are independent, we get

$$\mathcal{L}\left(\{\sigma_k Y_k,\ k \geq 1\}\right) = \mathcal{L}\left(\{W(s_k) - W(s_{k-1}),\ k \geq 1\}\right).$$

Therefore, via using Berkes and Philipp [5, Lemma A1], we can redefine $\{X_n,\ n \geq 1\}$ together with $\{W(t),\ t \geq 0\}$ on a common probability space, such that

$$\sum_{k=1}^{n} X_k - W(s_n) = o\left(n^{1/2}\right) \quad a.s. \quad (n \to \infty).$$

Let $a_n = s_n - n$. We have

$$\begin{aligned}\sup_{-|a_n| \le s \le |a_n|} |W(n-s) - W(n)| &\le \sup_{0 \le s \le |a_n|} |W(n+s) - W(n)| \\ &\quad + \sup_{0 \le s \le |a_n|} |W(n - a_n + s) - W(n)| \\ &= I_1 + I_2.\end{aligned}$$

By display (1.2.3) of Csörgő and Révész [26, Theorem 1.2.1], we have

$$I_1 = o\left(\sqrt{n \log\log n}\right) \quad a.s.,$$

where $(a_n/n)\log(n/a_n) \to 0 \quad (n \to \infty)$ was applied. Moreover,

$$\begin{aligned}I_2 &\le \sup_{0 \le s \le |a_n|} |W((n-a_n)+s) - W((n-a_n))| \\ &\quad + |W(n-a_n) - W((n-a_n)+a_n)| \\ &\le 2 \sup_{0 \le t \le n-|a_n|} \sup_{0 \le s \le |a_n|} |W(t+s) - W(t)|\end{aligned}$$

Now display (1.2.4) of Csörgő and Révész [26, Theorem 1.2.1], implies

$$I_2 = o\left(\sqrt{n \log\log n}\right) \quad a.s.$$

Therefore we get

$$\sup_{-|a_n| \le s \le |a_n|} |W(n-s) - W(n)| = o\left(\sqrt{n \log\log n}\right) \quad a.s.$$

□

The next example states that the result of Major [78] even implies a convergence-in-probability version of Donsker's invariance principle.

EXAMPLE 4.3. *Let* $\{X_n,\ n \ge 1\}$ *be a sequence of independent identically distributed random variables with zero mean and* $EX_1^2 = \sigma^2$ *for some constant* $0 < \sigma^2 < \infty$. *There exists a probability space supporting* $\{X_n,\ n \ge 1\}$ *together with a sequence of independent identically distributed standard normal random variables* $\{Y_n,\ n \ge 1\}$, *such that*

$$n^{-1/2} \sup_{0<t<1} \left| \sum_{k=1}^{[nt]} X_k - \sigma \sum_{k=1}^{[nt]} Y_k \right| = o_P(1) \quad (n \to \infty).$$

PROOF. Let $0 < \delta < 1$. Using Major [78], we have

$$n^{-1/2} \sup_{\delta \le t < 1} \left| \sum_{k=1}^{[nt]} (X_k - \sigma_k Y_k) \right| = o(1) \quad a.s. \quad (n \to \infty)$$

As a consequence of the Hàjek-Rényi inequality, cf. e.g. Petrov [85, Theorem 2.5], we have

$$n^{-1/2} \sup_{1/n \le t < \delta} \left| \sum_{k=1}^{[nt]} (X_k - \sigma_k Y_k) \right| = O_P(1) \sup_{1/n \le t < \delta} \sqrt{t} \quad (n \to \infty).$$

Since $n^{-1} \sum_{k=1}^{n} (\sigma_k - \sigma)^2 \to 0$, an application of Lévy's inequality, cf. e.g. Petrov [85], yields

$$n^{-1/2} \sup_{0<t<1} \left| \sum_{k=1}^{[nt]} (\sigma_k - \sigma) Y_k \right| = o_P(1) \quad (n \to \infty).$$

Since $\delta > 0$ can be as small as we wish, the assertion follows. □

If the probability space also contains a sequence of independent Brownian bridges, the construction in Csörgő and Révész [26, Proposition 1.4.1] yields a Wiener process $\{W(t),\ t \ge 0\}$, such that

$$\sum_{k=1}^{\ell} Y_k = W(\ell) \quad (\ell = 1, 2, \dots)$$

holds almost surely. Then, via using Csörgő-Révész estimates for the increments of a Wiener process, we have

$$n^{-1/2} \sup_{0<t<1} |S(nt) - \sigma W(nt)| = o_P(1) \quad (n \to \infty).$$

Remark. *An alternative approximation method for independent random variables with finite second moments which implies Strassen's invariance principle in the law of the iterated logarithm and also a convergence-in-probability version of Donsker's invariance principle is due to Einmahl [41, Section 3].*

Given a probability law μ on the real line with distribution function F_μ, the left-continuous inverse is defined by

$$F_\mu^{-1}(p) = \inf\{x \in \mathbb{R} \mid F_\mu(x) \ge p\}, \quad p \in (0,1).$$

Consider a Wiener process $\{W(t),\ t \ge 0\}$. A stopping time T *embeds* the law μ into the Wiener process if $\mathcal{L}(W(T)) = \mu$.

Example 4.4. *For each fixed and real a consider the distribution function $F_a(x) = I\{x \ge a\}$. Suppose U is a uniformly distributed random variable with values in $(0,1)$. If U is independent from the Wiener process $\{W(t),\ t \ge 0\}$, then the random variable*

$$T_a = \inf\left\{t \ge 0 \mid W(t) = F_a^{-1}(U)\right\}$$

embeds the discrete probability measure F_a *into the Wiener process. But* T_a *is not integrable in general.*

PROOF. The left-continuous inverse of the point mass is constant, i.e., for all $p \in (0,1)$, $F_a^{-1}(p) = a$. Hence

$$P\left[W\left(T_a\right) \leq x\right] = I\left\{x \geq a\right\}$$

which yields the first assertion. Moreover, we have

$$T_a = \inf\left\{t \geq 0 \mid W(t) = a\right\} \quad a.s.$$

Therefore, $ET_0 = 0$. But if $a > 0$, the passage time is not integrable, i.e $ET_a = \infty$, cf. e.g. Karatzas and Shreve [62, Remark 8.3]. □

The example above is motivated by an example due to J.L. Doob of a nonintegrable stopping time which embeds every given law, see Obłój [84, p. 331].

Skorokhod [96] introduced the *representation* of a whole random walk $\sum_{i=1}^n X_i$, consisting of independent and identically distributed random variables, as a randomly stopped Wiener process $W(\tau_n)$. Since this so-called "Skorokhod representation" constructs a family of integrable stopping times, it has become one of the main tools to establish invariance principles.

Freedman [48] proved the "Skorokhod representation" via representing a given distribution function F on the real line with mean zero as an average of two-point, mean zero distributions.

THEOREM (Freedman(1971, Lemma 108)). *For each* $u > 0$ *and* $v > 0$, *let* $G(u,v)$ *the distribution function which assigns measure* 1 *to the two points* $\{-u, v\}$ *and has mean zero; and let* $G(0,0)$ *be the distribution function assigning mass* 1 *to the point* 0. *Then there exists a probability space* (Ξ, Σ, m) *with nonnegative random variables* U *and* V *such that:* $U = 0$ *iff* $V = 0$ *and*

$$F(x) = \int_{t \in \Xi} G\left(U(t), V(t)\right)(x) m(dt).$$

One can choose Ξ *as interval and* Σ *as the Borel subsets of* Ξ.

Freedman embeds F by defining the stopping time to be the least t with $W(t) \notin (U, V)$. The random walk can be represented on a suitably enriched probability space.

THEOREM ("Skorokhod representation", Freedman(1971, p. 73)). *Let* $\{W(t),\ t \geq 0\}$ *be a Wiener process on the probability space* $(\Omega, \mathcal{F}, P)$. *Suppose* $(U_1, V_1), (U_2, V_2), \ldots$ *are independent and identically distributed random vectors on* $(\Omega, \mathcal{F}, P)$, *independent of* $\{W(t),\ t \geq 0\}$. *For all* $n \in \mathbb{N}$, *suppose* $U_n \geq 0$ *and* $V_n \geq 0$ *and* $U_n = 0$ *iff* $V_n = 0$, *and* $EG(U_n, V_n)) = F$. *Then there are nonnegative random variables* $\tau_1 \le \tau_2 \le \dots$ *on* $(\Omega, \mathcal{F}, P)$ *such that*

(1) $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$ *are independent, identically distributed and finite;*
(2) $E\tau_1 = \int_{-\infty}^{\infty} x^2 F(dx)$*;*
(3) $W(\tau_1), W(\tau_2) - W(\tau_1), W(\tau_3) - W(\tau_2), \dots$ *are independent and identically distributed and*
(4) $W(\tau_1)$ *has distribution function* F.

See Obłój [84, Chapter 3] for a discussion of the slight differences between the constructions of Skorokhod [96], Freedman [48] and others.

Strassen [101] established the invariance principle with rate for martingales via using the martingale generalization of the "Skorokhod representation." Philipp and Stout [86] established martingale approximations and blocking techniques as general method to derive almost sure invariance principles for sequences of weakly dependent random variables via using the "Skorokhod-Strassen representation" and its refinements. For instance, Bradley [12] used the "embedding" results of Jain et al [61] to establish a strong invariance principle under the strong mixing condition. Recently, Balan and Zamfirescu [2] established a strong approximation result for mixing sequences with infinite variance. Their approach rests upon a refinement of the "Skorokhod-Strassen representation" due to Shao [95]. One main tool in the proof of Shao's refinement are Csörgő-Révész type increment results for the Wiener process due to Hanson and Russo [56].

As an example and for further use, we will prove a variant of Shao [95, Theorem 2.1] for independent, not necessarily identically distributed random variables.

Theorem 4.1.1. *Suppose there exists on the same* $(\Omega, \mathcal{A}, P)$ *an uniformly distributed random variable* U, *independent from the infinite sequence* $\{X_k,\ k \ge 1\}$ *of independent random variables with zero means and*

$$\sum_{k}^{\infty} a_k^{-\nu} E\left|X_k\right|^{2\nu} < \infty \tag{4.1.1}$$

holds for some $1 \le \nu \le 2$ *and a nondecreasing sequence* $\{a_k,\ k \ge 1\}$ *with* $0 < a_n \uparrow \infty \quad (n \to \infty)$. *Let* $b_n = E\left(\sum_{k=1}^{n} X_k\right)^2$ *and assume that for some positive and finite constants* C_1, C_2 *and* n_0 *the following is satisfied*

$$C_1 < \frac{b_n}{a_n} < C_2 \tag{4.1.2}$$

for all $n > n_0$. *Then there exists a Wiener process* $\{W(t), t \geq 0\}$ *such that, as* $n \to \infty$,

$$S(n) - W(b_n) = o\left(\sqrt{a_n \log\log a_n}\right) \quad a.s. \tag{4.1.3}$$

REMARK. *Recognizing Shao [95, Theorem 2.1], one might expect an additional logarithmic term* $\log(b_n/a_n)$ *in the rate. The proof will show that this logarithmic term vanishes due to the extra condition on the fraction* b_n/a_n. *Since Hanson and Russo [56] established upper bounds for certain weighted increments, the crucial point in the proof of the approximation is to derive the small-"oh" rate.*

PROOF. Using the "Skorokhod-Strassen representation", cf. e.g. Hall and Heyde [53, Theorem A.1], there exists on another probability space $(\Omega_1, \mathcal{A}_1, P_1)$ a sequence $\{\tau_k,\ k \geq 1\}$ of independent nonnegative random variables and a Wiener process $\{W(t), t \geq 0\}$ such that

$$E\tau_k^r \leq C_r E|X_k|^{2r} \quad 1 \leq r \leq \nu \tag{4.1.4}$$

for some positive constant C_r depending on r only. Moreover,

$$\mathcal{L}(\{S_n,\ n \geq 1\}) = \mathcal{L}\left(\left\{W\left(\sum_{k=1}^n \tau_k\right),\ n \geq 1\right\}\right) \tag{4.1.5}$$

and $E\tau_k = EX_k^2$. Hence

$$|E\tau_k|^\nu = \|X_k\|_2^{2\nu} \leq \|X_k\|_{2\nu}^{2\nu}. \tag{4.1.6}$$

By (4.1.4)

$$E|\tau_k - E\tau_k|^\nu \leq 2^\nu C_\nu E|X_k|^{2\nu} + 2^\nu |E\tau_k|^\nu. \tag{4.1.7}$$

Now (4.1.6) and (4.1.7), together with (4.1.1) yields

$$\sum_{k=1}^\infty a_k^{-\nu} E|\tau_k - E\tau_k|^\nu < \infty. \tag{4.1.8}$$

By a generalized Kolmogorov three-series theorem from Petrov [85, Theorem 6.4, Assumption (a)] and Kronecker's lemma, we have

$$\sum_{k=1}^n \tau_k - \sum_{k=1}^n E\tau_k = o(a_n) \quad a.s. \quad (n \to \infty). \tag{4.1.9}$$

Thus there is a measurable set $A \in \mathcal{A}_1$ with $P_1(A) = 1$ and a sequence of nonnegative random variables $\{c_n,\ n \geq 1\}$ such that

$$\frac{1}{a_n}\left|\sum_{k=1}^n \tau_k(\omega) - \sum_{k=1}^n E\tau_k\right| \leq c_n(\omega) \tag{4.1.10}$$

for all $\omega \in A$ and $c_n(\omega) \to 0$ $(n \to \infty)$. Choose finite random variables $d_n > 0$ so that for each $\omega \in A$

$$d_n(\omega) \geq c_n(\omega), \quad d_n(\omega) \downarrow 0 \quad \text{but} \quad a_n d_n(\omega) \uparrow \infty \quad (n \to \infty). \tag{4.1.11}$$

Thus for each $\omega \in A$

$$-a_n d_n(\omega) \leq \sum_{k=1}^{n} \tau_k(\omega) - b_n \leq a_n d_n(\omega) \tag{4.1.12}$$

Since $a_n d_n(\omega) \uparrow \infty$ and we are interested in the almost sure growth rate of the increment $W(b_n + (\sum_{k=1}^n \tau_k - b_n)) - W(b_n)$ as $n \to \infty$ in terms of a_n only, it suffices to fix a deterministic sequence $d_n > 0$ so that

$$d_n \downarrow 0 \quad \text{and} \quad a_n d_n \uparrow \infty \quad (n \to \infty) \tag{4.1.13}$$

and consider the increment $\sup_{-d_n a_n \leq s \leq d_n a_n} |W(b_n + s) - W(b_n)|$. Observe

$$\begin{aligned}
&\sup_{-d_n a_n \leq s \leq d_n a_n} |W(b_n + s) - W(b_n)| \\
&\leq \sup_{0 \leq s \leq d_n a_n} |W(b_n + s) - W(b_n)| \\
&\quad + \sup_{0 \leq s \leq d_n a_n} |W(b_n - c_n a_n + s) - W(b_n)| \\
&= I_1 + I_2.
\end{aligned} \tag{4.1.14}$$

By display (3.10b) of Hanson and Russo [56, Theorem 3.2A] we have

$$I_1 = O\left(\sqrt{d_n a_n \left(\log\left(\frac{b_n}{d_n a_n} + 1\right) + \log\log(d_n a_n)\right)}\right) \quad a.s. \tag{4.1.15}$$

Next, observe that by (4.1.2)

$$\frac{d_n \log\left(\frac{b_n}{d_n a_n} + 1\right)}{\log\log a_n} \to 0 \tag{4.1.16}$$

and

$$\frac{d_n \log\log(d_n a_n)}{\log\log a_n} = d_n + \frac{d_n \log\left(1 + \frac{\log d_n}{\log a_n}\right)}{\log\log a_n} \to 0 \tag{4.1.17}$$

as $n \to \infty$. Therefore we arrive at

$$I_1 = o\left(\sqrt{a_n \log\log a_n}\right) \quad a.s. \tag{4.1.18}$$

Next, we have

$$\begin{aligned}
I_2 &\leq \sup_{0 \leq s \leq d_n a_n} |W(b_n - d_n a_n + s) - W(b_n - d_n a_n)| \\
&\quad + |W(b_n - d_n a_n + d_n a_n) - W(b_n - d_n a_n)| \\
&\leq 2 \sup_{0 \leq s \leq d_n a_n} |W(b_n - d_n a_n + s) - W(b_n - d_n a_n)|.
\end{aligned} \tag{4.1.19}$$

By display (3.10b) of Hanson and Russo [56, Theorem 3.2A] we have

$$I_2 = O\left(\sqrt{d_n a_n \left(\log\left(\frac{b_n}{d_n a_n}\right) + \log\log\left(d_n a_n\right)\right)}\right) \quad a.s. \tag{4.1.20}$$

Similarly, we get

$$I_2 = o\left(\sqrt{a_n \log\log a_n}\right) \quad a.s. \tag{4.1.21}$$

By (4.1.18) and (4.1.21) we have

$$W\left(\sum_{k=1}^{n} \tau_k\right) - W\left(b_n\right) = o\left(\sqrt{a_n \log\log a_n}\right) \quad a.s. \quad (n \to \infty). \tag{4.1.22}$$

Consider the following law on the polish space $\mathbb{R}^{\mathbb{N}} \times D[0,\infty)$

$$\nu = \mathcal{L}\left(\left\{W\left(\sum_{k=1}^{n} \tau_k\right),\ k \geq 1\right\}, \left\{W\left(t\right),\ t \geq 0\right\}\right) \tag{4.1.23}$$

Therefore, Billingsley [11, Lemma 21.1] together with (4.1.5) yields a Wiener process $\{W_0\left(t\right),\ t \geq 0\}$ on $(\Omega, \mathcal{A}, P)$ being a function of $\{X_k,\ k \geq 1\}$ and U and having the correct joint distribution, i.e.

$$\mathcal{L}\left(\left\{S_n,\ n \geq 1\right\}, \left\{W_0\left(t\right),\ t \geq 0\right\}\right) = \nu. \tag{4.1.24}$$

hence

$$\sum_{k=1}^{n} X_k - W_0\left(b_n\right) = o\left(\sqrt{a_n \log\log a_n}\right) \quad P - a.s. \quad (n \to \infty). \tag{4.1.25}$$

□

On the one hand it is known that certain rates of strong approximation results which are based on the “Skorokhod representation” cannot be improved any further with this method, cf. Csörgő and Révész [26, p. 95]. The so-called “Hungarian construction” with its “Csörgő-Révész quantile transformation” and “Komlós-Major-Tusnády-Theorems” are powerful techniques to obtain optimal rates, see Csörgő and Révész [25] and Komlós et al. [64]. We refer to Csörgő and Horváth [23] for an introduction and further development of these important results. Lifshits [71, p. 3] pointed out that the “KMT-construction” provides estimates for the approximation of $\sum_{i=1}^{n} X_i$ with a partial sum of n Gaussian random variables for each *fixed* n. In order to get asymptotic results as $n \to \infty$, this technicality can be remedied by the construction of a suitable common probability space, see Lifshits [71, Corollary 2.3]. On the other hand, “Skorokhod's representation” embeds the whole partial sum process in one Wiener process directly.

It is known that the “Skorokhod representation” is limited to prove strong invariance theorems for multidimensional random vectors, see e.g.

Mathematical Review MR515811 written by M. Csörgő. Concerning the independent case, we refer to Zaitsev [111] for a recent contribution and an overview on multidimensional extension of the "KMT-construction". Concerning the dependent case, the coupling results of Berkes and Philipp [5] are seminal for the approximation of weakly dependent random vectors. We refer to Bulinski and Shashkin [19] for a recent application and an elaborated proof of a version of the coupling technique due to Berkes and Philipp [5].

One tool used in the proof of the approximation theorems of Berkes and Philipp [5] is the so-called "Strassen-Dudley Theorem". We refer to Dudley [35] for an exposition of the origins and developments of this important result. Suppose two probability measures ν and μ on the Borel σ-field of a polish space (S, d). If the two measures are "close" with respect to the Prohorov distance, then, according to the "Strassen-Dudley Theorem", there exists a probability measure τ on the Borel σ-field of $S \times S$ with marginals ν and μ; and the measure τ concentrates the mass closely around the diagonal of $S \times S$. In terms of random variables, there exists a vector (X, Y) with prescribed marginals such that $d(X, Y)$ is small.

As an example we use a one-dimensional version of "Yurinskii's estimate" due to Pollard [88] and derive, via using the "Strassen-Dudley Theorem", a convergence-in-probability invariance principle with rate. Reworking Chapter 10.4 in Pollard [88] we can extract the following estimate.

Theorem ("Yurinskii's estimate"). *Let $X_1, \ldots, X_n$ be independent random variables with zero means. Let $\Gamma_n := \sum_{k=1}^n E|X_k|^3 < \infty$ and T_n be a normal random variable with zero mean and variance ES_n^2. Then for every $\delta > 0$ satisfying*

$$\Gamma_n \leq e^{-1}\delta^3 \tag{4.1.26}$$

and every Borel set A of $\mathbb{R}$ the following inequality holds

$$\begin{aligned} P[S_n \in A] \leq & P[T_n \in A^{3\delta}] && (4.1.27)\\ & + \frac{90\Gamma_n}{\delta^3\,(1-e^{-1})}\left(1 + |\log(\delta^3/\Gamma_n)|\right), && (4.1.28)\end{aligned}$$

where $A^{3\delta}$ is the closed set $\{x \in \mathbb{R} : |x-y| \leq 3\delta \text{ for some } y \in A\}$.

Remark. *Pollard [88, p. 245] attributes "Yurinskii's estimate" above to Yurinskii [110]. Moreover, Pollard remarked that the proof of the estimate in [88] is close to a method established by Le Cam [67].*

Example 4.5. *Suppose there exists on the same $(\Omega, \mathcal{A}, P)$ an uniformly distributed random variable U independent from the sequence $\{X_k,\ k \geq 1\}$*

of independent random variables with zero means and $EX_k^2 = 1 \quad (k \geq 1)$. *Assume that for some* $s \in (2,3)$ *the following two statements are true:*

$$\sup_{1 \leq n < \infty} \frac{1}{n} \sum_{k=1}^{n} E\left|X_k\right|^s I\left\{\left|X_k\right|^s \leq k\right\} < \infty \tag{4.1.29}$$

and

$$\sum_{k=1}^{\infty} \frac{1}{k^{2/s}} E\left|X_k\right|^2 I\left\{\left|X_k\right|^s > k\right\} < \infty \tag{4.1.30}$$

Then there exists a double array of rowwise independent standard normal random variables $\{Y_{nk},\ 1 \leq k \leq n\}$ *on the same probability space such that, as* $n \to \infty$,

$$\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} Y_{nk} = O_P\left(n^{1/s}\right). \tag{4.1.31}$$

This example is related to the results of Einmahl [39]. Therein invariance principle for independent random vectors satisfying Lindeberg type conditions are established. Moreover, the choice of the approximation rate $n^{1/s}$ in (4.1.31) is related to the optimal approximation construction given by Major [77] in case of independent, identically distributed random variables with no third moment. The truncation arguments in the proof of the example are motivated by truncation methods in Einmahl [38, Kapitel 3].

PROOF. Let

$$\bar{X}_k = X_k I\left\{\left|X_k\right| \leq H^{-1}(k)\right\}, \tag{4.1.32}$$

where H^{-1} denotes the inverse of $H(t) = t^s,\ t \geq 0$. Put

$$\tilde{X}_k = \bar{X}_k - E\bar{X}_k \quad \text{and} \quad \sigma_k^2 = E\tilde{X}_k^2. \tag{4.1.33}$$

Since $1/t$ is non-increasing as $t \to \infty$, we obtain

$$\begin{aligned} \frac{1}{H^{-1}(k)} E\left|X_k - \bar{X}_k\right| &= \frac{1}{H^{-1}(k)} E\left|X_k\right|^2 \frac{1}{\left|X_k\right|} I\left\{H\left(\left|X_k\right|\right) > k\right\} \\ &\leq \frac{1}{k^{2/s}} E\left|X_k\right|^2 I\left\{H\left(\left|X_k\right|\right) > k\right\}. \end{aligned} \tag{4.1.34}$$

Therefore by (4.1.30), as $n \to \infty$,

$$\sum_{k=1}^{n} E\left|X_k - \bar{X}_k\right| = O\left(H^{-1}(n)\right). \tag{4.1.35}$$

Moreover, since $\left|E\bar{X}_k\right| = \left|EX_k - E\bar{X}_k\right|$, we have similarly to (4.1.34)

$$\sum_{k=1}^{n} \left|E\bar{X}_k\right| = O\left(H^{-1}(n)\right). \tag{4.1.36}$$

From (4.1.35), (4.1.36) together with Markov's inequality, we have

$$S_n - \tilde{S}_n := \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \tilde{X}_k = O_P\left(H^{-1}(n)\right). \tag{4.1.37}$$

Let

$$\Gamma_n := \sum_{k=1}^{n} E|\tilde{X}_k|^3 \quad \text{and} \quad \delta_n := (3M)^{1/3}\, n^{1/s}, \tag{4.1.38}$$

where

$$M = 16 \sup_{1\le n<\infty} \frac{1}{n} \sum_{k=1}^{n} E\,|X_k|^s\, I\left\{|X_k|^s \le k\right\} < \infty. \tag{4.1.39}$$

From Jensen inequality

$$E|\tilde{X}_k|^3 \le 8E\left|\bar{X}_k\right|^3 + 8\left|E\bar{X}_k\right|^3 \le 16E\left|\bar{X}_k\right|^3. \tag{4.1.40}$$

Hence

$$\begin{aligned}
\Gamma_n &\le 16 \sum_{k=1}^{n} E|X_k|^s |X_k|^{3-s} I\left\{|X_k|^s \le k\right\} \\
&\le 16 \sum_{k=1}^{n} k^{(3-s)/s} E|X_k|^s I\left\{|X_k|^s \le k\right\} \\
&\le M n^{3/s} e^{-1} 3 = \delta_n^3 e^{-1}.
\end{aligned} \tag{4.1.41}$$

Thus, from (4.1.26) and (4.1.27), for each normal random variable T_n with variance $E\tilde{S}_n^2$ and each Borel set A of $\mathbb{R}$ and $\delta \ge \delta_n$ we have

$$P[\tilde{S}_n \in A] \le P[T_n \in A^{3\delta}] \tag{4.1.42}$$

$$+ \frac{90\Gamma_n}{\delta^3\left(1-e^{-1}\right)}\left(1 + |\log(\delta^3/\Gamma_n)|\right). \tag{4.1.43}$$

Using the Strassen-Dudley theorem, cf. Dudley [35], for each $\delta \ge \delta_n$ there is a law ν_1^δ on $\mathcal{B}\otimes\mathcal{B}$ with margins

$$\nu_1^\delta(\cdot\times\mathbb{R}) = \mathcal{L}\left(\tilde{S}_n\right) \quad \textit{and} \quad \nu_1^\delta(\mathbb{R}\times\cdot) = \mathcal{L}(T_n) \tag{4.1.44}$$

satisfying

$$\nu_1^\delta\{(s,t) \,:\, |s-t| > 3\delta\} \le \frac{90}{1-e^{-1}}\delta^{-3}\Gamma_n\left(1 + |\log(\delta^3/\Gamma_n)|\right). \tag{4.1.45}$$

We will construct a deconvolution of T_n on the initial probability space. Starting from ν_1^δ, the following construction depends on $\delta \ge \delta_n$, hence on n. Thus, we can construct another probability space $(\Omega_{2n}, \mathcal{A}_{2n}, P_{2n})$

together with P_{2n}-independent standard normal random variables $y_1, \ldots, y_n$ and define the measure ν_{2n} on $\mathcal{B} \otimes \mathcal{B}^n$ by

$$\nu_{2n} = \mathcal{L}\left(\sum_{k=1}^{n} \sigma_k y_k,\ \{y_1, \ldots, y_n\}\right). \tag{4.1.46}$$

Since

$$\nu_1^{\delta}\left(\mathbb{R} \times \cdot\right) = \nu_{2n}\left(\cdot \times \mathbb{R}^n\right), \tag{4.1.47}$$

we can glue together ν_1^{δ} and ν_{2n} via an application of Berkes and Philipp [5, Lemma A1]. This yields a law ν_{3n} on $\mathcal{B} \otimes \mathcal{B} \otimes \mathcal{B}^n$ satisfying

$$\nu_1^{\delta} = \nu_{3n}\left(\cdot \times \cdot \times \mathbb{R}^n\right) \quad \text{and} \quad \nu_{3n}\left(\mathbb{R} \times \cdot \times \cdot\right) = \nu_{2n}. \tag{4.1.48}$$

Observe

$$\mathcal{L}\left(\tilde{S}_n\right) = \nu_{3n}\left(\cdot \times \mathbb{R} \times \mathbb{R}^n\right), \tag{4.1.49}$$

whence, via using Billingsley [11, Lemma 21.1], there is an n-dimensional random vector $\mathbb{Y}_n = (\mathbb{Y}_{n1}, \ldots, \mathbb{Y}_{nn})$, on the initial probability space $(\Omega, \mathcal{A}, P)$ being a measurable function of $\left(\tilde{X}_1, \ldots, \tilde{X}_n, U\right)$, such that

$$\mathcal{L}\left(\sum_{k=1}^{n} \sigma_k Y_{nk}, \mathbb{Y}_n\right) = \nu_{3n}\left(\mathbb{R} \times \cdot \times \cdot\right) \tag{4.1.50}$$

and with joint law

$$\mathcal{L}\left(\tilde{S}_n, \sum_{k=1}^{n} \sigma_k Y_{nk}\right) = \nu_1^{\delta} \tag{4.1.51}$$

i.e.

$$\begin{aligned} &P\left[\left|\sum_{k=1}^{n} \tilde{X}_k - \sum_{k=1}^{n} \sigma_k Y_{nk}\right| > 3\delta\right] \\ &\leq \frac{90\Gamma_n}{\delta^3\left(1 - e^{-1}\right)}\left(1 + \left|\log(\delta^3/\Gamma_n)\right|\right). \end{aligned} \tag{4.1.52}$$

For each n, set $\delta = (4M)^{1/3} H^{-1}(n)$. Then, putting together (4.1.38), (4.1.41) and (4.1.52) we have

$$\sum_{k=1}^{n} \tilde{X}_k - \sum_{k=1}^{n} \sigma_k Y_{nk} = O_P\left(H^{-1}\left(n\right)\right). \tag{4.1.53}$$

Towards this end we have

$$\begin{aligned} (1 - \sigma_n)^2 &= \frac{|\sigma_n - 1|}{\sigma_n + 1}\left|\sigma_n^2 - 1\right| \\ &\leq \left(1 + \frac{2}{\sigma_n + 1}\right)\left|\sigma_n^2 - 1\right| \leq 3\left|\sigma_n^2 - 1\right|. \end{aligned} \tag{4.1.54}$$

For each $t > 0$ an application of the Chebyshev inequality gives

$$\begin{aligned} P & \left[\left| \sum_{k=1}^{n} \sigma_k Y_{nk} - \sum_{k=1}^{n} Y_{nk} \right| > tH^{-1}(n) \right] \\ & \leq \frac{1}{\left(tH^{-1}(n)\right)^2} \sum_{k=1}^{n} EY_{nk}^2 (1-\sigma_k)^2 \\ & \leq \frac{3}{\left(tH^{-1}(n)\right)^2} \sum_{k=1}^{n} \left| \sigma_k^2 - 1 \right| . \end{aligned} \tag{4.1.55}$$

From

$$E\left(X_n - \bar{X}_n\right)^2 = EX_n^2 - E\bar{X}_n^2 \tag{4.1.56}$$

and

$$\left(E\bar{X}_n\right)^2 \leq \left(E\left|X_n - \bar{X}_n\right|\right)^2 \tag{4.1.57}$$

together with

$$1 - \sigma_n^2 = EX_n^2 - E\bar{X}_n^2 + \left(E\bar{X}_n\right)^2 \tag{4.1.58}$$

an application of the Cauchy-Schwarz inequality implies

$$\begin{aligned} \left|1 - \sigma_n^2\right| & \leq E\left(X_n - \bar{X}_n\right)^2 + \left(E\left|X_n - \bar{X}_n\right|\right)^2 \\ & \leq 2E\left|X_n\right|^2 I\left\{H\left(\left|X_n\right|\right) > n\right\} . \end{aligned} \tag{4.1.59}$$

Therefore

$$\begin{aligned} & \frac{1}{\left(H^{-1}(n)\right)^2} \sum_{k=1}^{n} \left|1 - \sigma_k^2\right| \\ & \leq \sum_{k=1}^{n} \frac{1}{k^{2/s}} E\left|X_k\right|^2 I\left\{H\left(\left|X_k\right|\right) > k\right\} . \end{aligned} \tag{4.1.60}$$

Therefore, by (4.1.30) and (4.1.55), we obtain, as $n \to \infty$,

$$\sum_{k=1}^{n} \sigma_k Y_{nk} - \sum_{k=1}^{n} Y_{nk} = O_P\left(H^{-1}(n)\right) . \tag{4.1.61}$$

The assertion flows from (4.1.37), (4.1.53) and (4.1.61). □

Einmahl [39] proved an inequality for the so-called "δ-distance" between two probability measures on $C_d[0,1]$, that is, the space of all continuous and $\mathbb{R}^d$-valued functions on $[0,1]$. Using the "Strassen-Dudley Theorem", several strong approximation results for partial sums of independent not necessarily identically distributed random vectors were derived. These particular results from Einmahl [39] are crucial tools in the strong invariance principles for stationary vector-valued processes recently established by Liu and Lin [76]. Here, we focus on the *one-dimensional version* of one of Einmahl's multidimensional results.

THEOREM (Einmahl 1987, Theorem 2). *Let* $\{X_n,\ n \geq 1\}$ *be a sequence of independent random variables with zero means and* $EX_n^2 = \sigma_n^2,\ n \in \mathbb{N}$, *for some constants* $\sigma_n^2 > 0$. *Assume that for some* $s \in (2,4)$ *the following two statements are true:*

$$\sum_{k=1}^{\infty} a_k^{-s} E\,|X_k|^s\, I\,\{|X_k| \leq a_k\} < \infty \tag{4.1.62}$$

and

$$\sum_{k=1}^{\infty} a_k^{-2} E\,|X_k|^2\, I\,\{|X_k| > a_k\} < \infty, \tag{4.1.63}$$

where $\{a_k,\ k \geq 1\}$ *is a positive and monotone increasing sequence. Then a construction of independent standard normal random variables* $Y_1, Y_2, \ldots$ *is possible such that, as* $n \to \infty$,

$$\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \sigma_k Y_k = o\,(a_n) \quad \textit{almost surely.} \tag{4.1.64}$$

This version implies the following invariance principle, whose multidimensional version was used by Liu and Lin [76] to prove strong invariance principles for stationary vector-valued processes.

THEOREM (Due to Einmahl 1987). *Let* $\{X_n,\ n \geq 1\}$ *be a sequence of independent random variables with zero means and* $EX_n^2 = \sigma_n^2,\ n \in \mathbb{N}$, *for some constants* $\sigma_n^2 > 0$. *Assume that for some* $s \in (2,4)$ *the following statement is true:*

$$\sum_{k=1}^{\infty} a_k^{-s} E\,|X_k|^s < \infty. \tag{4.1.65}$$

where $\{a_k,\ k \geq 1\}$ *is a positive and monotone increasing sequence. Then a construction of independent standard normal random variables* $Y_1, Y_2, \ldots$ *is possible such that, as* $n \to \infty$,

$$\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \sigma_k Y_k = o\,(a_n) \quad \textit{almost surely.} \tag{4.1.66}$$

Moreover, Einmahl [39, p. 84] remarked that conditions (4.1.62) and (4.1.63) are fulfilled, if

$$\sum_{k=1}^{\infty} \frac{1}{H\,(a_k)} EH\,(|X_k|) < \infty, \tag{4.1.67}$$

where $H : [0, \infty) \to [0, \infty)$ is a continuous function, such that $t^{-2}H(t)$ is non-decreasing and $t^{-4+r}H(t)$ is non-increasing for some $r > 0$.

One can verify that (4.1.65) implies (4.1.62) and (4.1.63). But (4.1.67) is not sufficient in general. Let us consider an easy counterexample: suppose that independent identically distributed random variables $X_1, X_2, \ldots$ with zero means and $E|X_1|^s < \infty$ for some $s \in (2,4)$ and some $\Phi \in \Phi_C$ satisfy:

$$E\Phi\left(|X_1|^s\right) < \infty,$$

where Φ_C will denote the set of all nonnegative, continuous and monotone increasing functions Φ on $[0,\infty)$, satisfying:

$$\sum_{n=1}^{\infty} \frac{1}{\Phi(n)} < \infty.$$

Now, let $a_k = k^{1/s}$ and $H(t) = \Phi(t^s)$, then (4.1.67) is fulfilled. Moreover, since $t^2/H(t)$ is decreasing as $t \to \infty$, we have

$$\begin{aligned}
&a_k^{-2} E|X_k|^2 I\{|X_k| > a_k\} \\
&= k^{-2/s} E\Phi\left(|X_1|^s\right) \frac{|X_1|^2}{H(|X_1|)} I\{|X_k| > a_k\} \\
&\le \frac{1}{\Phi(k)} EH(|X_1|).
\end{aligned}$$

Hence (4.1.63) is also satisfied. But from the inequality

$$E|X_1|^s I\{|X_1|^s \le k\} \ge E|X_1|^s I\{|X_1|^s \le 1\},$$

condition (4.1.62) can not hold.

Moreover, this easy counterexample indicates that even under a stronger moment condition, e.g. $E|X_1|^s \log^2\left(|X_1|^s + 1\right) < \infty$, that is $\Phi(t) = t\log^2(t+1)$, it is not possible to derive rate $a_n = n^{1/s}$ in the almost sure invariance principle from Einmahl [39, Theorem 2]. Nevertheless, as an example we will show that in a convergence-in-probability version of the invariance principle these rates can be attained. The example is based on the following construction.

Theorem (Einmahl (1987, Proposition 1, (3.3))). *Let $X_1, \ldots, X_n$ be independent random variables with zero means and variance $EX_k^2 = \sigma_k^2$. If $E|X_k|^s < \infty$ for some $2 < s < 4$ then, for any given $\delta > 0$, one can redefine the finite sequence on another probability space $(\Omega_1, \mathcal{A}_1, P_1)$ together* *with independent standard normal random variables* $Y_1, \ldots, Y_n$ *such that*

$$P_1\left[\max_{1\le\ell\le n}\left|\sum_{k=1}^{\ell} X_k - \sum_{k=1}^{\ell}\sigma_k Y_k\right| > C_1\delta\right]$$
$$\le \frac{C_2}{\delta^s}\sum_{k=1}^{n} E\left|X_k\right|^s I\left\{\left|X_k\right| \le \delta\right\}$$
$$+\frac{C_2}{\delta^2}\sum_{k=1}^{n} E|X_k|^2 I\left\{|X_k|^2 > \delta\right\}. \tag{4.1.68}$$

where C_1 *and* C_2 *are positive constants depending on* s *only.*

EXAMPLE 4.6. *Suppose there exists on the same* $(\Omega, \mathcal{A}, P)$ *an uniformly distributed random variable* U *independent from the sequence* $\{X_k,\ k\ge 1\}$ *of independent random variables with zero means and* $EX_k^2 = 1 \quad (k\ge 1)$*. Assume that for some* $s\in(2,3)$ *the following two statements are true:*

$$\sup_{1\le n<\infty}\frac{1}{n}\sum_{k=1}^{n} E\left|X_k\right|^s I\left\{\left|X_k\right|^s \le k\right\} < \infty \tag{4.1.69}$$

and

$$\sum_{k=1}^{\infty}\frac{1}{k^{2/s}} E\left|X_k\right|^2 I\left\{\left|X_k\right|^s > k\right\} < \infty \tag{4.1.70}$$

Then there exists a double array of rowwise independent standard normal random variables $\{Y_{nk},\ 1\le k\le n\}$ *on the same probability space is possible such that, as* $n\to\infty$*,*

$$\max_{1\le\ell\le n}\left|\sum_{k=1}^{\ell} X_k - \sum_{k=1}^{\ell} Y_{nk}\right| = O_P\left(n^{1/s}\right). \tag{4.1.71}$$

REMARK. *Suppose a sequence* $\{X_k,\ k\ge 1\}$ *and some* $\Phi\in\Phi_C$ *satisfying* $\sup_{1\le k<\infty} E\Phi\left(\left|X_k\right|^s\right) < \infty$*, then conditions* (4.1.69) *and* (4.1.70) *are fulfilled.*

The truncation arguments in the proof of the example are motivated by truncation methods in Einmahl [38, Kapitel 3].

PROOF OF REMARK. The inequality

$$E\left|X_k\right|^s I\left\{\left|X_k\right|^s \le k\right\}$$
$$\le E\Phi\left(\left|X_k\right|^s\right)\frac{\left|X_k\right|^s}{\Phi\left(\left|X_k\right|^s\right)} \le const.\ E\Phi\left(\left|X_k\right|^s\right)$$

implies (4.1.69). And since $t/\phi\left(t^{s/2}\right)$ is decreasing as $t \to \infty$, we have

$$\begin{aligned} &E\,|X_k|^2\, I\left\{|X_k|^s > k\right\} \\ &= E\Phi\left(|X_k|^s\right) \frac{|X_k|^2}{\Phi\left(\left(|X_k|^2\right)^{s/2}\right)} I\left\{|X_k|^s > k\right\} \\ &\leq \frac{k^{2/s}}{\Phi(k)} E\Phi\left(|X_k|^s\right) \end{aligned}$$

hence (4.1.70) is fulfilled. □

PROOF. Let

$$\bar{X}_k = X_k I\left\{|X_k| \leq H^{-1}(k)\right\}, \tag{4.1.72}$$

where H^{-1} denotes the inverse of $H(t) = t^s,\ t \geq 0$. Put

$$\tilde{X}_k = \bar{X}_k - E\bar{X}_k \quad \text{and} \quad \sigma_k^2 = E\tilde{X}_k^2. \tag{4.1.73}$$

Similarly to the proof of Example 4.5 we have, via using a Kolmogorov type inequality, cf. e.g. Hall and Heyde [53, Corollary 2.1],

$$\max_{1\leq k\leq n}\left|S_k - \tilde{S}_k\right| = O_P\left(H^{-1}(n)\right). \tag{4.1.74}$$

Using Einmahl [39, Proposition 1] one can redefine the sequence $\tilde{X}_1, \ldots, \tilde{X}_n$ on another probability space $(\Omega_1, \mathcal{A}_1, P_1)$ together with independent standard normal random variables $Y_1, \ldots, Y_n$ such that for some positive constants C_1 and C_2 and any given $t > 0$:

$$\begin{aligned} &P_1\left[\max_{1\leq \ell\leq n}\left|\sum_{k=1}^{\ell}\tilde{X}_k - \sum_{k=1}^{\ell}\sigma_k Y_k\right| > C_1 t H^{-1}(n)\right] \\ &\leq \frac{C_2}{H(t)n}\sum_{k=1}^{n} EH(|\tilde{X}_k|)I\{|\tilde{X}_k| \leq tH^{-1}(n)\} \\ &+ \frac{C_2}{\left(tH^{-1}(n)\right)^2}\sum_{k=1}^{n} E|\tilde{X}_k|^2 I\{|\tilde{X}_k| > tH^{-1}(n)\} \\ &= I_1 + I_2. \end{aligned} \tag{4.1.75}$$

From the inequality

$$EH(|\tilde{X}_k|) \leq H(2)EH(|\bar{X}_k|) + H(2)H\left(E|\bar{X}_k|\right), \tag{4.1.76}$$

we obtain, via Jensen's inequality,

$$I_1 \leq \frac{C_3}{H(t)n}\sum_{k=1}^{n} EH(|\bar{X}_k|). \tag{4.1.77}$$

Since $t^2/H(t)$ is non-increasing as $t \to \infty$, we obtain

$$\begin{aligned} I_2 &\le \frac{C_2}{H(t)n} \sum_{k=1}^{n} EH(|\tilde{X}_k|) \\ &\le \frac{C_3}{H(t)n} \sum_{k=1}^{n} EH(|\bar{X}_k|). \end{aligned} \tag{4.1.78}$$

It follows that

$$P_1 \left[\max_{1\le\ell\le n} \left| \sum_{k=1}^{\ell} \tilde{X}_k - \sum_{k=1}^{\ell} \sigma_k Y_k \right| > C_1 t H^{-1}(n) \right] \le \frac{C_3}{H(t)n} \sum_{k=1}^{n} EH(|\bar{X}_k|). \tag{4.1.79}$$

By (4.1.69), we get, as $n \to \infty$,

$$\max_{1\le\ell\le n} \left| \sum_{k=1}^{\ell} \tilde{X}_k - \sum_{k=1}^{\ell} \sigma_k Y_k \right| = O_{P_1}\left(H^{-1}(n)\right). \tag{4.1.80}$$

Using the same construction as in the proof of Example 4.5 we can construct a double array of rowwise independent standard normal random variables $\{Y_{nk},\ 1 \le k \le n\}$ on the initial probability space such that

$$\max_{1\le\ell\le n} \left| \sum_{k=1}^{\ell} \tilde{X}_k - \sum_{k=1}^{\ell} \sigma_k Y_{nk} \right| = O_{P}\left(H^{-1}(n)\right). \tag{4.1.81}$$

Towards this end, observe

$$\begin{aligned} (1-\sigma_n)^2 &= \frac{|\sigma_n - 1|}{\sigma_n + 1} \left|\sigma_n^2 - 1\right| \\ &\le \left(1 + \frac{2}{\sigma_n + 1}\right) \left|\sigma_n^2 - 1\right| \le 3\left|\sigma_n^2 - 1\right|. \end{aligned} \tag{4.1.82}$$

An application of the Chebyshev inequality gives, for each $t > 0$,

$$\begin{aligned} &P\left[\left| \sum_{k=1}^{n} \sigma_k Y_{nk} - \sum_{k=1}^{n} Y_{nk} \right| > tH^{-1}(n) \right] \\ &\le \frac{1}{\left(tH^{-1}(n)\right)^2} \sum_{k=1}^{n} EY_{nk}^2 (1-\sigma_k)^2 \\ &\le \frac{3}{\left(tH^{-1}(n)\right)^2} \sum_{k=1}^{n} \left|\sigma_k^2 - 1\right|. \end{aligned} \tag{4.1.83}$$

From

$$E\left(X_n - \bar{X}_n\right)^2 = EX_n^2 - E\bar{X}_n^2 \tag{4.1.84}$$

and

$$\left(E\bar{X}_n\right)^2 \leq \left(E\left|X_n - \bar{X}_n\right|\right)^2 \tag{4.1.85}$$

together with

$$1 - \sigma_n^2 = EX_n^2 - E\bar{X}_n^2 + \left(E\bar{X}_n\right)^2 \tag{4.1.86}$$

an application of the Cauchy-Schwarz inequality implies

$$\begin{aligned} \left|1 - \sigma_n^2\right| &\leq E\left(X_n - \bar{X}_n\right)^2 + \left(E\left|X_n - \bar{X}_n\right|\right)^2 \\ &\leq 2E\left|X_n\right|^2 I\left\{H\left(\left|X_n\right|\right) > n\right\}. \end{aligned} \tag{4.1.87}$$

This implies

$$\begin{aligned} &\frac{1}{\left(H^{-1}(n)\right)^2} \sum_{k=1}^{n} \left|1 - \sigma_k^2\right| \\ &\leq \sum_{k=1}^{n} \frac{1}{k^{2/s}} E\left|X_k\right|^2 I\left\{H\left(\left|X_k\right|\right) > k\right\}. \end{aligned} \tag{4.1.88}$$

Therefore, by (4.1.70) and Levy's inequality, cf. e.g. Petrov [85, Theorem 2.2], we obtain

$$\begin{aligned} &P\left[\max_{1\leq \ell\leq n}\left|\sum_{k=1}^{\ell} \sigma_k Y_{nk} - \sum_{k=1}^{\ell} Y_{nk}\right| > tH^{-1}(n)\right] \\ &\leq 2P\left[\left|\sum_{k=1}^{n} \sigma_k Y_{nk} - \sum_{k=1}^{n} Y_{nk}\right| > tH^{-1}(n)\right] \\ &\leq C_4 t^{-2} \end{aligned} \tag{4.1.89}$$

for some positive constant C_4. Hence

$$\max_{1\leq \ell\leq n}\left|\sum_{k=1}^{\ell} \sigma_k Y_{nk} - \sum_{k=1}^{\ell} Y_{nk}\right| = O_P\left(H^{-1}(n)\right). \tag{4.1.90}$$

The assertion flows from (4.1.74), (4.1.81) and (4.1.90). □

4.2. Strong Approximations for Linear Processes

Philipp and Stout [86] established martingale approximations and blocking techniques as general method to derive almost sure invariance principles for sequences of weakly dependent random variables. This methodology was seminal for the theory of strong approximations for dependent random variables, cf. Lin and Lu [72, Part III] for an account. Recently, Wu [109] presented a kind of new version of martingale approximation as a tool to prove strong approximation results for dependent random variables. Liu and Lin [76] developed an approximation by partial sums of m-dependent random vectors. Here we will use an approximation by mixingales, see Appendix A, together with the blocking argument of Aue et al. [1] to establish a strong approximation for linear processes with dependent errors. We will show that the dependent blocks can be approximated by independent ones via a coupling result due to Bradley [12].

THEOREM (Bradley (1983, Theorem 3)). *Suppose* X *and* Y *are random variables taking their values on* $\mathcal{S}$ *and* $\mathbb{R}$*, respectively, where* $\mathcal{S}$ *is a Borel space; suppose* U *is a uniform-*$[0,1]$ *random variable independent of* (X,Y)*; and suppose* q *and* γ *are positive numbers such that* $q \le \|Y\|_\gamma < \infty$*. Then there exists a real-valued random variable* $Y^* = f(X,Y,U)$ *where* f *is a measurable function from* $\mathcal{S}\times\mathbb{R}\times[0,1]$ *into* $\mathbb{R}$*, such that* Y^* *is independent from* X*; the probability distributions of* Y^* *and* Y *are identical and*

$$P\left[|Y^*-Y|\ge q\right] \le 18\left(\|Y\|_\gamma/q\right)^{\gamma/(2\gamma+1)}\left(\alpha\left(\sigma\left(X\right),\sigma\left(Y\right)\right)\right)^{2\gamma/(2\gamma+1)}.$$

We refer to Bradley [12], Doukhan [33, Section 1.2.1] and Merlevède and Peligrad [81] for a broad discussion of this coupling result and related contributions. Concerning coupling methods see also Chapter 2.2 above.

Let us turn to the approximation theorem. Here we consider the following model: Let $\{\eta_k,\ k\in\mathbb{Z}\}$ be a sequence of independent and identically distributed random variables with mean zero and $\mathcal{F}_{k-1}$ denotes the σ-algebra generated by the family $\{\dots,\eta_{k-2},\eta_{k-1}\}$. We consider dependent errors

$$\epsilon_k = \sigma_k\eta_k, \quad k\in\mathbb{Z}, \tag{4.2.1}$$

where σ_k is measurable with respect to $\mathcal{F}_{k-1}$ for every $k\in\mathbb{Z}$.

We assume that $\{\epsilon_k,\ k\in\mathbb{Z}\}$ is a strictly stationary and strongly mixing sequence, i.e $\alpha(n)\to 0 \quad (n\to\infty)$.

Under these considerations the sequence $\{y_k,\ k\ge 1\}$ is defined as stationary solution of the autoregressive scheme

$$y_k = \phi y_{k-1} + \epsilon_k, \quad k = 1,2,\dots, \tag{4.2.2}$$

where $-1<\phi<1$ is a fixed parameter.

Our assumptions on the dependent errors are fulfilled for geometrically ergodic and strictly stationary augmented GARCH processes, see Chaper 1 above.

The main result is an invariance principle for the law of the iterated logarithm.

THEOREM 4.2.1. *Assume that* (4.2.1) *and* (4.2.2) *hold,* $\{\eta_k,\ k\in\mathbb{Z}\}$ *is a sequence of independent and identically distributed random variables with mean zero and* $\{\epsilon_k,\ k\in\mathbb{Z}\}$ *is a strictly stationary and strongly mixing* *sequence, such that* $\alpha(n) = O(\theta^n)$ $(n \to \infty)$ *for some* $0 < \theta < 1$ *and* $E|\epsilon_1|^r < \infty$ *for some* $r > 2$. *Then we can redefine the sequence* $\{y_k,\ k \geq 1\}$ *without changing its distribution on a new probability space together with a Wiener process* $\{W(t),\ t \geq 0\}$ *such that*

$$\frac{\left|\sum_{\ell=1}^n y_\ell - \Gamma^{1/2}W(n)\right|}{\sqrt{n \log\log n}} \to 0 \quad a.s. \quad (n \to \infty), \tag{4.2.3}$$

where $\Gamma = (1-\phi^2)^{-1}E\epsilon_1^2 + (2\phi E\epsilon_1^2)/((1+\phi)(1-\phi)^2)$.

Since this result is based on Theorem 4.1.1, which is essentially a representation for partial sums in the sense of Skorokhod-Strassen, the strong approximation above can be regarded as Strassen type almost sure invariance principle.

Using a stronger approximation result due to Einmahl [39], which was already employed by Liu and Lin [76], we can improve the rate of approximation under the same assumptions.

THEOREM 4.2.2. *Assume that* (4.2.1) *and* (4.2.2) *hold,* $\{\eta_k,\ k \in \mathbb{Z}\}$ *is a sequence of independent and identically distributed random variables with mean zero and* $\{\epsilon_k,\ k \in \mathbb{Z}\}$ *is a strictly stationary and strongly mixing sequence, such that* $\alpha(n) = O(\theta^n)$ $(n \to \infty)$ *for some* $0 < \theta < 1$ *and* $E|\epsilon_1|^r < \infty$ *for some* $r > 2$. *Then we can redefine the sequence* $\{y_k,\ k \geq 1\}$ *without changing its distribution on a new probability space together with a Wiener process* $\{W(t),\ t \geq 0\}$ *such that*

$$\left|\sum_{\ell=1}^n y_\ell - \Gamma^{1/2}W(n)\right| \ll n^{\frac{1}{2}-\kappa_0} \quad a.s. \quad (n \to \infty), \tag{4.2.4}$$

where $\Gamma = (1-\phi^2)^{-1}E\epsilon_1^2 + (2\phi E\epsilon_1^2)/((1+\phi)(1-\phi)^2)$ *and* κ_0 *is a constant depending on* r *only.*

REMARK. *If* $2 < r \leq 4$, *one can put* $\eta = 3$ *in the proof and derive that for each* $0 < \kappa < 1 - 2/r$ *the rate of approximation in* (4.2.4) *is* $O(n^{\frac{1}{2}-\frac{\kappa}{6}})$ *almost surely as* $n \to \infty$. *If* $r > 4$, *one can put* $\eta = 3$ *in the proof and derive that the order obeys* $O(n^{\frac{5}{12}}\sqrt{\log n})$ *almost surely as* $n \to \infty$.

The proof of both theorems will use two essential features from the method of Aue et al. [1]. In order to approximate $\sum_{\ell=1}^n y_\ell$, it suffices to consider truncated versions $\tilde{y}_\ell$ defined by

$$\tilde{y}_k = \sum_{\ell=0}^{k^\rho} \phi^\ell \epsilon_{k-\ell} \quad \text{for some fixed } \rho,\ (0 < \rho < 1). \tag{4.2.5}$$

The resulting sequence $\tilde{y}_1, \tilde{y}_2 \ldots$ is then merged into a sequence of consecutive "blocks" $X_1, Y_1, X_2, Y_2, \ldots$. Following the method in Aue et al. [1], we introduce "big" blocks

$$X_k = \sum_{\ell=k^\eta}^{(k+1)^\eta-(k+2)^\tau} \tilde{y}_\ell \quad (k = 1, 2, \dots) \tag{4.2.6}$$

and "small" blocks

$$Y_k = \sum_{\ell=(k+1)^\eta-(k+2)^\tau+1}^{(k+1)^\eta-1} \tilde{y}_\ell \quad (k = 1, 2, \dots), \tag{4.2.7}$$

where η, ρ and τ are nonnegative constants such that

$$\tau < \eta - 1, \quad 1 + \rho < \eta \quad \text{and} \quad \eta\rho < \tau. \tag{4.2.8}$$

The proofs are based on a series of lemmas.

LEMMA 4.2.1. *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_k,\ k \geq 1\}$ *be defined as in* (4.2.5). *Then for all* $k \geq 1,\ m \geq 0$

$$\|E(\tilde{y}_k|\mathcal{F}_{k-m})\|_r \leq \phi^m \frac{\|\varepsilon_1\|_r}{1-\phi}. \tag{4.2.9}$$

REMARK. *The inequality implies that the sequence* $\{(\tilde{y}_k, \mathcal{F}_k),\ k \geq 1\}$ *is an* $\mathcal{L}^r$*-mixingale (see Appendix A) with respect to the filtration* $\{\mathcal{F}_k,\ k \geq 1\}$, *where* $\mathcal{F}_k$ *denotes the* σ*-algebra generated by the family* $\{\dots, \eta_{k-1}, \eta_k\}$.

PROOF. If $\ell \geq m$, then $\sigma_{k-\ell}$ and $\eta_{k-\ell}$ are measurable with respect to $\mathcal{F}_{k-m}$. Hence

$$E(\sigma_{k-\ell}\eta_{k-\ell}|\mathcal{F}_{k-m}) = \sigma_{k-\ell}\eta_{k-\ell} \quad a.s.$$

Next, consider the case $0 \leq \ell \leq m-1$. The sigma-algebra generated by $\eta_{k-\ell}$ is independent from $\mathcal{F}_{k-m}$. This implies

$$E(\sigma_{k-\ell}\eta_{k-\ell}|\mathcal{F}_{k-m}) = E(\sigma_{k-\ell}E(\eta_{k-\ell}|\mathcal{F}_{k-l-1})|\mathcal{F}_{k-m}) = 0 \quad a.s.$$

We obtain for $m \leq k^\rho$ that

$$E(\tilde{y}_k|\mathcal{F}_{k-m}) = \sum_{\ell=m}^{k^\rho} \phi^\ell \epsilon_{k-\ell} \quad a.s.,$$

and $E(\tilde{y}_k|\mathcal{F}_{k-m}) = 0$ if $k^\rho < m$. Thus, Minkowski's inequality implies the assertion. $\square$

LEMMA 4.2.2. *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_k,\ k \geq 1\}$ *be defined as in* (4.2.5). *Then*

$$P\left[\max_{1\leq j<\infty}\left|\sum_{k=1}^{j} y_k - \sum_{k=1}^{j}\tilde{y}_k\right| < \infty\right] = 1. \tag{4.2.10}$$

PROOF. The assertion follows from

$$\left\| \max_{1\le j<\infty} \left| \sum_{k=1}^{j} (y_k - \tilde{y}_k) \right| \right\|_r \le \|\epsilon_1\|_r \sum_{k=1}^{\infty} \sum_{\ell=k^\rho+1}^{\infty} |\phi|^\ell ,$$

where Minkowski's inequality was applied. The right-hand side series converges due to Cauchy's condensation test, i.e.

$$2^k \exp\left\{2^{k\rho} \ln|\phi|\right\} = \left(2\exp\left\{2^{k\rho} k^{-1} \ln|\phi|\right\}\right)^k \le (2\exp\{-1\})^k$$

implies

$$\sum_{k=1}^{\infty} |\phi|^{k^\rho} < \infty \quad \text{if} \ \ \rho > 0 \ \ \text{and} \ \ 0 < |\phi| < 1. \tag{4.2.11}$$

□

LEMMA 4.2.3. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *Let the sequence* $\{X_k,\ k \ge 1\}$ *be defined as in* (4.2.6). *Then there exists some* $k_0 > 0$ *such that for all* $k \ge k_0$

$$\alpha\left(\sigma\left(X_1, \ldots, X_{k-1}\right), \sigma\left(X_k\right)\right) \le \alpha_\epsilon\left(k^\tau\right). \tag{4.2.12}$$

PROOF. Observe that X_k is a function g_k (say) of a finite subset of $\{\epsilon_k,\ k \ge 1\}$. By (4.2.6), in terms of indices, the last one is $(k+1)^\eta - (k+2)^\tau$. We claim that $k^\eta - k^{\eta\rho}$ is the first one. It suffices to establish the inequality

$$k^\eta - k^{\eta\rho} < (k+1)^\eta - (k+2)^\tau - \left((k+1)^\eta - (k+2)^\tau\right)^\rho, \tag{4.2.13}$$

which is equivalent to

$$k^{\eta\rho}\left(\left(1+\frac{1}{k}\right)^\eta - \frac{(k+2)^\rho}{k^\eta}\right)^\rho - k^{\eta\rho} + (k+2)^\tau < (k+1)^\eta - (k)^\eta.$$

Since the right-hand side is asymptotically equivalent to $\eta k^{\eta-1}$ $(k \to \infty)$, the claim flows from (4.2.8). Therefore, for some $k_0 > 0$, the following representation holds for all $k \ge k_0$

$$X_k = g_k\left(\epsilon_{k^\eta - k^{\eta\rho}}, \ldots, \epsilon_{(k+1)^\eta - (k+2)^\tau}\right).$$

This implies the desired time lag difference between X_{k-1} and X_k, i.e. $(k+1)^\tau - k^{\eta\rho}$. □

LEMMA 4.2.4. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *Let the sequence* $\{Y_k,\ k \ge 1\}$ *be defined as in* (4.2.7). *Then there exists some* $k_0 > 0$ *such that for all* $k \ge k_0$

$$\alpha\left(\sigma\left(Y_1, \ldots, Y_{k-1}\right), \sigma\left(Y_k\right)\right) \le \alpha_\epsilon\left(k^{\eta-1}\right). \tag{4.2.14}$$

PROOF. Observe that Y_k is a function g_k (say) of a finite subset of $\{\epsilon_k,\ k \geq 1\}$. By (4.2.7), in terms of indices, the last one is $(k+1)^\eta - 1$. We claim that $(k+1)^\eta - (k+2)^\tau + 1 - ((k+1)^\eta - (k+2)^\tau + 1)^\rho$ is the first one. It suffices to establish the inequality

$$\begin{aligned}&(k+1)^\eta - (k+2)^\tau + 1 - ((k+1)^\eta - (k+2)^\tau + 1)^\rho \\ &< (k+1)^\eta - 1 - ((k+1)^\eta - 1)^\rho,\end{aligned} \tag{4.2.15}$$

which is equivalent to

$$((k+1)^\eta - 1)^\rho - ((k+1)^\eta - (k+2)^\tau + 1)^\rho < (k+2)^\tau - 2.$$

The claim flows from (4.2.8). Therefore, for some $k_0 > 0$, the following representation holds for all $k \geq k_0$

$$Y_k = g_k\left(\epsilon_{(k+1)^\eta - (k+2)^\tau + 1 - ((k+1)^\eta - (k+2)^\tau + 1)^\rho}, \ldots, \epsilon_{(k+1)^\eta - 1}\right).$$

This implies the time lag difference between Y_{k-1} and Y_k is $(k+1)^\eta - (k+2)^\tau + 1 - k^\eta + 1$ which is asymptotically equivalent to $\eta k^{\eta-1}$ $(k \to \infty)$. □

LEMMA 4.2.5. *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_k,\ k \geq 1\}$ *be defined as in* (4.2.5). *Then there are positive constants* C_1, C_2, n_0 *and* m_0, *such that*

$$C_1 m \leq E\left(\sum_{k=n+1}^{n+m} \tilde{y}_k\right)^2 \leq C_2 m \tag{4.2.16}$$

holds for all $n \geq n_0$ *and* $m \geq m_0$.

PROOF. For each $k \geq 1$, by monotone convergence and Cauchy-Schwarz inequality

$$E\sum_{i,j=0}^{\infty} \phi^{i+j} \left|\epsilon_{1-i}\epsilon_{k-j}\right| < \infty.$$

Therefore, noticing that $\{\varepsilon_k,\ k \in \mathbb{Z}\}$ are strictly stationary martingale differences, we have

$$Ey_1y_k = \sum_{i=0}^{\infty} \phi^{2i+k-1} E\epsilon_1^2. \tag{4.2.17}$$

Moreover, strict stationarity of $\{y_k,\ k \in \mathbb{Z}\}$ yields

$$E\left(\sum_{\ell=n+1}^{n+m} y_\ell\right)^2 = mEy_1^2 + 2\sum_{\ell=2}^{m}(m-\ell+1)Ey_1y_\ell \tag{4.2.18}$$

and via dominated convergence

$$\lim_{m\to\infty}\frac{1}{m}E\left(\sum_{\ell=n+1}^{n+m}y_\ell\right)^2 = Ey_1^2+2\sum_{\ell=2}^{\infty}Ey_1y_\ell$$
$$=\frac{E\epsilon_1^2}{1-\phi^2}+\frac{2\phi E\epsilon_1^2}{(1+\phi)(1-\phi)^2}>0 \qquad (4.2.19)$$

Finally, an application of the Cauchy-Schwarz inequality implies

$$\left|E\left(\sum_{\ell=n+1}^{n+m}y_\ell\right)^2-E\left(\sum_{\ell=n+1}^{n+m}\tilde{y}_\ell\right)^2\right|$$
$$\le\sum_{k=n+1}^{n+m}\sum_{\ell=n+1}^{n+m}\sum_{i=k^\rho+1}^{\infty}\sum_{j=\ell^\rho+1}^{\infty}\phi^{i+j}\left\|\epsilon_1\right\|_2^2$$
$$\le\sum_{k,\ell=n+1}^{\infty}\sum_{i=k^\rho+1}^{\infty}\sum_{j=\ell^\rho+1}^{\infty}\phi^{i+j}\left\|\epsilon_1\right\|_2^2\to 0 \quad (n\to\infty). \qquad (4.2.20)$$

Via using (4.2.19) and (4.2.20), we find n_0 and m_0 such that (4.2.16) holds. □

LEMMA 4.2.6. *Let η,ρ and τ be nonnegative constants satisfying* (4.2.8). *Let the sequence $\{X_k,\ k\ge 1\}$ be defined as in* (4.2.6). *Then there exist nonnegative constants C_3,C_4,C_5,C_6 and $k_0>0$ such that for all $k\ge k_0$*

$$C_3k^{\eta-1}\le EX_k^2\le C_4k^{\eta-1} \quad \textit{and} \qquad (4.2.21)$$

$$C_5k^\eta\le E\left(\sum_{\ell=1}^{k}X_\ell\right)^2\le C_6k^\eta. \qquad (4.2.22)$$

PROOF. Observe that the order of $\tilde{y}$-summands of X_k obeys

$$((k+1)^\eta-(k+2)^\tau-k^\eta+1)\sim\eta k^{\eta-1} \quad (k\to\infty). \qquad (4.2.23)$$

Thus both inequalities in (4.2.21) follow from (4.2.16). Since $\{\varepsilon_k,\ k\in\mathbb{Z}\}$ is a martingale difference sequence, $\{X_k,\ k\ge 1\}$ is a sequence of uncorrelated random variables. Hence

$$C_3\sum_{\ell=1}^{k}\ell^{\eta-1}\le E\left(\sum_{\ell=1}^{k}X_\ell\right)^2\le C_4\sum_{\ell=1}^{k}\ell^{\eta-1}.$$

We claim $\lim_{k\to\infty}k^{-\eta}\sum_{\ell=1}^{k}\ell^{\eta-1}$ exists and is nonnegative. Since

$$\sum_{\ell=1}^{k}(\ell/k)^{\eta-1}\le k,$$

the sequence is bounded. Since

$$\begin{aligned}(k+1)^{-\eta}\sum_{\ell=1}^{k+1}\ell^{\eta-1}-k^{-\eta}\sum_{\ell=1}^{k}\ell^{\eta-1}\\ =\left(\left(\frac{k}{k+1}\right)^{\eta}+\frac{1}{k+1}-1\right)k^{-\eta}\sum_{\ell=1}^{k}\ell^{\eta-1},\end{aligned}$$

the difference is negative because $\eta > 1$ implies the strict inequality $\left(\frac{k}{k+1}\right)^{\eta} < \frac{k}{k+1}$. Whence the sequence is monotonically nondecreasing. This proves the claim and implies (4.2.22). □

LEMMA 4.2.7. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *Let the sequence* $\{Y_k,\ k \geq 1\}$ *be defined as in* (4.2.7). *Then there exist nonnegative constants* C_7, C_8, C_8, C_{10} *and* $k_0 > 0$ *such that for all* $k \geq k_0$

$$C_7 k^{\tau} \leq EY_k^2 \leq C_8 k^{\tau} \quad \textit{and} \tag{4.2.24}$$

$$C_9 k^{\tau+1} \leq E\left(\sum_{\ell=1}^{k} Y_\ell\right)^2 \leq C_{10} k^{\tau+1}. \tag{4.2.25}$$

PROOF. Observe that the order of

$$\tilde{y}\text{-summands of } Y_k \text{ is } (k+2)^{\tau}-1. \tag{4.2.26}$$

The proof of (4.2.24) and (4.2.25) mimics the proof of Lemma 4.2.6. □

LEMMA 4.2.8. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *Let the sequences* $\{X_k,\ k \geq 1\}$ *and* $\{Y_k,\ k \geq 1\}$ *be defined as in* (4.2.6) *and* (4.2.7). *Then there exist nonnegative constants* C_{11}, C_{12} *and* $k_0 > 0$ *such that for all* $k \geq k_0$

$$\|X_k\|_r \leq C_{11} k^{(\eta-1)/2} \quad \textit{and} \tag{4.2.27}$$

$$\|Y_k\|_r \leq C_{12} k^{\tau/2}. \tag{4.2.28}$$

PROOF. Putting together (4.2.23), (4.2.26) and Lemma 4.2.1, the assertions follow from Theorem A.1. □

LEMMA 4.2.9. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_k,\ k \geq 1\}$ *be defined as in* (4.2.5). *Put* $N_k = (k+1)^{\eta}$ $(k = 1, 2, \ldots)$. *Then*

$$\max_{N_k < n \leq N_{k+1}} \left|\sum_{\ell=1}^{n} \tilde{y}_\ell - \sum_{\ell=1}^{N_k} \tilde{y}_\ell\right| = O\left(k^{\eta/2}\right) \quad \textit{a.s.} \quad (k \to \infty). \tag{4.2.29}$$

Proof. Let $\{a_k,\ k \geq 1\}$ be a nonnegative and nondecreasing sequence. Lemma 4.2.1 together with Theorem A.1 imply

$$P\left[\max_{N_k<n\leq N_{k+1}}\left|\sum_{\ell=N_k+1}^{n}\tilde{y}_\ell\right|>a_k\right]\leq\left\|\max_{N_k<n\leq N_{k+1}}\left|\sum_{\ell=N_k+1}^{n}\tilde{y}_\ell\right|\right\|_r^r\Big/a_k^r$$
$$\leq C\left(\frac{N_{k+1}-N_k}{a_k^2}\right)^{r/2},$$

for some nonnegative constant C. Moreover, via using the mean value theorem, the right-hand side of the inequality is asymptotically equivalent to $C\left(\eta k^{\eta-1}/a_k^2\right)^{r/2}$ $(k\to\infty)$. Since $r/2>1$, let $a_k=k^{\eta/2}$, and the Borel-Cantelli lemma yields the assertion. □

Corollary 4.2.1. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *For each* $0<\rho<1$ *let the sequence* $\{\tilde{y}_k,\ k\geq 1\}$ *be defined as in* (4.2.5). *Put* $N_k=(k+1)^\eta$ $(k=1,2,\dots)$. *If* $0<\kappa<1-2/r$, *then*

$$\max_{N_k<n\leq N_{k+1}}\left|\sum_{\ell=1}^{n}\tilde{y}_\ell-\sum_{\ell=1}^{N_k}\tilde{y}_\ell\right|=O\left(k^{(\eta-\kappa)/2}\right)\quad a.s.\quad (k\to\infty). \tag{4.2.30}$$

Proof. Let $\{a_k,\ k \geq 1\}$ be a nonnegative and nondecreasing sequence. Lemma 4.2.1 together with Theorem A.1 imply

$$P\left[\max_{N_k<n\leq N_{k+1}}\left|\sum_{\ell=N_k+1}^{n}\tilde{y}_\ell\right|>a_k\right]\leq\left\|\max_{N_k<n\leq N_{k+1}}\left|\sum_{\ell=N_k+1}^{n}\tilde{y}_\ell\right|\right\|_r^r\Big/a_k^r$$
$$\leq C\left(\frac{N_{k+1}-N_k}{a_k^2}\right)^{r/2},$$

for some nonnegative constant C. Moreover, via using the mean value theorem, the right-hand side of the inequality is asymptotically equivalent to $C\left(\eta k^{\eta-1}/a_k^2\right)^{r/2}$ $(k\to\infty)$. Let $a_k=k^{(\eta-\kappa)/2}$, then $(1-\kappa)r/2>1$ and the Borel-Cantelli Lemma yields the assertion. □

The next lemma is an analogue of Aue et al. [1, Lemma 5.13].

Lemma 4.2.10. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *For each* $0<\rho<1$ *let the sequence* $\{\tilde{y}_k,\ k\geq 1\}$ *be defined as in* (4.2.5). *Let* $N_k=(k+1)^\eta$ $(k=1,2,\dots)$ *and put*

$$R_X(k)=E\left(\sum_{\ell=1}^{k}X_\ell\right)^2\quad\text{and}\quad R_Y(k)=E\left(\sum_{\ell=1}^{k}Y_\ell\right)^2; \tag{4.2.31}$$

$$T_n=E\left(\sum_{\ell=1}^{n}y_\ell\right)^2\quad\text{and}\quad \tilde{T}_n=E\left(\sum_{\ell=1}^{n}\tilde{y}_\ell\right)^2. \tag{4.2.32}$$

Then there exist nonnegative constants C_{13}, C_{14}, C_{15} *and* $k_0 > 0$ *such that for all* $k \geq k_0$ *and* $N_k < n \leq N_{k+1}$

$$\left|\tilde{T}_{N_k} - R_X(k)\right| \leq C_{13} R_X^{1/2}(k) R_Y^{1/2}(k), \tag{4.2.33}$$

$$\left|\tilde{T}_n - \tilde{T}_{N_k}\right| \leq C_{14} k^{(2\eta-1)/2} \quad \textit{and} \tag{4.2.34}$$

$$\left|T_n - \tilde{T}_n\right| \leq C_{15} n^{1/2}. \tag{4.2.35}$$

Proof. Observe

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell = X_1 + Y_1 + \cdots + X_k + Y_k + \tilde{y}_{N_k}. \tag{4.2.36}$$

Furthermore

$$\left|\|\tilde{y}_{N_k}\|_2 - \|y_{N_k}\|_2\right| \leq \sum_{i=N_k^\rho+1}^{\infty} \phi^i \left\|\epsilon_{N_k-i}\right\|_2 < \infty. \tag{4.2.37}$$

Since $\epsilon_{N_k-i} \stackrel{\mathcal{D}}{=} \epsilon_1$ and $y_{N_k} \stackrel{\mathcal{D}}{=} y_1$, we derive

$$\left\|\sum_{\ell=1}^{N_k} \tilde{y}_\ell - \sum_{\ell=1}^{k} X_\ell\right\|_2 \leq R_Y^{1/2}(k) + C, \tag{4.2.38}$$

for some nonnegative and finite constant C. Hence

$$\left|\tilde{T}_{N_k}^{1/2} - R_X^{1/2}(k)\right| \leq R_Y^{1/2}(k) + C. \tag{4.2.39}$$

Moreover, by (4.2.36), we have

$$\left|\tilde{T}_{N_k}^{1/2} + R_X^{1/2}(k)\right| \leq 2R_X^{1/2}(k) + R_Y^{1/2}(k) + C. \tag{4.2.40}$$

Combining (4.2.39) and (4.2.40) with (4.2.22) and (4.2.25) we derive for some k_0 and for each $k \geq k_0$

$$\left|\tilde{T}_{N_k}^{1/2} - R_X^{1/2}(k)\right| \left|\tilde{T}_{N_k}^{1/2} + R_X^{1/2}(k)\right| \leq C_{14} R_X^{1/2}(k) R_Y^{1/2}(k), \tag{4.2.41}$$

where $\tau + 1 < \eta$ was applied. (4.2.41) implies (4.2.33). Concerning the second assertion, observe

$$\left|\tilde{T}_n^{1/2} - \tilde{T}_{N_k}^{1/2}\right| \leq \left\| \max_{N_k < n \leq N_{k+1}} \left| \sum_{\ell=N_k+1}^{n} \tilde{y}_\ell \right| \right\|_2 \quad \text{and} \tag{4.2.42}$$

$$\left|\tilde{T}_n^{1/2} + \tilde{T}_{N_k}^{1/2}\right| \leq 2\tilde{T}_{N_k}^{1/2} + \left\| \max_{N_k < n \leq N_{k+1}} \left| \sum_{\ell=N_k+1}^{n} \tilde{y}_\ell \right| \right\|_2. \tag{4.2.43}$$

Putting together Lemma 4.2.1 with Theorem A.1, we obtain from (4.2.42) and (4.2.43), for some k_0 and for each $k \geq k_0$,

$$\left|\tilde{T}_n^{1/2} - \tilde{T}_{N_k}^{1/2}\right| \leq Ck^{(\eta-1)/2} \quad \text{and} \tag{4.2.44}$$

$$\left|\tilde{T}_n^{1/2} + \tilde{T}_{N_k}^{1/2}\right| \leq Ck^{\eta/2}. \tag{4.2.45}$$

Combining (4.2.44) and (4.2.45) yields (4.2.34). Finally, (4.2.35) is a direct consequence of (4.2.20). □

LEMMA 4.2.11. *Let the sequence* $\{X_k,\ k \geq 1\}$ *be defined as in* (4.2.6). *There exists a sequence* $\{X_k^*,\ k \geq 1\}$ *of independent random variables so that* $\mathcal{L}(X_k) = \mathcal{L}(X_k^*)$ *and*

$$\sum_{\ell=1}^{k} X_\ell - \sum_{\ell=1}^{k} X_\ell^* = O\left(R_X^{1/2}(k)\right) \quad a.s. \quad (k \to \infty), \tag{4.2.46}$$

where $R_X(k)$ *is defined as in* (4.2.31).

PROOF. It follows from (4.2.21) and Cauchy-Schwarz inequality that there exists some $k_0 > 0$ such that for all $k \geq k_0$

$$Ck^{(\eta-1)/2} \leq \|X_k\|_2 \leq \|X_k\|_r\,, \tag{4.2.47}$$

where C is some positive constant. Hence

$$Ck^{\frac{\eta}{2}-1} \leq \|X_k\|_r\,. \tag{4.2.48}$$

Next, we can enrich the probability space with a sequence $\{U_k, k \geq 1\}$ of independent, uniformly distributed random variables and we can apply Bradley [12, Theorem 3], i.e. for each $k \geq k_0$ we can construct a random variables X_k^* being a measurable functions of $(X_1, \ldots, X_k, U_k)$ with the same distribution as X_k and independent of $(X_1, \ldots, X_{k-1})$ and

$$P\left[|X_k - X_k^*| > Ck^{\frac{\eta}{2}-1}\right] \leq 18\left(\frac{\|X_k\|_r}{Ck^{\frac{\eta}{2}-1}}\alpha_k^2\right)^{r/(2r+1)},$$

where $\alpha_k = \alpha\left(\sigma\left(X_1, \ldots, X_{k-1}\right), \sigma\left(X_k\right)\right)$. By (4.2.12) and (4.2.27),

$$P\left[|X_k - X_k^*| > Ck^{\frac{\eta}{2}-1}\right] \leq C\left(k^{1/4}\theta^{k^\tau}\right)^{2r/(2r+1)}.$$

Thus $\{X_k^*, k \geq 1\}$ is a sequence of independent random variables and it follows from the Borel-Cantelli lemma that for almost surely all ω

$$\left|\sum_{\ell=1}^{k} X_\ell(\omega) - \sum_{\ell=1}^{k} X_\ell^*(\omega)\right| = c(\omega) + \sum_{\ell=1}^{k} \ell^{\frac{\eta}{2}-1},$$

where $c(\omega) \geq 0$ is finite. Finally, (4.2.22) yields the assertion. □

COROLLARY 4.2.2. *Let the sequence* $\{X_k,\ k \geq 1\}$ *be defined as in* (4.2.6). *If* $\kappa > 0$, *then there exists a sequence* $\{X_k^*,\ k \geq 1\}$ *of independent random variables so that* $\mathcal{L}(X_k) = \mathcal{L}(X_k^*)$ *and*

$$\sum_{\ell=1}^{k} X_\ell - \sum_{\ell=1}^{k} X_\ell^* = O\left(k^{(\eta-\kappa)/2}\right) \quad a.s. \quad (k \to \infty). \tag{4.2.49}$$

PROOF. It follows from (4.2.21) and Cauchy-Schwarz inequality that there exists some $k_0 > 0$ such that for all $k \geq k_0$

$$Ck^{(\eta-1)/2} \leq \|X_k\|_2 \leq \|X_k\|_r\,, \tag{4.2.50}$$

where C is some positive constant. Hence

$$Ck^{\frac{\eta-\kappa}{2}-1} \leq Ck^{\frac{\eta}{2}-1} \leq \|X_k\|_r\,. \tag{4.2.51}$$

Next, we can enrich the probability space with a sequence $\{U_k, k \geq 1\}$ of independent uniformly distributed random variables and we can apply Bradley [12, Theorem 3], i.e. for each $k \geq k_0$ we can construct a random variables X_k^* being a measurable functions of $(X_1, \ldots, X_k, U_k)$ with the same distribution as X_k and independent of $(X_1, \ldots, X_{k-1})$ and

$$P\left[|X_k - X_k^*| > Ck^{\frac{\eta-\kappa}{2}-1}\right] \leq 18\left(\frac{\|X_k\|_r}{Ck^{\frac{\eta-\kappa}{2}-1}}\alpha_k^2\right)^{r/(2r+1)},$$

where $\alpha_k = \alpha\left(\sigma\left(X_1, \ldots, X_{k-1}\right), \sigma\left(X_k\right)\right)$. By (4.2.12) and (4.2.27),

$$P\left[|X_k - X_k^*| > Ck^{\frac{\eta-\kappa}{2}-1}\right] \leq C\left(k^{\frac{1+\kappa}{4}}\theta^{k^\tau}\right)^{2r/(2r+1)}.$$

Thus $\{X_k^*, k \geq 1\}$ is a sequence of independent random variables and it follows from the Borel-Cantelli lemma that we have for almost surely all ω

$$\left|\sum_{\ell=1}^{k} X_\ell(\omega) - \sum_{\ell=1}^{k} X_\ell^*(\omega)\right| = c(\omega) + \sum_{\ell=1}^{k} \ell^{\frac{\eta-\kappa}{2}-1},$$

where $c(\omega) \geq 0$ is finite. □

LEMMA 4.2.12. *Let the sequence* $\{Y_k,\ k \geq 1\}$ *be defined as in* (4.2.7). *There exists a sequence* $\{Y_k^*,\ k \geq 1\}$ *of independent random variables so that* $\mathcal{L}(Y_k) = \mathcal{L}(Y_k^*)$ *and*

$$\sum_{\ell=1}^{k} Y_\ell - \sum_{\ell=1}^{k} Y_\ell^* = O\left(R_Y^{1/2}(k)\right) \quad a.s. \quad (k \to \infty), \tag{4.2.52}$$

where $R_Y(k)$ *is defined as in* (4.2.31).

Proof. It follows from (4.2.24) and the Cauchy-Schwarz inequality that there exists some $k_0 > 0$ such that for all $k \geq k_0$

$$Ck^{\tau/2} \leq \|Y_k\|_2 \leq \|Y_k\|_r \,, \tag{4.2.53}$$

where C is some positive constant. Hence

$$Ck^{\frac{\tau-1}{2}} \leq \|Y_k\|_r \,. \tag{4.2.54}$$

Next, we can enrich the probability space with a sequence $\{U_k, k \geq 1\}$ of independent, uniformly distributed random variables and we can apply Bradley [12, Theorem 3], i.e. for each $k \geq k_0$ we can construct a random variables Y_k^* being a measurable functions of $(Y_1, \ldots, Y_k, U_k)$ with the same distribution as Y_k and independent of $(Y_1, \ldots, Y_{k-1})$ and

$$P\left[|Y_k - Y_k^*| > Ck^{\frac{\tau-1}{2}}\right] \leq 18 \left(\frac{\|Y_k\|_r}{Ck^{\frac{\tau-1}{2}}} \alpha_k^2\right)^{r/(2r+1)} ,$$

where $\alpha_k = \alpha\left(\sigma\left(Y_1, \ldots, Y_{k-1}\right), \sigma\left(Y_k\right)\right)$. By (4.2.14) and (4.2.28),

$$P\left[|Y_k - Y_k^*| > Ck^{\frac{\tau-1}{2}}\right] \leq C\left(k^{1/4}\theta^{k^{\eta-1}}\right)^{2r/(2r+1)} .$$

Thus $\{Y_k^*, k \geq 1\}$ is a sequence of independent random variables and it follows from the Borel-Cantelli lemma that we have for almost surely all ω

$$\left|\sum_{\ell=1}^{k} Y_\ell(\omega) - \sum_{\ell=1}^{k} Y_\ell^*(\omega)\right| = c(\omega) + \sum_{\ell=1}^{k} \ell^{\frac{\tau-1}{2}},$$

where $c(\omega) \geq 0$ is finite. Finally, (4.2.25) yields the assertion. □

Corollary 4.2.3. *Let the sequence* $\{Y_k,\ k \geq 1\}$ *as in* (4.2.7). *If* $\kappa > 0$, *then there exists a sequence* $\{Y_k^*,\ k \geq 1\}$ *of independent random variables so that* $\mathcal{L}(Y_k) = \mathcal{L}(Y_k^*)$ *and*

$$\sum_{\ell=1}^{k} Y_\ell - \sum_{\ell=1}^{k} Y_\ell^* = O\left(k^{(\tau-\kappa+1)/2}\right) \quad a.s. \quad (k \to \infty). \tag{4.2.55}$$

Proof. It follows from (4.2.24) and Cauchy-Schwarz inequality that there exists some $k_0 > 0$ such that for all $k \geq k_0$

$$Ck^{\tau/2} \leq \|Y_k\|_2 \leq \|Y_k\|_r \,, \tag{4.2.56}$$

where C is some positive constant. Hence

$$Ck^{(\tau-\kappa-1)/2} \leq Ck^{\frac{\tau-1}{2}} \leq \|Y_k\|_r \,. \tag{4.2.57}$$

Next, we can enrich the probability space with a sequence $\{U_k, k \geq 1\}$ of independent, uniformly distributed random variables and we can apply Bradley [12, Theorem 3], i.e. for each $k \geq k_0$ we can construct a random variables Y_k^* being a measurable functions of $(Y_1, \dots, Y_k, U_k)$ with the same distribution as Y_k and independent of $(Y_1, \dots, Y_{k-1})$ and

$$P\left[|Y_k - Y_k^*| > Ck^{\frac{\tau-\kappa-1}{2}}\right] \leq 18\left(\frac{\|Y_k\|_r}{Ck^{\frac{\tau-\kappa-1}{2}}}\alpha_k^2\right)^{r/(2r+1)},$$

where $\alpha_k = \alpha\left(\sigma\left(Y_1, \dots, Y_{k-1}\right), \sigma\left(Y_k\right)\right)$. By (4.2.14) and (4.2.28),

$$P\left[|Y_k - Y_k^*| > Ck^{\frac{\tau-\kappa-1}{2}}\right] \leq C\left(k^{\frac{1+\kappa}{4}}\theta^{k^{\eta-1}}\right)^{2r/(2r+1)}.$$

Thus $\{Y_k^*, k \geq 1\}$ is a sequence of independent random variables and it follows from the Borel-Cantelli lemma that we have for almost surely all ω

$$\left|\sum_{\ell=1}^{k} Y_\ell(\omega) - \sum_{\ell=1}^{k} Y_\ell^*(\omega)\right| = c(\omega) + \sum_{\ell=1}^{k} \ell^{\frac{\tau-\kappa-1}{2}},$$

where $c(\omega) \geq 0$ is finite. □

Lemma 4.2.13. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_k,\ k \geq 1\}$ *be defined as in* (4.2.5). *Let* $N_k = (k+1)^\eta$. *Then we can redefine the sequence* $\{\tilde{y}_k,\ k \geq 1\}$ *without changing its distribution on a new probability space together with a Wiener process* $\{W(t),\ t \geq 0\}$ *such that, as* $k \to \infty$,

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell - W\left(R_X(k)\right) = o\left(\left(R_X(k) L_2 R_X(k)\right)^{1/2}\right) \quad a.s., \tag{4.2.58}$$

where $R_X(k)$ *is defined as in* (4.2.31) *and* $Ln = \log n$ *and* $L_2 n = \log Ln$.

Proof. Observe

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell = X_1 + Y_1 + \cdots + X_k + Y_k + \tilde{y}_{N_k}. \tag{4.2.59}$$

Furthermore

$$\left|\|\tilde{y}_{N_k}\|_2 - \|y_{N_k}\|_2\right| \leq \sum_{i=N_k^\rho+1}^{\infty} \phi^i \|\epsilon_{N_k-i}\|_2 < \infty. \tag{4.2.60}$$

Since $\|\epsilon_{N_k-i}\|_2 = \|\epsilon_1\|_2$, $\|y_{N_k}\|_2 = \|y_1\|_2$ and $\eta > 1$, an application of the Borel-Cantelli lemma implies

$$P\left[|\tilde{y}_{N_k}| > k^{\eta/2},\ i.o.\right] = 0. \tag{4.2.61}$$

Hence, as $k \to \infty$,

$$|\tilde{y}_{N_k}| = o\left(\left(R_X(k) L_2 R_X(k)\right)^{1/2}\right) \quad a.s. \tag{4.2.62}$$

Observing (4.2.25) and $\tau + 1 < \eta$, we derive from (4.2.46) and (4.2.52)

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell - \sum_{\ell=1}^{k} X_\ell^* - \sum_{\ell=1}^{k} Y_\ell^* = o\left((R_X(k) L_2 R_X(k))^{1/2} \right) \quad a.s. \tag{4.2.63}$$

By (4.2.27) and (4.2.28)

$$\sum_{k=1}^{\infty} k^{-\eta r/2} E|X_k|^r < \infty \tag{4.2.64}$$

and

$$\sum_{k=1}^{\infty} k^{-(\tau+1)r/2} E|Y_k|^r < \infty, \tag{4.2.65}$$

where $r/2 > 1$ was applied. Moreover, by (4.2.22) and (4.2.25)

$$C_5 \le k^{-\eta} R_X(k) = k^{-\eta} R_{X^*}(k) \le C_6 \tag{4.2.66}$$

and

$$C_9 \le k^{-(\tau+1)} R_Y(k) = k^{-(\tau+1)} R_{Y^*}(k) \le C_{10}. \tag{4.2.67}$$

Therefore an application of Theorem 4.1.1 yields two Wiener processes $\{W_1(t),\ t \ge 0\}$ and $\{W_2(t),\ t \ge 0\}$ such that

$$\sum_{\ell=1}^{k} X_\ell^* - W_1(R_X(k)) = o\left((R_X(k) L_2 R_X(k))^{1/2} \right) \quad a.s., \tag{4.2.68}$$

and

$$\sum_{\ell=1}^{k} Y_\ell^* - W_2(R_Y(k)) = o\left((R_Y(k) L_2 R_Y(k))^{1/2} \right) \quad a.s. \tag{4.2.69}$$

as $k \to \infty$. Using Petrov [85, Theorem 6.17], we have

$$W_2(R_Y(k)) = o\left(\left(R_Y(k) L^2 R_Y(k) \right)^{1/2} \right) \quad a.s. \quad (k \to \infty). \tag{4.2.70}$$

Since

$$\lim_{k\to\infty} \left(R_Y(k) L^2 R_Y(k) \right) / (R_X(k) L_2 R_X(k)) = 0, \tag{4.2.71}$$

the assertion follows immediately from (4.2.63),(4.2.68) and (4.2.69). □

COROLLARY 4.2.4. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (4.2.8). *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_k,\ k \ge 1\}$ *be defined as in* (4.2.5). *Let* $N_k = (k+1)^\eta$. *If* $0 < \kappa < (\eta - 1) \wedge (1 - 2/r)$ *and* $0 < \tau < \eta - 1 - \kappa$ *then we can redefine the sequence* $\{\tilde{y}_k,\ k \ge 1\}$ *without* *changing its distribution on a new probability space together with a Wiener process* $\{W(t),\ t \geq 0\}$ *such that, as* $k \to \infty$,

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell - W\left(R_X(k)\right) = O\left(k^{(\eta-\kappa)/2}\right) \quad a.s., \tag{4.2.72}$$

where $R_X(k)$ *is defined as in* (4.2.31) *and* $Ln = \log n$ *and* $L_2 n = \log Ln$.

Proof. Observe

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell = X_1 + Y_1 + \cdots + X_k + Y_k + \tilde{y}_{N_k}. \tag{4.2.73}$$

Furthermore

$$\left|\|\tilde{y}_{N_k}\|_2 - \|y_{N_k}\|_2\right| \leq \sum_{i=N_k^\rho+1}^{\infty} \phi^i \|\epsilon_{N_k-i}\|_2 < \infty. \tag{4.2.74}$$

Since $\|\epsilon_{N_k-i}\|_2 = \|\epsilon_1\|_2$, $\|y_{N_k}\|_2 = \|y_1\|_2$ and $0 < \kappa < \eta - 1$, an application of the Borel-Cantelli lemma implies

$$P\left[|\tilde{y}_{N_k}| > k^{(\eta-\kappa)/2},\ i.o.\right] = 0. \tag{4.2.75}$$

Hence, as $k \to \infty$,

$$|\tilde{y}_{N_k}| = O\left(k^{(\eta-\kappa)/2}\right) \quad a.s. \tag{4.2.76}$$

Since $\tau + 1 < \eta$, we derive from (4.2.49) and (4.2.55)

$$\sum_{\ell=1}^{N_k} \tilde{y}_\ell - \sum_{\ell=1}^{k} X_\ell^* - \sum_{\ell=1}^{k} Y_\ell^* = O\left(k^{(\eta-\kappa)/2}\right) \quad a.s. \tag{4.2.77}$$

By (4.2.27) and (4.2.28)

$$\sum_{k=1}^{\infty} k^{-(\eta-\kappa)r/2} E|X_k|^r < \infty \tag{4.2.78}$$

and

$$\sum_{k=1}^{\infty} k^{-(\tau-\kappa+1)r/2} E|Y_k|^r < \infty, \tag{4.2.79}$$

where $(1-\kappa)r/2 > 1$ was applied. Therefore from (4.1.66), i.e. Einmahl's result, there are two sequences of independent, standard normal random variables $\{Z_{1k},\ k \geq 1\}$ and $\{Z_{2k},\ k \geq 1\}$ such that

$$\sum_{\ell=1}^{k} X_\ell^* - \sum_{\ell=1}^{k} \|X_\ell\|_2 Z_{1\ell} = o\left(k^{(\eta-\kappa)/2}\right) \quad a.s., \tag{4.2.80}$$

and

$$\sum_{\ell=1}^{k} Y_\ell^* - \sum_{\ell=1}^{k} \|Y_\ell\|_2 Z_{2\ell} = o\left(k^{(\tau+1-\kappa)/2}\right) \quad a.s. \tag{4.2.81}$$

as $k \to \infty$. Using Csörgő and Révész [26, Proposition 1.4.1] we can interpolate the sum of normal random variables with independent Brownian bridge processes. This construction yields two Wiener processes $\{W_1(t),\ t \geq 0\}$ and $\{W_2(t),\ t \geq 0\}$ such that

$$\sum_{\ell=1}^{k} X_\ell^* - W_1\left(R_X(k)\right) = o\left(k^{(\eta-\kappa)/2}\right) \quad a.s., \tag{4.2.82}$$

and

$$\sum_{\ell=1}^{k} Y_\ell^* - W_2\left(R_Y(k)\right) = o\left(k^{(\tau+1-\kappa)/2}\right) \quad a.s. \tag{4.2.83}$$

as $k \to \infty$. Using Petrov [85, Theorem 6.17], we have

$$W_2\left(R_Y(k)\right) = o\left(\left(R_Y(k) L^2 R_Y(k)\right)^{1/2}\right) \quad a.s. \quad (k \to \infty). \tag{4.2.84}$$

Since $\tau + 1 < \eta - \kappa$,

$$\lim_{k \to \infty} \left(R_Y(k) L^2 R_Y(k)\right) / k^{(\eta-\kappa)} = 0, \tag{4.2.85}$$

the assertion follows immediately from (4.2.77), (4.2.82) and (4.2.83). □

PROOF OF THEOREM 4.2.1. Put $N_k = (k+1)^\eta$ $(k = 1, 2, \dots)$ and let $N_k < n \leq N_{k+1}$. By (4.2.29)

$$I_{1k} = \sum_{\ell=1}^{n} \tilde{y}_\ell - \sum_{\ell=1}^{N_k} \tilde{y}_\ell \ll k^{\eta/2} \quad a.s. \quad (k \to \infty). \tag{4.2.86}$$

Since

$$N_k < n \leq N_{k+1} \quad \text{implies} \quad k \sim n^{1/\eta} \quad (n \to \infty), \tag{4.2.87}$$

we have

$$I_{1n} \ll n^{1/2} \quad a.s. \quad (n \to \infty). \tag{4.2.88}$$

By (4.2.58) and (4.2.22)

$$I_{2k} = \sum_{\ell=1}^{N_k} \tilde{y}_\ell - W\left(R_X(k)\right) = o\left(\left(k^\eta L_2 k^\eta\right)^{1/2}\right) \quad a.s., \tag{4.2.89}$$

which implies

$$I_{2n} = o\left(\left(n L_2 n\right)^{1/2}\right) \quad a.s. \quad (n \to \infty). \tag{4.2.90}$$

We need to estimate the remaining increments. Consider a nonnegative and nondecreasing sequence $\{a_k,\ k \geq 1\}$. Then for each $t \geq a_k$

$$\begin{aligned}
&\sup_{-a_k \leq s \leq a_k} |W(t+s) - W(t)| \\
&\leq \sup_{0 \leq s \leq a_k} |W(t+s) - W(t)| \\
&\quad + \sup_{0 \leq s \leq a_k} |W(t - a_k + s) - W(t)| \\
&\leq \sup_{0 \leq s \leq a_k} |W(t+s) - W(t)| \\
&\quad + \sup_{0 \leq s \leq a_k} |W(t - a_k + s) - W(t - a_k)| \\
&\qquad + |W(t - a_k + a_k) - W(t - a_k)| \\
&\leq \sup_{0 \leq s \leq a_k} |W(t+s) - W(t)| \\
&\quad + 2 \sup_{0 \leq s \leq a_k} |W(t - a_k + s) - W(t - a_k)|.
\end{aligned}$$

Hence for another nonnegative and nondecreasing sequence $\{b_k,\ k \geq 1\}$ for which $b_k \geq a_k$, we have

$$\begin{aligned}
&\sup_{-a_k \leq s \leq a_k} |W(b_k + s) - W(b_k)| \\
&\leq 3 \sup_{0 \leq t \leq b_k} \sup_{0 \leq s \leq a_k} |W(t+s) - W(t)|.
\end{aligned} \tag{4.2.91}$$

Therefore, using Hanson and Russo [56, display (3.10b)], the increment

$$I_{3k} = \left| W(R_X(k)) - W\left(R_X(k) + \left(\tilde{T}_{N_k} - R_X(k)\right)\right)\right|$$

is almost surely bounded, as $k \to \infty$, by

$$I_{3k} \ll \left(k^{\frac{\eta+\tau+1}{2}} \left(L\left(k^{\frac{\eta-\tau-1}{2}} + 1\right) + L_2 k^{\frac{\eta+\tau+1}{2}}\right)\right)^{\frac{1}{2}}, \tag{4.2.92}$$

where (4.2.22), (4.2.25) and (4.2.33) was applied. Observe that (4.2.8), i.e

$$\tau < \eta - 1, \quad 1 + \rho < \eta \quad \text{and} \quad \eta\rho < \tau,$$

remains true under the additional assumption

$$\tau + 1 < \eta - 1. \tag{4.2.93}$$

Fix $\tau > 0$, let $\eta > 2$ and choose ρ sufficiently small. Hence

$$\begin{aligned}
I_{3n} &\ll \left(k^{\frac{2\eta-1}{2}} \left(L\left(k^{\frac{\eta}{2}} + 1\right) + L_2 k^{\eta}\right)\right)^{\frac{1}{2}} \\
&= \left(n^{1-\frac{1}{2\eta}} \left(L\left(n^{\frac{1}{2}} + 1\right) + L_2 n\right)\right)^{\frac{1}{2}} \\
&= n^{\frac{1}{2}-\frac{1}{4\eta}} \left(L\left(n^{\frac{1}{2}} + 1\right) + L_2 n\right)^{\frac{1}{2}} \quad a.s. \quad (n \to \infty).
\end{aligned}$$

Next

$$I_{4k} = \left| W\left(\tilde{T}_{N_k}\right) - W\left(\tilde{T}_n\right) \right| \tag{4.2.94}$$

is almost surely bounded, as $k \to \infty$, by

$$I_{4k} \ll \left(k^{\frac{2\eta-1}{2}} \left(L\left(k^{\frac{1}{2}} + 1\right) + L_2 k^{\frac{2\eta-1}{2}} \right) \right)^{\frac{1}{2}}, \tag{4.2.95}$$

where (4.2.16) and (4.2.34) was applied. Therefore

$$\begin{aligned} I_{4n} &\ll \left(k^{\frac{2\eta-1}{2}} \left(L\left(k^{\frac{1}{2}} + 1\right) + L_2 k^{\frac{2\eta-1}{2}} \right) \right)^{\frac{1}{2}} \\ &= n^{\frac{1}{2}-\frac{1}{4\eta}} \left(L\left(n^{\frac{1}{2\eta}} + 1\right) + L_2 n^{1-\frac{1}{2\eta}} \right)^{\frac{1}{2}} \quad a.s. \quad (n \to \infty). \end{aligned} \tag{4.2.96}$$

Finally,

$$I_{5n} = \left| W_1\left(\tilde{T}_n\right) - W_1\left(T_n\right) \right|$$

is almost surely bounded, as $n \to \infty$, by

$$I_{5n} \ll \left(18 n^{\frac{1}{2}} \left(L\left(n^{\frac{1}{2}} + 1\right) + L_2 n^{\frac{1}{2}} \right) \right)^{\frac{1}{2}}, \tag{4.2.97}$$

where (4.2.35) was applied. Next, from (4.2.19)

$$\Gamma = \lim_{n\to\infty} \frac{1}{n} T_n = \frac{E\epsilon_1^2}{1-\phi^2} + \frac{2\phi E\epsilon_1^2}{(1+\phi)(1-\phi)^2}. \tag{4.2.98}$$

Towards this end, observe that via (4.2.17) and (4.2.18)

$$\begin{aligned} &T_n - n\Gamma \\ &= 2\|\epsilon_1\|_2^2 \sum_{\ell=2}^{n} (n-\ell+1) \frac{\phi^{\ell-1}}{1-\phi^2} - \frac{2\phi\|\epsilon_1\|_2^2 n}{(1+\phi)(1-\phi)^2} \\ &= \frac{2\|\epsilon_1\|_2^2}{(1-\phi)^2} \left(\sum_{\ell=2}^{n} (n-\ell+1)\phi^{\ell-1} - n \sum_{\ell=2}^{\infty} \phi^{\ell-1} \right) \\ &= \frac{2\|\epsilon_1\|_2^2}{(1-\phi)^2} \left(-\sum_{\ell=2}^{n} (\ell-1)\phi^{\ell-1} - n \sum_{\ell=n+1}^{\infty} \phi^{\ell-1} \right) \end{aligned} \tag{4.2.99}$$

Hence, for every $\lambda > 0$

$$|T_n - n\Gamma| = o\left(n^{\lambda}\right) \quad (n \to \infty). \tag{4.2.100}$$

Together with (4.2.10) we arrive at

$$\sum_{\ell=1}^{n} y_\ell - W(n\Gamma) = o\left((nL_2 n)^{1/2} \right) \quad a.s. \quad (n \to \infty). \tag{4.2.101}$$

$\square$

PROOF OF THEOREM 4.2.1. Put $N_k = (k+1)^\eta$ $(k = 1, 2, \dots)$ and let $N_k < n \leq N_{k+1}$. By (4.2.30)

$$I_{1k} = \sum_{\ell=1}^{n} \tilde{y}_\ell - \sum_{\ell=1}^{N_k} \tilde{y}_\ell \ll k^{\frac{\eta-\kappa}{2}} \quad a.s. \quad (k \to \infty). \tag{4.2.102}$$

Since

$$N_k < n \leq N_{k+1} \quad \text{implies} \quad k \sim n^{1/\eta} \quad (n \to \infty), \tag{4.2.103}$$

we have

$$I_{1n} \ll n^{\frac{1}{2} - \frac{\kappa}{2\eta}} \quad a.s. \quad (n \to \infty). \tag{4.2.104}$$

By (4.2.72)

$$I_{2k} = \sum_{\ell=1}^{N_k} \tilde{y}_\ell - W\left(R_X(k)\right) \ll k^{\frac{\eta-\kappa}{2}} \quad a.s., \tag{4.2.105}$$

which implies

$$I_{2n} \ll n^{\frac{1}{2} - \frac{\kappa}{2\eta}} \quad a.s. \quad (n \to \infty). \tag{4.2.106}$$

These two estimates for the increments I_{1n} and I_{2n} together with the remaining increments, which are estimated exactly along the lines in the proof of Theorem 4.2.1, we derive assertion (4.2.4). $\square$

CHAPTER 5

Time-Reversibility and Invariance

In this chapter we will consider linear processes with dependent errors. In the first section we will derive a new "backward" strong approximation results. In the second section we will discuss some applications of these approximations in change-point analysis. In particular, we will further develop the range of applicability of certain weighted change-point statistics with respect to structural breaks in financial time series.

5.1. Reversed Approximations

A stationary process $\{X_k,\ k \in \mathbb{Z}\}$ is said to be time-reversible, in the sense of Cheng [21], if for each $n \in \mathbb{N}$ and integers $k_1 < \cdots < k_n$ the vectors $(X_{k_1}, \dots, X_{k_n})$ and $(X_{-k_1}, \dots, X_{-k_n})$ have the same distribution. This condition is fulfilled for sequences of independent and identically distributed random variables. Cheng [21] established necessary and sufficient conditions for stationary Gaussian linear processes to be time-reversible, see also [18, p. 546] for related references. As a consequence of time-reversibility: large sample properties established for the (forward) partial sum process hold in the same way for the "backward" version $\{\sum_{k=1}^n X_{-k},\ n \ge 1\}$. However, Brockwell and Davis [18, p. 546] argued convincingly that in certain applications a Gaussian framework is too limited.

Reversed partial sums appear in change-point analysis in the context of cumulative sum (CUSUM) statistics within the at most one change-point model (AMOC). For instance, given a sample of length n, the so-called CUSUM statistic $n^{-1/2}(S(k) - \frac{k}{n}S(n))$ can be decomposed as a sum of forward and backward sums, i.e.

$$T_n(k) = \left(1 - \frac{k}{n}\right) \frac{\sum_{\ell=1}^k X_\ell}{\sqrt{n}} - \frac{k}{n} \frac{\sum_{\ell=k+1}^n X_\ell}{\sqrt{n}} \quad (1 \le k < n).$$

Observe that under strict stationarity, for each fixed n,

$$\left\{ \sum_{\ell=k+1}^{n} X_\ell,\ k = n-1, \dots, 1 \right\} \stackrel{\mathcal{D}}{=} \left\{ \sum_{\ell=1}^{k} X_{-\ell},\ k = 1, \dots, n-1 \right\}.$$

The decomposition above is essential when dealing with weighted versions of the CUSUM. These weighted versions with weight function g (say) arise from a (quasi) maximum-likelihood approach under the AMOC alternative, see Chaper 3 above. Depending on the particular choice of g, the limiting distribution of $\sup_{0<t<1} |T_n(nt)|/g(nt)$ can not be derived from a weak convergence of $T_n(nt)$ in the space $D[0,1]$. Although finite-dimensional limit distributions may exists, it turns out that, in certain cases, the usual tightness conditions fail to hold. This can be remedied by a suitable renormalization and proving a Darling-Erdős limit theorem. To pursue this approach weighted Brownian bridge type approximations in the sense of Csörgő and Horváth [23, Chapter 5.1] are essential, and hence, especially in the case of dependent observations, strong invariance principles for backward sums become necessary.

Horvath et al. [59] considered multivariate dependent observations. They studied the asymptotic behavior of weighted CUSUM statistics under the no-change null hypothesis and derived consistency results under the AMOC alternative. In particular, they considered m-dependent observation, that is, two observations are independent if their time lag exceeds m. Due to this kind of asymptotic independence, they were able to establish the approximation for the "backward" sum with a Gaussian process G (say), i.e. $\max_{(n/2)\leq k<n} |\sum_{\ell=k+1}^{n} X_\ell - G(n-k)|/\sqrt{n-k}$ is "small", as $n \to \infty$, (cf. ibid. display (4.10)). Moreover, a Brownian bridge type approximation in the sense of [24] was derived.

Davis et. al. [32] consider the AMOC model for parameter changes in autoregressive time series models. Since they worked in a strongly mixing framework and the reversed process is still strongly mixing, it suffices to assume conditions such that a strong (forward) invariance principle due to Kuelbs and Philipp [65] is valid. In Csörgő and Horváth [24, Theorem 4.1.3] similar arguments concerning the backward approximation are implicitly used (loc. cit. display (4.1.52)). They studied at most one location change in the model $X_\ell = \mu_\ell + e_\ell$, where the noise process $\{e_\ell,\ \ell \geq 1\}$ is a strongly mixing (one-sided) linear process $e_\ell = \sum_{j=0}^{\infty} a_j \varepsilon_{\ell-j}$ on an independent and identically distributed error sequence, requiring appropriate smoothness condition on the density function of ε_1.

Recently, this particular location AMOC model was studied again in Berkes et al. [6]. Therein strong forward and backward approximations are established without imposing mixing condition on the linear process, hence without any restrictions on the density function of the underlying error sequence. Their approach rests upon the so-called "Beveridge-Nelson" decomposition for partial sums of linear processes which reduces the partial sums to a sum of (independent) errors and negligible remainder terms. Thus they were able to derive the strong approximations via using "Komlós-Major-Tusnády" results (loc. cit. displays (3.2), (3.7)). Nevertheless, compared with the latter mixing approach, this approach is more restrictive with respect to possible weight sequences $\{a_\ell,\ \ell \geq 1\}$ whose admission is stated in terms of summability conditions. Further strong approximation results for sums of linear processes based on "KMT-construction" are derived in Wu [109, Proposition 2] under rather general summability conditions.

Ling [74] regarded the issue of "backward" limit theorems for dependent sequences as a phenomena on its own. He then established, within a general dependence framework, new "backward" versions of a strong law of large numbers and a strong invariance principle. The latter one holds for vector-valued martingale difference sequences satisfying a near-epoch dependence (NED) condition. This approach is based on forward and reversed martingale approximations and strong approximations due to Eberlein [37]. Although his method is quite different from our blocking and coupling approach below, we also need an additional NED condition. Moreover, in order to pursue the proof of the main result, this additional NED condition enables us to prove a backward maximal inequality .

Within a general dependence situation, beyond the scope of [21] and even beyond mixing conditions, the NED condition seems to be indispensable to assure an approximately kind of time-reversibility, which in turn guarantees strong "backward" approximations. In conclusion we believe that our result is the only contribution towards Ling's new direction so far.

The near-epoch dependence concept was originally introduced in a "functions of mixing processes" context, see McLeish [79], to prove limit theorems for dependent "heterogeneous" processes. For instance, [79, Section 3] proved strong laws of large numbers and Billingsley [11, Section 19] established a functional central limit theorem (FCLT). See also Davidson [29] for a comprehensive exposition. In a time series context, when dealing with an underlying independent white noise process, the NED condition is used implicitly or explicitly to prove limit theorems. For instance, Berkes et al. [8] proved a FCLT for augmented GARCH models based on Billingsley's FCLT. In addition, Ling and Li [75, Theorem 3.2] obtained a generalized version of Billingsley's result. Recently, Davidson [30] derived conditions for a broad class of non-linear time series models to obey the NED property and conditions for a FCLT.

For further use we formalize, in the manner of [74], what is meant for a sequence to be near-epoch dependent on an underlying independent sequence.

Let $(\Omega, \mathcal{F}, P)$ denote a probability space on which there is a sequence of *independent* random variables $\{\eta_k,\ -\infty < k < \infty\}$. For each integer k and $m \geq 0$, let $\mathcal{F}_{km}$ be the σ-algebra generated by the family $\{\eta_{k-m}, \ldots, \eta_k\}$

and $\mathcal{F}_k = \sigma(\ldots, \eta_{k-1}, \eta_k)$. Suppose $\{X_k,\ -\infty < k < \infty\}$ is a sequence of $\mathcal{F}_k$-measurable random variables.

DEFINITION. *The sequence* $\{(X_k, \mathcal{F}_k),\ -\infty < k < \infty\}$ *is said to be* $\mathcal{L}^r$*-NED in terms of* $\{\eta_k,\ -\infty < k < \infty\}$, *if* $\sup_{-\infty<k<\infty} \|X_k\|_r < \infty$ $(r \geq 1)$, *and, for sequences of finite nonnegative constants* c_k *and* ψ_m, *where* $\psi_m \downarrow 0 \quad (m \to \infty)$, *we have for all integers* k *and* $m \geq 0$

$$\|X_k - E[X_k|\mathcal{F}_{km}]\|_r \leq c_k \psi_m.$$

Here we consider the following model: Let $\{\eta_k,\ k \in \mathbb{Z}\}$ be a sequence of independent random variables with mean zero and $\mathcal{F}_{k-1}$ denotes the σ-algebra generated by the family $\{\ldots, \eta_{k-2}, \eta_{k-1}\}$. We consider dependent errors

$$\epsilon_k = \sigma_k \eta_k, \quad k \in \mathbb{Z}, \tag{5.1.1}$$

where σ_k is measurable with respect to $\mathcal{F}_{k-1}$ for every $k \in \mathbb{Z}$.

We assume that $\{\varepsilon_k,\ k \in \mathbb{Z}\}$ is $\mathcal{L}^r$-NED in terms of $\{\eta_k,\ k \in \mathbb{Z}\}$ for some $r > 2$, i.e.

$$\|\epsilon_k - E[\epsilon_k|\eta_{k-m}, \ldots, \eta_k]\|_r \leq c_k \psi_k. \tag{5.1.2}$$

Under these considerations the sequence $\{y_k,\ k \geq 1\}$ is defined as solution of the autoregressive scheme

$$y_k = \phi y_{k-1} + \epsilon_k, \quad k = 1, 2, \ldots, \tag{5.1.3}$$

where $-1 < \phi < 1$ is a fixed parameter.

Before we turn to the reversed invariance principle, we state a "backward" maximal inequality for "truncated" versions of the solution of the autoregressive scheme.

THEOREM 5.1.1. *Assume that* (5.1.1), (5.1.2) *and* (5.1.3) *hold. Let* $\tilde{y}_{-k} = \sum_{\ell=0}^{k^\rho} \phi^\ell \epsilon_{-k-\ell}$ $(k = 1, 2, \ldots)$ *for some* $\rho > 0$. *If* $\sum_{k=0}^\infty \psi_k < \infty$, *then for each integers* m, n, *satisfying* $1 \leq m \leq n$, *we have*

$$\left\| \max_{m \leq j \leq n} \left| \sum_{i=m}^{j} \tilde{y}_{-i} \right| \right\|_r \leq C_1 \sum_{k=0}^{\infty} \psi_k \left(\sum_{i=1}^{n-m+1} \left(\sum_{\ell=0}^{k \wedge (i+m-1)^\rho} \phi^\ell c_{k-\ell} \right)^2 \right)^{1/2}, \tag{5.1.4}$$

where C_1 *is a nonnegative constant.*

Let us state the set of assumptions for the reversed approximation result.

ASSUMPTION A. *Suppose that* (5.1.1) *holds and* $\{\eta_k,\ k \in \mathbb{Z}\}$ *is a sequence of independent, identically distributed random variables with mean zero;* (5.1.2) *is satisfied for some* $r > 2$ *with* $\sup_{k \in \mathbb{Z}} c_k < \infty$ *and*

$\sum_{k=0}^{\infty} \psi_k < \infty$. *Further, assume that* $\{\epsilon_k,\ k \in \mathbb{Z}\}$ *is a strictly stationary and strongly mixing sequence, such that* $\alpha(n) = O(\theta^n) \quad (n \to \infty)$ *for some* $0 < \theta < 1$; *and* $\{y_k,\ k \geq 1\}$ *is defined as strictly stationary solution of* (5.1.3).

Let us state the reversed approximation results.

THEOREM 5.1.2. *If Assumption A holds, then we can redefine the sequence* $\{y_{-k},\ k \geq 1\}$ *without changing its distribution on a new probability space together with a Wiener process* $\{W(t),\ t \geq 0\}$ *such that*

$$\left| \sum_{\ell=1}^{n} y_{-\ell} - \Gamma^{1/2} W(n) \right| \ll n^{\frac{1}{2}-\kappa_0} \quad a.s. \quad (n \to \infty), \tag{5.1.5}$$

where $\Gamma = (1-\phi^2)^{-1} E\epsilon_1^2 + (2\phi E\epsilon_1^2)/((1+\phi)(1-\phi)^2)$ *and* κ_0 *is a constant depending on* r *only.*

REMARK. *If* $2 < r \leq 4$, *one can put* $\eta = 3$ *in the proof and derive that for each* $0 < \kappa < 1 - 2/r$ *the rate of approximation in* (5.1.5) *is* $O(n^{\frac{1}{2}-\frac{\kappa}{6}})$ *almost surely as* $n \to \infty$. *If* $r > 4$, *one can put* $\eta = 3$ *in the proof and derive that the order obeys* $O(n^{\frac{5}{12}} \sqrt{\log n})$ *almost surely as* $n \to \infty$.

The proof of the theorem will use two essential features from the method of Aue et al. [1]. In order to approximate $\sum_{\ell=1}^{n} y_{-\ell}$, it suffices to consider truncated versions $\tilde{y}_{-\ell}$ defined by

$$\tilde{y}_{-k} = \sum_{\ell=0}^{k^\rho} \phi^\ell \epsilon_{-k-\ell} \quad \text{for some fixed } \rho,\ (0 < \rho < 1). \tag{5.1.6}$$

The resulting sequence $\ldots, \tilde{y}_{-2}, \tilde{y}_{-1}$ is then merged into a sequence of consecutive "blocks" $\ldots, Y_{-2}, X_{-2}, Y_{-1}, X_{-1}$. Following the method in Aue et al. [1], we introduce "big" blocks

$$X_{-k} = \sum_{\ell=k^\eta}^{(k+1)^\eta - (k+2)^\tau} \tilde{y}_{-\ell} \quad (k = 1, 2, \ldots) \tag{5.1.7}$$

and "small" blocks

$$Y_{-k} = \sum_{\ell=(k+1)^\eta-(k+2)^\tau+1}^{(k+1)^\eta - 1} \tilde{y}_{-\ell} \quad (k = 1, 2, \ldots), \tag{5.1.8}$$

where η, ρ and τ are nonnegative constants so that

$$\tau < \eta - 1, \quad 1 + \rho < \eta \quad \text{and} \quad \eta\rho < \tau. \tag{5.1.9}$$

The proof is based on a series of lemmas.

LEMMA 5.1.1. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (5.1.9). *Let the sequence* $\{X_{-k},\ k \geq 1\}$ *be defined as in* (5.1.7). *Then there exists some* $k_0 > 0$ *such that for all* $k \geq k_0$

$$\alpha\left(\sigma\left(X_{-k}\right), \sigma\left(X_{-(k-1)}, \ldots, X_{-1}\right),\right) \leq \alpha_\epsilon\left(k^\tau\right). \tag{5.1.10}$$

PROOF. Observe that X_{-k} is a function g_k (say) of a finite subset of $\{\epsilon_{-k},\ k \geq 1\}$. By (5.1.7), in terms of indices, since $k^\eta < (k+1)^\eta - (k+2)^\tau$, the last index is $-k^\eta$. We claim that $-((k+1)^\eta - (k+2)^\tau - ((k+1)^\eta - (k+2)^\tau)^\rho)$ is the first one. It suffices to establish the inequality

$$k^\eta - k^{\eta\rho} < (k+1)^\eta - (k+2)^\tau - ((k+1)^\eta - (k+2)^\tau)^\rho,$$

which flows from the proof of (4.2.13). Therefore, for some $k_0 > 0$, the following representation holds for all $k \geq k_0$

$$X_{-k} = g_k\left(\epsilon_{-((k+1)^\eta-(k+2)^\tau-((k+1)^\eta-(k+2)^\tau)^\rho)}, \ldots, \epsilon_{-k^\eta}\right).$$

This implies the desired time lag difference between X_{-k} and $X_{-(k-1)}$, i.e. $(k+1)^\tau + k^{\eta\rho} \sim k^\tau$. □

LEMMA 5.1.2. *Let* η, ρ *and* τ *be nonnegative constants satisfying* (5.1.9). *Let the sequence* $\{Y_{-k},\ k \geq 1\}$ *be defined as in* (5.1.8). *Then there exists some* $k_0 > 0$ *such that for all* $k \geq k_0$

$$\alpha\left(\sigma\left(Y_{-k}\right), \sigma\left(Y_{-(k-1)}, \ldots, Y_{-1}\right),\right) \leq \alpha_\epsilon\left(k^{\eta-1}\right). \tag{5.1.11}$$

PROOF. Observe that Y_{-k} is a function g_k (say) of a finite subset of $\{\epsilon_{-k},\ k \geq 1\}$. By (5.1.8), in terms of indices, the last one is $-((k+1)^\eta - 1)$. We claim that $-((k+1)^\eta - 1 - ((k+1)^\eta - 1)^\rho)$ is the first one. It suffices to establish the inequality

$$\begin{aligned}&(k+1)^\eta - (k+2)^\tau + 1 - ((k+1)^\eta - (k+2)^\tau + 1)^\rho \\ &< (k+1)^\eta - 1 - ((k+1)^\eta - 1)^\rho,\end{aligned}$$

which flows from the proof of (4.2.15).Therefore, for some $k_0 > 0$, the following representation holds for all $k \geq k_0$

$$Y_{-k} = g_k\left(\epsilon_{-((k+1)^\eta-1-((k+1)^\eta-1)^\rho)}, \ldots, \epsilon_{-((k+1)^\eta-1)}\right).$$

This implies that the time lag difference between Y_{-k} and $Y_{-(k-1)}$ is $(k+1)^\eta - 1 - k^\eta - 1 - (k^\eta - 1)^\rho$ which is asymptotically equivalent to $\eta k^{\eta-1}$ $(k \to \infty)$. □

PROOF OF THEOREM 5.1.1. Let $\{\mathcal{G}_i,\ i \geq 1\}$ be the sequence of increasing σ-algebras, i.e. $\mathcal{G}_i \subset \mathcal{G}_{i+1}$, defined by

$$\mathcal{G}_i = \sigma\left(\eta_{-i}, \eta_{-i+1}, \ldots\right) \quad (i = 1, 2, \ldots). \tag{5.1.12}$$

We will represent the conditional expectation $E[X|\mathcal{G}_i]$ by $E_i[X]$. Observe that, for each fixed $i \geq 1$,

$$\{(E_{i+n}[\tilde{y}_{-i}],\, \mathcal{G}_{i+n}),\, n \geq 1\} \tag{5.1.13}$$

is an uniformly integrable martingale. Hence the martingale convergence theorem implies

$$E_{i+n}[\tilde{y}_{-i}] \to E_\infty[\tilde{y}_{-i}] \quad a.s. \quad (n \to \infty). \tag{5.1.14}$$

Moreover, using Minkowski's inequality and the NED property, we have

$$\begin{aligned} &\|E_{i+n}[\tilde{y}_{-i}] - \tilde{y}_{-i}\|_r \\ &\leq \sum_{\ell=0}^{i^\rho \wedge n} \phi^\ell \|E_{i+\ell+n-\ell}[\epsilon_{-i-\ell}] - \epsilon_{-i-\ell}\|_r \\ &\leq \sum_{\ell=0}^{i^\rho \wedge n} \phi^\ell c_{-i-\ell} \psi_{n-\ell}, \end{aligned} \tag{5.1.15}$$

which implies

$$\|E_{i+n}[\tilde{y}_{-i}] - \tilde{y}_{-i}\|_r \to 0 \quad (n \to \infty). \tag{5.1.16}$$

Therefore, using Elstrodt [44, Korollar VI.2.7], which J. Elstrodt attributes to Weyl [107], there exists some subsequence $\{n_k,\, k \geq 1\}$ so that

$$E_{i+n_k}[\tilde{y}_{-i}] \to \tilde{y}_{-i} \quad a.s. \quad (k \to \infty). \tag{5.1.17}$$

Combining (5.1.14) and (5.1.17) we arrive at

$$E_{i+n}[\tilde{y}_{-i}] \to \tilde{y}_{-i} \quad a.s. \quad (n \to \infty). \tag{5.1.18}$$

Observe that, for each fixed $i \geq 1$,

$$\{(E_{i-m-1}[\tilde{y}_{-i}],\, \mathcal{G}_{i-m}),\, m \geq 1\} \tag{5.1.19}$$

is a reverse martingale. Nevertheless, if $m \geq i$, then

$$E_{i-m-1}[\tilde{y}_{-i}] = 0 \quad a.s. \tag{5.1.20}$$

Since

$$\tilde{y}_{-i} = \sum_{k=-m}^{n} E_{i+k}[\tilde{y}_{-i}] - E_{i+k-1}[\tilde{y}_{-i}] = E_{i+n}[\tilde{y}_{-i}] - E_{i-m-1}[\tilde{y}_{-i}], \tag{5.1.21}$$

we have from (5.1.18) and (5.1.20) the series representation

$$\tilde{y}_{-i} = \sum_{k=-\infty}^{\infty} E_{i+k}[\tilde{y}_{-i}] - E_{i+k-1}[\tilde{y}_{-i}] \quad a.s. \tag{5.1.22}$$

Similar to the proof of Theorem A.1, an application of Minkowski's inequality together with (5.1.22) yields

$$\left\| \max_{m\le j\le n} \left| \sum_{i=m}^{j} \tilde{y}_{-i} \right| \right\|_r \le \sum_{k=-\infty}^{\infty} \left\| \max_{m\le j\le n} \left| \sum_{i=m}^{j} E_{i+k}\left[\tilde{y}_{-i}\right] - E_{i+k-1}\left[\tilde{y}_{-i}\right] \right| \right\|_r . \tag{5.1.23}$$

Adjusting indices, for each fixed k we have

$$\begin{aligned} &\left\| \max_{m\le j\le n} \left| \sum_{i=m}^{j} E_{i+k}\left[\tilde{y}_{-i}\right] - E_{i+k-1}\left[\tilde{y}_{-i}\right] \right| \right\|_r \\ &= \left\| \max_{1\le j\le n-m+1} \left| \sum_{i=1}^{j} E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right] - E_{i+m-1+k-1}\left[\tilde{y}_{-(i+m-1)}\right] \right| \right\|_r . \end{aligned} \tag{5.1.24}$$

Observe that for each fixed k, m

$$\left\{ \left(E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right] - E_{i+m-1+k-1}\left[\tilde{y}_{-(i+m-1)}\right], \mathcal{G}_{i+m-1+k} \right),\ i \ge 1 \right\} \tag{5.1.25}$$

is a martingale difference sequence. Similar to the proof of Theorem A.1, a combination of the inequalities of Doob, Burkeholder and Minkowski yields

$$\begin{aligned} &\left\| \max_{1\le j\le n-m+1} \left| \sum_{i=1}^{j} E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right] - E_{i+m-1+k-1}\left[\tilde{y}_{-(i+m-1)}\right] \right| \right\|_r \\ &\le C_0 \left(\sum_{i=1}^{n-m+1} \left\| E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right] - E_{i+m-1+k-1}\left[\tilde{y}_{-(i+m-1)}\right] \right\|_r^2 \right)^{1/2}, \end{aligned} \tag{5.1.26}$$

where C_0 is a nonnegative constant. Obviously, for $k < 0$

$$E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right] = 0 \quad a.s. \tag{5.1.27}$$

and for $k > 0$ via NED property

$$\begin{aligned} &\left\| \tilde{y}_{-(i+m-1)} - E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right] \right\|_r \\ \le &\sum_{\ell=0}^{k\wedge(i+m-1)^\rho} \phi^\ell \left\| \epsilon_{-(i+m-1)} - E_{i+m-1+k}\left[\epsilon_{-(i+m-1)}\right] \right\|_r \\ \le &\sum_{\ell=0}^{k\wedge(i+m-1)^\rho} \phi^\ell c_{k-\ell}\psi_k . \end{aligned} \tag{5.1.28}$$

Since

$$\begin{aligned}
&\left\|E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right]-E_{i+m-1+k-1}\left[\tilde{y}_{-(i+m-1)}\right]\right\|_r \\
&\leq\left\|\tilde{y}_{-(i+m-1)}-E_{i+m-1+k}\left[\tilde{y}_{-(i+m-1)}\right]\right\|_r \\
&\quad+\left\|\tilde{y}_{-(i+m-1)}-E_{i+m-1+k-1}\left[\tilde{y}_{-(i+m-1)}\right]\right\|_r,
\end{aligned}$$

together with (5.1.23),(5.1.24), (5.1.26), (5.1.27) and (5.1.28) we arrive at (5.1.4). □

PROOF OF THEOREM 5.1.2. Under Assumption A, for each integer m and n satisfying $1 \leq m \leq n$, Theorem 5.1.1 implies

$$\left\|\max_{m\leq j\leq n}\left|\sum_{i=m}^{j}\tilde{y}_{-i}\right|\right\|_r \leq C_1\,(n-m)^{1/2}\,, \tag{5.1.29}$$

where C_1 is an absolute constant. Since the underlying error sequences are strictly stationary, the maximal inequality (5.1.29) together with the block estimates (5.1.10) and (5.1.11) are sufficient to derive the assertions via following exactly the proof of Theorem 4.2.2 with $y_k, \tilde{y}_k, X_k, Y_k$ replaced by $y_{-k}, \tilde{y}_{-k}, X_{-k}$ and Y_{-k}, respectively. □

.

5.2. Applications in Change-Point Analysis

In certain financial time series, for instance foreign exchange rates, one usually observes Y_t in discrete time points, for example daily closing prices or daily exchange rates. Therefore it seems reasonable to choose a discrete time series model $\{Y_k,\ k \geq 1\}$. Taking into account that the autocorrelation coefficient is close to one at lag one, cf. e.g. [45, Chapter 4.2.7], it is common practice to model returns of the observed time series. Typically, logarithms of the return series are considered, i.e. $y_k = \log Y_k - \log Y_{k-1}$. In order to capture so-called stylized features like, for instance, heavy tails and volatility clustering of financial return series, returns are modeled in form of generalized autoregressive conditional heteroscedasticity (GARCH). Notice that heavy tails, in the sense of Fan and Yao [45, p. 169], means heavier tails than the tails of the normal distribution. They also remarked that the finite second moment assumption for daily returns is widely accepted. Nevertheless, Basrak et al. [3, p. 96] pointed out that for instance log-returns of foreign exchange rates can have infinite fifth, fourth or even third moments.

Following the common practice to fit returns according to the random walk hypothesis, Francq and Zakoïan [47, Section 5.2] considered the autoregressive model $y_k = \phi y_{k-1} + \epsilon_k$, where the errors belong to the GARCH class. This particular AR-GARCH model received attention in establishing asymptotics under the unit-root hypothesis, i.e. $\phi = 1$, for the Dickey-Fuller test, see Francq and Zakoïan [47] and Berkes et al. [8, Example 3.3] and the references therein.

Francq and Zakoïan [46] introduced a quasi-maximum likelihood estimator for the parameter vector of autoregressive moving average time series with GARCH errors, i.e. ARMA-GARCH models, and established its almost sure consistency. Based on their results, Lee and Song [68] used a cumulative sum (CUSUM) type statistic to test structural stability of the ARMA-GARCH model parameters. The limit distribution under the no-change null hypothesis was derived via a functional central limit theorem for martingales. With a view towards weighted CUSUM statistics, Aue et al. [1] derived a strong approximation result for partial sums of (squared) augmented GARCH observations. Motivated by the results in Csörgő and Horváth [24], they discussed the following weighted CUSUM version

$$\frac{1}{\sqrt{n}} \max_{1 \le k \le n} \left(\frac{n}{k}\right)^{\alpha} \left| \sum_{i=1}^{k} y_i^2 - \frac{k}{n} \sum_{i=1}^{n} y_i^2 \right| .$$

Suitably normalized, under the no-change null-hypothesis, the limiting distribution was shown to be $\sup_{0<t<1} |B_t|/t^{\alpha}$, for each $0 \le \alpha < 1/2$, where $\{B(t),\ 0 \le t \le 1\}$ denotes a Brownian bridge process, cf. (ibid., Example 3.3). In contrast to the (unweighted) statistic in Lee and Song [68], this weighted version has better power for detecting changes that occur rather early in the observed sample.

As a first example we discuss the case $\alpha = 1/2$, which is excluded above, and consider the analogous version to detect rather late change-points in AR(1)-GARCH(1,1) models under a low moment condition. Let us restate the set of assumptions.

Assumption A. *Suppose that* (5.1.1) *holds and* $\{\eta_k,\ k \in \mathbb{Z}\}$ *is a sequence of independent, identically distributed random variables with mean zero;* (5.1.2) *is satisfied for some* $r > 2$ *with* $\sup_{k\in\mathbb{Z}} c_k < \infty$ *and* $\sum_{k=0}^{\infty} \psi_k < \infty$. *Further assume that* $\{\epsilon_k,\ k \in \mathbb{Z}\}$ *is a strictly stationary and strongly mixing sequence, such that* $\alpha(n) = O(\theta^n)$ $(n \to \infty)$ *for some* $0 < \theta < 1$*; and* $\{y_k,\ k \ge 1\}$ *is defined as strictly stationary solution of* (5.1.3).

Among the recent contributions, Francq and Zakoïan [47] established geometric ergodicity for a class of GARCH$(1,1)$ models under a less restrictive moment condition. These models include the augmented GARCH class, see Chapter 1 above. Under their setup the geometric decay of the strong mixing coefficient is clearly satisfied. Hansen [55] established the $\mathcal{L}^r$-NED

$(r > 1)$ property for GARCH$(1,1)$ models. Actually, in [55] a kind of geometrically $\mathcal{L}^r$-NED property, in the sense of [30], was established which is sufficient for our efforts. Ling [74, Remark 3.1] verified also NED properties in AR-GARCH models.

THEOREM 5.2.1. *Suppose Assumption A holds. Let*

$$A(x) = (2\log x)^{1/2},$$

$$D^*(x) = 2\log x + \frac{1}{2}\log\log x - \frac{1}{2}\log(4\pi) \quad \textit{and}$$

$$T_n(k) = \left(1 - \frac{k}{n}\right)^{-1/2}\left(\sum_{i=1}^{k} y_i - \frac{k}{n}\sum_{i=1}^{n} y_i\right) \qquad (k = 1, \ldots, n-1).$$

Then we have, as $n \to \infty$,

$$A(\log n)\,\Gamma^{-1/2} \max_{1\le k\le n-1} \frac{|T_n(k)|}{\sqrt{n}} - D^*(\log n) \xrightarrow{\mathcal{D}} E, \tag{5.2.1}$$

where E *is a Gumbel distributed random variable and* $\Gamma = (1-\phi^2)^{-1}E\epsilon_1^2 + (2\phi E\epsilon_1^2)/((1+\phi)(1-\phi)^2)$.

The next example concerns the detection of a structural break without any prior information whether it occurs early or late within the observed sample. Similar to the proof of the last statement, the backward approximations, i.e. Lemma 5.2.4 - Lemma 5.2.6, are crucial. In addition, a kind of asymptotic independence between the early and late observations is established in order to derive the following extreme-value asymptotic.

THEOREM 5.2.2. *Suppose Assumption A holds. Let*

$$A(x) = (2\log x)^{1/2},$$

$$D(x) = 2\log x + \frac{1}{2}\log\log x - \frac{1}{2}\log\pi \quad \textit{and}$$

$$G_n(k) = \left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{-1/2}\left(\sum_{i=1}^{k} y_i - \frac{k}{n}\sum_{i=1}^{n} y_i\right) \qquad (k = 1, \ldots, n-1).$$

Then we have, as $n \to \infty$,

$$A(\log n)\,\Gamma^{-1/2} \max_{1\le k\le n-1} \frac{|G_n(k)|}{\sqrt{n}} - D(\log n) \xrightarrow{\mathcal{D}} E \vee E', \tag{5.2.2}$$

where E *and* E' *are independent and identically Gumbel distributed random variables and* $\Gamma = (1-\phi^2)^{-1}E\epsilon_1^2 + (2\phi E\epsilon_1^2)/((1+\phi)(1-\phi)^2)$.

REMARK. *For an application of both results in practice an estimator for the unknown parameter* Γ *is needed. The invariance-principle-based approach of Steinebach [99] yields estimators which are suitable with respect to the rate of convergence.*

The proof of the main results is based on a series of lemmas.

LEMMA 5.2.1. *Let* $\{u_k,\ k \geq 1\}$ *be a sequence for which* $1 \leq u_n \leq n-1$. *Then for each* $\rho > 0$, *as* $n \to \infty$,

$$\max_{n-u_n \leq k \leq n-1} \left| \sum_{i=k+1}^{n} y_i - \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell} \right| = O(1) \quad a.s. \tag{5.2.3}$$

PROOF. The following estimate yields immediately the assertion. By Minkowski's inequality

$$\begin{aligned} &\left\| \max_{n-u_n \leq k \leq n-1} \left| \sum_{i=k+1}^{n} y_i - \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell} \right| \right\|_r \\ &\leq \|\epsilon_1\|_r \sum_{i=n-u_n}^{n} \sum_{\ell=(n-i+1)^\rho+1}^{\infty} |\phi|^\ell < \infty, \end{aligned} \tag{5.2.4}$$

where the second inequality flows from

$$\sum_{i=n-u_n}^{n} |\phi|^{(n-i+1)^\rho} \leq \sum_{i=1}^{n} |\phi|^{i^\rho} < \infty. \tag{5.2.5}$$

□

LEMMA 5.2.2. *Let* $\{u_k,\ k \geq 1\}$ *be a sequence for which* $1 \leq u_n \leq n-1$. *Then for each* $0 < \rho < 1$ *and* $n = 1, 2, \ldots$ *we have*

$$\begin{aligned} &\left\{ \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell},\ k = n-1, \ldots, n-u_n \right\} \\ &\stackrel{\mathcal{D}}{=} \left\{ \sum_{i=1}^{k} \tilde{y}_{-i},\ k = 1, \ldots, u_n \right\}. \end{aligned} \tag{5.2.6}$$

PROOF. Since the errors form a stationary sequence, we have

$$\begin{aligned} &\left\{ \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell},\ k = n-1, \ldots, n-u_n \right\} \\ &\stackrel{\mathcal{D}}{=} \left\{ \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell-n-1},\ k = n-1, \ldots, n-u_n \right\}. \end{aligned} \tag{5.2.7}$$

Hence, rephrasing the right-hand side, we have

$$\begin{aligned}
&\left\{ \sum_{i=1}^{k} \sum_{\ell=0}^{i^\rho} \phi^\ell \epsilon_{-1-\ell},\ k = 1, \ldots, u_n \right\} \\
&= \left\{ \sum_{i=1}^{k} \tilde{y}_{-i},\ k = 1, \ldots, u_n \right\}, \qquad (5.2.8)
\end{aligned}$$

where $\{\tilde{y}_{-k},\ k \geq 1\}$ are the truncated versions as defined in (5.1.6). □

LEMMA 5.2.3. *For each* $0 < \rho < 1$ *let the sequence* $\{\tilde{y}_{-k},\ k \geq 1\}$ *be defined as in* (5.1.6). *Then*

$$P\left[\max_{1 \leq j < \infty} \left| \sum_{k=1}^{j} y_{-k} - \sum_{k=1}^{j} \tilde{y}_{-k} \right| < \infty \right] = 1. \qquad (5.2.9)$$

PROOF. Follow the same pattern as in the proof of (4.2.10). □

LEMMA 5.2.4. *Let* $\{u_k,\ k \geq 1\}$ *and* $\{v_k,\ k \geq 1\}$ *be sequences for which* $1 \leq n - u_n \leq v_n$ *and* $n - v_n \uparrow \infty$ $(n \to \infty)$. *Then, for each* $0 < \rho < 1$, *we can enrich the probability space with a uniformly distributed random variable* U, *such that, as* $n \to \infty$,

$$\max_{n-u_n \leq k \leq v_n} \left(\frac{\log\log(n-k)}{n-k} \right)^{1/2} \left| \sum_{i=k+1}^{n} y_i - \Gamma^{1/2} W_{2n}(n-k) \right| = o_P(1), \qquad (5.2.10)$$

where, for each n, *the Wiener process* $\{W_{2n}(t),\ 0 \leq t \leq u_n\}$ *is a measurable function of* $\left\{\epsilon_{n-u_n+1-u_n^\rho}, \ldots, \epsilon_n, U\right\}$.

PROOF. We apply Theorem 5.1.2, that is, we can redefine the sequence $\{y_\ell,\ l \geq 1\}$ on a possibly different probability space together with a Wiener process, such that

$$\left| \sum_{\ell=1}^{n} y_{-\ell} - \Gamma^{1/2} W(n) \right| \ll n^{\frac{1}{2}-\kappa_0} \quad a.s. \quad (n \to \infty). \qquad (5.2.11)$$

Lemma 5.2.3 yields immediately

$$\left| \sum_{\ell=1}^{n} \tilde{y}_{-\ell} - \Gamma^{1/2} W(n) \right| \ll n^{\frac{1}{2}-\kappa_0} \quad a.s. \quad (n \to \infty). \qquad (5.2.12)$$

This strong approximation result constitutes a law on the Borel sets of $\mathbb{R}^{u_n} \times C[0, u_n]$, i.e.

$$\mathcal{L}\left(\left\{ \sum_{\ell=1}^{k} \tilde{y}_{-\ell},\ k = 1, \ldots, u_n \right\},\ \{W(t),\ t \in [0, u_n]\} \right). \qquad (5.2.13)$$

In light of Lemma 5.2.2 we can apply Billingsley [11, Lemma 21.1], see also Chapter 2.2 above, that is, we enrich the probability space with a uniformly distributed random variable U and we can construct a sequence of approximating Wiener processes W_{2n} (say) on the initial probability space, such that, for each $n \in \mathbb{N}$,

$$\left\{ \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell} - \Gamma^{1/2} W_{2n}(n-k),\ k = n-1, \ldots, n-u_n \right\}$$
$$\stackrel{\mathcal{D}}{=} \left\{ \sum_{\ell=1}^{n-k} \tilde{y}_{-\ell} - \Gamma^{1/2} W(n-k),\ k = n-1, \ldots, n-u_n \right\}. \tag{5.2.14}$$

This implies

$$\max_{n-u_n \le k \le v_n} \left(\frac{\log\log(n-k)}{n-k} \right)^{1/2} \left| \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell} - \Gamma^{1/2} W_{2n}(n-k) \right|$$
$$= o_P(1) \quad (n \to \infty), \tag{5.2.15}$$

The application of [11, Lemma 21.1] rests upon the observation that the approximating Wiener process $\{W(t),\, t \in [0, u_n]\}$ is a random element in the polish space $C[0, u_n]$ and upon the fact that $C[0, u_n]$ is Borel isomorphic to $[0, 1]$, cf. e.g. Dudley [35, Theorem 13.1.1] and [11, p. 212]. As a consequence of the coupling construction of [11, Lemma 21.1], which involves the Borel isomorphism and regular conditional distributions, each random element W_{2n} is a measurable function of U and

$$\left\{ \sum_{i=k+1}^{n} \sum_{\ell=0}^{(n-i+1)^\rho} \phi^\ell \epsilon_{i-\ell},\ k = n-1, \ldots, n-u_n \right\} \tag{5.2.16}$$

Since the latter vector is measurable with respect to $\left\{\epsilon_{n-u_n+1-u_n^\rho}, \ldots, \epsilon_n\right\}$, the assertion follows via an application of Lemma 5.2.1. □

LEMMA 5.2.5. *Let* $\{u_k,\ k \ge 1\}$ *be a sequence for which* $1 \le u_n \le n-1$ *and* $u_n \uparrow \infty$ $(n \to \infty)$. *Then, for each* $0 < \rho < 1$, *we can enrich the probability space with a uniformly distributed random variable* U, *such that, as* $n \to \infty$,

$$\max_{n-u_n \le k \le n-1} \frac{\left| \sum_{i=k+1}^{n} y_i - \Gamma^{1/2} W_{2n}(n-k) \right|}{\sqrt{n-k}} = o_P\left((\log\log n)^{1/2} \right), \tag{5.2.17}$$

where, for each n, *the Wiener process* $\{W_{2n}(t),\ 0 \le t \le u_n\}$ *is a measurable function of* $\left\{\epsilon_{n-u_n+1-u_n^\rho}, \ldots, \epsilon_n, U\right\}$.

PROOF. Choose a sequence $\{v_k,\ k \geq 1\}$ such that $n - u_n \leq v_n \leq n-1$ and

$$n - v_n \uparrow \infty \quad \text{and} \quad (\log\log n)^{-1/2}\sqrt{n - v_n} \downarrow 0 \quad (n \to \infty). \tag{5.2.18}$$

By Lemma 5.2.4, we have

$$\max_{n-u_n \leq k \leq v_n} \frac{\left|\sum_{i=k+1}^n y_i - \Gamma^{1/2} W_{2n}(n-k)\right|}{\sqrt{n-k}} = o_P\left((\log\log n)^{1/2}\right) \tag{5.2.19}$$

as $n \to \infty$. Using (4.2.10) and (4.2.9) together with the maximal inequality in Theorem A.1, we derive for each $\epsilon > 0$

$$\lim_{n\to\infty} P\left[(\log\log n)^{-1/2} \max_{v_n \leq k \leq n-1} \frac{\left|\sum_{i=k+1}^n y_i\right|}{\sqrt{n-k}} > \epsilon\right] = 0. \tag{5.2.20}$$

□

LEMMA 5.2.6. *For each $0 < \rho < 1$ we can enrich the probability space with a uniformly distributed random variable U, such that, as $n \to \infty$,*

$$\begin{aligned}&\Big| \max_{n-\frac{n}{\log n} \leq k \leq n-1} \frac{\left|\sum_{i=k+1}^n y_i\right|}{\sqrt{n-k}} - \max_{n-\frac{n}{\log n} \leq k \leq n-1} \frac{\left|\Gamma^{1/2} W_{2n}(n-k)\right|}{\sqrt{n-k}} \Big| \\ &= o_P\left((\log\log n)^{-1/2}\right),\end{aligned} \tag{5.2.21}$$

where, for each n, the Wiener process $\{W_{2n}(t),\ 0 \leq t \leq n/\log n\}$ is a measurable function of $\{\epsilon_{n-(n/\log n)+1-(n/\log n)^\rho}, \ldots, \epsilon_n, U\}$.

PROOF. Follow the pattern in the proof of Lemma 2.1.8 via using Lemma 5.2.4 and Lemma 5.2.5. □

With similar arguments, via using the approximation of Theorem 5.1.2 only, we derive the following approximation.

LEMMA 5.2.7. *For each $0 < \rho < 1$ we can enrich the probability space with a uniformly distributed random variable V, such that, as $n \to \infty$,*

$$\begin{aligned}&\Big| \max_{1 \leq k \leq \frac{n}{\log n}} \frac{\left|\sum_{i=1}^k y_i\right|}{\sqrt{k}} - \max_{1 \leq k \leq \frac{n}{\log n}} \frac{\left|\Gamma^{1/2} W_{1n}(k)\right|}{\sqrt{k}} \Big| \\ &= o_P\left((\log\log n)^{-1/2}\right),\end{aligned} \tag{5.2.22}$$

where, for each n, the Wiener process $\{W_{1n}(t),\ 0 \leq t \leq n/\log n\}$ is a measurable function of $\{\epsilon_1, \ldots, \epsilon_{n/\log n}, V\}$.

PROOF OF THEOREM 5.2.1. Consider the decomposition

$$\frac{T_n(k)}{\sqrt{n}} = \left(1 - \frac{k}{n}\right)^{1/2} \frac{S(k)}{\sqrt{n}} - \left(\frac{k}{n}\right) \frac{S(n) - S(k)}{\sqrt{n-k}}, \tag{5.2.23}$$

where $S(k)$ denotes the k-th partial sum. Using Theorem 4.2.2 we arrive at

$$\max_{1\le k\le n/\log n}\left|\frac{|T_n(k)|}{\sqrt{n}}-\left(1-\frac{k}{n}\right)^{1/2}\frac{|S(k)|}{\sqrt{n}}\right|=O_P\left((\log n)^{-1}\right). \tag{5.2.24}$$

Therefore

$$A(\log n)\,\Gamma^{-1/2}\max_{1\le k\le n/\log n}\frac{|T_n(k)|}{\sqrt{n}}-D(\log n)\xrightarrow{\mathcal{P}}-\infty. \tag{5.2.25}$$

Since the law of the iterated logarithm at zero implies

$$\sup_{1/\log n\le t\le 1-(1/\log n)}\frac{|W(t)|}{\sqrt{t}}=O_P\left((\log\log\log n)^{1/2}\right), \tag{5.2.26}$$

an application of (5.2.10) together with scaling properties of the Wiener process yields

$$\max_{n/\log n\le k\le n-(n/\log n)}\frac{|T_n(k)|}{\sqrt{n}}=O_P\left((\log\log\log n)^{1/2}\right). \tag{5.2.27}$$

Hence

$$A(\log n)\,\Gamma^{-1/2}\max_{n/\log n\le k\le n-(n/\log n)}\frac{|T_n(k)|}{\sqrt{n}}-D(\log n)\xrightarrow{\mathcal{P}}-\infty. \tag{5.2.28}$$

By (5.2.17), we have

$$\max_{n-(n/\log n)\le k\le n-1}\left(1-\frac{k}{n}\right)\frac{|S(n)-S(k)|}{\sqrt{n-k}}=O_P\left(\frac{(\log\log n)^{1/2}}{\log n}\right). \tag{5.2.29}$$

Hence

$$\max_{n-(n/\log n)\le k\le n-1}\left|\frac{|T_n(k)|}{\sqrt{n}}-\frac{|S(n)-S(k)|}{\sqrt{n-k}}\right|=O_P\left(\frac{(\log\log n)^{1/2}}{\log n}\right). \tag{5.2.30}$$

The final assertion follows by standard arguments as in the proof of Theorem 2.1.2 via using (5.2.21). □

PROOF OF THEOREM 5.2.2. Using Theorem 5.1.2 and following the pattern in the proof of Theorem 2.1.4, we derive

$$\max_{\frac{n}{\log n}\le k\le n-\frac{n}{\log n}}|G_n(k)|=O_P(\log\log\log n)\quad(n\to\infty). \tag{5.2.31}$$

Therefore, using Lemma 5.2.6 and Lemma 5.2.7 and reproving (2.1.40) and (2.1.45), it suffices to prove

$$\lim_{n\to\infty} P\left[\max\left\{\max_{1\le k\le n/\log n}|W_{1n}(k)|, \max_{n-(n/\log n)\le k\le n-1}|W_{2n}(n-k)|\right\}\right.$$
$$\left.\le (y+D(\log n))/A(\log n)\right]$$
$$= \exp\{-2\exp\{-y\}\}. \tag{5.2.32}$$

Observe the construction of the approximating Wiener processes in Lemma 5.2.6 and Lemma 5.2.7. Since the error sequence is α-mixing, it suffices to prove

$$\lim_{n\to\infty} P\left[\max_{1\le k\le n/\log n}\frac{|W_{1n}(k)|}{\sqrt{k}} \le (y+D(\log n))/A(\log n)\right]$$
$$\times P\left[\max_{n-(n/\log n)\le k\le n-1}\frac{|W_{2n}(n-k)|}{\sqrt{n-k}} \le (y+D(\log n))/A(\log n)\right]$$
$$= \exp\{-2\exp\{-y\}\}. \tag{5.2.33}$$

The final assertion follows now by the same arguments as in the proof of Theorem 2.1.5. $\square$

APPENDIX A

Maximal Inequalities for Mixingales

A mixingale sequence can be viewed as an asymptotic equivalent of a martingale difference sequence. Let $(\Omega, \mathcal{F}, P)$ denote a probability space on which there is a sequence of random variables $\{X_k,\ -\infty \le k \le \infty\}$. Let $\{\mathcal{F}_k,\ -\infty \le k \le \infty\}$ be a sequence of sub-σ-algebras which are increasing in k. We will represent the conditional expectation $E\left[X|\mathcal{F}_k\right]$ by $E_k X$.

Definition A.1. *The sequence $\{(X_k, \mathcal{F}_k),\ k \ge 1\}$ is an $\mathcal{L}^r$-mixingale $(r \ge 1)$ if, for sequences of finite nonnegative constants c_n and ψ_m where $\psi_m \downarrow 0 \quad (m \to \infty)$, we have for all $n \ge 1,\ m \ge 0$*

$$\|E_{n-m}X_n\|_r \le c_n\psi_m \quad \text{and} \quad \|X_n - E_{n+m}X_n\|_r \le c_n\psi_m. \tag{A.1}$$

McLeish [79] originally introduced $\mathcal{L}^2$-mixingales and presented a telescoping sum representation which is valid in $\mathcal{L}^r$ for $\mathcal{L}^r$-mixingales $(r \ge 1)$. Observing that $\{(E_{i-n-1}X_i, \mathcal{F}_{i-n}),\ n \ge 1\}$ is a reversed martingale, the reversed martingale convergence theorem implies

$$E_{i-n-1}X_i \to E_{-\infty}X_i \quad a.s. \quad (n \to \infty).$$

Since $\|E_{-\infty}X_i\|_r \le c_i\psi_m$ for all m, and hence is 0. Moreover, $\{(E_{i+m}X_i, \mathcal{F}_{i+m}),\ m \ge 1\}$ is a uniformly integrable martingale. The martingale convergence theorem yields

$$E_{i+m}X_i \to E_{\infty}X_i \quad a.s. \quad (m \to \infty),$$

and by the mixingale property $X_i = E_{+\infty}X_i$ almost surely for all n. Therefore, for each $i \ge 1$, the telescoping sum

$$X_i = \sum_{k=-m}^{n} E_{i-k}X_i - E_{i-k-1}X_i = E_{i+m}X_i - E_{i-n-1}X_i$$

yields the following representation

$$X_i = \sum_{k=-\infty}^{\infty} E_{i-k}X_i - E_{i-k-1}X_i \quad a.s. \tag{A.2}$$

which is a crucial tool in the proof of McLeish's $\mathcal{L}^2$-mixingale analogue of Doob's inequality.

Hansen [54] contributes a maximal inequality for $\mathcal{L}^r$-mixingales $(r > 1)$. Recently, Meng and Ling [80, Remark 2] raised doubts whether Hansen's maximal inequality is true for $1 < r < 2$ or not. They proved, via using estimates by Li [70], another maximal inequality for triangular $\mathcal{L}^r$-mixingale arrays $(1 < r \leq 2)$. Here, we will present a direct proof, mainly along the lines of Hansen [54] via using a standard convexity inequality, to estimate $\| \max_{m \leq j \leq n} | \sum_{i=m}^{j} X_i | \|_r$. For $m = 1$ our estimates coincide with the aforementioned results.

THEOREM A.1. *Let the sequence* $\{(X_k, \mathcal{F}_k)\,,\ k \geq 1\}$ *be an* $\mathcal{L}^r$*-mixingale* $(r \geq 1)$ *such that* $\sum_{k=0}^{\infty} \psi_k < \infty$. *If* $1 < r < 2$, *then there exists a finite constant* C_1, *such that for all* $n \geq 1$ *and* $1 \leq m \leq n$,

$$\Big\| \max_{m \leq j \leq n} \Big| \sum_{i=m}^{j} X_i \Big| \Big\|_r \leq C_1 \sum_{k=0}^{\infty} \psi_k \left(\sum_{i=m}^{n} c_i^r \right)^{1/r} \tag{A.3}$$

and, if $r \geq 2$,

$$\Big\| \max_{m \leq j \leq n} \Big| \sum_{i=m}^{j} X_i \Big| \Big\|_r \leq C_1 \sum_{k=0}^{\infty} \psi_k \left(\sum_{i=m}^{n} c_i^2 \right)^{1/2}. \tag{A.4}$$

PROOF. For each $i \geq 1$ and integer k let

$$X_{ki} = E_{i-k} X_i - E_{i-k-1} X_i. \tag{A.5}$$

Moreover, put $S_j = \sum_{i=1}^{j} X_i$. From (A.2) and Minkowski's inequality we have

$$\| \max_{m \leq j \leq n} |S_j - S_{m-1}| \, \|_r \leq \sum_{k=-\infty}^{\infty} \Big\| \max_{m \leq j \leq n} \Big| \sum_{i=m}^{j} X_{ki} \Big| \Big\|_r. \tag{A.6}$$

The sequence $\{(X_{k,i+m-1}, \mathcal{F}_{i+m-1-k})\,,\ i \geq 1\}$ is a martingale difference sequence. Since $\max_{m \leq j \leq n} \sum_{i=m}^{j} X_{ki} = \max_{1 \leq j \leq n-m+1} \sum_{i=1}^{j} X_{k,i+m-1}$, an application of Doob's inequality yields

$$\sum_{k=-\infty}^{\infty} \Big\| \max_{m \leq j \leq n} \Big| \sum_{i=m}^{j} X_{ki} \Big| \Big\|_r \leq \frac{r}{r-1} \sum_{k=-\infty}^{\infty} \Big\| \sum_{i=1}^{n-m+1} X_{k,i+m-1} \Big\|_r. \tag{A.7}$$

From Burkholder's inequality, cf. e.g. Hall and Heyde [53, Theorem 2.10], we arrive via (A.6) and (A.7) at

$$\| \max_{m \leq j \leq n} |S_j - S_{m-1}| \, \|_r \leq 18r \left(\frac{r}{r-1} \right)^{1/2} \sum_{k=-\infty}^{\infty} \left(\Big\| \sum_{i=1}^{n-m+1} X_{k,i+m-1}^2 \Big\|_{r/2}^{r/2} \right)^{1/r} \tag{A.8}$$

Consider the case $r \geq 2$. We derive from (A.8) via an application of Minkowski's inequality

$$\| \max_{m \leq j \leq n} |S_j - S_{m-1}| \, \|_r \leq C_0 \sum_{k=-\infty}^{\infty} \left(\sum_{i=1}^{n-m+1} \|X_{k,i+m-1}\|_r^2 \right)^{1/2}, \tag{A.9}$$

where $C_0 = 18r\,(r/(r-1))^{1/2}$. Observe, from (A.1) and (A.5), for $k \geq 0$

$$\begin{aligned} \|X_{k,i+m-1}\|_r &\leq \|E_{i+m-1-k}X_{i+m-1}\|_r \\ &\quad + \|E_{i+m-1-k-1}X_{i+m-1}\|_r \leq 2c_{i+m-1}\psi_k \end{aligned} \tag{A.10}$$

and for $k < 0$

$$\begin{aligned} \|X_{k,i+m-1}\|_r &\leq \|X_{i+m-1} - E_{i+m-1-k}X_{i+m-1}\|_r \\ &\quad + \|X_{i+m-1} - E_{i+m-1-k-1}X_{i+m-1}\|_r \\ &\leq 2c_{i+m-1}\psi_{|k|}. \end{aligned} \tag{A.11}$$

Hence, we have

$$\| \max_{m \leq j \leq n} |S_j - S_{m-1}| \, \|_r \leq 4C_0 \sum_{k=0}^{\infty} \psi_k \left(\sum_{i=m}^{n} c_i^2 \right)^{1/2}, \tag{A.12}$$

which yields (A.4). Towards this end, consider the case $1 < r < 2$. Since $2/r > 1$, the convexity inequality

$$\Big(\sum_{i=1}^{n-m+1} (|X_{k,i+m-1}|^r)^{2/r} \Big)^{r/2} \leq \sum_{i=1}^{n-m+1} |X_{k,i+m-1}|^r \quad a.s. \tag{A.13}$$

holds. We derive via (A.8) and (A.13)

$$\| \max_{m \leq j \leq n} |S_j - S_{m-1}| \, \|_r \leq C_0 \sum_{k=-\infty}^{\infty} \left(E \sum_{i=1}^{n-m+1} |X_{k,i+m-1}|^r \right)^{1/r}. \tag{A.14}$$

Finally, we arrive via (A.10) and (A.11) at

$$\| \max_{m \leq j \leq n} |S_j - S_{m-1}| \, \|_r \leq 4C_0 \sum_{k=0}^{\infty} \psi_k \left(\sum_{i=m}^{n} c_i^r \right)^{1/r}, \tag{A.15}$$

where $C_0 = 18r\,(r/(r-1))^{1/2}$. This implies (A.3) $\square$

Bibliography

[1] Aue, A.; Berkes, I.; Horváth, L.: Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli* 12, 583–608 (2006).

[2] Balan, R.; Zamfirescu, I.M.: Strong approximation for mixing sequences with infinite variance. *Electron. Comm. Probab.* 11, 11–23 (2006).

[3] Basrak, B.; Davis, R.A.; Mikosch, T.: Regular variation of GARCH processes. *Stochastic Process. Appl.* 99, 95–115 (2002).

[4] Berbee, H.C.P.: *Random walks with stationary increments and renewal theory.* Mathematical Centre Tracts 112. Amsterdam: Mathematisch Centrum 1979.

[5] Berkes, I.; Philipp, W.: Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* 7, 29-54 (1979).

[6] Berkes, I.; Gombay, E.; Horváth, L.: Testing for changes in the covariance structure of linear processes. *J. Statist. Plann. Inference* 139, 2044–2063 (2009).

[7] Berkes, I.; Horváth, L.; Kokoszka, P.: GARCH processes: structure and estimation. *Bernoulli* 9, 201–227 (2003).

[8] Berkes, I.; Hörmann, S.; Horváth, L.: The functional central limit theorem for a family of GARCH observations with applications. *Statist. Probab. Lett.* 78, 2725–2730 (2008).

[9] Bickel, P.J.; Rosenblatt, M.: On some global measures of the deviations of density function estimates. *Ann. Statist.* 1, 1071–1095 (1973).

[10] Billingsley, P.: *Probability and measure.* 3rd ed. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons 1995.

[11] Billingsley, P.: *Convergence of probability measures.* 2nd ed. Wiley Series in Probability and Statistics. New York: John Wiley & Sons 1999.

[12] Bradley, R.C.: Approximation theorems for strongly mixing random variables. *Michigan Math. J.* 30, 69–81 (1983).

[13] Bradley, R.C.: Basic properties of strong mixing conditions. In: *Dependence in probability and statistics.* 165-192, Ed. by E. Eberlein and M.S. Taqqu, Boston: Birkhäuser 1985.

[14] Bradley, R.C.: Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* 2, 107-144 (2005).

[15] Bradley, R.C.: *Introduction to strong mixing conditions. Vol. 1.* Heber City, Utah: Kendrick Press 2007.

[16] Bradley, R.C.: On a theorem of Rosenblatt. *Acta Sci. Math. (Szeged)* 75, 347-359 (2009).

[17] Breiman, L.: On the tail behavior of sums of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 9, 20–25 (1967).

[18] Brockwell, P.J.; Davis, R.A.: *Time series: theory and methods.* 2nd ed. Springer Series in Statistics. New York: Springer-Verlag 1991.

[19] Bulinski, A.; Shashkin A.: *Limit theorems for associated random fields and related systems.* Advanced Series on Statistical Science & Applied Probability 10. Singapore: World Scientific Publishing 2007.

[20] Carrasco, M.; Chen, X.: Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17–39 (2002).

[21] Cheng, Q. On time-reversibility of linear processes. *Biometrika* 86, 483–486 (1999).

[22] Csörgő, M.: A glimpse of the impact of Pál Erdős on probability and statistics. *Canad. J. Statist.* 30, 493–556 (2002).

[23] Csörgő, M.; Horváth, L.: *Weighted approximations in probability and statistics.* Wiley Series in Probability and Mathematical Statistics. Chichester: John Wiley & Sons 1993.

[24] Csörgő, M.; Horváth, L.: *Limit theorems in change-point analysis.* Wiley Series in Probability and Mathematical Statistics. Chichester: John Wiley & Sons 1997.

[25] Csörgő, M.; Révész, P.: A new method to prove Strassen type laws of invariance principle. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 31, 255–259 (1975).

[26] Csörgő, M.; Révész, P.: *Strong approximations in probability and statistics.* Volume in Probability and Mathematical Statistics. New York-London: Academic Press 1980.

[27] Darling, D.A.; Erdős, P.: A limit theorem for the maximum of normalized sums of independent random variables. *Duke Math. J.* 23, 143-155 (1956).

[28] DasGupta, A.: *Asymptotic theory of statistics and probability.* Springer Texts in Statistics. New York: Springer 2008.

[29] Davidson, J.: *Stochastic limit theory.* Advanced Texts in Econometrics. New York: Oxford University Press 1994.

[30] Davidson, J.: Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. *J. Econometrics* 106, 243–269 (2002).

[31] Davis, R.A.; Yao, Y.C.: The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates. *Sankhyā Ser. A* 48, 339–353 (1986).

[32] Davis, R.A.; Huang, D.W.; Yao, Y.C.: Testing for a change in the parameter values and order of an autoregressive model. *Ann. Statist.* 23, 282–304 (1995).

[33] Doukhan, P.: *Mixing. Properties and examples.* Lecture Notes in Statistics 85. New York: Springer-Verlag 1994.

[34] Duan, J.C.: Augmented GARCH(p, q) process and its diffusion limit. *J. Econometrics* 79, 97–127 (1997).

[35] Dudley, R.M.: *Real analysis and probability.* Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics 74. Cambridge: Cambridge University Press 2002.

[36] Durrett, R.: *Probability: theory and examples.* 2nd ed. Belmont, California: Duxbury Press 1996.

[37] Eberlein, E.: On strong invariance principles under dependence assumptions. *Ann. Probab.* 14, 260–270 (1986).

[38] Einmahl, U.: *Starke Approximationen von Partialsummen unabhängiger Zufallsvektoren.* Dissertation. Universität zu Köln 1984.

[39] Einmahl, U.: A useful estimate in the multidimensional invariance principle. *Probab. Theory Related Fields* 76, 81-101 (1987).

[40] Einmahl, U.: Strong invariance principles for partial sums of independent random vectors. *Ann. Probab.* 15, 1419-1440 (1987).

[41] Einmahl, U.: The Darling-Erdős theorem for sums of i.i.d. random variables. *Probab. Theory Related Fields* 82, 241–257 (1989).

[42] Einmahl, U.: Extensions of results of Komlós, Major, and Tusnády to the multivariate case. *J. Multivariate Anal.* 28, 20-68 (1989).

[43] Einmahl, U.; Mason, D.M.: Darling-Erdős theorems for martingales. *J. Theoret. Probab.* 2, 437–460 (1989).

[44] Elstrodt, J.: *Maß- und Integrationstheorie.* 4., korrigierte Auflage. Reihe: Grundwissen Mathematik. Berlin: Springer-Verlag 2005.

[45] Fan, J.; Yao, Q.: *Nonlinear time series. Nonparametric and parametric methods.* Springer Series in Statistics, New-York etc.: Springer-Verlag 2003.

[46] Francq, C.; Zakoïan, J.M.: Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10, 605–637 (2004).

[47] Francq, C.; Zakoïan, J.M.: Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* 22, 815–834 (2006).

[48] Freedman, D.: *Brownian motion and diffusion.* Holden-Day Series in Probability and Statistics. San Francisco: Holden-Day 1971.

[49] Gänssler, P. and Stute, W.: *Wahrscheinlichkeitstheorie.* Reihe: Hochschultext. Berlin-New York: Springer-Verlag 1977.

[50] Gombay, E.: Change detection in linear regression with time series errors. *Canad. J. Statist.* 38, 65–79 (2010).

[51] Gombay, E.; Horváth, L.: Asymptotic distributions of maximum likelihood tests for change in the mean. *Biometrika* 77, 411–414 (1990).

[52] Gorodetskii, V.V.: On the Strong Mixing Property for Linear Sequences. *Theory Probab. Appl.* 22, 411-413 (1978).

[53] Hall, P.; Heyde, C.C.: *Martingale limit theory and its application.* Volume in Probability and Mathematical Statistics. New York-London: Academic Press 1980.

[54] Hansen, B.E.: Strong laws for dependent heterogeneous processes. *Econometric Theory* 7, 213–221 (1991).

[55] Hansen, B.E.: GARCH$(1,1)$ processes are near epoch dependent. *Econom. Lett.* 36, 181–186 (1991).

[56] Hanson, D.L.; Russo, R.P.: Some results on increments of the Wiener process with applications to lag sums of i.i.d. random variables. *Ann. Probab.* 11, 609–623 (1983).

[57] Horváth, L.: The maximum likelihood method for testing changes in the parameters of normal observations. *Ann. Statist.* 21, 671-680 (1993).

[58] Horváth, L.: Detection of changes in linear sequences. *Ann. Inst. Statist. Math.* 49, 271–283 (1997).

[59] Horváth, L.; Kokoszka, P.; Steinebach, J.: Testing for changes in multivariate dependent observations with an application to temperature changes. *J. Multivariate Anal.* 68, 96–119 (1999).

[60] Hušková, M.; Prášková, Z.; Steinebach, J.: On the detection of changes in autoregressive time series. I. Asymptotics. *J. Statist. Plann. Inference* 137, 1243–1259 (2007).

[61] Jain, N.C.; Jogdeo, K.; Stout W.F.: Upper and lower functions for martingales and mixing processes. *Ann. Probab.* 3, 119–145 (1975).

[62] Karatzas, I.; Shreve, S.E.: *Brownian motion and stochastic calculus.* 2nd ed. Graduate Texts in Mathematics 113. New York: Springer-Verlag 1991.

[63] Kirch, C.: *Resampling Methods for the Change Analysis of Dependent Data.* Dissertation. Universität zu Köln 2006.

[64] Komlós, J.; Major, P.; Tusnády, G.: An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 32, 111–131 (1975).

[65] Kuelbs, J.; Philipp, W.: Almost sure invariance principles for partial sums of mixing B-valued random variables. *Ann. Probab.* 8, 1003–1036 (1980).

[66] Leadbetter, M.R.; Lindgren, G; Rootzén, H.: *Extremes and related properties of random sequences and processes.* Springer Series in Statistics. New York-Berlin: Springer-Verlag.

[67] Le Cam, L.: *On the Prohorov distance between the empirical process and the associated Gaussian bridge.* Technical report 170, Berkeley: Department of Statistics, U.C. Berkeley 1988.

[68] Lee, S.; Song, J.: Test for parameter change in ARMA models with GARCH innovations. *Statist. Probab. Lett.* 78, 1990–1998 (2008).

[69] Lehmann, E.L.; Romano, J.P.: *Testing statistical hypotheses.* 3rd ed. Springer Texts in Statistics. New York: Springer.

[70] Li, Y.: A martingale inequality and large deviations. *Statist. Probab. Lett.* 62, 317–321 (2003).

[71] Lifshits, M.A.: Lecture Notes on Strong Approximation. *Publ. IRMA Lille* 53-XIII, 1-25 (2000).

[72] Lin, Z.; Lu, C.: *Limit theory for mixing dependent random variables.* Mathematics and its Applications 378. Dordrecht: Kluwer Academic Publishers, New York: Science Press 1996.

[73] Lindvall, T.: *Lectures on the coupling method.* Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons 1992. Reprint: Mineola, New York: Dover Publications.

[74] Ling, S.: Testing for change points in time series models and limiting theorems for NED sequences. *Ann. Statist.* 35, 1213–1237 (2007).

[75] Ling, S.; Li, W.K.: Limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with general autoregressive heteroscedastic errors. *Ann. Statist.* 26, 84–125 (1998).

[76] Liu, W.; Lin, Z.: Strong approximation for a class of stationary processes. *Stochastic Process. Appl.* 119, 249–280 (2009).

[77] Major, P.: The approximation of partial sums of independent RV's. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 35, 213–220 (1976).

[78] Major, P.: An improvement of Strassen's invariance principle. *Ann. Probab.* 7, 55-61 (1979).

[79] McLeish, D.L.: A maximal inequality and dependent strong laws. *Ann. Probab.* 3, 829–839 (1975).

[80] Meng, Y.; Lin, Z.: Maximal inequalities and laws of large numbers for L_q-mixingale arrays. *Statist. Probab. Lett.* 79, 1539–1547 (2009).

[81] Merlevède, F.; Peligrad, M.: On the coupling of dependent random variables and applications. In: *Empirical process techniques for dependent data.* 171–193, Ed. by H. Dehling, T. Mikosch and M. Sørensen, Boston: Birkhäuser 2002.

[82] Meyn, S.; Tweedie, R.L.: *Markov chains and stochastic stability.* 2nd ed. Cambridge: Cambridge University Press 2009.

[83] Miller, R.; Siegmund, D.: Maximally selected chi square statistics. *Biometrics* 38, 1011–1016 (1982).

[84] Obłój, J.: The Skorokhod embedding problem and its offspring. *Probab. Surv.* 1, 321–390 (2004).

[85] Petrov, V.V.: *Limit theorems of probability theory. Sequences of independent random variables.* Oxford Studies in Probability 4. New York: Oxford University Press 1995.

[86] Philipp, W.; Stout, W.: Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* no. 161 (1975).

[87] Philipp, W.: Invariance principles for independent and weakly dependent random variables. In: *Dependence in probability and statistics.* 225-268, Ed. by E. Eberlein and M.S. Taqqu, Boston: Birkhäuser 1985.

[88] Pollard, D.: *A user's guide to measure theoretic probability.* Cambridge Series in Statistical and Probabilistic Mathematics 8. Cambridge: Cambridge University Press 2002.

[89] Révész, P.: *The laws of large numbers.* Volume in Probability and Mathematical Statistics. New York-London: Academic Press 1968.

[90] Rosenblatt, M.: A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.* 42, 43-47 (1956).

[91] Rosenblatt, M.: *Markov processes. Structure and asymptotic behavior.* Die Grundlehren der mathematischen Wissenschaften 184. New York-Heidelberg: Springer-Verlag 1971.

[92] Rosenblatt, M.: *Gaussian and non-Gaussian linear time series and random fields.* Springer Series in Statistics. New York: Springer-Verlag 2000.

[93] Schwarz, G.: Finitely determined processes - an indiscrete approach. *J. Math. Anal. Appl.* 76, 146-158 (1980).

[94] Serfling, R.J.: Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist.* 41, 1227–1234 (1970).

[95] Shao, Q.M.: Almost sure invariance principles for mixing sequences of random variables. *Stochastic Process. Appl.* 48, 319–334 (1993).

[96] Skorokhod, A.V. *Studies in the theory of random processes.* English Translation. Reading, Massachusetts: Addison-Wesley Publishing 1965.

[97] Skorohod, A.V.: On a representation of random variables. *Theory Probab. Appl.* 21, 628-632 (1977).

[98] Sotres, D.A.; Ghosh, M.: Strong convergence of linear rank statistics for mixing processes. *Sankhyā Ser. B* 39, 1–11 (1977).

[99] Steinebach, J.: Variance estimation based on invariance principles. *Statistics* 27, 15–25 (1995).

[100] Strassen, V.: An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 3, 211-226 (1964).

[101] Strassen, V.: Almost sure behavior of sums of independent random variables and martingales. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.* Vol. 2, 315–343, Ed. by J. Neyman and L. Le Cam, Berkeley-Los Angeles: University of California Press 1967.

[102] Szyszkowicz, B.: Weighted stochastic processes under contiguous alternatives. *C. R. Math. Rep. Acad. Sci. Canada* 13, 211–216 (1991).

[103] Varadhan, S.R.S.: *Probability Theory.* Courant Lecture Notes in Mathematics 7. New York, New York: Courant Institute of Mathematical Sciences; Providence, Rhode Island: American Mathematical Society 2001.

[104] Volkonskii, V.A.; Rozanov, Yu.A.: Some limit theorems for random functions. I. *Theory Probab. Appl.* 4, 178-197 (1959).

[105] Volkonskii, V.A.; Rozanov, Yu.A.: Some limit theorems for random functions. II. *Theory Probab. Appl.* 6, 186-199 (1961).

[106] Vostrikova, L.Yu.: Detection of "disorder" of a Wiener process. *Theory Probab. Appl.* 26, 356-362 (1981).

[107] Weyl, H.: Über die Konvergenz von Reihen, die nach Orthogonalfunktionen fortschreiten. *Math. Ann.* 67, 225-245 (1909).

[108] Worsley, K.J.: Confidence regions and test for a change-point in a sequence of exponential family random variables. *Biometrika* 73, 91–104 (1986).

[109] Wu, W.B.: Strong invariance principles for dependent random variables. *Ann. Probab.* 35, 2294–2320 (2007).

[110] Yurinskii, V.V.: On the error of the Gaussian approximation for convolutions. *Theory Probab. Appl.* 22, 236-247 (1977).

[111] Zaitsev, A.Yu.: Estimates for the strong approximation in multidimensional central limit theorem. In: *Proceedings of the International Congress of Mathematicians.* Vol. III, 107-116. Ed. by L. Tatsien (L. Daqian), Beijing: Higher Ed. Press 2002.

[112] Zygmund, A.: *Trigonometric series.* 3rd ed. Volumes I & II combinded. Cambridge: Cambridge University Press 2002.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; daß sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, daß ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Josef G. Steinebach betreut worden.

Köln, 21. März 2011

Alexander Schmitz

Lebenslauf

Name:	Schmitz
Vorname:	Alexander
Geburtsort:	Lindlar
Geburtsjahr:	1980

24.07.2000	Abitur am Erzbischöflichen St.-Angela-Gymnasium in Wipperfürth
31.08.2001	Beendigung des Grundwehrdienst
28.08.2003	Vordiplom in Wirtschaftsmathematik an der Universität Ulm
24.07.2007	Diplom in Mathematik an der Universität zu Köln
09.10.2007	Einstellung als wissenschaftlicher Mitarbeiter an der Universität zu Köln

Köln, 21. März 2011

Alexander Schmitz