On the limiting Pitman asymptotic relative efficiency of two Cramér-von Mises tests

A dissertation presented

by

Eva Sabina Hörmann

to

The Faculty of Mathematics and Computer Science, Physics, Geography

of the Justus-Liebig-University Giessen

in partial fulfillment of the requirements

for the degree of

Doctor rerum naturalium

Justus-Liebig-University Giessen

June 2013

To Elisabeth Sophia Langsdorf and Karl Hörmann

Contents

List of Symbols

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$	natural, integer, real, complex numbers
$\mathbb{R}^*$	$\mathbb{R} \setminus \{0\}$
$\mathbb{R}^{\mathbb{N}}$	infinite product space $\mathbb{R} \times \mathbb{R} \times \ldots$
$\mathbb{R}^{n \times n}$	real $n \times n$-matrices
$\mathcal{B}^*, \bigotimes_{i\in\mathbb{N}} \mathcal{B}^*$	Borel σ-algebra on $\mathbb{R}$, $\mathbb{R}^{\mathbb{N}}$
$\lambda\!\!\lambda$	one-dimensional Lebesgue measure
$\sim$	equality in distribution
$\xrightarrow[n]{\mathcal{L}}$	weak convergence, convergence in distribution
B°	Brownian bridge on [0,1]
$\mathcal{N}(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$\Vert f\Vert_\infty = \sup_{x\in\mathbb{R}} \vert f(x)\vert$	supremum norm of a bounded function f
$\Vert f\Vert_{w,\infty} = \Vert wf\Vert_\infty$	weighted supremum (semi-)norm of f with weight function w, p. 103
$\Vert f\Vert_{s,\infty}$	weighted supremum (semi-)norm of f with weight function $w(\cdot) = \vert\cdot\vert^s$, p. 129
$x \wedge y$, $x \vee y$	minimum of x and y, maximum of x and y
$\lfloor x\rfloor$, $\lceil x\rceil$	largest integer not greater than x, smallest integer not less than x
$o_{P_F}(a_n)$, $O_{P_F}(a_n)$	stochastic order symbols with respect to the probability measure P_F
$o_P^u(a_n)$, $O_P^u(a_n)$	uniform stochastic order symbols, p. 17, p. 55
x^T	transpose of vector or matrix x
$\Vert\cdot\Vert_{\mathrm{Fr}}$	Frobenius matrix norm, p. 55
B^+	Moore-Penrose pseudoinverse of matrix B
$[h]_\gamma$	γ-Hölder coefficient of function h, p. 103
$C[-\infty, \infty]$	space of continuous functions on $[-\infty, \infty]$
$D[-\infty, \infty]$	space of càdlàg functions on $[-\infty, \infty]$
$\overline{A}$	complement of the set A
$m_\tau(F)$	p. 40
$\mathcal{G}_q$	set of all centered continuous distribution functions F on $\mathbb{R}$ with finite absolute q-th moment, p. 28
$\widetilde{\mathcal{G}}_q$	elements of $\mathcal{G}_q$ with absolute τ-th moment equal to $1/\tau$, p. 41
$\mathcal{G}_q^u$	elements of $\mathcal{G}_q$ having uniformly continuous Lebesgue density f, p. 100
$\mathcal{G}_{q,\gamma,w}$	elements of $\mathcal{G}_q^u$ with $[f]_\gamma$ and $\Vert f\Vert_{w,\infty}$ finite, p. 103
$\mathcal{G}_{q,\gamma,s}$	$\mathcal{G}_{q,\gamma,w}$ with weight function $w(\cdot) = \vert\cdot\vert^s$, p. 129
$\widetilde{\mathcal{G}}_{q,\gamma,s}$	elements of $\mathcal{G}_{q,\gamma,s}$ with absolute τ-th moment equal to $1/\tau$, p. 132
d_q	Kantorovich-Wasserstein metric, p. 28
d_K	Kolmogorov metric, p. 28
$d_{q,\gamma,w}$, $d_{q,\gamma,s}$	metrics, p. 104, p. 129
$\square$, $\blacklozenge$	end of proof, end of remark

1 Introduction

The classical empirical distribution function F_n of a sample of n independent and identically distributed observations is the nonparametric maximum likelihood estimator of the underlying distribution function F if this is completely unknown, see e.g. section 2.1 in Owen [25]. For this and a considerable number of other reasons the empirical distribution function plays a prominent role in statistical inference. For example, many classical goodness-of-fit statistics such as the Kolmogorov-Smirnov statistic or the Cramér-von Mises statistic are based on it.

Suppose now that it is additionally known that the underlying distribution function is centered. It may be considered as a drawback of the empirical distribution function F_n that it does not take this auxiliary information into account, since $\int_{\mathbb{R}} xF_n(dx) \neq 0$ in general. By using nonparametric maximum likelihood estimation under constraints, a centered empirical distribution function $\widetilde{F}_n$ can be constructed, see Owen [22–25] and Qin and Lawless [27]. Zhang [31] has established a functional central limit theorem for the empirical process $\sqrt{n}(\widetilde{F}_n - F)$ based on $\widetilde{F}_n$. The asymptotic variance of this process is pointwise not greater than that of the classical empirical process $\sqrt{n}(F_n - F)$, whose limit is a time-transformed Brownian bridge by the classical functional central limit theorem of Donsker. A corresponding result holds for the covariance matrices of the finite-dimensional distributions of the limit processes, see inequality (1.12) in Genz and Häusler [12]. Furthermore, it follows from Example 2 in section 5.3 of Bickel et al. [5] in combination with Zhang's result that the estimator $\widetilde{F}_n$ is asymptotically efficient for F in the sense of the Hájek-Le Cam convolution theorem. For F belonging to a parametric family $\{F(\cdot, \vartheta) \colon \vartheta \in \Theta\}$ of centered distribution functions, a functional central limit theorem for the empirical process with estimated parameter based on $\widetilde{F}_n$, i.e., for $\sqrt{n}(\widetilde{F}_n(\cdot) - F(\cdot, \hat{\vartheta}_n))$ with a suitable estimator $\hat{\vartheta}_n$ for ϑ, was derived in [12], see also Genz [11]. If ϑ is estimated appropriately, e.g., by maximum likelihood, the asymptotic variance of this modified empirical process is again seen to be pointwise less than or equal to the one of the classical empirical process with estimated parameter $\sqrt{n}(F_n(\cdot) - F(\cdot, \hat{\vartheta}_n))$, for which a functional central limit theorem was proven in the fundamental work of Durbin [7]. In this case a corresponding result also holds again for the covariance matrices of the finite-dimensional distributions of the limit processes, see inequality (2.16) in [12]. Note that in [22–25], [27], [31] and [12] more general auxiliary information than $\int_{\mathbb{R}} xF(dx) = 0$ is considered, but we will restrict our attention to the case of centered distributions.

While the model of independent and identically distributed centered data may not be of great relevance in practice, in various other important statistical models like in many regression and time series models the centeredness of the error variables is part of the model. Hence, in order to estimate the error distribution function F in such models nonparametrically at sample size n, instead of the standard empirical distribution function $F_{n,res}$ of the residuals one can use a centered version $\widetilde{F}_{n,res}$ in the spirit above, which includes the model assumption explicitly. Some investigations in this direction have already been made. For example, Genz [11] studied the estimation of the error distribution by $\widetilde{F}_{n,res}$ for autoregressive processes of order one in the case that $F = F(\cdot, \vartheta)$ for some $\vartheta \in \Theta$ and derived a functional central limit theorem for $\sqrt{n}(\widetilde{F}_{n,res}(\cdot) - F(\cdot, \hat{\vartheta}_{n,res}))$, the residual empirical process with estimated parameter based on $\widetilde{F}_{n,res}$. He showed that the distributional limit of this process is the same as that of the process $\sqrt{n}(\widetilde{F}_n(\cdot) - F(\cdot, \hat{\vartheta}_n))$ based on independent and identically distributed observations with common distribution function $F = F(\cdot, \vartheta)$ for suitable estimators $\hat{\vartheta}_{n,res}$ and $\hat{\vartheta}_n$ of ϑ. Since the ordinary residual empirical process with estimated parameter $\sqrt{n}(F_{n,res}(\cdot) - F(\cdot, \hat{\vartheta}_{n,res}))$ converges weakly to the same limit as the process $\sqrt{n}(F_n(\cdot) - F(\cdot, \hat{\vartheta}_n))$ in the model of independent and identically distributed data with distribution function $F = F(\cdot, \vartheta)$ when suitable estimators for ϑ are used, see section 3 in Genz [11] and the references therein, it follows again that if ϑ is estimated appropriately, the asymptotic variance of the residual empirical process with estimated parameter based on $\widetilde{F}_{n,res}$ is pointwise not greater than the one of the process based on $F_{n,res}$, and the analogous result also holds for the covariance matrices of the finite-dimensional distributions of the limit processes. For estimating the error distribution in a nonparametric homoscedastic regression model, Kiwitt et al. [17] consider inter alia the centered empirical distribution function $\widetilde{F}_{n,res}$ of the residuals and establish a functional central limit theorem for a corresponding stochastic process. They also compare the resulting asymptotic mean squared error with the analogous term for the ordinary empirical distribution function of the residuals and show for some examples of underlying error distributions that the former is considerably smaller than the latter due to a reduction of bias, see Example 4.1 in [17].

In models such as those above, for goodness-of-fit testing for F it is natural to consider the classical goodness-of-fit statistics with F_n and $F_{n,res}$ replaced by $\widetilde{F}_n$ and $\widetilde{F}_{n,res}$, respectively, so that each of the classical test statistics based on the ordinary (residual) empirical distribution function has a counterpart based on the centered (residual) empirical distribution function. In view of the above, it seems reasonable to presume that the goodness-of-fit tests based on $\widetilde{F}_n$ and $\widetilde{F}_{n,res}$ exhibit a better performance than their classical counterparts. To the best of our knowledge, up to now this has only been studied in a few cases. For independent and identically distributed observations, Genz and Häusler [12] considered testing the composite null hypothesis $\mathrm{H}_0\colon F \in \{F_0(\cdot/\sigma)\colon \sigma \in (0,\infty)\}$ for certain centered distribution functions F_0 and simulated the power of the asymptotic bootstrap test based on the classical Kolmogorov-Smirnov statistic with estimated parameter and of its counterpart using the centered empirical distribution function against some fixed alternatives. Their results show that the tests based on $\widetilde{F}_n$ lead to a higher power even for small sample sizes in most of the examples. Analogous results are derived by Genz [11] also for autoregressive processes of order one. In [15] the asymptotic power of the asymptotic tests based on the classical Cramér-von Mises statistic and on its modified version using $\widetilde{F}_n$ for testing the simple null hypothesis $\mathrm{H}_0\colon F = F_0$ for certain centered distribution functions F_0 in the case of independent and identically distributed data is computed numerically against a sequence of contiguous scale alternatives. It is found that in all of the investigated cases the test based on $\widetilde{F}_n$ has substantially better asymptotic power than the one based on F_n.

The object of this thesis is to provide further mathematical evidence that in the presence of centered distributions the use of Cramér-von Mises statistics based on the centered (residual) empirical distribution function instead of classical Cramér-von Mises statistics leads to improved asymptotic test procedures for goodness-of-fit testing. We will investigate these tests not only in the model of independent and identically distributed centered data, but also for certain stable autoregressive processes of arbitrary order with independent and identically distributed centered errors.

For comparing the performance of two sequences of tests for a given testing problem there are various concepts of asymptotic relative efficiency discussed in the literature. The relative efficiency of two sequences of tests is the ratio of the sample sizes needed with the two tests to obtain a given power β at the significance level α. Then clearly the sequence of tests is preferable that needs less observations to attain a power of β. As the relative efficiency will generally depend on the values of α, β, and on the alternative under which the power is considered, it is hardly possible to determine its value except in simple cases. For this reason several asymptotic procedures concerning the relative efficiency have been proposed, see e.g. Nikitin [21] for a comprehensive account. Since the quality of a sequence of tests can be assessed by its power at alternatives that are close to the null hypothesis and at small significance levels, the limit of the relative efficiency when the alternative approaches the null hypothesis and the level tends to zero is studied. In case of its existence, this quantity is called the limiting (as $\alpha \to 0$) Pitman asymptotic relative efficiency. Wieand [30] established a condition under which it is possible to equate the limiting Pitman asymptotic relative efficiency to the limit of the approximate Bahadur asymptotic relative efficiency, which is another concept for the comparison of two sequences of tests introduced by Bahadur [1]. As the approximate Bahadur asymptotic relative efficiency is in general easy to compute, this provides a means to determine the value of the limiting Pitman asymptotic relative efficiency. Using this approach, we will compare the performance of the two competing Cramér-von Mises tests in this thesis by examining their limiting Pitman asymptotic relative efficiency. In section 2 we will describe the aforementioned concepts of asymptotic relative efficiency in more detail and adjust Wieand's results to our setting, which differs from the one considered in [30].

The explicit definition of the centered empirical distribution function $\widetilde{F}_n$ based on a sample of independent and identically distributed centered random variables is given in section 3 and some results concerning its asymptotic stochastic behavior uniformly with respect to the underlying distribution of the data are proven. These uniform results are then used in the next section to verify Wieand's condition for the Cramér-von Mises statistics based on $\widetilde{F}_n$.

In section 4 we consider observations that are independent and identically distributed according to a centered distribution function F and determine the limiting Pitman asymptotic relative efficiency of the asymptotic tests based on the classical Cramér-von Mises statistics and on their counterparts using $\widetilde{F}_n$ for testing the simple null hypothesis $\mathrm{H}_0\colon F = F_0$ against $\mathrm{H}_1\colon F \in \mathcal{G}\backslash\{F_0\}$, where $\mathcal{G}$ is an appropriate set of continuous centered distribution functions, and for testing the composite null hypothesis $\mathrm{H}_0\colon F \in \mathcal{F}_\tau$ against $\mathrm{H}_1\colon F \in \mathcal{G} \setminus \mathcal{F}_\tau$, where $\mathcal{F}_\tau$ is the scale family generated by the exponential power distribution with fixed parameter $\tau \in (0,\infty)$. The class of exponential power distributions, whose explicit definition is given in subsection 4.2, includes both the normal and the double exponential distribution as special cases. The scale parameter of the scale family $\mathcal{F}_\tau$ will be estimated by maximum likelihood. For both of the above testing problems we will show in section 4 that the limiting Pitman asymptotic relative efficiency of the classical Cramér-von Mises test with respect to the modified test based on $\widetilde{F}_n$ is equal to the ratio of the largest eigenvalues of those Hilbert-Schmidt integral operators whose kernels are the (time-transformed) covariance functions of the limit processes under the null hypothesis of the empirical processes the test statistics are based upon. By results from [15] we will deduce that this ratio is strictly less than one in all of the cases considered, so that the sequence of tests based on the modified Cramér-von Mises statistic is preferable to the standard one in both testing problems.

A paper prior to our investigations which studies the limiting Pitman asymptotic relative efficiency of Cramér-von-Mises-type tests based on suitably weighted classical empirical processes with and without estimated parameter in the case of independent and identically distributed data is Wells [29]. Using the results of Wieand, Wells determined the limiting Pitman asymptotic relative efficiency of the test statistics with estimated parameter relative to their counterparts with fully specified distribution function under some regularity conditions in a model of parametric alternatives. Similar to the results above, he showed that the efficiency equals the ratio of the largest eigenvalues of certain Hilbert-Schmidt integral operators and is less than or equal to one, whence he concluded that the test procedure based on the statistic with estimated parameter is better than the one with a fully specified distribution function.

An important basic model in time series analysis is the autoregressive process. We will restrict our attention to certain stable autoregressive processes with independent and identically distributed centered errors in section 5 and section 6. More specifically, we will investigate strictly stationary stable autoregressive processes as well as stable autoregressive processes with fixed distribution of the starting values that does not vary with the error distribution. For these processes we will then consider goodness-of-fit tests for the error distribution using the classical Cramér-von Mises statistics based on the residual empirical distribution function $F_{n,res}$ and the modified statistics based on the centered residual empirical distribution function $\widetilde{F}_{n,res}$.

In section 5 we will discuss the residual empirical distribution functions $F_{n,res}$ and $\widetilde{F}_{n,res}$ for the aforementioned autoregressive processes in some detail and study in particular their asymptotic stochastic behavior uniformly with respect to the underlying distribution of the errors. Moreover, we will investigate the uniform stochastic behavior of the least squares estimator for the autoregressive parameter. These uniform results will then be used in section 6 to verify Wieand's condition for the Cramér-von Mises statistics based on $F_{n,res}$ and $\widetilde{F}_{n,res}$.

The limiting Pitman asymptotic relative efficiency of the asymptotic tests based on the aforementioned Cramér-von Mises statistics is studied in section 6 for testing the same simple and composite null hypotheses as in section 4, with F denoting the distribution function of the error variables of the autoregressive processes here. The set $\mathcal{G}$ of possible distribution functions is adjusted in this section to the model under consideration. The unknown autoregressive parameter will be estimated by least squares. For testing the composite null hypothesis $\mathrm{H}_0\colon F \in \mathcal{F}_\tau$ we will confine our investigations to strictly stationary stable autoregressive processes and stable autoregressive processes that start in zero. The scale parameter of the parametric family $\mathcal{F}_\tau$ will be estimated by the residual-based version of the maximum likelihood estimator for the scale parameter in the model of independent and identically distributed observations. Using Wieand's approach again, we will show that in both testing problems the limiting Pitman asymptotic relative efficiency of the asymptotic tests based on the Cramér-von Mises statistics using $F_{n,res}$ and $\widetilde{F}_{n,res}$ respectively is the same as the one of the respective tests based on F_n and $\widetilde{F}_n$ in the model of independent and identically distributed data determined in section 4. Hence, also for the stable autoregressive processes under consideration the goodness-of-fit tests based on the Cramér-von Mises statistics using $\widetilde{F}_{n,res}$ lead to better test procedures than the tests based on the classical statistics.

2 Asymptotic relative efficiency of two sequences of tests

There are various concepts of asymptotic relative efficiency for comparing the performance of two sequences of statistical tests for a given hypothesis testing problem. In this section, we will describe the concepts of approximate Bahadur asymptotic relative efficiency and Pitman asymptotic relative efficiency and extend a result of Wieand that specifies conditions under which the limit (as the alternative approaches the hypothesis) of the former efficiency coincides with the limit (as the level tends to zero) of the latter.

To begin with, let us introduce some notation. Throughout this thesis, the end of a proof is signaled by the symbol □ and the end of a remark by ⧫. Moreover, the minimum and maximum of two real numbers x and y will be denoted by $x \wedge y$ and $x \vee y$, respectively.
Now let $(\mathcal{G}, d)$ be a metric space. For every nonempty set $A \subset \mathcal{G}$, point $\gamma \in \mathcal{G}$ and $\epsilon > 0$ we set, as usual, $d(\gamma, A) := \inf\{d(\gamma, \hat{\gamma})\colon \hat{\gamma} \in A\}$ and $U_\epsilon(A) := \{\gamma \in \mathcal{G}\colon d(\gamma, A) < \epsilon\}$. If the set A is a singleton, say $A = \{\gamma_0\}$, we will write $U_\epsilon(\gamma_0)$ instead of $U_\epsilon(\{\gamma_0\})$.

Let $(\Omega, \mathcal{A})$ be a measurable space and let $\gamma \mapsto P_\gamma$ be an injective mapping from $\mathcal{G}$ into the set of probability measures on $\mathcal{A}$. Consider now the statistical model $(\Omega, \mathcal{A}, \{P_\gamma\colon \gamma \in \mathcal{G}\})$. It is required to test

$$\mathrm{H}_0\colon \gamma \in \mathcal{G}_0 \quad \text{versus} \quad \mathrm{H}_1\colon \gamma \in \mathcal{G} \setminus \mathcal{G}_0, \tag{2.1}$$

where $\mathcal{G}_0$ is a nonempty subset of $\mathcal{G}$ with

$$U_\epsilon(\mathcal{G}_0) \cap (\mathcal{G} \setminus \mathcal{G}_0) \neq \emptyset \quad \forall\, \epsilon > 0. \tag{2.2}$$

The foregoing condition ensures that the set $\mathcal{G}_0$ is not isolated in $\mathcal{G}$, but can be approximated by elements in $\mathcal{G} \setminus \mathcal{G}_0$. For each $n \in \mathbb{N}$, let T_n be a real-valued test statistic on $(\Omega, \mathcal{A})$ for testing (2.1) such that H_0 is being rejected if and only if $T_n > k(\alpha)$ with $k(\alpha) \in \mathbb{R}$ such that

$$P_\gamma\big(T_n > k(\alpha)\big) \underset{n}{\longrightarrow} \alpha \quad \forall\, \gamma \in \mathcal{G}_0$$

for every $\alpha \in (0, 1)$. Thus, the sequence of tests corresponding to $(T_n)_{n\in\mathbb{N}}$ is asymptotically of level α, and $k(\alpha)$ is the asymptotic critical value.

The following definition is due to Bahadur, cf. page 276 in Bahadur [1].

Definition 2.1
The sequence $(T_n)_{n\in\mathbb{N}}$ is said to be a *standard sequence* if the following conditions are satisfied.

(BI) For each $\gamma \in \mathcal{G}_0$,

$$P_\gamma(T_n \le x) \underset{n}{\longrightarrow} G(x) \quad \forall\, x \in \mathbb{R},$$

where G is a continuous distribution function.

(BII) There is a constant $a > 0$ such that

$$\lim_{x\to\infty} \frac{\log\big(1 - G(x)\big)}{x^2} = -\frac{a}{2}.$$

(BIII) There is a function $b\colon \mathcal{G} \setminus \mathcal{G}_0 \to (0, \infty)$ with

$$\frac{T_n}{\sqrt{n}} - b(\gamma) \underset{n}{\longrightarrow} 0 \quad \text{in } P_\gamma\text{-probability} \quad \forall\, \gamma \in \mathcal{G} \setminus \mathcal{G}_0.$$

For a standard sequence $(T_n)_{n\in\mathbb{N}}$ we set $k(\alpha) = G^{-1}(1-\alpha)$ because of (BI), where G^{-1} is the quantile function of G. Condition (BIII) implies that $T_n \to \infty$ in probability under H_1, so that the sequence of tests corresponding to $(T_n)_{n\in\mathbb{N}}$ is consistent.

In [1] Bahadur studies the behavior of $1 - G(T_n)$, the *approximate p-value* or *approximate level attained by* T_n, for any standard sequence $(T_n)_{n\in\mathbb{N}}$. He considers the random variable

$$K_n := -2\log\bigl(1 - G(T_n)\bigr) \tag{2.3}$$

and shows that for each $\gamma \in \mathcal{G}_0$

$$\lim_{n\to\infty} P_\gamma(K_n \le x) = F_{\chi_2^2}(x) \ \ \forall\, x \in \mathbb{R},$$

where $F_{\chi_2^2}$ is the distribution function of the chi-square distribution with two degrees of freedom. Moreover, he notes that

$$\frac{K_n}{n} \quad \underset{n}{\longrightarrow} \quad a\, b(\gamma)^2 \quad \text{in } P_\gamma\text{-probability} \quad \forall\, \gamma \in \mathcal{G}\setminus\mathcal{G}_0.$$

The function $c(\gamma) := a\, b(\gamma)^2$, $\gamma \in \mathcal{G}\setminus\mathcal{G}_0$, is called the *asymptotic* or *approximate slope* of the sequence $(T_n)_{n\in\mathbb{N}}$. For two standard sequences $(T_{1n})_{n\in\mathbb{N}}$ and $(T_{2n})_{n\in\mathbb{N}}$ with approximate slopes $c_1(\gamma)$ and $c_2(\gamma)$ respectively Bahadur compares the approximate attained levels for fixed $n \in \mathbb{N}$. He argues that the test based on T_{in} is less successful than that based on T_{jn} if the approximate level attained by T_{in} exceeds that of T_{jn}, which is equivalent to $K_{in} < K_{jn}$, where K_{in} and K_{jn} are as in (2.3), $i \neq j \in \{1, 2\}$. Since

$$\frac{K_{1n}}{K_{2n}} \quad \underset{n}{\longrightarrow} \quad \frac{c_1(\gamma)}{c_2(\gamma)} \quad \text{in } P_\gamma\text{-probability} \quad \forall\, \gamma \in \mathcal{G}\setminus\mathcal{G}_0,$$

with P_γ-probability tending to one the test corresponding to T_{1n} is less successful than that corresponding to T_{2n} if $c_1(\gamma)/c_2(\gamma) < 1$ and more successful if $c_1(\gamma)/c_2(\gamma) > 1$. The ratio $c_1(\gamma)/c_2(\gamma)$ is thus called the *approximate Bahadur asymptotic relative efficiency (approximate Bahadur ARE)* of the sequence $(T_{1n})_{n\in\mathbb{N}}$ relative to the sequence $(T_{2n})_{n\in\mathbb{N}}$.

A drawback of the concept of approximate Bahadur ARE is that the approximate slope of a standard sequence is not a very trustworthy measure for the performance of the corresponding test, as Bahadur himself notes at the end of section 4 in [1], see also section 6 and 7 in Bahadur [2]. Nevertheless, the approximate Bahadur ARE has its merits. For example, it is generally easy to compute and under certain conditions its limit as the alternative approaches the hypothesis equals the limit as $\alpha \to 0$ of the Pitman asymptotic relative efficiency, a different efficiency concept which we will describe next.

The concept of Pitman asymptotic relative efficiency is based on the notion of relative efficiency of two sequences of tests. For this, let $(T_{in})_{n\in\mathbb{N}}$, $i = 1, 2$, be sequences of statistics for testing the hypothesis testing problem (2.1). The index n here denotes the size of the random sample the statistic T_{in} is based on. As before, we assume that the sequence of tests corresponding to $(T_{in})_{n\in\mathbb{N}}$ is asymptotically of level α and that $\{T_{in} > k_i(\alpha)\}$ is the rejection region of the test based on T_{in}, where $k_i(\alpha)$ is the asymptotic critical value, $i = 1, 2$. Furthermore, we assume that the test sequences based on $(T_{1n})_{n\in\mathbb{N}}$ and $(T_{2n})_{n\in\mathbb{N}}$ are consistent. For fixed $\alpha, \beta \in (0, 1)$ and $\gamma \in \mathcal{G}\setminus\mathcal{G}_0$ we define

$$N_i(\alpha, \beta, \gamma) := \min\bigl\{n \in \mathbb{N} \colon P_\gamma\bigl(T_{im} > k_i(\alpha)\bigr) \ge \beta \ \forall\, m \ge n\bigr\}, \quad i = 1, 2. \tag{2.4}$$

Note that the consistency of the respective test sequence ensures that $N_i(\alpha, \beta, \gamma) \in \mathbb{N}$ for $i = 1, 2$. The number $N_i(\alpha, \beta, \gamma)$ is the smallest sample size such that the power of the test based on

$(T_{in})_{n\in\mathbb{N}}$ under the alternative γ and the asymptotic significance level α is not less than β for all sample sizes larger or equal to it. Hence, for given α, β and γ, the sequence of tests based on $(T_{in})_{n\in\mathbb{N}}$ is preferable to the one based on $(T_{jn})_{n\in\mathbb{N}}$ if $N_i(\alpha,\beta,\gamma) < N_j(\alpha,\beta,\gamma)$, because it needs less observations to attain a power of at least β at the alternative γ and the asymptotic significance level α. We will call the ratio $N_2(\alpha,\beta,\gamma)/N_1(\alpha,\beta,\gamma)$ the *relative efficiency* of the sequence $(T_{1n})_{n\in\mathbb{N}}$ with respect to the sequence $(T_{2n})_{n\in\mathbb{N}}$.

In general, the relative efficiency depends on all three arguments α, β and γ, and its explicit computation is often very difficult. Since from a practical point of view small significance levels, high powers and alternatives close to the hypothesis are especially relevant, several limiting procedures have been proposed. One approach is to investigate the limit of the relative efficiency as the alternative tends to H_0. If

$$\lim_{\substack{\gamma\in\mathcal{G}\setminus\mathcal{G}_0,\\ d(\gamma,\mathcal{G}_0)\to 0}} \frac{N_2(\alpha,\beta,\gamma)}{N_1(\alpha,\beta,\gamma)}$$

exists, we will call it the *Pitman asymptotic relative efficiency (Pitman ARE)* of the sequence $(T_{1n})_{n\in\mathbb{N}}$ with respect to the sequence $(T_{2n})_{n\in\mathbb{N}}$. The concept of Pitman ARE was introduced by E. J. G. Pitman at the end of the 1940s in his unpublished lecture notes on nonparametric statistical inference and has since then become one of the most popular types of asymptotic relative efficiency.

In the literature, there are several variants of the notion of relative efficiency. For example, other definitions of $N_i(\alpha,\beta,\gamma)$ are used. Sometimes $N_i(\alpha,\beta,\gamma)$ is defined to be the first sample size such that the power of the test at the alternative γ and the significance level α is larger than or equal to β, without requiring the power to remain at this level for sample sizes larger than $N_i(\alpha,\beta,\gamma)$. If the power is an increasing function of the sample size, this definition of $N_i(\alpha,\beta,\gamma)$ coincides of course with the one above. Note moreover that often the sequence of exact level α tests corresponding to $(T_{in})_{n\in\mathbb{N}}$ is considered. In this case, the asymptotic critical value is replaced by the exact critical value in the definition of $N_i(\alpha,\beta,\gamma)$. Since we are only interested in comparing sequences of tests as described above that are asymptotically of level α, the definition of $N_i(\alpha,\beta,\gamma)$ as given in (2.4) is the most suitable for our purposes, and we will henceforth only consider the relative efficiency and Pitman asymptotic relative efficiency as defined above. A comprehensive description of the aforementioned and other notions of asymptotic relative efficiency and related results can be found in the book of Nikitin [21].

As the Pitman ARE may depend on the values of α and β, it is in general still difficult to determine its value. Because of this, its limit as $\alpha\to 0$ is investigated. For $\mathcal{G}$ being an interval and $\mathcal{G}_0=\{\gamma_0\}$, Wieand [30] gives conditions ensuring that the limit as $\alpha\to 0$ of an extended version of Pitman asymptotic relative efficiency agrees with the limit of the approximate Bahadur asymptotic relative efficiency as the alternative γ approaches $\mathcal{G}_0$. He shows that for this equality to hold, it is sufficient to strengthen condition (BIII) locally. In what follows, we adjust Wieand's results to our definition of Pitman ARE and extend them to the general hypothesis testing problem (2.1). Another extension of Wieand's results was done by Kallenberg and Koning [16].

The following definition extends Wieand's additional Condition III*.

Definition 2.2
The sequence $(T_n)_{n\in\mathbb{N}}$ is said to fulfill *Wieand's condition* (WIII) if there exists a function $b\colon \mathcal{G}\setminus\mathcal{G}_0\to(0,\infty)$ so that there is an $\epsilon^*>0$ such that for each $\epsilon>0$ and $\delta\in(0,\frac{1}{2})$ there is a positive constant $C(\epsilon,\delta)$ with

$$P_\gamma\Big(\Big|\frac{T_n}{\sqrt{n}}-b(\gamma)\Big|\geq\epsilon\, b(\gamma)\Big) \;<\; \delta$$

for all $\gamma\in U_{\epsilon^*}(\mathcal{G}_0)\setminus\mathcal{G}_0$ and for all $n\in\mathbb{N}$ with $\sqrt{n}\,b(\gamma)>C(\epsilon,\delta)$.

Note that condition (WIII) implies (BIII) locally, i.e., for all $\gamma \in U_{\epsilon^*}(\mathcal{G}_0)\setminus\mathcal{G}_0$. Hence, the function b in Wieand's condition is locally unique, that is, if the sequence $(T_n)_{n\in\mathbb{N}}$ satisfies (WIII) with two functions b_1 and b_2, there is a $\varrho > 0$ such that $b_1(\gamma) = b_2(\gamma)$ for all $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$.

We will now state and prove a version of the theorem on page 1005 in Wieand [30] that is adjusted to our setting. For this, let us consider two sequences $(T_{1n})_{n\in\mathbb{N}}$ and $(T_{2n})_{n\in\mathbb{N}}$ of test statistics again. For functions and symbols such as G_i, a_i, b_i, c_i, the subscript i refers to the sequence $(T_{in})_{n\in\mathbb{N}}$, $i = 1, 2$.

Theorem 2.3
Let $(T_{in})_{n\in\mathbb{N}}$, $i = 1, 2$, be two sequences such that

(i) $(T_{in})_{n\in\mathbb{N}}$ *fulfills conditions* (BI), (BII) *and* (WIII) *for $i = 1, 2$,*

(ii) G_1 *and* G_2 *are strictly increasing on (z, ∞) for some $z \in \mathbb{R}$,*

(iii) $b_i(\gamma) \to 0$ *as $d(\gamma, \mathcal{G}_0) \to 0$, $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$, for $i = 1, 2$,*

(iv) there exists

$$\lim_{\substack{\gamma\in\mathcal{G}\setminus\mathcal{G}_0,\\ d(\gamma,\mathcal{G}_0)\to 0}} \frac{c_1(\gamma)}{c_2(\gamma)} =: B(\mathcal{G}_0) \in \mathbb{R}.$$

Then for all $\beta \in (0, 1)$

$$B(\mathcal{G}_0) = \lim_{\alpha\to 0} \liminf_{\substack{\gamma\in\mathcal{G}\setminus\mathcal{G}_0,\\ d(\gamma,\mathcal{G}_0)\to 0}} \frac{N_2(\alpha,\beta,\gamma)}{N_1(\alpha,\beta,\gamma)} = \lim_{\alpha\to 0} \limsup_{\substack{\gamma\in\mathcal{G}\setminus\mathcal{G}_0,\\ d(\gamma,\mathcal{G}_0)\to 0}} \frac{N_2(\alpha,\beta,\gamma)}{N_1(\alpha,\beta,\gamma)}. \tag{2.5}$$

The common value in (2.5) is called the *limiting (as $\alpha \to 0$) Pitman ARE* of the sequence $(T_{1n})_{n\in\mathbb{N}}$ with respect to the sequence $(T_{2n})_{n\in\mathbb{N}}$. Note that it is independent of β, since $B(\mathcal{G}_0)$ does not depend on it.

As from a practical point of view the performance of a test under small significance levels and alternatives close to H_0 is of special importance, the limiting (as $\alpha \to 0$) Pitman ARE is an appropriate means for choosing between the two test sequences for the hypothesis testing problem (2.1). If $B(\mathcal{G}_0) > 1$, then the sequence of tests based on $(T_{1n})_{n\in\mathbb{N}}$ is preferable to the one based on $(T_{2n})_{n\in\mathbb{N}}$, and the test sequence based on $(T_{2n})_{n\in\mathbb{N}}$ is preferable if $B(\mathcal{G}_0) < 1$.

The following proof of Theorem 2.3 is mainly based on the proof on page 1006 in Wieand [30] but also borrows some ideas from the proof of Theorem 1 in Kallenberg and Koning [16].

Proof. For any fixed $\beta \in (0, 1)$ we can choose a $\delta \in (0, \frac{1}{2})$ with $\beta \in [\delta, 1-\delta]$. For every $\epsilon \in (0, 1)$ by (BII) there is an $x_0 = x_0(\epsilon) > z$ such that

$$(1+\epsilon)\Big(-\frac{a_i}{2}\Big)x^2 \le \log(1 - G_i(x)) \le (1-\epsilon)\Big(-\frac{a_i}{2}\Big)x^2$$

for all $x \ge x_0$ and $i = 1, 2$, where z is from (ii). Let $\alpha' := \min_{i=1,2}(1 - G_i(x_0))$ (note that $\alpha' = \alpha'(\epsilon) \in (0, 1)$). Then for $\alpha \in (0, \alpha']$ it is $k_i(\alpha) = G_i^{-1}(1-\alpha) \ge x_0$ for $i = 1, 2$, and therefore

$$(1+\epsilon)\Big(-\frac{a_i}{2}\Big)k_i(\alpha)^2 \le \log(\alpha) \le (1-\epsilon)\Big(-\frac{a_i}{2}\Big)k_i(\alpha)^2,$$

which is equivalent to

$$\left[\frac{-2\log(\alpha)}{a_i(1+\epsilon)}\right]^{1/2} \;\le\; k_i(\alpha) \;\le\; \left[\frac{-2\log(\alpha)}{a_i(1-\epsilon)}\right]^{1/2}. \tag{2.6}$$

Define $\alpha'' := \min_{i=1,2} \exp(-a_i C_i(\epsilon,\delta)^2)$, where $C_i(\epsilon,\delta)$ is as in (WIII) (note that $\alpha'' = \alpha''(\epsilon,\delta)$). For $\alpha \in (0,\alpha'')$ we have for $i = 1,2$ and for all $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ that for $n \in \mathbb{N}$

$$n \geq \frac{-\log(\alpha)}{a_i b_i(\gamma)^2} \quad \Longrightarrow \quad \sqrt{n}\, b_i(\gamma) \geq \Big(\frac{-\log(\alpha)}{a_i}\Big)^{1/2} > C_i(\epsilon,\delta). \tag{2.7}$$

We will now derive a lower bound for $N_i(\alpha,\beta,\gamma)$.

For $\epsilon \in (0,1)$ set $M(\epsilon) := 2 \cdot (1-\epsilon)/(1+\epsilon)^4 - 1$ and fix $\epsilon_0 \in (0,1)$ with $M(\epsilon_0) > 0$. Then $2(1-\epsilon)/(1+\epsilon)^4 > 1$ for every $\epsilon \in (0,\epsilon_0]$ because the function $M(\epsilon)$ is strictly decreasing in $\epsilon \in (0,1)$. Further set $K := \exp(-(a_1 \vee a_2)/M(\epsilon_0))$. Note that by (iii) there is an $\tilde{\eta} > 0$ such that $b_i(\gamma) \leq 1$, $i = 1,2$, for every $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma,\mathcal{G}_0) < \tilde{\eta}$. For such γ it follows for $\alpha \in (0,K]$ that

$$\alpha \leq \exp\Big(-\frac{a_i b_i(\gamma)^2}{M(\epsilon)}\Big)$$

for $i = 1,2$ and every $0 < \epsilon \leq \epsilon_0$, so that

$$\frac{-\log(\alpha)}{a_i b_i(\gamma)^2} M(\epsilon) \geq 1. \tag{2.8}$$

Thus, there is an $n \in \mathbb{N}$ with

$$\frac{-\log(\alpha)}{a_i b_i(\gamma)^2} \leq n < \frac{-2\log(\alpha)(1-\epsilon)}{a_i b_i(\gamma)^2 (1+\epsilon)^4}, \tag{2.9}$$

because the difference of the bounds is at least one, as was shown in (2.8). So for $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma,\mathcal{G}_0) < \tilde{\eta}$, $\epsilon \in (0,\epsilon_0]$, $\alpha \in (0,\min(K,\alpha',\alpha''))$ and such n we have

$$\sqrt{n}\, b_i(\gamma)(1+\epsilon) < \left[\frac{-2\log(\alpha)(1-\epsilon)}{a_i(1+\epsilon)^2}\right]^{1/2} \leq k_i(\alpha)\Big(\frac{1-\epsilon}{1+\epsilon}\Big)^{1/2} < k_i(\alpha)$$

for $i = 1,2$ using (2.6), and therefore

$$P_\gamma\big(T_{in} > k_i(\alpha)\big) \leq P_\gamma\big(T_{in} > \sqrt{n}\, b_i(\gamma)(1+\epsilon)\big).$$

Now set $\epsilon^* := \epsilon_1^* \wedge \epsilon_2^*$ with ϵ_i^* from (WIII) and take $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma,\mathcal{G}_0) < \epsilon^* \wedge \tilde{\eta}$. Since $n \geq (-\log(\alpha))/(a_i b_i(\gamma)^2)$ and γ and α are such that Wieand's condition (WIII) holds, we have

$$\begin{aligned}
\beta \geq \delta > P_\gamma\Big(\Big|\frac{T_{in}}{\sqrt{n}\, b_i(\gamma)} - 1\Big| \geq \epsilon\Big) \\
&= P_\gamma\Big(\Big\{\frac{T_{in}}{\sqrt{n}\, b_i(\gamma)} \geq 1+\epsilon\Big\} \cup \Big\{\frac{T_{in}}{\sqrt{n}\, b_i(\gamma)} \leq 1-\epsilon\Big\}\Big) \\
&\geq P_\gamma\big(T_{in} \geq \sqrt{n}\, b_i(\gamma)(1+\epsilon)\big) \\
&\geq P_\gamma\big(T_{in} > k_i(\alpha)\big).
\end{aligned}$$

Thus, for $i = 1,2$

$$N_i(\alpha,\beta,\gamma) \geq \frac{-2\log(\alpha)(1-\epsilon)}{a_i b_i(\gamma)^2(1+\epsilon)^4} \tag{2.10}$$

for every $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma,\mathcal{G}_0) < \epsilon^* \wedge \tilde{\eta}$, $\epsilon \in (0,\epsilon_0]$ and $\alpha \in (0,\min(K,\alpha',\alpha''))$.

Next, we want to find an upper bound for $N_i(\alpha, \beta, \gamma)$.

For this, let $\epsilon \in (0, \epsilon_0]$ and $\alpha \in (0, \min(K, \alpha', \alpha''))$ again. Note that by (iii) there is an $\widehat{\eta} > 0$ with $b_i(\gamma) < \sqrt{\epsilon}\, C_i(\epsilon, \delta)$ for $\gamma \in U_{\widehat{\eta}}(\mathcal{G}_0) \setminus \mathcal{G}_0$ and $i = 1, 2$ (note that $\widehat{\eta} = \widehat{\eta}(\epsilon, \delta)$). For the following investigations let $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma, \mathcal{G}_0) < \min(\epsilon^*, \tilde{\eta}, \widehat{\eta})$.

Then for $n \in \mathbb{N}$ such that

$$n \geq \frac{-2\log(\alpha)(1+\epsilon)}{a_i b_i(\gamma)^2 (1-\epsilon)^3} > \frac{-\log(\alpha)}{a_i b_i(\gamma)^2}$$

it follows with (2.6) that for $i = 1, 2$

$$\sqrt{n}\, b_i(\gamma)(1-\epsilon) \geq \Big(\frac{-2\log(\alpha)(1+\epsilon)}{a_i(1-\epsilon)}\Big)^{1/2} \geq k_i(\alpha)\sqrt{1+\epsilon} > k_i(\alpha).$$

Hence, by the monotonicity of the distribution of T_{in} under P_γ we have

$$P_\gamma\big(T_{in} > k_i(\alpha)\big) \geq P_\gamma\big(T_{in} > \sqrt{n}\, b_i(\gamma)(1-\epsilon)\big).$$

As $\alpha < \alpha''$ and n is such that the left side of (2.7) holds, it follows from Wieand's condition (WIII) that

$$\begin{aligned} \beta \leq 1 - \delta < P_\gamma\Big(\Big|\frac{T_{in}}{\sqrt{n}\, b_i(\gamma)} - 1\Big| < \epsilon\Big) &= P_\gamma\Big(1 - \epsilon < \frac{T_{in}}{\sqrt{n}\, b_i(\gamma)} < 1 + \epsilon\Big) \\ &\leq P_\gamma\big(T_{in} > \sqrt{n}\, b_i(\gamma)(1-\epsilon)\big) \\ &\leq P_\gamma\big(T_{in} > k_i(\alpha)\big). \end{aligned}$$

This implies

$$N_i(\alpha, \beta, \gamma) \leq \Big\lceil \frac{-2\log(\alpha)(1+\epsilon)}{a_i b_i(\gamma)^2(1-\epsilon)^3} \Big\rceil \tag{2.11}$$

for $i = 1, 2$, where $\lceil y \rceil := \min\{m \in \mathbb{Z} \colon m \geq y\}$ for $y \in \mathbb{R}$. Now note that for $i = 1, 2$

$$\sqrt{N_i(\alpha, \beta, \gamma)}\, b_i(\gamma) > C_i(\epsilon, \delta)$$

using (2.10), (2.9) and (2.7). But since $b_i(\gamma) < \sqrt{\epsilon}\, C_i(\epsilon, \delta)$, this yields $N_i(\alpha, \beta, \gamma)\, \epsilon > 1$. Hence,

$$N_i(\alpha, \beta, \gamma)(1-\epsilon) = N_i(\alpha, \beta, \gamma) - N_i(\alpha, \beta, \gamma)\, \epsilon < N_i(\alpha, \beta, \gamma) - 1 < \frac{-2\log(\alpha)(1+\epsilon)}{a_i b_i(\gamma)^2(1-\epsilon)^3},$$

where the last inequality follows from (2.11). Thus,

$$N_i(\alpha, \beta, \gamma) < \frac{-2\log(\alpha)(1+\epsilon)}{a_i b_i(\gamma)^2(1-\epsilon)^4} \tag{2.12}$$

for every $\epsilon \in (0, \epsilon_0]$, $\alpha \in (0, \min(K, \alpha', \alpha''))$, $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma, \mathcal{G}_0) < \min(\epsilon^*, \tilde{\eta}, \widehat{\eta})$ and $i = 1, 2$.

A combination of (2.10) and (2.12) yields

$$\frac{c_1(\gamma)}{c_2(\gamma)} \cdot \Big(\frac{1-\epsilon}{1+\epsilon}\Big)^5 < \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} < \frac{c_1(\gamma)}{c_2(\gamma)} \cdot \Big(\frac{1+\epsilon}{1-\epsilon}\Big)^5 \tag{2.13}$$

for every $\epsilon \in (0, \epsilon_0]$, $\alpha \in (0, \min(K, \alpha', \alpha''))$ and $\gamma \in \mathcal{G} \setminus \mathcal{G}_0$ with $d(\gamma, \mathcal{G}_0) < \min(\epsilon^*, \tilde{\eta}, \widehat{\eta})$, whence it follows that

$$\limsup_{\substack{\gamma \in \mathcal{G}\setminus\mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} \leq B(\mathcal{G}_0) \cdot \Big(\frac{1+\epsilon}{1-\epsilon}\Big)^5$$

for these values of ϵ and α. Now taking the limit superior as $\alpha \to 0$ of both sides of this inequality first and letting ϵ tend to zero afterward, we get

$$\limsup_{\alpha \to 0} \limsup_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} \leq B(\mathcal{G}_0).$$

In the same way it follows from (2.13) that

$$\liminf_{\alpha \to 0} \liminf_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} \geq B(\mathcal{G}_0).$$

Hence, it is

$$\begin{aligned} B(\mathcal{G}_0) &\leq \liminf_{\alpha \to 0} \liminf_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} \leq \limsup_{\alpha \to 0} \liminf_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} \\ &\leq \limsup_{\alpha \to 0} \limsup_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} \leq B(\mathcal{G}_0), \end{aligned}$$

and this implies

$$\lim_{\alpha \to 0} \liminf_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} = B(\mathcal{G}_0).$$

Analogously, we get

$$\lim_{\alpha \to 0} \limsup_{\substack{\gamma \in \mathcal{G} \setminus \mathcal{G}_0, \\ d(\gamma, \mathcal{G}_0) \to 0}} \frac{N_2(\alpha, \beta, \gamma)}{N_1(\alpha, \beta, \gamma)} = B(\mathcal{G}_0),$$

which completes the proof. □

Oftentimes the verification of Wieand's condition (WIII) is not straightforward, because in order to establish it, it is necessary to study the behavior of the test statistics under H_1, and the knowledge of this behavior is often limited. The following proposition thus sometimes facilitates the verification of (WIII). It extends the lemma on page 1007 in Wieand [30] to composite null hypotheses in an arbitrary metric space.

Proposition 2.4
Let $\{(V_{n,\gamma})_{n \in \mathbb{N}} \colon \gamma \in \mathcal{G}\}$ be a family of sequences of real-valued test statistics on $(\Omega, \mathcal{A})$. Suppose that there is a $\varrho > 0$ such that

(i) for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ there is a continuous distribution function Q_γ with

$$\sup_{\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0} \left| P_\gamma(V_{n,\gamma} \leq x) - Q_\gamma(x) \right| \underset{n}{\longrightarrow} 0 \quad \forall\, x \in \mathbb{R},$$

(ii) $\displaystyle\sup_{\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0} |Q_\gamma^{-1}(\alpha)| < \infty$ *for all $\alpha \in (0,1)$, where Q_γ^{-1} is the quantile function of Q_γ.*

Let $g \colon U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0 \to (0,1]$ be an arbitrary function. Then for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a constant $C = C(\epsilon, \delta)$ such that for all $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ and all $n \in \mathbb{N}$ with $n > C/g(\gamma)^2$

$$P_\gamma\Big(\Big|\frac{V_{n,\gamma}}{\sqrt{n}}\Big| \leq \epsilon \cdot g(\gamma)\Big) > 1 - \delta.$$

Proof. Let $\epsilon > 0$ and $\delta \in (0,1)$. Choose $M_1 \in (0,\infty)$ so that

$$\frac{1}{\epsilon} \cdot \sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} Q_\gamma^{-1}\Big(1 - \frac{\delta}{4}\Big) < M_1.$$

Then $Q_\gamma(\epsilon M_1) \geq 1 - \frac{\delta}{4}$ for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$. Moreover, choose a constant $C_1 \geq M_1^2$ such that $n > C_1$ implies

$$\sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} \big|P_\gamma(V_{n,\gamma} \leq \epsilon M_1) - Q_\gamma(\epsilon M_1)\big| < \frac{\delta}{4}.$$

Now $0 < g \leq 1$ implies $C_1/g^2 \geq C_1$ and thus it follows that for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ and $n \in \mathbb{N}$ with $n > C_1/g(\gamma)^2$ we have

$$P_\gamma(V_{n,\gamma} \leq \epsilon M_1) > Q_\gamma(\epsilon M_1) - \frac{\delta}{4} \geq 1 - \frac{\delta}{2}.$$

Because of $M_1^2 \leq C_1 < ng(\gamma)^2$ this implies

$$P_\gamma\big(V_{n,\gamma} \leq \epsilon\sqrt{n}g(\gamma)\big) > 1 - \frac{\delta}{2}$$

for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ and $n \in \mathbb{N}$ with $n > C_1/g(\gamma)^2$.

Next, choose $M_2 \in (0,\infty)$ such that

$$\Big(-\frac{1}{\epsilon}\Big) \cdot \inf_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} Q_\gamma^{-1}\Big(\frac{\delta}{4}\Big) < M_2.$$

Then $Q_\gamma(-\epsilon M_2) \leq \frac{\delta}{4}$ for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$. Let $C_2 \geq M_2^2$ such that $n > C_2$ implies

$$\begin{aligned}
&\sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} \big|P_\gamma(V_{n,\gamma} > -\epsilon M_2) - (1 - Q_\gamma(-\epsilon M_2))\big| \\
&= \sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} \big|P_\gamma(V_{n,\gamma} \leq -\epsilon M_2) - Q_\gamma(-\epsilon M_2)\big| < \frac{\delta}{4}.
\end{aligned}$$

For all $n \in \mathbb{N}$ with $n > C_2$ and all $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ we then have

$$P_\gamma(V_{n,\gamma} \geq -\epsilon M_2) \geq P_\gamma(V_{n,\gamma} > -\epsilon M_2) > 1 - Q_\gamma(-\epsilon M_2) - \frac{\delta}{4} \geq 1 - \frac{\delta}{2}.$$

As above, $C_2/g^2 \geq C_2$ because of $0 < g \leq 1$. Thus, for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ and all $n \in \mathbb{N}$ with $n > C_2/g(\gamma)^2$ it is

$$P_\gamma\big(V_{n,\gamma} \geq -\epsilon\sqrt{n}g(\gamma)\big) > 1 - \frac{\delta}{2}$$

because $M_2^2 \leq C_2 < ng(\gamma)^2$.

Combining these results, with $C := \max(C_1, C_2)$ we have for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ and all $n \in \mathbb{N}$ such that $n > C/g(\gamma)^2$

$$\begin{aligned}
P_\gamma\Big(\Big|\frac{V_{n,\gamma}}{\sqrt{n}}\Big| \leq \epsilon \cdot g(\gamma)\Big) &= P_\gamma\big(V_{n,\gamma} \leq \sqrt{n}\epsilon g(\gamma)\big) + P_\gamma\big(V_{n,\gamma} \geq -\sqrt{n}\epsilon g(\gamma)\big) \\
&\quad - P_\gamma\big(\{V_{n,\gamma} \leq \sqrt{n}\epsilon g(\gamma)\} \cup \{V_{n,\gamma} \geq -\sqrt{n}\epsilon g(\gamma)\}\big) \\
&> 1 - \delta.
\end{aligned}$$

□

Obviously, Proposition 2.4 can be extended to a finite sum of test statistics.

Corollary 2.5
For fixed $K \in \mathbb{N}$, let $(V_{n,\gamma}^{(1)})_{n\in\mathbb{N}}, \dots, (V_{n,\gamma}^{(K)})_{n\in\mathbb{N}}$ be sequences of test statistics, each fulfilling the assumptions of Proposition 2.4. Then there is a $\varrho > 0$ such that for an arbitrary function $g\colon U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0 \to (0,1]$ and for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a constant $C = C(\epsilon, \delta)$ with

$$P_\gamma\Big(\Big|\sum_{j=1}^{K} \frac{V_{n,\gamma}^{(j)}}{\sqrt{n}}\Big| \le \epsilon \cdot g(\gamma)\Big) > 1 - \delta$$

for all $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ and all $n \in \mathbb{N}$ with $n > C/g(\gamma)^2$.

The next result states conditions under which it is possible to obtain convergence in distribution uniformly in γ as required in assumption (i) of Proposition 2.4 if every element of the sequence of test statistics can be decomposed in a main term and a remainder term that converges to zero in probability.

Proposition 2.6
Let $\{(V_{n,\gamma})_{n\in\mathbb{N}}\colon \gamma \in \mathcal{G}\}$ and $\{(R_{n,\gamma})_{n\in\mathbb{N}}\colon \gamma \in \mathcal{G}\}$ be families of sequences of real-valued measurable functions on $(\Omega, \mathcal{A})$. Suppose there is a $\varrho > 0$ such that

(i) for every $\gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0$ there is a continuous distribution function Q_γ with

$$\sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} \big|P_\gamma(V_{n,\gamma} \le x) - Q_\gamma(x)\big| \underset{n}{\longrightarrow} 0 \quad \forall\, x \in \mathbb{R},$$

(ii) the family $\{Q_\gamma\colon \gamma \in U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0\}$ is pointwise equicontinuous, i.e., for every $x \in \mathbb{R}$ and $\epsilon > 0$ there is a $\delta = \delta(x,\epsilon) > 0$ with

$$\sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} |Q_\gamma(x) - Q_\gamma(y)| \le \epsilon \quad \text{for all } y \in \mathbb{R} \text{ with } |x-y| \le \delta,$$

(iii) $$\sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} P_\gamma(|R_{n,\gamma}| > \epsilon) \underset{n}{\longrightarrow} 0 \quad \forall\, \epsilon > 0.$$

Then

$$\sup_{\gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} \big|P_\gamma(V_{n,\gamma} + R_{n,\gamma} \le x) - Q_\gamma(x)\big| \underset{n}{\longrightarrow} 0 \quad \forall\, x \in \mathbb{R}.$$

Proof. For simplicity of notation, set $U_\varrho(\mathcal{G}_0) \setminus \mathcal{G}_0 =: M$. For every constant $c > 0$, every $x \in \mathbb{R}$ and $\gamma \in M$ we have

$$\begin{aligned}
&P_\gamma(V_{n,\gamma} + R_{n,\gamma} \le x) - Q_\gamma(x) \\
&\le P_\gamma\big(V_{n,\gamma} + R_{n,\gamma} \le x, |R_{n,\gamma}| \le c\big) + P_\gamma(|R_{n,\gamma}| > c) - Q_\gamma(x) \\
&\le P_\gamma(V_{n,\gamma} \le x + c) - Q_\gamma(x+c) + Q_\gamma(x+c) + P_\gamma(|R_{n,\gamma}| > c) - Q_\gamma(x) \\
&\le \sup_{\gamma\in M} |P_\gamma(V_{n,\gamma} \le x + c) - Q_\gamma(x+c)| + \sup_{\gamma\in M} |Q_\gamma(x+c) - Q_\gamma(x)| + \sup_{\gamma\in M} P_\gamma(|R_{n,\gamma}| > c). \quad (2.14)
\end{aligned}$$

Since

$$\{V_{n,\gamma} \le x - c\} \subset \{V_{n,\gamma} + R_{n,\gamma} \le x\} \cup \{|R_{n,\gamma}| > c\}$$

for every $x \in \mathbb{R}, \gamma \in M$ and $c > 0$, it also holds that

$$P_\gamma(V_{n,\gamma} + R_{n,\gamma} \le x) \ge P_\gamma(V_{n,\gamma} \le x - c) - P_\gamma(|R_{n,\gamma}| > c),$$

whence it follows that

$$
\begin{aligned}
& Q_\gamma(x) - P_\gamma(V_{n,\gamma} + R_{n,\gamma} \le x) \\
\le\; & Q_\gamma(x) + Q_\gamma(x-c) - Q_\gamma(x-c) - P_\gamma(V_{n,\gamma} \le x-c) + P_\gamma(|R_{n,\gamma}| > c) \\
\le\; & \sup_{\gamma\in M} |P_\gamma(V_{n,\gamma} \le x-c) - Q_\gamma(x-c)| + \sup_{\gamma\in M} |Q_\gamma(x) - Q_\gamma(x-c)| + \sup_{\gamma\in M} P_\gamma(|R_{n,\gamma}| > c). \quad (2.15)
\end{aligned}
$$

Now let $x \in \mathbb{R}$ and $\epsilon > 0$ be arbitrary, but fixed. Because of (ii), there is a $\delta = \delta(x,\epsilon) > 0$ such that

$$\sup_{\gamma\in M} |Q_\gamma(x) - Q_\gamma(y)| \;\le\; \frac{\epsilon}{3}$$

for all $y \in \mathbb{R}$ with $|y-x| \le \delta$. Keep this δ fixed for the rest of the proof. It follows from (iii) that there is an $N = N(x,\epsilon) \in \mathbb{N}$ such that

$$\sup_{\gamma\in M} P_\gamma(|R_{n,\gamma}| > \delta) \;\le\; \frac{\epsilon}{3} \quad \forall\;\; n \ge N(x,\epsilon).$$

By (i) there are $K' = K'(x,\epsilon)$, $K'' = K''(x,\epsilon) \in \mathbb{N}$ with

$$\sup_{\gamma\in M} |P_\gamma(V_{n,\gamma} \le x+\delta) - Q_\gamma(x+\delta)| \;\le\; \frac{\epsilon}{3} \quad \forall\;\; n \ge K'$$

and

$$\sup_{\gamma\in M} |P_\gamma(V_{n,\gamma} \le x-\delta) - Q_\gamma(x-\delta)| \;\le\; \frac{\epsilon}{3} \quad \forall\;\; n \ge K''.$$

Using (2.14) and (2.15) with $c = \delta$, it thus follows that for all $n \ge \max(N, K', K'')$ and all $\gamma \in M$ we have

$$|P_\gamma(V_{n,\gamma} + R_{n,\gamma} \le x) - Q_\gamma(x)| \;\le\; \epsilon. \qquad \square$$

We conclude this section with the following remarks.

Remark 2.7: Due to the monotonicity of every Q_γ the family $\{Q_\gamma \colon \gamma \in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0\}$ is pointwise equicontinuous if and only if for every $x \in \mathbb{R}$ and $\epsilon > 0$ there is a $\delta = \delta(\epsilon, x) > 0$ such that

$$\sup_{\gamma\in U_\varrho(\mathcal{G}_0)\setminus\mathcal{G}_0} \big(Q_\gamma(x+\delta) - Q_\gamma(x-\delta)\big) \;\le\; \epsilon. \qquad \blacklozenge$$

Remark 2.8: Note that all results of this section still hold true when replacing the index set $\mathbb{N}$ of the sequences of test statistics by the subset $\{n \in \mathbb{N} \colon n \ge n_0\}$ for fixed $n_0 \in \mathbb{N}$. While this is trivial for most of the results, the proof of Theorem 2.3 requires a simple modification in this case:
Because the function $M(\epsilon) = 2\cdot(1-\epsilon)/(1+\epsilon)^4 - 1$ is continuous and strictly decreasing on $[0,1]$ with $M(0) = 1$ and $M(1) = -1$, we can fix $\epsilon_0 \in (0,1)$ such that $0 < M(\epsilon_0) \le 1/n_0$. By (2.8) this ensures that the lower bound in (2.9) is larger than or equal to n_0. The rest of the proof remains unchanged, except of the substitution of the index set $\mathbb{N}$ by $\{n \in \mathbb{N} \colon n \ge n_0\}$, of course. $\blacklozenge$

3 Preparatory results for independent and identically distributed centered random variables

In this section we will present the definition of the centered empirical distribution function $\widetilde{F}_n$ based on a sample of independent and identically distributed centered random variables. This centered empirical distribution function is an estimator of the underlying distribution function that takes the additional information about the mean into account. Moreover, we will investigate the stochastic behavior of $\widetilde{F}_n$ and its components uniformly with respect to the distribution of the data. These uniform results will be used in the next section to verify Wieand's condition (WIII) for the Cramér-von Mises statistics based on $\widetilde{F}_n$.

3.1 The centered empirical distribution function

Let $(\Omega, \mathcal{A}, P)$ be a probability space and $(X_i)_{i\in\mathbb{N}}$ a sequence of random variables on it such that $X_1, X_2, \dots$ are independent and identically distributed according to a distribution function F with

$$\int_{\mathbb{R}} xF(dx) = 0 \quad \text{and} \quad 0 < \int_{\mathbb{R}} x^2 F(dx) < \infty. \tag{3.1}$$

For every $n \in \mathbb{N}$, $n \geq 2$, set

$$\Omega_n := \{\min_{1\leq i\leq n} X_i < 0 < \max_{1\leq i\leq n} X_i\} \in \mathcal{A}.$$

On Ω_n by Lemma A.1 there is a unique $t_n = t(X_1, \dots, X_n) \in \mathbb{R}$ with

$$\Big(\frac{1}{n} - 1\Big)\frac{1}{\max\limits_{1\leq i\leq n} X_i} < t_n < \Big(\frac{1}{n} - 1\Big)\frac{1}{\min\limits_{1\leq i\leq n} X_i} \tag{3.2}$$

and

$$\sum_{i=1}^{n} \frac{X_i}{1 + t_n X_i} = 0. \tag{3.3}$$

It follows from Lemma A.2 that for every $n \geq 2$ the function

$$t_n \colon \Omega_n \ni \omega \mapsto t(X_1(\omega), \dots, X_n(\omega)) \in \mathbb{R}$$

is $\Omega_n \cap \mathcal{A}, \mathcal{B}^*$-measurable, where $\Omega_n \cap \mathcal{A}$ is the trace σ-algebra of $\mathcal{A}$ on Ω_n, and $\mathcal{B}^*$ denotes the Borel σ-Algebra on $\mathbb{R}$. In order to extend t_n to a measurable function on Ω, we have to define it measurably on $\overline{\Omega}_n := \Omega \setminus \Omega_n$. But the set $\overline{\Omega}_n$ is asymptotically negligible in the following sense: Under the moment conditions (3.1) it is

$$P(\overline{\Omega}_n) = P\Big(0 \notin (\min_{1\leq i\leq n} X_i, \max_{1\leq i\leq n} X_i)\Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{3.4}$$

To verify this, note that

$$\begin{aligned} P\Big(0 \notin (\min_{1\leq i\leq n} X_i, \max_{1\leq i\leq n} X_i)\Big) &\leq P\big(X_i \geq 0,\ i = 1, \dots, n\big) + P\big(X_i \leq 0,\ i = 1, \dots, n\big) \\ &= P(X_1 \geq 0)^n + P(X_1 \leq 0)^n \underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

because under (3.1) neither $X_1 \geq 0$ almost everywhere (a.e.) nor $X_1 \leq 0$ a.e. is possible.

It follows from this that for asymptotic considerations such as the investigation of convergence in probability and convergence in distribution, the definition of t_n on $\overline{\Omega}_n$ is irrelevant, we can let t_n

be any $\overline{\Omega}_n \cap \mathcal{A}, \mathcal{B}^*$-measurable function on $\overline{\Omega}_n$. For this reason we will not specify the definition of t_n on $\overline{\Omega}_n$, and we will assume henceforth that Ω_n holds whenever investigating t_n or functions thereof. Then t_n is well-defined through (3.2) and (3.3).

For every $n \geq 2$ set

$$p_{ni} := \frac{1}{n(1+t_n X_i)}, \quad 1 \leq i \leq n,$$

and

$$\widetilde{F}_n(x) := \sum_{i=1}^{n} p_{ni} 1_{\{X_i \leq x\}} = \sum_{i=1}^{n} \frac{1}{n(1+t_n X_i)} 1_{\{X_i \leq x\}}, \quad x \in \mathbb{R}.$$

Then $p_{ni} > 0$, $i = 1, \ldots, n$, as shown in the proof of Lemma A.1, and

$$\sum_{i=1}^{n} p_{ni} = \sum_{i=1}^{n} \frac{1+t_n X_i}{n(1+t_n X_i)} - \frac{t_n}{n} \sum_{i=1}^{n} \frac{X_i}{1+t_n X_i} = 1$$

by (3.3). Hence, $\widetilde{F}_n$ is a discrete distribution function that puts random mass p_{ni} on each data point X_i. Moreover,

$$\int_{\mathbb{R}} x \widetilde{F}_n(dx) = \sum_{i=1}^{n} p_{ni} X_i = 0$$

because of (3.3), so that $\widetilde{F}_n$ is centered. Thus, if F is assumed to satisfy (3.1) but to be otherwise unknown, $\widetilde{F}_n$ can be used as an estimator for F that takes the additional information about the mean into account. We will call $\widetilde{F}_n$ the *centered empirical distribution function* of $X_1, \ldots, X_n$.

The function $\widetilde{F}_n$ can also be derived by an empirical likelihood approach as developed by Owen [22–24], see also Owen [25] for a comprehensive account. Using ideas from this concept for the nonparametric estimation of distribution functions under auxiliary information, Qin and Lawless [27] gave a closed-form expression of the nonparametric maximum likelihood estimator (MLE) $\hat{F}_n$ for the underlying but unknown F in the presence of some auxiliary information about F, but in a more general setting than considered here. Zhang [31] studied some asymptotic properties of this $\hat{F}_n$. The function $\widetilde{F}_n$ as defined above is just the nonparametric MLE $\hat{F}_n$ in the special case that the additional information we have about F is $\int_{\mathbb{R}} xF(dx) = 0$.

Note that in contrast to $\widetilde{F}_n$, the classical empirical distribution function F_n of $X_1, \ldots, X_n$, i.e.,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq x\}}, \quad x \in \mathbb{R},$$

which is well known to be the nonparametric MLE for F in the absence of additional information, does not incorporate the additional information that the true distribution is centered, since

$$\int_{\mathbb{R}} x F_n(dx) = \frac{1}{n} \sum_{i=1}^{n} X_i \neq 0$$

in general.

3.2 Uniform asymptotic results

Consider now a nonempty set M of distribution functions $F\colon \mathbb{R} \to [0,1]$. Let $(\Omega, \mathcal{A})$ be a measurable space and $F \mapsto P_F$ be an injective mapping from M into the set of probability measures on $\mathcal{A}$.

Definition 3.1
Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers, and for every $F \in M$ let $(Y_{n,F})_{n\in\mathbb{N}}$ be a sequence of random variables on $(\Omega, \mathcal{A})$. We say that $Y_{n,F} = o_P^u(a_n)$ *in* M *as* $n \to \infty$ if and only if $Y_{n,F}/a_n$ converges to zero in P_F-probability uniformly in M, i.e., if and only if

$$\sup_{F\in M} P_F\Big(\Big|\frac{Y_{n,F}}{a_n}\Big| \geq \epsilon\Big) \underset{n\to\infty}{\longrightarrow} 0 \quad \forall\, \epsilon > 0.$$

We say that $Y_{n,F} = O_P^u(a_n)$ *in* M *as* $n \to \infty$ if and only if $Y_{n,F}/a_n$ is stochastically bounded with respect to P_F uniformly in M, i.e., if and only if

$$\lim_{K\to\infty} \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\Big|\frac{Y_{n,F}}{a_n}\Big| \geq K\Big) = 0.$$

Of course this definition covers the special case that the sequence of variables $(Y_{n,F})_{n\in\mathbb{N}}$ is the same for every $F \in M$, i.e., $(Y_{n,F})_{n\in\mathbb{N}} \equiv (Y_n)_{n\in\mathbb{N}}$, say, for every $F \in M$.

As in the usual case, the following rules apply, where the convergence of every term is understood to be uniform in the same M as $n \to \infty$:

$$o_P^u(1) + o_P^u(1) = o_P^u(1),\ o_P^u(1)\cdot o_P^u(1) = o_P^u(1),\ o_P^u(1)\cdot O_P^u(1) = o_P^u(1),\ O_P^u(1)\cdot O_P^u(1) = O_P^u(1).$$

For the rest of this subsection, let M now be a set of continuous distribution functions having zero mean and finite variance. It follows from these assumptions that the variance of every $F \in M$ is strictly positive. Moreover, we assume that the model $(\Omega, \mathcal{A}, \{P_F\colon F \in M\})$ is such that there is a sequence $(X_i)_{i\in\mathbb{N}}$ of random variables on $(\Omega, \mathcal{A})$ such that under P_F the X_i are independent and identically distributed with common distribution function F.

Note that for a given set M such a model always exists, e.g., we can always use the infinite product measure space $(\mathbb{R}^{\mathbb{N}}, \bigotimes_{i\in\mathbb{N}} \mathcal{B}^*, \bigotimes_{i\in\mathbb{N}} Q_F) =: (\Omega, \mathcal{A}, P_F)$, where Q_F denotes the probability measure on the Borel σ-algebra $\mathcal{B}^*$ induced by F, and let X_i be the i-th coordinate projection.

Here and in the following, the subscript F in functionals such as the expectation E_F and the variance Var_F signifies that the respective term is understood to be with respect to the measure P_F. Note that the above assumptions imply that $\mathrm{E}_F(X_i) = \mathrm{E}_F(X_1) = 0$ and $\mathrm{Var}_F(X_i) = \mathrm{Var}_F(X_1) = \mathrm{E}_F(X_1^2) =: \sigma_F^2 \in (0, \infty)$ for every $i \in \mathbb{N}$ and $F \in M$.

We will now examine the asymptotic stochastic behavior of the centered empirical distribution function $\widetilde{F}_n$ of $X_1, \ldots, X_n$ and of its components uniformly in $F \in M$. If M is a singleton, i.e., if the distribution of the X_i is fixed, this has already been studied for example by Owen [23], Qin and Lawless [27], and Zhang [31] in a more general setting than considered here. Based on these works, we will investigate in the following under which assumptions about M certain results concerning the stochastic behavior of $\widetilde{F}_n$ and its components hold uniformly in M if it contains arbitrarily many elements. For these investigations we introduce the following collection of conditions:

$$\inf_{F\in M} \int_{\mathbb{R}} x^2 F(dx) > 0, \tag{3.5}$$

$$\sup_{F\in M} \int_{\mathbb{R}} x^2 F(dx) < \infty, \tag{3.6}$$

$$g(c) := \sup_{F\in M} \int_{\{x\in\mathbb{R}:\,|x|>c\}} x^2 F(dx) \to 0 \quad \text{for} \quad c\to\infty, \tag{3.7}$$

$$\inf_{F\in M} \int_{\mathbb{R}} |x| F(dx) > 0, \tag{3.8}$$

$$\sup_{F\in M} \int_{\mathbb{R}} |x| F(dx) < \infty. \tag{3.9}$$

Note that the following implications hold

$$(3.7) \implies (3.6) \implies (3.9) \quad \text{and} \quad (3.8) \implies (3.5).$$

Observe moreover that if M is a singleton, i.e., $M = \{F\}$ with a centered continuous distribution function F that has finite second moment, then M obviously satisfies conditions (3.7) and (3.8).

Lemma 3.2
Assume the set M is such that (3.7) *holds. Then*

(i) $\max_{1\le i\le n} |X_i| = o_P^u(\sqrt{n})$ *in M as $n\to\infty$,*

(ii) $\sum_{i=1}^n X_i = O_P^u(\sqrt{n})$ *in M as $n\to\infty$,*

(iii) $\frac{1}{n}\sum_{i=1}^n X_i^2 - \sigma_F^2 = o_P^u(1)$ *in M as $n\to\infty$.*

Proof. Keep in mind that (3.7) implies (3.6).

First, we show (i). For every $F\in M$ and $\epsilon>0$ it is

$$\begin{aligned} P_F\Big(\frac{1}{\sqrt{n}}\max_{1\le i\le n}|X_i|>\epsilon\Big) &= P_F\Big(\bigcup_{i=1}^n\{|X_i|>\sqrt{n}\epsilon\}\Big) \;\le\; nP_F\big(|X_1|>\sqrt{n}\epsilon\big) \\ &= n\int_{\{x\in\mathbb{R}:\,|x|>\sqrt{n}\epsilon\}} F(dx) \;\le\; \frac{1}{\epsilon^2}\int_{\{x\in\mathbb{R}:\,|x|>\sqrt{n}\epsilon\}} x^2F(dx) \;\le\; \frac{1}{\epsilon^2}\, g(\sqrt{n}\epsilon), \end{aligned}$$

and the right-hand side of the last inequality does not depend on F and converges to zero as n tends to infinity because of (3.7).

Next, we prove (ii). By using Markov's inequality we see that for every $F\in M$ and $K>0$

$$P_F\Big(\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\Big|>K\Big) \le \frac{1}{K^2}\,\mathrm{E}_F\Big(\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\Big|^2\Big) = \frac{\mathrm{E}_F(X_1^2)}{K^2} \le \frac{1}{K^2}\sup_{F\in M}\mathrm{E}_F(X_1^2),$$

and the supremum of the second moments is finite because of (3.6). Therefore

$$\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\Big|>K\Big) \;\le\; \frac{1}{K^2}\sup_{F\in M}\mathrm{E}_F(X_1^2) \underset{K\to\infty}{\longrightarrow} 0.$$

The proof of (iii) is based on ideas from the proof of the Kolmogorov-Feller weak law of large numbers for independent and identically distributed random variables without finite mean, see for example section VII.7 in Feller [9]. Define new random variables Z_i by truncating X_i^2 at an arbitrary, but fixed, level $b>0$, i.e.,

$$Z_i := X_i^2\cdot 1_{\{X_i^2\le b\}} \quad \forall\, i\in\mathbb{N}.$$

Then we have for all $F \in M$ and $y > 0$ that

$$\begin{aligned}
P_F\Big(\Big|\sum_{i=1}^n X_i^2 - \sum_{i=1}^n \mathrm{E}_F(Z_i)\Big| > y\Big) &\le P_F\Big(\Big|\sum_{i=1}^n Z_i - \sum_{i=1}^n \mathrm{E}_F(Z_i)\Big| > y\Big) + P_F\Big(\sum_{i=1}^n X_i^2 \neq \sum_{i=1}^n Z_i\Big) \\
&\le \frac{1}{y^2}\operatorname{Var}_F\Big(\sum_{i=1}^n Z_i\Big) + \sum_{i=1}^n P_F(X_i^2 \neq Z_i) \\
&\le \frac{n}{y^2}\mathrm{E}_F(Z_1^2) + nP_F(X_1^2 \neq Z_1),
\end{aligned}$$

where the second-to-last inequality follows from Chebychev's inequality and the fact that $\{\sum_{i=1}^n X_i^2 \neq \sum_{i=1}^n Z_i\} \subset \bigcup_{i=1}^n \{X_i^2 \neq Z_i\}$. In the last inequality the Bienaymé formula and the fact that the variance is bounded by the second moment were used. But

$$P_F(X_1^2 \neq Z_1) = P_F(X_1^2 1_{\{X_1^2 > b\}} \neq 0) = P_F(X_1^2 > b).$$

Since $Z_1 \ge 0$, we have

$$\begin{aligned}
\mathrm{E}_F(Z_1^2) = 2 \cdot \int_0^\infty xP_F(Z_1 > x)dx &= 2 \cdot \int_0^\infty xP_F\big(X_1^2 1_{\{X_1^2 \le b\}} > x\big)dx \\
&= 2 \cdot \int_0^b xP_F\big(X_1^2 1_{\{X_1^2 \le b\}} > x\big)dx \\
&\le 2 \cdot \int_0^b xP_F\big(X_1^2 > x\big)dx.
\end{aligned}$$

Now set $b = n$ and $y = n\epsilon$ for arbitrary, but fixed, $\epsilon > 0$. Then using the above it is

$$\begin{aligned}
P_F\Big(\Big|\frac{1}{n}\sum_{i=1}^n X_i^2 - \frac{1}{n}\sum_{i=1}^n \mathrm{E}_F(Z_i)\Big| > \epsilon\Big) &\le \frac{2}{n\epsilon^2} \cdot \int_0^n xP_F\big(X_1^2 > x\big)dx + nP_F(X_1^2 > n) \\
&\le \frac{2}{n\epsilon^2} \cdot \int_0^n x \sup_{F\in M} P_F\big(X_1^2 > x\big)dx + n \sup_{F\in M} P_F(X_1^2 > n)
\end{aligned}$$

for every $F \in M$. But

$$x \cdot \sup_{F\in M} P_F(X_1^2 > x) = \sup_{F\in M} \int_{\{y\in\mathbb{R}:\, y^2 > x\}} xF(dy) \le \sup_{F\in M} \int_{\{y\in\mathbb{R}:\, |y| > \sqrt{x}\}} y^2 F(dy) = g(\sqrt{x})$$

for every $x > 0$, and $g(\sqrt{x}) \to 0$ as $x \to \infty$ because of (3.7). This yields

$$n \sup_{F\in M} P_F(X_1^2 > n) \le g(\sqrt{n}) \underset{n\to\infty}{\longrightarrow} 0 \quad \text{and} \quad \frac{1}{n}\int_0^n x \sup_{F\in M} P_F\big(X_1^2 > x\big)dx \underset{n\to\infty}{\longrightarrow} 0. \tag{3.10}$$

To see the latter, define $f(x) := x \cdot \sup_{F\in M} P_F(X_1^2 > x)$, $x \ge 0$. Then $f \ge 0$ and $f(x) \to 0$ as x tends to infinity. Hence, for every $\tilde{\epsilon} > 0$ there is a $K(\tilde{\epsilon}) > 0$ with $f(x) \le \tilde{\epsilon}$ for all $x \ge K(\tilde{\epsilon})$. For all $n > K(\tilde{\epsilon})$ we now have

$$\frac{1}{n}\int_0^n f(x)dx = \frac{1}{n}\Big(\int_0^{K(\tilde{\epsilon})} f(x)dx + \int_{K(\tilde{\epsilon})}^n f(x)dx\Big) \le \frac{1}{n}\Big(\int_0^{K(\tilde{\epsilon})} f(x)dx + \tilde{\epsilon}\big(n - K(\tilde{\epsilon})\big)\Big).$$

Since $\int_0^{K(\tilde{\epsilon})} f(x)dx < \infty$ and $\tilde{\epsilon}$ is arbitrary, the second statement in (3.10) follows. Thus, we have

$$\sup_{F\in M} P_F\Big(\Big|\frac{1}{n}\sum_{i=1}^n X_i^2 - \frac{1}{n}\sum_{i=1}^n \mathrm{E}_F(Z_i)\Big| > \epsilon\Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{3.11}$$

Moreover, for every $F \in M$ it is

$$\Big|\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(Z_i) - \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(X_i^2)\Big| = \big|\mathrm{E}_F(Z_1) - \mathrm{E}_F(X_1^2)\big| = \mathrm{E}_F(X_1^2 \cdot 1_{\{X_1^2 > n\}}),$$

and

$$\mathrm{E}_F(X_1^2 \cdot 1_{\{X_1^2 > n\}}) = \int_{\{x\in\mathbb{R}:\, |x|>\sqrt{n}\}} x^2 F(dx) \;\le\; \sup_{F\in M}\int_{\{x\in\mathbb{R}:\, |x|>\sqrt{n}\}} x^2 F(dx) = g(\sqrt{n}).$$

Hence, (3.7) implies that

$$\sup_{F\in M}\Big|\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(Z_i) - \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(X_i^2)\Big| \underset{n\to\infty}{\longrightarrow} 0,$$

and it obviously follows from this that

$$\Big|\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(Z_i) - \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(X_i^2)\Big| = o_P^u(1) \text{ in } M \text{ as } n\to\infty. \tag{3.12}$$

A combination of (3.11) and (3.12) now yields the statement because of

$$\Big|\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \sigma_F^2\Big| \le \Big|\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(Z_i)\Big| + \Big|\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}_F(Z_i) - \sigma_F^2\Big|. \qquad \square$$

Next, we want to examine the uniform asymptotic behavior of t_n and functions thereof. Recall that t_n is defined through (3.2) and (3.3) only on the set Ω_n, and its definition on the complement $\overline{\Omega}_n$ does usually not matter for asymptotic considerations, since this set is an asymptotic P_F-nullset for every fixed $F \in M$, cf. (3.4). If we want to study the asymptotic behavior of t_n under the measure P_F uniformly in $F \in M$, however, we cannot neglect the set $\overline{\Omega}_n$ a priori, since $P_F(\overline{\Omega}_n)$ will not converge to zero uniformly in $F \in M$ in general.
There are several ways to overcome this problem. For one, we could of course explicitly define t_n on $\overline{\Omega}_n$ and then study its uniform asymptotic behavior on Ω. Here, a natural definition would certainly be to set $t_n = 0$ on $\overline{\Omega}_n$, as $\widetilde{F}_n$ would equal F_n in this case. The uniform behavior of t_n would then of course depend on the respective definition chosen on $\overline{\Omega}_n$.
Alternatively, we can impose additional conditions on the set M that ensure that $P_F(\overline{\Omega}_n)$ will converge to zero uniformly in F. Then, as before, there is no need to specify t_n on $\overline{\Omega}_n$. Since similar to the proof of (3.4) we have

$$P_F(\overline{\Omega}_n) = P_F\big(\min_{1\le i\le n} X_i \ge 0\big) + P_F\big(\max_{1\le i\le n} X_i \le 0\big) = \big(1 - F(0)\big)^n + F(0)^n,$$

the conditions

$$\inf_{F\in M} F(0) > 0 \quad \text{and} \quad \sup_{F\in M} F(0) < 1 \tag{3.13}$$

imply that $\sup_{F\in M} P_F(\overline{\Omega}_n) \to 0$ as $n \to \infty$. Hence, if M satisfies (3.13), then the set $\overline{\Omega}_n$ is irrelevant for uniform asymptotic considerations. We will therefore in the following always work under the assumption (3.13) and continue to assume that Ω_n holds for every $n \ge 2$ when studying t_n or functions thereof.

Lemma 3.3
If the set M is such that (3.5), (3.7) and (3.13) are satisfied, then
(i) $\sqrt{n}\, t_n = O_P^u(1)$ *in M as $n \to \infty$,*

(ii) $\max\limits_{1\le i\le n} \dfrac{1}{1+t_nX_i} = O_P^u(1)$ *in* M *as* $n\to\infty$,

(iii) $t_n = \dfrac{1}{\sigma_F^2}\dfrac{1}{n}\sum\limits_{i=1}^n X_i + o_P^u(1/\sqrt{n})$ *in* M *as* $n\to\infty$.

Proof. First we show (i). For every $n\ge 2$, $K>0$ and $F\in M$ it is

$$\begin{aligned} P_F(|\sqrt{n}t_n|\ge K) \;&\le\; P_F\Big(\sqrt{n}|t_n|\cdot\Big[\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2-\frac{K}{\sqrt{n}}\max_{1\le i\le n}|X_i|\Big]\ge K\Big)\\ &\quad + P_F\Big(\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2-\frac{K}{\sqrt{n}}\max_{1\le i\le n}|X_i|\le 1\Big). \end{aligned} \tag{3.14}$$

To handle the first term on the right-hand side of the above inequality, we see as in Owen [23], page 101, that

$$\frac{|t_n|}{1+|t_n|\max\limits_{1\le i\le n}|X_i|}\cdot\frac{1}{n}\sum_{i=1}^n X_i^2 \;\le\; \Big|\frac{1}{n}\sum_{i=1}^n X_i\Big|, \tag{3.15}$$

and the last term is $O_P^u(1/\sqrt{n})$ in M as $n\to\infty$ by Lemma 3.2 (ii). Now

$$\begin{aligned} &P_F\Big(\sqrt{n}|t_n|\cdot\Big[\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2-\frac{K}{\sqrt{n}}\max_{1\le i\le n}|X_i|\Big]\ge K\Big)\\ &= P_F\Big(\sqrt{n}|t_n|\cdot\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2\ge K\big(1+|t_n|\max_{1\le i\le n}|X_i|\big)\Big)\\ &= P_F\Big(\frac{\sqrt{n}|t_n|}{1+|t_n|\max\limits_{1\le i\le n}|X_i|}\cdot\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2\ge K\Big)\\ &\le P_F\Big(\frac{\sqrt{n}|t_n|}{1+|t_n|\max\limits_{1\le i\le n}|X_i|}\cdot\frac{1}{n}\sum_{i=1}^n X_i^2\ge \frac{K}{2}\inf_{F\in M}\sigma_F^2\Big). \end{aligned}$$

Using (3.5), (3.15) and Lemma 3.2 (ii), this yields

$$\lim_{K\to\infty}\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sqrt{n}|t_n|\cdot\Big[\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2-\frac{K}{\sqrt{n}}\max_{1\le i\le n}|X_i|\Big]\ge K\Big)=0.$$

It remains to investigate the second term on the right-hand side of (3.14). For simplicity of notation, set

$$L_n := \frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2 \quad\text{and}\quad M_n := \frac{K}{\sqrt{n}}\max_{1\le i\le n}|X_i|.$$

Then

$$P_F\big(L_n-M_n\le 1\big) \;\le\; P_F\Big(L_n-M_n\le 1, |M_n|<\frac{1}{2}\Big)+P_F\Big(|M_n|\ge\frac{1}{2}\Big),$$

and

$$\sup_{F\in M} P_F\Big(|M_n|\ge\frac{1}{2}\Big) = \sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\max_{1\le i\le n}|X_i|\ge\frac{1}{2K}\Big) \underset{n\to\infty}{\longrightarrow} 0$$

by Lemma 3.2 (i). Also,

$$P_F\Big(L_n-M_n\le 1, |M_n|<\frac{1}{2}\Big) \;\le\; P_F\Big(L_n\le\frac{3}{2}\Big)=P_F\Big(L_n-2\le-\frac{1}{2}\Big) \;\le\; P_F\Big(|L_n-2|\ge\frac{1}{2}\Big),$$

and

$$P_F\Big(|L_n - 2| \geq \frac{1}{2}\Big) \;\leq\; \sup_{F\in M} P_F\Big(\Big|\frac{2}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i^2 - 2\Big| \geq \frac{1}{2}\Big)$$
$$\leq\; \sup_{F\in M} P_F\Big(\Big|\frac{1}{n}\sum_{i=1}^n X_i^2 - \sigma_F^2\Big| \geq \frac{1}{4}\inf_{F\in M}\sigma_F^2\Big) \;\underset{n\to\infty}{\longrightarrow}\; 0$$

because of (3.5) and Lemma 3.2 (iii). Hence, the proof of (i) is finished.

To see that (ii) holds, note that if $|t_n|\max_{1\leq i\leq n}|X_i| \leq 1/2$, then we have for every $i \in \{1,\dots,n\}$

$$1 + t_nX_i \geq 1 - |t_n||X_i| \geq 1 - |t_n|\max_{1\leq i\leq n}|X_i| \geq 1 - \frac{1}{2} = \frac{1}{2},$$

whence it follows that

$$\max_{1\leq i\leq n}\frac{1}{1+t_nX_i} \leq 2.$$

Thus it is for every $K \in (2,\infty)$, $n \geq 2$ and every $F \in M$

$$P_F\Big(\max_{1\leq i\leq n}\frac{1}{1+t_nX_i} \geq K\Big) \leq P_F\Big(|t_n|\max_{1\leq i\leq n}|X_i| > \frac{1}{2}\Big),$$

and the statement follows from (i) and Lemma 3.2 (i).

It remains to show (iii). Using the equality $1/(1+y) = 1 - y + y^2/(1+y)$ for $y \neq -1$, we have by (3.3) for every $n \geq 2$

$$0 = \frac{1}{n}\sum_{i=1}^n\frac{X_i}{1+t_nX_i} = \frac{1}{n}\sum_{i=1}^n X_i - t_n\frac{1}{n}\sum_{i=1}^n X_i^2 + t_n^2\frac{1}{n}\sum_{i=1}^n\frac{X_i^3}{1+t_nX_i}$$
$$= \frac{1}{n}\sum_{i=1}^n X_i - t_n\frac{1}{n}\sum_{i=1}^n(X_i^2 - \sigma_F^2) - t_n\sigma_F^2 + t_n^2\frac{1}{n}\sum_{i=1}^n\frac{X_i^3}{1+t_nX_i},$$

which is equivalent to

$$t_n = \frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{\sigma_F^2}t_n\frac{1}{n}\sum_{i=1}^n(X_i^2-\sigma_F^2) + \frac{1}{\sigma_F^2}t_n^2\frac{1}{n}\sum_{i=1}^n\frac{X_i^3}{1+t_nX_i}.$$

Using (3.5), it is

$$\Big|\frac{1}{\sigma_F^2}t_n\frac{1}{n}\sum_{i=1}^n(X_i^2-\sigma_F^2)\Big| \leq \frac{1}{\inf_{F\in M}\sigma_F^2}|t_n|\Big|\frac{1}{n}\sum_{i=1}^n(X_i^2-\sigma_F^2)\Big|$$
$$= O_P^u(1/\sqrt{n})o_P^u(1) = o_P^u(1/\sqrt{n}) \text{ in } M \text{ as } n\to\infty$$

by (i) and Lemma 3.2 (iii). Because of (3.6) we obviously have $\sup_{F\in M}\sigma_F^2 = O_P^u(1)$ in M as $n \to \infty$. Therefore it follows with (3.5) and (3.6) that

$$\Big|\frac{1}{\sigma_F^2}t_n^2\frac{1}{n}\sum_{i=1}^n\frac{X_i^3}{1+t_nX_i}\Big| \leq \frac{1}{\inf_{F\in M}\sigma_F^2}|t_n|^2\max_{1\leq i\leq n}|X_i|\cdot\max_{1\leq i\leq n}\frac{1}{1+t_nX_i}\cdot\frac{1}{n}\sum_{i=1}^n X_i^2$$
$$\leq \frac{1}{\inf_{F\in M}\sigma_F^2}|t_n|^2\max_{1\leq i\leq n}|X_i|\cdot\max_{1\leq i\leq n}\frac{1}{1+t_nX_i}\cdot\Big(\frac{1}{n}\sum_{i=1}^n X_i^2 - \sigma_F^2 + \sup_{F\in M}\sigma_F^2\Big)$$
$$= O_P^u(1/n)o_P^u(\sqrt{n})O_P^u(1)O_P^u(1) = o_P^u(1/\sqrt{n}) \text{ in } M \text{ as } n\to\infty$$

because of (i), (ii) and Lemma 3.2 (i), (iii). Hence, we have shown that

$$t_n = \frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i + o_P^u(1/\sqrt{n}) \text{ in } M \text{ as } n \to \infty. \qquad \square$$

For every $n \in \mathbb{N}$ and $F \in M$ define

$$U_n(x) := \frac{1}{n}\sum_{i=1}^n X_i 1_{\{X_i \le x\}}, \quad U_F(x) := \mathrm{E}_F\big(X_1 1_{\{X_1 \le x\}}\big) = \int_{-\infty}^x y F(dy), \quad x \in \mathbb{R}.$$

Lemma 3.4
Let M be such that (3.6) *and* (3.8) *hold. Then*

$$\sup_{x\in\mathbb{R}} |U_n(x) - U_F(x)| = o_P^u(1) \quad \text{in } M \text{ as } n \to \infty.$$

Proof. Recall that (3.6) implies (3.9). With $f^+ := f \vee 0$ being the positive and $f^- := -(f \wedge 0)$ being the negative part of the function f, set

$$U_n^+(x) := \frac{1}{n}\sum_{i=1}^n X_i^+ 1_{\{X_i \le x\}}, \quad U_F^+(x) := \mathrm{E}_F\big(X_1^+ 1_{\{X_1 \le x\}}\big) = \int_{-\infty}^x y^+ F(dy),$$

for $x \in \mathbb{R}$, $n \in \mathbb{N}$ and $F \in M$, and define U_n^- and U_F^- analogously. Then $U_n = U_n^+ - U_n^-$ and $U_F = U_F^+ - U_F^-$. Obviously, $0 \le U_F^+, U_F^- \le \mathrm{E}_F(X_1^+)$ with $U_n^+(x) = U_F^+(x) = 0$ for $x \le 0$ and $U_F^-(x) = \mathrm{E}_F(X_1^-)$ for $x > 0$. An application of Lebesgue's dominated convergence theorem shows that U_F^+ and U_F^- are continuous on $\mathbb{R}$ with

$$\lim_{x\to-\infty} U_F^-(x) = 0 \quad \text{and} \quad \lim_{x\to\infty} U_F^+(x) = \mathrm{E}_F(X_1^+).$$

Therefore we extend U_F^+ and U_F^- continuously to $[-\infty, \infty]$ by defining $U_F^\pm(-\infty) := 0$ and $U_F^\pm(\infty) := \mathrm{E}_F(X_1^+) = \mathrm{E}_F(X_1^-)$. Moreover, set

$$U_n^\pm(-\infty) := 0, \quad U_n^\pm(\infty) := \frac{1}{n}\sum_{i=1}^n X_i^\pm.$$

The centeredness of F implies that $\mathrm{E}_F(|X_1|) = \mathrm{E}_F(X_1^+) + \mathrm{E}_F(X_1^-) = 2\,\mathrm{E}_F(X_1^+)$. Now define

$$a := \frac{1}{2}\sup_{F\in M} \mathrm{E}_F(|X_1|) = \sup_{F\in M} \mathrm{E}_F(X_1^+) \quad \text{and} \quad b := \frac{1}{2}\inf_{F\in M} \mathrm{E}_F(|X_1|) = \inf_{F\in M} \mathrm{E}_F(X_1^+).$$

Then $0 < b \le a < \infty$ under the assumptions.

For every $m \in \mathbb{N}$ with $m \ge 2$ and $m \ge a/b$, $0 < a/m < 2a/m < \ldots < (m-1)a/m < a$ is an equidistant partition of $[0, a]$ with mesh a/m. Since we have for every $F \in M$ that $a/m \le b \le \mathrm{E}_F(X_1^+) \le a$, it is

$$k_F := \max\{z \in \mathbb{Z} \colon z \le \frac{m}{a}\,\mathrm{E}_F(X_1^+)\} = \max\{z \in \mathbb{Z} \colon z\frac{a}{m} \le \mathrm{E}_F(X_1^+)\} \ \in \ \{1, \ldots, m\}.$$

Now fix an $F \in M$. Then either $k_F = \mathrm{E}_F(X_1^+) \cdot m/a$ or $k_F < \mathrm{E}_F(X_1^+) \cdot m/a$.

First, we consider the case $k_F \cdot a/m = \mathrm{E}_F(X_1^+)$.
Because of $U_F^+(-\infty) = 0$, $U_F^+(\infty) = \mathrm{E}_F(X_1^+) = k_F a/m$ and the continuity of U_F^+, for every $k = 1, \ldots, k_F - 1$ there is a point $x_k \in \mathbb{R}$ with $U_F^+(x_k) = k \cdot a/m$ by the intermediate value theorem, and the monotonicity of U_F^+ implies that $x_1 < \ldots < x_{k_F-1}$.

Thus, $-\infty =: x_0 < x_1 < \ldots < x_{k_F-1} < x_{k_F} := \infty$ is a partition of $[-\infty, \infty]$, so that for every $x \in \mathbb{R}$ there is exactly one $k \in \{0, \ldots, k_F - 1\}$ with $x \in [x_k, x_{k+1})$. Using the monotonicity of U_n^+ and U_F^+, this implies

$$\begin{aligned} U_n^+(x) - U_F^+(x) &\leq U_n^+(x_{k+1}) - U_F^+(x_k) = U_n^+(x_{k+1}) - k \cdot \frac{a}{m} \\ &= U_n^+(x_{k+1}) - U_F^+(x_{k+1}) + \frac{a}{m} \\ &\leq \max_{1 \leq k \leq k_F} |U_n^+(x_k) - U_F^+(x_k)| + \frac{a}{m} \end{aligned}$$

and

$$\begin{aligned} U_F^+(x) - U_n^+(x) &\leq U_F^+(x_{k+1}) - U_n^+(x_k) = U_F^+(x_k) - U_n^+(x_k) + \frac{a}{m} \\ &\leq \max_{0 \leq k \leq k_F - 1} |U_n^+(x_k) - U_F^+(x_k)| + \frac{a}{m} \\ &\leq \max_{1 \leq k \leq k_F} |U_n^+(x_k) - U_F^+(x_k)| + \frac{a}{m}. \end{aligned}$$

Together, this yields

$$\sup_{x \in \mathbb{R}} |U_n^+(x) - U_F^+(x)| \leq \max_{1 \leq k \leq k_F} |U_n^+(x_k) - U_F^+(x_k)| + \frac{a}{m}. \tag{3.16}$$

Analogously, it is

$$\sup_{x \in \mathbb{R}} |U_n^-(x) - U_F^-(x)| \leq \max_{1 \leq k \leq k_F} |U_n^-(x_k) - U_F^-(x_k)| + \frac{a}{m}. \tag{3.17}$$

Next, we investigate the case $k_F a/m < \mathrm{E}_F(X_1^+)$. Just as in the first case, it follows from the intermediate value theorem that for every $k = 1, \ldots, k_F$ there is an $x_k \in \mathbb{R}$ with $U_F^+(x_k) = k \cdot a/m$, and $-\infty =: x_0 < x_1 < \ldots < x_{k_F} < x_{k_F+1} := \infty$ partitions $[-\infty, \infty]$. Using the same monotonicity arguments as before, we see that

$$\sup_{x \in \mathbb{R}} |U_n^\pm(x) - U_F^\pm(x)| \leq \max_{1 \leq k \leq k_F+1} |U_n^\pm(x_k) - U_F^\pm(x_k)| + \frac{a}{m}.$$

Let $\epsilon > 0$. In both of the aforementioned cases we have for $x \in (-\infty, \infty]$ that

$$P_F\big(|U_n^\pm(x) - U_F^\pm(x)| > \epsilon\big) \leq \frac{1}{n\epsilon^2} \mathrm{E}_F(X_1^2) \tag{3.18}$$

by using Chebychev's inequality, the Bienaymé formula and the fact that the variance is bounded by the second moment. Now choose m so that $\epsilon > a/m$. Then in case $k_F a/m = \mathrm{E}_F(X_1^+)$ it follows from (3.16) and (3.17) that

$$\begin{aligned} P_F\big(\sup_{x \in \mathbb{R}} |U_n^\pm(x) - U_F^\pm(x)| > \epsilon\big) &\leq P_F\Big(\max_{1 \leq k \leq k_F} |U_n^\pm(x_k) - U_F^\pm(x_k)| > \epsilon - \frac{a}{m}\Big) \\ &\leq \sum_{k=1}^{k_F} P_F\Big(|U_n^\pm(x_k) - U_F^\pm(x_k)| > \epsilon - \frac{a}{m}\Big) \\ &\underset{(3.18)}{\leq} \frac{k_F}{n \cdot (\epsilon - a/m)^2} \mathrm{E}_F(X_1^2) \\ &\leq \frac{m}{n \cdot (\epsilon - a/m)^2} \cdot \sup_{F \in M} \mathrm{E}_F(X_1^2), \end{aligned}$$

where $k_F \leq m$ was used in the last inequality. For $k_F a/m < \mathrm{E}_F(X_1^+)$ it follows along the same lines that

$$P_F\big(\sup_{x\in\mathbb{R}} |U_n^{\pm}(x) - U_F^{\pm}(x)| > \epsilon\big) \;\leq\; \frac{m+1}{n\cdot(\epsilon - a/m)^2} \cdot \sup_{F\in M} \mathrm{E}_F(X_1^2).$$

Since $\sup_{F\in M} \mathrm{E}_F(X_1^2) < \infty$, we have in both cases

$$\sup_{x\in\mathbb{R}} |U_n^{\pm}(x) - U_F^{\pm}(x)| = o_P^u(1) \text{ in } M \text{ as } n\to\infty.$$

Using

$$\sup_{x\in\mathbb{R}} |U_n(x) - U_F(x)| \;\leq\; \sup_{x\in\mathbb{R}} |U_n^{+}(x) - U_F^{+}(x)| + \sup_{x\in\mathbb{R}} |U_n^{-}(x) - U_F^{-}(x)|,$$

this concludes the proof of the lemma. □

We are now ready to state and prove a uniform asymptotic expansion of $\widetilde{F}_n - F_n$. Before we do this, let us set, as usual, $\|f\|_\infty := \sup_{x\in\mathbb{R}} |f(x)|$ for any bounded function f.

Proposition 3.5
Assume the set M satisfies (3.7), (3.8) *and* (3.13). *Then*

$$\widetilde{F}_n(x) - F_n(x) = -U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i + R_{n,F}(x), \quad x\in\mathbb{R},$$

with $\|R_{n,F}\|_\infty = o_P^u(1/\sqrt{n})$ in M as $n\to\infty$.

Proof. By using again that $1/(1+y) = 1 - y + y^2/(1+y)$ for $y \neq -1$, we see that for every $x\in\mathbb{R}$ and $n\geq 2$ the following expansion of $\widetilde{F}_n$ is valid:

$$\widetilde{F}_n(x) = F_n(x) - t_n \cdot \frac{1}{n}\sum_{i=1}^n X_i \cdot 1_{\{X_i\leq x\}} + t_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{X_i^2}{1+t_nX_i}\cdot 1_{\{X_i\leq x\}}.$$

This implies that

$$\begin{aligned}
\widetilde{F}_n(x) - F_n(x) &= -t_n \cdot \frac{1}{n}\sum_{i=1}^n X_i \cdot 1_{\{X_i\leq x\}} + t_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{X_i^2}{1+t_nX_i}\cdot 1_{\{X_i\leq x\}} \\
&= -t_nU_F(x) - t_n\big(U_n(x) - U_F(x)\big) - U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i + U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i \\
&\quad + t_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{X_i^2}{1+t_nX_i}\cdot 1_{\{X_i\leq x\}} \\
&= -U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i + R_{1n,F}(x) + R_{2n}(x) + R_{3n,F}(x)
\end{aligned}$$

for every $F\in M$, where

$$R_{1n,F}(x) := t_n\big(U_F(x) - U_n(x)\big), \quad R_{2n}(x) := t_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{X_i^2}{1+t_nX_i}\cdot 1_{\{X_i\leq x\}}$$

and

$$R_{3n,F}(x) := U_F(x)\Big(\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i - t_n\Big).$$

Now

$$\begin{aligned}\sup_{x\in\mathbb{R}}|R_{1n,F}(x)| &= |t_n|\cdot\sup_{x\in\mathbb{R}}|U_F(x)-U_n(x)| \\ &= O_P^u(1/\sqrt{n})\cdot o_P^u(1) = o_P^u(1/\sqrt{n}) \text{ in } M \text{ as } n\to\infty\end{aligned}$$

by Lemma 3.3 (i) and Lemma 3.4. Moreover,

$$\begin{aligned}\sup_{x\in\mathbb{R}}|R_{2n}(x)| &\le |t_n|^2\cdot\max_{1\le i\le n}\frac{1}{1+t_nX_i}\cdot\frac{1}{n}\sum_{i=1}^n X_i^2 \\ &= O_P^u(1/n)\cdot O_P^u(1)\cdot O_P^u(1) = O_P^u(1/n) = o_P^u(1/\sqrt{n}) \text{ in } M \text{ as } n\to\infty\end{aligned}$$

by Lemma 3.3 (i), (ii) and Lemma 3.2 (iii). Next, note that it follows from the proof of Lemma 3.4 that U_F is continuous on $\mathbb{R}$. Observe moreover that U_F is monotonically decreasing on $(-\infty,0]$, monotonically increasing on $[0,\infty)$, and non-positive on $\mathbb{R}$. This and the centeredness of F imply $\sup_{x\in\mathbb{R}}|U_F(x)| = |U_F(0)| = |-\mathrm{E}_F(X_1^-)| = \mathrm{E}_F(X_1^-) = \mathrm{E}_F(X_1^+) = \frac{1}{2}\mathrm{E}_F(|X_1|)$. Therefore

$$\begin{aligned}\sup_{x\in\mathbb{R}}|R_{3n,F}(x)| &= \sup_{x\in\mathbb{R}}|U_F(x)|\cdot\Big|\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n X_i - t_n\Big| \\ &\le \frac{1}{2}\sup_{F\in M}\mathrm{E}_F(|X_1|)\cdot o_P^u(1/\sqrt{n}) = o_P^u(1/\sqrt{n}) \text{ in } M \text{ as } n\to\infty\end{aligned}$$

because of Lemma 3.3 (iii). □

The next result is a uniform central limit theorem and follows from Theorem 3 on page 441 of Eicker [8].

Lemma 3.6
Let the set M be such that (3.5) *and* (3.7) *hold. Then*

$$\sup_{F\in M}\sup_{x\in\mathbb{R}}\Big|P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{X_i}{\sigma_F}\le x\Big)-\Phi(x)\Big| \underset{n\to\infty}{\longrightarrow} 0, \tag{3.19}$$

with Φ denoting the distribution function of the standard normal distribution.

As a direct consequence of this, we get the following corollary.

Corollary 3.7
Under the assumptions of Lemma 3.6,

$$\sup_{F\in M}\sup_{x\in\mathbb{R}}\Big|P_F\Big(\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{X_i}{\sigma_F}\Big|\le x\Big)-H(x)\Big| \underset{n\to\infty}{\longrightarrow} 0,$$

where $H(x) := (2\Phi(x)-1)1_{[0,\infty)}(x)$, $x\in\mathbb{R}$, is the distribution function of the standard half-normal distribution.

Proof. For every $x\ge 0$ and $F\in M$ it is

$$\begin{aligned}&\Big|P_F\Big(\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{X_i}{\sigma_F}\Big|\le x\Big)-H(x)\Big| \\ &= \Big|P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{X_i}{\sigma_F}\le x\Big)-P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{X_i}{\sigma_F}< -x\Big)-\Phi(x)+\Phi(-x)\Big|\end{aligned}$$

$$\leq \Big|P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i}{\sigma_F} \leq x\Big) - \Phi(x)\Big| + \Big|P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i}{\sigma_F} < -x\Big) - \Phi(-x)\Big|.$$

Thus, (3.19) yields the statement. □

Lemma 3.8
Assume the set M satisfies (3.7) *and* (3.8). *Then*

(i) $\sup\limits_{F\in M}\sup\limits_{x\in\mathbb{R}}\Big|P_F\Big(\frac{\|U_F\|_\infty}{\sigma_F}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i}{\sigma_F}\Big| \leq x\Big) - Q_F(x)\Big| \underset{n\to\infty}{\longrightarrow} 0,$

where $Q_F(x) := H\big(x\cdot\sigma_F/\|U_F\|_\infty\big) = 1_{[0,\infty)}(x)\Big(2\Phi\Big(x\cdot\frac{\sigma_F}{\|U_F\|_\infty}\Big) - 1\Big),\ x\in\mathbb{R},$

(ii) *the family* $\{Q_F\colon F\in M\}$ *is uniformly equicontinuous,*

(iii) $\sup\limits_{F\in M}|Q_F^{-1}(\alpha)| < \infty$ *for all* $\alpha\in(0,1)$.

Proof. (i) Since

$$\sup_{x\in\mathbb{R}}\Big|P_F\Big(\frac{\|U_F\|_\infty}{\sigma_F}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i}{\sigma_F}\Big| \leq x\Big) - H\Big(\frac{x\cdot\sigma_F}{\|U_F\|_\infty}\Big)\Big| = \sup_{x\in\mathbb{R}}\Big|P_F\Big(\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i}{\sigma_F}\Big| \leq x\Big) - H(x)\Big|$$

for every $F\in M$, the statement follows from Corollary 3.7.

(ii) To see that $\{Q_F\colon F\in M\}$ is uniformly equicontinuous, note that Q_F has Lebesgue density

$$Q_F'(x) = 1_{[0,\infty)}(x)\cdot 2\cdot\varphi\Big(x\cdot\frac{\sigma_F}{\|U_F\|_\infty}\Big)\cdot\frac{\sigma_F}{\|U_F\|_\infty},\quad x\in\mathbb{R},$$

with $\varphi(x) = 1/\sqrt{2\pi}\cdot\exp(-x^2/2)$, so that

$$\sup_{x\in\mathbb{R}}\big|Q_F'(x)\big| \leq \sqrt{\frac{2}{\pi}}\cdot\frac{\big(\sup_{F\in M}\sigma_F^2\big)^{1/2}}{\frac{1}{2}\inf_{F\in M}\mathrm{E}_F(|X_1|)} =: K < \infty$$

for every $F\in M$, using $\|U_F\|_\infty = \frac{1}{2}\,\mathrm{E}_F(|X_1|)$ and (3.6) and (3.8). Thus, we have for $\epsilon > 0$ and $\delta = \epsilon/K$

$$|Q_F(x) - Q_F(y)| \ \leq\ K|x-y| \ <\ \epsilon$$

for every $x,y\in\mathbb{R}$ with $|x-y|<\delta$ and every $F\in M$.

(iii) For every $F\in M$ it is

$$Q_F^{-1}(\alpha) = \Phi^{-1}\Big(\frac{\alpha+1}{2}\Big)\cdot\frac{\|U_F\|_\infty}{\sigma_F} = \Phi^{-1}\Big(\frac{\alpha+1}{2}\Big)\cdot\frac{\mathrm{E}_F(|X_1|)}{2\sigma_F},\quad \alpha\in(0,1),$$

so that by (3.5) and (3.9) we have

$$\sup_{F\in M}\big|Q_F^{-1}(\alpha)\big| \ \leq\ \Phi^{-1}\Big(\frac{\alpha+1}{2}\Big)\cdot\frac{\sup_{F\in M}\mathrm{E}_F(|X_1|)}{2\big(\inf_{F\in M}\sigma_F^2\big)^{1/2}} \ <\ \infty.$$ □

4 The limiting Pitman ARE of the two tests for independent and identically distributed centered observations

Let us now consider testing the fit of a sample of independent and identically distributed data with zero mean. In the following we will examine testing both the simple null hypothesis that the underlying distribution function is equal to some fully specified centered distribution function as well as the composite null hypothesis that the true distribution function of the data belongs to a certain scale family against appropriate nonparametric alternatives. For these testing problems we will consider the respective classical Cramér-von Mises test as well as its modified counterpart where the test statistic is based on the centered empirical distribution function $\widetilde{F}_n$ instead of on F_n. In order to compare the performance of the two competing tests for the respective testing problem, we will determine in the following their limiting Pitman asymptotic relative efficiency using the results of the previous sections. To begin with, we will introduce the set of distribution functions that will be considered in the testing problems and equip it with a suitable metric.

For fixed $q \in [2, \infty)$, let $\mathcal{G}_q$ denote the set of all continuous distribution functions on the real line with finite absolute q-th moment and zero mean, i.e.,

$$\mathcal{G}_q = \Big\{ F \colon F \text{ is a continuous distribution function with} \int_{\mathbb{R}} |x|^q F(dx) < \infty \text{ and} \int_{\mathbb{R}} x F(dx) = 0 \Big\}.$$

Note that $0 < \int_{\mathbb{R}} x^2 F(dx) < \infty$ for every $F \in \mathcal{G}_q$. By setting $F(-\infty) := 0$ and $F(\infty) := 1$ for every distribution function F, it is obvious that $\mathcal{G}_q \subset C[-\infty, \infty]$, which is the space of all continuous real-valued functions on the extended real line $[-\infty, \infty]$. In the following, we will equip $\mathcal{G}_q$ with a metric d_q and derive some results concerning the metric space $(\mathcal{G}_q, d_q)$.

We will measure the distance between two elements F and G of $\mathcal{G}_q$ with the *Kantorovich-Wasserstein* or *minimal L_q metric*

$$d_q(F, G) := \inf\big\{ \mathrm{E}\big(|X - Y|^q\big)^{1/q} \colon (X, Y) \in S(F, G) \big\} \in [0, \infty), \tag{4.1}$$

where $S(F, G)$ is the collection of all pairs (X, Y) of random variables X and Y defined on the same probability space such that F and G are the distribution functions of X and Y, respectively. The function d_2 is also known as *Mallows metric.* The following properties of d_q hold for $1 \le q < \infty$ and can be found for example in the mathematical appendix of Bickel and Freedman [4]: The function d_q is a metric on the set of all distribution functions on $\mathbb{R}$ with finite absolute q-th moment and admits the following representation in terms of quantile functions

$$d_q(F, G) = \Big(\int_0^1 \big| F^{-1}(u) - G^{-1}(u) \big|^q du \Big)^{1/q}.$$

Furthermore, the convergence of a sequence $(G_n)_{n \in \mathbb{N}}$ to G with respect to d_q is equivalent to $G_n \xrightarrow{\mathcal{L}} G$ in addition to the convergence of $\int |x|^q G_n(dx)$ to $\int |x|^q G(dx)$, where the symbol $\xrightarrow{\mathcal{L}}$ denotes weak convergence.

Here and in the following, we will not distinguish between a metric on some set of distribution functions and its restriction to subsets thereof.

It follows from the above that $(\mathcal{G}_q, d_q)$ is a metric space for all $q \in [2, \infty)$.

Let us denote by d_K the *Kolmogorov* or *supremum metric* on $\mathcal{G}_q$, i.e.,

$$d_K(F, G) := \|F - G\|_\infty = \sup_{x \in \mathbb{R}} |F(x) - G(x)|$$

for F, $G \in \mathcal{G}_q$. Then $d_K(G_n, G) \to 0$ is equivalent to the weak convergence $G_n \xrightarrow{\mathcal{L}} G$, since all $G \in \mathcal{G}_q$ are continuous.

Lemma 4.1

(i) *For every $r \in [1, q]$ the function*

$$(\mathcal{G}_q, d_q) \ \ni \ F \ \mapsto \ \int_{\mathbb{R}} |x|^r F(dx) \ \in (\mathbb{R}, |\cdot|)$$

is continuous.

(ii) *The identity function*

$$id\colon (\mathcal{G}_q, d_q) \ \ni \ F \ \mapsto \ F \ \in (\mathcal{G}_q, d_K)$$

is continuous.

Proof.
(i) Let F_1, $F_2 \in \mathcal{G}_q$. Using Lyapunov's inequality, we see that $d_q(F_1, F_2) \geq d_r(F_1, F_2)$ for every $1 \leq r \leq q$. Thus, $d_q(F_n^*, F) \to 0$ as $n \to \infty$ implies $d_r(F_n^*, F) \to 0$ as $n \to \infty$ for all $(F_n^*)_{n\in\mathbb{N}}, F \in \mathcal{G}_q$. But since convergence of F_n^* to F with respect to d_r implies $\int |x|^r F_n^*(dx) \to \int |x|^r F(dx)$ as $n \to \infty$, this completes the proof.
(ii) Let $(F_n^*)_{n\in\mathbb{N}}$, $F \in \mathcal{G}_q$ with $d_q(F_n^*, F) \to 0$ as $n \to \infty$. Then $F_n^* \xrightarrow{\mathcal{L}} F$, and this implies $d_K(id(F_n^*), id(F)) = d_K(F_n^*, F) \to 0$ as $n \to \infty$. □

Now let $(\Omega, \mathcal{A})$ be a measurable space and $\{P_F \colon F \in \mathcal{G}_q\}$ be a family of probability measures on $\mathcal{A}$ such that on $(\Omega, \mathcal{A})$ there is a sequence $X_1, X_2, \ldots$ of random variables that are under each P_F, $F \in \mathcal{G}_q$, independent and identically distributed according to the distribution function F. Note that under these assumptions the mapping $F \mapsto P_F$ is injective, as $P_{F_1} = P_{F_2}$ implies that $F_1(x) = P_{F_1}(X_1 \leq x) = P_{F_2}(X_1 \leq x) = F_2(x)$ for every $x \in \mathbb{R}$.

4.1 Simple null hypothesis

Assume now that we have observed a sample $X_1, \ldots, X_n$, $n \geq 2$, with distribution function $F \in \mathcal{G}_q$, but that F is unknown to us. Then we consider testing the simple null hypothesis

$$\mathrm{H}_0\colon F = F_0 \ \text{ versus } \ \mathrm{H}_1\colon F \in \mathcal{G}_q \setminus \{F_0\} \tag{4.2}$$

for some fixed $F_0 \in \mathcal{G}_q$. To test this hypothesis, we will use the classical Cramér-von Mises statistic

$$\mathcal{W}_n^2 = n \cdot \int_{-\infty}^{\infty} \big(F_n(x) - F_0(x)\big)^2 F_0(dx)$$

and its counterpart

$$\mathcal{V}_n^2 = n \cdot \int_{-\infty}^{\infty} \big(\widetilde{F}_n(x) - F_0(x)\big)^2 F_0(dx)$$

based on the centered empirical distribution function $\widetilde{F}_n$. Both of these test statistics are measurable mappings from Ω to $[0, \infty)$.

By Donsker's theorem for the empirical process we have

$$\sqrt{n}(F_n - F) \quad \xrightarrow[n]{\mathcal{L}} \quad B^\circ(F) \quad \text{in} \quad D[-\infty, \infty]$$

under the measure P_F, where $D[-\infty, \infty]$ denotes the space of càdlàg functions from $[-\infty, \infty]$ to $\mathbb{R}$ equipped with the Skorokhod metric s_∞, B° is the Brownian bridge on $[0, 1]$, and "$\xrightarrow[n]{\mathcal{L}}$

in $D[-\infty,\infty]$" denotes convergence in distribution in $D[-\infty,\infty]$. Hence, it follows with the continuous mapping theorem that under H_0

$$\mathcal{W}_n \quad \xrightarrow[n]{\mathcal{L}} \quad \Big(\int_0^1 B^\circ(t)^2 dt\Big)^{1/2} =: \mathcal{W},$$

where $\mathcal{W}_n = \big(\mathcal{W}_n^2\big)^{1/2}$. The distribution function of $\mathcal{W}^2$ was derived by Smirnov and others and has been tabulated. Since the distribution function of $\mathcal{W}^2$ is continuous, so is that of $\mathcal{W}$. Thus, for every $\alpha \in (0,1)$

$$P_{F_0}\big(\mathcal{W}_n > k(\alpha)\big) \quad \underset{n\to\infty}{\longrightarrow} \quad P_{F_0}\big(\mathcal{W} > k(\alpha)\big) = \alpha,$$

where $k(\alpha)$ denotes the $(1-\alpha)$-quantile of the distribution of $\mathcal{W}$. Hence, the classical Cramér-von Mises test defined by the decision rule

$$\text{Reject } \mathrm{H}_0 \quad \Longleftrightarrow \quad \mathcal{W}_n > k(\alpha)$$

has asymptotic level α for the testing problem (4.2).

Moreover, as $\sigma_F^2 = \int_{\mathbb{R}} x^2 F(dx) \in (0,\infty)$ for every $F \in \mathcal{G}_q$, we have by Theorem 3.3. of Zhang [31] that

$$\sqrt{n}(\widetilde{F}_n - F) \quad \xrightarrow[n]{\mathcal{L}} \quad W \quad \text{in} \quad D[-\infty,\infty] \tag{4.3}$$

under P_F, where $W = (W(x))_{x\in[-\infty,\infty]}$ is a centered Gaussian process with continuous sample paths and covariance function

$$\mathrm{cov}_F\big(W(x), W(y)\big) = F(x\wedge y) - F(x)F(y) - \frac{U_F(x)U_F(y)}{\sigma_F^2}, \quad x, y \in \mathbb{R}, \tag{4.4}$$

with $U_F(x) = \int_{-\infty}^x yF(dy)$ as in the previous section, see also Theorem B in Genz and Häusler [12] and the remark thereafter. An application of the continuous mapping theorem to this functional central limit theorem now yields that under H_0

$$\mathcal{V}_n \quad \xrightarrow[n]{\mathcal{L}} \quad \Big(\int_0^1 W\big(F_0^{-1}(t)\big)^2 dt\Big)^{1/2} =: \mathcal{V}, \tag{4.5}$$

where $\mathcal{V}_n = \big(\mathcal{V}_n^2\big)^{1/2}$. Since $\mathcal{V}$ is the L_2 norm of the process $W \circ F_0^{-1}$, it follows directly from the Karhunen-Loève expansion of $W \circ F_0^{-1}$ that $\mathcal{V}$ is equal in distribution to

$$\Big(\sum_{j=1}^\infty \lambda_j^* N_j^2\Big)^{1/2}, \tag{4.6}$$

where the N_j are independent and identically $\mathcal{N}(0,1)$-distributed random variables (with $\mathcal{N}(\mu,\sigma^2)$ denoting the normal distribution with mean μ and variance σ^2) and $(\lambda_j^*)_{j\in\mathbb{N}}$ is the decreasing sequence of positive eigenvalues of the Hilbert-Schmidt integral operator having kernel

$$\begin{aligned} k(s,t) &= \mathrm{cov}_{F_0}\big(W\big(F_0^{-1}(s)\big), W\big(F_0^{-1}(t)\big)\big) \\ &= s\wedge t - s\cdot t - \frac{1}{\sigma_{F_0}^2}\int_0^s F_0^{-1}(u)du \int_0^t F_0^{-1}(u)du, \quad s,t \in [0,1], \end{aligned} \tag{4.7}$$

such that each positive eigenvalue is repeated in the sequence $(\lambda_j^*)_{j\in\mathbb{N}}$ as many times as its multiplicity.

It is obvious that the distribution function of the random variable in (4.6) is continuous, and its distribution function is just the one of $\mathcal{V}$. Using this, it follows that

$$P_{F_0}\big(\mathcal{V}_n > c(\alpha)\big) \underset{n\to\infty}{\longrightarrow} P_{F_0}\big(\mathcal{V} > c(\alpha)\big) = \alpha$$

for every $\alpha \in (0,1)$, where $c(\alpha)$ denotes the $(1-\alpha)$-quantile of the distribution of $\mathcal{V}$. Hence, the modified Cramér-von Mises test given by the decision rule

$$\text{Reject } \mathrm{H}_0 \iff \mathcal{V}_n > c(\alpha)$$

is also asymptotically of level α for testing H_0 versus H_1 in (4.2).

Note that since the kernel k depends on F_0, so do the eigenvalues λ_j^* and hence the distribution of $\mathcal{V}$. For some distribution functions F_0, the quantiles of $\mathcal{V}^2$ were computed and tabulated in section 6 of Hörmann [15].

Remark 4.2: Evidently, $\mathcal{G}_q \setminus \{F_0\} \neq \emptyset$, and for any $F \in \mathcal{G}_q \setminus \{F_0\}$ and $t \in (0,1)$ it is $F_t := tF + (1-t)F_0 \in \mathcal{G}_q \setminus \{F_0\}$ again. Now let $(t_n)_{n\in\mathbb{N}}$ be a sequence in $(0,1)$ with $t_n \to 0$ as $n \to \infty$. Then we have $\lim_{n\to\infty} F_{t_n}(x) = F_0(x)\ \forall\, x \in \mathbb{R}$ and additionally

$$\int_{\mathbb{R}} |x|^q F_{t_n}(dx) = t_n \int_{\mathbb{R}} |x|^q F(dx) + (1-t_n)\int_{\mathbb{R}} |x|^q F_0(dx) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} |x|^q F_0(dx),$$

whence it follows that $d_q(F_{t_n}, F_0) \to 0$ as $n \to \infty$. Thus, for every $\epsilon > 0$ there is an $n_\epsilon \in \mathbb{N}$ with $d_q(F_{t_n}, F_0) < \epsilon$ for all $n \geq n_\epsilon$, and this implies that

$$U_\epsilon(F_0) \cap (\mathcal{G}_q \setminus \{F_0\}) \neq \emptyset \quad \forall\, \epsilon > 0.$$

Hence, condition (2.2) holds for the testing problem (4.2). ♦

As we are in the framework of section 2, we can use the results derived there to determine the limiting (as $\alpha \to 0$) Pitman ARE of the classical Cramér-von Mises test based on $(\mathcal{W}_n)_{n\geq 2}$ with respect to the modified Cramér-von Mises test based on $(\mathcal{V}_n)_{n\geq 2}$. We will proceed by verifying that both of the two sequences of test statistics are Bahadur standard sequences in the sense of Definition 2.1.

It is well known and easy to see that $(\mathcal{W}_n)_{n\geq 2}$ is a standard sequence. Here, the constant a in condition (BII) is equal to $1/\lambda_1$ with $\lambda_1 := 1/\pi^2$, and the function in (BIII) is

$$b\colon \mathcal{G}_q \setminus \{F_0\} \ni F \longmapsto \Big(\int_{-\infty}^{\infty} \big(F(x) - F_0(x)\big)^2 F_0(dx)\Big)^{1/2} \in (0,1], \tag{4.8}$$

cf. Table 1 in section 5 of Wieand [30]. The fact that $(\mathcal{W}_n)_{n\geq 2}$ satisfies condition (BIII) with b as in (4.8) is an immediate consequence of the Glivenko-Cantelli theorem. To see this, note that by Minkowski's inequality we have for every $F \in \mathcal{G}_q \setminus \{F_0\}$

$$\frac{\mathcal{W}_n}{\sqrt{n}} = \Big(\int_{-\infty}^{\infty} \big(F_n(x) - F_0(x)\big)^2 F_0(dx)\Big)^{1/2} \leq \Big(\int_{-\infty}^{\infty} \big(F_n(x) - F(x)\big)^2 F_0(dx)\Big)^{1/2} + b(F)$$

and

$$b(F) = \Big(\int_{-\infty}^{\infty} \big(F(x) - F_0(x)\big)^2 F_0(dx)\Big)^{1/2} \leq \Big(\int_{-\infty}^{\infty} \big(F_n(x) - F(x)\big)^2 F_0(dx)\Big)^{1/2} + \frac{\mathcal{W}_n}{\sqrt{n}},$$

so that

$$\Big|\frac{\mathcal{W}_n}{\sqrt{n}} - b(F)\Big| \leq \Big(\int_{-\infty}^{\infty} \big(F_n(x) - F(x)\big)^2 F_0(dx)\Big)^{1/2} \leq \|F_n - F\|_\infty, \tag{4.9}$$

and the statement follows from the Glivenko-Cantelli theorem. Hence, the approximate slope of $(\mathcal{W}_n)_{n\geq 2}$ is $b(F)^2/\lambda_1$, $F \in \mathcal{G}_q \setminus \{F_0\}$.

Let us show now that $(\mathcal{V}_n)_{n\geq 2}$ is a standard sequence as well, i.e., that conditions (BI), (BII) and (BIII) of Definition 2.1 also hold for $(\mathcal{V}_n)_{n\geq 2}$. This follows similarly to the verification of the very conditions for $(\mathcal{W}_n)_{n\geq 2}$.

To check (BI) and (BII), we have to investigate the distribution of $\mathcal{V}$. As was mentioned before, the random variable $\mathcal{V}$ has a continuous distribution function, so that (BI) holds for $(\mathcal{V}_n)_{n\geq 2}$. In order to see that this distribution function also satisfies condition (BII), we use the following tail probability approximation, which follows from Remark 1 on page 1274 in Linde [19], see also Theorem 2 in Beran [3] and Lemma 2.4 and the remark on page 121 in Gregory [13].

Lemma 4.3
Let $(N_i)_{i\in\mathbb{N}}$ be a sequence of independent and identically $\mathcal{N}(0,1)$-distributed random variables on a probability space $(\mathfrak{X}, \mathcal{X}, P)$, and let $(a_i)_{i\in\mathbb{N}}$ be a monotonically decreasing sequence of non-negative real constants with $a_1 > 0$ and $\sum_{i\in\mathbb{N}} a_i < \infty$. Then

$$\lim_{x\to\infty} \frac{\log P\Big(\sum_{i=1}^{\infty} a_i N_i^2 > x^2\Big)}{x^2} = -\frac{1}{2a_1}.$$

As the kernel k in (4.7) is continuous, Mercer's theorem implies that $\sum_{j\in\mathbb{N}} \lambda_j^* < \infty$. Hence, a direct application of the above lemma to the random variable in (4.6) shows that condition (BII) is satisfied for $(\mathcal{V}_n)_{n\geq 2}$ with $a = 1/\lambda_1^*$. It remains to verify condition (BIII). But using Minkowski's inequality again, it follows analogously to before that

$$\Big|\frac{\mathcal{V}_n}{\sqrt{n}} - b(F)\Big| \;\leq\; \Big(\int_{-\infty}^{\infty} \big(\widetilde{F}_n(x) - F(x)\big)^2 F_0(dx)\Big)^{1/2} \;\leq\; \|\widetilde{F}_n - F\|_\infty \tag{4.10}$$

for every $F \in \mathcal{G}_q \setminus \{F_0\}$, where b is as in (4.8). Now by Theorem 3.1 in Zhang [31] we have

$$\|\widetilde{F}_n - F\|_\infty \underset{n\to\infty}{\longrightarrow} 0 \text{ in } P_F\text{-probability,}$$

so that

$$\frac{\mathcal{V}_n}{\sqrt{n}} - b(F) \underset{n\to\infty}{\longrightarrow} 0 \text{ in } P_F\text{-probability}$$

for every $F \in \mathcal{G}_q \setminus \{F_0\}$.

To sum up, we have shown the following proposition.

Proposition 4.4
The sequence $(\mathcal{V}_n)_{n\geq 2}$ is a standard sequence with approximate slope $b(F)^2/\lambda_1^*$, $F \in \mathcal{G}_q \setminus \{F_0\}$. The approximate Bahadur ARE of $(\mathcal{W}_n)_{n\geq 2}$ relative to $(\mathcal{V}_n)_{n\geq 2}$ is thus λ_1^*/λ_1.

Note that the approximate Bahadur ARE of $(\mathcal{W}_n)_{n\geq 2}$ relative to $(\mathcal{V}_n)_{n\geq 2}$ is independent of the alternative distribution $F \in \mathcal{G}_q \setminus \{F_0\}$, because the function b in condition (BIII) is the same for both sequences of test statistics. Hence, the approximate Bahadur ARE of these two sequences does only depend on F_0, namely through the eigenvalue λ_1^*. Moreover, it is $\lambda_1^* < \lambda_1$, as is shown in Example B.1 in the appendix. Consequently, the ratio λ_1^*/λ_1 is always less than one. For some specific distribution functions F_0 the values of λ_1^* and λ_1^*/λ_1 are given in Table 1 on page 35.

In order to equate the approximate Bahadur ARE of $(\mathcal{W}_n)_{n\geq 2}$ with respect to $(\mathcal{V}_n)_{n\geq 2}$ to the limiting Pitman ARE with the aid of Theorem 2.3, we have to verify that the two sequences $(\mathcal{W}_n)_{n\geq 2}$ and $(\mathcal{V}_n)_{n\geq 2}$ also meet Wieand's condition (WIII). To show this for the latter sequence, we require $q \in (2, \infty)$ from now on.

Theorem 4.5
The sequence $(\mathcal{V}_n)_{n\geq 2}$ *fulfills Wieand's condition* (WIII) *with* b *as in* (4.8).

Proof. Let $F \in \mathcal{G}_q \setminus \{F_0\}$ and $n \geq 2$. Using (4.10), we see that

$$\Big|\frac{\mathcal{V}_n}{\sqrt{n}} - b(F)\Big| \;\leq\; \|\widetilde{F}_n - F\|_\infty \;\leq\; \|\widetilde{F}_n - F_n\|_\infty + \|F_n - F\|_\infty.$$

We will now examine both summands on the right-hand side of the above inequality separately.

We start by showing that the Kolmogorov-Smirnov statistic $\sqrt{n}\|F_n - F\|_\infty$ fulfills the assumptions of Proposition 2.4. By the classical asymptotic theory of the Kolmogorov-Smirnov statistic,

$$P_F(\sqrt{n}\|F_n - F\|_\infty \leq x) \underset{n\to\infty}{\longrightarrow} P_F\big(\|B^\circ(F)\|_\infty \leq x\big), \quad x \in \mathbb{R},$$

and $P_F(\sqrt{n}\|F_n - F\|_\infty \leq \cdot\,)$ as well as $P_F(\|B^\circ(F)\|_\infty \leq \cdot\,)$ do not depend on F anymore, since F is continuous. Moreover, the distribution function of $\|B^\circ(F)\|_\infty$ is continuous. This shows that condition (i) of Proposition 2.4 holds for every $\varrho > 0$. But since $Q(\cdot) := P_F(\|B^\circ(F)\|_\infty \leq \cdot\,)$ does not depend on F, condition (ii) of the very proposition is trivially met for every $\varrho > 0$.

Next, we investigate the term $\|\widetilde{F}_n - F_n\|_\infty$.

Let $K := \int_\mathbb{R} |x| F_0(dx)/2$. Then $K \in (0,\infty)$ since $F_0 \in \mathcal{G}_q$, and Lemma 4.1 (i) implies that there are $\delta_1, \delta_2 > 0$ such that

$$\Big|\int_\mathbb{R} |x| F_0(dx) - \int_\mathbb{R} |x| F(dx)\Big| < K \;\text{ for all } F \in \mathcal{G}_q \text{ with } d_q(F, F_0) < \delta_1 \tag{4.11}$$

and

$$\Big|\int_\mathbb{R} |x|^q F_0(dx) - \int_\mathbb{R} |x|^q F(dx)\Big| < K \;\text{ for all } F \in \mathcal{G}_q \text{ with } d_q(F, F_0) < \delta_2. \tag{4.12}$$

Now set $K' := \min\big(F_0(0), 1 - F_0(0)\big)/2$. Note that $K' > 0$ because neither $F_0(0) = 0$ nor $F_0(0) = 1$ is possible since $F_0 \in \mathcal{G}_q$. By part (ii) of Lemma 4.1 there is a $\delta_3 > 0$ with

$$d_K(F, F_0) = \sup_{x\in\mathbb{R}} |F(x) - F_0(x)| < K' \;\text{ for all } F \in \mathcal{G}_q \text{ with } d_q(F, F_0) < \delta_3. \tag{4.13}$$

Define $\tilde{\varrho} := \min(\delta_1, \delta_2, \delta_3)$ and $M := U_{\tilde{\varrho}}(F_0) \setminus \{F_0\}$. Then the set M is such that (3.7), (3.8) and (3.13) hold. To see that (3.7) is satisfied, we have to show that

$$g(c) := \sup_{F\in M} \int_{\{x\in\mathbb{R}:\,|x|>c\}} x^2 F(dx) \to 0 \;\text{ for }\; c \to \infty.$$

But for every $F \in M$ and $c \in (0,\infty)$ it is

$$\begin{aligned}\int_{\{|x|>c\}} x^2 F(dx) &= \int_{\{(\frac{|x|}{c})^{q-2} > 1\}} x^2 F(dx) \;\leq\; \int_{\{(\frac{|x|}{c})^{q-2} > 1\}} x^2 \cdot \Big(\frac{|x|}{c}\Big)^{q-2} F(dx) \\ &\leq\; c^{2-q} \int_\mathbb{R} |x|^q F(dx) \;<\; \frac{1}{c^{q-2}}\big(K + \int_\mathbb{R} |x|^q F_0(dx)\big)\end{aligned}$$

because of $q > 2$ and (4.12). This implies

$$0 \;\leq\; g(c) \;\leq\; \frac{1}{c^{q-2}}\big(K + \int_\mathbb{R} |x|^q F_0(dx)\big) \underset{c\to\infty}{\longrightarrow} 0.$$

Next, we verify (3.8). Because of (4.11) we have $\int_\mathbb{R} |x| F(dx) > K$ for every $F \in M$, and therefore $\inf\{\int_\mathbb{R} |x| F(dx) \colon F \in M\} \geq K > 0$.

It remains to show (3.13). It follows from (4.13) and by definition of K' that

$$0 < F_0(0) - K' < F(0) < F_0(0) + K' < 1$$

for every $F \in M$, and this yields (3.13).

Since M satisfies all assumptions of Proposition 3.5, we get

$$\|\widetilde{F}_n - F_n\|_\infty \;\leq\; \|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{n}\sum_{i=1}^n X_i\Big| + \|R_{n,F}\|_\infty$$

with $\sqrt{n}\|R_{n,F}\|_\infty = o_P^u(1)$ in M as $n \to \infty$. By parts (i) and (ii) of Lemma 3.8 we see that all assumptions of Proposition 2.6 are met for

$$\overline{V}_{n,F} := \|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\Big| \qquad \text{and} \qquad \overline{R}_{n,F} := \sqrt{n}\|R_{n,F}\|_\infty$$

with $\varrho = \tilde{\varrho}$, so that

$$\sup_{F\in M}\big|P_F(\overline{V}_{n,F} + \overline{R}_{n,F} \leq x) - Q_F(x)\big| \;\underset{n}{\longrightarrow}\; 0 \quad \forall\, x \in \mathbb{R}, \tag{4.14}$$

where Q_F is as in Lemma 3.8 (i) the distribution function of a half-normal distribution. Now (4.14) and part (iii) of Lemma 3.8 imply that all assumptions of Proposition 2.4 hold for the family of sequences $\{(\overline{V}_{n,F} + \overline{R}_{n,F})_{n\geq 2} \colon F \in \mathcal{G}_q\}$.

Hence, by Corollary 2.5 there is a $\varrho > 0$ such that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a positive constant $C(\epsilon, \delta)$ with

$$\begin{aligned} &P_F\Big(\Big|\frac{\mathcal{V}_n}{\sqrt{n}} - b(F)\Big| \geq \epsilon\, b(F)\Big) \\ &\leq P_F\Big(\|F_n - F\|_\infty + \|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{n}\sum_{i=1}^n X_i\Big| + \|R_{n,F}\|_\infty \geq \epsilon\, b(F)\Big) \;<\; \delta \end{aligned}$$

for all $F \in U_\varrho(F_0) \setminus \{F_0\}$ and for all $n \geq 2$ with $\sqrt{n}\, b(F) > C(\epsilon, \delta)$, but this is just (WIII) for $(\mathcal{V}_n)_{n\geq 2}$. □

An analog of Theorem 4.5 holds for $(\mathcal{W}_n)_{n\geq 2}$ as well. Wieand [30] showed in Example 3 on page 1008 that $\mathcal{W}_n$ satisfies condition (WIII) for parametric alternatives. This is easily seen to be true also in the case of nonparametric alternatives considered here, since it follows immediately from inequality (4.9) and the fact that the Kolmogorov-Smirnov statistic $\sqrt{n}\|F_n - F\|_\infty$ satisfies the assumptions of Proposition 2.4 for every $\varrho > 0$, as was mentioned in the previous proof. Hence, condition (WIII) also holds for $(\mathcal{W}_n)_{n\geq 2}$ with b as in (4.8).

Our aim is now to equate the limit (as the alternative F approaches F_0) of the approximate Bahadur ARE of $(\mathcal{W}_n)_{n\geq 2}$ relative to $(\mathcal{V}_n)_{n\geq 2}$ with the limiting (as $\alpha \to 0$) Pitman ARE of these sequences using Theorem 2.3. To see that this theorem is applicable, we will check its assumptions first:
By what we have already shown it follows that the sequences $(\mathcal{W}_n)_{n\geq 2}$ and $(\mathcal{V}_n)_{n\geq 2}$ satisfy condition (i) of Theorem 2.3. Moreover, the random variables $\mathcal{W}$ and $\mathcal{V}$, to which $\mathcal{W}_n$ and $\mathcal{V}_n$ respectively converge to in distribution under H_0, have distribution functions that are strictly increasing on $(0, \infty)$, since the distribution functions of $\mathcal{W}^2$ and $\mathcal{V}^2$ are strictly increasing on $(0, \infty)$, see e.g. Lemma 5.1 in Hörmann [15]. Thus, condition (ii) also holds. In addition, as $0 < b(F) \leq \|F - F_0\|_\infty = d_K(F, F_0)$ for all $F \in \mathcal{G}_q \setminus \{F_0\}$ and $d_K(F, F_0) \to 0$ as $d_q(F, F_0) \to 0$,

$F \in \mathcal{G}_q \backslash \{F_0\}$, by Lemma 4.1 (ii), assumption (iii) is met as well. It remains to verify condition (iv) of Theorem 2.3. But as mentioned before, the approximate Bahadur ARE of $(\mathcal{W}_n)_{n\geq 2}$ relative to $(\mathcal{V}_n)_{n\geq 2}$ is independent of the alternative F, and therefore its limit as the alternative F approaches F_0 trivially exists and is equal to λ_1^*/λ_1.

Now a direct application of Theorem 2.3 yields

Theorem 4.6
For $T_{1n} = \mathcal{W}_n$, $T_{2n} = \mathcal{V}_n$, $n \geq 2$, and every $\beta \in (0,1)$ it is

$$\lim_{\alpha \to 0} \liminf_{\substack{F \in \mathcal{G}_q \backslash \{F_0\}, \\ d_q(F,F_0) \to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \lim_{\alpha \to 0} \limsup_{\substack{F \in \mathcal{G}_q \backslash \{F_0\}, \\ d_q(F,F_0) \to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \frac{\lambda_1^*}{\lambda_1}. \tag{4.15}$$

Recall that $\lambda_1^* < \lambda_1$, as is shown in Example B.1 in the appendix. Hence, the limiting Pitman ARE in (4.15) is strictly less than one, so that the sequence of tests based on $(\mathcal{V}_n)_{n\geq 2}$ is preferable to the one based on $(\mathcal{W}_n)_{n\geq 2}$. Furthermore, by Remark 5.4 in [15] it is $\lambda_1^* \geq 1/(2\pi)^2$, which implies that $\lambda_1^*/\lambda_1 \geq 0.25$.

We will now explicitly specify λ_1^*/λ_1 for some distribution functions F_0. For this, keep in mind that the distribution of $\mathcal{W}$ does not depend on F_0, so that λ_1, which is the largest eigenvalue of the Hilbert-Schmidt integral operator with kernel $k(s,t) = s \wedge t - st$, is the same for every F_0 and equals $1/\pi^2$. In contrast to this, the kernel in (4.7) depends on F_0, hence the value of λ_1^* may vary for different null-distributions.

In subsection 6.1 of [15] the numerical computation of λ_1^* is described for F_0 being the distribution function of one of the following distributions:

- the standard normal distribution $\mathcal{N}(0,1)$,
- the double exponential distribution (denoted by Dexp) having Lebesgue density $f(x) = 0.5\exp(-|x|)$, $x \in \mathbb{R}$,
- the logistic distribution (denoted by Logistic) having distribution function $F(x) = 1/(1 + \exp(-x))$, $x \in \mathbb{R}$.

Observe that since all of these distributions have finite moments of all order and zero mean, their distribution functions are elements of $\mathcal{G}_q$ regardless of the value of q. Using the R-function `eigenvalues()` of appendix A.1 in [15], we determined λ_1^* for these distributions, the result of which can be found in Table 1.

F_0	λ_1	λ_1^*	λ_1^*/λ_1
$\mathcal{N}(0,1)$	$1/\pi^2$	$1/(2\pi)^2$	0.25
Dexp	$1/\pi^2$	0.02983768	0.2944861
Logistic	$1/\pi^2$	$1/(2\pi)^2$	0.25

Table 1: Values of λ_1 and λ_1^* for some distributions.

4.2 Composite null hypothesis

Assume again that the observations $X_1, \dots, X_n$, $n \geq 2$, are independent and identically distributed with common distribution function $F \in \mathcal{G}_q$, $q \geq 2$ fixed, but that F is unknown to us. Consider now the testing problem

$$\mathrm{H}_0\colon F \in \mathcal{F}_\tau := \left\{F_\tau\left(\frac{\cdot}{\sigma}\right) : \sigma \in (0,\infty)\right\} \quad \text{versus} \quad \mathrm{H}_1\colon F \in \mathcal{G}_q \setminus \mathcal{F}_\tau, \tag{4.16}$$

where F_τ is the distribution function of the *generalized normal* or *exponential power distribution* having Lebesgue density

$$f_\tau(x) = \frac{\tau}{2\Gamma\big(1/\tau\big)} \cdot \exp\big(-|x|^\tau\big), \quad x \in \mathbb{R},$$

with fixed $\tau > 0$ and

$$\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} dy, \quad x > 0,$$

being the Gamma function. For $\tau = 2$, this yields the $\mathcal{N}(0, 1/2)$ distribution, and for $\tau = 1$ the Laplace or double exponential distribution. If $\tau < 2$, the tails of the distribution with density f_τ are heavier than those of the normal distribution, whereas for $\tau > 2$ the tails are lighter. In Figure 1 the density f_τ is depicted for different values of τ. Note that since F_τ is continuous, centered, and has finite moments of all order for every $\tau > 0$, it is indeed $\mathcal{F}_\tau \subset \mathcal{G}_q$ for every $q \geq 2$. For more details on the generalized normal distribution see e.g. Nadarajah [20].

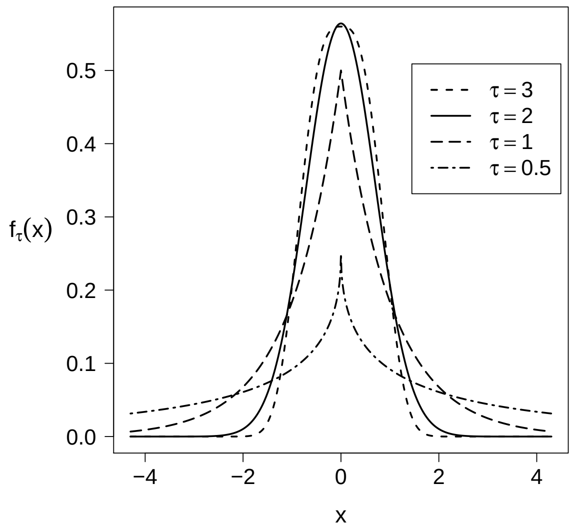

Figure 1: Density f_τ for different values of τ.

Let us introduce some notation. For every $\sigma \in (0,\infty)$ and $x \in \mathbb{R}$ set

$$F(x,\sigma) := F_\tau\left(\frac{x}{\sigma}\right) \quad \text{and} \quad f(x,\sigma) := \frac{1}{\sigma} f_\tau\left(\frac{x}{\sigma}\right).$$

Then $f(\cdot,\sigma)$ is the continuous Lebesgue density of $F(\cdot,\sigma)$. Because τ is kept fixed, we will not mention the dependency of $F(x,\sigma)$ and $f(x,\sigma)$ on τ in this notation.

For testing the composite null hypothesis H_0 against H_1, we will use the Cramér-von Mises statistics

$$\widehat{\mathcal{W}}_n^2 = n\int_{-\infty}^{\infty}\big(F_n(x) - F(x,\hat\sigma_n)\big)^2 F(dx,\hat\sigma_n)$$

and

$$\widehat{\mathcal{V}}_n^2 = n\int_{-\infty}^{\infty}\big(\widetilde{F}_n(x) - F(x,\hat\sigma_n)\big)^2 F(dx,\hat\sigma_n),$$

where $(\hat\sigma_n)_{n\geq 2}$ is the sequence of maximum likelihood estimators (MLE) for the scale parameter σ in $\mathcal{F}_\tau$, i.e.,

$$\hat\sigma_n = \hat\sigma_n(X_1,\dots,X_n) = \tau^{1/\tau}\Big(\frac{1}{n}\sum_{i=1}^n |X_i|^\tau\Big)^{1/\tau} \tag{4.17}$$

for all $n \geq 2$, as is easily seen. Note that $\{\hat\sigma_n = 0\}$ is a P_F-nullset because of the continuity of F. Thus, we can and will always assume that $\hat\sigma_n \in (0,\infty)$. Note moreover that the MLE $\hat\sigma_n$ is scale equivariant, i.e., $\hat\sigma_n(X_1,\dots,X_n) = c\cdot\hat\sigma_n(X_1/c,\dots,X_n/c)$ for every $c\in(0,\infty)$.

It is well known and easy to see that the scale equivariance of $\hat\sigma_n$ implies that $\widehat{\mathcal{W}}_n^2$ is scale invariant, i.e.,

$$\widehat{\mathcal{W}}_n^2(X_1,\dots,X_n) = \widehat{\mathcal{W}}_n^2(X_1/c,\dots,X_n/c)\ \ \forall\ c\in(0,\infty).$$

Now note that by the scale equivariance of $\hat\sigma_n$ the statistic $\widehat{\mathcal{V}}_n^2$ is scale invariant on the set $\Omega_n = \{\min_{1\leq i\leq n} X_i < 0 < \max_{1\leq i\leq n} X_i\}$. In order to verify this, recall that for every $n\geq 2$

$$\widetilde{F}_n(\cdot) = \sum_{i=1}^n \frac{1}{n(1+t_nX_i)}1_{\{X_i\leq\,\cdot\,\}}$$

on Ω_n with $t_n = t_n(X_1,\dots,X_n)$ being the unique solution of the equation $\sum_{i=1}^n X_i/(1+tX_i) = 0$ in the open interval

$$\Big(\Big(\frac{1}{n}-1\Big)\frac{1}{\max\limits_{1\leq i\leq n} X_i}, \Big(\frac{1}{n}-1\Big)\frac{1}{\min\limits_{1\leq i\leq n} X_i}\Big),$$

see (3.2) and (3.3). Then for arbitrary $c>0$

$$\sum_{i=1}^n \frac{1}{n(1+t_nX_i)}1_{\{X_i\leq x\}} = \sum_{i=1}^n \frac{1}{n(1+\tilde t_nY_i)}1_{\{Y_i\leq x/c\}},\quad x\in\mathbb{R},$$

where $\tilde t_n := c\,t_n$ and $Y_i := X_i/c$, $i=1,\dots,n$. But $\tilde t_n = c\,t_n(X_1,\dots,X_n)$ is just the unique solution of $\sum_{i=1}^n Y_i/(1+tY_i) = 0$ in the set

$$\Big(\Big(\frac{1}{n}-1\Big)\frac{1}{\max\limits_{1\leq i\leq n} Y_i}, \Big(\frac{1}{n}-1\Big)\frac{1}{\min\limits_{1\leq i\leq n} Y_i}\Big),$$

denoted by $t_n(Y_1,\dots,Y_n)$. Thus, $c\,t_n(X_1,\dots,X_n) = t_n(Y_1,\dots,Y_n)$, and in combination with the scale equivariance of $\hat\sigma_n$ this yields the statement. Let us assume henceforth that $\widehat{\mathcal{V}}_n^2$ is defined on $\overline{\Omega}_n$ in such a way that it is scale invariant on this set as well (for example, set $\widehat{\mathcal{V}}_n^2 = \widehat{\mathcal{W}}_n^2$ on $\overline{\Omega}_n$). On Ω we then have

$$\widehat{\mathcal{V}}_n^2(X_1,\dots,X_n) = \widehat{\mathcal{V}}_n^2(X_1/c,\dots,X_n/c)\ \ \forall\ c\in(0,\infty).$$

In the following, we will construct asymptotic level α tests for the testing problem (4.16) based on the test statistics $\widehat{\mathcal{W}}_n = \big(\widehat{\mathcal{W}}_n^2\big)^{1/2}$ and $\widehat{\mathcal{V}}_n = \big(\widehat{\mathcal{V}}_n^2\big)^{1/2}$, $n \geq 2$. Note that in order to determine the asymptotic null distributions of these statistics, we can assume that $F = F_\tau$ under H_0 because of the aforementioned scale invariance.

To start with, we observe that the following regularity conditions hold:

The mapping

$$\big((0,\infty), |\cdot|\big) \ni \sigma \mapsto F(\cdot,\sigma) \in \big(C[-\infty,\infty], \|\cdot\|_\infty\big) \tag{4.18}$$

is differentiable at $\sigma = 1$, i.e., there is a function $\Delta \in C[-\infty,\infty]$ with

$$\|F(\cdot, 1+h) - F(\cdot,1) - \Delta(\cdot)h\|_\infty = o(|h|) \quad \text{as } h \to 0, \tag{4.19}$$

namely $\Delta(x) = -xf_\tau(x)$, $x \in \mathbb{R}$, $\Delta(-\infty) = \Delta(\infty) = 0$. Note that the differentiability of $\sigma \mapsto F(\cdot,\sigma)$ at $\sigma = 1$ implies its differentiability on $(0,\infty)$.

Moreover, the MLE $\hat\sigma_n$ admits the expansion

$$\hat\sigma_n(X_1,\dots,X_n) - 1 = \frac{1}{n}\sum_{i=1}^n \Big(|X_i|^\tau - \frac{1}{\tau}\Big) + R_n \tag{4.20}$$

with $\sqrt{n}R_n$ converging to zero in P_{F_τ}-probability as $n \to \infty$. Let $L(x) = |x|^\tau - 1/\tau$, $x \in \mathbb{R}$. Then L is a measurable function with $\mathrm{E}_{F_\tau}(L(X_1)) = 0$ because of $\int_{\mathbb{R}} |x|^\tau F_\tau(dx) = 1/\tau$, and $\mathrm{E}_{F_\tau}(L(X_1)^2) = \mathrm{Var}_{F_\tau}(|X_1|^\tau) = \mathrm{E}_{F_\tau}(|X_1|^{2\tau}) - \mathrm{E}_{F_\tau}(|X_1|^\tau)^2 = 1/\tau < \infty$ since $\int_{\mathbb{R}} |x|^{2\tau} F_\tau(dx) = 1/\tau^2 + 1/\tau$.

These regularity conditions now imply that

$$\sqrt{n}\big(F_n(\cdot) - F(\cdot,\hat\sigma_n)\big) \quad \xrightarrow[n]{\mathcal{L}} \quad Z \quad \text{in } D[-\infty,\infty] \tag{4.21}$$

under P_{F_τ}, where $Z = (Z(x))_{x\in[-\infty,\infty]}$ is a centered Gaussian process with continuous sample paths and covariance function

$$\begin{aligned}\mathrm{cov}_{F_\tau}\big(Z(x), Z(y)\big) &= F_\tau(x\wedge y) - F_\tau(x)F_\tau(y) + xf_\tau(x)\frac{1}{\tau}yf_\tau(y) \\ &\quad + xf_\tau(x)\,\mathrm{E}_{F_\tau}\big(L(X_1)1_{\{X_1\le y\}}\big) + yf_\tau(y)\,\mathrm{E}_{F_\tau}\big(L(X_1)1_{\{X_1\le x\}}\big)\end{aligned} \tag{4.22}$$

for all $x, y \in \mathbb{R}$, cf. Theorem 1 in Durbin [7], see also Theorem A in Genz and Häusler [12] and the remark thereafter. Since for every $x \in \mathbb{R}$ it is

$$\mathrm{E}_{F_\tau}\big(L(X_1)1_{\{X_1\le x\}}\big) = \int_{-\infty}^x |y|^\tau f_\tau(y)dy - \frac{1}{\tau}F_\tau(x) = -\frac{1}{\tau}xf_\tau(x),$$

where the last equality follows by integration by parts, the covariance function (4.22) reduces to

$$\mathrm{cov}_{F_\tau}\big(Z(x), Z(y)\big) = F_\tau(x\wedge y) - F_\tau(x)F_\tau(y) - xf_\tau(x)\frac{1}{\tau}yf_\tau(y), \quad x, y \in \mathbb{R}. \tag{4.23}$$

Additionally, since $\int_{\mathbb{R}} xF_\tau(dx) = 0$ and $\sigma^2_{F_\tau} = \int_{\mathbb{R}} x^2F_\tau(dx) \in (0,\infty)$, it follows from Theorem 1 in Genz and Häusler [12] that under P_{F_τ}

$$\sqrt{n}\big(\widetilde{F}_n(\cdot) - F(\cdot,\hat\sigma_n)\big) \quad \xrightarrow[n]{\mathcal{L}} \quad V \quad \text{in } D[-\infty,\infty] \tag{4.24}$$

with a centered Gaussian processs $V = (V(x))_{x\in[-\infty,\infty]}$ having continuous sample paths and covariance function

$$\mathrm{cov}_{F_\tau}\big(V(x), V(y)\big) = F_\tau(x\wedge y) - F_\tau(x)F_\tau(y) - xf_\tau(x)\frac{1}{\tau}yf_\tau(y) - \frac{U_{F_\tau}(x)U_{F_\tau}(y)}{\sigma^2_{F_\tau}} \tag{4.25}$$

for $x, y \in \mathbb{R}$. Note that the additional two summands appearing in the covariance function of V in the general situation of Theorem 1 in [12] do vanish here because $\mathrm{E}_{F_\tau}\big(X_1 L(X_1)\big) = 0$.

Now observe that the density f_τ is infinitely often differentiable for all $x \in \mathbb{R}^* := \mathbb{R} \setminus \{0\}$ with

$$f'_\tau(x) = f_\tau(x)\tau|x|^{\tau-1} \cdot \big(-\operatorname{sgn}(x)\big),$$

where $\operatorname{sgn} = 1_{(0,\infty)} - 1_{(-\infty,0)}$ is the sign function. Thus,

$$\int_{\mathbb{R}^*} |x f'_\tau(x)| dx = \int_{-\infty}^{\infty} |x|^\tau f_\tau(x)\tau dx = \tau \int_{-\infty}^{\infty} |x|^\tau F_\tau(dx) = 1, \tag{4.26}$$

and by Lemma 2.5 in Hörmann [15] in combination with Example 2.6 from [15] and the continuous mapping theorem this implies that

$$\widehat{\mathcal{W}}_n \quad \xrightarrow[n]{\mathcal{L}} \quad \Big(\int_{-\infty}^{\infty} Z(x)^2 F_\tau(dx)\Big)^{1/2} =: \widehat{\mathcal{W}} \tag{4.27}$$

and

$$\widehat{\mathcal{V}}_n \quad \xrightarrow[n]{\mathcal{L}} \quad \Big(\int_{-\infty}^{\infty} V(x)^2 F_\tau(dx)\Big)^{1/2} =: \widehat{\mathcal{V}} \tag{4.28}$$

under P_{F_τ}. The Karhunen-Loève expansion of the processes $Z \circ F_\tau^{-1}$ and $V \circ F_\tau^{-1}$ then yields

$$\widehat{\mathcal{W}} \sim \Big(\sum_{j=1}^{\infty} \tilde{\lambda}_j N_j^2\Big)^{1/2} \quad \text{and} \quad \widehat{\mathcal{V}} \sim \Big(\sum_{j=1}^{\infty} \tilde{\lambda}_j^* N_j^2\Big)^{1/2} \tag{4.29}$$

with $(N_j)_{j\in\mathbb{N}}$ independent and identically $\mathcal{N}(0,1)$-distributed and $(\tilde{\lambda}_j)_{j\in\mathbb{N}}$ and $(\tilde{\lambda}_j^*)_{j\in\mathbb{N}}$ being the decreasing sequences of positive eigenvalues of the Hilbert-Schmidt integral operators having kernels $k(s,t) = \operatorname{cov}_{F_\tau}\big(Z(F_\tau^{-1}(s)), Z(F_\tau^{-1}(t))\big)$ and $k^*(s,t) = \operatorname{cov}_{F_\tau}\big(V(F_\tau^{-1}(s)), V(F_\tau^{-1}(t))\big)$, respectively, where each positive eigenvalue is repeated as many times as its multiplicity. The symbol $\sim$ in (4.29) signifies equality in distribution.

It is obvious by (4.29) that the distribution functions of $\widehat{\mathcal{W}}$ and $\widehat{\mathcal{V}}$ are continuous. Hence, for every $\alpha \in (0,1)$

$$P_{F_\tau}\big(\widehat{\mathcal{W}}_n > \tilde{k}(\alpha)\big) \underset{n\to\infty}{\longrightarrow} P_{F_\tau}\big(\widehat{\mathcal{W}} > \tilde{k}(\alpha)\big) = \alpha$$

and

$$P_{F_\tau}\big(\widehat{\mathcal{V}}_n > \tilde{c}(\alpha)\big) \underset{n\to\infty}{\longrightarrow} P_{F_\tau}\big(\widehat{\mathcal{V}} > \tilde{c}(\alpha)\big) = \alpha,$$

where $\tilde{k}(\alpha)$ and $\tilde{c}(\alpha)$ denote the $(1-\alpha)$-quantiles of the distributions of $\widehat{\mathcal{W}}$ and $\widehat{\mathcal{V}}$, respectively. For $\tau = 1$ and $\tau = 2$, the quantiles of the distributions of $\widehat{\mathcal{W}}^2$ and $\widehat{\mathcal{V}}^2$ were computed and tabulated in section 7 of Hörmann [15].

Because of the above, both the classical Cramér-von Mises test having decision rule

$$\text{Reject } \mathrm{H}_0 \quad \Longleftrightarrow \quad \widehat{\mathcal{W}}_n > \tilde{k}(\alpha) \tag{4.30}$$

and the modified Cramér-von Mises test with decision rule

$$\text{Reject } \mathrm{H}_0 \quad \Longleftrightarrow \quad \widehat{\mathcal{V}}_n > \tilde{c}(\alpha) \tag{4.31}$$

are asymptotically of level α for testing (4.16).

For the following considerations we assume that $q \geq \max(2,\tau)$. Then

$$m_\tau(F) := \Big(\tau \int_{\mathbb{R}} |x|^\tau F(dx)\Big)^{1/\tau} \in (0,\infty)$$

for all $F \in \mathcal{G}_q$. Note that $m_\tau(F_\tau) = 1$, since $\mathrm{E}_{F_\tau}(|X_1|^\tau) = 1/\tau$. On $\mathcal{G}_q$ we can now define a relation $\sim_R$ by

$$F \sim_R G \quad \Longleftrightarrow \quad F(m_\tau(F)\,\cdot\,) = G(m_\tau(G)\,\cdot\,).$$

Obviously $\sim_R$ is reflexive, symmetric and transitive, so that it is an equivalence relation. For every $F \in \mathcal{G}_q$ its equivalence class under $\sim_R$ is just the scale family generated by F, i.e.,

$$[F]_R := \{G \in \mathcal{G}_q \colon G \sim_R F\} = \big\{F\big(\cdot/c\big) \colon c \in (0,\infty)\big\}.$$

Now the aforementioned scale invariance of the test statistics under consideration yields that the mappings

$$F \mapsto P_F \circ \widehat{\mathcal{W}}_n^{-1} \quad \text{and} \quad F \mapsto P_F \circ \widehat{\mathcal{V}}_n^{-1}$$

from $\mathcal{G}_q$ into the set of probability measures on $\mathcal{B}^*$ are compatible with $\sim_R$, i.e.,

$$F \sim_R G \quad \Longrightarrow \quad P_F \circ \widehat{\mathcal{W}}_n^{-1} = P_G \circ \widehat{\mathcal{W}}_n^{-1} \;\text{ and }\; P_F \circ \widehat{\mathcal{V}}_n^{-1} = P_G \circ \widehat{\mathcal{V}}_n^{-1}.$$

We have already made use of this fact when determining the asymptotic null distributions of $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ and $(\widehat{\mathcal{V}}_n)_{n\geq 2}$. But note that this also implies that the power of the tests in (4.30) and (4.31) is invariant with respect to the scale of the underlying data. Because of this, the quantities

$$N_1(\alpha,\beta,F) := \min\big\{n \geq 2 \colon P_F\big(\widehat{\mathcal{W}}_m > \tilde{k}(\alpha)\big) \geq \beta \;\forall\, m \geq n\big\}$$

and

$$N_2(\alpha,\beta,F) := \min\big\{n \geq 2 \colon P_F\big(\widehat{\mathcal{V}}_m > \tilde{c}(\alpha)\big) \geq \beta \;\forall\, m \geq n\big\},$$

$\alpha,\beta \in (0,1)$, $F \in \mathcal{G}_q \setminus \mathcal{F}_\tau$, are compatible with $\sim_R$ as well, since $N_i(\alpha,\beta,F) = N_i(\alpha,\beta,F(\cdot/c))$ for all $c > 0$, $i = 1,2$, by what has just been mentioned. Hence, for every fixed $\alpha,\beta \in (0,1)$ the relative efficiency $N_2(\alpha,\beta,F)/N_1(\alpha,\beta,F)$ of $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ with respect to $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ is invariant on the equivalence classes of $\sim_R$, and therefore a reasonable investigation of the asymptotic behavior of the relative efficiency when the alternative approaches the null hypothesis requires the identification of distribution functions deriving from the same scale family. Because of this, we will consider in the following the mappings

$$[F]_R \mapsto P_F \circ \widehat{\mathcal{W}}_n^{-1} \quad \text{and} \quad [F]_R \mapsto P_F \circ \widehat{\mathcal{V}}_n^{-1}$$

on the quotient set $\mathcal{G}_q/{\sim_R} := \{[F]_R \colon F \in \mathcal{G}_q\}$. Note that these mappings are well-defined because of the above.

We will now equip the quotient set $\mathcal{G}_q/{\sim_R}$ with a suitable metric.

For this, observe that the test statistics $\widehat{\mathcal{W}}_n$ and $\widehat{\mathcal{V}}_n$ can be written as

$$\widehat{\mathcal{W}}_n = \sqrt{n}\,\Big\|F_n\Big(\frac{\hat{\sigma}_n}{c}\,\cdot\Big) - F_\tau\Big(\frac{\cdot}{c}\Big)\Big\|_{L_2(F_\tau(\cdot/c))} \quad \text{and} \quad \widehat{\mathcal{V}}_n = \sqrt{n}\,\Big\|\widetilde{F}_n\Big(\frac{\hat{\sigma}_n}{c}\,\cdot\Big) - F_\tau\Big(\frac{\cdot}{c}\Big)\Big\|_{L_2(F_\tau(\cdot/c))}$$

for every $c \in (0,\infty)$. Thus, for fixed $c > 0$ they compare the unknown underlying distribution function F deriving from a scale family $[F]_R$ to $[F_\tau]_R$, the scale family under H_0, by measuring the distance of an estimator of $F(m_\tau(F)/c\,\cdot\,)$ to $F_\tau(\cdot/c) = F_\tau(m_\tau(F_\tau)/c\,\cdot\,)$ in the $L_2(F_\tau(\cdot/c))$ metric.

In analogy to this, let us introduce the following metrics on $\mathcal{G}_q/{\sim_R}$:

Lemma 4.7
For every fixed $c > 0$, set

$$\tilde{d}_{q,c}([F]_R, [G]_R) := d_q\big(F(m_\tau(F)/c\,\cdot\,), G(m_\tau(G)/c\,\cdot\,)\big)$$

for every $[F]_R, [G]_R \in \mathcal{G}_q/{\sim_R}$, where d_q is the Kantorovich-Wasserstein metric defined in (4.1). *Then*

(i) $\tilde{d}_{q,c}$ *is well-defined and a metric on* $\mathcal{G}_q/{\sim_R}$,

(ii) *for any two constants $c_1, c_2 \in (0,\infty)$ the metrics $\tilde{d}_{q,c_1}$ and $\tilde{d}_{q,c_2}$ are uniformly equivalent.*

Proof.
(i) To see that $\tilde{d}_{q,c}$ is well-defined, note that for every F_i, $G_i \in \mathcal{G}_q$, $i = 1,2$, with $F_1 \sim_R F_2$, $G_1 \sim_R G_2$ it is

$$F_1(m_\tau(F_1)\,\cdot\,) = F_2(m_\tau(F_2)\,\cdot\,) \quad \text{and} \quad G_1(m_\tau(G_1)\,\cdot\,) = G_2(m_\tau(G_2)\,\cdot\,)$$

by definition of $\sim_R$. Hence

$$d_q\big(F_1(m_\tau(F_1)/c\,\cdot\,), G_1(m_\tau(G_1)/c\,\cdot\,)\big) = d_q\big(F_2(m_\tau(F_2)/c\,\cdot\,), G_2(m_\tau(G_2)/c\,\cdot\,)\big),$$

which yields the statement.

Moreover,

$$\begin{aligned}
\tilde{d}_{q,c}([F]_R, [G]_R) = 0 \quad &\Leftrightarrow \quad d_q\big(F\big(m_\tau(F)/c\,\cdot\,\big), G\big(m_\tau(G)/c\,\cdot\,\big)\big) = 0 \\
&\Leftrightarrow \quad F\big(m_\tau(F)/c\,\cdot\,\big) = G\big(m_\tau(G)/c\,\cdot\,\big) \\
&\Leftrightarrow \quad F\big(m_\tau(F)\,\cdot\,\big) = G\big(m_\tau(G)\,\cdot\,\big) \\
&\Leftrightarrow \quad F \sim_R G \quad \Leftrightarrow \quad [F]_R = [G]_R.
\end{aligned}$$

The fact that $\tilde{d}_{q,c}$ is symmetric and satisfies the triangle inequality follows directly from the respective properties of d_q.

(ii) By the scaling properties of d_q it is

$$\tilde{d}_{q,c_1}([F]_R, [G]_R) = \frac{c_1}{c_2}\, \tilde{d}_{q,c_2}([F]_R, [G]_R)$$

for all $[F]_R, [G]_R \in \mathcal{G}_q/{\sim_R}$, whence the assertion follows. □

By what has just been shown, any two of the metric spaces $(\mathcal{G}_q/{\sim_R}, \tilde{d}_{q,c})$, $c > 0$, are uniformly, and therefore topologically, isomorphic. Since we are only interested in topological properties of these metric spaces such as convergence of sequences in them and continuity of mappings on them, we will not differentiate between these spaces and therefore always work on $(\mathcal{G}_q/{\sim_R}, \tilde{d}_{q,1})$.

Now observe that the set

$$\widetilde{\mathcal{G}}_q := \big\{F\big(m_\tau(F)\,\cdot\,\big) \colon F \in \mathcal{G}_q\big\} = \big\{F \in \mathcal{G}_q \colon \tau \textstyle\int_{\mathbb{R}} |x|^\tau F(dx) = 1\big\}$$

is a complete set of equivalence class representatives, i.e., it contains exactly one element from each equivalence class of $\sim_R$. Because of this, the well-defined mapping

$$h\colon\ \mathcal{G}_q/{\sim_R}\ \ni\ [F]_R\ \mapsto\ F\big(m_\tau(F)\,\cdot\,\big)\ \in\ \widetilde{\mathcal{G}}_q$$

is obviously a bijection. Furthermore, it is

$$\tilde{d}_{q,1}\big([F]_R, [G]_R\big) = d_q\big(h([F]_R), h([G]_R)\big) \quad \forall\ [F]_R,\ [G]_R\ \in\ \mathcal{G}_q/{\sim_R},$$

so that $(\mathcal{G}_q/\sim_R, \tilde{d}_{q,1})$ and $(\widetilde{\mathcal{G}}_q, d_q)$ are isometrically isomorphic. Hence, we will identify these two metric spaces and assume from now on that the distribution function F of the data $X_1, \ldots, X_n$, $n \geq 2$, is in $(\widetilde{\mathcal{G}}_q, d_q)$. The appropriate hypotheses for the investigation of the asymptotic behavior of $N_2(\alpha, \beta, F)/N_1(\alpha, \beta, F)$ now are

$$\mathrm{H}_0 \colon F = F_\tau \quad \text{ versus } \quad \mathrm{H}_1 \colon F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}. \tag{4.32}$$

Trivially, the tests in (4.30) and (4.31) are asymptotic level α tests for this testing problem as well. Henceforth we will only consider the testing problem (4.32).

Remark 4.8: Note that the above testing problem is such that condition (2.2) is satisfied. To verify this, observe that since the set of alternatives $\widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$ is obviously not empty, there is an $F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$. Now set $F_t := tF + (1-t)F_\tau$ for every $t \in (0,1)$. It is easy to see that $F_t \in \mathcal{G}_q \setminus \{F_\tau\}$. Moreover,

$$\int_{\mathbb{R}} |x|^\tau F_t(dx) = t \int_{\mathbb{R}} |x|^\tau F(dx) + (1-t) \int_{\mathbb{R}} |x|^\tau F_\tau(dx) = \frac{1}{\tau}(t + 1 - t) = \frac{1}{\tau},$$

so that $F_t \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$ for every $t \in (0,1)$. As in Remark 4.2 we can show that $d_q(F_{t_n}, F_\tau) \to 0$ as $n \to \infty$ for any sequence $(t_n)_{n\in\mathbb{N}}$ in $(0,1)$ converging to zero, whence it follows that

$$U_\epsilon(F_\tau) \cap \big(\widetilde{\mathcal{G}}_q \setminus \{F_\tau\}\big) \neq \emptyset \quad \forall\, \epsilon > 0. \qquad \blacklozenge$$

We proceed by showing that $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ and $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ are standard sequences, the notion of which was introduced in Definition 2.1.

It follows from (4.27), (4.28) and (4.29) that both sequences of test statistics fulfill condition (BI). Furthermore, as the kernels k and k^* are continuous on $[0,1]\times[0,1]$, by Mercer's theorem we get $\sum_{j\in\mathbb{N}} \tilde{\lambda}_j < \infty$ and $\sum_{j\in\mathbb{N}} \tilde{\lambda}_j^* < \infty$. Hence, we can use Lemma 4.3 again to see that condition (BII) holds for both sequences of test statistics as well, where $a = 1/\tilde{\lambda}_1$ for $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ and $a = 1/\tilde{\lambda}_1^*$ for $(\widehat{\mathcal{V}}_n)_{n\geq 2}$.

It remains to verify (BIII). For every $F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$ let

$$b(F) = \Big(\int_{-\infty}^{\infty} \big(F(x) - F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2}$$

and

$$b_n(F) = \Big(\int_{-\infty}^{\infty} \big(F(x) - F(x, \hat{\sigma}_n)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2}, \quad n \geq 2.$$

Then

$$\Big|\frac{\widehat{\mathcal{W}}_n}{\sqrt{n}} - b(F)\Big| \;\leq\; \Big|\frac{\widehat{\mathcal{W}}_n}{\sqrt{n}} - b_n(F)\Big| + \big|b_n(F) - b(F)\big|. \tag{4.33}$$

Let us first examine the first term on the right-hand side of (4.33). Using Minkowski's inequality, it is

$$\Big|\frac{\widehat{\mathcal{W}}_n}{\sqrt{n}} - b_n(F)\Big| \;\leq\; \Big(\int_{-\infty}^{\infty} \big(F_n(x) - F(x)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2} \;\leq\; \|F_n - F\|_\infty. \tag{4.34}$$

Now the Glivenko-Cantelli theorem ensures that $\|F_n - F\|_\infty = o_{P_F}(1)$ as $n \to \infty$, where $o_{P_F}(1)$ signifies convergence to zero in P_F-probability.

Analogously to the above we have

$$\Big|\frac{\widehat{\mathcal{V}}_n}{\sqrt{n}} - b(F)\Big| \;\leq\; \|\widetilde{F}_n - F\|_\infty + \big|b_n(F) - b(F)\big|. \tag{4.35}$$

Again, Theorem 3.1 of Zhang [31] gives $\|\widetilde{F}_n - F\|_\infty = o_{P_F}(1)$ as $n \to \infty$.

It remains to show that $\big|b_n(F) - b(F)\big| \underset{n}{\to} 0$ in P_F-probability. To do this, we will use the next lemma.

Lemma 4.9
For every fixed continuous distribution function F the function

$$T_F\colon (0,\infty) \ni \sigma \mapsto \Big(\int_{-\infty}^{\infty} \big(F(x) - F_\tau(x/\sigma)\big)^2 F_\tau(dx/\sigma)\Big)^{1/2} \in [0,1]$$

is continuous.

Proof. We will show the continuity of T_F^2, whence the assertion follows. Fix some arbitrary $\bar\sigma \in (0,\infty)$ and let $(\sigma_n)_{n\in\mathbb{N}}$ be a sequence in $(0,\infty)$ that tends to $\bar\sigma$ as n tends to infinity. Then

$$F(x,\sigma_n) = F_\tau(x/\sigma_n) \quad \underset{n}{\longrightarrow} \quad F_\tau(x/\bar\sigma) = F(x,\bar\sigma) \quad \forall \;\; x\in\mathbb{R}$$

because of the continuity of F_τ. This shows that the sequence $\big(F(\cdot,\sigma_n)\big)_{n\in\mathbb{N}}$ converges weakly to $F(\cdot,\bar\sigma)$. Now

$$\begin{aligned}
\big|T_F(\sigma_n)^2 - T_F(\bar\sigma)^2\big| &= \Big|\int_{-\infty}^{\infty} \big(F(x) - F(x,\sigma_n)\big)^2 F(dx,\sigma_n) - \int_{-\infty}^{\infty} \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\bar\sigma)\Big| \\
&= \Big|\int_{-\infty}^{\infty} \big(F(x) - F(x,\sigma_n)\big)^2 - \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\sigma_n) \\
&\quad + \int_{-\infty}^{\infty} \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\sigma_n) - \int_{-\infty}^{\infty} \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\bar\sigma)\Big| \\
&\leq \Big|\int_{-\infty}^{\infty} \big(F(x) - F(x,\sigma_n)\big)^2 - \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\sigma_n)\Big| \\
&\quad + \Big|\int_{-\infty}^{\infty} \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\sigma_n) - \int_{-\infty}^{\infty} \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\bar\sigma)\Big|.
\end{aligned}$$

The second term converges to zero as $n\to\infty$ because of the weak convergence of $F(\cdot,\sigma_n)$ to $F(\cdot,\bar\sigma)$, using that $(F(\cdot) - F(\cdot,\bar\sigma))^2$ is a continuous and bounded function on $\mathbb{R}$. It remains to investigate the first term on the right-hand side of the above inequality. It is

$$\begin{aligned}
&\Big|\int_{-\infty}^{\infty} \big(F(x) - F(x,\sigma_n)\big)^2 - \big(F(x) - F(x,\bar\sigma)\big)^2 F(dx,\sigma_n)\Big| \\
&\leq \int_{-\infty}^{\infty} \big|2F(x)\big(F(x,\bar\sigma) - F(x,\sigma_n)\big) + \big(F(x,\sigma_n) - F(x,\bar\sigma)\big)\big(F(x,\sigma_n) + F(x,\bar\sigma)\big)\big| F(dx,\sigma_n) \\
&\leq 4\cdot \|F(\cdot,\sigma_n) - F(\cdot,\bar\sigma)\|_\infty = o(1) \text{ as } n\to\infty,
\end{aligned}$$

which follows again from the weak convergence of $F(\cdot,\sigma_n)$ to $F(\cdot,\bar\sigma)$ and the continuity of the latter function. □

Next, observe that by the strong law of large numbers it is for every $F \in \widetilde{\mathcal{G}}_q$

$$\hat\sigma_n(X_1,\ldots,X_n) = \tau^{1/\tau}\Big(\frac{1}{n}\sum_{i=1}^{n}|X_i|^\tau\Big)^{1/\tau} \quad \underset{n}{\longrightarrow} \quad \tau^{1/\tau}\Big(\int_{\mathbb{R}} |x|^\tau F(dx)\Big)^{1/\tau} = m_\tau(F) = 1$$

P_F-almost everywhere. Together with the preceding lemma this implies that $b_n(F) - b(F) = T_F(\hat\sigma_n) - T_F(1) = o_{P_F}(1)$ as $n\to\infty$, which shows that both $\big(\widehat{\mathcal{W}}_n/\sqrt{n}\big)_{n\geq 2}$ and $\big(\widehat{\mathcal{V}}_n/\sqrt{n}\big)_{n\geq 2}$ converge in P_F-probability to $b(F)$ under H_1, but this is just (BIII).

To sum up, we have shown the following proposition.

Proposition 4.10
The sequences $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ *and* $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ *are standard sequences with approximate slopes* $b(F)^2/\tilde{\lambda}_1$ *and* $b(F)^2/\tilde{\lambda}_1^*$ *respectively. Thus, the approximate Bahadur ARE of* $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ *relative to* $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ *is* $\tilde{\lambda}_1^*/\tilde{\lambda}_1$.

It is shown in Example B.2 in the appendix that $\tilde{\lambda}_1 = 1/\pi^2$ for every $\tau \in (0,\infty)$ and that $\tilde{\lambda}_1^* < \tilde{\lambda}_1$. Because of the latter, the approximate Bahadur ARE of $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ relative to $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ is strictly less than one. Moreover, it is independent of the alternative distribution $F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$ but depends on the parameter τ of the null distribution F_τ through the eigenvalue $\tilde{\lambda}_1^*$. For $\tau = 1$ and $\tau = 2$, the values of $\tilde{\lambda}_1^*$ and $\tilde{\lambda}_1^*/\tilde{\lambda}_1$ are given in Table 2 on page 52.

Next, we want to verify that for both sequences $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ and $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ Wieand's condition (WIII) holds. In order to do this, we have to strengthen the condition on the moments of the distribution of the data and require from now on that $q = 2\tau$ if $\tau > 1$, otherwise q shall be fix in $(2,\infty)$.

Theorem 4.11
The sequences $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ *and* $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ *fulfill Wieand's condition* (WIII) *with*

$$b\colon \widetilde{\mathcal{G}}_q \setminus \{F_\tau\} \ni F \;\longmapsto\; \Big(\int_{-\infty}^{\infty} \big(F(x) - F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2} \in (0,1].$$

To prove this theorem we need some additional results. For the following investigations, let us introduce the condition

$$\sup_{F\in M} \int_{\mathbb{R}} |x|^{2\tau} F(dx) < \infty \tag{4.36}$$

for a set M of distribution functions. We will take a closer look now at the uniform asymptotic behavior of the sequence $(\hat{\sigma}_n)_{n\geq 2}$.

Lemma 4.12
Let $\emptyset \neq M \subset \widetilde{\mathcal{G}}_q$ *with* (4.36)*. Then*

(i) $\sup_{F\in M} \mathrm{E}_F\big(|\hat{\sigma}_n^\tau - 1|^2\big) = O(1/n)$ *as* $n \to \infty$,

(ii) $\hat{\sigma}_n - 1 = o_P^u(1)$ *in* M *as* $n \to \infty$.

Proof. (i) Recall that $\tau \int_{\mathbb{R}} |x|^\tau F(dx) = 1$ for every $F \in \widetilde{\mathcal{G}}_q$. Thus it is

$$\begin{aligned}\mathrm{E}_F\big(|\hat{\sigma}_n^\tau - 1|^2\big) &= \mathrm{E}_F\Big(\Big(\frac{1}{n}\sum_{i=1}^n \big(\tau|X_i|^\tau - 1\big)\Big)^2\Big) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{E}_F\big((\tau|X_i|^\tau - 1)^2\big) = \frac{\mathrm{Var}_F\big(\tau|X_1|^\tau\big)}{n} \\ &= \frac{\tau^2\,\mathrm{E}_F\big(|X_1|^{2\tau}\big) - 1}{n} \;\leq\; \frac{\tau^2}{n}\sup_{F\in M}\int_{\mathbb{R}} |x|^{2\tau} F(dx) < \infty\end{aligned}$$

for every $F \in M$, which yields the statement.

To verify (ii), it obviously suffices to show

$$\sup_{F\in M} P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\big) \;\underset{n\to\infty}{\longrightarrow}\; 0$$

for $\epsilon \in (0,1)$. Now observe that there is a $K_\tau \in (0,\infty)$ such that for every $\epsilon \in (0,1)$ the following inequalities hold

$$(1+\epsilon)^\tau \;\geq\; 1 + K_\tau\epsilon \quad\text{and}\quad (1-\epsilon)^\tau \;\leq\; 1 - K_\tau\epsilon. \tag{4.37}$$

Using this, we have for every $\epsilon \in (0,1)$ and $F \in M$

$$\begin{aligned} P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\big) &= 1 - P_F\big(1-\epsilon < \hat{\sigma}_n < 1+\epsilon\big) = 1 - P_F\big((1-\epsilon)^\tau < \hat{\sigma}_n^\tau < (1+\epsilon)^\tau\big) \\ &\leq 1 - P_F\big(1 - K_\tau\epsilon < \hat{\sigma}_n^\tau < 1 + K_\tau\epsilon\big) = 1 - P_F\big(|\hat{\sigma}_n^\tau - 1| < K_\tau\epsilon\big) \\ &= P_F\big(|\hat{\sigma}_n^\tau - 1| \geq K_\tau\epsilon\big). \end{aligned}$$

Hence, with Markov's inequality and (i) it is

$$\sup_{F\in M} P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\big) \leq \sup_{F\in M} P_F\big(|\hat{\sigma}_n^\tau - 1| \geq K_\tau\epsilon\big) \leq \frac{1}{K_\tau^2\epsilon^2} \sup_{F\in M} \mathrm{E}_F\big(|\hat{\sigma}_n^\tau - 1|^2\big) \underset{n\to\infty}{\longrightarrow} 0. \qquad \square$$

Moreover, we get the following result concerning the sequence $(\hat{\sigma}_n)_{n\geq 2}$.

Lemma 4.13
If the nonempty set $M \subset \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$ satisfies (4.36), then for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C(\epsilon,\delta) > 0$ such that

$$P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\, b(F)\big) < \delta$$

for each $F \in M$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C(\epsilon,\delta)/b(F)$.

Proof. Obviously, it is sufficient to show the statement for every $\epsilon \in (0,1)$. Thus, let $0 < \epsilon, \delta < 1$. Since $b(F) \in (0,1]$ for every $F \in M$, it is $\epsilon\, b(F) \in (0,1)$, and using (4.37) we get

$$P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\, b(F)\big) \leq P_F\big(|\hat{\sigma}_n^\tau - 1| \geq K_\tau\, \epsilon\, b(F)\big) \leq \frac{\widetilde{K}}{nK_\tau^2\epsilon^2 b(F)^2}$$

for every $F \in M$, employing the fact that by Lemma 4.12 (i) there is a $\widetilde{K} \in (0,\infty)$ such that

$$\sup_{F\in M} \mathrm{E}_F\big(|\hat{\sigma}_n^\tau - 1|^2\big) \leq \frac{1}{n}\widetilde{K}$$

for every $n \geq 2$. Now set $C(\epsilon,\delta) := \big(\widetilde{K}/(K_\tau^2\epsilon^2\delta)\big)^{1/2}$. Then for every $F \in M$ it is

$$P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\, b(F)\big) < \delta$$

for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C(\epsilon,\delta)/b(F)$. $\square$

Now note again that the density f_τ is infinitely often differentiable for all $x \in \mathbb{R}^* = \mathbb{R}\setminus\{0\}$ with

$$f_\tau'(x) = f_\tau(x)\tau|x|^{\tau-1} \cdot \big(-\operatorname{sgn}(x)\big)$$

and

$$f_\tau''(x) = f_\tau(x)\tau|x|^{\tau-1} \cdot \Big(\tau|x|^{\tau-1} - (\tau-1)|x|^{-1}\Big),$$

where $\operatorname{sgn} = 1_{(0,\infty)} - 1_{(-\infty,0)}$ is as before the sign function. Hence, for all $(x,\sigma) \in \mathbb{R}\times(0,\infty)$ the first-order and the second-order partial derivative of $f(\cdot,\cdot)$ with respect to σ exist and are given by

$$\frac{\partial f}{\partial \sigma}(x,\sigma) = -\frac{1}{\sigma^2} f_\tau\Big(\frac{x}{\sigma}\Big) + \frac{1}{\sigma^2}\tau f_\tau\Big(\frac{x}{\sigma}\Big)\Big|\frac{x}{\sigma}\Big|^\tau \tag{4.38}$$

and

$$\frac{\partial^2 f}{\partial \sigma^2}(x,\sigma) = \frac{1}{\sigma^3}\bigg(2\cdot f_\tau\Big(\frac{x}{\sigma}\Big) - f_\tau\Big(\frac{x}{\sigma}\Big)\cdot\Big|\frac{x}{\sigma}\Big|^\tau \cdot \big[3\tau + \tau^2\big] + \tau^2 f_\tau\Big(\frac{x}{\sigma}\Big)\cdot\Big|\frac{x}{\sigma}\Big|^{2\tau}\bigg). \tag{4.39}$$

We will now use the foregoing results to prove the following proposition.

Proposition 4.14
Let $\emptyset \neq M \subset \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$ *be such that* (4.36) *holds. Then for each* $\epsilon > 0$ *and* $\delta \in (0,1)$ *there is a* $C(\epsilon, \delta) > 0$ *with*

$$P_F\big(\big|b_n(F) - b(F)\big| \geq \epsilon\, b(F)\big) \;<\; \delta$$

for all $F \in M$ *and for all* $n \in \mathbb{N}$, $n \geq 2$, *with* $\sqrt{n} > C(\epsilon, \delta)/b(F)$.

Proof. Let $\epsilon > 0$, $\delta \in (0,1)$, $n \geq 2$ and $F \in M$. By an application of Minkowski's inequality we see that

$$\begin{aligned}
b_n(F) \;&\leq\; \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2} + \Big(\int_{\mathbb{R}} \big(F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2} \\
&= \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 F_\tau(dx) + \int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 \cdot \big(f(x, \hat{\sigma}_n) - f_\tau(x)\big) dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2} \\
&\leq\; \Big(b(F)^2 + \int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 \cdot \big|f(x, \hat{\sigma}_n) - f_\tau(x)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2} \\
&\leq\; b(F) + \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 \cdot \big|f(x, \hat{\sigma}_n) - f_\tau(x)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2},
\end{aligned}$$

where the last inequality is due to the fact that $\sqrt{a+b} \;\leq\; \sqrt{a} + \sqrt{b}$ for all $a, b \geq 0$. Analogously, it is

$$\begin{aligned}
b(F) \;&\leq\; b_n(F) + \Big(\int_{\mathbb{R}} \big(F(x) - F(x, \hat{\sigma}_n)\big)^2 \cdot \big|f_\tau(x) - f(x, \hat{\sigma}_n)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F(x, \hat{\sigma}_n) - F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\big|b_n(F) - b(F)\big| \;&\leq\; \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 \cdot \big|f(x, \hat{\sigma}_n) - f_\tau(x)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 F(dx, \hat{\sigma}_n)\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F(x) - F(x, \hat{\sigma}_n)\big)^2 \cdot \big|f_\tau(x) - f(x, \hat{\sigma}_n)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F(x, \hat{\sigma}_n) - F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2} \\
&=\; \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 \cdot \big|f(x, \hat{\sigma}_n) - f_\tau(x)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 \big(f(x, \hat{\sigma}_n) - f_\tau(x) + f_\tau(x)\big) dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x) + F_\tau(x) - F(x, \hat{\sigma}_n)\big)^2 \cdot \big|f_\tau(x) - f(x, \hat{\sigma}_n)\big| dx\Big)^{1/2} \\
&\quad + \Big(\int_{\mathbb{R}} \big(F(x, \hat{\sigma}_n) - F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2}
\end{aligned}$$

$$\begin{aligned}
&\le\ 2\cdot\Big(\int_{\mathbb{R}}\big(F(x,\hat\sigma_n)-F_\tau(x)\big)^2F_\tau(dx)\Big)^{1/2}\\
&\quad+(\sqrt{2}+1)\Big(\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\big|f_\tau(x)-f(x,\hat\sigma_n)\big|dx\Big)^{1/2}\\
&\quad+(\sqrt{2}+1)\Big(\int_{\mathbb{R}}\big(F_\tau(x)-F(x,\hat\sigma_n)\big)^2\cdot\big|f_\tau(x)-f(x,\hat\sigma_n)\big|dx\Big)^{1/2}\\
&=:2\ \mathrm{I}_n^{1/2}+(\sqrt{2}+1)\ \mathrm{II}_n^{1/2}+(\sqrt{2}+1)\ \mathrm{III}_n^{1/2}.
\end{aligned}$$

Using Taylor's theorem, we get

$$\begin{aligned}
\mathrm{II}_n&=\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\big|f_\tau(x)-f(x,\hat\sigma_n)\big|dx\\
&=\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\Big|\frac{\partial f}{\partial\sigma}(x,1)\cdot\big(\hat\sigma_n-1\big)+\frac{1}{2}\cdot\frac{\partial^2 f}{\partial\sigma^2}(x,\xi_n)\cdot\big(\hat\sigma_n-1\big)^2\Big|dx
\end{aligned}$$

with $\xi_n=\xi_n(x)$ between $\hat\sigma_n$ and 1. Now

$$\begin{aligned}
\mathrm{II}_n\ &\le\ \big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\Big|\frac{\partial f}{\partial\sigma}(x,1)\Big|dx\\
&\quad+\frac{1}{2}\cdot\big(\hat\sigma_n-1\big)^2\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\Big|\frac{\partial^2 f}{\partial\sigma^2}(x,\xi_n)\Big|dx\\
&=:\ \mathrm{II}_{n,1}+\frac{1}{2}\mathrm{II}_{n,2}.
\end{aligned}\tag{4.40}$$

We will first investigate $\mathrm{II}_{n,1}$. It is

$$\begin{aligned}
\mathrm{II}_{n,1}&=\big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\Big|\frac{\partial f}{\partial\sigma}(x,1)\Big|dx\\
&=\big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\cdot\big|-f_\tau(x)+\tau f_\tau(x)|x|^\tau\big|dx\qquad\text{by (4.38)}\\
&\le\big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2f_\tau(x)dx+\big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2\tau f_\tau(x)|x|^\tau dx\\
&=:\mathrm{II}^*_{n,1}+\mathrm{II}^{**}_{n,1}.
\end{aligned}$$

Now $\mathrm{II}^*_{n,1}=\big|\hat\sigma_n-1\big|\cdot b(F)^2$, so that

$$P_F\big(\mathrm{II}^*_{n,1}\ge\epsilon\,b(F)^2\big)=P_F\big(\big|\hat\sigma_n-1\big|\cdot b(F)^2\ge\epsilon\,b(F)^2\big)=P_F\big(\big|\hat\sigma_n-1\big|\ge\epsilon\big).$$

By part (ii) of Lemma 4.12 it is $\hat\sigma_n-1=o_P^u(1)$ in M as $n\to\infty$. Hence, there is an $n_1(\epsilon,\delta)\in\mathbb{N}$ such that

$$\sup_{F\in M}P_F\big(\big|\hat\sigma_n-1\big|\ge\epsilon\big)<\delta$$

for all $n>n_1(\epsilon,\delta)$. With $C_1(\epsilon,\delta):=\sqrt{n_1(\epsilon,\delta)}$ we therefore have

$$P_F\big(\mathrm{II}^*_{n,1}\ge\epsilon\,b(F)^2\big)<\delta$$

for every $F\in M$ and every $n\in\mathbb{N}$ with $\sqrt{n}>C_1(\epsilon,\delta)/b(F)\ge\sqrt{n_1(\epsilon,\delta)}$.

Next, we investigate $\mathrm{II}^{**}_{n,1}$. It is

$$\begin{aligned}
\mathrm{II}^{**}_{n,1}&=\big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2f_\tau(x)\tau|x|^\tau dx\\
&=\big|\hat\sigma_n-1\big|\cdot\int_{\mathbb{R}}\big(F(x)-F_\tau(x)\big)^2f_\tau(x)^{1/2}\cdot\tau|x|^\tau f_\tau(x)^{1/2}dx
\end{aligned}$$

$$\leq \big|\hat{\sigma}_n - 1\big| \cdot \Big(\int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^4 f_\tau(x)dx\Big)^{1/2} \cdot \Big(\int_{\mathbb{R}} \tau^2 f_\tau(x)|x|^{2\tau}dx\Big)^{1/2}$$
$$\leq \big|\hat{\sigma}_n - 1\big| \cdot b(F) \cdot c_\tau,$$

where $c_\tau := \tau\big(\int_{\mathbb{R}} |x|^{2\tau} F_\tau(dx)\big)^{1/2} \in (0,\infty)$. Thus,

$$P_F\big(\mathrm{II}^{**}_{n,1} \geq \epsilon\, b(F)^2\big) \;\leq\; P_F\big(|\hat{\sigma}_n - 1| \geq \epsilon\, b(F)/c_\tau\big),$$

and by Lemma 4.13 there is a $C_2(\epsilon,\delta) > 0$ such that the right-hand side of this inequality is less than δ for every $F \in M$ and $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C_2(\epsilon,\delta)/b(F)$.

We now take a look at $\mathrm{II}_{n,2}$ defined in (4.40). Recall that

$$\mathrm{II}_{n,2} = \big(\hat{\sigma}_n - 1\big)^2 \cdot \int_{\mathbb{R}} \big(F(x) - F_\tau(x)\big)^2 \cdot \Big|\frac{\partial^2 f}{\partial \sigma^2}(x,\xi_n)\Big| dx$$

with $\xi_n = \xi_n(x)$ lying between $\hat{\sigma}_n$ and 1.

Fix some $h \in (0,1)$ for the rest of the proof. The absolute value of the second-order partial derivative $\partial^2 f/\partial\sigma^2$ is bounded above on $\mathbb{R} \times [1-h, 1+h]$ by some integrable function $H\colon \mathbb{R} \to [0,\infty)$, that is,

$$\Big|\frac{\partial^2 f}{\partial \sigma^2}(x,\sigma)\Big| \;\leq\; H(x) \qquad \forall\; (x,\sigma) \in \mathbb{R} \times [1-h, 1+h], \tag{4.41}$$

and the majorant H can be taken to be

$$H(x) = \frac{1}{(1-h)^3}\Big(2 \cdot f_\tau\Big(\frac{x}{1+h}\Big) + f_\tau\Big(\frac{x}{1+h}\Big) \cdot \Big|\frac{x}{1-h}\Big|^\tau \cdot \big[3\tau + \tau^2\big] + \tau^2 f_\tau\Big(\frac{x}{1+h}\Big) \cdot \Big|\frac{x}{1-h}\Big|^{2\tau}\Big)$$

for all $x \in \mathbb{R}$, cf. (4.39). Obviously, $0 < \int_{\mathbb{R}} H(x)dx < \infty$.

Coming back to $\mathrm{II}_{n,2}$, on the event $\{|\hat{\sigma}_n - 1| \leq h\}$ we have $(x,\xi_n) \in \mathbb{R} \times [1-h, 1+h]$ for all $x \in \mathbb{R}$, and so it follows in this case directly from (4.41) that

$$\mathrm{II}_{n,2} \;\leq\; \big(\hat{\sigma}_n - 1\big)^2 \cdot \int_{\mathbb{R}} H(x)dx.$$

Thus, it is

$$P_F\big(\mathrm{II}_{n,2} \geq \epsilon\, b(F)^2\big) \;\leq\; P_F\big(\mathrm{II}_{n,2} \geq \epsilon\, b(F)^2, |\hat{\sigma}_n - 1| \leq h\big) + P_F(|\hat{\sigma}_n - 1| > h)$$
$$\leq\; P_F\big(|\hat{\sigma}_n - 1| \geq \big(\epsilon \,/ \textstyle\int_{\mathbb{R}} H(x)dx\big)^{1/2} b(F)\big) + P_F(|\hat{\sigma}_n - 1| > h).$$

Again by Lemma 4.13 there is a $C_3(\epsilon,\delta) > 0$ such that for every $F \in M$ it is

$$P_F\big(|\hat{\sigma}_n - 1| \geq b(F)\big(\epsilon \,/ \textstyle\int_{\mathbb{R}} H(x)dx\big)^{1/2}\big) < \frac{\delta}{2}$$

for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C_3(\epsilon,\delta)/b(F)$. Moreover, it follows from Lemma 4.12 (ii) that there is an $n_2(\delta) \in \mathbb{N}$ with

$$\sup_{F \in M} P_F(|\hat{\sigma}_n - 1| > h) < \frac{\delta}{2} \quad \forall\; n > n_2(\delta). \tag{4.42}$$

For all $F \in M$ and $n \in \mathbb{N}$ such that $\sqrt{n} > \max\big(C_3(\epsilon,\delta), \sqrt{n_2(\delta)}\big)/b(F)$ we then have

$$P_F\big(\mathrm{II}_{n,2} \geq \epsilon\, b(F)^2\big) \;<\; \delta.$$

Next, we investigate

$$\mathrm{I}_n = \int_{\mathbb{R}} \big(F(x,\hat\sigma_n) - F_\tau(x)\big)^2 F_\tau(dx).$$

Since F_τ is differentiable on $\mathbb{R}$, the partial derivative of $F(\cdot,\cdot)$ with respect to σ exists for all $(x,\sigma) \in \mathbb{R}\times(0,\infty)$ with

$$\frac{\partial F}{\partial \sigma}(x,\sigma) = \Big(-\frac{x}{\sigma^2}\Big) f_\tau\Big(\frac{x}{\sigma}\Big).$$

By the mean value theorem for every $x \in \mathbb{R}$ there is a $\bar\xi_n = \bar\xi_n(x)$ between $\hat\sigma_n$ and 1 such that

$$F(x,\hat\sigma_n) - F_\tau(x) = F(x,\hat\sigma_n) - F(x,1) = \frac{\partial F}{\partial \sigma}(x,\bar\xi_n)\cdot\big(\hat\sigma_n - 1\big). \tag{4.43}$$

Now let $h \in (0,1)$ be as before. The partial derivative $\partial F/\partial\sigma$ is bounded on $\mathbb{R}\times[1-h,1+h]$ with

$$\Big|\frac{\partial F}{\partial \sigma}(x,\sigma)\Big| \;\le\; \frac{|x|}{(1-h)^2}\cdot\|f_\tau\|_\infty \quad \text{for } (x,\sigma)\in\mathbb{R}\times[1-h,1+h],$$

and $\|f_\tau\|_\infty$ is in $(0,\infty)$. On $\{|\hat\sigma_n - 1| \le h\}$ we have $(x,\bar\xi_n) \in \mathbb{R}\times[1-h,1+h]$ for all $x\in\mathbb{R}$, so that in this case it follows from the above that

$$\mathrm{I}_n = \int_{\mathbb{R}} \Big(\frac{\partial F}{\partial\sigma}(x,\bar\xi_n)\Big)^2 F_\tau(dx)\cdot\big(\hat\sigma_n-1\big)^2 \;\le\; \frac{\|f_\tau\|_\infty^2}{(1-h)^4}\cdot\int_{\mathbb{R}} x^2 F_\tau(dx)\cdot\big(\hat\sigma_n - 1\big)^2.$$

Thus, it is

$$\begin{aligned} P_F\big(\,\mathrm{I}_n \ge \epsilon\, b(F)^2\big) \;&\le\; P_F\big(\,\mathrm{I}_n \ge \epsilon\, b(F)^2, |\hat\sigma_n - 1| \le h\big) + P_F\big(|\hat\sigma_n - 1| > h\big)\\ &\le\; P_F\Big(|\hat\sigma_n - 1| \ge \frac{\epsilon^{1/2}\, b(F)(1-h)^2}{\|f_\tau\|_\infty \sigma_{F_\tau}}\Big) + P_F\big(|\hat\sigma_n - 1| > h\big). \end{aligned}$$

Using Lemma 4.13 once more, we see that there is a $C_4(\epsilon,\delta) > 0$ so that the first term on the right-hand side of the last inequality is less than $\delta/2$ for every $F \in M$ and $n\in\mathbb{N}$, $n\ge 2$, with $n > C_4(\epsilon,\delta)^2/b(F)^2$. Hence, for every $F\in M$ we have

$$P_F\big(\,\mathrm{I}_n \ge \epsilon\, b(F)^2\big) < \delta$$

if $n \in \mathbb{N}$ is such that $\sqrt{n} > \max\big(C_4(\epsilon,\delta), \sqrt{n_2(\delta)}\big)/b(F)$, where $n_2(\delta)$ is as in (4.42).

The last term to investigate is

$$\mathrm{III}_n = \int_{\mathbb{R}} \big(F_\tau(x) - F(x,\hat\sigma_n)\big)^2\cdot\big|f_\tau(x) - f(x,\hat\sigma_n)\big|dx.$$

Again, the mean value theorem assures that for every $x\in\mathbb{R}$ there is a $\tilde\xi_n = \tilde\xi_n(x)$ between $\hat\sigma_n$ and 1 such that

$$f_\tau(x) - f(x,\hat\sigma_n) = f(x,1) - f(x,\hat\sigma_n) = \frac{\partial f}{\partial\sigma}(x,\tilde\xi_n)\cdot\big(1-\hat\sigma_n\big).$$

Now consider $h\in(0,1)$ again. The partial derivative $\partial f/\partial\sigma$ is bounded on $\mathbb{R}\times[1-h,1+h]$ with

$$\Big|\frac{\partial f}{\partial\sigma}(x,\sigma)\Big| \;\le\; \frac{1}{(1-h)^2} f_\tau\Big(\frac{x}{1+h}\Big) + \frac{1}{(1-h)^2}\tau f_\tau\Big(\frac{x}{1+h}\Big)\Big|\frac{x}{1-h}\Big|^\tau$$

for $(x,\sigma)\in\mathbb{R}\times[1-h,1+h]$, cf. (4.38). Moreover, let $\bar\xi_n$ be as in (4.43). Then $|\hat\sigma_n - 1| \le h$ implies $(x,\bar\xi_n),(x,\tilde\xi_n)\in\mathbb{R}\times[1-h,1+h]$ for all $x\in\mathbb{R}$, and on the event $\{|\hat\sigma_n-1|\le h\}$ the following holds

$$\mathrm{III}_n = \int_{\mathbb{R}}\Big(\frac{\partial F}{\partial\sigma}(x,\bar\xi_n)\Big)^2\cdot\Big|\frac{\partial f}{\partial\sigma}(x,\tilde\xi_n)\Big|dx\cdot\big|\hat\sigma_n - 1\big|^3$$

$$
\begin{aligned}
&\le \bigg(\int_{\mathbb{R}} \frac{x^2}{(1-h)^6} f_\tau\Big(\frac{x}{1+h}\Big)dx + \tau\int_{\mathbb{R}} \frac{1}{(1-h)^4} f_\tau\Big(\frac{x}{1+h}\Big)\Big|\frac{x}{1-h}\Big|^{\tau+2}dx\bigg)\cdot\|f_\tau\|_\infty^2\cdot\big|\hat\sigma_n-1\big|^3\\
&=: \mathcal{I}(h)\cdot\|f_\tau\|_\infty^2\cdot\big|\hat\sigma_n-1\big|^3,
\end{aligned}
$$

where obviously both $\mathcal{I}(h)$ and $\|f_\tau\|_\infty\in(0,\infty)$. It follows that

$$
\begin{aligned}
P_F\big(\mathrm{III}_n\ge\epsilon\, b(F)^2\big) &\le P_F\big(\mathrm{III}_n\ge\epsilon\, b(F)^3\big)\\
&\le P_F\big(\mathrm{III}_n\ge\epsilon\, b(F)^3,\ |\hat\sigma_n-1|\le h\big)+P_F\big(|\hat\sigma_n-1|>h\big)\\
&\le P_F\Big(|\hat\sigma_n-1|\ge\frac{\epsilon^{1/3}b(F)}{(\mathcal{I}(h)\|f_\tau\|_\infty^2)^{1/3}}\Big)+P_F\big(|\hat\sigma_n-1|>h\big).
\end{aligned}
$$

Again, Lemma 4.13 ensures the existence of a $C_5(\epsilon,\delta)>0$ such that

$$
P_F\Big(|\hat\sigma_n-1|\ge\frac{\epsilon^{1/3}b(F)}{(\mathcal{I}(h)\|f_\tau\|_\infty^2)^{1/3}}\Big)<\frac{\delta}{2}
$$

for all $F\in M$ and $n\ge2$ with $\sqrt{n}>C_5(\epsilon,\delta)/b(F)$. Hence,

$$
P_F\big(\mathrm{III}_n\ge\epsilon\, b(F)^2\big)\ <\ \delta
$$

for every $F\in M$ and $n\in\mathbb{N}$ with $\sqrt{n}>\max\big(C_5(\epsilon,\delta),\sqrt{n_2(\delta)}\big)/b(F)$, where $n_2(\delta)$ is as in (4.42). Combining all of the above yields the assertion. □

We will now verify the statement of Theorem 4.11.

Proof of Theorem 4.11. Combining (4.33) and (4.34), we see that

$$
\Big|\frac{\widehat{\mathcal{W}}_n}{\sqrt{n}}-b(F)\Big|\ \le\ \|F_n-F\|_\infty+\big|b_n(F)-b(F)\big| \tag{4.44}
$$

for every $F\in\widetilde{\mathcal{G}}_q\setminus\{F_\tau\}$. Analogously we have

$$
\begin{aligned}
\Big|\frac{\widehat{\mathcal{V}}_n}{\sqrt{n}}-b(F)\Big| &\le \|\widetilde{F}_n-F\|_\infty+\big|b_n(F)-b(F)\big|\\
&\le \|\widetilde{F}_n-F_n\|_\infty+\|F_n-F\|_\infty+\big|b_n(F)-b(F)\big|
\end{aligned}
$$

for every $F\in\widetilde{\mathcal{G}}_q\setminus\{F_\tau\}$, where the first inequality is just (4.35).

As was shown in the proof of Theorem 4.5, the Kolmogorov-Smirnov statistic $\sqrt{n}\|F_n-F\|_\infty$ fulfills the assumptions of Proposition 2.4 for every $\varrho>0$.

Analogously to the proof of Theorem 4.5, let $K:=\int_{\mathbb{R}}|x|F_\tau(dx)/2\in(0,\infty)$. It follows from Lemma 4.1 (i) that there are $\delta_1,\delta_2>0$ such that

$$
\Big|\int_{\mathbb{R}}|x|^qF_\tau(dx)-\int_{\mathbb{R}}|x|^qF(dx)\Big|<K\ \text{ for all } F\in\widetilde{\mathcal{G}}_q \text{ with } d_q(F,F_\tau)<\delta_1 \tag{4.45}
$$

and

$$
\Big|\int_{\mathbb{R}}|x|F_\tau(dx)-\int_{\mathbb{R}}|x|F(dx)\Big|<K\ \text{ for all } F\in\widetilde{\mathcal{G}}_q \text{ with } d_q(F,F_\tau)<\delta_2. \tag{4.46}
$$

Recall that for $\tau>1$ we set $q=2\tau$ (>2), and for $\tau\le1$ we set $q>2$ $(\ge2\tau)$. Now since

$$
0<\int_{\mathbb{R}}|x|^qF(dx)<\int_{\mathbb{R}}|x|^qF_\tau(dx)+K<\infty
$$

for all $F \in U_{\delta_1}(F_\tau)$, see (4.45), we have for both of the above cases

$$0 < \sup_{F \in U_{\delta_1}(F_\tau)} \int_{\mathbb{R}} |x|^{2\tau} F(dx) < \infty. \tag{4.47}$$

Define $\varrho_1 := \delta_1$ and $M_1 := U_{\varrho_1}(F_\tau) \setminus \{F_\tau\}$. Then the set M_1 fulfills the assumption of Proposition 4.14, as we have just shown. Moreover, the Kolmogorov-Smirnov statistic obviously fulfills the assumptions of Proposition 2.4 for $\varrho = \varrho_1$. As a result of this, it follows with inequality (4.44) that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C_1(\epsilon, \delta) > 0$ such that

$$P_F\Big(\Big|\frac{\widehat{\mathcal{W}}_n}{\sqrt{n}} - b(F)\Big| \geq \epsilon b(F)\Big) \;\leq\; P_F\Big(\|F_n - F\|_\infty + \big|b_n(F) - b(F)\big| \geq \epsilon b(F)\Big) \;<\; \delta$$

for all $F \in M_1$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C_1(\epsilon, \delta)/b(F)$. Thus, we have verified Wieand's condition (WIII) for the sequence $(\widehat{\mathcal{W}}_n)_{n\geq 2}$.

Now set $K' := \min\big(F_\tau(0), 1 - F_\tau(0)\big)/2$. Then $K' > 0$, and by Lemma 4.1 (ii) there is a $\delta_3 > 0$ such that

$$d_K(F, F_\tau) = \sup_{x \in \mathbb{R}} |F(x) - F_\tau(x)| < K' \quad \text{for all } F \in \widetilde{\mathcal{G}}_q \text{ with } d_q(F, F_\tau) < \delta_3. \tag{4.48}$$

Let $\varrho_2 := \min(\delta_1, \delta_2, \delta_3)$ and $M_2 := U_{\varrho_2}(F_\tau) \setminus \{F_\tau\}$. Since $M_2 \subset U_{\delta_1}(F_\tau)$, (4.47) implies that the assumption of Proposition 4.14 holds for $M = M_2$. Additionally, the Kolmogorov-Smirnov statistic fulfills the assumptions of Proposition 2.4 on M_2, i.e., with $\varrho = \varrho_2$. Moreover, observe that the set M_2 is such that (3.7), (3.8) and (3.13) hold. The verification of (3.7) for M_2 follows along the same lines as in the proof of Theorem 4.5, using $q > 2$ and (4.45). Condition (3.8) is also easy to check since

$$\int_{\mathbb{R}} |x| F(dx) \;>\; K \;> 0$$

for every $F \in M_2$ because of (4.46), implying that $\inf\{\int_{\mathbb{R}} |x| F(dx) \colon F \in M_2\} \geq K > 0$. Furthermore, condition (3.13) is satisfied because

$$0 < F_\tau(0) - K' < F(0) < F_\tau(0) + K' < 1$$

for every $F \in M_2$ by (4.48).

Now we can handle the term $\|\widetilde{F}_n - F_n\|_\infty$ just as in the proof of Theorem 4.5. Combining all of the above, this shows that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C_2(\epsilon, \delta) > 0$ such that

$$P_F\Big(\Big|\frac{\widehat{\mathcal{V}}_n}{\sqrt{n}} - b(F)\Big| \geq \epsilon b(F)\Big) \;\leq\; P_F\Big(\|\widetilde{F}_n - F_n\|_\infty + \|F_n - F\|_\infty + \big|b_n(F) - b(F)\big| \geq \epsilon b(F)\Big) \;<\; \delta$$

for every $F \in M_2$ and every $n \geq 2$ such that $\sqrt{n} > C_2(\epsilon, \delta)/b(F)$, which is just (WIII) for the sequence $(\widehat{\mathcal{V}}_n)_{n\geq 2}$. □

We have now gathered all results that are necessary to show that the approximate Bahadur ARE of $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ with respect to $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ is equal to the limiting (as $\alpha \to 0$) Pitman ARE. This follows directly from Theorem 2.3 once we have verified that its assumptions hold. But as we have shown above, the sequences $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ and $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ fulfill conditions (BI), (BII) and (WIII). Moreover, the distribution functions of $\widehat{\mathcal{W}}$ in (4.27) and $\widehat{\mathcal{V}}$ in (4.28) are strictly increasing in their right tails, cf. Lemma 5.1 in Hörmann [15]. This shows that conditions (i) and (ii) of Theorem 2.3 hold. Observe that assumption (iii) is also satisfied, because $0 < b(F) \leq \|F - F_\tau\|_\infty = d_K(F, F_\tau)$ for all $F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$, and $d_K(F, F_\tau) \to 0$ as $d_q(F, F_\tau) \to 0$, $F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}$, by Lemma 4.1 (ii).

The remaining condition (iv) of Theorem 2.3 is again trivially fulfilled, as the ratio of approximate slopes does not depend on the alternative distribution anymore.

As a result of this, we get the following theorem.

Theorem 4.15
For every $\beta \in (0,1)$ it is

$$\lim_{\alpha \to 0} \liminf_{\substack{F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}, \\ d_q(F,F_\tau) \to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \lim_{\alpha \to 0} \limsup_{\substack{F \in \widetilde{\mathcal{G}}_q \setminus \{F_\tau\}, \\ d_q(F,F_\tau) \to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \frac{\tilde{\lambda}_1^*}{\tilde{\lambda}_1}. \tag{4.49}$$

As already mentioned, the eigenvalue $\tilde{\lambda}_1^*$ is strictly less than $\tilde{\lambda}_1$, see Example B.2 in the appendix. Thus, the limiting Pitman ARE of $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ with respect to $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ in (4.49) is strictly less than one, so that the test based on the latter sequence of test statistics is to be preferred.

For $\tau = 1$ and $\tau = 2$, we have computed the ratio $\tilde{\lambda}_1^*/\tilde{\lambda}_1$ explicitly:

It is shown in Example B.2 that $\tilde{\lambda}_1 = 1/\pi^2$ for every F_τ, $\tau > 0$. The value of $\tilde{\lambda}_1^*$, however, will vary with F_τ. For $\tau = 1$ and $\tau = 2$, i.e., for the double exponential distribution and for the normal distribution $\mathcal{N}(0, 1/2)$, the numerical computation of $\tilde{\lambda}_1^*$ is described in subsection 7.2 of [15]. The entries for $\tilde{\lambda}_1^*$ in the following table were computed using the R-function `eigenvalues.parmod()` of appendix A.2 in [15].

F_τ	$\tilde{\lambda}_1$	$\tilde{\lambda}_1^*$	$\tilde{\lambda}_1^*/\tilde{\lambda}_1$
Dexp ($\tau = 1$)	$1/\pi^2$	0.029837676	0.2944861
$\mathcal{N}(0,1/2)$ ($\tau = 2$)	$1/\pi^2$	0.01834741	0.1810817

Table 2: Values of $\tilde{\lambda}_1$ and $\tilde{\lambda}_1^*$ for $\tau = 1$ and $\tau = 2$.

5 Preparatory results for stable autoregressive models

In this section we will consider certain stable autoregressive processes with independent and identically distributed centered errors and discuss the classical empirical distribution function $F_{n,res}$ of the residuals and its centered version $\widetilde{F}_{n,res}$. Similar to section 3, we will study the asymptotic behavior of these functions and their components uniformly with respect to the distribution of the error variables. Moreover, the uniform behavior of the least squares estimator of the autoregressive parameter is investigated. We will use these results in the following section to verify Wieand's condition (WIII) for the Cramér-von Mises statistics based on $F_{n,res}$ and $\widetilde{F}_{n,res}$.

Let us now state the setting of this section. As before, let $M \neq \emptyset$ be a set of continuous distribution functions F that are centered and have finite second moments. Moreover, let $(\Omega, \mathcal{A})$ be a measurable space and $\{P_F \colon F \in M\}$ be a family of probability measures on $\mathcal{A}$ such that for some fixed $p \in \mathbb{N}$ there are random variables $S_0, S_{-1}, \ldots, S_{1-p}$ and e_i, $i \in \mathbb{Z}$, on $(\Omega, \mathcal{A})$ with the following properties:

- Under P_F the $(e_i)_{i\in\mathbb{Z}}$ are independent and identically distributed with common distribution function F for every $F \in M$.
- The variables $S_0, S_{-1}, \ldots, S_{1-p}$ are jointly distributed according to some fixed distribution Q, say, under every P_F, i.e.,
$$P_F \circ (S_0, S_{-1}, \ldots, S_{1-p})^{-1} = Q \quad \forall\ F \in M,$$
where the left-hand side denotes the joint distribution of $S_0, S_{-1}, \ldots, S_{1-p}$ under P_F. Furthermore, we assume that $S_0, S_{-1}, \ldots, S_{1-p}$ have finite second moments.
- The variables $S_0, S_{-1}, \ldots, S_{1-p}$ are independent of $(e_i)_{i\in\mathbb{N}}$ under every P_F.

Note that by Kolmogorov's consistency theorem such a model always exists. Moreover, under the above assumptions the mapping $F \mapsto P_F$ from M into the set of probability measures on $\mathcal{A}$ is injective.

Now let $\rho_1, \ldots, \rho_p$ be some real constants with $\rho_p \neq 0$ that satisfy

$$\{z \in \mathbb{C} \colon z^p - \rho_1 z^{p-1} - \rho_2 z^{p-2} - \ldots - \rho_{p-1} z - \rho_p = 0\} \ \subset\ \{z \in \mathbb{C} \colon |z| < 1\}, \tag{5.1}$$

which is equivalent to the condition $\{z \in \mathbb{C} \colon 1 - \rho_1 z - \ldots - \rho_p z^p = 0\} \subset \{z \in \mathbb{C} \colon |z| > 1\}$.

In this section we consider the following autoregressive models:

<u>Model 1</u>:
Let $(X_i)_{i\geq 1-p}$ be the autoregressive process of order p ($AR(p)$ for short) on $(\Omega, \mathcal{A})$ defined by

$$X_i = \rho_1 X_{i-1} + \ldots + \rho_p X_{i-p} + e_i, \quad i \geq 1, \tag{5.2}$$

with $\rho_1, \ldots, \rho_p \in \mathbb{R}$ as above and *starting values* $X_0 := S_0, \ldots, X_{1-p} := S_{1-p}$.

<u>Model 2</u>:
Let $(X_i)_{i\geq 1-p}$ be the stationary $AR(p)$ process on $(\Omega, \mathcal{A})$, i.e., $(X_i)_{i\geq 1-p}$ satisfies the model equation (5.2) and is strictly stationary under every P_F. It is well known that the stability condition (5.1) implies the existence of this stationary process, and that it is unique, cf. Remark 2 on page 86 in Brockwell and Davis [6]. The stationary process can be expressed as

$$X_i = \sum_{j=0}^{\infty} \psi_j e_{i-j} \quad \forall\ i \geq 1-p, \tag{5.3}$$

where the coefficients $\psi_j \in \mathbb{R}$ are uniquely determined, depend on $\rho_1, \ldots, \rho_p$ alone, and satisfy

$$|\psi_j| < K \cdot c^{-j}, \quad j \geq 0, \tag{5.4}$$

for some $K > 0$ and $c > 1$, cf. Theorem 3.1.1 in [6] and its proof. As follows from Proposition 3.1.1 in [6], the series in (5.3) converges in mean square as well as absolutely with probability one under every P_F, $F \in M$. The representation (5.3) is called the *MA*(∞)*-representation* of $(X_i)_{i\geq 1-p}$. Note that in contrast to the process of model 1, the distribution of the starting values $X_0, \ldots, X_{1-p}$ of the stationary process under the measure P_F does vary with F.

From now on, let $(X_i)_{i\geq 1-p}$ be either one of these $AR(p)$ processes. Since we want to study both of them simultaneously without having to differentiate between them, we will in general not make use of the stationarity of the process in the second model and derive all results by using the recursion formula (5.2) instead. Moreover, when explicitly referring to the process of model i, we will simply write process i, $i = 1, 2$.

When studying functionals such as the expectation E_F and the variance Var_F, the subscript F will denote as before that the respective term is understood to be with respect to the measure P_F. Hence, by the above assumptions it is $\mathrm{E}_F(e_i) = \mathrm{E}_F(e_1) = 0$ and $\mathrm{Var}_F(e_i) = \mathrm{Var}_F(e_1) = \mathrm{E}_F(e_1^2) =: \sigma_F^2 \in (0, \infty)$ for every $i \in \mathbb{Z}$ and $F \in M$.

It is useful and common practice to express the process $(X_i)_{i\geq 1-p}$ in matrix notation. In order to do this, we have to introduce some more notation. Set

$$\begin{aligned}
\mathbf{X}_i &:= \big(X_i, X_{i-1}, \ldots, X_{i-p+1}\big)^T \in \mathbb{R}^p \ \forall\, i \geq 0, \\
\mathbf{e}_i &:= \big(e_i, 0, \ldots, 0\big)^T \in \mathbb{R}^p \ \forall\, i \geq 1,
\end{aligned}$$

where x^T denotes the transpose of the vector or matrix x. Moreover, let

$$A := \begin{pmatrix}
\rho_1 & \rho_2 & \rho_3 & \cdots & \rho_{p-2} & \rho_{p-1} & \rho_p \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & & \cdots & & 1 & 0 & 0 \\
0 & & \cdots & & 0 & 1 & 0
\end{pmatrix} \in \mathbb{R}^{p\times p}.$$

Then the recursion formula (5.2) can be rewritten as

$$\mathbf{X}_i = A\mathbf{X}_{i-1} + \mathbf{e}_i, \quad i \geq 1.$$

By induction this implies that

$$\mathbf{X}_i = A^i \mathbf{X}_0 + \sum_{j=1}^{i} A^{j-1} \mathbf{e}_{i+1-j} \quad \forall\, i \geq 0. \tag{5.5}$$

Note that since the characteristic polynomial of A is given by

$$\det(A - zI_p) = (-1)^p (z^p - \rho_1 z^{p-1} - \ldots - \rho_{p-1} z - \rho_p),$$

the set $\{z \in \mathbb{C} \colon z^p - \rho_1 z^{p-1} - \rho_2 z^{p-2} - \ldots - \rho_{p-1} z - \rho_p = 0\}$ is just the set of eigenvalues of A. Thus, condition (5.1) states that the spectral radius of A is less than one, i.e.,

$$\max\{|z| \colon z \text{ is eigenvalue of } A\} < 1.$$

Because of this, there exists a matrix norm $\|\cdot\|_A$ such that $\|A\|_A < 1$, see e.g. Theorem 4.20 in Schott [28]. Since the transpose A^T of A has the same spectral radius as A, there is also a matrix norm $\|\cdot\|_{A^T}$ with $\|A^T\|_{A^T} < 1$. Moreover, for every real $n \times n$-matrix $B = (b_{jk})_{1\leq j,k\leq n}$ we will denote by $\|B\|_{\mathrm{Fr}}$ its Frobenius norm, i.e.,

$$\|B\|_{\mathrm{Fr}} = \sqrt{\sum_{1\leq j,k\leq n} b_{jk}^2}\,.$$

The Frobenius norm is compatible with the Euclidean norm $\|\cdot\|$ in $\mathbb{R}^n$, so that $\|Bx\| \leq \|B\|_{\mathrm{Fr}}\|x\|$ for all $B \in \mathbb{R}^{n\times n}$ and all $x \in \mathbb{R}^n$. Because of the equivalence of norms in finite-dimensional vector spaces there are positive numbers c_A and c_{A^T} such that

$$\|B\|_{\mathrm{Fr}} \;\leq\; c_A\,\|B\|_A \quad \text{and} \quad \|B\|_{\mathrm{Fr}} \;\leq\; c_{A^T}\,\|B\|_{A^T}$$

for every $B \in \mathbb{R}^{n\times n}$.

For the following considerations we need a generalization of the notions of Definition 3.1 to random vectors and random matrices.

Definition 5.1
Let $d \in \mathbb{N}$. Let E be either $\mathbb{R}^d$ or $\mathbb{R}^{d\times d}$, and $[\![\cdot]\!]$ be a norm on E. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers, and for every $F \in M$ let $(Y_{n,F})_{n\in\mathbb{N}}$ be a sequence of random elements on $(\Omega, \mathcal{A})$ with values in E. We say that $Y_{n,F} = o_P^u(a_n)$ in M as $n \to \infty$ if and only if $[\![Y_{n,F}]\!] = o_P^u(a_n)$ in M as $n \to \infty$. Analogously, we say that $Y_{n,F} = O_P^u(a_n)$ in M as $n \to \infty$ if and only if $[\![Y_{n,F}]\!] = O_P^u(a_n)$ in M as $n \to \infty$.

Note that since all norms in finite-dimensional vector spaces are equivalent, the notions defined above do not depend on the actual choice of $[\![\cdot]\!]$ and thus are well-defined.

Next, we will study the stochastic behavior of certain functions of the process $(X_i)_{i\geq 1-p}$ uniformly in $F \in M$. The following results are well known for stable autoregressive processes if $M = \{F\}$, i.e., if the distribution function F of the errors is fixed. We will investigate again under which assumptions these results hold uniformly in $F \in M$ if the set M contains arbitrarily many elements. For this, let us consider once more the following conditions:

$$\inf_{F\in M} \int_{\mathbb{R}} x^2 F(dx) > 0, \tag{3.5}$$

$$\sup_{F\in M} \int_{\mathbb{R}} x^2 F(dx) < \infty, \tag{3.6}$$

$$g(c) := \sup_{F\in M} \int_{\{x\in\mathbb{R}:\,|x|>c\}} x^2 F(dx) \to 0 \quad \text{for} \quad c \to \infty, \tag{3.7}$$

$$\inf_{F\in M} \int_{\mathbb{R}} |x| F(dx) > 0, \tag{3.8}$$

$$\sup_{F\in M} \int_{\mathbb{R}} |x| F(dx) < \infty. \tag{3.9}$$

Note again that (3.7) $\Longrightarrow$ (3.6) $\Longrightarrow$ (3.9) and (3.8) $\Longrightarrow$ (3.5).

Remark 5.2: If $\mathbf{X}_0$ is the vector of starting values of process 1, then $\mathrm{E}_F(\|\mathbf{X}_0\|) = \int_{\mathbb{R}^p} \|x\| Q(dx)$ and $\mathrm{E}_F(\|\mathbf{X}_0\|^2) = \int_{\mathbb{R}^p} \|x\|^2 Q(dx)$ for every $F \in M$, and both quantities are finite by assumption. Hence,

$$\sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|) < \infty \quad \text{and} \quad \sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|^2) < \infty$$

in this case. Using Markov's inequality, this implies that $\mathbf{X}_0 = O_P^u(1)$ in M as $n \to \infty$ and $\mathbf{X}_0 = o_P^u(a_n)$ in M as $n \to \infty$ for every positive real sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \underset{n}{\to} \infty$.

If $\mathbf{X}_0$ is the vector of starting values of process 2, then

$$\mathrm{E}_F(\|\mathbf{X}_0\|) \leq \mathrm{E}_F\Big(\sum_{i=1}^p |X_{1-i}|\Big) = p\,\mathrm{E}_F(|X_0|) = p \lim_{n\to\infty} \mathrm{E}_F\Big(\Big|\sum_{j=0}^n \psi_j e_{-j}\Big|\Big) \leq p\,\mathrm{E}_F(|e_1|) \sum_{j=0}^\infty |\psi_j| < \infty$$

and

$$\mathrm{E}_F(\|\mathbf{X}_0\|^2) = \sum_{i=1}^p \mathrm{E}_F(X_{1-i}^2) = p\,\mathrm{E}_F(X_0^2) = p\,\sigma_F^2 \sum_{j=0}^\infty \psi_j^2 < \infty$$

using the stationarity of $(X_i)_{i\geq 1-p}$, its $MA(\infty)$-representation (5.3) and inequality (5.4). Hence,

$$(3.9) \;\Rightarrow\; \sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|) < \infty \quad \text{and} \quad (3.6) \;\Rightarrow\; \sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|^2) < \infty.$$

It follows immediately from this and Markov's inequality that (3.9) implies $\mathbf{X}_0 = O_P^u(1)$ in M as $n \to \infty$ and $\mathbf{X}_0 = o_P^u(a_n)$ in M as $n \to \infty$ for every positive real sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \underset{n}{\to} \infty$. ♦

Lemma 5.3
If the set M is such that (3.7) *holds, then* $\max_{1\leq i\leq n} \|\mathbf{X}_{i-1}\| = o_P^u(\sqrt{n})$ *in M as $n \to \infty$.*

Proof. Because of $\|A\|_A < 1$ it is

$$K := \sum_{i=0}^\infty \|A\|_A^i = \frac{1}{1-\|A\|_A} \in (0,\infty).$$

By using (5.5) we see that for every $i \in \mathbb{N}$

$$\begin{aligned} \|\mathbf{X}_{i-1}\| \;&\leq\; \|A^{i-1}\mathbf{X}_0\| + \sum_{j=1}^{i-1} \|A^{j-1}\mathbf{e}_{i-j}\| \;\leq\; \|A^{i-1}\|_{\mathrm{Fr}}\, \|\mathbf{X}_0\| + \sum_{j=1}^{i-1} \|A^{j-1}\|_{\mathrm{Fr}}\, \|\mathbf{e}_{i-j}\| \\ &\leq\; c_A\|A^{i-1}\|_A\, \|\mathbf{X}_0\| + \sum_{j=1}^{i-1} c_A\|A^{j-1}\|_A\, |e_{i-j}| \\ &\leq\; c_A\|A\|_A^{i-1}\, \|\mathbf{X}_0\| + \sum_{j=1}^{i-1} c_A\|A\|_A^{j-1}\, |e_{i-j}|, \end{aligned}$$

where we used the sub-multiplicativity of $\|\cdot\|_A$ in the last inequality. Thus, it is for every $n \in \mathbb{N}$

$$\max_{1\leq i\leq n} \|\mathbf{X}_{i-1}\| \;\leq\; c_A\,\|\mathbf{X}_0\| + c_A \max_{1\leq i\leq n} |e_i|\, K.$$

Now $\|\mathbf{X}_0\| = o_P^u(\sqrt{n})$ in M by Remark 5.2, and since M satisfies (3.7) we have by Lemma 3.2 (i) that $\max_{1\leq i\leq n} |e_i| = o_P^u(\sqrt{n})$ in M as well, which yields the statement. □

Lemma 5.4
Assume the set M is such that (3.9) *holds. Then*

(i) $\sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_i\|) = O(1)$ *as $i \to \infty$,*

(ii) $\frac{1}{n}\sum_{i=1}^n \|\mathbf{X}_{i-1}\| = O_P^u(1)$ *in M as $n \to \infty$.*

Proof. (i) Recall that $\sum_{i=0}^{\infty} \|A\|_A^i =: K < \infty$. Using (5.5), we have for every $i \geq 0$ and $F \in M$

$$\begin{aligned} \mathrm{E}_F\big(\|\mathbf{X}_i\|\big) \;&\leq\; \mathrm{E}_F\Big(c_A\|A\|_A^i\|\mathbf{X}_0\| + \sum_{j=1}^{i} c_A\|A\|_A^{j-1}|e_{i+1-j}|\Big) \\ &\leq\; c_A\,\mathrm{E}_F(\|\mathbf{X}_0\|) + c_A K\,\mathrm{E}_F(|e_1|) \\ &\leq\; c_A \sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|) + c_A K \sup_{F\in M}\int_{\mathbb{R}} |x| F(dx) < \infty \end{aligned}$$

by the assumption and Remark 5.2, and the right-hand side does not depend on F anymore.

(ii) The statement follows with Markov's inequality and part (i). $\square$

Lemma 5.5
If the set M is such that (3.6) *holds, then*

(i) $\sup_{F\in M} \mathrm{E}_F\big(\|\mathbf{X}_i\|^2\big) = O(1)$ *as* $i \to \infty$,

(ii) $\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{X}_{i-1}\|^2 = O_P^u(1)$ *in* M *as* $n \to \infty$,

(iii) $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \mathbf{X}_{i-1}e_i = O_P^u(1)$ *in* M *as* $n \to \infty$,

(iv) $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \mathbf{X}_{i-1} = O_P^u(1)$ *in* M *as* $n \to \infty$.

Proof. Set $K := \sum_{i=0}^{\infty} \|A\|_A^i < \infty$ again.

(i) We have to show that the sequence $\big(\mathrm{E}_F(\|\mathbf{X}_i\|^2)\big)_{i\geq 0}$ is uniformly bounded in $F \in M$. To do this, we will investigate for every $F \in M$ the sequence of L_2 norms

$$\|\mathbf{X}_i\|_{L_2,F} = \big(\mathrm{E}_F(\|\mathbf{X}_i\|^2)\big)^{1/2}$$

instead. Using (5.5) we have

$$\|\mathbf{X}_i\|_{L_2,F} \;\leq\; \|A^i\mathbf{X}_0\|_{L_2,F} + \sum_{j=1}^{i} \|A^{j-1}\mathbf{e}_{i+1-j}\|_{L_2,F}$$

for every $i \geq 0$. But for all $i \geq 1$ it is

$$\mathrm{E}_F\big(\|A^i\mathbf{X}_0\|^2\big) \;\leq\; \|A^i\|_{\mathrm{Fr}}^2\,\mathrm{E}_F\big(\|\mathbf{X}_0\|^2\big) \;\leq\; c_A^2\,\|A\|_A^{2i}\;\mathrm{E}_F\big(\|\mathbf{X}_0\|^2\big) \;\leq\; c_A^2 \sup_{F\in M}\mathrm{E}_F\big(\|\mathbf{X}_0\|^2\big)$$

since $\|A\|_A < 1$. Because the right-hand side of the above display is finite by Remark 5.2 and does not depend on F anymore, this shows that the sequence $\big(\|A^i\mathbf{X}_0\|_{L_2,F}\big)_{i\geq 0}$ is uniformly bounded in F.

Moreover, for every $i \in \mathbb{N}$

$$\begin{aligned} \sum_{j=1}^{i} \|A^{j-1}\mathbf{e}_{i+1-j}\|_{L_2,F} &\leq \sum_{j=1}^{i} \|A^{j-1}\|_{\mathrm{Fr}}\|\mathbf{e}_{i+1-j}\|_{L_2,F} \leq c_A \sum_{j=1}^{i} \|A\|_A^{j-1}\,\mathrm{E}_F(e_1^2)^{1/2} \\ &\leq c_A K\Big(\sup_{F\in M}\int_{\mathbb{R}} x^2 F(dx)\Big)^{1/2} < \infty, \end{aligned}$$

and the statement follows.

(ii) Using Markov's inequality, the result follows from (i).

(iii) Let $C \in (0, \infty)$ and $F \in M$. Then

$$P_F\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\| \geq C\Big) \;\leq\; \frac{1}{C^2}\,\mathrm{E}_F\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\|^2\Big),$$

and using the fact that under every P_F, $F \in M$, the sequence of random vectors $(Z_i)_{i\geq 1}$, $Z_i := \mathbf{X}_{i-1}e_i$, is a square-integrable martingale difference sequence with respect to the filtration $\mathcal{F}_i := \sigma(X_{1-p}, \ldots, X_0, e_1, \ldots, e_i)$, $i \geq 1$, $\mathcal{F}_0 := \{\emptyset, \Omega\}$, it follows that

$$\begin{aligned}\mathrm{E}_F\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\|^2\Big) &= \frac{1}{n}\sum_{i=1}^n \mathrm{E}_F\big(\|\mathbf{X}_{i-1}e_i\|^2\big) = \frac{1}{n}\sum_{i=1}^n \mathrm{E}_F(\|\mathbf{X}_{i-1}\|^2)\,\mathrm{E}_F(e_1^2)\\ &\leq \frac{1}{n}\sum_{i=1}^n \sup_{F\in M}\mathrm{E}_F(\|\mathbf{X}_{i-1}\|^2)\cdot\sup_{F\in M}\int_{\mathbb{R}} x^2 F(dx).\end{aligned}$$

Using (i), this yields

$$\lim_{C\to\infty}\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\| \geq C\Big) = 0.$$

(iv) For every $n \in \mathbb{N}$ it follows from (5.5) that

$$\begin{aligned}\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}\Big\| &= \frac{1}{\sqrt{n}}\Big\|\sum_{i=1}^n\Big(A^{i-1}\mathbf{X}_0 + \sum_{j=1}^{i-1}A^{j-1}\mathbf{e}_{i-j}\Big)\Big\|\\ &\leq \frac{1}{\sqrt{n}}\sum_{i=1}^n\|A^{i-1}\mathbf{X}_0\| + \frac{1}{\sqrt{n}}\Big\|\sum_{i=1}^n\sum_{j=1}^{i-1}A^{j-1}\mathbf{e}_{i-j}\Big\|.\end{aligned}\tag{5.6}$$

Now it is

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\|A^{i-1}\mathbf{X}_0\| \leq \frac{1}{\sqrt{n}}\sum_{i=1}^n c_A\|A^{i-1}\|_A\|\mathbf{X}_0\| \leq \frac{1}{\sqrt{n}}c_A\sum_{i=1}^n\|A\|_A^{i-1}\|\mathbf{X}_0\| \leq \frac{1}{\sqrt{n}}c_A K\|\mathbf{X}_0\|,$$

and since $\|\mathbf{X}_0\| = O_P^u(1)$ in M as $n \to \infty$, cf. Remark 5.2, the right-hand side (and thus the left-hand side) of the above inequality is $O_P^u(n^{-1/2})$. Moreover, the second term in (5.6) is $O_P^u(1)$ because for every $n \in \mathbb{N}$, $C \in (0, \infty)$ and $F \in M$ it is

$$\begin{aligned}P_F\Big(\frac{1}{\sqrt{n}}\Big\|\sum_{i=1}^n\sum_{j=1}^{i-1}A^{j-1}\mathbf{e}_{i-j}\Big\| \geq C\Big) &= P_F\Big(\Big\|\sum_{r=1}^{n-1}\Big(\sum_{j=0}^{n-r-1}A^j\Big)\mathbf{e}_r\Big\| \geq \sqrt{n}C\Big)\\ &\leq \frac{1}{nC^2}\sum_{r=1}^{n-1}\mathrm{E}_F\Big(\Big\|\Big(\sum_{j=0}^{n-r-1}A^j\Big)\mathbf{e}_r\Big\|^2\Big) \leq \frac{1}{nC^2}\sum_{r=1}^{n-1}\Big\|\sum_{j=0}^{n-r-1}A^j\Big\|_{\mathrm{Fr}}^2\,\mathrm{E}_F(e_1^2)\\ &\leq \frac{1}{nC^2}\sum_{r=1}^{n-1}\Big(c_A\Big\|\sum_{j=0}^{n-r-1}A^j\Big\|_A\Big)^2\,\mathrm{E}_F(e_1^2) \leq \frac{c_A^2K^2}{C^2}\cdot\sup_{F\in M}\int_{\mathbb{R}}x^2F(dx).\end{aligned}$$

□

Let us consider solely the stationary process $(X_i)_{i\geq 1-p}$ from model 2 on page 53 for the moment. Since every $F \in M$ is centered, it follows that $\mathrm{E}_F(X_i) = \sum_{j=0}^\infty \psi_j\,\mathrm{E}_F(e_1) = 0$ for all $i \geq 1-p$, using the mean square convergence of the series in (5.3) and the continuity of the inner product. Now set

$$\mathbf{\Sigma} := \frac{1}{\sigma_F^2}\,\mathrm{E}_F\big(\mathbf{X}_0\mathbf{X}_0^T\big),$$

where $\mathrm{E}_F\big(\mathbf{X}_0\mathbf{X}_0^T\big)$ is the covariance matrix of $\mathbf{X}_0$ under P_F. Using the stationarity of the process and the recursion formula (5.5), we see that for every $k \in \mathbb{N}$

$$\sigma_F^2 \cdot \mathbf{\Sigma} = \mathrm{E}_F\big(\mathbf{X}_0\mathbf{X}_0^T\big) = \mathrm{E}_F\big(\mathbf{X}_k\mathbf{X}_k^T\big) = A^k\,\mathrm{E}_F\big(\mathbf{X}_0\mathbf{X}_0^T\big)(A^T)^k + \sum_{j=1}^{k} A^{j-1}\cdot\widetilde{E}\cdot(A^T)^{j-1}\cdot\sigma_F^2,$$

where $\widetilde{E}$ is the $p \times p$-matrix with a one in the upper left corner and zeros elsewhere. Since $\|A^k\|_A \le \|A\|_A^k \to 0$ as $k \to \infty$ and $\|(A^T)^k\|_{A^T} \le \|A^T\|_{A^T}^k \to 0$ as $k \to \infty$, both A^k and $(A^T)^k$ converge to the zero matrix as $k \to \infty$ in any norm on $\mathbb{R}^{p\times p}$. Hence,

$$\mathbf{\Sigma} = \sum_{j=1}^{\infty} A^{j-1}\cdot\widetilde{E}\cdot(A^T)^{j-1}, \tag{5.7}$$

where the series in (5.7) converges to $\mathbf{\Sigma}$ in any norm on $\mathbb{R}^{p\times p}$. It is evident from this representation that $\mathbf{\Sigma}$ is a function of $\rho_1,\dots,\rho_p$ alone and does not depend on F. Moreover, $\mathbf{\Sigma}$ is symmetric and positive definite, see e.g. Proposition 5.1.1 and Example 3.3.4 in [6].

Now let $(X_i)_{i\ge 1-p}$ be either one of the $AR(p)$ processes from model 1 or model 2 again. We are now ready to formulate and prove the following lemma.

Lemma 5.6
If the set M satisfies (3.7)*, then*

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{i-1}\mathbf{X}_{i-1}^T - \sigma_F^2\mathbf{\Sigma} = o_P^u(1) \quad \textit{in } M \textit{ as } n\to\infty.$$

Proof. Using (5.5) it is for every $n \in \mathbb{N}$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{i-1}\mathbf{X}_{i-1}^T &= \frac{1}{n}\sum_{i=1}^{n}A^{i-1}\mathbf{X}_0\mathbf{X}_0^T(A^T)^{i-1} + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}A^{i-1}\mathbf{X}_0\mathbf{e}_{i-j}^T(A^T)^{j-1}\\ &\quad+\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}A^{j-1}\mathbf{e}_{i-j}\mathbf{X}_0^T(A^T)^{i-1} + \frac{1}{n}\sum_{i=1}^{n}\Big(\sum_{j=1}^{i-1}A^{j-1}\mathbf{e}_{i-j}\Big)\Big(\sum_{j=1}^{i-1}A^{j-1}\mathbf{e}_{i-j}\Big)^T\\ &=: \mathrm{I}_n + \mathrm{II}_n + \mathrm{III}_n + \mathrm{IV}_n.\end{aligned}$$

Set $a := \max\{\|A\|_A, \|A^T\|_{A^T}\} \in (0,1)$. We will show first that $\mathrm{I}_n = o_P^u(1)$. For each $F \in M$ we have

$$\begin{aligned}\mathrm{E}_F\big(\|\mathrm{I}_n\|_{\mathrm{Fr}}\big) &\le \frac{1}{n}\,\mathrm{E}_F\Big(\sum_{i=1}^{n}\|A^{i-1}\mathbf{X}_0\mathbf{X}_0^T(A^T)^{i-1}\|_{\mathrm{Fr}}\Big) \le \frac{c_Ac_{A^T}}{n}\sum_{i=1}^{n}\mathrm{E}_F\big(\|A\|_A^{i-1}\|\mathbf{X}_0\mathbf{X}_0^T\|_{\mathrm{Fr}}\|A^T\|_{A^T}^{i-1}\big)\\ &= \frac{c_Ac_{A^T}}{n}\,\mathrm{E}_F\big(\|\mathbf{X}_0\mathbf{X}_0^T\|_{\mathrm{Fr}}\big)\sum_{i=1}^{n}\|A\|_A^{i-1}\|A^T\|_{A^T}^{i-1} \le \frac{c_Ac_{A^T}}{n}\sup_{F\in M}\mathrm{E}_F\big(\|\mathbf{X}_0\|^2\big)\sum_{i=0}^{\infty}a^{2i} < \infty\end{aligned}$$

using $\|\mathbf{X}_0\mathbf{X}_0^T\|_{\mathrm{Fr}} = \|\mathbf{X}_0\|^2$ and Remark 5.2. Thus, the statement follows with Markov's inequality.

Next, we investigate II_n. Let $F \in M$. As $\|\mathbf{X}_0\mathbf{e}_k^T\|_{\mathrm{Fr}} = \|\mathbf{X}_0\|\cdot|e_k|$ for every $k \in \mathbb{N}$, and $\mathbf{X}_0$ and $e_1,\dots,e_n$ are independent under P_F, it is

$$\mathrm{E}_F\big(\|\mathrm{II}_n\|_{\mathrm{Fr}}\big) \le \frac{1}{n}\,\mathrm{E}_F\Big(\sum_{i=1}^{n}\sum_{j=1}^{i-1}\|A^{i-1}\|_{\mathrm{Fr}}\|\mathbf{X}_0\mathbf{e}_{i-j}^T\|_{\mathrm{Fr}}\|(A^T)^{j-1}\|_{\mathrm{Fr}}\Big)$$

$$
\begin{aligned}
&= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{i-1} \|A^{i-1}\|_{\mathrm{Fr}} \, \mathrm{E}_F\big(\|\mathbf{X}_0\|\big) \, \mathrm{E}_F(|e_1|) \|(A^T)^{j-1}\|_{\mathrm{Fr}} \\
&\le \frac{c_A c_{A^T}}{n} \, \mathrm{E}_F\big(\|\mathbf{X}_0\|\big) \, \mathrm{E}_F(|e_1|) \sum_{i=1}^{n} \sum_{j=1}^{i-1} \|A\|_A^{i-1} \|A^T\|_{A^T}^{j-1} \\
&\le \frac{c_A c_{A^T}}{n} \cdot \sup_{F \in M} \mathrm{E}_F\big(\|\mathbf{X}_0\|\big) \cdot \sup_{F \in M} \int_{\mathbb{R}} |x| F(dx) \cdot \sum_{i=0}^{\infty} \|A\|_A^i \cdot \sum_{j=0}^{\infty} \|A^T\|_{A^T}^j < \infty
\end{aligned}
$$

using Remark 5.2. Again by Markov's inequality it follows that $\mathrm{II}_n = o_P^u(1)$ in M as $n \to \infty$. Since $\mathrm{III}_n = \mathrm{II}_n^T$, this immediately implies that $\mathrm{III}_n = o_P^u(1)$ as well.

It remains to show that $\mathrm{IV}_n - \sigma_F^2 \mathbf{\Sigma} = o_P^u(1)$. It is

$$
\begin{aligned}
\mathrm{IV}_n &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{i-1} A^{j-1} \mathbf{e}_{i-j} \mathbf{e}_{i-j}^T (A^T)^{j-1} + \frac{1}{n} \sum_{i=1}^{n} \sum_{1 \le j < m \le i-1} A^{j-1} \mathbf{e}_{i-j} \mathbf{e}_{i-m}^T (A^T)^{m-1} \\
&+ \frac{1}{n} \sum_{i=1}^{n} \sum_{1 \le m < j \le i-1} A^{j-1} \mathbf{e}_{i-j} \mathbf{e}_{i-m}^T (A^T)^{m-1} \\
&=: \mathrm{IV}_{1,n} + \mathrm{IV}_{2,n} + \mathrm{IV}_{3,n}.
\end{aligned}
$$

By changing the order of summation we have

$$
\mathrm{IV}_{2,n} = \frac{1}{n} \sum_{s=2}^{n-1} \sum_{t=1}^{s-1} \sum_{k=s+1}^{n} A^{k-s-1} \cdot \widetilde{E} \cdot (A^T)^{k-t-1} e_s e_t = \frac{1}{n} \sum_{s=2}^{n-1} \sum_{t=1}^{s-1} S_n(s,t) e_s e_t,
$$

say, where $\widetilde{E}$ is as before the matrix with a one in the upper left corner and zeros elsewhere. Since the variables $e_1, \dots, e_n$ are independent and centered under every P_F, it is

$$
\mathrm{E}_F\big(\|\mathrm{IV}_{2,n}\|_{\mathrm{Fr}}^2\big) = \frac{1}{n^2} \mathrm{E}_F(e_1^2)^2 \sum_{s=2}^{n-1} \sum_{t=1}^{s-1} \|S_n(s,t)\|_{\mathrm{Fr}}^2
$$

for each $F \in M$. Now for every $1 \le t < s \le n-1$ we have

$$
\begin{aligned}
\|S_n(s,t)\|_{\mathrm{Fr}} &\le \sum_{k=s+1}^{n} \|A^{k-s-1} \cdot \widetilde{E} \cdot (A^T)^{k-t-1}\|_{\mathrm{Fr}} \le c_A c_{A^T} \sum_{k=s+1}^{n} \|A\|_A^{k-s-1} \cdot \|\widetilde{E}\|_{\mathrm{Fr}} \cdot \|A^T\|_{A^T}^{k-t-1} \\
&\le c_A c_{A^T} \sum_{k=s+1}^{n} a^{2k-s-t-2} \le c_A c_{A^T} a^{s-t} \frac{1}{1-a^2},
\end{aligned}
$$

so that

$$
\sum_{s=2}^{n-1} \sum_{t=1}^{s-1} \|S_n(s,t)\|_{\mathrm{Fr}}^2 \le \Big(\frac{c_A c_{A^T}}{1-a^2}\Big)^2 \sum_{s=2}^{n-1} \sum_{t=1}^{s-1} (a^2)^{s-t} \le \frac{\big(c_A c_{A^T}\big)^2}{(1-a^2)^3} \cdot n.
$$

Hence,

$$
\mathrm{E}_F\big(\|\mathrm{IV}_{2,n}\|_{\mathrm{Fr}}^2\big) \le \frac{1}{n^2} \mathrm{E}_F(e_1^2)^2 \cdot \frac{\big(c_A c_{A^T}\big)^2}{(1-a^2)^3} \cdot n \le \frac{1}{n} \Big(\sup_{F \in M} \int_{\mathbb{R}} x^2 F(dx)\Big)^2 \cdot \frac{\big(c_A c_{A^T}\big)^2}{(1-a^2)^3}.
$$

This shows that $\mathrm{IV}_{2,n} = o_P^u(1)$. As $\mathrm{IV}_{3,n} = \mathrm{IV}_{2,n}^T$, it follows from this that $\mathrm{IV}_{3,n} = o_P^u(1)$ as well.

To complete the proof, it remains to show that $\mathrm{IV}_{1,n} - \sigma_F^2 \mathbf{\Sigma} = o_P^u(1)$ in M as $n \to \infty$. Using the representation (5.7), i.e.,

$$\mathbf{\Sigma} = \sum_{j=1}^{\infty} A^{j-1} \cdot \widetilde{E} \cdot (A^T)^{j-1},$$

we can write

$$\begin{aligned}
\mathrm{IV}_{1,n} &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{i-1} A^{j-1} \cdot \widetilde{E} \cdot (A^T)^{j-1} e_{i-j}^2 = \frac{1}{n} \sum_{i=1}^{n-1} \Big\{ \mathbf{\Sigma} - \sum_{j=n-i}^{\infty} A^j \cdot \widetilde{E} \cdot (A^T)^j \Big\} e_i^2 \\
&= \mathbf{\Sigma} \frac{1}{n} \sum_{i=1}^{n-1} e_i^2 - \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=n-i}^{\infty} A^j \cdot \widetilde{E} \cdot (A^T)^j e_i^2 \\
&=: \mathrm{IV}_{1,n}^{(i)} - \mathrm{IV}_{1,n}^{(ii)},
\end{aligned}$$

and we have for every $F \in M$

$$\begin{aligned}
\mathrm{E}_F\big(\|\mathrm{IV}_{1,n}^{(ii)}\|_{\mathrm{Fr}}\big) &\leq \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=n-i}^{\infty} \|A^j \cdot \widetilde{E} \cdot (A^T)^j\|_{\mathrm{Fr}} \, \mathrm{E}_F(e_1^2) \\
&\leq \mathrm{E}_F(e_1^2) \frac{c_A c_{A^T}}{n} \sum_{i=1}^{n-1} \sum_{j=n-i}^{\infty} \|A\|_A^j \cdot \|\widetilde{E}\|_{\mathrm{Fr}} \cdot \|A^T\|_{A^T}^j \\
&\leq \sup_{F \in M} \int_{\mathbb{R}} x^2 F(dx) \, \frac{c_A c_{A^T}}{n} \sum_{i=1}^{n-1} \sum_{j=n-i}^{\infty} a^{2j} \leq \sup_{F \in M} \int_{\mathbb{R}} x^2 F(dx) \, \frac{c_A c_{A^T}}{n(1-a^2)^2} \underset{n \to \infty}{\longrightarrow} 0,
\end{aligned}$$

so that $\mathrm{IV}_{1,n}^{(ii)} = o_P^u(1)$. But for $\mathrm{IV}_{1,n}^{(i)}$ we get

$$\mathrm{IV}_{1,n}^{(i)} - \sigma_F^2 \mathbf{\Sigma} = \mathbf{\Sigma} \frac{1}{n} \sum_{i=1}^{n} \big(e_i^2 - \sigma_F^2\big) - \mathbf{\Sigma} \frac{e_n^2}{n}.$$

Now recall that we have shown in Lemma 3.2 (iii) that

$$\frac{1}{n} \sum_{i=1}^{n} \big(e_i^2 - \sigma_F^2\big) = o_P^u(1) \quad \text{in } M \text{ as } n \to \infty.$$

Moreover, $e_n^2/n = o_P^u(1)$ in M as $n \to \infty$, as is easily seen by Markov's inequality. This concludes the proof of this lemma. □

Let now $\rho := \big(\rho_1, \ldots, \rho_p\big)^T \in \mathbb{R}^p$ be the autoregressive parameter in equation (5.2). Then we can write (5.2) as

$$X_i = \rho^T \mathbf{X}_{i-1} + e_i, \; i \geq 1.$$

Since ρ is assumed to be unknown, we have to estimate it by a sequence of estimators $(\hat{\rho}_n)_{n \in \mathbb{N}}$. We will always assume that this sequence is such that

$$\sqrt{n}\big(\hat{\rho}_n - \rho\big) = O_P^u(1) \quad \text{in } M \text{ as } n \to \infty. \tag{5.8}$$

One of the most prominent estimators for ρ is the *least squares estimator*

$$\hat{\rho}_n^{LS} = \hat{\rho}_n^{LS}(X_{1-p}, \ldots, X_n) := \Big(\sum_{i=1}^{n} \mathbf{X}_{i-1} \mathbf{X}_{i-1}^T\Big)^+ \sum_{i=1}^{n} \mathbf{X}_{i-1} X_i, \quad n \in \mathbb{N},$$

where B^+ denotes the Moore-Penrose pseudoinverse of the matrix B. Recall that if B is nonsingular, then B^+ and B^{-1} coincide.

We will show next that the sequence of least squares estimators satisfies condition (5.8) under certain assumptions.

Proposition 5.7
Let the set M be such that (3.7) *and* (3.5) *hold. Then*

$$\sqrt{n}(\hat{\rho}_n^{LS} - \rho) = O_P^u(1) \quad \text{in } M \text{ as } n \to \infty.$$

Proof. Set $\Lambda_n := \frac{1}{n}\sum_{i=1}^n \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T$. On the event $\{\det(\Lambda_n) \neq 0\} =: D_n$ we have

$$\hat{\rho}_n^{LS} = \Big(\sum_{i=1}^n \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T\Big)^{-1} \sum_{i=1}^n \mathbf{X}_{i-1}\big(\mathbf{X}_{i-1}^T\rho + e_i\big) = \rho + \Big(\sum_{i=1}^n \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T\Big)^{-1} \sum_{i=1}^n \mathbf{X}_{i-1}e_i,$$

so that

$$\sqrt{n}\big(\hat{\rho}_n^{LS} - \rho\big) = \Big(\frac{1}{n}\sum_{i=1}^n \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T\Big)^{-1} \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i = \Lambda_n^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i. \tag{5.9}$$

Let $GL(p)$ denote the general linear group of degree p over $\mathbb{R}$ and

$$\mathrm{inv}\colon GL(p) \ni B \mapsto B^{-1} \in GL(p)$$

be the inverse operator on $GL(p)$. Consider the matrix $\mathbf{\Sigma}$ from (5.7), which is positive definite and thus nonsingular. Because of the continuity of inv in $\mathbf{\Sigma}$ there is a $\delta = \delta(\mathbf{\Sigma}) > 0$ such that for all $B \in GL(p)$

$$\|B - \mathbf{\Sigma}\|_{\mathrm{Fr}} < \delta \quad \Rightarrow \quad \|\mathrm{inv}(B) - \mathrm{inv}(\mathbf{\Sigma})\|_{\mathrm{Fr}} < \|\mathrm{inv}\big(\mathbf{\Sigma}\big)\|_{\mathrm{Fr}}. \tag{5.10}$$

Set $v := 2\|\mathbf{\Sigma}^{-1}\|_{\mathrm{Fr}}/\inf_{F\in M}\sigma_F^2 \in (0,\infty)$. Now for every $C \in (0,\infty)$ and every $F \in M$ it is

$$P_F\big(\|\sqrt{n}\big(\hat{\rho}_n^{LS} - \rho\big)\| \geq C\big) \;\leq\; P_F\big(\{\|\sqrt{n}\big(\hat{\rho}_n^{LS} - \rho\big)\| \geq C\} \cap D_n\big) + P_F(\overline{D}_n) =: \mathrm{I}_{1,n} + \mathrm{I}_{2,n},$$

and with (5.9) we have

$$\begin{aligned}
\mathrm{I}_{1,n} &= P_F\big(\{\|\sqrt{n}\big(\hat{\rho}_n^{LS} - \rho\big)\| \geq C\} \cap D_n\big) \\
&\leq P_F\Big(\Big\{\Big\|\mathrm{inv}(\Lambda_n)\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\| \geq C\Big\} \cap D_n \cap \big\{\|\mathrm{inv}(\Lambda_n)\|_{\mathrm{Fr}} \geq v\big\}\Big) \\
&\quad + P_F\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\| \geq \frac{C}{v}\Big) \\
&\leq P_F\big(\{\|\mathrm{inv}(\Lambda_n)\|_{\mathrm{Fr}} \geq v\} \cap D_n\big) + P_F\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i\Big\| \geq \frac{C}{v}\Big).
\end{aligned}$$

As was shown in Lemma 5.5 (iii),

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_{i-1}e_i = O_P^u(1) \text{ in } M \text{ as } n \to \infty.$$

Moreover, using (5.10) we see that

$$P_F\big(\{\|\mathrm{inv}(\Lambda_n)\|_{\mathrm{Fr}} \geq v\} \cap D_n\big) = P_F\Big(\Big\{\Big\|\frac{1}{\sigma_F^2}\mathrm{inv}\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big)\Big\|_{\mathrm{Fr}} \geq v\Big\} \cap D_n\Big)$$

$$\begin{aligned}
&\le P_F\Big(\Big\{\Big\|\mathrm{inv}\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big)\Big\|_{\mathrm{Fr}} \ge v \cdot \inf_{F\in M} \sigma_F^2\Big\} \cap D_n\Big)\\
&\le P_F\Big(\Big\{\Big\|\mathrm{inv}\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big) - \mathrm{inv}(\mathbf{\Sigma})\Big\|_{\mathrm{Fr}} \ge \|\mathrm{inv}(\mathbf{\Sigma})\|_{\mathrm{Fr}}\Big\} \cap D_n\Big)\\
&\le P_F\Big(\Big\|\frac{1}{\sigma_F^2}\Lambda_n - \mathbf{\Sigma}\Big\|_{\mathrm{Fr}} \ge \delta\Big).
\end{aligned}$$

Now by Lemma 5.6 and (3.5) we have

$$\Big\|\frac{1}{\sigma_F^2}\Lambda_n - \mathbf{\Sigma}\Big\|_{\mathrm{Fr}} \le \frac{1}{\inf\limits_{F\in M} \sigma_F^2} \Big\|\frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T - \sigma_F^2\mathbf{\Sigma}\Big\|_{\mathrm{Fr}} = o_P^u(1). \tag{5.11}$$

Combining all this, it follows that

$$\lim_{C\to\infty} \limsup_{n\to\infty} \sup_{F\in M} \mathrm{I}_{1,n} = 0.$$

It remains to investigate $\mathrm{I}_{2,n}$. Since $\mathbf{\Sigma}$ is nonsingular, it is $\det(\mathbf{\Sigma}) \neq 0$. The continuity of the determinant function in $\mathbf{\Sigma}$ implies that there is an $\eta = \eta(\mathbf{\Sigma}) > 0$ such that for all $B \in \mathbb{R}^{p\times p}$

$$\|B - \mathbf{\Sigma}\|_{\mathrm{Fr}} < \eta \quad \Rightarrow \quad |\det(B) - \det(\mathbf{\Sigma})| < \frac{1}{2}|\det(\mathbf{\Sigma})|.$$

So for every $F \in M$ we have

$$\begin{aligned}
P_F(\overline{D}_n) &= P_F\Big(\det\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big) = 0\Big)\\
&= P_F\Big(\Big\{\det\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big) = 0\Big\} \cap \Big\{\Big|\det\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big) - \det(\mathbf{\Sigma})\Big| \ge \frac{1}{2}|\det(\mathbf{\Sigma})|\Big\}\Big)\\
&\quad + P_F\Big(\Big\{\det\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big) = 0\Big\} \cap \Big\{\Big|\det\Big(\frac{1}{\sigma_F^2}\Lambda_n\Big) - \det(\mathbf{\Sigma})\Big| < \frac{1}{2}|\det(\mathbf{\Sigma})|\Big\}\Big)\\
&\le P_F\Big(\Big\|\frac{1}{\sigma_F^2}\Lambda_n - \mathbf{\Sigma}\Big\|_{\mathrm{Fr}} \ge \eta\Big),
\end{aligned}$$

and by (5.11) this shows that

$$\limsup_{n\to\infty} \sup_{F\in M} P_F(\overline{D}_n) = 0. \qquad \square$$

Let $(\hat\rho_n)_{n\in\mathbb{N}}$ be any sequence of estimators for ρ now. Then the *residuals* $\hat e_{n1}, \dots, \hat e_{nn}$, $n \in \mathbb{N}$, with respect to this sequence of estimators are defined as

$$\hat e_{ni} := X_i - \hat\rho_n^T \mathbf{X}_{i-1}, \quad 1 \le i \le n,\ n \in \mathbb{N}.$$

We will from now on always work under the assumption that $\hat\rho_n$ is such that (5.8) is satisfied.

Lemma 5.8
If the set M is such that (3.7) *holds and the sequence of estimators $(\hat\rho_n)_{n\ge1}$ for ρ fulfills* (5.8), *then*

(i) $\sum\limits_{i=1}^{n} (\hat e_{ni} - e_i) = O_P^u(1)$ *in M as $n \to \infty$,*

(ii) $\max\limits_{1\le i\le n} |\hat e_{ni} - e_i| = o_P^u(1)$ *in M as $n \to \infty$,*

(iii) $\max\limits_{1\le i\le n} |\hat e_{ni}| = o_P^u(\sqrt{n})$ *in M as $n \to \infty$,*

(iv) $\displaystyle \frac{1}{n}\sum_{i=1}^{n}\big|\hat{e}_{ni}^2 - e_i^2\big| = o_P^u(1)$ *in* M *as* $n \to \infty$,

(v) $\displaystyle \frac{1}{n}\sum_{i=1}^{n}\hat{e}_{ni}^2 - \sigma_F^2 = o_P^u(1)$ *in* M *as* $n \to \infty$.

Proof. (i) We have

$$\sum_{i=1}^{n}\big(\hat{e}_{ni} - e_i\big) = -\big(\hat{\rho}_n^T - \rho^T\big)\sum_{i=1}^{n}\mathbf{X}_{i-1} = -\sqrt{n}\big(\hat{\rho}_n - \rho\big)^T\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{X}_{i-1} = O_P^u(1)O_P^u(1) = O_P^u(1)$$

because of (5.8) and Lemma 5.5 (iv).

(ii) Using the Cauchy-Schwarz inequality, (5.8) and Lemma 5.3, it is

$$\max_{1\le i\le n}|\hat{e}_{ni} - e_i| = \max_{1\le i\le n}\big|(\hat{\rho}_n - \rho)^T\mathbf{X}_{i-1}\big| \le \|\hat{\rho}_n - \rho\|\max_{1\le i\le n}\|\mathbf{X}_{i-1}\| = O_P^u\Big(\frac{1}{\sqrt{n}}\Big)\cdot o_P^u(\sqrt{n}) = o_P^u(1).$$

(iii) By the previous result and part (i) of Lemma 3.2 we have

$$\max_{1\le i\le n}|\hat{e}_{ni}| \le \max_{1\le i\le n}|\hat{e}_{ni} - e_i| + \max_{1\le i\le n}|e_i| = o_P^u(1) + o_P^u(\sqrt{n}) = o_P^u(\sqrt{n}).$$

(iv) It is

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\big|\hat{e}_{ni}^2 - e_i^2\big| &= \frac{1}{n}\sum_{i=1}^{n}\big|\hat{e}_{ni} - e_i\big|\cdot\big|\hat{e}_{ni} + e_i\big| \le \max_{1\le i\le n}\big|\hat{e}_{ni} - e_i\big|\cdot\Big(\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni} - e_i| + \frac{2}{n}\sum_{i=1}^{n}|e_i|\Big)\\ &\le \max_{1\le i\le n}\big|\hat{e}_{ni} - e_i\big|\cdot\Big(\max_{1\le i\le n}\big|\hat{e}_{ni} - e_i\big| + \frac{2}{n}\sum_{i=1}^{n}|e_i|\Big).\end{aligned}$$

Now it obviously is $\sum_{i=1}^{n}|e_i| = O_P^u(n)$ since

$$P_F\Big(\frac{1}{n}\sum_{i=1}^{n}|e_i| > C\Big) \le \frac{1}{nC}\sum_{i=1}^{n}\mathrm{E}_F(|e_1|) \le \frac{1}{C}\sup_{F\in M}\int_{\mathbb{R}}|x|F(dx) < \infty$$

for every $C \in (0,\infty)$ and $F \in M$. Thus, the result follows with part (ii).

(v) Since

$$\Big|\frac{1}{n}\sum_{i=1}^{n}\hat{e}_{ni}^2 - \sigma_F^2\Big| \le \frac{1}{n}\sum_{i=1}^{n}\big|\hat{e}_{ni}^2 - e_i^2\big| + \Big|\frac{1}{n}\sum_{i=1}^{n}e_i^2 - \sigma_F^2\Big|,$$

the result follows from (iv) and Lemma 3.2 (iii). □

5.1 The empirical distribution function of the residuals

Let $(\hat{\rho}_n)_{n\in\mathbb{N}}$ be a sequence of estimators for the autoregressive parameter $\rho = (\rho_1, \dots, \rho_p)^T$, and for every $n \in \mathbb{N}$ let

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{e_i \le x\}}, \quad x \in \mathbb{R},$$

be the empirical distribution function of the errors $e_1, \dots, e_n$. In analogy to this, set

$$F_{n,res}(x) := \frac{1}{n}\sum_{i=1}^{n} 1_{\{\hat{e}_{ni} \le x\}} = \frac{1}{n}\sum_{i=1}^{n} 1_{\{e_i \le x + (\hat{\rho}_n - \rho)^T \mathbf{X}_{i-1}\}}, \quad x \in \mathbb{R}.$$

The function $F_{n,res}$ is called the *empirical distribution function of the residuals* $\hat{e}_{n1}, \dots, \hat{e}_{nn}$.

In this subsection we will assume that

$$M \subset \Big\{F\colon F \text{ is a distribution function that has uniformly continuous Lebesgue density } f \\ \text{and satisfies} \int_{\mathbb{R}} x^2 F(dx) < \infty \text{ and } \int_{\mathbb{R}} x F(dx) = 0\Big\}.$$

The uniformly continuous Lebesgue density f of F will also be denoted by F'. For the following investigations we will impose the following assumptions on the set M:

$$\{F'\colon F \in M\} \text{ is uniformly equicontinuous,} \tag{5.12}$$

$$\sup_{F\in M} F'(x) \underset{|x|\to\infty}{\longrightarrow} 0, \tag{5.13}$$

$$\sup_{F\in M} \|F'\|_\infty < \infty. \tag{5.14}$$

We are now interested in the asymptotic stochastic behavior of $\sqrt{n}\|F_n - F_{n,res}\|_\infty$ uniformly with respect to the set of underlying probability measures $\{P_F\colon F \in M\}$. An answer to this gives the next theorem. Its proof is based on the one for the classical, non-uniform case with a fixed underlying probability measure P_F, $F \in M$, which is described in a more general setting in section 7.2 of Koul [18].

Theorem 5.9
Suppose that the set M satisfies conditions (5.12), (5.13) and (5.14). Moreover, assume that

(i) $\sqrt{n}(\hat{\rho}_n - \rho) = O_P^u(1)$ *in M as $n \to \infty$,*

(ii) $\frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_{i-1} = o_P^u(1)$ *in M as $n \to \infty$,*

(iii) $\max_{1\le i\le n} \|\mathbf{X}_{i-1}\| = o_P^u(\sqrt{n})$ *in M as $n \to \infty$,*

(iv) $\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{X}_{i-1}\| = O_P^u(1)$ *in M as $n \to \infty$,*

(v) $\sup_{F\in M} \mathrm{E}_F\Big(\sum_{i=1}^{n} \|\mathbf{X}_{i-1}\|\Big) = o(n^{3/2})$ *as $n \to \infty$.*

Then

$$\sqrt{n}\sup_{x\in\mathbb{R}} |F_n(x) - F_{n,res}(x)| = o_P^u(1) \quad \text{in } M \text{ as } n \to \infty.$$

Note that condition (i) is just (5.8), and as (3.7) ⇒ (3.6) ⇒ (3.9), it follows with Lemma 5.3, Lemma 5.4 and part (iv) of Lemma 5.5 that (3.7) implies conditions (ii)–(v).

The proof of Theorem 5.9 uses a more general result, which is formulated in the following theorem. Hence, we will first prove the next result, and give the proof of the above theorem thereafter. The proof of the theorem below is based again on the ideas of the proof in the non-uniform case, as described in subsection 2.2.2 of Koul [18]. Moreover, it uses a typical chaining argument as discussed for example in Pollard [26], pages 160–162.

Theorem 5.10
Let M be a nonempty set of distribution functions such that all $F \in M$ possess a uniformly continuous Lebesgue density $F' = f$ and (5.12), (5.13) and (5.14) hold.
Moreover, let $(\Omega, \mathcal{A})$ be a measurable space and $\{P_F : F \in M\}$ be a family of probability measures on $\mathcal{A}$. Let $e_1, e_2, \dots$ be random variables and Y be a random element on $(\Omega, \mathcal{A})$ such that for each $F \in M$ the variables $(e_i)_{i\in\mathbb{N}}$ are independent and identically distributed with common distribution function F under P_F and Y is independent of $(e_i)_{i\in\mathbb{N}}$ under P_F. Set

$$\mathcal{F}_0 := \sigma(Y), \quad \mathcal{F}_n := \sigma(Y, e_1, \dots, e_n), \quad n \geq 1.$$

For all $n \in \mathbb{N}$ let $\delta_{n1}, \dots, \delta_{nn}$ be random variables on $(\Omega, \mathcal{A})$ with

(i) $\delta_{n1}, \dots, \delta_{nn}$ *is predictable with respect to* $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_n$,

(ii) $\sup\limits_{F\in M} \mathrm{E}_F\big(\frac{1}{n}\sum_{i=1}^n |\delta_{ni}|\big) = o(1)$ *for* $n \to \infty$,

(iii) $\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| = O_P^u(1)$ *in M as* $n \to \infty$,

(iv) $\max\limits_{1\leq i\leq n} |\delta_{ni}| = o_P^u(1)$ *in M as* $n \to \infty$.

Then

$$\sup_{x\in\mathbb{R}} \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n \big[1_{\{e_i \leq x+\delta_{ni}\}} - F(x+\delta_{ni})\big] - \sum_{i=1}^n \big[1_{\{e_i\leq x\}} - F(x)\big]\Big| = o_P^u(1) \quad \text{in } M \text{ as } n \to \infty.$$

Proof. Let G be a continuously differentiable, strictly increasing distribution function. For every $x, y \in \mathbb{R}$ and every $F \in M$ set

$$d_F(x,y) := |F(x) - F(y)|^{1/2}, \qquad d_G(x,y) := |G(x) - G(y)|^{1/2},$$

and

$$\bar{d}_F(x,y) := d_F(x,y) + d_G(x,y) = |F(x)-F(y)|^{1/2} + |G(x)-G(y)|^{1/2}.$$

It is easy to see that $\bar{d}_F$ is symmetric and fulfills the triangle inequality, so that $\bar{d}_F$ is a pseudo-metric. For every $F \in M$ define

$$\omega_F(\delta) := \sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\leq\delta}} |f(x) - f(y)|, \quad \delta > 0. \tag{5.15}$$

Note that because of (5.14) the function ω_F is uniformly bounded. Since

$$\omega_F(\delta) \leq \sup_{\substack{x,y\in\mathbb{R}\\ d_G(x,y)\leq\delta}} |f(x)-f(y)| = \sup_{\substack{x,y\in\mathbb{R}\\ |G(x)-G(y)|\leq\delta^2}} |f(x)-f(y)|$$

for every $F \in M$, it follows from Proposition A.3 that

$$\lim_{\delta\downarrow 0} \sup_{F\in M} \omega_F(\delta) = 0. \tag{5.16}$$

We will now prove the statement of the theorem in several steps.

Step 1:

We will show first that for every $\epsilon > 0$

$$\sup_{F\in M}\sup_{x\in\mathbb{R}} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big[1_{\{e_i\leq x+\delta_{ni}\}} - F(x+\delta_{ni})\big] - \sum_{i=1}^{n}\big[1_{\{e_i\leq x\}} - F(x)\big]\Big| \geq \epsilon\Big) \underset{n\to\infty}{\longrightarrow} 0. \quad (5.17)$$

Proof of Step 1. First note that for every $F \in M$, $x \in \mathbb{R}$, $n \in \mathbb{N}$ and $1 \leq i \leq n$

$$\mathrm{E}_F(1_{\{e_i\leq x+\delta_{ni}\}}|\mathcal{F}_{i-1}) = \mathrm{E}_F(1_{\{e_i\leq x+\delta_{ni}\}}|Y=\cdot, e_1=\cdot,\ldots,e_{i-1}=\cdot)\circ(Y,e_1,\ldots,e_{i-1}).$$

Moreover, by the factorization lemma there is a measurable function g_{ni} such that $\delta_{ni} = g_{ni}(Y, e_1, \ldots, e_{i-1})$. Using this and the independence of e_i and $Y, e_1, \ldots, e_{i-1}$ under P_F, it is

$$\begin{aligned}&\mathrm{E}_F(1_{\{e_i\leq x+\delta_{ni}\}}|Y=y, e_1=x_1,\ldots,e_{i-1}=x_{i-1})\\ &= \mathrm{E}_F(1_{\{e_i\leq x+g_{ni}(Y,e_1,\ldots,e_{i-1})\}}|Y=y, e_1=x_1,\ldots,e_{i-1}=x_{i-1}) = F(x+g_{ni}(y,x_1,\ldots,x_{i-1})).\end{aligned}$$

Thus,

$$\mathrm{E}_F(1_{\{e_i\leq x+\delta_{ni}\}}|\mathcal{F}_{i-1}) = F(x+g_{ni}(Y,e_1,\ldots,e_{i-1})) = F(x+\delta_{ni}). \quad (5.18)$$

Because of (5.18) the random variables

$$\zeta_i := 1_{\{e_i\leq x+\delta_{ni}\}} - F(x+\delta_{ni}) - \big[1_{\{e_i\leq x\}} - F(x)\big], \quad 1\leq i\leq n,$$

form a martingale difference sequence (MDS) with respect to $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n$ under P_F. Therefore we have for every $\epsilon > 0$

$$\begin{aligned}&P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big[1_{\{e_i\leq x+\delta_{ni}\}} - F(x+\delta_{ni})\big] - \sum_{i=1}^{n}\big[1_{\{e_i\leq x\}} - F(x)\big]\Big| \geq \epsilon\Big)\\ &= P_F\Big(\Big|\sum_{i=1}^{n}\zeta_i\Big| \geq \epsilon\sqrt{n}\Big) \leq \frac{1}{\epsilon^2 n}\sum_{i=1}^{n}\mathrm{E}_F(\zeta_i^2). \quad (5.19)\end{aligned}$$

In order to handle the expectations $\mathrm{E}_F(\zeta_i^2)$, we will use the following result, which is proven here in a more general form for later use.

Let $x, y \in \mathbb{R}$, $n \in \mathbb{N}$ and $1 \leq i \leq n$. Let $\tilde{\delta}_{ni}$ be a random variable on $(\Omega, \mathcal{A})$ that is measurable with respect to $\mathcal{F}_{i-1}$. Then

$$\begin{aligned}&\mathrm{E}_F\Big(\big(1_{\{e_i\leq x+\delta_{ni}\}} - F(x+\delta_{ni}) - \big[1_{\{e_i\leq y+\tilde{\delta}_{ni}\}} - F(y+\tilde{\delta}_{ni})\big]\big)^2|\mathcal{F}_{i-1}\Big)\\ &\leq |F(x+\delta_{ni}) - F(y+\tilde{\delta}_{ni})|. \quad (5.20)\end{aligned}$$

To prove (5.20), set

$$\xi := 1_{\{e_i\leq x+\delta_{ni}\}} - 1_{\{e_i\leq y+\tilde{\delta}_{ni}\}}.$$

Then with (5.18), which holds of course for $\tilde{\delta}_{ni}$ as well, we have

$$\begin{aligned}&\mathrm{E}_F\Big(\big(1_{\{e_i\leq x+\delta_{ni}\}} - F(x+\delta_{ni}) - \big[1_{\{e_i\leq y+\tilde{\delta}_{ni}\}} - F(y+\tilde{\delta}_{ni})\big]\big)^2|\mathcal{F}_{i-1}\Big)\\ &= \mathrm{E}_F\big([\xi - \mathrm{E}_F(\xi|\mathcal{F}_{i-1})]^2|\mathcal{F}_{i-1}\big) = \mathrm{E}_F\big(\xi^2|\mathcal{F}_{i-1}\big) - \mathrm{E}_F\big(\xi|\mathcal{F}_{i-1}\big)^2\\ &\leq \mathrm{E}_F\big(\xi^2|\mathcal{F}_{i-1}\big) = \mathrm{E}_F\big(|1_{\{e_i\leq x+\delta_{ni}\}} - 1_{\{e_i\leq y+\tilde{\delta}_{ni}\}}|\,|\mathcal{F}_{i-1}\big)\\ &= \mathrm{E}_F\big(1_{\{(x+\delta_{ni})\wedge(y+\tilde{\delta}_{ni}) < e_i \leq (x+\delta_{ni})\vee(y+\tilde{\delta}_{ni})\}}|\mathcal{F}_{i-1}\big)\end{aligned}$$

$$= F\big((x+\delta_{ni}) \vee (y+\tilde{\delta}_{ni})\big) - F\big((x+\delta_{ni}) \wedge (y+\tilde{\delta}_{ni})\big)$$
$$= |F(x+\delta_{ni}) - F(y+\tilde{\delta}_{ni})|,$$

where the second-to-last equality is shown along the same lines as (5.18).

Now we come back to (5.19). Using (5.20) with $y = x$ and $\tilde{\delta}_{ni} = 0$, we see that

$$\frac{1}{\epsilon^2 n}\sum_{i=1}^n \mathrm{E}_F(\zeta_i^2) \le \frac{1}{\epsilon^2 n}\sum_{i=1}^n \mathrm{E}_F\big(|F(x+\delta_{ni}) - F(x)|\big) \le \frac{1}{\epsilon^2 n}\sum_{i=1}^n \|f\|_\infty \,\mathrm{E}_F\big(|\delta_{ni}|\big)$$
$$\le \frac{1}{\epsilon^2}\sup_{F\in M}\|f\|_\infty \sup_{F\in M}\mathrm{E}_F\Big(\frac{1}{n}\sum_{i=1}^n |\delta_{ni}|\Big).$$

Now (5.17) follows with (5.14) and (ii).

Step 2:

For every $n \in \mathbb{N}$, $x \in \mathbb{R}$ and $F \in M$ set

$$U_{n,F}(x) := \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le x+\delta_{ni}\}} - F(x+\delta_{ni})\big] \quad \text{and} \quad V_{n,F}(x) := \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le x\}} - F(x)\big].$$

We will show now that the condition

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\le\delta}} |U_{n,F}(x) - U_{n,F}(y)| \ge \epsilon\Big) = 0 \;\; \forall \epsilon > 0 \tag{5.21}$$

implies the statement of the theorem.

Proof of Step 2. Assume that (5.21) holds. Then by setting $\delta_{ni} = 0$ for every $n \in \mathbb{N}$ and $1 \le i \le n$ it follows directly that

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\le\delta}} |V_{n,F}(x) - V_{n,F}(y)| \ge \epsilon\Big) = 0 \;\; \forall \epsilon > 0. \tag{5.22}$$

Let $k \in \mathbb{N}$ and $F \in M$. Because of the continuity of F, for every $j \in \{1,\dots,k-1\}$ there is an $x_j \in \mathbb{R}$ such that $F(x_j) = j/k$. Moreover we have $-\infty < x_1 < \ldots < x_{k-1} < \infty$ because of the monotonicity of F. The same arguments ensure the existence of $-\infty < y_1 < \ldots < y_{k-1} < \infty$ with $G(y_j) = j/k$, $j = 1,\dots,k-1$. Let $z_1,\dots,z_l$ be the common refinement of these two partitions, i.e., $\{z_1,\dots,z_l\} = \{x_1,\dots,x_{k-1}\} \cup \{y_1,\dots,y_{k-1}\}$ and $-\infty < z_1 < \ldots < z_l < \infty$. Obviously, $l \le 2(k-1)$. Note that since $|F(x_{j+1}) - F(x_j)| = 1/k$ and $|G(y_{j+1}) - G(y_j)| = 1/k$ for every $j = 1,\dots,k-2$ by construction and $z_1,\dots,z_l$ is a refinement of the two partitions, it follows from the monotonicity of F and G that

$$|F(z_{j+1}) - F(z_j)| \le 1/k, \quad |G(z_{j+1}) - G(z_j)| \le 1/k \quad \forall j = 1,\dots,l-1. \tag{5.23}$$

We will show now that

$$\text{for every } x \in \mathbb{R} \text{ there is a } j(x) \in \{1,\dots,l\} \text{ such that } \bar{d}_F(x, z_{j(x)}) \le \frac{2}{\sqrt{k}}. \tag{5.24}$$

For the proof of (5.24), let $x \in \mathbb{R}$. We investigate the following cases:
Case 1: $-\infty < x \le z_1$. Since $z_1 = x_1 \wedge y_1$, it follows by the monotonicity of F and G that

$$\bar{d}_F(x, z_1) = |F(x) - F(z_1)|^{1/2} + |G(x) - G(z_1)|^{1/2} \le F(z_1)^{1/2} + G(z_1)^{1/2}$$

$$\leq F(x_1)^{1/2} + G(y_1)^{1/2} = \frac{2}{\sqrt{k}}.$$

Case 2: $z_1 < x \leq z_l$. Then there is a $j \in \{1, \dots, l-1\}$ with $z_j < x \leq z_{j+1}$. Using (5.23) we see that

$$\bar{d}_F(x, z_{j+1}) \leq |F(z_j) - F(z_{j+1})|^{1/2} + |G(z_j) - G(z_{j+1})|^{1/2} \leq \frac{2}{\sqrt{k}}.$$

Case 3: $z_l < x < \infty$. Then $z_l = x_{k-1} \vee y_{k-1}$, and

$$\bar{d}_F(x, z_l) \leq |1 - F(z_l)|^{1/2} + |1 - G(z_l)|^{1/2} \leq |1 - F(x_{k-1})|^{1/2} + |1 - G(y_{k-1})|^{1/2} = \frac{2}{\sqrt{k}}.$$

This shows (5.24).

Thus, for every $F \in M$, $x \in \mathbb{R}$ and $n, k \in \mathbb{N}$ we have

$$\begin{aligned}
A_{n,F}(x) &:= \frac{1}{\sqrt{n}} \Big| \sum_{i=1}^{n} \big[1_{\{e_i \leq x + \delta_{ni}\}} - F(x + \delta_{ni}) \big] - \sum_{i=1}^{n} \big[1_{\{e_i \leq x\}} - F(x) \big] \Big| = \big| U_{n,F}(x) - V_{n,F}(x) \big| \\
&\leq \big| U_{n,F}(x) - U_{n,F}(z_{j(x)}) \big| + \big| U_{n,F}(z_{j(x)}) - V_{n,F}(z_{j(x)}) \big| + \big| V_{n,F}(z_{j(x)}) - V_{n,F}(x) \big| \\
&\leq \sup_{\substack{x,y \in \mathbb{R} \\ \bar{d}_F(x,y) \leq 2/\sqrt{k}}} \big| U_{n,F}(x) - U_{n,F}(y) \big| + \max_{1 \leq j \leq l} \big| U_{n,F}(z_j) - V_{n,F}(z_j) \big| \\
&\quad + \sup_{\substack{x,y \in \mathbb{R} \\ \bar{d}_F(x,y) \leq 2/\sqrt{k}}} \big| V_{n,F}(x) - V_{n,F}(y) \big|,
\end{aligned}$$

where $j(x)$ is as in (5.24). But for every $k \in \mathbb{N}$ it is

$$\begin{aligned}
&\sup_{F \in M} P_F \Big(\max_{1 \leq j \leq l} \big| U_{n,F}(z_j) - V_{n,F}(z_j) \big| \geq \epsilon \Big) \leq \sum_{j=1}^{l} \sup_{F \in M} P_F \Big(\big| U_{n,F}(z_j) - V_{n,F}(z_j) \big| \geq \epsilon \Big) \\
&\leq \sum_{j=1}^{k-1} \sup_{F \in M} P_F \Big(\big| U_{n,F}(x_j) - V_{n,F}(x_j) \big| \geq \epsilon \Big) + \sum_{j=1}^{k-1} \sup_{F \in M} P_F \Big(\big| U_{n,F}(y_j) - V_{n,F}(y_j) \big| \geq \epsilon \Big)
\end{aligned}$$

for every $\epsilon > 0$, and every summand of these two sums converges to zero as $n \to \infty$ because of (5.17). Thus it is for every $\epsilon > 0$ and $k \in \mathbb{N}$

$$\begin{aligned}
\limsup_{n \to \infty} \sup_{F \in M} P_F \big(\sup_{x \in \mathbb{R}} A_{n,F}(x) \geq \epsilon \big) \;\leq\; & \limsup_{n \to \infty} \sup_{F \in M} P_F \Big(\sup_{\substack{x,y \in \mathbb{R} \\ \bar{d}_F(x,y) \leq 2/\sqrt{k}}} \big| U_{n,F}(x) - U_{n,F}(y) \big| \geq \frac{\epsilon}{3} \Big) \\
& + \limsup_{n \to \infty} \sup_{F \in M} P_F \Big(\sup_{\substack{x,y \in \mathbb{R} \\ \bar{d}_F(x,y) \leq 2/\sqrt{k}}} \big| V_{n,F}(x) - V_{n,F}(y) \big| \geq \frac{\epsilon}{3} \Big),
\end{aligned}$$

and the desired result follows because the right-hand side of this inequality converges to zero as $k \to \infty$ by (5.21) and (5.22).

Step 3: It remains to show (5.21), which is equivalent to

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \sup_{F \in M} P_F \big(\sup_{\substack{x,y \in \mathbb{R} \\ \bar{d}_F(x,y) \leq \delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 64\epsilon \big) = 0 \;\; \forall \epsilon > 0. \tag{5.25}$$

Proof of Step 3. For arbitrary $u \in (0,1)$ set

$$j(u) := \max\{ j \in \mathbb{N} \colon j < 1/u^2 \}.$$

Then $0 < u^2 \leq ju^2 < 1$ for $j = 1, \ldots, j(u)$.
Let $F \in M$ be arbitrary, but fixed for the moment. Because of the continuity and monotonicity of F there is a partition $-\infty < x_1(u) < \ldots < x_{j(u)}(u) < \infty$ such that $F(x_j(u)) = ju^2$ for $j = 1, \ldots, j(u)$.
Analogously, there is a partition $-\infty < y_1(u) < \ldots < y_{j(u)}(u) < \infty$ such that $G(y_j(u)) = ju^2$ for $j = 1, \ldots, j(u)$. Let $D(u) := \{z_1(u), \ldots, z_{l(u)}(u)\}$ be the common refinement of these two partitions, i.e.,

$$D(u) = \{x_1(u), \ldots, x_{j(u)}(u)\} \cup \{y_1(u), \ldots, y_{j(u)}(u)\} \tag{5.26}$$

and $z_1(u) < \ldots < z_{l(u)}(u)$. Note that $l(u) \leq 2j(u)$. By construction, $|F(x_{j+1}(u)) - F(x_j(u))| = u^2 = |G(y_{j+1}(u)) - G(y_j(u))|$ for all $j = 1, \ldots, j(u) - 1$. Since $z_1(u), \ldots, z_{l(u)}(u)$ is a refinement of the two partitions, we have

$$|F(z_{j+1}(u)) - F(z_j(u))| \leq u^2, \quad |G(z_{j+1}(u)) - G(z_j(u))| \leq u^2 \quad \forall j = 1, \ldots, l(u) - 1 \tag{5.27}$$

by the monotonicity of F and G. Now consider the mapping $J_u \colon \mathbb{R} \to D(u)$ with

$$J_u(x) = \begin{cases} z_1(u) & \text{if } x \in (-\infty, z_1(u)], \\ z_j(u) & \text{if } x \in (z_{j-1}(u), z_j(u)] \text{ for a } j \in \{2, \ldots, l(u)\}, \\ z_{l(u)}(u) & \text{if } x \in (z_{l(u)}(u), \infty). \end{cases}$$

Then we have

$$\bar{d}_F(x, J_u(x)) \leq 2u \quad \forall x \in \mathbb{R}. \tag{5.28}$$

The proof of (5.28) follows along the same lines as the proof of (5.24) and is therefore omitted here.
Now let $\epsilon > 0$ and $\delta \in (0, 1)$. Let $n \in \mathbb{N}$ with

$$n \geq \Big(\frac{2\epsilon^{1/2}}{\delta}\Big)^4. \tag{5.29}$$

Then $\sqrt{n} \geq 4\epsilon/\delta^2 \geq 4\epsilon$, and therefore $0 < \epsilon/\sqrt{n} \leq 1/4 < 1$. Thus, $D_n := D\big((\epsilon/\sqrt{n})^{1/2}\big)$ and $J_n := J_{(\epsilon/\sqrt{n})^{1/2}} \colon \mathbb{R} \to D_n$ are well-defined. We set

$$j(n) := j\big((\epsilon/\sqrt{n})^{1/2}\big), \quad x_j(n) := x_j\big((\epsilon/\sqrt{n})^{1/2}\big) \text{ and } y_j(n) := y_j\big((\epsilon/\sqrt{n})^{1/2}\big) \ \forall j = 1, \ldots, j(n),$$

and $l(n)$ and $z_j(n)$ are defined accordingly. Because of (5.28) it is

$$\bar{d}_F(x, J_n(x)) \ \leq \ 2\Big(\frac{\epsilon}{\sqrt{n}}\Big)^{1/2} = \frac{2\epsilon^{1/2}}{n^{1/4}} \quad \forall x \in \mathbb{R}. \tag{5.30}$$

Moreover, we have for all $x, y \in \mathbb{R}$

$$\bar{d}_F(x, y) \leq \delta \quad \Rightarrow \quad \bar{d}_F\big(J_n(x), J_n(y)\big) \leq 3\delta. \tag{5.31}$$

To see that this is true, recall that $\bar{d}_F$ is a pseudometric, so that

$$\bar{d}_F(J_n(x), J_n(y)) \ \leq \ \bar{d}_F(J_n(x), x) + \bar{d}_F(x, y) + \bar{d}_F(y, J_n(y)) \ \leq \ 2\frac{2\epsilon^{1/2}}{n^{1/4}} + \delta \ \leq \ 3\delta$$

by (5.30) and (5.29).

Now for every $x, y \in \mathbb{R}$ with $\bar{d}_F(x,y) \leq \delta$ it is

$$\begin{aligned}
|U_{n,F}(x) - U_{n,F}(y)| &\leq |U_{n,F}(x) - U_{n,F}(J_n(x))| + |U_{n,F}(J_n(x)) - U_{n,F}(J_n(y))| \\
&\quad + |U_{n,F}(J_n(y)) - U_{n,F}(y)| \\
&\leq 2 \sup_{x \in \mathbb{R}} |U_{n,F}(x) - U_{n,F}(J_n(x))| + \max_{\substack{x,y \in D_n \\ \bar{d}_F(x,y) \leq 3\delta}} |U_{n,F}(x) - U_{n,F}(y)|
\end{aligned}$$

because of (5.31). This implies that for every $\epsilon > 0$ and $\delta \in (0,1)$

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R} \\ \bar{d}_F(x,y)\leq\delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 64\epsilon\Big) \\
\leq\ &\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{x\in\mathbb{R}} |U_{n,F}(x) - U_{n,F}(J_n(x))| \geq 16\epsilon\Big) \\
&+ \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n \\ \bar{d}_F(x,y)\leq 3\delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\epsilon\Big). \qquad (5.32)
\end{aligned}$$

We will show next that
Step 3a:

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{x\in\mathbb{R}} |U_{n,F}(x) - U_{n,F}(J_n(x))| \geq 16\epsilon\Big) = 0 \quad \forall \epsilon > 0,\ \delta \in (0,1).$$

Proof of Step 3a. It is

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{x\in\mathbb{R}} |U_{n,F}(x) - U_{n,F}(J_n(x))| \geq 16\epsilon\Big) \\
\leq\ &\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{-\infty < x \leq z_1(n)} |U_{n,F}(x) - U_{n,F}(J_n(x))| \geq 16\epsilon\Big) \qquad (5.33) \\
&+ \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{z_1(n) < x \leq z_{l(n)}(n)} |U_{n,F}(x) - U_{n,F}(J_n(x))| \geq 16\epsilon\Big) \qquad (5.34) \\
&+ \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{z_{l(n)}(n) < x < \infty} |U_{n,F}(x) - U_{n,F}(J_n(x))| \geq 16\epsilon\Big). \qquad (5.35)
\end{aligned}$$

We will first investigate the term in (5.34).
As before, let $F \in M$, $\epsilon > 0$, $\delta \in (0,1)$ and $n \in \mathbb{N}$ with (5.29). For $x \in (z_1(n), z_{l(n)}(n)]$ there is a unique $l \in \{2, \ldots, l(n)\}$ such that $x \in (z_{l-1}(n), z_l(n)]$. Thus, $J_n(x) = z_l(n)$ by definition. Hence,

$$\begin{aligned}
&U_{n,F}(x) - U_{n,F}(J_n(x)) = U_{n,F}(x) - U_{n,F}(z_l(n)) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq x + \delta_{ni}\}} - F(x+\delta_{ni})\big] - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq z_l(n) + \delta_{ni}\}} - F(z_l(n)+\delta_{ni})\big] \\
&\geq \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq z_{l-1}(n) + \delta_{ni}\}} - F(z_{l-1}(n)+\delta_{ni})\big] - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq z_l(n) + \delta_{ni}\}} - F(z_l(n)+\delta_{ni})\big] \\
&\quad - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big] \\
&= U_{n,F}(z_{l-1}(n)) - U_{n,F}(z_l(n)) - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big] \\
&\geq -|U_{n,F}(z_{l-1}(n)) - U_{n,F}(z_l(n))| - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big].
\end{aligned}$$

Analogously we see that

$$\begin{aligned} &U_{n,F}(J_n(x)) - U_{n,F}(x) = U_{n,F}(z_l(n)) - U_{n,F}(x) \\ &\le |U_{n,F}(z_l(n)) - U_{n,F}(z_{l-1}(n))| + \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big]. \end{aligned}$$

This implies that

$$\begin{aligned} &|U_{n,F}(J_n(x)) - U_{n,F}(x)| \\ &\le |U_{n,F}(z_l(n)) - U_{n,F}(z_{l-1}(n))| + \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big] \\ &\le \max_{2\le l\le l(n)} |U_{n,F}(z_l(n)) - U_{n,F}(z_{l-1}(n))| + \max_{2\le l\le l(n)} \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big], \end{aligned}$$

whence it follows that

$$\begin{aligned} &\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{z_1(n)<x\le z_{l(n)}(n)} |U_{n,F}(x) - U_{n,F}(J_n(x))| \ge 16\epsilon\Big) \\ \le\ &\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{2\le l\le l(n)} |U_{n,F}(z_l(n)) - U_{n,F}(z_{l-1}(n))| \ge 8\epsilon\Big) \\ &+ \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{2\le l\le l(n)} \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big] \ge 8\epsilon\Big). \end{aligned} \tag{5.36}$$

We will first investigate the second term on the right-hand side of the above inequality. It is

$$\Delta_{n,F} := \sup\{|f(x) - f(y)|\colon x,y\in\mathbb{R}, |x-y| \le \max_{1\le i\le n} |\delta_{ni}|\} = o_P^u(1) \text{ in } M \text{ as } n\to\infty. \tag{5.37}$$

To prove (5.37), note that because of (5.12), i.e., because of the uniform equicontinuity of the family $\{f\colon F\in M\}$, for every $\eta > 0$ there is a $\tilde{\delta} > 0$ such that for all $x, y \in \mathbb{R}$

$$|x-y| \le \tilde{\delta} \quad \Rightarrow \quad |f(x) - f(y)| \le \eta \ \ \forall F \in M,$$

so that $\max_{1\le i\le n} |\delta_{ni}| \le \tilde{\delta}$ implies that $\Delta_{n,F} \le \eta$ for all $F \in M$. Hence

$$\sup_{F\in M} P_F(\Delta_{n,F} > \eta) \ \le\ \sup_{F\in M} P_F(\max_{1\le i\le n} |\delta_{ni}| > \tilde{\delta}) \ \underset{n\to\infty}{\longrightarrow}\ 0$$

by assumption (iv), and this shows (5.37).

For any $l \in \{2,\ldots,l(n)\}$ we now have

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F(z_l(n)+\delta_{ni}) - F(z_{l-1}(n)+\delta_{ni})\big] \\ &\le \frac{1}{\sqrt{n}} \sum_{i=1}^n \big|F(z_l(n)+\delta_{ni}) - F(z_l(n)) - \big[F(z_{l-1}(n)+\delta_{ni}) - F(z_{l-1}(n))\big]\big| \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \big|F(z_l(n)) - F(z_{l-1}(n))\big| \\ &\le \frac{1}{\sqrt{n}} \sum_{i=1}^n \big|f(\xi_{1i})\delta_{ni} - f(\xi_{2i})\delta_{ni}\big| + \epsilon, \end{aligned} \tag{5.38}$$

where ξ_{1i} lies between $z_l(n)+\delta_{ni}$ and $z_l(n)$, and ξ_{2i} lies between $z_{l-1}(n)+\delta_{ni}$ and $z_{l-1}(n)$. Moreover, in the last inequality we used (5.27) with $u=(\epsilon/\sqrt{n})^{1/2}$. But

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big|f(\xi_{1i})\delta_{ni}-f(\xi_{2i})\delta_{ni}\big| \\
&\le \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\big|f(\xi_{1i})-f(z_l(n))\big|+\big|f(z_l(n))-f(z_{l-1}(n))\big|+\big|f(z_{l-1}(n))-f(\xi_{2i})\big|\Big)|\delta_{ni}| \\
&\le \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\Delta_{n,F}+\omega_F\big(2(\epsilon/\sqrt{n})^{1/2}\big)+\Delta_{n,F}\Big)|\delta_{ni}|, \qquad (5.39)
\end{aligned}$$

where $\Delta_{n,F}$ is as in (5.37) and ω_F was defined in (5.15). To see that $|f(z_l(n))-f(z_{l-1}(n))|\le\omega_F\big(2(\epsilon/\sqrt{n})^{1/2}\big)$ is indeed true, note that $\bar{d}_F(z_l(n),z_{l-1}(n))\le 2(\epsilon/\sqrt{n})^{1/2}$ by (5.27) with $u=(\epsilon/\sqrt{n})^{1/2}$.

Combining (5.38) and (5.39), it follows that

$$\begin{aligned}
&\max_{2\le l\le l(n)}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[F(z_l(n)+\delta_{ni})-F(z_{l-1}(n)+\delta_{ni})\big] \\
&\le\Big(2\Delta_{n,F}+\omega_F\big(2(\epsilon/\sqrt{n})^{1/2}\big)\Big)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|\delta_{ni}|+\epsilon=o_P^u(1)+\epsilon \qquad (5.40)
\end{aligned}$$

by (5.16), (5.37) and assumption (iii). Hence

$$\begin{aligned}
&\limsup_{n\to\infty}\sup_{F\in M}P_F\Big(\max_{2\le l\le l(n)}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[F(z_l(n)+\delta_{ni})-F(z_{l-1}(n)+\delta_{ni})\big]\ge 8\epsilon\Big) \\
&\le\limsup_{n\to\infty}\sup_{F\in M}P_F\Big(o_P^u(1)\ge 7\epsilon\Big)=0,
\end{aligned}$$

so that the second term on the right-hand side of inequality (5.36) equals zero. We will next investigate the first term on the right-hand side of this inequality.

For every $l\in\{2,\dots,l(n)\}$ it is

$$\begin{aligned}
&U_{n,F}(z_l(n))-U_{n,F}(z_{l-1}(n)) \\
&=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(1_{\{e_i\le z_l(n)+\delta_{ni}\}}-F(z_l(n)+\delta_{ni})-\big[1_{\{e_i\le z_{l-1}(n)+\delta_{ni}\}}-F(z_{l-1}(n)+\delta_{ni})\big]\big) \\
&=:\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\zeta_i(l).
\end{aligned}$$

Using (5.18) we see that $\zeta_1(l),\dots,\zeta_n(l)$ is a MDS with respect to $\mathcal{F}_0\subset\mathcal{F}_1\subset\ldots\subset\mathcal{F}_n$ under P_F, and $|\zeta_i(l)|\le 1$ for all $i=1,\dots,n$. For every $l=2,\dots,l(n)$ it follows from (5.20) with $x=z_l(n)$, $y=z_{l-1}(n)$ and $\tilde{\delta}_{ni}=\delta_{ni}$ that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathrm{E}_F(\zeta_i(l)^2|\mathcal{F}_{i-1})\le\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|F(z_l(n)+\delta_{ni})-F(z_{l-1}(n)+\delta_{ni})|,$$

so that by (5.40)

$$\max_{2\le l\le l(n)}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathrm{E}_F(\zeta_i(l)^2|\mathcal{F}_{i-1})\le o_P^u(1)+\epsilon.$$

Then for every $n \in \mathbb{N}$ with (5.29) it is

$$\begin{aligned}
&P_F\Big(\max_{2\le l\le l(n)} |U_{n,F}(z_l(n)) - U_{n,F}(z_{l-1}(n))| \ge 8\epsilon\Big)\\
&\le P_F\Big(\Big\{\max_{2\le l\le l(n)} \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n \zeta_i(l)\Big| \ge 8\epsilon\Big\} \cap \Big\{\max_{2\le l\le l(n)} \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathrm{E}_F(\zeta_i(l)^2|\mathcal{F}_{i-1}) \le 2\epsilon\Big\}\Big)\\
&\quad + P_F\Big(\max_{2\le l\le l(n)} \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathrm{E}_F(\zeta_i(l)^2|\mathcal{F}_{i-1}) > 2\epsilon\Big)\\
&\le \sum_{l=2}^{l(n)} P_F\Big(\Big\{\Big|\sum_{i=1}^n \zeta_i(l)\Big| \ge 8\sqrt{n}\epsilon\Big\} \cap \Big\{\sum_{i=1}^n \mathrm{E}_F(\zeta_i(l)^2|\mathcal{F}_{i-1}) \le 2\sqrt{n}\epsilon\Big\}\Big) + P_F\big(o_p^u(1) + \epsilon > 2\epsilon\big)\\
&\le \sum_{l=2}^{l(n)} 2\exp\Big(8\sqrt{n}\epsilon - 8\sqrt{n}\epsilon\log\Big(1 + \frac{8\sqrt{n}\epsilon}{2\sqrt{n}\epsilon}\Big)\Big) + P_F\big(o_p^u(1) > \epsilon\big)\\
&\le 2\cdot l(n)\cdot \exp\Big(8\sqrt{n}\epsilon[1-\log(5)]\Big) + P_F\big(o_p^u(1) > \epsilon\big)\\
&\le 4\cdot \frac{\sqrt{n}}{\epsilon}\cdot \exp\Big(8\sqrt{n}\epsilon[1-\log(5)]\Big) + \sup_{F\in M} P_F\big(o_p^u(1) > \epsilon\big),
\end{aligned}$$

where the third-to-last inequality follows by inequality (A.3) of Lemma A.4, and in the last inequality we used that $l(n) \le 2j(n) \le 2\sqrt{n}/\epsilon$ by construction. Now obviously both summands on the right-hand side of the last inequality do not depend on F anymore and converge to zero as $n \to \infty$. Thus, we have shown that

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{2\le l\le l(n)} |U_{n,F}(z_l(n)) - U_{n,F}(z_{l-1}(n))| \ge 8\epsilon\Big) = 0,$$

and this concludes the proof of

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{z_1(n) < x \le z_{l(n)}(n)} |U_{n,F}(x) - U_{n,F}(J_n(x))| \ge 16\epsilon\Big) = 0$$

from (5.34).

Next, we will investigate the term in (5.33).
Again, let $F \in M$, $\epsilon > 0$, $\delta \in (0,1)$ and $n \in \mathbb{N}$ with (5.29). Let $x \in (-\infty, z_1(n)]$. Then $J_n(x) = z_1(n)$ by definition of J_n. Thus

$$\begin{aligned}
&U_{n,F}(x) - U_{n,F}(J_n(x)) = U_{n,F}(x) - U_{n,F}(z_1(n))\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}\}} - F(x+\delta_{ni})\big] - \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le z_1(n)+\delta_{ni}\}} - F(z_1(n)+\delta_{ni})\big]\\
&\le \frac{1}{\sqrt{n}}\sum_{i=1}^n 1_{\{e_i \le z_1(n)+\delta_{ni}\}} - \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le z_1(n)+\delta_{ni}\}} - F(z_1(n)+\delta_{ni})\big]\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n F(z_1(n)+\delta_{ni}),
\end{aligned}$$

and similarly

$$\begin{aligned}
&U_{n,F}(J_n(x)) - U_{n,F}(x) = U_{n,F}(z_1(n)) - U_{n,F}(x)\\
&\le \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le z_1(n)+\delta_{ni}\}} - F(z_1(n)+\delta_{ni})\big] + \frac{1}{\sqrt{n}}\sum_{i=1}^n F(z_1(n)+\delta_{ni}).
\end{aligned}$$

This yields

$$\sup_{-\infty<x\leq z_1(n)} |U_{n,F}(x) - U_{n,F}(J_n(x))|$$
$$\leq \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big[1_{\{e_i\leq z_1(n)+\delta_{ni}\}} - F(z_1(n)+\delta_{ni})\big]\Big| + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} F(z_1(n)+\delta_{ni}). \tag{5.41}$$

Now for all $n \in \mathbb{N}$ with (5.29) we have $F(z_1(n)) \leq F(x_1(n)) = \epsilon/\sqrt{n}$ by definition, and

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} F(z_1(n)+\delta_{ni}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} f(\xi_i)\delta_{ni} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} F(z_1(n))\\ &\leq \frac{1}{\sqrt{n}}\sum_{i=1}^{n} |f(\xi_i) - f(z_1(n))||\delta_{ni}| + f(z_1(n))\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|\delta_{ni}| + \epsilon,\end{aligned} \tag{5.42}$$

where ξ_i lies between $z_1(n)+\delta_{ni}$ and $z_1(n)$. It is $|f(\xi_i) - f(z_1(n))| \leq \Delta_{n,F}$ because $|\xi_i - z_1(n)| \leq |\delta_{ni}|$. Also,

$$\sup_{F\in M} f(z_1(n)) \underset{n\to\infty}{\longrightarrow} 0. \tag{5.43}$$

To see that (5.43) is true, note that

$$0 \leq f(z_1(n)) \leq |f(z_1(n)) - f(z_1(n)\wedge(-n))| + f(z_1(n)\wedge(-n)).$$

Because of $z_1(n)\wedge(-n) \leq -n \underset{n\to\infty}{\longrightarrow} -\infty$ it follows from assumption (5.13) that

$$\sup_{F\in M} f(z_1(n)\wedge(-n)) \underset{n\to\infty}{\longrightarrow} 0.$$

Moreover, since $-\infty < z_1(n)\wedge(-n) \leq z_1(n) = x_1(n)\wedge y_1(n)$ we have

$$F(z_1(n)\wedge(-n)) \leq F(z_1(n)) \leq F(x_1(n)) = \frac{\epsilon}{\sqrt{n}}$$

and

$$G(z_1(n)\wedge(-n)) \leq G(z_1(n)) \leq G(y_1(n)) = \frac{\epsilon}{\sqrt{n}},$$

so that

$$\begin{aligned}\bar{d}_F(z_1(n)\wedge(-n), z_1(n)) &= |F(z_1(n)\wedge(-n)) - F(z_1(n))|^{1/2} + |G(z_1(n)\wedge(-n)) - G(z_1(n))|^{1/2}\\ &\leq F(z_1(n))^{1/2} + G(z_1(n))^{1/2} \leq 2\left(\frac{\epsilon}{\sqrt{n}}\right)^{1/2}.\end{aligned}$$

Therefore,

$$\sup_{F\in M} |f(z_1(n)) - f(z_1(n)\wedge(-n))| \leq \sup_{F\in M} \omega_F\Big(2\Big(\frac{\epsilon}{\sqrt{n}}\Big)^{1/2}\Big) \underset{n\to\infty}{\longrightarrow} 0$$

because of (5.16). This completes the proof of (5.43).

Combining all this, it follows from inequality (5.42) that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} F(z_1(n)+\delta_{ni}) \leq \big(\Delta_{n,F} + f(z_1(n))\big)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|\delta_{ni}| + \epsilon = o_P^u(1) + \epsilon \tag{5.44}$$

by (5.37), (5.43) and assumption (iii).

Now using (5.41) it follows that

$$\begin{aligned}
&\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sup_{-\infty<x\le z_1(n)} |U_{n,F}(x)-U_{n,F}(J_n(x))|\ge 16\epsilon\Big)\\
&\le \limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n\big[1_{\{e_i\le z_1(n)+\delta_{ni}\}}-F(z_1(n)+\delta_{ni})\big]\Big|\ge 8\epsilon\Big)\\
&\quad+\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n F(z_1(n)+\delta_{ni})\ge 8\epsilon\Big)\\
&= \limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n\big[1_{\{e_i\le z_1(n)+\delta_{ni}\}}-F(z_1(n)+\delta_{ni})\big]\Big|\ge 8\epsilon\Big).
\end{aligned}$$

Set

$$\zeta_i := 1_{\{e_i\le z_1(n)+\delta_{ni}\}}-F(z_1(n)+\delta_{ni}),\quad i=1,\dots,n.$$

Because of (5.18) it follows that $\zeta_1,\dots,\zeta_n$ is a MDS with respect to $\mathcal{F}_0\subset\mathcal{F}_1\subset\dots\subset\mathcal{F}_n$ under P_F, and obviously $|\zeta_i|\le 1$ for all $i=1,\dots,n$. Also by (5.18) it is

$$\begin{aligned}
\sum_{i=1}^n \mathrm{E}_F(\zeta_i^2|\mathcal{F}_{i-1}) &= \sum_{i=1}^n \mathrm{E}_F\Big(\big[1_{\{e_i\le z_1(n)+\delta_{ni}\}}-\mathrm{E}_F(1_{\{e_i\le z_1(n)+\delta_{ni}\}}|\mathcal{F}_{i-1})\big]^2\Big|\mathcal{F}_{i-1}\Big)\\
&\le \sum_{i=1}^n \mathrm{E}_F\big(1_{\{e_i\le z_1(n)+\delta_{ni}\}}\big|\mathcal{F}_{i-1}\big)=\sum_{i=1}^n F(z_1(n)+\delta_{ni})\\
&\le o_P^u(\sqrt{n})+\sqrt{n}\epsilon,
\end{aligned}$$

the last inequality following from (5.44). Thus, for all $n\in\mathbb{N}$ with (5.29) this yields

$$\begin{aligned}
&P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n\big[1_{\{e_i\le z_1(n)+\delta_{ni}\}}-F(z_1(n)+\delta_{ni})\big]\Big|\ge 8\epsilon\Big)=P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n\zeta_i\Big|\ge 8\epsilon\Big)\\
&\le P_F\Big(\Big\{\Big|\sum_{i=1}^n\zeta_i\Big|\ge 8\sqrt{n}\epsilon\Big\}\cap\Big\{\sum_{i=1}^n \mathrm{E}_F(\zeta_i^2|\mathcal{F}_{i-1})\le 2\sqrt{n}\epsilon\Big\}\Big)\\
&\quad+P_F\Big(\sum_{i=1}^n \mathrm{E}_F(\zeta_i^2|\mathcal{F}_{i-1})>2\sqrt{n}\epsilon\Big)\\
&\le 2\cdot\exp\Big(8\sqrt{n}\epsilon-8\sqrt{n}\epsilon\log\Big(1+\frac{8\sqrt{n}\epsilon}{2\sqrt{n}\epsilon}\Big)\Big)+P_F\big(o_P^u(\sqrt{n})+\sqrt{n}\epsilon>2\sqrt{n}\epsilon\big)\\
&\le 2\cdot\exp\Big(8\sqrt{n}\epsilon\big[1-\log(5)\big]\Big)+\sup_{F\in M}P_F\big(o_P^u(1)>\epsilon\big),
\end{aligned}$$

where the second-to-last inequality follows again by (A.3) of Lemma A.4. Now the right-hand side of the last inequality does not depend on F anymore and converges to zero as $n\to\infty$. This yields

$$\begin{aligned}
&\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sup_{-\infty<x\le z_1(n)} |U_{n,F}(x)-U_{n,F}(J_n(x))|\ge 16\epsilon\Big)\\
&\le \limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n\big[1_{\{e_i\le z_1(n)+\delta_{ni}\}}-F(z_1(n)+\delta_{ni})\big]\Big|\ge 8\epsilon\Big)=0.
\end{aligned}$$

Thus, the term in (5.33) is zero.

To conclude the proof of Step 3a, it remains to show that the term in (5.35) equals zero. As before, let $F \in M$, $\epsilon > 0$, $\delta \in (0,1)$ and $n \in \mathbb{N}$ with (5.29). For $x \in (z_{l(n)}(n), \infty)$ we have $J_n(x) = z_{l(n)}(n)$ by definition. It is

$$\begin{aligned}
&U_{n,F}(x) - U_{n,F}(J_n(x)) = U_{n,F}(x) - U_{n,F}(z_{l(n)}(n)) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(1_{\{z_{l(n)}(n)+\delta_{ni} < e_i \le x+\delta_{ni}\}} - \Big[F(x+\delta_{ni}) - F(z_{l(n)}(n)+\delta_{ni}) \Big] \Big) \\
&\le \frac{1}{\sqrt{n}} \sum_{i=1}^{n} 1_{\{z_{l(n)}(n)+\delta_{ni} < e_i\}} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(1_{\{z_{l(n)}(n)+\delta_{ni} < e_i\}} - \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big] \Big) + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big]
\end{aligned}$$

and

$$\begin{aligned}
&U_{n,F}(J_n(x)) - U_{n,F}(x) = U_{n,F}(z_{l(n)}(n)) - U_{n,F}(x) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(1_{\{e_i \le z_{l(n)}(n)+\delta_{ni}\}} - 1_{\{e_i \le x+\delta_{ni}\}} + F(x+\delta_{ni}) - F(z_{l(n)}(n)+\delta_{ni}) \Big) \\
&\le \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big].
\end{aligned}$$

This yields

$$\begin{aligned}
&|U_{n,F}(J_n(x)) - U_{n,F}(x)| \\
&\le \frac{1}{\sqrt{n}} \Big| \sum_{i=1}^{n} \Big(1_{\{z_{l(n)}(n)+\delta_{ni} < e_i\}} - \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big] \Big) \Big| + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big],
\end{aligned}$$

so that

$$\begin{aligned}
&\sup_{z_{l(n)}(n) < x < \infty} |U_{n,F}(J_n(x)) - U_{n,F}(x)| \\
&\le \frac{1}{\sqrt{n}} \Big| \sum_{i=1}^{n} \Big(1_{\{e_i \le z_{l(n)}(n)+\delta_{ni}\}} - F(z_{l(n)}(n)+\delta_{ni}) \Big) \Big| + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big].
\end{aligned}$$

For all $n \in \mathbb{N}$ with (5.29) we have $F(z_{l(n)}(n)) \ge F(x_{j(n)}(n)) = j(n) \cdot \epsilon/\sqrt{n}$ and $1 - j(n) \cdot \epsilon/\sqrt{n} \le \epsilon/\sqrt{n}$ by definition of $j(n)$, so that

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[1 - F(z_{l(n)}(n)+\delta_{ni}) \big] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[1 - F(z_{l(n)}(n)) \big] + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[F(z_{l(n)}(n)) - F(z_{l(n)}(n)+\delta_{ni}) \big] \\
&\le \epsilon + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(\xi_i)|\delta_{ni}| \le \epsilon + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |f(\xi_i) - f(z_{l(n)}(n))||\delta_{ni}| + f(z_{l(n)}(n)) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |\delta_{ni}| \quad (5.45)
\end{aligned}$$

with ξ_i lying between $z_{l(n)}(n) + \delta_{ni}$ and $z_{l(n)}(n)$. Now $|\xi_i - z_{l(n)}(n)| \le |\delta_{ni}|$ implies that $|f(\xi_i) - f(z_{l(n)}(n))| \le \Delta_{n,F}$. Moreover,

$$\sup_{F \in M} f(z_{l(n)}(n)) \underset{n\to\infty}{\longrightarrow} 0, \qquad (5.46)$$

which can be shown similarly to (5.43).

Hence, by (5.45) we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[1-F(z_{l(n)}(n)+\delta_{ni})\big] \leq \big(\Delta_{n,F}+f(z_{l(n)}(n))\big)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|\delta_{ni}|+\epsilon = o_P^u(1)+\epsilon \tag{5.47}$$

using (5.37), (5.46) and assumption (iii). Combining all this yields

$$\begin{aligned}
&\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sup_{z_{l(n)}(n)<x<\infty}|U_{n,F}(x)-U_{n,F}(J_n(x))|\geq 16\epsilon\Big)\\
&\leq \limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\Big(1_{\{e_i\leq z_{l(n)}(n)+\delta_{ni}\}}-F(z_{l(n)}(n)+\delta_{ni})\Big)\Big|\geq 8\epsilon\Big)\\
&\quad+\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[1-F(z_{l(n)}(n)+\delta_{ni})\big]\geq 8\epsilon\Big)\\
&=\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\Big(1_{\{e_i\leq z_{l(n)}(n)+\delta_{ni}\}}-F(z_{l(n)}(n)+\delta_{ni})\Big)\Big|\geq 8\epsilon\Big)
\end{aligned}$$

because of (5.47). Define

$$\bar{\zeta}_i := 1_{\{e_i\leq z_{l(n)}(n)+\delta_{ni}\}}-F(z_{l(n)}(n)+\delta_{ni}),\quad i=1,\ldots,n.$$

Using (5.18) we see that $\bar{\zeta}_1,\ldots,\bar{\zeta}_n$ is a MDS with respect to $\mathcal{F}_0\subset\mathcal{F}_1\subset\ldots\subset\mathcal{F}_n$ under P_F. Moreover, all of the $|\bar{\zeta}_i|$ are less than or equal to one. Again by (5.18) it is

$$\begin{aligned}
\sum_{i=1}^{n}\mathrm{E}_F(\bar{\zeta}_i^2|\mathcal{F}_{i-1}) &= \sum_{i=1}^{n}\mathrm{E}_F\Big(\big[1_{\{e_i\leq z_{l(n)}(n)+\delta_{ni}\}}-F(z_{l(n)}(n)+\delta_{ni})\big]^2\Big|\mathcal{F}_{i-1}\Big)\\
&=\sum_{i=1}^{n}\mathrm{E}_F\Big(\big[1_{\{z_{l(n)}(n)+\delta_{ni}<e_i\}}-\big(1-F(z_{l(n)}(n)+\delta_{ni})\big)\big]^2\Big|\mathcal{F}_{i-1}\Big)\\
&=\sum_{i=1}^{n}\mathrm{E}_F\Big(\big[1_{\{z_{l(n)}(n)+\delta_{ni}<e_i\}}-\mathrm{E}_F\big(1_{\{z_{l(n)}(n)+\delta_{ni}<e_i\}}|\mathcal{F}_{i-1}\big)\big]^2\Big|\mathcal{F}_{i-1}\Big)\\
&\leq\sum_{i=1}^{n}\mathrm{E}_F\Big(1_{\{z_{l(n)}(n)+\delta_{ni}<e_i\}}\Big|\mathcal{F}_{i-1}\Big)=\sum_{i=1}^{n}\Big[1-F(z_{l(n)}(n)+\delta_{ni})\Big]\\
&\leq o_P^u(\sqrt{n})+\sqrt{n}\epsilon,
\end{aligned}$$

and the last inequality follows from (5.47). Hence,

$$\begin{aligned}
&P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\Big(1_{\{e_i\leq z_{l(n)}(n)+\delta_{ni}\}}-F(z_{l(n)}(n)+\delta_{ni})\Big)\Big|\geq 8\epsilon\Big)\\
&\leq P_F\Big(\Big\{\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\bar{\zeta}_i\Big|\geq 8\epsilon\Big\}\cap\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathrm{E}_F(\bar{\zeta}_i^2|\mathcal{F}_{i-1})\leq 2\epsilon\Big\}\Big)+P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathrm{E}_F(\bar{\zeta}_i^2|\mathcal{F}_{i-1})>2\epsilon\Big)\\
&\leq P_F\Big(\Big\{\Big|\sum_{i=1}^{n}\bar{\zeta}_i\Big|\geq 8\sqrt{n}\epsilon\Big\}\cap\Big\{\sum_{i=1}^{n}\mathrm{E}_F(\bar{\zeta}_i^2|\mathcal{F}_{i-1})\leq 2\sqrt{n}\epsilon\Big\}\Big)+P_F\big(o_P^u(1)+\epsilon>2\epsilon\big)\\
&\leq 2\exp\big(8\sqrt{n}\epsilon\big[1-\log(5)\big]\big)+\sup_{F\in M}P_F\big(o_P^u(1)>\epsilon\big),
\end{aligned}$$

where the last inequality follows again by (A.3) of Lemma A.4. Since both summands on the right-hand side of the last inequality are independent of F and converge to zero with $n\to\infty$, it follows that

$$\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\sup_{z_{l(n)}(n)<x<\infty}|U_{n,F}(x)-U_{n,F}(J_n(x))|\geq 16\epsilon\Big)$$

$$\leq \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\Big(1_{\{e_i\leq z_{l(n)}(n)+\delta_{ni}\}} - F(z_{l(n)}(n)+\delta_{ni})\Big)\Big| \geq 8\epsilon\Big) = 0,$$

and this completes the proof of Step 3a.

It follows now from (5.32) that for every $\epsilon > 0$ and $\delta \in (0,1)$

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\leq\delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 64\epsilon\Big)\\
&\leq \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n\\ \bar{d}_F(x,y)\leq 3\delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\epsilon\Big).
\end{aligned} \tag{5.48}$$

We will next prove the following statement:

If for every $\epsilon > 0$ there is a $\delta_\epsilon \in (0,1)$ such that

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n\\ \bar{d}_F(x,y)\leq 3\delta_\epsilon}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\epsilon\Big) \leq \epsilon,$$

then (5.25) holds.

For the proof of this, let $\epsilon > 0$. Note that the function $h\colon (0,\infty) \to [0,1]$,

$$h(\delta) := \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\leq\delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 64\epsilon\Big),$$

is monotonically increasing. This ensures the existence of $\lim_{\delta\downarrow 0} h(\delta) =: h(0)$ in $[0,1]$. Because of the monotonicity it is $h(\delta) \geq h(0)$ for all $\delta \in (0,\infty)$. Now let $\eta \in (0,\epsilon)$ be arbitrary. By the assumptions there is a $\delta_\eta \in (0,1)$ with

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n\\ \bar{d}_F(x,y)\leq 3\delta_\eta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\eta\Big) \leq \eta.$$

Using (5.48), this yields

$$\begin{aligned}
\eta &\geq \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\leq\delta_\eta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 64\eta\Big)\\
&\geq \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ \bar{d}_F(x,y)\leq\delta_\eta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 64\epsilon\Big) = h(\delta_\eta) \geq h(0).
\end{aligned}$$

Since η was chosen arbitrarily, this shows that $h(0) = 0$, and this is just (5.25).

To conclude the proof of Step 3, it therefore remains to show:

Step 3b: For every $\epsilon > 0$ there is a $\delta_\epsilon \in (0,1)$ such that

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n\\ \bar{d}_F(x,y)\leq 3\delta_\epsilon}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\epsilon\Big) \leq \epsilon. \tag{5.49}$$

Proof of Step 3b: Let $\epsilon > 0$, $\delta \in (0,1/3)$, $F \in M$ and $n \in \mathbb{N}$ with

$$n \geq \Big(\frac{4\epsilon^{1/2}}{\delta}\Big)^4. \tag{5.50}$$

Set $\alpha := (\epsilon/\sqrt{n})^{1/2}$. Then $0 < \alpha < 4(\epsilon/\sqrt{n})^{1/2} \leq \delta$ because of (5.50). Let

$$k = k(n) := \min\{j \in \mathbb{N} \colon \delta < 3^j \alpha\}.$$

Note that k is well-defined and greater than or equal to two. Now set

$$\delta_l := 3^{k-l} \cdot \alpha, \quad l = 0, \ldots, k.$$

Then

$$0 < \alpha = \delta_k < \delta_{k-1} < \ldots < \delta_1 = 3^{k-1}\alpha \leq \delta < \delta_0 = 3^k \alpha \leq 3\delta < 1, \tag{5.51}$$

where we used that $3^k\alpha \leq 3\delta$ because $3^{k-1}\alpha \leq \delta$ by definition. Then for every $l = 0, \ldots, k$ the set $D(\delta_l)$ is well-defined, where $D(\delta_l)$ is as in (5.26) the common refinement of the partitions $x_1(\delta_l), \ldots, x_{j(\delta_l)}(\delta_l)$ and $y_1(\delta_l), \ldots, y_{j(\delta_l)}(\delta_l)$. Since $\delta_k = \alpha = (\epsilon/\sqrt{n})^{1/2}$, we have

$$D(\delta_k) = D(\alpha) = D\big((\epsilon/\sqrt{n})^{1/2}\big) = D_n.$$

For every $l = 1, 2, \ldots, k$ let the mapping

$$N_l \colon D(\delta_l) \longrightarrow D(\delta_{l-1})$$

be such that every $z \in D(\delta_l)$ is mapped onto an $N_l(z) \in D(\delta_{l-1})$ so, that $\bar{d}_F(z, N_l(z)) \leq \bar{d}_F(z, x)$ for all $x \in D(\delta_{l-1})$. Note that such an element always exists because of the finiteness of $D(\delta_{l-1})$, but it need not be unique. For our purposes, however, it is irrelevant onto which of these elements z is mapped. By construction,

$$\bar{d}_F(z, N_l(z)) \leq \bar{d}_F(z, x) \;\; \forall\, z \in D(\delta_l),\, x \in D(\delta_{l-1}),$$

and especially, since $J_{\delta_{l-1}}(z) \in D(\delta_{l-1})$,

$$\bar{d}_F(z, N_l(z)) \leq \bar{d}_F(z, J_{\delta_{l-1}}(z)) \;\; \forall\, z \in D(\delta_l) \tag{5.52}$$

for all $l = 1, \ldots, k$. Now for every $z \in D_n = D(\delta_k)$ set

$$s_k(z) := z, \quad s_{l-1}(z) := N_l(s_l(z)) \;\text{ for } l = k, \ldots, 1. \tag{5.53}$$

Then by construction $s_l(z) \in D(\delta_l)$ for all $z \in D_n$ and all $l = 0, \ldots, k$, and

$$\bar{d}_F(s_l(z), s_{l-1}(z)) = \bar{d}_F(s_l(z), N_l(s_l(z))) \leq \bar{d}_F(s_l(z), J_{\delta_{l-1}}(s_l(z))) \leq 2 \cdot \delta_{l-1} \tag{5.54}$$

for all $l = k, \ldots, 1$ using (5.28). Moreover, for every $x, y \in D_n$ with $\bar{d}_F(x, y) \leq 3\delta$ it is

$$\bar{d}_F(s_0(x), s_0(y)) \leq 21\delta. \tag{5.55}$$

To see that (5.55) is true, note that

$$\begin{aligned}
\bar{d}_F(s_0(x), s_0(y)) &\leq \sum_{l=0}^{k-1} \bar{d}_F(s_l(x), s_{l+1}(x)) + \bar{d}_F(s_k(x), s_k(y)) + \sum_{l=0}^{k-1} \bar{d}_F(s_l(y), s_{l+1}(y)) \\
&\leq \sum_{l=0}^{k-1} 2\delta_l + \bar{d}_F(x, y) + \sum_{l=0}^{k-1} 2\delta_l \leq 4 \sum_{l=0}^{k-1} 3^{k-l}\alpha + 3\delta \\
&\leq 4 \cdot 3^k \alpha \sum_{l=0}^{\infty} 3^{-l} + 3\delta \leq 4 \cdot 3\delta \sum_{l=0}^{\infty} 3^{-l} + 3\delta = 21\delta,
\end{aligned}$$

where we used (5.54) and again the fact that $3^k\alpha \leq 3\delta$.

Consequently, for every $\epsilon > 0$, $\delta \in (0, 1/3)$, $F \in M$ and $n \in \mathbb{N}$ with (5.50) we have

$$\begin{aligned}
&\max_{\substack{x,y\in D_n \\ \bar{d}_F(x,y)\le 3\delta}} |U_{n,F}(x) - U_{n,F}(y)| \\
&\le \max_{\substack{x,y\in D_n \\ \bar{d}_F(x,y)\le 3\delta}} \Big(|U_{n,F}(x) - U_{n,F}(s_0(x))| + |U_{n,F}(s_0(x)) - U_{n,F}(s_0(y))| + |U_{n,F}(s_0(y)) - U_{n,F}(y)|\Big) \\
&\le 2 \max_{x\in D_n} |U_{n,F}(x) - U_{n,F}(s_0(x))| + \max_{\substack{x,y\in D(\delta_0) \\ \bar{d}_F(x,y)\le 21\delta}} |U_{n,F}(x) - U_{n,F}(y)|,
\end{aligned}$$

and this yields

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n \\ \bar{d}_F(x,y)\le 3\delta}} |U_{n,F}(x) - U_{n,F}(y)| \ge 32\epsilon\Big) \\
&\le \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{x\in D_n} |U_{n,F}(x) - U_{n,F}(s_0(x))| \ge 8\epsilon\Big) \qquad (5.56) \\
&\quad + \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D(\delta_0) \\ \bar{d}_F(x,y)\le 21\delta}} |U_{n,F}(x) - U_{n,F}(y)| \ge 16\epsilon\Big) \qquad (5.57)
\end{aligned}$$

for $\epsilon > 0$, $\delta \in (0, 1/3)$.

We will first investigate the term in (5.57). Let $F \in M$. For every $x, y \in D(\delta_0)$ with $\bar{d}_F(x, y) \le 21\delta$ set

$$\zeta_i(x, y) := 1_{\{e_i\le x+\delta_{ni}\}} - F(x + \delta_{ni}) - \big(1_{\{e_i\le y+\delta_{ni}\}} - F(y + \delta_{ni})\big).$$

Then $|U_{n,F}(x) - U_{n,F}(y)| = \frac{1}{\sqrt{n}}\big|\sum_{i=1}^n \zeta_i(x, y)\big|$, and because of (5.18) the random variables $\zeta_1(x, y), \ldots, \zeta_n(x, y)$ form a MDS with respect to $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n$ under P_F. Also,

$$|\zeta_i(x, y)| = \big|1_{\{(x\wedge y)+\delta_{ni}<e_i\le(x\vee y)+\delta_{ni}\}} - \big(F((x \vee y) + \delta_{ni}) - F((x \wedge y) + \delta_{ni})\big)\big| \le 1 \qquad (5.58)$$

for all $i = 1, \ldots, n$. By (5.20) it is

$$\begin{aligned}
\sum_{i=1}^n \mathrm{E}_F\big(\zeta_i(x, y)^2|\mathcal{F}_{i-1}\big) &\le \sum_{i=1}^n |F(x + \delta_{ni}) - F(y + \delta_{ni})| \\
&\le \sup_{\substack{x,y\in\mathbb{R} \\ \bar{d}_F(x,y)\le 21\delta}} \sum_{i=1}^n |F(x + \delta_{ni}) - F(y + \delta_{ni})|.
\end{aligned}$$

Let $x, y \in \mathbb{R}$ with $\bar{d}_F(x, y) \le 21\delta$. Then $d_F(x, y) = |F(x) - F(y)|^{1/2} \le 21\delta$ as well, and

$$\begin{aligned}
&\sum_{i=1}^n |F(x + \delta_{ni}) - F(y + \delta_{ni})| \\
&\le \sum_{i=1}^n |F(x + \delta_{ni}) - F(x)| + \sum_{i=1}^n |F(x) - F(y)| + \sum_{i=1}^n |F(y) - F(y + \delta_{ni})| \\
&\le \|f\|_\infty \sum_{i=1}^n |\delta_{ni}| + n \cdot d_F(x, y)^2 + \|f\|_\infty \sum_{i=1}^n |\delta_{ni}| \\
&\le \Big(2 \sup_{F\in M} \|f\|_\infty \frac{1}{n} \sum_{i=1}^n |\delta_{ni}|\Big) \cdot n + n(21\delta)^2 \\
&= O_P^u(1/\sqrt{n}) + n \cdot 441\delta^2
\end{aligned}$$

because of (5.14) and assumption (iii).

It follows that

$$\max_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i(x,y)^2|\mathcal{F}_{i-1}\big) \le \sup_{\substack{x,y\in\mathbb{R}\\ \bar d_F(x,y)\le 21\delta}} \sum_{i=1}^n |F(x+\delta_{ni})-F(y+\delta_{ni})|$$
$$\le O_P^u(n^{-1/2}) + n\cdot 441\delta^2.$$

So for every $\epsilon>0$, $\delta\in(0,1/3)$, $F\in M$ and $n\in\mathbb{N}$ with (5.50) it is

$$\begin{aligned}
&P_F\Big(\max_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} |U_{n,F}(x)-U_{n,F}(y)|\ge 16\epsilon\Big)\\
&\le P_F\Big(\Big\{\max_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n \zeta_i(x,y)\Big|\ge 16\epsilon\Big\}\cap\Big\{\max_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i(x,y)^2|\mathcal{F}_{i-1}\big)\le n\cdot 442\delta^2\Big\}\Big)\\
&\quad + P_F\Big(\max_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i(x,y)^2|\mathcal{F}_{i-1}\big) > n\cdot 442\delta^2\Big)\\
&\le \sum_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} P_F\Big(\Big\{\Big|\sum_{i=1}^n \zeta_i(x,y)\Big|\ge 16\sqrt{n}\epsilon\Big\}\cap\Big\{\sum_{i=1}^n \mathrm{E}_F\big(\zeta_i(x,y)^2|\mathcal{F}_{i-1}\big)\le n\cdot 442\delta^2\Big\}\Big)\\
&\quad + P_F\big(O_P^u(1/\sqrt{n}) + n\cdot 441\delta^2 > n\cdot 442\delta^2\big)\\
&\le \sum_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} 2\cdot\exp\Big(-\frac{(16\sqrt{n}\epsilon)^2}{2n\cdot 442\delta^2}+\frac{1}{2}\frac{(16\sqrt{n}\epsilon)^3}{(n442\delta^2)^2}\Big) + P_F\big(o_P^u(1)>\delta^2\big) =: I
\end{aligned}$$

by (A.4) of Lemma A.4. Now note that for every $F\in M$ we have $|D(\delta_0)|^2\le(2\cdot j(\delta_0))^2\le(2\cdot\delta_0^{-2})^2<4\cdot\delta^{-4}$ by definition of δ_0, $D(\delta_0)$ and $j(\delta_0)$. Thus,

$$\begin{aligned}
I &\le |D(\delta_0)|^2\cdot 2\cdot\exp\Big(-\frac{64\epsilon^2}{221\delta^2}\Big)\cdot\exp\Big(\frac{512\epsilon^3}{48841\sqrt{n}\delta^4}\Big) + P_F\big(o_P^u(1)>\delta^2\big)\\
&\le 8\cdot\frac{1}{\delta^4}\exp\Big(-\frac{64\epsilon^2}{221\delta^2}\Big)\cdot\exp\Big(\frac{\epsilon^3}{\sqrt{n}\delta^4}\Big) + \sup_{F\in M} P_F\big(o_P^u(1)>\delta^2\big).
\end{aligned}$$

This yields

$$\limsup_{n\to\infty}\sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D(\delta_0)\\ \bar d_F(x,y)\le 21\delta}} |U_{n,F}(x)-U_{n,F}(y)|\ge 16\epsilon\Big)\le 8\cdot\frac{1}{\delta^4}\exp\Big(-\frac{64\epsilon^2}{221\delta^2}\Big) \tag{5.59}$$

for every $\epsilon>0$ and $\delta\in(0,1/3)$.

Next, we will study the term in (5.56). For this, let $\epsilon>0$, $\delta\in(0,1/3)$, $F\in M$ and $n\in\mathbb{N}$ with (5.50). For every $x\in D_n$ it is

$$\begin{aligned}
&|U_{n,F}(x)-U_{n,F}(s_0(x))| = |U_{n,F}(s_k(x))-U_{n,F}(s_0(x))| \le \sum_{l=1}^k \big|U_{n,F}(s_l(x))-U_{n,F}(s_{l-1}(x))\big|\\
&= \sum_{l=1}^k \big|U_{n,F}(s_l(x))-U_{n,F}\big(N_l(s_l(x))\big)\big| \le \sum_{l=1}^k \max_{x\in D(\delta_l)}\big|U_{n,F}(x)-U_{n,F}\big(N_l(x)\big)\big|,
\end{aligned}$$

so that

$$\max_{x\in D_n}|U_{n,F}(x)-U_{n,F}(s_0(x))|\le\sum_{l=1}^k\max_{x\in D(\delta_l)}\big|U_{n,F}(x)-U_{n,F}\big(N_l(x)\big)\big|.$$

For every $l=1,\dots,k$ and $x\in D(\delta_l)$ set

$$\zeta_i^l(x):=1_{\{e_i\le x+\delta_{ni}\}}-F(x+\delta_{ni})-\big(1_{\{e_i\le N_l(x)+\delta_{ni}\}}-F(N_l(x)+\delta_{ni})\big),\quad i=1,\dots,n.$$

Then $\big|U_{n,F}(x)-U_{n,F}\big(N_l(x)\big)\big|=\frac{1}{\sqrt{n}}\big|\sum_{i=1}^n\zeta_i^l(x)\big|$. As before, it follows from (5.18) that under P_F the random variables $\zeta_1^l(x),\dots,\zeta_n^l(x)$ form a MDS with respect to $\mathcal{F}_0\subset\mathcal{F}_1\subset\ldots\subset\mathcal{F}_n$, and we see as in (5.58) that $|\zeta_i^l(x)|\le 1$ for all $i=1,\dots,n$. Moreover, by (5.20) it is

$$\max_{x\in D(\delta_l)}\sum_{i=1}^n\mathrm{E}_F\big(\zeta_i^l(x)^2\big|\mathcal{F}_{i-1}\big)\le\max_{x\in D(\delta_l)}\sum_{i=1}^n|F(x+\delta_{ni})-F(N_l(x)+\delta_{ni})|.$$

Since for each $l=1,\dots,k$ and $x\in D(\delta_l)$ we have

$$|F(x)-F(N_l(x))|=d_F(x,N_l(x))^2\le\bar d_F(x,N_l(x))^2\le\bar d_F(x,J_{\delta_{l-1}}(x))^2\le(2\cdot\delta_{l-1})^2$$

by (5.52) and (5.28), it follows that

$$\begin{aligned}&\sum_{i=1}^n|F(x+\delta_{ni})-F(N_l(x)+\delta_{ni})|\\&\le\sum_{i=1}^n\big|F(x+\delta_{ni})-F(x)-\big(F(N_l(x)+\delta_{ni})-F(N_l(x))\big)\big|+n\big|F(x)-F(N_l(x))\big|\\&\le\sum_{i=1}^n\big|f(\xi_{1i})\delta_{ni}-f(\xi_{2i})\delta_{ni}\big|+4n\delta_{l-1}^2\\&\le\sum_{i=1}^n\Big(|f(\xi_{1i})-f(x)|+|f(x)-f(N_l(x))|+|f(N_l(x))-f(\xi_{2i})|\Big)\big|\delta_{ni}\big|+4n\delta_{l-1}^2\end{aligned}$$

with ξ_{1i} lying between $x+\delta_{ni}$ and x, and ξ_{2i} lying between $N_l(x)+\delta_{ni}$ and $N_l(x)$. Because of $|\xi_{1i}-x|\le|\delta_{ni}|$ and $|\xi_{2i}-N_l(x)|\le|\delta_{ni}|$ it is

$$|f(\xi_{1i})-f(x)|\le\Delta_{n,F}\quad\text{and}\quad|f(N_l(x))-f(\xi_{2i})|\le\Delta_{n,F}.$$

Furthermore, we have

$$\bar d_F(x,N_l(x))\le\bar d_F(x,J_{\delta_{l-1}}(x))\le 2\cdot\delta_{l-1}\le 2\cdot 3\cdot\delta=6\delta,$$

where in the last inequality we used (5.51). This implies

$$|f(x)-f(N_l(x))|\le\sup_{\substack{x,y\in\mathbb{R}\\ \bar d_F(x,y)\le 6\delta}}|f(x)-f(y)|=\omega_F(6\delta).$$

Combining all this, we get for every $l=1,\dots,k$

$$\max_{x\in D(\delta_l)}\sum_{i=1}^n\mathrm{E}_F\big(\zeta_i^l(x)^2\big|\mathcal{F}_{i-1}\big)\le 2\Delta_{n,F}\frac{1}{\sqrt{n}}\sum_{i=1}^n|\delta_{ni}|\sqrt{n}+\omega_F(6\delta)\frac{1}{\sqrt{n}}\sum_{i=1}^n|\delta_{ni}|\sqrt{n}+4n\delta_{l-1}^2.$$

By assumption (iii) for $\epsilon>0$ there is a $c_\epsilon\in(0,\infty)$ such that

$$\limsup_{n\to\infty}\sup_{F\in M}P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n|\delta_{ni}|>c_\epsilon\Big)\le\frac{\epsilon}{2}.\tag{5.60}$$

Additionally, by (5.16) there is a $\delta'_\epsilon > 0$ with

$$\sup_{F\in M} \omega_F(6\delta) \le \frac{\epsilon}{c_\epsilon} \quad \forall\, \delta \in (0, \delta'_\epsilon).$$

Now if $2\Delta_{n,F}\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| \le \epsilon$ and $\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| \le c_\epsilon$, then for every $l = 1, \dots, k$

$$\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \le 2\epsilon\sqrt{n} + 4n\delta_{l-1}^2 = 2\alpha^2 n + 4n\delta_{l-1}^2 \le 2\delta_{l-1}^2 n + 4n\delta_{l-1}^2 = 6n\delta_{l-1}^2$$

for $\delta \in (0, \delta'_\epsilon \wedge 1/3)$ by definition of α and δ_l. Thus, for all $F \in M$ it is

$$\begin{aligned}
&P_F\Big(\max_{x\in D_n} |U_{n,F}(x) - U_{n,F}(s_0(x))| \ge 8\epsilon\Big) \\
&\le P_F\Big(\Big\{\sum_{l=1}^k \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \ge 8\epsilon\Big\} \cap \Big\{2\Delta_{n,F}\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| \le \epsilon\Big\} \\
&\qquad \cap \Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| \le c_\epsilon\Big\}\Big) + P_F\Big(2\Delta_{n,F}\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| > \epsilon\Big) + P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| > c_\epsilon\Big) \\
&\le P_F\Big(\Big\{\sum_{l=1}^k \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \ge 8\epsilon\Big\} \cap \\
&\qquad \bigcap_{l=1}^k \Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \le 6n\delta_{l-1}^2\Big\}\Big) \\
&\quad + P_F\Big(2\Delta_{n,F}\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| > \epsilon\Big) + P_F\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| > c_\epsilon\Big).
\end{aligned}$$

Now $\Delta_{n,F}\frac{1}{\sqrt{n}}\sum_{i=1}^n |\delta_{ni}| = o_P^u(1)$ because of (5.37) and assumption (iii). Using this and (5.60), it follows that for every $\delta \in (0, \delta'_\epsilon \wedge 1/3)$

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{x\in D_n} |U_{n,F}(x) - U_{n,F}(s_0(x))| \ge 8\epsilon\Big) \\
&\le \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\Big\{\sum_{l=1}^k \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \ge 8\epsilon\Big\} \cap \\
&\qquad \bigcap_{l=1}^k \Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \le 6n\delta_{l-1}^2\Big\}\Big) + \frac{\epsilon}{2}.
\end{aligned} \tag{5.61}$$

Now set

$$v_l := 6 \cdot \delta_{l-1} \cdot |\log(\delta_{l-1})|^{1/2} \quad \forall\, l = 1, \dots, k.$$

Then

$$\begin{aligned}
\sum_{l=1}^k v_l &= 6 \cdot \sum_{l=1}^k \delta_{l-1}^{1/2}\big(\delta_{l-1}|\log(\delta_{l-1})|\big)^{1/2} \le 6 \cdot \sum_{l=1}^k \delta_{l-1}^{1/2} = 6 \cdot \sum_{l=0}^{k-1} (3^{k-l}\alpha)^{1/2} \\
&= 6 \cdot (3^k\alpha)^{1/2} \sum_{l=0}^{k-1} (3^{-l})^{1/2} \le 6 \cdot (3\delta)^{1/2} \sum_{l=0}^{\infty} (\sqrt{3})^{-l} = \frac{18\sqrt{\delta}}{\sqrt{3}-1}.
\end{aligned}$$

Here we used that $x|\log(x)| \le 1$ for $x \in (0,1]$, the definition of δ_l and the fact that $3^k\alpha \le 3\delta$. Thus, there is a $\delta''_\epsilon > 0$ such that

$$\sum_{l=1}^k v_l \le 8\epsilon \quad \forall\, \delta \in (0, \delta''_\epsilon \wedge 1/3).$$

This implies that for every $\delta \in \big(0, \delta'_\epsilon \wedge \delta''_\epsilon \wedge 1/3\big)$, $F \in M$ and all $n \in \mathbb{N}$ with (5.50) we have

$$\begin{aligned}
&P_F\Big(\Big\{\sum_{l=1}^k \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \geq 8\epsilon\Big\} \cap \bigcap_{l=1}^k \Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big) \\
&\leq P_F\Big(\Big\{\sum_{l=1}^k \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \geq \sum_{l=1}^k v_l\Big\} \cap \\
&\qquad \bigcap_{l=1}^k \Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big) \\
&\leq \sum_{l=1}^k P_F\Big(\Big\{\max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \geq v_l\Big\} \cap \Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big) \\
&\leq \sum_{l=1}^k \sum_{x\in D(\delta_l)} P_F\Big(\Big\{\big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \geq v_l\Big\} \cap \Big\{\sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big).
\end{aligned}$$

Now recall that $\big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| = \frac{1}{\sqrt{n}}\big|\sum_{i=1}^n \zeta_i^l(x)\big|$. By using (A.4) of Lemma A.4 we get for every $l = 1, \ldots, k$ and $x \in D(\delta_l)$

$$\begin{aligned}
&P_F\Big(\Big\{\Big|\sum_{i=1}^n \zeta_i^l(x)\Big| \geq \sqrt{n}v_l\Big\} \cap \Big\{\sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big) \\
&\leq 2\exp\Big(-\frac{(\sqrt{n}v_l)^2}{2\cdot 6n\delta_{l-1}^2} + \frac{1}{2}\frac{(\sqrt{n}v_l)^3}{(6n\delta_{l-1}^2)^2}\Big) = 2\cdot\exp\Big(-\frac{v_l^2}{12\delta_{l-1}^2}\Big)\cdot\exp\Big(\frac{v_l^3}{72\sqrt{n}\delta_{l-1}^4}\Big).
\end{aligned}$$

Observe that because of (5.51) it is for every $l = 1, \ldots, k$

$$\begin{aligned}
\frac{v_l^3}{\sqrt{n}\delta_{l-1}^4} &= \frac{6^3}{\sqrt{n}} \cdot \frac{|\log(\delta_{l-1})|^{3/2}}{\delta_{l-1}} \leq \frac{6^3}{\sqrt{n}} \cdot \frac{|\log(\delta_k)|^{3/2}}{\delta_k} = \frac{6^3}{\sqrt{n}} \cdot \frac{|\log((\epsilon/\sqrt{n})^{1/2})|^{3/2}}{(\epsilon/\sqrt{n})^{1/2}} \\
&= \frac{6^3}{\sqrt{\epsilon}} \cdot \frac{1}{n^{1/4}} \cdot \Big(\frac{1}{2}\Big)^{3/2} |\log(\epsilon) - \log(\sqrt{n})|^{3/2} =: r_n(\epsilon) \underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Also, for every $F \in M$ we have $|D(\delta_l)| \leq 2j(\delta_l) \leq 2/\delta_l^2$ by definition of $D(\delta_l)$ and $j(\delta_l)$. Combining all this shows that

$$\begin{aligned}
&P_F\Big(\Big\{\sum_{l=1}^k \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \geq 8\epsilon\Big\} \cap \bigcap_{l=1}^k \Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^n \mathrm{E}_F\big(\zeta_i^l(x)^2 \big| \mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big) \\
&\leq 4\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) \sum_{l=1}^k \frac{1}{\delta_l^2}\cdot\exp\Big(-\frac{v_l^2}{12\delta_{l-1}^2}\Big) \\
&= 4\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) \sum_{l=1}^k \frac{1}{\delta_l^2}\cdot\exp\Big(-\frac{(6\cdot\delta_{l-1}\cdot|\log(\delta_{l-1})|^{1/2})^2}{12\delta_{l-1}^2}\Big) = 4\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) \sum_{l=1}^k \frac{1}{\delta_l^2}\cdot\delta_{l-1}^3 \\
&= 4\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) \sum_{l=1}^k \Big(\frac{\delta_{l-1}}{\delta_l}\Big)^2 \cdot \delta_{l-1} = 4\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) \sum_{l=1}^k 3^2\cdot\delta_{l-1} \\
&= 36\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) \sum_{l=0}^{k-1} 3^{k-l}\alpha \leq 36\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big) 3^k\alpha \sum_{l=0}^{\infty} 3^{-l} \\
&\leq 36\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big)\cdot 3\delta\cdot\frac{3}{2} = 162\cdot\exp\Big(\frac{r_n(\epsilon)}{72}\Big)\cdot\delta.
\end{aligned}$$

Thus it is for every $\delta \in \big(0, \delta'_\epsilon \wedge \delta''_\epsilon \wedge 1/3\big)$

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\Big\{\sum_{l=1}^{k} \max_{x\in D(\delta_l)} \big|U_{n,F}(x) - U_{n,F}\big(N_l(x)\big)\big| \geq 8\epsilon\Big\} \cap \bigcap_{l=1}^{k}\Big\{\max_{x\in D(\delta_l)} \sum_{i=1}^{n} \mathrm{E}_F\big(\zeta_i^l(x)^2 \big|\mathcal{F}_{i-1}\big) \leq 6n\delta_{l-1}^2\Big\}\Big) \leq 162\delta,$$

and so

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{x\in D_n} |U_{n,F}(x) - U_{n,F}(s_0(x))| \geq 8\epsilon\Big) \leq 162\delta + \frac{\epsilon}{2} \tag{5.62}$$

because of (5.61).

Using the bounds in (5.62) and (5.59), it follows now from (5.56) and (5.57) that for every $\delta \in \big(0, \delta'_\epsilon \wedge \delta''_\epsilon \wedge 1/3\big)$

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n\\ \bar{d}_F(x,y)\leq 3\delta}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\epsilon\Big) \leq 162\delta + \frac{\epsilon}{2} + 8\cdot\frac{1}{\delta^4}\exp\Big(-\frac{64\epsilon^2}{221\delta^2}\Big).$$

But since

$$162\delta + 8\cdot\frac{1}{\delta^4}\exp\Big(-\frac{64\epsilon^2}{221\delta^2}\Big) \underset{\delta\downarrow 0}{\longrightarrow} 0,$$

there is obviously a $\delta_\epsilon \in (0,1)$ such that

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\max_{\substack{x,y\in D_n\\ \bar{d}_F(x,y)\leq 3\delta_\epsilon}} |U_{n,F}(x) - U_{n,F}(y)| \geq 32\epsilon\Big) \leq \epsilon,$$

which is just the statement of Step 3b in (5.49). This concludes the proof of the theorem. □

We are now able to formulate the proof of Theorem 5.9.

Proof of Theorem 5.9. Let $F \in M$, $n \in \mathbb{N}$ and $x \in \mathbb{R}$. Then

$$\begin{aligned} F_{n,res}(x) - F_n(x) &= \frac{1}{n}\sum_{i=1}^{n}\big[1_{\{e_i\leq x+e_i-\hat{e}_{ni}\}} - F(x+e_i-\hat{e}_{ni}) - \big(1_{\{e_i\leq x\}} - F(x)\big)\big] \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\big[F(x+e_i-\hat{e}_{ni}) - F(x)\big], \end{aligned}$$

and by the mean value theorem there is a ξ_i between $x + e_i - \hat{e}_{ni}$ and x such that

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\big[F(x+e_i-\hat{e}_{ni}) - F(x)\big] &= \frac{1}{n}\sum_{i=1}^{n} f(\xi_i)(e_i-\hat{e}_{ni}) \\ &= \frac{1}{n}\sum_{i=1}^{n}\big(f(\xi_i)-f(x)\big)(e_i-\hat{e}_{ni}) + f(x)\frac{1}{n}\sum_{i=1}^{n}(e_i-\hat{e}_{ni}), \end{aligned}$$

so that

$$\sqrt{n}\sup_{x\in\mathbb{R}}|F_{n,res}(x) - F_n(x)| \leq \sup_{x\in\mathbb{R}}\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big[1_{\{e_i\leq x+e_i-\hat{e}_{ni}\}} - F(x+e_i-\hat{e}_{ni}) - \big(1_{\{e_i\leq x\}} - F(x)\big)\big]\Big|$$

$$+ \sup_{\substack{x,y\in\mathbb{R}\\ |x-y|\le \max\limits_{1\le i\le n}|e_i-\hat{e}_{ni}|}} |f(x)-f(y)| \cdot \frac{1}{\sqrt{n}}\sum_{i=1}^n |e_i-\hat{e}_{ni}|$$

$$+ \|f\|_\infty \cdot \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n (e_i-\hat{e}_{ni})\Big|. \tag{5.63}$$

Since

$$e_i-\hat{e}_{ni} = (\hat{\rho}_n-\rho)^T \mathbf{X}_{i-1} \ \ \forall\, i=1,\ldots,n,\ n\in\mathbb{N},$$

using the Cauchy-Schwarz inequality we get

$$\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n (e_i-\hat{e}_{ni})\Big| = \frac{1}{\sqrt{n}}\Big|(\hat{\rho}_n-\rho)^T\sum_{i=1}^n \mathbf{X}_{i-1}\Big| \le \big\|\sqrt{n}(\hat{\rho}_n-\rho)\big\| \cdot \Big\|\frac{1}{n}\sum_{i=1}^n \mathbf{X}_{i-1}\Big\| = o_P^u(1)$$

by assumptions (i) and (ii). Hence, it follows with (5.14) that

$$\|f\|_\infty \cdot \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n (e_i-\hat{e}_{ni})\Big| \le \sup_{F\in M}\|f\|_\infty \cdot \frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n (e_i-\hat{e}_{ni})\Big| = o_P^u(1).$$

Also,

$$\max_{1\le i\le n}|e_i-\hat{e}_{ni}| \le \|\hat{\rho}_n-\rho\| \max_{1\le i\le n}\|\mathbf{X}_{i-1}\| = O_P^u(1/\sqrt{n})o_P^u(\sqrt{n}) = o_P^u(1) \tag{5.64}$$

with (i) and (iii). This yields

$$\sup_{\substack{x,y\in\mathbb{R}\\ |x-y|\le \max\limits_{1\le i\le n}|e_i-\hat{e}_{ni}|}} |f(x)-f(y)| = o_P^u(1) \text{ in } M \text{ as } n\to\infty. \tag{5.65}$$

To see that (5.65) is true, note that because of the uniform equicontinuity (5.12) of the family $\{f\colon F\in M\}$ for every $\epsilon>0$ there is a $\delta>0$ such that for every $x,y\in\mathbb{R}$ with $|x-y|\le\delta$

$$\sup_{F\in M}|f(x)-f(y)|\le\epsilon.$$

So if $\max_{1\le i\le n}|e_i-\hat{e}_{ni}|\le\delta$, then

$$\sup_{\substack{x,y\in\mathbb{R}\\ |x-y|\le \max\limits_{1\le i\le n}|e_i-\hat{e}_{ni}|}} |f(x)-f(y)| \le \epsilon$$

for all $F\in M$. Thus,

$$\sup_{F\in M} P_F\Big(\sup_{\substack{x,y\in\mathbb{R}\\ |x-y|\le \max\limits_{1\le i\le n}|e_i-\hat{e}_{ni}|}} |f(x)-f(y)|>\epsilon\Big) \le \sup_{F\in M} P_F\Big(\max_{1\le i\le n}|e_i-\hat{e}_{ni}|>\delta\Big) \underset{n\to\infty}{\longrightarrow} 0$$

because of (5.64). This shows (5.65).
Moreover, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n |e_i-\hat{e}_{ni}| \le \sqrt{n}\|\hat{\rho}_n-\rho\|\frac{1}{n}\sum_{i=1}^n\|\mathbf{X}_{i-1}\| = O_P^u(1)$$

because of assumptions (i) and (iv). This shows that the second summand on the right-hand side of inequality (5.63) is a $o_P^u(1)$ in M as $n\to\infty$ as well. Therefore it remains to show that

$$\sup_{x\in\mathbb{R}}\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^n \big[1_{\{e_i\le x+e_i-\hat{e}_{ni}\}} - F(x+e_i-\hat{e}_{ni}) - \big(1_{\{e_i\le x\}}-F(x)\big)\big]\Big| = o_P^u(1) \text{ in } M \text{ as } n\to\infty. \tag{5.66}$$

For the proof of (5.66), let $F \in M$. For every $x \in \mathbb{R}$, $n \in \mathbb{N}$ and $t \in \mathbb{R}^p$ set

$$W_n(x,t) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \le x + (t-\rho)^T \mathbf{X}_{i-1}\}} - F\big(x + (t-\rho)^T \mathbf{X}_{i-1}\big) \big].$$

Then we can write

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \le x + e_i - \hat{e}_{ni}\}} - F(x + e_i - \hat{e}_{ni}) - \big(1_{\{e_i \le x\}} - F(x)\big)\big] = W_n(x, \hat{\rho}_n) - W_n(x,\rho)$$
$$= W_n\big(x, \rho + \frac{1}{\sqrt{n}}\big[\sqrt{n}(\hat{\rho}_n - \rho)\big]\big) - W_n(x,\rho).$$

Now define

$$V_n(x,s) := W_n\big(x, \rho + \frac{1}{\sqrt{n}} s\big) - W_n(x,\rho), \quad x \in \mathbb{R},\ s \in \mathbb{R}^p.$$

Then (5.66) will follow from

$$\sup_{\substack{x \in \mathbb{R} \\ s \in \mathbb{R}^p, \|s\| \le C}} |V_n(x,s)| = o_P^u(1) \ \text{ in } M \text{ as } n \to \infty \quad \forall\, C \in (0,\infty). \tag{5.67}$$

For showing that (5.67) implies (5.66), let $F \in M$ and $C \in (0,\infty)$. Then if $\sqrt{n}\|\hat{\rho}_n - \rho\| \le C$, we have for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$

$$\Big| W_n\big(x, \rho + \frac{1}{\sqrt{n}}\big[\sqrt{n}(\hat{\rho}_n - \rho)\big]\big) - W_n(x,\rho) \Big| = \big|V_n\big(x, \sqrt{n}(\hat{\rho}_n - \rho)\big)\big| \le \sup_{\substack{x \in \mathbb{R} \\ s \in \mathbb{R}^p, \|s\| \le C}} |V_n(x,s)|.$$

So for every $\epsilon > 0$, $n \in \mathbb{N}$ and $C \in (0,\infty)$

$$\sup_{F \in M} P_F\Big(\sup_{x \in \mathbb{R}} \Big| W_n\big(x, \rho + \frac{1}{\sqrt{n}}\big[\sqrt{n}(\hat{\rho}_n - \rho)\big]\big) - W_n(x,\rho) \Big| \ge \epsilon \Big)$$
$$\le \sup_{F \in M} P_F\Big(\sup_{\substack{x \in \mathbb{R} \\ s \in \mathbb{R}^p, \|s\| \le C}} |V_n(x,s)| \ge \epsilon \Big) + \sup_{F \in M} P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > C\big),$$

and this yields

$$\limsup_{n \to \infty} \sup_{F \in M} P_F\Big(\sup_{x \in \mathbb{R}} \Big| W_n\big(x, \rho + \frac{1}{\sqrt{n}}\big[\sqrt{n}(\hat{\rho}_n - \rho)\big]\big) - W_n(x,\rho) \Big| \ge \epsilon \Big)$$
$$\le \limsup_{n \to \infty} \sup_{F \in M} P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > C\big) \underset{C \to \infty}{\longrightarrow} 0$$

by (5.67) and assumption (i).

Next, we will show that the two conditions

$$\sup_{x \in \mathbb{R}} |V_n(x,s)| = o_P^u(1) \text{ in } M \text{ as } n \to \infty \ \ \forall\, s \in \mathbb{R}^p \tag{5.68}$$

and

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \sup_{F \in M} P_F\Big(\sup_{\substack{x \in \mathbb{R} \\ s,t \in \mathbb{R}^p \\ \|s\|,\|t\| \le C \\ \|s-t\| \le \delta}} |V_n(x,s) - V_n(x,t)| \ge \epsilon \Big) = 0 \quad \forall\, \epsilon > 0,\ C \in (0,\infty) \tag{5.69}$$

together imply (5.67), and thus (5.66).

To simplify notation, we denote in the following by $B_C(0)$ the closed ball $\{s \in \mathbb{R}^p \colon \|s\| \leq C\}$.

For the proof of the above statement, assume that (5.68) and (5.69) hold. Let $C \in (0,\infty)$ and $\delta > 0$. Since $B_C(0)$ is compact, there are $s_1, \ldots, s_k \in B_C(0)$, $k \in \mathbb{N}$, such that for every $s \in B_C(0)$ there is a $j \in \{1, \ldots, k\}$ with $\|s - s_j\| < \delta$.

Now let $x \in \mathbb{R}$, $s \in \mathbb{R}^p$ with $\|s\| \leq C$ and $j \in \{1, \ldots, k\}$ such that $\|s - s_j\| < \delta$. Then for every $F \in M$ and $n \in \mathbb{N}$

$$\begin{aligned} |V_n(x,s)| &\leq |V_n(x,s) - V_n(x,s_j)| + |V_n(x,s_j)| \\ &\leq \sup_{\substack{x\in\mathbb{R} \\ s,t\in B_C(0) \\ \|s-t\|\leq\delta}} |V_n(x,s) - V_n(x,t)| + \max_{1\leq j\leq k} \sup_{x\in\mathbb{R}} |V_n(x,s_j)|, \end{aligned}$$

whence it follows that for every $\epsilon, \delta > 0$, $n \in \mathbb{N}$ and $F \in M$

$$\begin{aligned} &P_F\big(\sup_{\substack{x\in\mathbb{R} \\ s\in B_C(0)}} |V_n(x,s)| \geq \epsilon \big) \\ &\leq P_F\Big(\sup_{\substack{x\in\mathbb{R} \\ s,t\in B_C(0) \\ \|s-t\|\leq\delta}} |V_n(x,s) - V_n(x,t)| \geq \frac{\epsilon}{2} \Big) + P_F\Big(\max_{1\leq j\leq k} \sup_{x\in\mathbb{R}} |V_n(x,s_j)| \geq \frac{\epsilon}{2} \Big) \\ &\leq P_F\Big(\sup_{\substack{x\in\mathbb{R} \\ s,t\in B_C(0) \\ \|s-t\|\leq\delta}} |V_n(x,s) - V_n(x,t)| \geq \frac{\epsilon}{2} \Big) + \sum_{j=1}^k P_F\Big(\sup_{x\in\mathbb{R}} |V_n(x,s_j)| \geq \frac{\epsilon}{2} \Big). \end{aligned}$$

Using (5.68), this yields

$$\limsup_{n\to\infty} \sup_{F\in M} P_F\big(\sup_{\substack{x\in\mathbb{R} \\ s\in B_C(0)}} |V_n(x,s)| \geq \epsilon \big) \leq \limsup_{n\to\infty} \sup_{F\in M} P_F\Big(\sup_{\substack{x\in\mathbb{R} \\ s,t\in B_C(0) \\ \|s-t\|\leq\delta}} |V_n(x,s) - V_n(x,t)| \geq \frac{\epsilon}{2} \Big),$$

and the term on the right-hand side of this inequality converges to zero as $\delta \to 0$ by (5.69). This concludes the proof of (5.68), (5.69) $\Rightarrow$ (5.67).

It remains to show that (5.68) and (5.69) hold.

First, we investigate condition (5.68). For every $x \in \mathbb{R}$, $s \in \mathbb{R}^p$, $n \in \mathbb{N}$ and $F \in M$ it is

$$\begin{aligned} V_n(x,s) &= W_n\big(x, \rho + \frac{1}{\sqrt{n}} s\big) - W_n(x,\rho) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq x + s^T \mathbf{X}_{i-1}/\sqrt{n}\}} - F\big(x + s^T \mathbf{X}_{i-1}/\sqrt{n}\big) \big] - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq x\}} - F(x) \big] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \leq x + \delta_{ni}\}} - F\big(x + \delta_{ni}\big) - \big(1_{\{e_i \leq x\}} - F(x)\big) \big] \end{aligned}$$

with

$$\delta_{ni} := \frac{s^T \mathbf{X}_{i-1}}{\sqrt{n}}, \quad 1 \leq i \leq n.$$

Set $\mathcal{F}_0 := \sigma(\mathbf{X}_0)$, $\mathcal{F}_n := \sigma(\mathbf{X_0}, e_1, \ldots, e_n)$ for $n \geq 1$. Then the random variables $\delta_{n1}, \ldots, \delta_{nn}$ are predictable with respect to $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n$. It also is

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n |\delta_{ni}| = \frac{1}{n} \sum_{i=1}^n |s^T \mathbf{X}_{i-1}| \leq \|s\| \frac{1}{n} \sum_{i=1}^n \|\mathbf{X}_{i-1}\| = O_P^u(1)$$

by assumption (iv). Moreover,

$$\max_{1\le i\le n} |\delta_{ni}| \le \frac{1}{\sqrt{n}}\|s\| \max_{1\le i\le n} \|\mathbf{X}_{i-1}\| = o_P^u(1)$$

by (iii). Additionally we have

$$\mathrm{E}_F\Big(\frac{1}{n}\sum_{i=1}^n |\delta_{ni}|\Big) = \frac{1}{n^{3/2}}\,\mathrm{E}_F\Big(\sum_{i=1}^n |s^T\mathbf{X}_{i-1}|\Big) \le \|s\|\frac{1}{n^{3/2}}\,\mathrm{E}_F\Big(\sum_{i=1}^n \|\mathbf{X}_{i-1}\|\Big),$$

so that

$$\sup_{F\in M} \mathrm{E}_F\Big(\frac{1}{n}\sum_{i=1}^n |\delta_{ni}|\Big) \le \|s\|\frac{1}{n^{3/2}} \sup_{F\in M} \mathrm{E}_F\Big(\sum_{i=1}^n \|\mathbf{X}_{i-1}\|\Big) = o(1) \text{ as } n\to\infty$$

because of (v). Thus, all assumptions of Theorem 5.10 are satisfied, and it follows for every $s \in \mathbb{R}^p$ that

$$\sup_{x\in\mathbb{R}} |V_n(x,s)| = o_P^u(1) \text{ in } M \text{ as } n\to\infty,$$

which is just (5.68).

Next, we show (5.69). Let $F \in M$, $n \in \mathbb{N}$, $x \in \mathbb{R}$, $C \in (0,\infty)$, $\delta > 0$ and $s, t \in \mathbb{R}^p$ with $\|s\|, \|t\| \le C$ and $\|s-t\| \le \delta$. It is

$$\begin{aligned} V_n(x,s) - V_n(x,t) &= W_n\big(x, \rho + \tfrac{1}{\sqrt{n}}s\big) - W_n\big(x, \rho + \tfrac{1}{\sqrt{n}}t\big) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le x + s^T\mathbf{X}_{i-1}/\sqrt{n}\}} - F\big(x + s^T\mathbf{X}_{i-1}/\sqrt{n}\big)\big] \\ &\quad - \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[1_{\{e_i \le x + t^T\mathbf{X}_{i-1}/\sqrt{n}\}} - F\big(x + t^T\mathbf{X}_{i-1}/\sqrt{n}\big)\big]. \end{aligned}$$

As shown before, there is a $k \in \mathbb{N}$ and $s_1, \dots, s_k \in B_C(0)$ such that for every $s \in \mathbb{R}^p$ with $\|s\| \le C$ there is a $j \in \{1, \dots, k\}$ with $\|s - s_j\| < \delta$.
Now let s be as above and $j \in \{1,\dots,k\}$ such that $\|s - s_j\| < \delta$. Then it is for every $x \in \mathbb{R}$ and $i = 1, \dots, n$

$$x + \frac{s^T\mathbf{X}_{i-1}}{\sqrt{n}} = x + \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} + \frac{(s-s_j)^T\mathbf{X}_{i-1}}{\sqrt{n}} \le x + \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} + \frac{\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}}$$

and

$$x + \frac{s^T\mathbf{X}_{i-1}}{\sqrt{n}} \ge x + \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} - \frac{\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}}.$$

For t as above we have $\|s-t\| \le \delta$, and so $\|t - s_j\| \le \|t-s\| + \|s-s_j\| \le 2\delta$. This implies for every $x \in \mathbb{R}$ and $i = 1, \dots, n$ that

$$x + \frac{t^T\mathbf{X}_{i-1}}{\sqrt{n}} \le x + \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} + \frac{2\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}}$$

and

$$x + \frac{t^T\mathbf{X}_{i-1}}{\sqrt{n}} \ge x + \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} - \frac{2\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}}.$$

Set

$$\delta_{ni}^+(j) := \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} + \frac{2\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}} \quad\text{and}\quad \delta_{ni}^-(j) := \frac{s_j^T\mathbf{X}_{i-1}}{\sqrt{n}} - \frac{2\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, k$. Then obviously for every $x \in \mathbb{R}$ and $i = 1, \ldots, n$ it is

$$x + \delta_{ni}^-(j) \le x + \frac{s^T \mathbf{X}_{i-1}}{\sqrt{n}} \le x + \delta_{ni}^+(j)$$

and

$$x + \delta_{ni}^-(j) \le x + \frac{t^T \mathbf{X}_{i-1}}{\sqrt{n}} \le x + \delta_{ni}^+(j).$$

Therefore

$$\begin{aligned}
&V_n(x,s) - V_n(x,t) \\
&\le \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}^+(j)\}} - F\big(x + \delta_{ni}^-(j)\big)\big] - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}^-(j)\}} - F\big(x + \delta_{ni}^+(j)\big)\big] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}^+(j)\}} - F\big(x + \delta_{ni}^+(j)\big)\big] - \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}^-(j)\}} - F\big(x + \delta_{ni}^-(j)\big)\big] \\
&\quad + 2 \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F\big(x + \delta_{ni}^+(j)\big) - F\big(x + \delta_{ni}^-(j)\big)\big] \\
&\le \max_{1 \le j \le k} R_n^+(j) + \max_{1 \le j \le k} R_n^-(j) + 2 \max_{1 \le j \le k} \sup_{x \in \mathbb{R}} \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F\big(x + \delta_{ni}^+(j)\big) - F\big(x + \delta_{ni}^-(j)\big)\big],
\end{aligned}$$

where

$$R_n^+(j) := \sup_{x \in \mathbb{R}} \frac{1}{\sqrt{n}} \Big| \sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}^+(j)\}} - F\big(x + \delta_{ni}^+(j)\big) - \big(1_{\{e_i \le x\}} - F(x)\big)\big] \Big|$$

and

$$R_n^-(j) := \sup_{x \in \mathbb{R}} \frac{1}{\sqrt{n}} \Big| \sum_{i=1}^n \big[1_{\{e_i \le x + \delta_{ni}^-(j)\}} - F\big(x + \delta_{ni}^-(j)\big) - \big(1_{\{e_i \le x\}} - F(x)\big)\big] \Big|.$$

Analogously it follows that

$$\begin{aligned}
&V_n(x,t) - V_n(x,s) \\
&\le \max_{1 \le j \le k} R_n^+(j) + \max_{1 \le j \le k} R_n^-(j) + 2 \max_{1 \le j \le k} \sup_{x \in \mathbb{R}} \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F\big(x + \delta_{ni}^+(j)\big) - F\big(x + \delta_{ni}^-(j)\big)\big].
\end{aligned}$$

This yields

$$\begin{aligned}
&\sup_{\substack{x \in \mathbb{R} \\ s,t \in B_C(0) \\ \|s-t\| \le \delta}} |V_n(x,s) - V_n(x,t)| \\
&\le \max_{1 \le j \le k} R_n^+(j) + \max_{1 \le j \le k} R_n^-(j) + 2 \max_{1 \le j \le k} \sup_{x \in \mathbb{R}} \frac{1}{\sqrt{n}} \sum_{i=1}^n \big[F\big(x + \delta_{ni}^+(j)\big) - F\big(x + \delta_{ni}^-(j)\big)\big]. \qquad (5.70)
\end{aligned}$$

As before, let $\mathcal{F}_0 = \sigma(\mathbf{X}_0)$ and $\mathcal{F}_n = \sigma(\mathbf{X_0}, e_1, \ldots, e_n)$ for $n \ge 1$. Then for every $j = 1, \ldots, k$ the random variables $\delta_{n1}^\pm(j), \ldots, \delta_{nn}^\pm(j)$ are predictable with respect to $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n$. Also, we have for every $j = 1, \ldots, k$

$$\max_{1 \le i \le n} |\delta_{ni}^\pm(j)| = \max_{1 \le i \le n} \Big| \frac{s_j^T \mathbf{X}_{i-1}}{\sqrt{n}} \pm \frac{2\delta \|\mathbf{X}_{i-1}\|}{\sqrt{n}} \Big| \le \frac{1}{\sqrt{n}} \big(\|s_j\| + 2\delta\big) \max_{1 \le i \le n} \|\mathbf{X}_{i-1}\| = o_P^u(1)$$

by assumption (iii) of this theorem. Moreover,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|\delta_{ni}^{\pm}(j)| \le \big(\|s_j\|+2\delta\big)\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\| = O_P^u(1)$$

using (iv). Last, we see that

$$\sup_{F\in M}\mathrm{E}_F\Big(\frac{1}{n}\sum_{i=1}^{n}|\delta_{ni}^{\pm}(j)|\Big) \le \big(\|s_j\|+2\delta\big)\frac{1}{n^{3/2}}\sup_{F\in M}\mathrm{E}_F\Big(\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|\Big) = o(1) \text{ as } n\to\infty$$

because of (v). Combining all this, we see that all assumptions of Theorem 5.10 are met, and it follows from this that

$$R_n^{\pm}(j) = o_P^u(1) \quad \text{in } M \text{ as } n\to\infty$$

for all $j=1,\ldots,k$. But since k neither depends on F nor n, this implies

$$\max_{1\le j\le k} R_n^{\pm}(j) = o_P^u(1) \quad \text{in } M \text{ as } n\to\infty. \tag{5.71}$$

Furthermore, for every $j\in\{1,\ldots,k\}$, $x\in\mathbb{R}$, $n\in\mathbb{N}$, $i\in\{1,\ldots,n\}$ and $F\in M$ it is

$$0\le F\big(x+\delta_{ni}^{+}(j)\big)-F\big(x+\delta_{ni}^{-}(j)\big) \le \|f\|_\infty\big(\delta_{ni}^{+}(j)-\delta_{ni}^{-}(j)\big) \le \big(\sup_{F\in M}\|f\|_\infty\big)\frac{4\delta\|\mathbf{X}_{i-1}\|}{\sqrt{n}}.$$

Hence,

$$2\max_{1\le j\le k}\sup_{x\in\mathbb{R}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[F\big(x+\delta_{ni}^{+}(j)\big)-F\big(x+\delta_{ni}^{-}(j)\big)\big] \le 8\delta\,\big(\sup_{F\in M}\|f\|_\infty\big)\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|.$$

Now it follows from (5.70) that for every $n\in\mathbb{N}$ and $\delta>0$

$$\sup_{\substack{x\in\mathbb{R}\\ s,t\in B_C(0)\\ \|s-t\|\le\delta}} |V_n(x,s)-V_n(x,t)| \;\le\; R_n(\delta) +\; 8\delta\,\big(\sup_{F\in M}\|f\|_\infty\big)\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|$$

with $R_n(\delta)=o_P^u(1)$ in M as $n\to\infty$ by (5.71). So for every $\epsilon>0$, $\delta>0$, $K\in(0,\infty)\ni C$, $n\in\mathbb{N}$ and $F\in M$ it is

$$\begin{aligned}
&P_F\Big(\sup_{\substack{x\in\mathbb{R}\\ s,t\in B_C(0)\\ \|s-t\|\le\delta}} |V_n(x,s)-V_n(x,t)|\ge\epsilon\Big)\\
&\le P_F\Big(R_n(\delta)\ge\frac{\epsilon}{2}\Big)+P_F\Big(\Big\{8\delta\,\big(\sup_{F\in M}\|f\|_\infty\big)\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|\ge\frac{\epsilon}{2}\Big\}\cap\Big\{\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|\le K\Big\}\Big)\\
&\quad+P_F\Big(\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|>K\Big)\\
&\le P_F\Big(R_n(\delta)\ge\frac{\epsilon}{2}\Big)+P_F\Big(\delta\ge\frac{\epsilon}{16K}\big(\sup_{F\in M}\|f\|_\infty\big)^{-1}\Big)+P_F\Big(\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|>K\Big),
\end{aligned}$$

where we used that $\sup_{F\in M}\|f\|_\infty<\infty$ by (5.14). This implies that for every $\epsilon>0$ and $K>0$

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\sup_{F\in M}P_F\Big(\sup_{\substack{x\in\mathbb{R}\\ s,t\in B_C(0)\\ \|s-t\|\le\delta}} |V_n(x,s)-V_n(x,t)|\ge\epsilon\Big)\\
&\le \limsup_{n\to\infty}\sup_{F\in M}P_F\Big(\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|>K\Big),
\end{aligned}$$

and the last term converges to zero as $K\to\infty$ because of (iv). Thus, we have shown that (5.69) holds, and this concludes the proof of Theorem 5.9. □

5.2 The centered empirical distribution function of the residuals

Let us consider the probability space $(\Omega, \mathcal{A}, P_F)$ for some fixed $F \in M$ for the moment. Let $(\hat{\rho}_n)_{n\in\mathbb{N}}$ be a sequence of estimators for the autoregressive parameter $\rho = (\rho_1, \dots, \rho_p)^T$ in (5.2) such that $\sqrt{n}(\hat{\rho}_n - \rho) = O_{P_F}(1)$, and as before let $\hat{e}_{ni} = X_i - \hat{\rho}_n^T \mathbf{X}_{i-1}$, $1 \le i \le n$, $n \in \mathbb{N}$, be the residuals with respect to $(\hat{\rho}_n)_{n\in\mathbb{N}}$. Analogously to the case of independent and identically distributed data described in subsection 3.1, it follows from Lemma A.1 that for every $n \ge 2$ there is a unique $\hat{t}_n = t(\hat{e}_{n1}, \dots, \hat{e}_{nn}) \in \mathbb{R}$ such that

$$\Big(\frac{1}{n} - 1\Big)\frac{1}{\max\limits_{1\le i\le n} \hat{e}_{ni}} < \hat{t}_n < \Big(\frac{1}{n} - 1\Big)\frac{1}{\min\limits_{1\le i\le n} \hat{e}_{ni}} \tag{5.72}$$

and

$$\sum_{i=1}^{n} \frac{\hat{e}_{ni}}{1 + \hat{t}_n \hat{e}_{ni}} = 0 \tag{5.73}$$

on the event

$$\Omega_{n,res} := \{\min_{1\le i\le n} \hat{e}_{ni} < 0 < \max_{1\le i\le n} \hat{e}_{ni}\} \in \mathcal{A},$$

and by Lemma A.2 the mapping

$$\hat{t}_n \colon \Omega_{n,res} \ni \omega \mapsto t\big(\hat{e}_{n1}(\omega), \dots, \hat{e}_{nn}(\omega)\big) \in \mathbb{R}$$

is $\Omega_{n,res} \cap \mathcal{A}, \mathcal{B}^*$-measurable.

Just like the set $\overline{\Omega}_n$ in subsection 3.1, the complement $\overline{\Omega}_{n,res}$ of $\Omega_{n,res}$ is asymptotically negligible, as the following lemma shows. Its proof is a reformulation of the proof of Satz 3.2 in Genz [10] for the autoregressive processes of order p under investigation here, and is given in detail for the reader's convenience.

Lemma 5.11 (cf. Satz 3.2 in [10])
If $F \in M$ and $\sqrt{n}(\hat{\rho}_n - \rho) = O_{P_F}(1)$, then

$$P_F(\overline{\Omega}_{n,res}) = P_F\Big(0 \notin (\min_{1\le i\le n} \hat{e}_{ni}, \max_{1\le i\le n} \hat{e}_{ni})\Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{5.74}$$

Proof. Recall that F is continuous with $\int_{\mathbb{R}} xF(dx) = 0$ and $\int_{\mathbb{R}} x^2 F(dx) < \infty$. Hence, we have $F(0) \in (0,1)$, and by the continuity of F there are $x_1 \in (-\infty, 0)$ and $x_2 \in (0, \infty)$ such that $F(x_1) > 0$ and $F(x_2) < 1$. Now

$$P_F(\overline{\Omega}_{n,res}) \le P_F\Big(\bigcap_{i=1}^{n}\{\hat{e}_{ni} \ge 0\}\Big) + P_F\Big(\bigcap_{i=1}^{n}\{\hat{e}_{ni} \le 0\}\Big).$$

By using that $\hat{e}_{ni} = (\rho - \hat{\rho}_n)^T \mathbf{X}_{i-1} + e_i$ for $i = 1, \dots, n$, it is

$$\begin{aligned}
P_F\Big(\bigcap_{i=1}^{n}\{\hat{e}_{ni} \ge 0\}\Big) &\le P_F\Big(\bigcap_{i=1}^{n}\{e_i \ge x_1\}\Big) + P_F\Big(\bigcup_{i=1}^{n}\{(\hat{\rho}_n - \rho)^T \mathbf{X}_{i-1} < x_1\}\Big) \\
&\le P_F(e_1 \ge x_1)^n + P_F\Big(\bigcup_{i=1}^{n}\{\|\rho - \hat{\rho}_n\|\|\mathbf{X}_{i-1}\| > -x_1\}\Big) \\
&= \big(1 - F(x_1)\big)^n + P_F\big(\|\rho - \hat{\rho}_n\| \max_{1\le i\le n} \|\mathbf{X}_{i-1}\| > -x_1\big) \quad \underset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

since $F(x_1) > 0$ and because of the assumption $\sqrt{n}(\hat{\rho}_n - \rho) = O_{P_F}(1)$ and $\max_{1\le i\le n} \|\mathbf{X}_{i-1}\| = o_{P_F}(\sqrt{n})$, which follows directly from Lemma 5.3 by considering the singleton $M = \{F\}$ there.

Analogously, we have

$$P_F\Big(\bigcap_{i=1}^{n}\{\hat{e}_{ni} \le 0\}\Big) \le F(x_2)^n + P_F\big(\|\rho - \hat{\rho}_n\| \max_{1\le i\le n} \|\mathbf{X}_{i-1}\| > x_2\big) \underset{n\to\infty}{\longrightarrow} 0$$

using $F(x_2) < 1$. □

Because of the above, the set $\overline{\Omega}_{n,res}$ is irrelevant for asymptotic considerations such as the investigation of convergence in distribution and convergence in probability under the fixed measure P_F, and we need not specify the definition of $\hat{t}_n$ on $\overline{\Omega}_{n,res}$ for such investigations. Thus, when working with $\hat{t}_n$ we assume henceforth that $\Omega_{n,res}$ holds.

For every $n \ge 2$ define

$$\hat{p}_{ni} := \frac{1}{n(1 + \hat{t}_n \hat{e}_{ni})}, \quad 1 \le i \le n,$$

and

$$\widetilde{F}_{n,res}(x) := \sum_{i=1}^{n} \hat{p}_{ni} 1_{\{\hat{e}_{ni} \le x\}} = \sum_{i=1}^{n} \frac{1}{n(1 + \hat{t}_n \hat{e}_{ni})} 1_{\{\hat{e}_{ni} \le x\}}, \quad x \in \mathbb{R}.$$

As in subsection 3.1 it follows that $\widetilde{F}_{n,res}$ is a discrete distribution function with

$$\int_{\mathbb{R}} x \widetilde{F}_{n,res}(dx) = 0.$$

We will call $\widetilde{F}_{n,res}$ the *centered empirical distribution function of the residuals* $\hat{e}_{n1}, \dots, \hat{e}_{nn}$.

Now let F vary in M again. In order to study the asymptotic stochastic behavior of $\hat{t}_n$ and $\widetilde{F}_{n,res}$ uniformly with respect to the family of probability measures $\{P_F \colon F \in M\}$, we can again not neglect the set $\overline{\Omega}_{n,res}$ a priori, because $P_F(\overline{\Omega}_{n,res})$ will not converge to zero uniformly in $F \in M$ in general. As in the case of independent and identically distributed data, we will overcome this issue by making additional assumptions about the set M.

Let us introduce the following condition:

There are $x_1 \in (-\infty, 0)$, $x_2 \in (0, \infty)$, such that

$$\inf_{F\in M} F(x_1) > 0 \quad \text{and} \quad \sup_{F\in M} F(x_2) < 1. \tag{5.75}$$

Assume now that the set M satisfies (3.7) and (5.75), and that the sequence $(\hat{\rho}_n)_{n\in\mathbb{N}}$ is such that (5.8) holds. Then with Lemma 5.3 it is

$$\|\hat{\rho}_n - \rho\| \max_{1\le i\le n} \|\mathbf{X}_{i-1}\| = o_P^u(1) \quad \text{in } M \text{ as } n \to \infty.$$

Using this and (5.75), we see as in the proof of Lemma 5.11 that this implies

$$\sup_{F\in M} P_F(\overline{\Omega}_{n,res}) \underset{n\to\infty}{\longrightarrow} 0.$$

Hence, under these assumptions about M and $(\hat{\rho}_n)_{n\in\mathbb{N}}$ the set $\overline{\Omega}_{n,res}$ plays no role for uniform asymptotic considerations as well. We will therefore from now on always assume that these conditions hold and continue to work on the event $\Omega_{n,res}$ for every $n \geq 2$ when studying $\hat{t}_n$, $\widetilde{F}_{n,res}$ or functions thereof.

Note that the conditions (3.7) and (5.8) were used before when studying the uniform asymptotic behavior of $\mathbf{X}_i$ and the residuals. Thus, they are natural conditions for uniform asymptotic considerations and are by no means only needed for the uniform convergence of $P_F(\overline{\Omega}_{n,res})$.

The following results concerning the stochastic behavior of $\hat{t}_n$ and $\widetilde{F}_{n,res}$ have been shown for fixed $F \in M$ and autoregressive processes of order one by Genz [11]. Here, we will investigate again under which assumptions about M these results hold uniformly in $F \in M$ for the autoregressive processes of order $p \in \mathbb{N}$ described at the beginning of this section.

The next lemma is an analog of Lemma 3.3.

Lemma 5.12
If the set M is such that (3.5), (3.7) *and* (5.75) *hold and the sequence of estimators* $(\hat{\rho}_n)_{n\geq 1}$ *for ρ fulfills* (5.8), *then*

(i) $\sqrt{n}\,\hat{t}_n = O_P^u(1)$ *in M as $n \to \infty$,*

(ii) $\displaystyle\max_{1\leq i\leq n} \frac{1}{1+\hat{t}_n\hat{e}_{ni}} = O_P^u(1)$ *in M as $n \to \infty$,*

(iii) $\displaystyle\hat{t}_n = \frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^{n} e_i + o_P^u(1/\sqrt{n})$ *in M as $n \to \infty$.*

Proof. The statements are proven analogously to those of Lemma 3.3, using the results of Lemma 5.8. The proof is therefore omitted here. □

Now recall that for every $x \in \mathbb{R}$ and distribution function $F \in M$

$$\begin{aligned}
U_F(x) &= \int_{-\infty}^{x} yF(dy),\\
U_F^+(x) &= \int_{-\infty}^{x} y^+F(dy) = 1_{(0,\infty)}(x)\cdot\int_0^x yF(dy),\\
U_F^-(x) &= \int_{-\infty}^{x} y^-F(dy) = 1_{(0,\infty)}(x)\,\mathrm{E}_F(e_1^-) + 1_{(-\infty,0]}(x)\cdot\int_{-\infty}^{x}(-y)F(dy)
\end{aligned}$$

and $U_F = U_F^+ - U_F^-$. For the following investigations we assume that the set M is such that

$$\text{the family } \{U_F\colon F\in M\} \text{ is uniformly equicontinuous.} \tag{5.76}$$

Note that condition (5.76) is equivalent to the fact that both families $\{U_F^+\colon F \in M\}$ and $\{U_F^-\colon F\in M\}$ are uniformly equicontinuous.

Lemma 5.13
Suppose the set M fulfills conditions (3.7), (3.8) *and* (5.76) *and the sequence* $(\hat{\rho}_n)_{n\in\mathbb{N}}$ *satisfies* (5.8). *Then*

$$\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^{n} e_i 1_{\{\hat{e}_{ni}\leq x\}} - U_F(x)\Big| = o_P^u(1) \quad \textit{in } M \textit{ as } n\to\infty.$$

Proof. Since the family $\{U_F^+\colon F\in M\}$ is uniformly equicontinuous because of (5.76), for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$\sup_{x\in\mathbb{R}} |U_F^+(x+\delta) - U_F^+(x)| \leq \epsilon \quad \forall\ F \in M. \tag{5.77}$$

Now if $\|\hat{\rho}_n - \rho\| \cdot \max_{1\le i\le n} \|\mathbf{X}_{i-1}\| < \delta$, then

$$|\hat{e}_{ni} - e_i| = |(\hat{\rho}_n - \rho)^T \mathbf{X}_{i-1}| \le \|\hat{\rho}_n - \rho\| \cdot \max_{1\le i\le n} \|\mathbf{X}_{i-1}\| < \delta$$

for every $i = 1, \dots, n$. Thus, using (5.77) we have for every $x \in \mathbb{R}$ and $F \in M$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{\hat{e}_{ni}\le x\}} - U_F^+(x) &= \frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{e_i \le x + e_i - \hat{e}_{ni}\}} - U_F^+(x) \\ &\le \frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{e_i\le x+\delta\}} - U_F^+(x) = U_n^+(x+\delta) - U_F^+(x) \\ &\le \big|U_n^+(x+\delta) - U_F^+(x+\delta)\big| + \epsilon\end{aligned}$$

with $U_n^+(x) := \frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{e_i\le x\}}$, $x \in \mathbb{R}$, and analogously

$$\begin{aligned}U_F^+(x) - \frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{\hat{e}_{ni}\le x\}} &= U_F^+(x) - \frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{e_i\le x+e_i-\hat{e}_{ni}\}} \\ &\le U_F^+(x) - \frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{e_i\le x-\delta\}} = U_F^+(x) - U_n^+(x-\delta) \\ &\le \big|U_n^+(x-\delta) - U_F^+(x-\delta)\big| + \epsilon.\end{aligned}$$

This yields

$$\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{\hat{e}_{ni}\le x\}} - U_F^+(x)\Big| \;\le\; \sup_{x\in\mathbb{R}}\big|U_n^+(x) - U_F^+(x)\big| + \epsilon$$

for every $F \in M$ on the event $\{\|\hat{\rho}_n - \rho\| \cdot \max_{1\le i\le n}\|\mathbf{X}_{i-1}\| < \delta\}$. Because of this it is

$$\begin{aligned}&\sup_{F\in M} P_F\Big(\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{\hat{e}_{ni}\le x\}} - U_F^+(x)\Big| \ge 2\epsilon\Big) \\ \le\;& \sup_{F\in M} P_F\Big(\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{\hat{e}_{ni}\le x\}} - U_F^+(x)\Big| \ge 2\epsilon,\ \|\hat{\rho}_n - \rho\| \max_{1\le i\le n}\|\mathbf{X}_{i-1}\| < \delta\Big) \\ &+ \sup_{F\in M} P_F\big(\|\hat{\rho}_n - \rho\| \cdot \max_{1\le i\le n}\|\mathbf{X}_{i-1}\| \ge \delta\big) \\ \le\;& \sup_{F\in M} P_F\Big(\sup_{x\in\mathbb{R}}\big|U_n^+(x) - U_F^+(x)\big| \ge \epsilon\Big) + o(1),\end{aligned}$$

where the second term tends to zero as n tends to infinity because of Lemma 5.3 and (5.8). But the first summand on the right-hand side of the above inequality also tends to zero for $n \to \infty$, as was shown in the proof of Lemma 3.4. To sum up, we have shown that

$$\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^n e_i^+ 1_{\{\hat{e}_{ni}\le x\}} - U_F^+(x)\Big| = o_P^u(1) \text{ in } M \text{ as } n\to\infty.$$

Analogously we see that

$$\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^n e_i^- 1_{\{\hat{e}_{ni}\le x\}} - U_F^-(x)\Big| = o_P^u(1) \text{ in } M \text{ as } n\to\infty.$$

But this yields the statement, because

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^{n} e_i 1_{\{\hat{e}_{ni}\le x\}} - U_F(x)\Big| \\
&\le \sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^{n} e_i^+ 1_{\{\hat{e}_{ni}\le x\}} - U_F^+(x)\Big| + \sup_{x\in\mathbb{R}}\Big|\frac{1}{n}\sum_{i=1}^{n} e_i^- 1_{\{\hat{e}_{ni}\le x\}} - U_F^-(x)\Big|. \qquad \square
\end{aligned}$$

For the rest of this subsection let us assume again that

$$M \subset \Big\{F\colon F \text{ is a distribution function that has uniformly continuous Lebesgue density } f \text{ and satisfies } \int_{\mathbb{R}} x^2 F(dx) < \infty \text{ and } \int_{\mathbb{R}} xF(dx) = 0\Big\}.$$

As before, the uniformly continuous Lebesgue density f of F will also be denoted by F'.

The next lemma gives sufficient conditions for (5.76) to hold.

Lemma 5.14
If the set M satisfies conditions (3.6) *and* (5.14)*, then* (5.76) *holds, i.e., the family $\{U_F\colon F \in M\}$ is uniformly equicontinuous.*

Proof. For simplicity of notation, set $K := \sup_{F\in M}\int_{\mathbb{R}} x^2 F(dx)$ and $B := \sup_{F\in M}\|f\|_\infty$. Then $K, B \in (0,\infty)$ because of the assumptions. Now

$$|U_F(x)| \le \frac{K}{|x|} \quad \text{for every } x \in \mathbb{R}\setminus\{0\} \text{ and every } F \in M. \tag{$*$}$$

To see that this is true, let $x > 0$ first. Then

$$|U_F(x)| = \Big|\int_{\mathbb{R}} uF(du) - \int_x^\infty uF(du)\Big| = \int_x^\infty uF(du) \le \frac{1}{x}\int_{\mathbb{R}} u^2 F(du) \le \frac{K}{x}$$

for every $F \in M$. Analogously, for $x < 0$ it is

$$|U_F(x)| = -\int_{-\infty}^x uF(du) \le -\frac{1}{x}\int_{\mathbb{R}} u^2 F(du) \le -\frac{K}{x}$$

for every $F \in M$. This shows $(*)$.

Now let $\epsilon > 0$ be arbitrary, but fixed. Set $C_\epsilon := 4K/\epsilon$ and

$$\delta_\epsilon := \min\Big(\frac{\epsilon}{2BC_\epsilon}, 2C_\epsilon\Big) > 0.$$

Let $x, y \in \mathbb{R}$ with $|x-y| < \delta_\epsilon$. We have to investigate the following cases:

Case 1: $x, y \in [-C_\epsilon, C_\epsilon]$. Then

$$|U_F(x) - U_F(y)| = \Big|\int_x^y uF(du)\Big| \le C_\epsilon \cdot \|f\|_\infty \cdot |y-x| \le C_\epsilon \cdot B \cdot \delta_\epsilon \le \frac{\epsilon}{2} < \epsilon$$

for every $F \in M$.

Case 2: $x > C_\epsilon$. Then either $y \le C_\epsilon$, or $y > C_\epsilon$, too.
If $y \le C_\epsilon$, then $|y-x| < \delta_\epsilon \le 2C_\epsilon$ implies that $y > -C_\epsilon$. Thus, using $(*)$ and Case 1 we see that

$$|U_F(x) - U_F(y)| \le |U_F(x)| + |U_F(C_\epsilon)| + |U_F(C_\epsilon) - U_F(y)| \le \frac{K}{x} + \frac{K}{C_\epsilon} + \frac{\epsilon}{2} < \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{2} = \epsilon$$

for every $F \in M$.
Now suppose $y > C_\epsilon$ as well. Then for every $F \in M$

$$|U_F(x) - U_F(y)| \leq |U_F(x)| + |U_F(y)| \leq \frac{K}{x} + \frac{K}{y} \leq \frac{\epsilon}{4} + \frac{\epsilon}{4} < \epsilon$$

using $(*)$.

Case 3: $x < -C_\epsilon$. This case follows analogously to Case 2.

To sum up, we have shown that

$$x, y \in \mathbb{R},\ |y - x| < \delta_\epsilon \quad \Longrightarrow \quad |U_F(x) - U_F(y)| < \epsilon\ \forall\ F \in M,$$

which yields the statement. □

We will now combine the previous results to show the following expansion of the difference $\widetilde{F}_{n,res} - F_n$. Compare this to Proposition 3.5.

Proposition 5.15
Suppose the set M is such that conditions (3.7), (3.8) *as well as* (5.12), (5.13), (5.14) *and* (5.75) *hold. Also, let the sequence $(\hat{\rho}_n)_{n\in\mathbb{N}}$ of estimators of ρ be such that* (5.8) *is satisfied. Then*

$$\widetilde{F}_{n,res}(x) - F_n(x) = -U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i + R_{n,F}(x), \quad x \in \mathbb{R},$$

with $\|R_{n,F}\|_\infty = o_P^u(1/\sqrt{n})$ in M as $n \to \infty$.

Proof. Employing once more that $1/(1+y) = 1 - y + y^2/(1+y)$ for every $y \in \mathbb{R} \setminus \{-1\}$, we can expand the function $\widetilde{F}_{n,res}$ for every $x \in \mathbb{R}$ and $n \geq 2$ in the following way

$$\widetilde{F}_{n,res}(x) = F_{n,res}(x) - \hat{t}_n\frac{1}{n}\sum_{i=1}^n \hat{e}_{ni}1_{\{\hat{e}_{ni}\leq x\}} + \hat{t}_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{\hat{e}_{ni}^2}{1 + \hat{t}_n\hat{e}_{ni}} \cdot 1_{\{\hat{e}_{ni}\leq x\}}.$$

Using this, we have

$$\begin{aligned}
\widetilde{F}_{n,res}(x) - F_n(x) &= F_{n,res}(x) - F_n(x) - U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i + U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i \\
&\quad - \hat{t}_n\frac{1}{n}\sum_{i=1}^n \hat{e}_{ni}1_{\{\hat{e}_{ni}\leq x\}} + \hat{t}_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{\hat{e}_{ni}^2}{1 + \hat{t}_n\hat{e}_{ni}} \cdot 1_{\{\hat{e}_{ni}\leq x\}} \\
&= F_{n,res}(x) - F_n(x) - U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i + U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i \\
&\quad - \hat{t}_n\frac{1}{n}\sum_{i=1}^n e_i1_{\{\hat{e}_{ni}\leq x\}} + \hat{t}_n\frac{1}{n}\sum_{i=1}^n \big(e_i - \hat{e}_{ni}\big)1_{\{\hat{e}_{ni}\leq x\}} \\
&\quad + \hat{t}_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{\hat{e}_{ni}^2}{1 + \hat{t}_n\hat{e}_{ni}} \cdot 1_{\{\hat{e}_{ni}\leq x\}} \\
&= F_{n,res}(x) - F_n(x) - U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i + U_F(x)\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i \\
&\quad - \hat{t}_n\Big(\frac{1}{n}\sum_{i=1}^n e_i1_{\{\hat{e}_{ni}\leq x\}} - U_F(x)\Big) - \hat{t}_nU_F(x) + \hat{t}_n\frac{1}{n}\sum_{i=1}^n \big(e_i - \hat{e}_{ni}\big)1_{\{\hat{e}_{ni}\leq x\}}
\end{aligned}$$

$$
\begin{aligned}
&+ \hat{t}_n^2 \cdot \frac{1}{n} \sum_{i=1}^n \frac{\hat{e}_{ni}^2}{1+\hat{t}_n \hat{e}_{ni}} \cdot 1_{\{\hat{e}_{ni} \le x\}} \\
&= -U_F(x) \frac{1}{\sigma_F^2} \frac{1}{n} \sum_{i=1}^n e_i + R_{1n}(x) + R_{2n,F}(x) + R_{3n,F}(x) + R_{4n}(x) + R_{5n}(x),
\end{aligned}
$$

where

$$
R_{1n}(x) := F_{n,res}(x) - F_n(x), \qquad R_{2n,F}(x) := U_F(x)\Big(\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i - \hat{t}_n\Big),
$$

$$
R_{3n,F}(x) := \hat{t}_n\Big(U_F(x) - \frac{1}{n}\sum_{i=1}^n e_i 1_{\{\hat{e}_{ni}\le x\}}\Big), \qquad R_{4n}(x) := \hat{t}_n \frac{1}{n}\sum_{i=1}^n (e_i - \hat{e}_{ni}) 1_{\{\hat{e}_{ni}\le x\}},
$$

$$
R_{5n}(x) := \hat{t}_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{\hat{e}_{ni}^2}{1+\hat{t}_n\hat{e}_{ni}} \cdot 1_{\{\hat{e}_{ni}\le x\}}.
$$

Since all assumptions of Theorem 5.9 are satisfied, it follows from this that

$$
\|R_{1n}\|_\infty = \sup_{x\in\mathbb{R}} |F_{n,res}(x) - F_n(x)| = o_P^u(1/\sqrt{n}).
$$

Moreover, we have seen before that $\sup_{x\in\mathbb{R}} |U_F(x)| = |U_F(0)| = \mathrm{E}_F(e_1^+) = \mathrm{E}_F(e_1^-) = \frac{1}{2}\,\mathrm{E}_F(|e_1|)$. Thus, we have

$$
\begin{aligned}
\|R_{2n,F}\|_\infty &= \sup_{x\in\mathbb{R}} |U_F(x)| \cdot \Big|\frac{1}{\sigma_F^2}\frac{1}{n}\sum_{i=1}^n e_i - \hat{t}_n\Big| \\
&\le \frac{1}{2} \sup_{F\in M} \int_{\mathbb{R}} |x| F(dx) \cdot o_P^u(1/\sqrt{n}) = o_P^u(1/\sqrt{n}) \ \text{ in } M \text{ as } n\to\infty
\end{aligned}
$$

using Lemma 5.12 (iii).

It also is

$$
\|R_{3n,F}\|_\infty = |\hat{t}_n| \cdot \sup_{x\in\mathbb{R}} \Big| U_F(x) - \frac{1}{n}\sum_{i=1}^n e_i 1_{\{\hat{e}_{ni}\le x\}}\Big|,
$$

and the desired result follows from Lemma 5.12 (i) and Lemma 5.13. Note here that Lemma 5.14 ensures that condition (5.76) of Lemma 5.13 indeed holds.

Next, we have

$$
\|R_{4n}\|_\infty \le |\hat{t}_n| \frac{1}{n}\sum_{i=1}^n |e_i - \hat{e}_{ni}| \le |\hat{t}_n| \max_{1\le i\le n} |e_i - \hat{e}_{ni}| = o_P^u(1/\sqrt{n})
$$

by Lemma 5.8 (ii) and Lemma 5.12 (i).

Last, it is

$$
\begin{aligned}
\|R_{5n}\|_\infty &\le \hat{t}_n^2 \cdot \frac{1}{n}\sum_{i=1}^n \frac{\hat{e}_{ni}^2}{1+\hat{t}_n\hat{e}_{ni}} \\
&\le \hat{t}_n^2 \cdot \max_{1\le i\le n} \frac{1}{1+\hat{t}_n\hat{e}_{ni}} \cdot \frac{1}{n}\sum_{i=1}^n \hat{e}_{ni}^2 = O_P^u(1/n) O_P^u(1) O_P^u(1) = o_P^u(1/\sqrt{n})
\end{aligned}
$$

with Lemma 5.8 (v) and Lemma 5.12 (i), (ii). □

6 The limiting Pitman ARE of the two tests in stable autoregressive models

We will now consider goodness-of-fit testing for the error distribution F in certain stable autoregressive models using the classical Cramér-von Mises statistics based on the standard empirical distribution function of the residuals as well as their modified counterparts based on the centered empirical distribution function of the residuals. As in the case of independent and identically distributed centered data investigated in section 4, we will consider testing the simple null hypothesis that F is equal to some centered continuous distribution function as well as the composite null hypothesis that F belongs to the scale family $\mathcal{F}_\tau$ against suitable nonparametric alternatives, respectively. Similar to before, we will compare the performance of the asymptotic test corresponding to the classical residual Cramér-von Mises statistic to that of the asymptotic test corresponding to the modified statistic for each of these testing problems by determining their limiting Pitman asymptotic relative efficiency.

First, we will describe the setting of this section. For every $q \in [2, \infty)$ set

$$\mathcal{G}_q^u := \Big\{F\colon F \text{ is a distribution function that has uniformly continuous Lebesgue density } f \\ \text{and satisfies } \int_{\mathbb{R}} |x|^q F(dx) < \infty \text{ and } \int_{\mathbb{R}} xF(dx) = 0\Big\}.$$

Then $\mathcal{G}_q^u \subset \mathcal{G}_q$, where the latter set is defined in section 4. In this section we consider a measurable space $(\Omega, \mathcal{A})$ and a family $\{P_F\colon F \in \mathcal{G}_q^u\}$ of probability measures on $\mathcal{A}$ such that the following requirements are satisfied:
On $(\Omega, \mathcal{A})$ there are random variables $S_0, \ldots, S_{1-p}$, $p \in \mathbb{N}$ fixed, and e_i, $i \in \mathbb{Z}$, such that

- for each $F \in \mathcal{G}_q^u$ the variables $(e_i)_{i\in\mathbb{Z}}$ are independent and identically distributed with distribution function F under P_F,
- the random vector $\mathbf{S}_0 := (S_0, \ldots, S_{1-p})^T$ has the same fixed distribution Q, say, under every P_F, $F \in \mathcal{G}_q^u$, and $\int_{\mathbb{R}^p} \|x\|^2 Q(dx) < \infty$,
- under every P_F, $F \in \mathcal{G}_q^u$, the vector $\mathbf{S}_0$ is independent of $(e_i)_{i\in\mathbb{N}}$.

As in section 5, we assume that $(X_i)_{i\geq 1-p}$ is either one of the $AR(p)$ processes of model 1 or model 2 on $(\Omega, \mathcal{A})$. Then $(X_i)_{i\geq 1-p}$ satisfies the model equation

$$X_i = \rho_1 X_{i-1} + \ldots + \rho_p X_{i-p} + e_i, \quad i \geq 1, \tag{6.1}$$

where $\rho_1, \ldots, \rho_p$ are real constants with $\rho_p \neq 0$ and

$$\{z \in \mathbb{C}\colon z^p - \rho_1 z^{p-1} - \rho_2 z^{p-2} - \ldots - \rho_{p-1} z - \rho_p = 0\} \ \subset \ \{z \in \mathbb{C}\colon |z| < 1\}.$$

The autoregressive parameter $\rho = (\rho_1, \ldots, \rho_p)^T \in \mathbb{R}^p$ is assumed to be unknown and therefore has to be estimated by a suitable sequence of estimators $(\hat{\rho}_n)_{n\in\mathbb{N}}$. Moreover, the error variables e_i, $i \in \mathbb{N}$, are not observable, whereas the process $(X_i)_{i\geq 1-p}$ is assumed to be so.

6.1 Simple null hypothesis

Suppose now that the error variables $e_1, e_2, \ldots$ have common distribution function $F \in \mathcal{G}_q^u$, $q \geq 2$ fixed, but that F is unknown. For this reason we consider the testing problem

$$\mathrm{H}_0\colon F = F_0 \ \text{ versus } \ \mathrm{H}_1\colon F \in \mathcal{G}_q^u \setminus \{F_0\} \tag{6.2}$$

for a fixed $F_0 \in \mathcal{G}_q^u$. Since unfortunately the errors are not observable, we cannot compute a test statistic for this testing problem based on a "sample" $e_1, \dots, e_n$. However, we can observe data $\mathbf{X}_0, X_1, \dots, X_n$, $n \geq 2$, and use this sample to study the residuals

$$\hat{e}_{ni} = X_i - \hat{\rho}_n^T \mathbf{X}_{i-1}, \quad 1 \leq i \leq n,$$

where $(\hat{\rho}_n)_{n \geq 2}$ is a suitable sequence of estimators for the unknown autoregressive parameter ρ. In the following, we will only consider estimators $\hat{\rho}_n$ with $\sqrt{n}(\hat{\rho}_n - \rho) = O_{P_F}(1)$, which holds for example for the least squares estimator $\hat{\rho}_n^{LS}$.

For testing H_0 versus H_1, we consider the residual Cramér-von Mises statistic

$$\mathcal{W}_{n,res}^2 = n \cdot \int_{-\infty}^{\infty} \big(F_{n,res}(x) - F_0(x)\big)^2 F_0(dx)$$

based on $\hat{e}_{n1}, \dots, \hat{e}_{nn}$ and its modified version

$$\mathcal{V}_{n,res}^2 = n \cdot \int_{-\infty}^{\infty} \big(\widetilde{F}_{n,res}(x) - F_0(x)\big)^2 F_0(dx),$$

which are both measurable mappings from Ω to $[0, \infty)$. The latter statistic is reasonable here since the true error distribution function F is centered, and so is $\widetilde{F}_{n,res}$.

It follows from Koul's results, see for example chapter 7 in Koul [18], that

$$\|\sqrt{n}(F_{n,res} - F_n)\|_\infty \underset{n \to \infty}{\longrightarrow} 0 \;\text{ in } P_F\text{-probability} \tag{6.3}$$

for every fixed $F \in \mathcal{G}_q^u$, where F_n is the empirical distribution function of the errors. This can also be deduced directly from Theorem 5.9, which is in fact just the uniform version of (6.3), by considering the singleton $M = \{F\}$ there. Note that for every $F \in \mathcal{G}_q^u$ the set $M = \{F\}$ satisfies (5.12)–(5.14) and assumptions (ii)–(v) of Theorem 5.9, since (5.12)–(5.14) are trivially fulfilled by the uniformly continuous Lebesgue density f of F, and the latter conditions follow from assumption (3.7), which holds for $M = \{F\}$ because F has finite second moment.

Thus, by Slutzky's theorem we have

$$\sqrt{n}\big(F_{n,res} - F\big) \;\; \underset{n}{\xrightarrow{\mathcal{L}}} \;\; B^\circ(F) \quad \text{in} \;\; D[-\infty, \infty] \tag{6.4}$$

under the measure P_F. Now a direct application of the continuous mapping theorem yields

$$\mathcal{W}_{n,res} \;\; \underset{n}{\xrightarrow{\mathcal{L}}} \;\; \Big(\int_0^1 B^\circ(t)^2 dt\Big)^{1/2} = \mathcal{W} \tag{6.5}$$

under H_0 for $\mathcal{W}_{n,res} = \big(\mathcal{W}_{n,res}^2\big)^{1/2}$, so that a test of asymptotic level $\alpha \in (0,1)$ based on $\mathcal{W}_{n,res}$ for the testing problem (6.2) can be constructed analogously to the case of independent and identically distributed data described in subsection 4.1, where the test statistic $\mathcal{W}_n$ was used.

Now let us investigate $\mathcal{V}_{n,res} = \big(\mathcal{V}_{n,res}^2\big)^{1/2}$. If we consider the singleton $M = \{F\}$ for some $F \in \mathcal{G}_q^u$, then all assumptions of Proposition 5.15 are satisfied and it follows that

$$\widetilde{F}_{n,res}(x) - F(x) = \frac{1}{n} \sum_{i=1}^n Y_i(x) + R_{n,F}(x), \quad x \in \mathbb{R}, \tag{6.6}$$

with

$$Y_i(x) := 1_{\{e_i \leq x\}} - F(x) - \frac{U_F(x)}{\sigma_F^2} e_i, \quad i \geq 1,$$

and $\|R_{n,F}\|_\infty = o_{P_F}(1/\sqrt{n})$. Hence, by Slutzky's theorem the processes

$$\sqrt{n}(\widetilde{F}_{n,res} - F) \quad \text{and} \quad \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$$

have the same limit distribution in $D[-\infty, \infty]$ under the measure P_F. Observe that the process $\big(n^{-1/2} \sum_{i=1}^{n} Y_i(x)\big)_{x \in \mathbb{R}}$ does only depend on the sequence $(e_i)_{i \in \mathbb{N}}$ of independent and identically distributed random variables, and not on the autoregressive process $(X_i)_{i \geq 1-p}$. The asymptotic distribution of $\big(n^{-1/2} \sum_{i=1}^{n} Y_i(x)\big)_{x \in \mathbb{R}}$ has been derived by Zhang [31] in a more general setting than considered here, see the proof of Theorem 3.3 in [31], and Genz [11] has established the above expansion of $\widetilde{F}_{n,res} - F$ for arbitrary $AR(1)$ processes, see Lemma 3.7 in [11]. It follows now with Zhang's result that under P_F

$$\sqrt{n}\big(\widetilde{F}_{n,res} - F\big) \ \underset{n}{\overset{\mathcal{L}}{\longrightarrow}} \ W \quad \text{in} \ D[-\infty, \infty], \tag{6.7}$$

where W is just the Gaussian process appearing in (4.3). By applying the continuous mapping theorem once more, this yields

$$\mathcal{V}_{n,res} \quad \underset{n}{\overset{\mathcal{L}}{\longrightarrow}} \quad \Big(\int_0^1 W\big(F_0^{-1}(t)\big)^2 dt\Big)^{1/2} = \mathcal{V} \tag{6.8}$$

under H_0, where the random variable $\mathcal{V}$ is the same as the limit in (4.5), given that F_0 is the same in both cases, of course. Hence, a test of asymptotic level $\alpha \in (0,1)$ for testing (6.2) with the test statistic $\mathcal{V}_{n,res}$ can be constructed analogously to the case of independent and identically distributed data described in subsection 4.1.

We proceed by showing that $(\mathcal{W}_{n,res})_{n \geq 2}$ and $(\mathcal{V}_{n,res})_{n \geq 2}$ are standard sequences.

Since the asymptotic distributions of the two sequences under the null hypothesis are the same as in the case of independent and identically distributed data discussed in subsection 4.1, conditions (BI) and (BII) of Definition 2.1 have already been verified there. We only need to show that (BIII) holds for both sequences. To see this, let $F \in \mathcal{G}_q^u \setminus \{F_0\}$. Then by Minkowski's inequality

$$\Big|\frac{\mathcal{W}_{n,res}}{\sqrt{n}} - b(F)\Big| \ \leq \ \Big(\int_{-\infty}^{\infty} \big(F_{n,res}(x) - F(x)\big)^2 F_0(dx)\Big)^{1/2} \ \leq \ \|F_{n,res} - F\|_\infty$$

and

$$\Big|\frac{\mathcal{V}_{n,res}}{\sqrt{n}} - b(F)\Big| \ \leq \ \Big(\int_{-\infty}^{\infty} \big(\widetilde{F}_{n,res}(x) - F(x)\big)^2 F_0(dx)\Big)^{1/2} \ \leq \ \|\widetilde{F}_{n,res} - F\|_\infty$$

for every $n \geq 2$, where

$$b \colon \mathcal{G}_q^u \setminus \{F_0\} \ni F \mapsto \Big(\int_{-\infty}^{\infty} \big(F(x) - F_0(x)\big)^2 F_0(dx)\Big)^{1/2} \in (0,1].$$

But it follows from (6.4) that

$$\|F_{n,res} - F\|_\infty \underset{n \to \infty}{\longrightarrow} 0 \ \text{ in } P_F\text{-probability},$$

and analogously

$$\|\widetilde{F}_{n,res} - F\|_\infty \underset{n \to \infty}{\longrightarrow} 0 \ \text{ in } P_F\text{-probability}$$

because of (6.7). Thus, (BIII) is satisfied.

Hence, the following result holds.

Proposition 6.1
The sequences $(\mathcal{W}_{n,res})_{n\geq 2}$ and $(\mathcal{V}_{n,res})_{n\geq 2}$ are standard sequences with approximate slopes $b(F)^2/\lambda_1$ and $b(F)^2/\lambda_1^*$, respectively, for every $F \in \mathcal{G}_q^u \setminus \{F_0\}$, where λ_1 and λ_1^* are as in Proposition 4.4. Hence, the approximate Bahadur ARE of $(\mathcal{W}_{n,res})_{n\geq 2}$ relative to $(\mathcal{V}_{n,res})_{n\geq 2}$ is λ_1^*/λ_1.

Note that the approximate Bahadur ARE of $(\mathcal{W}_{n,res})_{n\geq 2}$ relative to $(\mathcal{V}_{n,res})_{n\geq 2}$ is equal to the one of $(\mathcal{W}_n)_{n\geq 2}$ relative to $(\mathcal{V}_n)_{n\geq 2}$ when testing the same null hypothesis, see Proposition 4.4. In particular, it is $\lambda_1^* < \lambda_1$, as noted there.

Our next aim is to show that $(\mathcal{W}_{n,res})_{n\geq 2}$ and $(\mathcal{V}_{n,res})_{n\geq 2}$ also satisfy Wieand's condition (WIII), so that we can use Theorem 2.3 to determine the limiting Pitman ARE of these two sequences. For this we need to reduce the set of possible distribution functions F, since we have to require the alternatives to be sufficiently smooth. Let us introduce some notation first.

For any $0 < \gamma \leq 1$ and function $h\colon \mathbb{R} \to \mathbb{R}$ the *γ-Hölder coefficient* of h is defined as

$$[h]_\gamma := \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{|h(x)-h(y)|}{|x-y|^\gamma} \in [0,\infty].$$

The function h is said to be *Hölder continuous* with *exponent* γ (γ-Hölder continuous, for short) if $[h]_\gamma < \infty$. In particular, if $[h]_1 < \infty$, then h is *Lipschitz continuous*. Obviously, every Hölder continuous function is uniformly continuous. Note that if the function h is bounded and γ-Hölder continuous for some $\gamma \in (0,1]$, then $[h]_\kappa < \infty$ for all $0 < \kappa < \gamma$, since

$$[h]_\kappa = \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \Big[|h(x)-h(y)|^{1-\kappa/\gamma} \cdot \Big(\frac{|h(x)-h(y)|}{|x-y|^\gamma}\Big)^{\kappa/\gamma}\Big] \leq 2\cdot \|h\|_\infty^{1-\kappa/\gamma} \cdot [h]_\gamma^{\kappa/\gamma} < \infty.$$

Hence, if $f\colon \mathbb{R}\to\mathbb{R}$ is a density with $[f]_\gamma < \infty$ for some $\gamma \in (0,1]$, then $[f]_\kappa < \infty$ for all $0 < \kappa < \gamma$, as f is uniformly continuous because of $[f]_\gamma < \infty$ and therefore bounded. For example, if the density f is differentiable everywhere and has a bounded first derivative, then it follows by the mean value theorem that $[f]_1 < \infty$, so that $[f]_\gamma < \infty$ for all $\gamma \in (0,1]$ in this case.

Let $w\colon \mathbb{R} \to [0,\infty)$ be a function with $w(x) \to \infty$ as $|x| \to \infty$. We will call such a function w a *weight function*, and we set

$$\|h\|_{w,\infty} := \|wh\|_\infty = \sup_{x\in\mathbb{R}} |w(x)h(x)| \in [0,\infty]$$

for any function $h\colon \mathbb{R}\to\mathbb{R}$.

Now for every $q \in (2,\infty)$, $\gamma \in (0,1]$ and weight function w let

$$\mathcal{G}_{q,\gamma,w} := \Big\{F\colon F \text{ is a distribution function having a Lebesgue density } f \text{ that satisfies } [f]_\gamma + \|f\|_{w,\infty} + \int_{\mathbb{R}} |x|^q f(x)dx < \infty \text{ and } \int_{\mathbb{R}} x f(x)dx = 0\Big\}. \tag{6.9}$$

Note that such a Lebesgue density f of $F \in \mathcal{G}_{q,\gamma,w}$ is uniquely determined, as it is (uniformly) continuous because of $[f]_\gamma < \infty$. Henceforth we will always refer without further mention to this uniquely determined uniformly continuous density when considering a density of a distribution function in $\mathcal{G}_{q,\gamma,w}$. Observe furthermore that by what was said above, it is $\mathcal{G}_{q,\gamma,w} \subset \mathcal{G}_{q,\kappa,w}$ for every $\kappa \in (0,\gamma)$.

Throughout this subsection we will always assume that w is bounded away from zero, so that $\|1/w\|_\infty < \infty$. Possible choices of such weight functions are $w(x) = \exp(a|x|^s)$ for some $a, s > 0$ or $w(x) = |x|^s 1_{[1,\infty)}(|x|) + 1_{[0,1)}(|x|)$ for some $s > 0$, among others. Moreover, we suppose of course that $\mathcal{G}_{q,\gamma,w}$ contains more than one element. For the weight functions just mentioned, this is obviously the case for every value of $\gamma \in (0,1]$ and $q > 2$, as there are numerous examples of centered distributions with finite absolute q-th moment that possess a Lebesgue density f that is differentiable with bounded derivative ($\Rightarrow [f]_\gamma < \infty$ for every $\gamma \in (0,1]$) and satisfies $w(x)f(x) \to 0$ as $|x| \to \infty$, which implies $\|f\|_{w,\infty} < \infty$ because of the continuity of wf.

We will now equip $\mathcal{G}_{q,\gamma,w}$ with a suitable metric. It is easy to see that the function $d_{q,\gamma,w} \colon \mathcal{G}_{q,\gamma,w} \times \mathcal{G}_{q,\gamma,w} \to [0,\infty)$,

$$d_{q,\gamma,w}(F,G) := [f-g]_\gamma + \|f-g\|_{w,\infty} + \Big| \int_{\mathbb{R}} |x|^q F(dx) - \int_{\mathbb{R}} |x|^q G(dx) \Big|,$$

is a metric on $\mathcal{G}_{q,\gamma,w}$, where f and g denote the densities of F and G, respectively. Observe here that the fact that $d_{q,\gamma,w}(F,G) = 0$ implies $F = G$ follows for one because $\|f-g\|_{w,\infty} = 0 \Rightarrow f = g$ since w is strictly positive by assumption, but also because $[f-g]_\gamma = 0 \Rightarrow f = g$. To verify this, note that it follows from $[f-g]_\gamma = 0$ that $f = g + c$ for some $c \in \mathbb{R}$, but as f and g are densities, this is only true for $c = 0$. Hence, $(\mathcal{G}_{q,\gamma,w}, d_{q,\gamma,w})$ is a metric space.

For the rest of this subsection, we will only consider distribution functions $F \in \mathcal{G}_{q,\gamma,w}$ for some fixed $q > 2$, $\gamma \in (0,1]$ and weight function w, and we will measure the distance of two such distribution functions with the metric $d_{q,\gamma,w}$, unless stated otherwise. Since $\mathcal{G}_{q,\gamma,w} \subset \mathcal{G}_q^u$, all previously derived results of course still hold for $F \in \mathcal{G}_{q,\gamma,w}$.

Lemma 6.2
The identity function

$$id \colon (\mathcal{G}_{q,\gamma,w}, d_{q,\gamma,w}) \ \ni \ F \ \mapsto \ F \ \in \ (\mathcal{G}_{q,\gamma,w}, d_q)$$

is continuous, where d_q is the Kantorovich-Wasserstein metric defined in (4.1).

Proof. Let $(F_n^*)_{n\in\mathbb{N}}$, $F \in \mathcal{G}_{q,\gamma,w}$ with $d_{q,\gamma,w}(F_n^*, F) \to 0$ as $n \to \infty$, and let f_n^* and f denote the densities of F_n^* and F, respectively. Then

$$\|f_n^* - f\|_\infty \le \|1/w\|_\infty \cdot \|f_n^* - f\|_{w,\infty} \underset{n\to\infty}{\longrightarrow} 0,$$

using that $1/w$ is bounded. By Scheffé's theorem it follows from this that $F_n^* \xrightarrow{\mathcal{L}} F$. Moreover, $d_{q,\gamma,w}(F_n^*, F) \to 0$ implies $\int |x|^q F_n^*(dx) \to \int |x|^q F(dx)$ for $n \to \infty$. But convergence in distribution of F_n^* to F combined with convergence of the absolute q-th moments is equivalent to $d_q(F_n^*, F) \to 0$ as $n \to \infty$. ☐

As the composition of continuous functions is continuous again, the next result follows immediately from the foregoing lemma and Lemma 4.1.

Corollary 6.3

(i) *For every $r \in [1,q]$ the function*

$$(\mathcal{G}_{q,\gamma,w}, d_{q,\gamma,w}) \ \ni \ F \ \mapsto \ \int_{\mathbb{R}} |x|^r F(dx) \ \in (\mathbb{R}, |\cdot|)$$

is continuous.

(ii) *The identity function*

$$id \colon (\mathcal{G}_{q,\gamma,w}, d_{q,\gamma,w}) \ \ni \ F \ \mapsto \ F \ \in (\mathcal{G}_{q,\gamma,w}, d_K)$$

is continuous, where d_K is the Kolmogorov metric.

Since we assume now that $e_1, e_2, \dots$ are independent and identically distributed with distribution function $F \in \mathcal{G}_{q,\gamma,w}$, we have to adjust the testing problem (6.2) accordingly by letting all null and non-null distribution functions be in $\mathcal{G}_{q,\gamma,w}$, i.e., we consider henceforth the testing problem

$$\mathrm{H}_0\colon F = F_0 \quad \text{versus} \quad \mathrm{H}_1\colon F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\} \tag{6.10}$$

for some $F_0 \in \mathcal{G}_{q,\gamma,w}$.

Remark 6.4: Note that for this testing problem condition (2.2) holds. To verify this, let $F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}$ and set $F_t := tF + (1-t)F_0$ for every $t \in (0,1)$. It is easy to see that $F_t \in \mathcal{G}_q \setminus \{F_0\}$. Moreover, F_t has Lebesgue density $f_t := tf + (1-t)f_0$, where f and f_0 are the respective densities of F and F_0. It is

$$[f_t]_\gamma \;\le\; t\,[f]_\gamma + (1-t)\,[f_0]_\gamma < \infty$$

and

$$\|f_t\|_{w,\infty} = \|wf_t\|_\infty \;\le\; t\,\|wf\|_\infty + (1-t)\,\|wf_0\|_\infty = t\,\|f\|_{w,\infty} + (1-t)\,\|f_0\|_{w,\infty} < \infty,$$

so that $F_t \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}$ for all $t \in (0,1)$. Now note that

$$\begin{aligned} d_{q,\gamma,w}(F_t, F_0) &= [f_t - f_0]_\gamma + \|f_t - f_0\|_{w,\infty} + \Big| \int_{\mathbb{R}} |x|^q F_t(dx) - \int_{\mathbb{R}} |x|^q F_0(dx) \Big| \\ &= t \cdot \Big([f - f_0]_\gamma + \|f - f_0\|_{w,\infty} + \Big| \int_{\mathbb{R}} |x|^q F(dx) - \int_{\mathbb{R}} |x|^q F_0(dx) \Big| \Big) \\ &= t \cdot d_{q,\gamma,w}(F, F_0) \underset{t \to 0}{\longrightarrow} 0, \end{aligned}$$

whence it follows that for every $\epsilon > 0$ there is a $t(\epsilon) \in (0,1)$ with

$$F_t \in U_\epsilon(F_0) \cap (\mathcal{G}_{q,\gamma,w} \setminus \{F_0\}) \quad \forall\, 0 < t \le t(\epsilon),$$

which shows the claim. ♦

For the following investigations, we have to specify the sequence of estimators for the unknown autoregressive parameter ρ in (6.1). Henceforth, we will estimate ρ by the least squares estimator

$$\hat{\rho}_n^{LS} = \Big(\sum_{i=1}^n \mathbf{X}_{i-1} \mathbf{X}_{i-1}^T \Big)^+ \sum_{i=1}^n \mathbf{X}_{i-1} X_i, \quad n \ge 2.$$

Recall that B^+ is the Moore-Penrose pseudoinverse of the matrix B.

We are now able to verify Wieand's condition for our two sequences of test statistics.

Theorem 6.5
The sequences $(\mathcal{W}_{n,res})_{n\ge2}$ and $(\mathcal{V}_{n,res})_{n\ge2}$ fulfill Wieand's condition (WIII) *with*

$$b\colon \mathcal{G}_{q,\gamma,w} \setminus \{F_0\} \ni F \mapsto \Big(\int_{-\infty}^{\infty} \big(F(x) - F_0(x)\big)^2 F_0(dx) \Big)^{1/2} \in (0,1].$$

Proof. We start by verifying the statement for $(\mathcal{W}_{n,res})_{n\ge2}$.

Set $K := (\int_{\mathbb{R}} x^2 F_0(dx))/2 > 0$. Then by part (i) of Corollary 6.3 there is a $\varrho > 0$ such that

$$\Big| \int_{\mathbb{R}} x^2 F(dx) - \int_{\mathbb{R}} x^2 F_0(dx) \Big| < K \quad \text{for all } F \in \mathcal{G}_{q,\gamma,w} \text{ with } d_{q,\gamma,w}(F, F_0) < \varrho. \tag{6.11}$$

Now define

$$M := U_\varrho(F_0) \setminus \{F_0\} = \big\{F \in \mathcal{G}_{q,\gamma,w} \colon d_{q,\gamma,w}(F, F_0) < \varrho\big\} \setminus \{F_0\}. \tag{6.12}$$

We will show next that this set satisfies conditions (3.5), (3.7) and (5.12)–(5.14).

By (6.11) it is $\int_{\mathbb{R}} x^2 F(dx) > K > 0$ for every $F \in M$, which implies (3.5).

To check condition (3.7), note that

$$\Big|\int_{\mathbb{R}} |x|^q F(dx) - \int_{\mathbb{R}} |x|^q F_0(dx)\Big| < \varrho$$

for every $F \in M$ according to the definition of $d_{q,\gamma,w}$. Using this and the fact that $q > 2$, we see as before that for every $F \in M$ and $c \in (0, \infty)$ it is

$$\int_{\{|x|>c\}} x^2 F(dx) \le c^{2-q} \int_{\mathbb{R}} |x|^q F(dx) < c^{2-q}\big(\varrho + \int_{\mathbb{R}} |x|^q F_0(dx)\big) < \infty,$$

whence (3.7) is evident.

It follows moreover from the definition of $d_{q,\gamma,w}$ that $[f - f_0]_\gamma < \varrho$ and $\|f - f_0\|_{w,\infty} < \varrho$ for every $F \in M$, where f and f_0 are the densities of F and F_0, respectively. Hence,

$$[f]_\gamma \le [f - f_0]_\gamma + [f_0]_\gamma < \varrho + [f_0]_\gamma < \infty$$

for every $F \in M$, so that the set $\{[f]_\gamma \colon F \in M\}$ of γ-Hölder coefficients is bounded. But this just implies that the family $\{f \colon F \in M\}$ is uniformly equicontinuous, so that condition (5.12) holds.

Now keep in mind that the uniform continuity of f_0 implies that it tends to zero as $|x| \to \infty$, whence $\|f_0\|_\infty < \infty$ follows.

For every $F \in M$ and $x \in \mathbb{R}$ it is

$$f(x) \le |f(x) - f_0(x)| + f_0(x) \le \frac{1}{w(x)} \cdot \|f - f_0\|_{w,\infty} + f_0(x) < \frac{\varrho}{w(x)} + f_0(x),$$

so that

$$\sup_{F \in M} f(x) \le \frac{\varrho}{w(x)} + f_0(x) \quad \underset{|x|\to\infty}{\longrightarrow} \quad 0,$$

which is just condition (5.13).

Furthermore, because of $\|f - f_0\|_\infty \le \|1/w\|_\infty \cdot \|f - f_0\|_{w,\infty} < \varrho \cdot \|1/w\|_\infty < \infty$ we have

$$\|f\|_\infty \le \|f - f_0\|_\infty + \|f_0\|_\infty < \varrho \, \|1/w\|_\infty + \|f_0\|_\infty < \infty$$

for every $F \in M$. Hence, (5.14) also holds.

As we have already shown, it follows with Minkowski's inequality that

$$\Big|\frac{\mathcal{W}_{n,res}}{\sqrt{n}} - b(F)\Big| \;\le\; \|F_{n,res} - F\|_\infty \;\le\; \|F_{n,res} - F_n\|_\infty + \|F_n - F\|_\infty$$

for every $F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}$ and $n \ge 2$.

Now the set M is such that by Proposition 5.7 the least squares estimator $\hat{\rho}_n^{LS}$ satisfies $\sqrt{n}(\hat{\rho}_n^{LS} - \rho) = O_P^u(1)$ in M as $n \to \infty$. Moreover, as M satisfies (3.7), it follows that the assumptions (ii)–(v) of Theorem 5.9 hold as well. Hence, all assumptions of this theorem are satisfied for M as in (6.12), so that we have

$$\sqrt{n}\|F_n - F_{n,res}\|_\infty = o_P^u(1) \;\text{ in } M \text{ as } n \to \infty.$$

It remains to investigate $\|F_n - F\|_\infty$. As already mentioned, it is well known that the Kolmogorov-Smirnov statistic satisfies

$$P_F\big(\sqrt{n}\|F_n - F\|_\infty \le x\big) \underset{n\to\infty}{\longrightarrow} P_F\big(\|B^\circ(F)\|_\infty \le x\big), \quad x \in \mathbb{R},$$

for every $F \in \mathcal{G}_{q,\gamma,w}$. Moreover, by the continuity of F the distributions of $\sqrt{n}\|F_n - F\|_\infty$, $n \ge 2$, and $\|B^\circ(F)\|_\infty$ under the measure P_F do not depend on F anymore, and the distribution function $P_F\big(\|B^\circ(F)\|_\infty \le \cdot\,\big) =: Q(\cdot)$ is continuous.

Thus, if we set

$$V_{n,F} := \sqrt{n}\|F_n - F\|_\infty \quad \text{and} \quad R_{n,F} := \sqrt{n}\|F_n - F_{n,res}\|_\infty$$

for every $F \in \mathcal{G}_{q,\gamma,w}$ and $n \ge 2$, the assumptions of Proposition 2.6 are satisfied with ϱ as above. This yields

$$\sup_{F\in M}\big|P_F(V_{n,F} + R_{n,F} \le x) - Q(x)\big| \underset{n}{\longrightarrow} 0 \quad \forall\, x \in \mathbb{R},$$

which means that assumption (i) of Proposition 2.4 holds for the family

$$\big\{\big(\sqrt{n}\|F_n - F\|_\infty + \sqrt{n}\|F_n - F_{n,res}\|_\infty\big)_{n\ge 2} \colon F \in \mathcal{G}_{q,\gamma,w}\big\}.$$

But by what was said above, condition (ii) of the same proposition is trivially met, so that it follows from this that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C(\epsilon,\delta) > 0$ such that

$$P_F\Big(\Big|\frac{\mathcal{W}_{n,res}}{\sqrt{n}} - b(F)\Big| \ge \epsilon\, b(F)\Big) \le P_F\big(\|F_n - F\|_\infty + \|F_n - F_{n,res}\|_\infty \ge \epsilon\, b(F)\big) < \delta$$

for all $F \in M$ and all $n \in \mathbb{N}$, $n \ge 2$, with $n > C(\epsilon,\delta)/b(F)^2$, which is (WIII) for $(\mathcal{W}_{n,res})_{n\ge 2}$.

We will show next that (WIII) holds for $(\mathcal{V}_{n,res})_{n\ge 2}$ as well.

Set $K' := (\int_{\mathbb{R}} |x| F_0(dx))/2 > 0$. It follows again from Corollary 6.3 (i) that there is a $\varrho_1' > 0$ such that

$$\Big|\int_{\mathbb{R}} |x| F(dx) - \int_{\mathbb{R}} |x| F_0(dx)\Big| < K' \;\text{ for all } F \in \mathcal{G}_{q,\gamma,w} \text{ with } d_{q,\gamma,w}(F,F_0) < \varrho_1'. \tag{6.13}$$

Next, note that $F_0(0) \in (0,1)$ because F_0 is centered and has positive variance. Hence, by the continuity of F_0 there are $x_1 \in (-\infty,0)$ and $x_2 \in (0,\infty)$ such that $F_0(x_1) > 0$ and $F_0(x_2) < 1$. Now set $K'' := \min\big(F_0(x_1), 1 - F_0(x_2)\big)/2$. Part (ii) of Corollary 6.3 now states the existence of a $\varrho_2' > 0$ with

$$d_K(F,F_0) = \|F - F_0\|_\infty < K'' \;\text{ for all } F \in \mathcal{G}_{q,\gamma,w} \text{ with } d_{q,\gamma,w}(F,F_0) < \varrho_2'. \tag{6.14}$$

Define $\varrho' := \min(\varrho_1', \varrho_2')$ and

$$M' := U_{\varrho'}(F_0) \setminus \{F_0\} = \big\{F \in \mathcal{G}_{q,\gamma,w} \colon d_{q,\gamma,w}(F,F_0) < \varrho'\big\} \setminus \{F_0\}.$$

We proceed by verifying that this set fulfills conditions (3.7), (3.8), (5.12)–(5.14) and (5.75).

The verification of (3.7) and (5.12)–(5.14) for M' is analogous to that for the set M before, simply replace ϱ by ϱ' and M by M'.

To see that (3.8) holds, note that (6.13) implies $\int_{\mathbb{R}} |x| F(dx) > K' > 0$ for all $F \in M'$.

It remains to check (5.75). Because of (6.14) we have $|F(x) - F_0(x)| < K''$ for all $x \in \mathbb{R}$ and for every $F \in M'$. This obviously implies

$$F(x_1) > F_0(x_1) - K'' > 0 \quad \text{and} \quad F(x_2) < F_0(x_2) + K'' < 1$$

for every $F \in M'$, so that (5.75) holds as well.

Note that (3.8) implies (3.5), so that the least squares estimator $\hat{\rho}_n^{LS}$ fulfills $\sqrt{n}(\hat{\rho}_n^{LS} - \rho) = O_P^u(1)$ in M' as $n \to \infty$ by Proposition 5.7.

Now using Minkowski's inequality we get that

$$\Big|\frac{\mathcal{V}_{n,res}}{\sqrt{n}} - b(F)\Big| \;\le\; \|\widetilde{F}_{n,res} - F\|_\infty \;\le\; \|\widetilde{F}_{n,res} - F_n\|_\infty + \|F_n - F\|_\infty \tag{6.15}$$

for every $n \ge 2$ and $F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}$. But as the set M' satisfies the assumptions of Proposition 5.15, it follows from this that

$$\|\widetilde{F}_{n,res} - F_n\|_\infty \;\le\; \|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{n}\sum_{i=1}^n e_i\Big| + \|R_{n,F}\|_\infty \tag{6.16}$$

with $\sqrt{n}\|R_{n,F}\|_\infty = o_P^u(1)$ in M' as $n \to \infty$. Because of this and Lemma 3.8 (i), (ii) it is clear that the assumptions of Proposition 2.6 hold with $\varrho = \varrho'$ and

$$\overline{V}_{n,F} := \|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n e_i\Big| \quad \text{and} \quad \overline{R}_{n,F} := \sqrt{n}\|R_{n,F}\|_\infty,$$

so that

$$\sup_{F \in M'}\Big|P_F\Big(\|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n e_i\Big| + \sqrt{n}\|R_{n,F}\|_\infty \le x\Big) - Q_F(x)\Big| \;\underset{n}{\longrightarrow}\; 0 \quad \forall\, x \in \mathbb{R}$$

with Q_F as in Lemma 3.8 (i). Now it is evident by the above and by Lemma 3.8 (iii) that the family

$$\Big\{\Big(\|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n e_i\Big| + \sqrt{n}\|R_{n,F}\|_\infty\Big)_{n\ge 2} \colon F \in \mathcal{G}_{q,\gamma,w}\Big\}$$

satisfies the assumptions of Proposition 2.4 with $\varrho = \varrho'$. Moreover, as was mentioned above, the family $\{(\sqrt{n}\|F_n - F\|_\infty)_{n\ge 2} \colon F \in \mathcal{G}_{q,\gamma,w}\}$ fulfills the requirements of this proposition as well for any value of $\varrho > 0$.

Thus, by Corollary 2.5 there is a $\tilde{\varrho} > 0$ such that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a positive constant $\widetilde{C}(\epsilon, \delta)$ with

$$\begin{aligned} P_F\Big(\Big|\frac{\mathcal{V}_{n,res}}{\sqrt{n}} - b(F)\Big| \ge \epsilon\, b(F)\Big) &\le P_F\big(\|\widetilde{F}_{n,res} - F\|_\infty \ge \epsilon\, b(F)\big) \\ &\le P_F\Big(\|U_F\|_\infty \frac{1}{\sigma_F^2}\Big|\frac{1}{n}\sum_{i=1}^n e_i\Big| + \|R_{n,F}\|_\infty + \|F_n - F\|_\infty \ge \epsilon\, b(F)\Big) < \delta \end{aligned}$$

for all $F \in U_{\tilde{\varrho}}(F_0) \setminus \{F_0\}$ and all $n \in \mathbb{N}$, $n \ge 2$, with $n > \widetilde{C}(\epsilon, \delta)/b(F)^2$. Note that we may take $\tilde{\varrho} = \varrho'$ here. □

Finally we are ready to determine the limiting (as $\alpha \to 0$) Pitman ARE of $(\mathcal{W}_{n,res})_{n\ge 2}$ with respect to $(\mathcal{V}_{n,res})_{n\ge 2}$.

Theorem 6.6
For $T_{1n} = \mathcal{W}_{n,res}$ and $T_{2n} = \mathcal{V}_{n,res}$, $n \ge 2$, it is

$$\lim_{\alpha\to 0} \liminf_{\substack{F \in \mathcal{G}_{q,\gamma,w}\setminus\{F_0\},\\ d_{q,\gamma,w}(F,F_0)\to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \lim_{\alpha\to 0} \limsup_{\substack{F \in \mathcal{G}_{q,\gamma,w}\setminus\{F_0\},\\ d_{q,\gamma,w}(F,F_0)\to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \frac{\lambda_1^*}{\lambda_1} \;<\; 1$$

for every value of $\beta \in (0,1)$.

Proof. The statement follows from Theorem 2.3 once we have made sure that all of its assumptions are satisfied here. But we have already shown that the sequences $(\mathcal{W}_{n,res})_{n\geq 2}$ and $(\mathcal{V}_{n,res})_{n\geq 2}$ satisfy condition (i) of Theorem 2.3. Moreover, as was mentioned in subsection 4.1 the distribution functions of the random variables $\mathcal{W}$ in (6.5) and $\mathcal{V}$ in (6.8) are strictly increasing on $(0, \infty)$, so that condition (ii) also holds. To verify condition (iii) of Theorem 2.3, note that $0 < b(F) \leq \|F - F_0\|_\infty = d_K(F, F_0)$ for every $F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}$, and

$$\lim_{\substack{F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}, \\ d_{q,\gamma,w}(F,F_0) \to 0}} d_K(F, F_0) = 0$$

by Corollary 6.3 (ii). The last assumption of Theorem 2.3 is again trivially satisfied here, since the approximate Bahadur ARE of the two sequences does not depend on $F \in \mathcal{G}_{q,\gamma,w} \setminus \{F_0\}$, cf. Proposition 6.1, and this concludes the proof. □

As the limiting Pitman ARE of $(\mathcal{W}_{n,res})_{n\geq 2}$ with respect to $(\mathcal{V}_{n,res})_{n\geq 2}$ is strictly less than one, the sequence of tests based on $(\mathcal{V}_{n,res})_{n\geq 2}$ is preferable to the one based on $(\mathcal{W}_{n,res})_{n\geq 2}$. Recall that for certain distribution functions F_0 the explicit value of the ratio λ_1^*/λ_1 is given in Table 1 on page 35.

6.2 Composite null hypothesis

Fix $q \geq 2$ again. In this subsection we assume that Q is the Dirac measure on $0 \in \mathbb{R}^p$, so that $\mathbf{S}_0 = (0, \ldots, 0)^T \in \mathbb{R}^p$ P_F-almost surely for every $F \in \mathcal{G}_q^u$. Thus, $(X_i)_{i \geq 1-p}$ is either the $AR(p)$ process of model 1 with starting values $X_0 = \ldots = X_{1-p} = 0$, or the stationary $AR(p)$ process of model 2.

Suppose again that the errors $e_1, e_2, \ldots$ are independent and identically distributed with common distribution function $F \in \mathcal{G}_q^u$, but that F is unknown to us. In this subsection we consider the problem of testing the composite null hypothesis

$$\mathrm{H}_0 \colon F \in \mathcal{F}_\tau = \Big\{ F_\tau\Big(\frac{\cdot}{\sigma}\Big) \colon \sigma \in (0, \infty) \Big\} \quad \text{versus} \quad \mathrm{H}_1 \colon F \in \mathcal{G}_q^u \setminus \mathcal{F}_\tau, \tag{6.17}$$

where F_τ is as in subsection 4.2 the distribution function of the exponential power distribution, and $\tau > 0$ is kept fixed. Since F_τ is centered, has moments of all order, and its Lebesgue density f_τ is uniformly continuous, it obviously is an element of $\mathcal{G}_q^u$. But if $G \in \mathcal{G}_q^u$, then it is easily checked that every scale variant $G(\cdot/c)$, $c > 0$, is contained in $\mathcal{G}_q^u$ as well. Consequently, $\mathcal{F}_\tau$ is indeed a subset of $\mathcal{G}_q^u$.

Adopting the notation of subsection 4.2, we set

$$F(x, \sigma) := F_\tau\Big(\frac{x}{\sigma}\Big) \quad \text{and} \quad f(x, \sigma) := \frac{1}{\sigma} f_\tau\Big(\frac{x}{\sigma}\Big)$$

for every $\sigma \in (0, \infty)$ and $x \in \mathbb{R}$, suppressing again the dependency of $F(x, \sigma)$ and $f(x, \sigma)$ on τ, as this is held constant.

As before, we will use the residuals

$$\hat{e}_{ni} = X_i - \hat{\rho}_n^T \mathbf{X}_{i-1}, \quad 1 \leq i \leq n,$$

based on observed data $\mathbf{X}_0, X_1, \ldots, X_n$, $n \geq 2$, and some suitable sequence of estimators $(\hat{\rho}_n)_{n \geq 2}$ for the unknown autoregressive parameter ρ in equation (6.1) to construct test statistics for the above testing problem. Again, we will consider in the following only such estimators $\hat{\rho}_n$ that satisfy

$$\sqrt{n}\big(\hat{\rho}_n - \rho\big) = O_{P_F}(1) \quad \text{as } n \to \infty. \tag{6.18}$$

As already mentioned, the least squares estimator $\hat{\rho}_n^{LS}$ for example fulfills this assumption.

Analogously to the case of independent and identically distributed data discussed in subsection 4.2, we will estimate the scale parameter σ of the family $\mathcal{F}_\tau$ by

$$\hat{\sigma}_{n,res} = \hat{\sigma}_{n,res}(\hat{e}_{n1}, \ldots, \hat{e}_{nn}) := \tau^{1/\tau} \Big(\frac{1}{n} \sum_{i=1}^n |\hat{e}_{ni}|^\tau \Big)^{1/\tau}, \quad n \geq 2. \tag{6.19}$$

Thus, the estimator $\hat{\sigma}_{n,res}$ is the residual-based version of the maximum likelihood estimator $\hat{\sigma}_n$ for σ based on independent and identically distributed observations, cf. (4.17). Note that $\hat{\sigma}_{n,res} = 0$ is equivalent to $\hat{e}_{ni} = 0$ for all $i = 1, \ldots, n$. Thus, on the event $A_n := \{\hat{e}_{ni} = 0 \ \forall\, i = 1, \ldots, n\}$ the estimator $\hat{\sigma}_{n,res}$ is not contained in the parameter space $(0, \infty)$. But since $A_n \subset \overline{\Omega}_{n,res}$ with

$$\Omega_{n,res} = \{ \min_{1 \leq i \leq n} \hat{e}_{ni} < 0 < \max_{1 \leq i \leq n} \hat{e}_{ni} \},$$

it follows from Lemma 5.11 that $P_F(A_n) \leq P_F(\overline{\Omega}_{n,res}) \to 0$ as $n \to \infty$. Hence, for all asymptotic investigations under the fixed probability measure P_F it suffices to work on $\overline{A}_n$, and therefore we will assume henceforth that $\overline{A}_n$ holds, which is equivalent to $\hat{\sigma}_{n,res} \in (0, \infty)$.

For testing (6.17) we will use the residual Cramér-von Mises statistic with estimated parameter

$$\widehat{\mathcal{W}}_{n,res}^2 = n \int_{-\infty}^{\infty} \big(F_{n,res}(x) - F(x, \hat{\sigma}_{n,res})\big)^2 F(dx, \hat{\sigma}_{n,res})$$

and its counterpart

$$\widehat{\mathcal{V}}_{n,res}^2 = n \int_{-\infty}^{\infty} \big(\widetilde{F}_{n,res}(x) - F(x, \hat{\sigma}_{n,res})\big)^2 F(dx, \hat{\sigma}_{n,res})$$

based on the centered empirical distribution function of the residuals.

Analogously to the verification of the scale invariance of the test statistics $\widehat{\mathcal{W}}_n^2$ and $\widehat{\mathcal{V}}_n^2$ in the case of independent and identically distributed data discussed in subsection 4.2 we can show that both of the above test statistics are scale invariant with respect to the residuals, which means that

$$\widehat{\mathcal{W}}_{n,res}^2(\hat{e}_{n1}, \ldots, \hat{e}_{nn}) = \widehat{\mathcal{W}}_{n,res}^2\Big(\frac{\hat{e}_{n1}}{c}, \ldots, \frac{\hat{e}_{nn}}{c}\Big)$$

and

$$\widehat{\mathcal{V}}_{n,res}^2(\hat{e}_{n1}, \ldots, \hat{e}_{nn}) = \widehat{\mathcal{V}}_{n,res}^2\Big(\frac{\hat{e}_{n1}}{c}, \ldots, \frac{\hat{e}_{nn}}{c}\Big)$$

for every $c \in (0, \infty)$ and $n \geq 2$. For the latter equality to hold on Ω, we assume analogously to the case of independent and identically distributed data that $\widehat{\mathcal{V}}_{n,res}^2$ is defined on $\overline{\Omega}_{n,res}$ in such a way that it is scale invariant with respect to the residuals on this event as well.

Suppose henceforth that the estimator $\hat{\rho}_n$ is scale invariant, i.e.,

$$\hat{\rho}_n(X_{1-p}, \ldots, X_n) = \hat{\rho}_n\Big(\frac{X_{1-p}}{c}, \ldots, \frac{X_n}{c}\Big) \quad \forall\, c \in (0, \infty),\ n \geq 2. \tag{6.20}$$

Then the residuals are scale equivariant, since

$$\frac{1}{c}\hat{e}_{ni}(X_{1-p}, \ldots, X_n) = \frac{X_i}{c} - \hat{\rho}_n\Big(\frac{X_{1-p}}{c}, \ldots, \frac{X_n}{c}\Big)^T \frac{1}{c}\mathbf{X}_{i-1} = \hat{e}_{ni}\Big(\frac{X_{1-p}}{c}, \ldots, \frac{X_n}{c}\Big)$$

for every $1 \leq i \leq n$, $n \geq 2$, and $c > 0$. As a consequence of this, the test statistics $\widehat{\mathcal{W}}_{n,res}^2$ and $\widehat{\mathcal{V}}_{n,res}^2$ are scale invariant with respect to the underlying data $\mathbf{X}_0, X_1, \ldots, X_n$, $n \geq 2$.

Now consider $\widetilde{X}_i := X_i/c$, $i \geq 1-p$, for some fixed $c > 0$. The process $(\widetilde{X}_i)_{i \geq 1-p}$ obviously satisfies

$$\widetilde{X}_i = \rho_1 \widetilde{X}_{i-1} + \ldots + \rho_p \widetilde{X}_{i-p} + \tilde{e}_i, \quad i \geq 1,$$

with $\tilde{e}_i = e_i/c$ and $\rho_1, \ldots, \rho_p$ as in (6.1). Moreover, we have $\widetilde{X}_{1-p} = \ldots = \widetilde{X}_0 = 0$ if $(X_i)_{i \geq 1-p}$ is the $AR(p)$ process with starting values $X_{1-p} = \ldots = X_0 = 0$, so that in this case both $(X_i)_{i \geq 1-p}$ and $(\widetilde{X}_i)_{i \geq 1-p}$ are $AR(p)$ processes that start in zero, but the sequences of error variables differ by the scale factor c. If $(X_i)_{i \geq 1-p}$ is the stationary $AR(p)$ process instead, then by using its $MA(\infty)$-representation (5.3) we get that

$$\widetilde{X}_i = \sum_{j=0}^{\infty} \psi_j \tilde{e}_{i-j} \quad \forall\, i \geq 1-p, \tag{6.21}$$

where $\tilde{e}_i = e_i/c$ for all $i \in \mathbb{Z}$, and the convergence of the series is as before in mean square as well as absolutely with probability one under every P_F. It is evident by this representation that $(\widetilde{X}_i)_{i \geq 1-p}$ is the stationary $AR(p)$ process with respect to the sequence $(\tilde{e}_i)_{i \in \mathbb{Z}}$ of error variables. Thus, in both of the models considered here the differences between the processes $(X_i)_{i \geq 1-p}$ and $(\widetilde{X}_i)_{i \geq 1-p}$ solely result from changing the scale of the underlying sequence of error variables.

Remark 6.7: Note that the least squares estimator $\hat{\rho}_n^{LS}$ is scale invariant, i.e., it fulfills (6.20). To verify this, recall that

$$\hat{\rho}_n^{LS} = \hat{\rho}_n^{LS}(X_{1-p}, \ldots, X_n) = \Big(\sum_{i=1}^n \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T\Big)^+ \sum_{i=1}^n \mathbf{X}_{i-1}X_i, \quad n \geq 2,$$

with B^+ denoting the Moore-Penrose pseudoinverse of the matrix B. As the Moore-Penrose pseudoinverse satisfies $(cB)^+ = c^{-1}B^+$ for every $c \in \mathbb{R} \setminus \{0\}$, it follows that

$$\begin{aligned}\hat{\rho}_n^{LS}(X_{1-p}, \ldots, X_n) &= \frac{c^2}{c^2}\Big(\sum_{i=1}^n \mathbf{X}_{i-1}\mathbf{X}_{i-1}^T\Big)^+ \sum_{i=1}^n \mathbf{X}_{i-1}X_i = \Big(\sum_{i=1}^n \frac{1}{c}\mathbf{X}_{i-1}\frac{1}{c}\mathbf{X}_{i-1}^T\Big)^+ \sum_{i=1}^n \frac{1}{c}\mathbf{X}_{i-1}\frac{X_i}{c} \\ &= \hat{\rho}_n^{LS}\Big(\frac{X_{1-p}}{c}, \ldots, \frac{X_n}{c}\Big)\end{aligned}$$

for all $c > 0$. ♦

We will now construct asymptotic level α tests for the testing problem (6.17) based on the test statistics $\widehat{\mathcal{W}}_{n,res} = \big(\widehat{\mathcal{W}}^2_{n,res}\big)^{1/2}$ and $\widehat{\mathcal{V}}_{n,res} = \big(\widehat{\mathcal{V}}^2_{n,res}\big)^{1/2}$. By what was mentioned on the previous page, it suffices again to assume that $F = F_\tau$ under H_0 in order to determine the asymptotic null distributions of these statistics.

Let us start by showing that under the measure P_{F_τ} the scale estimator $\hat{\sigma}_{n,res}$ admits the same asymptotic linear expansion as $\hat{\sigma}_n$ in the case of independent and identically distributed data, see (4.20).

Proposition 6.8
Suppose that $F = F_\tau$ and the sequence $(\hat{\rho}_n)_{n\geq 2}$ satisfies (6.18). Then

$$\hat{\sigma}_{n,res}(\hat{e}_{n1}, \ldots, \hat{e}_{nn}) - 1 = \frac{1}{n}\sum_{i=1}^n L(e_i) + o_{P_{F_\tau}}(n^{-1/2}),$$

where $L(x) = |x|^\tau - 1/\tau$, $x \in \mathbb{R}$.

Proof. For every $n \in \mathbb{N}$, $n \geq 2$, it is

$$\hat{\sigma}_{n,res} - 1 = \Big(\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau\Big)^{1/\tau} - 1^{1/\tau} = g_\tau\Big(\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau\Big) - g_\tau(1)$$

with $g_\tau(x) := x^{1/\tau}$, $x \in [0, \infty)$. Since g_τ is continuously differentiable on $(0, \infty)$ with $g'_\tau(x) = (1/\tau)x^{1/\tau - 1}$, it follows from the mean value theorem that there is a ξ_n between $\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau$ and 1 such that

$$\begin{aligned}\hat{\sigma}_{n,res} - 1 &= g'_\tau(\xi_n)\Big(\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau - 1\Big) \\ &= g'_\tau(1)\Big(\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau - 1\Big) + \big(g'_\tau(\xi_n) - g'_\tau(1)\big)\Big(\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau - 1\Big) =: \mathrm{I}_n + \mathrm{II}_n.\end{aligned}$$

Now

$$\mathrm{I}_n = \frac{1}{\tau}\Big(\frac{\tau}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau - 1\Big) = \frac{1}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau - \frac{1}{\tau}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big(|e_i|^\tau - \frac{1}{\tau}\Big) + \Big(\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau - \frac{1}{n}\sum_{i=1}^{n}|e_i|^\tau\Big),$$

so we need to show that

$$\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau - \frac{1}{n}\sum_{i=1}^{n}|e_i|^\tau = o_{P_{F_\tau}}(n^{-1/2}) \quad \text{as} \quad n\to\infty. \tag{6.22}$$

Assume now that (6.22) holds. Then (6.22) and the strong law of large numbers imply

$$\frac{\tau}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau = \tau\Big(\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau - \frac{1}{n}\sum_{i=1}^{n}|e_i|^\tau + \frac{1}{n}\sum_{i=1}^{n}|e_i|^\tau\Big) \underset{n\to\infty}{\longrightarrow} \tau\,\mathrm{E}_{F_\tau}(|e_1|^\tau) = 1$$

in P_{F_τ}-probability. Thus, (6.22) yields

$$\begin{aligned}\mathrm{II}_n &= \big(g_\tau'(\xi_n) - g_\tau'(1)\big)\Big(\frac{\tau}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau - \frac{\tau}{n}\sum_{i=1}^{n}|e_i|^\tau + \frac{\tau}{n}\sum_{i=1}^{n}|e_i|^\tau - 1\Big)\\ &= o_{P_{F_\tau}}(1)\big(o_{P_{F_\tau}}(n^{-1/2}) + O_{P_{F_\tau}}(n^{-1/2})\big) = o_{P_{F_\tau}}(n^{-1/2}),\end{aligned}$$

where $\frac{\tau}{n}\sum_{i=1}^{n}|e_i|^\tau - 1 = O_{P_{F_\tau}}(n^{-1/2})$ is easily seen to be true with Chebychev's inequality. Thus, it only remains to verify (6.22), which is equivalent to

$$A_n := \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big) = o_{P_{F_\tau}}(1) \quad \text{as} \quad n\to\infty.$$

Since $P_{F_\tau}(e_i \neq 0 \text{ for all } i \in \mathbb{N}) = 1$, we will assume that $|e_i| > 0$ for all $i \in \mathbb{N}$. Now observe that the function $h_\tau(x) := x^\tau$, $x \in [0,\infty)$, is infinitely often differentiable on $(0,\infty)$ and $h_\tau'(x) = \tau x^{\tau-1}$. Thus, by the mean value theorem for every $i \in \{1,\dots,n\}$ there is a ζ_i between $|\hat{e}_{ni}|$ and $|e_i|$ in $(0,\infty)$ such that

$$\begin{aligned}A_n &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}h_\tau'(\zeta_i)\big(|\hat{e}_{ni}| - |e_i|\big)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}h_\tau'(|e_i|)\big(|\hat{e}_{ni}| - |e_i|\big) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(h_\tau'(\zeta_i) - h_\tau'(|e_i|)\big)\big(|\hat{e}_{ni}| - |e_i|\big)\\ &=: A_{n,1} + A_{n,2}.\end{aligned}$$

First, we will investigate $A_{n,1}$. It is

$$A_{n,1} = \frac{\tau}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\big(|\hat{e}_{ni}| - |e_i|\big).$$

Now for every $\delta > 0$ and $n \geq 2$ set $D_{n,\delta} := \{\max_{1\leq i\leq n}|\hat{e}_{ni} - e_i| < \delta\}$. Note that $P_{F_\tau}(\overline{D}_{n,\delta}) = o(1)$ as $n \to \infty$ for every $\delta > 0$ by part (ii) of Lemma 5.8 (consider the singleton $M = \{F_\tau\}$ there). On $D_{n,\delta}$ we have

$$\begin{aligned}A_{n,1} = &\frac{\tau}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\big(\hat{e}_{ni} - e_i\big)1_{\{e_i>\delta\}} + \frac{\tau}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\big(|\hat{e}_{ni}| - |e_i|\big)1_{\{-\delta\leq e_i\leq\delta\}}\\ &+ \frac{\tau}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\big(e_i - \hat{e}_{ni}\big)1_{\{e_i<-\delta\}}\end{aligned}$$

$$
\begin{aligned}
&= \frac{\tau}{\sqrt{n}} \sum_{i=1}^{n} |e_i|^{\tau-1}\big(\hat{e}_{ni} - e_i\big)\big(1_{\{e_i>\delta\}} - 1_{\{e_i<-\delta\}}\big) + \frac{\tau}{\sqrt{n}} \sum_{i=1}^{n} |e_i|^{\tau-1}\big(|\hat{e}_{ni}| - |e_i|\big)1_{\{-\delta\le e_i\le\delta\}} \\
&=: R_{n,\delta}^{(i)} + R_{n,\delta}^{(ii)}.
\end{aligned}
$$

It is

$$
R_{n,\delta}^{(i)} = \frac{\tau}{\sqrt{n}} \sum_{i=1}^{n} \big(\rho - \hat{\rho}_n\big)^T \mathbf{X}_{i-1}|e_i|^{\tau-1}\big(1_{\{e_i>\delta\}} - 1_{\{e_i<-\delta\}}\big) = \tau\big(\rho - \hat{\rho}_n\big)^T \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_{i-1}\varepsilon_i \tag{6.23}
$$

with $\varepsilon_i := |e_i|^{\tau-1}\big(1_{\{e_i>\delta\}} - 1_{\{e_i<-\delta\}}\big)$, $i \in \mathbb{N}$. Now note that under the measure P_{F_τ} we have that $\mathbf{X}_{i-1}$ and ε_i are independent and ε_i is square-integrable. To verify the latter, observe that $|\varepsilon_i| \le \delta^{\tau-1}$ for $\tau \in (0,1)$. If $\tau \ge 1$, then $\mathrm{E}_{F_\tau}(\varepsilon_i^2) < \infty$ since F_τ has moments of all order. Furthermore, we have

$$
\mathrm{E}_{F_\tau}(\varepsilon_i) = \mathrm{E}_{F_\tau}\big(|e_i|^{\tau-1}1_{\{e_i>\delta\}}\big) - \mathrm{E}_{F_\tau}\big(|-e_i|^{\tau-1}1_{\{-e_i>\delta\}}\big) = 0
$$

by the symmetry of F_τ. It follows from this that under P_{F_τ} the sequence of random vectors $(\mathbf{X}_{i-1}\varepsilon_i)_{i\ge1}$ is a square-integrable martingale difference sequence with respect to the filtration $\mathcal{F}_i := \sigma(X_{1-p},\ldots,X_0,e_1,\ldots,e_i)$, $i \ge 1$, $\mathcal{F}_0 := \{\emptyset,\Omega\}$. Hence, by Markov's inequality we get for every $C \in (0,\infty)$

$$
\begin{aligned}
P_{F_\tau}\Big(\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{X}_{i-1}\varepsilon_i\Big\| \ge C\Big) &\le \frac{\mathrm{E}_{F_\tau}\Big(\Big\|\sum_{i=1}^{n}\mathbf{X}_{i-1}\varepsilon_i\Big\|^2\Big)}{nC^2} = \frac{\sum_{i=1}^{n}\mathrm{E}_{F_\tau}\big(\|\mathbf{X}_{i-1}\varepsilon_i\|^2\big)}{nC^2} \\
&= \frac{\mathrm{E}_{F_\tau}(\varepsilon_1^2)\sum_{i=1}^{n}\mathrm{E}_{F_\tau}\big(\|\mathbf{X}_{i-1}\|^2\big)}{nC^2} \\
&\le \frac{\mathrm{E}_{F_\tau}(\varepsilon_1^2)K}{C^2}
\end{aligned}
$$

since $\big(\mathrm{E}_{F_\tau}\big(\|\mathbf{X}_{i-1}\|^2\big)\big)_{i\ge1}$ is bounded by some constant $K \in (0,\infty)$, which follows from Lemma 5.5 (i) by considering $M = \{F_\tau\}$ there. Thus, we have shown that

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{X}_{i-1}\varepsilon_i = O_{P_{F_\tau}}(1).
$$

It follows from this and (6.18) that

$$
R_{n,\delta}^{(i)} = O_{P_{F_\tau}}(n^{-1/2})O_{P_{F_\tau}}(1) = o_{P_{F_\tau}}(1). \tag{6.24}
$$

Combining all this, we have for every $\epsilon, \delta > 0$

$$
\begin{aligned}
P_{F_\tau}\big(|A_{n,1}| \ge \epsilon\big) &\le P_{F_\tau}\big(\{|A_{n,1}| \ge \epsilon\} \cap D_{n,\delta}\big) + P_{F_\tau}\big(\overline{D}_{n,\delta}\big) \\
&= P_{F_\tau}\big(\{|R_{n,\delta}^{(i)} + R_{n,\delta}^{(ii)}| \ge \epsilon\} \cap D_{n,\delta}\big) + P_{F_\tau}\big(\overline{D}_{n,\delta}\big) \\
&\le P_{F_\tau}\big(|R_{n,\delta}^{(i)}| \ge \epsilon/2\big) + P_{F_\tau}\big(|R_{n,\delta}^{(ii)}| \ge \epsilon/2\big) + P_{F_\tau}\big(\overline{D}_{n,\delta}\big) \\
&= P_{F_\tau}\big(|R_{n,\delta}^{(ii)}| \ge \epsilon/2\big) + o(1)
\end{aligned} \tag{6.25}
$$

as $n \to \infty$. But

$$
|R_{n,\delta}^{(ii)}| \le \frac{\tau}{\sqrt{n}}\|\hat{\rho}_n - \rho\| \sum_{i=1}^{n} |e_i|^{\tau-1}\|\mathbf{X}_{i-1}\|1_{\{-\delta\le e_i\le\delta\}}, \tag{6.26}
$$

so that for every $\epsilon, \delta > 0$ and $0 < c < \infty$ it is

$$\begin{aligned} P_{F_\tau}\big(|R_{n,\delta}^{(ii)}| \geq \epsilon\big) &\leq P_{F_\tau}\big(\{|R_{n,\delta}^{(ii)}| \geq \epsilon\} \cap \{\sqrt{n}\|\hat{\rho}_n - \rho\| \leq c\}\big) + P_{F_\tau}\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > c\big) \\ &\leq P_{F_\tau}\Big(\sum_{i=1}^n |e_i|^{\tau-1}\|\mathbf{X}_{i-1}\|1_{\{-\delta \leq e_i \leq \delta\}} \geq n\epsilon/(\tau c)\Big) + P_{F_\tau}\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > c\big) \\ &\leq \frac{\tau c}{n\epsilon}\sum_{i=1}^n \mathrm{E}_{F_\tau}\big(|e_1|^{\tau-1}1_{\{-\delta \leq e_1 \leq \delta\}}\big)\,\mathrm{E}_{F_\tau}\big(\|\mathbf{X}_{i-1}\|\big) + P_{F_\tau}\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > c\big). \end{aligned}$$

Now using Lemma 5.4 (i) with $M = \{F_\tau\}$, there is a $\widetilde{K} > 0$ such that

$$\mathrm{E}_{F_\tau}\big(\|\mathbf{X}_{i-1}\|\big) \leq \widetilde{K} \quad \forall\, i \in \mathbb{N}.$$

Moreover, it is

$$\mathrm{E}_{F_\tau}\big(|e_1|^{\tau-1}1_{\{-\delta \leq e_1 \leq \delta\}}\big) = 2\int_0^\delta x^{\tau-1}f_\tau(x)dx \leq 2\|f_\tau\|_\infty \int_0^\delta x^{\tau-1}dx = \frac{2}{\tau}\|f_\tau\|_\infty \delta^\tau.$$

Thus, we have

$$P_{F_\tau}\big(|R_{n,\delta}^{(ii)}| \geq \epsilon\big) \leq \frac{2c\|f_\tau\|_\infty \widetilde{K}}{\epsilon}\delta^\tau + P_{F_\tau}\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > c\big). \tag{6.27}$$

Combining (6.25) and (6.27) we get for every $\epsilon, \delta > 0$ and $0 < c < \infty$

$$\limsup_{n\to\infty} P_{F_\tau}\big(|A_{n,1}| \geq \epsilon\big) \leq \frac{4c\|f_\tau\|_\infty \widetilde{K}}{\epsilon}\delta^\tau + \limsup_{n\to\infty} P_{F_\tau}\big(\sqrt{n}\|\hat{\rho}_n - \rho\| > c\big).$$

By letting δ tend to zero first and then c tend to infinity, it follows with (6.18) that $A_{n,1} = o_{P_{F_\tau}}(1)$. It remains to study

$$A_{n,2} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \big(h'_\tau(\zeta_i) - h'_\tau(|e_i|)\big)\big(|\hat{e}_{ni}| - |e_i|\big).$$

First assume that $\tau \in [2, \infty)$. It is $h''_\tau(x) = \tau(\tau-1)x^{\tau-2}$, $x \in (0, \infty)$, and by the mean value theorem there is an η_i between ζ_i and $|e_i|$ such that

$$\begin{aligned} |A_{n,2}| &\leq \frac{\tau(\tau-1)}{\sqrt{n}}\sum_{i=1}^n \eta_i^{\tau-2}\big|\zeta_i - |e_i|\big| \cdot \big||\hat{e}_{ni}| - |e_i|\big| \\ &\leq \frac{\tau(\tau-1)}{\sqrt{n}}\sum_{i=1}^n \max(|\hat{e}_{ni}|, |e_i|)^{\tau-2}\big||\hat{e}_{ni}| - |e_i|\big|^2 \\ &\leq \frac{\tau(\tau-1)}{\sqrt{n}}\Big(\max_{1\leq i\leq n}|\hat{e}_{ni}| + \max_{1\leq i\leq n}|e_i|\Big)^{\tau-2}\sum_{i=1}^n \big|\hat{e}_{ni} - e_i\big|^2. \end{aligned} \tag{6.28}$$

Now note that for all $\tau \in (0, \infty)$

$$\sum_{i=1}^n \big|\hat{e}_{ni} - e_i\big|^2 \leq \|\hat{\rho}_n - \rho\|^2 \sum_{i=1}^n \|\mathbf{X}_{i-1}\|^2 = O_{P_{F_\tau}}(n^{-1})O_{P_{F_\tau}}(n) = O_{P_{F_\tau}}(1) \tag{6.29}$$

by (6.18) and Lemma 5.5 (ii) (consider $M = \{F_\tau\}$ again). For $\tau = 2$, this immediately implies that $A_{n,2} = o_{P_{F_\tau}}(1)$. To handle the case $\tau > 2$, we have to work some more.

For every $\tau \in (0, \infty)$ it is

$$\max_{1 \le i \le n} |e_i| = o_{P_{F_\tau}}(n^\delta) \quad \text{for every} \quad \delta > 0. \tag{6.30}$$

To see that this is true, note that for every $\epsilon, \delta > 0$ it is

$$P_{F_\tau}\big(\max_{1 \le i \le n} |e_i| \ge n^\delta \epsilon\big) \le n P_{F_\tau}\big(|e_1| \ge n^\delta \epsilon\big) = n P_{F_\tau}\big(|e_1|^{2/\delta} \ge n^2 \epsilon^{2/\delta}\big) \le \frac{n}{n^2 \epsilon^{2/\delta}} \operatorname{E}_{F_\tau}\big(|e_1|^{2/\delta}\big) \underset{n\to\infty}{\longrightarrow} 0$$

using that F_τ has finite moments of all order. It follows from this that

$$\max_{1 \le i \le n} |\hat{e}_{ni}| \le \max_{1 \le i \le n} |\hat{e}_{ni} - e_i| + \max_{1 \le i \le n} |e_i| = o_{P_{F_\tau}}(1) + o_{P_{F_\tau}}(n^\delta) = o_{P_{F_\tau}}(n^\delta) \tag{6.31}$$

for every $\delta > 0$ using Lemma 5.8 (ii) with $M = \{F_\tau\}$.

Coming back to (6.28) for $\tau > 2$, we see that by (6.30) and (6.31) it is

$$|A_{n,2}| \le \frac{\tau(\tau - 1)}{\sqrt{n}} o_{P_{F_\tau}}(n^\delta) O_{P_{F_\tau}}(1)$$

for every $\delta > 0$. Hence, choosing $\delta \in (0, 1/2]$, this yields $A_{n,2} = o_{P_{F_\tau}}(1)$ for $\tau > 2$ as well.

Now consider $\tau \in (1, 2)$. Then

$$\begin{aligned}
|A_{n,2}| &\le \frac{1}{\sqrt{n}} \sum_{i=1}^n \big|h'_\tau(\zeta_i) - h'_\tau(|e_i|)\big| \cdot \big||\hat{e}_{ni}| - |e_i|\big| \\
&\le \frac{\tau}{\sqrt{n}} \sum_{i=1}^n \big|\zeta_i^{\tau-1} - |e_i|^{\tau-1}\big| \cdot \big|\hat{e}_{ni} - e_i\big| \le \frac{\tau}{\sqrt{n}} \sum_{i=1}^n \big|\zeta_i - |e_i|\big|^{\tau-1} \cdot \big|\hat{e}_{ni} - e_i\big| \\
&\le \frac{\tau}{\sqrt{n}} \sum_{i=1}^n \big|\hat{e}_{ni} - e_i\big|^{\tau-1} \cdot \big|\hat{e}_{ni} - e_i\big| \le \big(\max_{1 \le i \le n} |\hat{e}_{ni} - e_i|\big)^{\tau-1} \frac{\tau}{\sqrt{n}} \sum_{i=1}^n \big|\hat{e}_{ni} - e_i\big| = o_{P_{F_\tau}}(1)
\end{aligned}$$

because of Lemma 5.8 (ii), $\tau > 1$ and

$$\sum_{i=1}^n \big|\hat{e}_{ni} - e_i\big| \le \|\hat{\rho}_n - \rho\| \sum_{i=1}^n \|\mathbf{X}_{i-1}\| = O_{P_{F_\tau}}(n^{-1/2}) O_{P_{F_\tau}}(n) = O_{P_{F_\tau}}(n^{1/2}),$$

where we used Lemma 5.4 (ii). Thus, $A_{n,2} = o_{P_{F_\tau}}(1)$ also for $\tau \in (1, 2)$. As $A_{n,2} = 0$ for $\tau = 1$, we now have verified (6.22) for every $\tau \in [1, \infty)$.

It remains to verify (6.22) for $\tau \in (0, 1)$, i.e., we have to show that

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big) = o_{P_{F_\tau}}(1), \quad \tau \in (0, 1).$$

For every $\delta > 0$ it is

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big) 1_{\{|e_i| \le \delta\}} + \frac{1}{\sqrt{n}} \sum_{i=1}^n \big(h_\tau(|\hat{e}_{ni}|) - h_\tau(|e_i|)\big) 1_{\{|e_i| > \delta\}} =: B_{n,\delta} + U_{n,\delta}.$$

Recall that $D_{n,\delta} = \{\max_{1 \le i \le n} |\hat{e}_{ni} - e_i| < \delta\}$. On $D_{n,\delta/2}$ we have

$$U_{n,\delta} = \frac{1}{\sqrt{n}} \sum_{i=1}^n h'_\tau(\xi_i) \big(|\hat{e}_{ni}| - |e_i|\big) 1_{\{|e_i| > \delta\}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n h'_\tau(|e_i|) \big(\hat{e}_{ni} - e_i\big) \big(1_{\{e_i > \delta\}} - 1_{\{e_i < -\delta\}}\big)$$

$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(h'_\tau(\xi_i)-h'_\tau(|e_i|)\big)\big(|\hat{e}_{ni}|-|e_i|\big)1_{\{|e_i|>\delta\}}$$

with ξ_i lying between $|\hat{e}_{ni}|$ and $|e_i|$, $i=1,\ldots,n$. Another application of the mean value theorem gives

$$\begin{aligned}U_{n,\delta}&=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}h'_\tau(|e_i|)\big(\hat{e}_{ni}-e_i\big)\big(1_{\{e_i>\delta\}}-1_{\{e_i<-\delta\}}\big)\\&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}h''_\tau(\zeta_i)\big(\xi_i-|e_i|\big)\big(|\hat{e}_{ni}|-|e_i|\big)1_{\{|e_i|>\delta\}}=:S_{n,\delta}+T_{n,\delta},\end{aligned}$$

where ζ_i lies between ξ_i and $|e_i|$ and thus between $|\hat{e}_{ni}|$ and $|e_i|$.

Now note that $S_{n,\delta}=R^{(i)}_{n,\delta}$ with $R^{(i)}_{n,\delta}$ from (6.23), and it has been shown in (6.24) that $R^{(i)}_{n,\delta}=o_{P_{F_\tau}}(1)$ for all $\tau\in(0,\infty)$.

Moreover, on $D_{n,\delta/2}$ we have

$$\begin{aligned}|T_{n,\delta}|&\le\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|h''_\tau(\zeta_i)|\big|\hat{e}_{ni}-e_i\big|^2 1_{\{|e_i|>\delta\}}\le\frac{\tau|\tau-1|}{\sqrt{n}}\sum_{i=1}^{n}\Big(\frac{\delta}{2}\Big)^{\tau-2}\big|\hat{e}_{ni}-e_i\big|^2 1_{\{|e_i|>\delta\}}\\&\le\Big(\frac{\delta}{2}\Big)^{\tau-2}\frac{\tau|\tau-1|}{\sqrt{n}}\sum_{i=1}^{n}\big|\hat{e}_{ni}-e_i\big|^2=o_{P_{F_\tau}}(1)\end{aligned}$$

using (6.29).

We still have to investigate $B_{n,\delta}$. A short computation shows that $|y^\tau-x^\tau|\le x^{\tau-1}|y-x|$ for every $x,y\in(0,\infty)$ and $\tau\in(0,1)$. Note that for $y=0$ this inequality is trivially fulfilled. By applying this inequality we get

$$\begin{aligned}|B_{n,\delta}|&\le\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\big||\hat{e}_{ni}|-|e_i|\big|1_{\{|e_i|\le\delta\}}\le\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\big|\hat{e}_{ni}-e_i\big|1_{\{|e_i|\le\delta\}}\\&\le\|\hat{\rho}_n-\rho\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\|\mathbf{X}_{i-1}\|1_{\{|e_i|\le\delta\}},\end{aligned}$$

and the right-hand side of the last inequality is equal to the upper bound in (6.26) up to the factor τ.

Combining all this, we finally get for every $\epsilon,\delta>0$ and $c\in(0,\infty)$

$$\begin{aligned}P_{F_\tau}\big(|A_n|\ge\epsilon\big)&\le P_{F_\tau}\big(|B_{n,\delta}|\ge\epsilon/2\big)+P_{F_\tau}\big(\{|U_{n,\delta}|\ge\epsilon/2\}\cap D_{n,\delta/2}\big)+P_{F_\tau}\big(\overline{D}_{n,\delta/2}\big)\\&\le P_{F_\tau}\Big(\|\hat{\rho}_n-\rho\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|e_i|^{\tau-1}\|\mathbf{X}_{i-1}\|1_{\{|e_i|\le\delta\}}\ge\epsilon/2\Big)+o(1)\\&\le\frac{4c\|f_\tau\|_\infty\widetilde{K}}{\tau\epsilon}\delta^\tau+P_{F_\tau}\big(\sqrt{n}\|\hat{\rho}_n-\rho\|>c\big)+o(1)\end{aligned}$$

analogously to the derivation of (6.27). By taking the limit superior of both sides and afterward letting first δ tend to zero and then c tend to infinity, the desired result follows. □

Note that the previous proposition implies the consistency of $\hat{\sigma}_{n,res}$ under H_0, since

$$P_{F_\tau(\cdot/\sigma)}\big(|\hat{\sigma}_{n,res}-\sigma|\ge\epsilon\big)=P_{F_\tau}\big(|\hat{\sigma}_{n,res}-1|\ge\epsilon/\sigma\big)\underset{n\to\infty}{\longrightarrow}0$$

for every $\epsilon,\sigma>0$ using the scaling properties of $\hat{\sigma}_{n,res}$ and of the residuals.

Now consider $F = F_\tau$ again. For every $x \in \mathbb{R}$ and $n \geq 2$ it is

$$\sqrt{n}\big(F_{n,res}(x) - F(x, \hat\sigma_{n,res})\big) = \sqrt{n}\big(F_{n,res}(x) - F_n(x)\big) + \sqrt{n}\big(F_n(x) - F(x, \hat\sigma_{n,res})\big),$$

and by using (6.3) it follows that $\|\sqrt{n}(F_{n,res} - F_n)\|_\infty \underset{n}{\to} 0$ in P_{F_τ}-probability. Additionally, we have

$$\sqrt{n}\big(F_n(\cdot) - F(\cdot, \hat\sigma_{n,res})\big) \quad \underset{n}{\xrightarrow{\mathcal{L}}} \quad Z \quad \text{in } D[-\infty, \infty]$$

under P_{F_τ}, where the process Z is as in (4.21). This is proven analogously to the functional central limit theorem (4.21) by using the differentiability of the mapping in (4.18) and the linear expansion of the estimator $\hat\sigma_{n,res}$, which has just been shown in Proposition 6.8. Hence, we get by Slutzky's theorem that

$$\sqrt{n}\big(F_{n,res}(\cdot) - F(\cdot, \hat\sigma_{n,res})\big) \quad \underset{n}{\xrightarrow{\mathcal{L}}} \quad Z \quad \text{in } D[-\infty, \infty]$$

under P_{F_τ}, and Z is a centered Gaussian process with continuous sample paths and covariance function given in (4.23).

Moreover, for every $x \in \mathbb{R}$ and $n \geq 2$ it is

$$\begin{aligned} \sqrt{n}\big(\widetilde{F}_{n,res}(x) - F(x, \hat\sigma_{n,res})\big) &= \sqrt{n}\big(\widetilde{F}_{n,res}(x) - F_\tau(x)\big) - \sqrt{n}\big(F(x, \hat\sigma_{n,res}) - F_\tau(x)\big) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(x) + R_{1n}(x) - \sqrt{n}\big(F(x, \hat\sigma_{n,res}) - F_\tau(x)\big) \end{aligned} \tag{6.32}$$

using (6.6) with $F = F_\tau$, where

$$Y_i(x) = 1_{\{e_i \leq x\}} - F_\tau(x) - \frac{U_{F_\tau}(x)}{\sigma^2_{F_\tau}} e_i, \quad i \in \mathbb{N},$$

and $\|R_{1n}\|_\infty$ converges to zero in P_{F_τ}-probability. Furthermore, it follows from the differentiability of the mapping in (4.18) and the linear expansion of $\hat\sigma_{n,res}$ shown in Proposition 6.8 that

$$\sqrt{n}\big(F(x, \hat\sigma_{n,res}) - F_\tau(x)\big) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Delta(x) L(e_i) + R_{2n}(x) \tag{6.33}$$

with Δ as in (4.19) and $\|R_{2n}\|_\infty$ converging to zero in P_{F_τ}-probability. Combining (6.32) and (6.33), Slutzky's theorem implies that the processes $\sqrt{n}\big(\widetilde{F}_{n,res}(\cdot) - F(\cdot, \hat\sigma_{n,res})\big)$ and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \big(Y_i(\cdot) - \Delta(\cdot) L(e_i)\big) \tag{6.34}$$

have the same asymptotic distribution in $D[-\infty, \infty]$ under P_{F_τ}. Now observe that the summand $Y_i(\cdot) - \Delta(\cdot)L(e_i)$ of the latter process is just the process in equation (2.17) in Genz [11] in the special case that $F(\cdot, \vartheta_0) = F_\tau$, $L(e_i, \vartheta_0) = L(e_i)$ and $\Delta(x, \vartheta_0) = \Delta(x)$. Hence, it follows from the proof of Satz 2.6 in [11] that under P_{F_τ} the process in (6.34) converges in distribution in $D[-\infty, \infty]$ to the process V from (4.24), which implies that

$$\sqrt{n}\big(\widetilde{F}_{n,res}(\cdot) - F(\cdot, \hat\sigma_{n,res})\big) \quad \underset{n}{\xrightarrow{\mathcal{L}}} \quad V \quad \text{in } D[-\infty, \infty]$$

under P_{F_τ}, and V is the centered Gaussian process that already appeared in (4.24). For autoregressive processes of order one, the functional central limit theorem for the residual empirical process with estimated parameter based on $\widetilde{F}_{n,res}$ has been established in a general setting by Genz [11], see Lemma 3.8 and Satz 3.9 in [11].

Now recall that we have shown in (4.26) that $\int_{\mathbb{R}^*} |xf_\tau'(x)|dx < \infty$. Hence, an application of Lemma 2.5 from Hörmann [15] in combination with Example 2.6 from [15] and the continuous mapping theorem yields

$$\widehat{\mathcal{W}}_{n,res} \quad \xrightarrow[n]{\mathcal{L}} \quad \Big(\int_{-\infty}^{\infty} Z(x)^2 F_\tau(dx)\Big)^{1/2} = \widehat{\mathcal{W}} \tag{6.35}$$

and

$$\widehat{\mathcal{V}}_{n,res} \quad \xrightarrow[n]{\mathcal{L}} \quad \Big(\int_{-\infty}^{\infty} V(x)^2 F_\tau(dx)\Big)^{1/2} = \widehat{\mathcal{V}} \tag{6.36}$$

under the measure P_{F_τ}. Note that the random variables $\widehat{\mathcal{W}}$ and $\widehat{\mathcal{V}}$ already appeared as the limits in (4.27) and (4.28). Hence, we can construct tests of asymptotic level $\alpha \in (0,1)$ for the testing problem (6.17) based on $\widehat{\mathcal{W}}_{n,res}$ and $\widehat{\mathcal{V}}_{n,res}$ just as described in subsection 4.2 for the case of independent and identically distributed observations, where we used $\widehat{\mathcal{W}}_n$ and $\widehat{\mathcal{V}}_n$ instead.

For the following investigations, we require again that $q \geq \max(2,\tau)$.

Recall that in subsection 4.2 we have studied the equivalence relation

$$F \sim_R G \quad \Longleftrightarrow \quad F(m_\tau(F)\,\cdot\,) = G(m_\tau(G)\,\cdot\,) \tag{6.37}$$

on $\mathcal{G}_q$, where $m_\tau(F) = (\tau \int_{\mathbb{R}} |x|^\tau F(dx))^{1/\tau} \in (0,\infty)$ for every $F \in \mathcal{G}_q$. Since $\mathcal{G}_q^u$ is a subset of $\mathcal{G}_q$, this relation is obviously an equivalence relation on $\mathcal{G}_q^u$ as well, and

$$[F]_R := \{G \in \mathcal{G}_q^u \colon G \sim_R F\} = \big\{F\big(\cdot/c\big) \colon c \in (0,\infty)\big\}$$

is the equivalence class of $F \in \mathcal{G}_q^u$ under it. Now note that it follows from the considerations on page 111 that the mappings

$$F \mapsto P_F \circ \widehat{\mathcal{W}}_{n,res}^{-1} \quad \text{and} \quad F \mapsto P_F \circ \widehat{\mathcal{V}}_{n,res}^{-1}$$

from $\mathcal{G}_q^u$ into the set of probability measures on $\mathcal{B}^*$ are compatible with $\sim_R$. As in subsection 4.2, this implies in particular that the power of the tests based on $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ is invariant with respect to the scale of the underlying error variables, so that for every fixed $\alpha, \beta \in (0,1)$ the relative efficiency $N_2(\alpha,\beta,F)/N_1(\alpha,\beta,F)$ of $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ with respect to $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ is invariant on the equivalence classes of $\sim_R$. Hence, for investigating the asymptotic behavior of the relative efficiency when the alternative approaches the null hypothesis we need to identify distribution functions deriving from the same scale family again. For this reason we will study the well-defined mappings

$$[F]_R \mapsto P_F \circ \widehat{\mathcal{W}}_{n,res}^{-1} \quad \text{and} \quad [F]_R \mapsto P_F \circ \widehat{\mathcal{V}}_{n,res}^{-1}$$

on the quotient set $\mathcal{G}_q^u/{\sim_R} := \big\{[F]_R \colon F \in \mathcal{G}_q^u\big\}$. Similar to before, the set

$$\widetilde{\mathcal{G}}_q^u := \big\{F(m_\tau(F)\,\cdot\,) \colon F \in \mathcal{G}_q^u\big\} = \big\{F \in \mathcal{G}_q^u \colon \tau \textstyle\int_{\mathbb{R}} |x|^\tau F(dx) = 1\big\}$$

is a complete set of equivalence class representatives of the relation $\sim_R$ on $\mathcal{G}_q^u$. Thus, for the investigation of the asymptotic behavior of $N_2(\alpha,\beta,F)/N_1(\alpha,\beta,F)$ we assume henceforth that the distribution function F of the variables $(e_i)_{i\in\mathbb{N}}$ is an element of $\widetilde{\mathcal{G}}_q^u$, and consider in the following the testing problem

$$\mathrm{H}_0 \colon F = F_\tau \quad \text{versus} \quad \mathrm{H}_1 \colon F \in \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}. \tag{6.38}$$

Note that the asymptotic level α tests for (6.17) based on $\widehat{\mathcal{W}}_{n,res}$ and $\widehat{\mathcal{V}}_{n,res}$ obviously are asymptotic level α tests for this testing problem as well.

We will show now that the sequences $\big(\widehat{\mathcal{W}}_{n,res}\big)_{n\geq 2}$ and $\big(\widehat{\mathcal{V}}_{n,res}\big)_{n\geq 2}$ are standard sequences. Recall that we have to verify conditions (BI), (BII) and (BIII) for this. But as the first two of these conditions only concern the distribution of $\widehat{\mathcal{W}}$ and $\widehat{\mathcal{V}}$, we have already checked them in subsection 4.2. Therefore it only remains to show that (BIII) holds for both sequences. For this, consider the functions

$$b\colon\ \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}\ \ni\ F\ \mapsto\ \Big(\int_{-\infty}^{\infty}\big(F(x)-F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2}\ \in\ (0,1]$$

and

$$b_{n,res}\colon\ \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}\ \ni\ F\ \mapsto\ \Big(\int_{-\infty}^{\infty}\big(F(x)-F(x,\hat\sigma_{n,res})\big)^2 F(dx,\hat\sigma_{n,res})\Big)^{1/2}\ \in\ (0,1],\quad n\geq 2.$$

Now let $F\in\widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}$ and $n\geq 2$. Then

$$\Big|\frac{\widehat{\mathcal{W}}_{n,res}}{\sqrt{n}}-b(F)\Big|\leq\Big|\frac{\widehat{\mathcal{W}}_{n,res}}{\sqrt{n}}-b_{n,res}(F)\Big|+\big|b_{n,res}(F)-b(F)\big|, \tag{6.39}$$

and by Minkowski's inequality we have

$$\Big|\frac{\widehat{\mathcal{W}}_{n,res}}{\sqrt{n}}-b_{n,res}(F)\Big|\leq\Big(\int_{-\infty}^{\infty}\big(F_{n,res}(x)-F(x)\big)^2F(dx,\hat\sigma_{n,res})\Big)^{1/2}\leq\|F_{n,res}-F\|_\infty, \tag{6.40}$$

where $\|F_{n,res}-F\|_\infty=o_{P_F}(1)$ because of (6.4). Analogously, we get

$$\begin{aligned}\Big|\frac{\widehat{\mathcal{V}}_{n,res}}{\sqrt{n}}-b(F)\Big|&\leq\Big|\frac{\widehat{\mathcal{V}}_{n,res}}{\sqrt{n}}-b_{n,res}(F)\Big|+\big|b_{n,res}(F)-b(F)\big|\\&\leq\|\widetilde{F}_{n,res}-F\|_\infty+\big|b_{n,res}(F)-b(F)\big|,\end{aligned} \tag{6.41}$$

and the functional central limit theorem (6.7) implies that $\|\widetilde{F}_{n,res}-F\|_\infty=o_{P_F}(1)$. Hence, to complete the verification of condition (BIII) for both sequences of test statistics, it remains to show that

$$\big|b_{n,res}(F)-b(F)\big|=o_{P_F}(1)\quad\text{as } n\to\infty \tag{6.42}$$

for every $F\in\widetilde{\mathcal{G}}_q^u\setminus\{F_\tau\}$. But since $b_{n,res}(F)-b(F)=T_F(\hat\sigma_{n,res})-T_F(1)$ with T_F as in Lemma 4.9, it follows by this lemma that (6.42) results from $\hat\sigma_{n,res}\to 1$ as $n\to\infty$ in P_F-probability for every $F\in\widetilde{\mathcal{G}}_q^u\setminus\{F_\tau\}$. In order to show this, we need some additional results.

The following lemma is a generalization of Lemma 5.4 (i) and Lemma 5.5 (i) to arbitrary powers.

Lemma 6.9
Let M be a nonempty set of continuous distribution functions that are centered and have finite second moments, and let $s\in(0,\infty)$. If $\sup_{F\in M}\int_{\mathbb{R}}|x|^sF(dx)<\infty$, then

$$\sup_{F\in M}\mathrm{E}_F\big(\|\mathbf{X}_i\|^s\big)=O(1)\ \text{as } i\to\infty.$$

Proof. Consider first $s\geq 1$. Let us investigate the sequence of L_s norms

$$\|\mathbf{X}_i\|_{L_s,F}=\mathrm{E}_F(\|\mathbf{X}_i\|^s)^{1/s},\quad i\geq 0.$$

The statement will follow if we show that the sequence $\big(\|\mathbf{X}_i\|_{L_s,F}\big)_{i\geq 0}$ is uniformly bounded in $F\in M$. By the representation (5.5) of $\mathbf{X}_i$ we get

$$\|\mathbf{X}_i\|_{L_s,F}\ \leq\ \|A^i\mathbf{X}_0\|_{L_s,F}+\sum_{j=1}^{i}\|A^{j-1}\mathbf{e}_{i+1-j}\|_{L_s,F}$$

$$\begin{aligned}
&\leq\ c_A \|A\|_A^i \|\mathbf{X}_0\|_{L_s,F} + \sum_{j=1}^{i} c_A \|A\|_A^{j-1} \|\mathbf{e}_{i+1-j}\|_{L_s,F} \\
&\leq\ c_A \sup_{F\in M} \|\mathbf{X}_0\|_{L_s,F} + c_A \Big(\sup_{F\in M} \int_{\mathbb{R}} |x|^s F(dx) \Big)^{1/s} \sum_{j=0}^{\infty} \|A\|_A^j
\end{aligned}$$

for every $i \geq 0$ and $F \in M$ using $\|A\|_A < 1$. It only remains to show that $\sup_{F\in M} \|\mathbf{X}_0\|_{L_s,F}$ is finite. Now recall that either $X_0 = \ldots = X_{1-p} = 0$, in which case there is nothing left to show, or $\mathbf{X}_0$ is the vector of starting values of the stationary $AR(p)$ process. Let us investigate the latter case. By the equivalence of norms in $\mathbb{R}^p$ there is a positive constant k_s such that $\|x\| \leq k_s \|x\|_s$ for all $x \in \mathbb{R}^p$, where

$$\|\cdot\|_s \colon \mathbb{R}^p \ni x \mapsto \Big(\sum_{i=1}^{p} |x_i|^s \Big)^{1/s} \in \mathbb{R}.$$

Hence,

$$\|\mathbf{X}_0\|_{L_s,F} = \mathrm{E}_F(\|\mathbf{X}_0\|^s)^{1/s} \leq k_s\, \mathrm{E}_F(\|\mathbf{X}_0\|_s^s)^{1/s} = k_s\, p^{1/s}\, \mathrm{E}_F(|X_0|^s)^{1/s},$$

using the strict stationarity of the process for the last equality. By the $MA(\infty)$-representation (5.3) we have

$$X_0 = \sum_{j=0}^{\infty} \psi_j e_{-j}, \tag{6.43}$$

where the series converges with probability one under every P_F, $F \in M$. We will show next that the series also converges to X_0 in s-th mean under every P_F. For this, set

$$Z_n := \sum_{j=0}^{n} \psi_j e_{-j}, \quad n \geq 0. \tag{6.44}$$

Obviously, $|Z_n| \leq \sum_{j=0}^{n} |\psi_j||e_{-j}|$. If $s > 1$, then by Hölder's inequality it is

$$|Z_n|^s \leq \Big(\sum_{j=0}^{n} |\psi_j||e_{-j}| \Big)^s \leq \Big(\sum_{j=0}^{n} |\psi_j|^r \Big)^{s/r} \sum_{j=0}^{n} |e_{-j}|^s$$

with $r = s/(s-1)$. Hence, for $s \geq 1$ we have $\mathrm{E}_F(|Z_n|^s) < \infty$ for all $n \geq 0$, since $\mathrm{E}_F(|e_1|^s) < \infty$ by assumption. Furthermore,

$$\sup_{m\geq n} \mathrm{E}_F\big(|Z_m - Z_n|^s\big)^{1/s} \leq \sup_{m>n} \sum_{j=n+1}^{m} |\psi_j|\, \mathrm{E}_F\big(|e_{-j}|^s\big)^{1/s} = \mathrm{E}_F\big(|e_1|^s\big)^{1/s} \sum_{j=n+1}^{\infty} |\psi_j| \underset{n\to\infty}{\longrightarrow} 0$$

since $\sum_{j=0}^{\infty} |\psi_j| < \infty$ because of (5.4), and so by the Cauchy criterion the sequence $(Z_n)_{n\geq 0}$ converges in s-th mean with respect to P_F towards a random variable Z with $\mathrm{E}_F(|Z|^s) < \infty$. Since $\lim_{n\to\infty} Z_n = X_0$ P_F-almost surely as well, it follows that $P_F(X_0 = Z) = 1$. This yields

$$\begin{aligned}
\mathrm{E}_F(|X_0|^s)^{1/s} = \lim_{n\to\infty} \mathrm{E}_F(|Z_n|^s)^{1/s} &\leq \lim_{n\to\infty} \sum_{j=0}^{n} |\psi_j|\, \mathrm{E}_F\big(|e_{-j}|^s\big)^{1/s} = \mathrm{E}_F\big(|e_1|^s\big)^{1/s} \sum_{j=0}^{\infty} |\psi_j| \\
&\leq \Big(\sup_{F\in M} \int_{\mathbb{R}} |x|^s F(dx) \Big)^{1/s} \sum_{j=0}^{\infty} |\psi_j| < \infty
\end{aligned}$$

for every $F \in M$, which concludes the proof of the case $s \geq 1$.

Now let $s \in (0,1)$. Again with (5.5) and $\|A\|_A < 1$ it is

$$\|\mathbf{X}_i\| \le c_A \|\mathbf{X}_0\| + \sum_{j=1}^{i} c_A \|A\|_A^{j-1} |e_{i+1-j}| \quad \forall\, i \ge 0,$$

and using that the function $h_s(x) = x^s$, $x \in [0,\infty)$, is subadditive, this yields

$$\begin{aligned}
\mathrm{E}_F(\|\mathbf{X}_i\|^s) &\le c_A^s \,\mathrm{E}_F(\|\mathbf{X}_0\|^s) + c_A^s \,\mathrm{E}_F(|e_1|^s) \sum_{j=1}^{i} \|A\|_A^{s(j-1)} \\
&\le c_A^s \sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|^s) + c_A^s \sup_{F\in M} \int_{\mathbb{R}} |x|^s F(dx) \sum_{j=0}^{\infty} \|A\|_A^{sj}
\end{aligned}$$

for every $i \ge 0$ and $F \in M$. Therefore it only remains to verify that $\sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_0\|^s)$ is finite. But if $X_0 = \ldots = X_{1-p} = 0$, there is nothing left to show. Hence, let us examine the case that $\mathbf{X}_0$ is the vector of starting values of the stationary $AR(p)$ process. Then it is for every $F \in M$

$$\mathrm{E}_F(\|\mathbf{X}_0\|^s) = \mathrm{E}_F\Big(\big(\sum_{j=1}^{p} X_{1-j}^2\big)^{s/2}\Big) \le \mathrm{E}_F\Big(\sum_{j=1}^{p} |X_{1-j}|^s\Big) = p\,\mathrm{E}_F(|X_0|^s),$$

using that $0 < s/2 < 1$ and the strict stationarity of the process. By the $MA(\infty)$-representation (6.43) of X_0 we get

$$\begin{aligned}
\mathrm{E}_F(|X_0|^s) &= \mathrm{E}_F(|\lim_{n\to\infty} Z_n|^s) = \mathrm{E}_F\big(\lim_{n\to\infty} |Z_n|^s\big) \le \mathrm{E}_F\Big(\lim_{n\to\infty} \sum_{j=0}^{n} |\psi_j|^s |e_{-j}|^s\Big) \\
&= \lim_{n\to\infty} \sum_{j=0}^{n} |\psi_j|^s \,\mathrm{E}_F(|e_1|^s) \le \sup_{F\in M} \int_{\mathbb{R}} |x|^s F(dx) \sum_{j=0}^{\infty} |\psi_j|^s < \infty
\end{aligned}$$

with Z_n as in (6.44), where we used the monotone convergence theorem and once more inequality (5.4). Hence, the statement follows for $s \in (0,1)$ as well. □

For the following investigations let us introduce the condition

$$\sup_{F\in M} \int_{\mathbb{R}} |x|^\tau F(dx) < \infty \tag{6.45}$$

for a set M of distribution functions. Obviously, (6.45) is just (3.9) if $\tau = 1$ and (3.6) if $\tau = 2$.

The next lemma provides a means to establish the convergence of $\hat{\sigma}_{n,res}$ to 1 in P_F-probability. Its statement is formulated to hold uniformly in a set $M \subset \widetilde{\mathcal{G}}_q^u$, and hence more general than needed here, because we will use the uniform result later on.

Lemma 6.10
Suppose the nonempty set $M \subset \widetilde{\mathcal{G}}_q^u$ is such that (6.45) *holds. Moreover, assume that the sequence of estimators $(\hat{\rho}_n)_{n\ge 2}$ for ρ fulfills* (5.8). *Then*

$$\frac{1}{n}\sum_{i=1}^{n} |\hat{e}_{ni}|^\tau - \frac{1}{n}\sum_{i=1}^{n} |e_i|^\tau = o_P^u(1) \quad \textit{in } M \textit{ as } n \to \infty.$$

Proof. First, note that by (5.8) it is $\|\hat{\rho}_n - \rho\| = O_P^u(n^{-1/2}) = o_P^u(1)$, whence it follows that

$$\|\hat{\rho}_n - \rho\|^s = o_P^u(1) \quad \forall\, s \in (0,\infty). \tag{6.46}$$

Consider now $\tau \in (0,1]$. Then $\big||x|^\tau - |y|^\tau\big| \le |x-y|^\tau$ for all $x, y \in \mathbb{R}$, so that

$$\Big|\frac{1}{n}\sum_{i=1}^n |\hat{e}_{ni}|^\tau - \frac{1}{n}\sum_{i=1}^n |e_i|^\tau\Big| \le \frac{1}{n}\sum_{i=1}^n \big||\hat{e}_{ni}|^\tau - |e_i|^\tau\big| \le \frac{1}{n}\sum_{i=1}^n \big|\hat{e}_{ni} - e_i\big|^\tau \le \|\hat{\rho}_n - \rho\|^\tau \frac{1}{n}\sum_{i=1}^n \|\mathbf{X}_{i-1}\|^\tau.$$

It is easily seen with Lemma 6.9 and Markov's inequality that $\frac{1}{n}\sum_{i=1}^n \|\mathbf{X}_{i-1}\|^\tau = O_P^u(1)$. In combination with (6.46) this yields the statement.

Now let $\tau > 1$. Note that the function $h_\tau(x) = x^\tau$, $x \in [0,\infty)$, is $\lfloor\tau\rfloor$-times continuously differentiable on $[0,\infty)$, where $\lfloor\tau\rfloor = \max\{n \in \mathbb{N} \colon n \le \tau\}$, and

$$h_\tau^{(k)}(x) = \tau(\tau-1)\cdot\ldots\cdot(\tau-k+1)x^{\tau-k}, \quad 1 \le k \le \lfloor\tau\rfloor.$$

By Taylor's theorem we get

$$\begin{aligned}
&h_\tau(y) - h_\tau(x)\\
&= \sum_{k=1}^{\lfloor\tau\rfloor-1} \frac{h_\tau^{(k)}(x)}{k!}(y-x)^k + \frac{1}{(\lfloor\tau\rfloor-1)!}\int_0^1 (1-u)^{\lfloor\tau\rfloor-1} h_\tau^{(\lfloor\tau\rfloor)}\big(x+(y-x)u\big)du\,(y-x)^{\lfloor\tau\rfloor}\\
&= \sum_{k=1}^{\lfloor\tau\rfloor} \frac{h_\tau^{(k)}(x)}{k!}(y-x)^k\\
&\quad + \frac{1}{(\lfloor\tau\rfloor-1)!}\int_0^1 (1-u)^{\lfloor\tau\rfloor-1}\Big(h_\tau^{(\lfloor\tau\rfloor)}\big(x+(y-x)u\big) - h_\tau^{(\lfloor\tau\rfloor)}(x)\Big)du\,(y-x)^{\lfloor\tau\rfloor}
\end{aligned}$$

for every $x, y \in [0,\infty)$. Hence,

$$\begin{aligned}
&\Big|\frac{1}{n}\sum_{i=1}^n \big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big)\Big| = \Big|\frac{1}{n}\sum_{i=1}^n \big(h_\tau(|\hat{e}_{ni}|) - h_\tau(|e_i|)\big)\Big|\\
&\le \Big|\frac{1}{n}\sum_{i=1}^n \sum_{k=1}^{\lfloor\tau\rfloor} \frac{h_\tau^{(k)}(|e_i|)}{k!}(|\hat{e}_{ni}| - |e_i|)^k\Big|\\
&\quad + \Big|\frac{1}{n}\sum_{i=1}^n \frac{1}{(\lfloor\tau\rfloor-1)!}\int_0^1 (1-u)^{\lfloor\tau\rfloor-1}\Big(h_\tau^{(\lfloor\tau\rfloor)}\big(|e_i| + (|\hat{e}_{ni}| - |e_i|)u\big) - h_\tau^{(\lfloor\tau\rfloor)}(|e_i|)\Big)du\\
&\qquad \cdot (|\hat{e}_{ni}| - |e_i|)^{\lfloor\tau\rfloor}\Big|\\
&=: \mathrm{I}_n + \mathrm{II}_n.
\end{aligned}$$

First, we investigate I_n. It is

$$\begin{aligned}
\mathrm{I}_n &\le \sum_{k=1}^{\lfloor\tau\rfloor} \frac{1}{k!}\frac{1}{n}\sum_{i=1}^n h_\tau^{(k)}(|e_i|)\cdot\big||\hat{e}_{ni}| - |e_i|\big|^k \le \sum_{k=1}^{\lfloor\tau\rfloor} \frac{1}{k!}\frac{1}{n}\sum_{i=1}^n h_\tau^{(k)}(|e_i|)\cdot\big|\hat{e}_{ni} - e_i\big|^k\\
&\le \sum_{k=1}^{\lfloor\tau\rfloor} c(k)\|\hat{\rho}_n - \rho\|^k \frac{1}{n}\sum_{i=1}^n |e_i|^{\tau-k}\cdot\|\mathbf{X}_{i-1}\|^k =: \sum_{k=1}^{\lfloor\tau\rfloor} A_{nk}
\end{aligned}$$

with $c(k) := \big(\tau(\tau-1)\cdot\ldots\cdot(\tau-k+1)\big)/k!$. Now condition (6.45) implies that $\sup_{F\in M}\int_{\mathbb{R}}|x|^r F(dx)$ is finite for all $r \in (0,\tau)$ by Lyapunov's inequality. Thus, it follows from Lemma 6.9 that for every $k = 1,\ldots,\lfloor\tau\rfloor$ there is a $\widetilde{K}(k) \in (0,\infty)$ with

$$\sup_{F\in M} \mathrm{E}_F(\|\mathbf{X}_{i-1}\|^k) \le \widetilde{K}(k) \quad \forall\, i \in \mathbb{N}.$$

Hence, for every $F \in M$, $C > 0$ and $1 \leq k \leq \lfloor \tau \rfloor$ it is

$$P_F\Big(\frac{1}{n}\sum_{i=1}^{n}|e_i|^{\tau-k}\cdot\|\mathbf{X}_{i-1}\|^k \geq C\Big) \leq \frac{1}{C}\,\widetilde{K}(k)\sup_{F\in M}\int_{\mathbb{R}}|x|^{\tau-k}F(dx) < \infty,$$

which implies that

$$\frac{1}{n}\sum_{i=1}^{n}|e_i|^{\tau-k}\cdot\|\mathbf{X}_{i-1}\|^k = O_P^u(1).$$

It follows from this and (6.46) that $A_{nk} = o_P^u(1)$ for all $k = 1,\dots,\lfloor\tau\rfloor$, so that $\mathrm{I}_n = o_P^u(1)$.

It remains to study II_n. Observe that for integer-valued τ it is $h_\tau^{(\lfloor\tau\rfloor)}(x) = \tau!$, so that $\mathrm{II}_n = 0$ in this case. If τ is not integer-valued, then $\tau - \lfloor\tau\rfloor \in (0,1)$, and using again $\big||x|^{\tau-\lfloor\tau\rfloor} - |y|^{\tau-\lfloor\tau\rfloor}\big| \leq |y-x|^{\tau-\lfloor\tau\rfloor}$ for all $x, y \in \mathbb{R}$, we get

$$\begin{aligned}
\mathrm{II}_n &\leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(\lfloor\tau\rfloor-1)!}\int_0^1(1-u)^{\lfloor\tau\rfloor-1}\Big|h_\tau^{(\lfloor\tau\rfloor)}\big(|e_i|+(|\hat{e}_{ni}|-|e_i|)u\big) - h_\tau^{(\lfloor\tau\rfloor)}(|e_i|)\Big|du \\
&\quad\cdot\big|\hat{e}_{ni}-e_i\big|^{\lfloor\tau\rfloor} \\
&= c(\lfloor\tau\rfloor)\cdot\lfloor\tau\rfloor\cdot\frac{1}{n}\sum_{i=1}^{n}\int_0^1(1-u)^{\lfloor\tau\rfloor-1}\Big|\big(|e_i|+(|\hat{e}_{ni}|-|e_i|)u\big)^{\tau-\lfloor\tau\rfloor} - |e_i|^{\tau-\lfloor\tau\rfloor}\Big|du\,\big|\hat{e}_{ni}-e_i\big|^{\lfloor\tau\rfloor} \\
&\leq c(\lfloor\tau\rfloor)\cdot\lfloor\tau\rfloor\cdot\frac{1}{n}\sum_{i=1}^{n}\int_0^1(1-u)^{\lfloor\tau\rfloor-1}\big|(|\hat{e}_{ni}|-|e_i|)u\big|^{\tau-\lfloor\tau\rfloor}du\,\big|\hat{e}_{ni}-e_i\big|^{\lfloor\tau\rfloor} \\
&\leq c(\lfloor\tau\rfloor)\cdot\lfloor\tau\rfloor\cdot\frac{1}{n}\sum_{i=1}^{n}\big|\hat{e}_{ni}-e_i\big|^{\tau} \\
&\leq c(\lfloor\tau\rfloor)\cdot\lfloor\tau\rfloor\cdot\|\hat{\rho}_n-\rho\|^\tau\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|^\tau = o_P^u(1)O_P^u(1) = o_P^u(1)
\end{aligned}$$

because of (6.46) and $\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\|^\tau = O_P^u(1)$, as is easily shown using Markov's inequality and the fact that $\sup_{F\in M}\mathrm{E}_F(\|\mathbf{X}_i\|^\tau) = O(1)$ as $i\to\infty$ by Lemma 6.9. □

We are now able to conclude the verification of (BIII) for the two sequences $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$. Note that by using the foregoing lemma with $M = \{F\}$ and by the strong law of large numbers it is

$$\Big|\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau - \mathrm{E}_F(|e_1|^\tau)\Big| \leq \Big|\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau - \frac{1}{n}\sum_{i=1}^{n}|e_i|^\tau\Big| + \Big|\frac{1}{n}\sum_{i=1}^{n}|e_i|^\tau - \mathrm{E}_F(|e_1|^\tau)\Big| = o_{P_F}(1)$$

for every $F \in \widetilde{\mathcal{G}}_q^u$. Thus,

$$\hat{\sigma}_{n,res}(\hat{e}_{n1},\dots,\hat{e}_{nn}) = \Big(\tau\frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{ni}|^\tau\Big)^{1/\tau} \quad\underset{n}{\longrightarrow}\quad \Big(\tau\int_{\mathbb{R}}|x|^\tau F(dx)\Big)^{1/\tau} = m_\tau(F) = 1$$

in P_F-probability for every $F \in \widetilde{\mathcal{G}}_q^u$, which implies (6.42).

To recapitulate, we have shown the following:

Proposition 6.11
The sequences $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ are standard sequences, and their approximate slopes are $b(F)^2/\tilde{\lambda}_1$ and $b(F)^2/\tilde{\lambda}_1^*$, respectively, for every $F \in \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}$, where $\tilde{\lambda}_1$ and $\tilde{\lambda}_1^*$ are as in Proposition 4.10. Hence, the approximate Bahadur ARE of $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ relative to $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ is $\tilde{\lambda}_1^*/\tilde{\lambda}_1$.

Observe again that the approximate Bahadur ARE of $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ relative to $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ is independent of the alternative distribution $F \in \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}$, as the function b is the same for both sequences. Moreover, it is equal to the approximate Bahadur ARE of $(\widehat{\mathcal{W}}_n)_{n\geq 2}$ relative to $(\widehat{\mathcal{V}}_n)_{n\geq 2}$ when testing the same null hypothesis in the case of independent and identically distributed data, cf. Proposition 4.10. As mentioned in subsection 4.2, it is $\tilde{\lambda}_1^* < \tilde{\lambda}_1 = 1/\pi^2$ for every $\tau \in (0,\infty)$, so that $\tilde{\lambda}_1^*/\tilde{\lambda}_1$ is always less than one. The values of $\tilde{\lambda}_1^*$ and $\tilde{\lambda}_1^*/\tilde{\lambda}_1$ for $\tau = 1$ and $\tau = 2$ are given in Table 2 on page 52.

Our next goal is to show that the sequences $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ also meet Wieand's condition (WIII). For this, we first need to examine the (uniform) asymptotic behavior of the scale estimator $\hat{\sigma}_{n,res}$ more closely, which we will do on the following pages.

Recall that in subsection 4.2 we used the condition

$$\sup_{F\in M} \int_{\mathbb{R}} |x|^{2\tau} F(dx) < \infty, \tag{4.36}$$

where M is a set of distribution functions. It is evident that condition (4.36) implies (6.45) for every $0 < \tau < \infty$ (given the set M is the same in both cases, of course).

Lemma 6.12
Assume that the nonempty set $M \subset \widetilde{\mathcal{G}}_q^u$ satisfies (4.36)*. Suppose further that the sequence of estimators $(\hat{\rho}_n)_{n\geq 2}$ for ρ fulfills* (5.8)*. Then $\hat{\sigma}_{n,res} - 1 = o_P^u(1)$ in M as $n \to \infty$.*

Proof. Obviously, it is sufficient to show

$$\sup_{F\in M} P_F\big(|\hat{\sigma}_{n,res} - 1| \geq \epsilon\big) \underset{n\to\infty}{\longrightarrow} 0$$

for every $\epsilon \in (0,1)$. But for every $F \in M$ and $\epsilon \in (0,1)$ it is

$$P_F\big(|\hat{\sigma}_{n,res} - 1| \geq \epsilon\big) \leq P_F\big(|\hat{\sigma}_{n,res}^\tau - 1| \geq K_\tau \epsilon\big)$$

using inequality (4.37), and

$$\big|\hat{\sigma}_{n,res}^\tau - 1\big| \leq \big|\hat{\sigma}_{n,res}^\tau - \hat{\sigma}_n^\tau\big| + \big|\hat{\sigma}_n^\tau - 1\big|.$$

Now note that the conditions of Lemma 6.10 are satisfied, so that

$$\big|\hat{\sigma}_{n,res}^\tau - \hat{\sigma}_n^\tau\big| = \tau\Big|\frac{1}{n}\sum_{i=1}^n \big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big)\Big| = o_P^u(1)$$

by this lemma. Moreover, with Markov's inequality and Lemma 4.12 (i) it is easily seen that $|\hat{\sigma}_n^\tau - 1| = o_P^u(1)$ as well, which concludes the proof. □

Let us continue our investigation of $\hat{\sigma}_{n,res}$ with the following lemma.

Lemma 6.13
Let $\emptyset \neq M \subset \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}$. If $\tau \geq 1$, suppose that (6.45) *holds, and if $\tau \in (0,1)$, assume that* (3.9) *and $\sup_{F\in M}\int_{\mathbb{R}^*} |x|^{\tau-1} F(dx) < \infty$ hold. Also assume that $(\hat{\rho}_n)_{n\geq 2}$ satisfies* (5.8)*. Then for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C(\epsilon,\delta) > 0$ such that*

$$P_F\Big(\Big|\frac{1}{n}\sum_{i=1}^n \big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big)\Big| \geq \epsilon\, b(F)\Big) < \delta$$

for every $F \in M$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C(\epsilon,\delta)/b(F)$.

Proof. Let $\epsilon > 0$, $\delta \in (0,1)$ and $F \in M$. Since $P_F(e_i \neq 0 \text{ for all } i \in \mathbb{N}) = 1$ because of the continuity of F, we will assume again that $|e_i| > 0$ for all $i \in \mathbb{N}$.

Observe that it follows from (5.8), i.e., from

$$\lim_{a\to\infty} \limsup_{n\to\infty} \sup_{F\in M} P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| \geq a\big) = 0,$$

that there are $a(\delta) \in (0,\infty)$ and $n_0(\delta) \in \mathbb{N}$ with

$$\sup_{F\in M} P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| \geq a(\delta)\big) < \frac{\delta}{2}$$

for all $n \geq n_0(\delta)$.

Consider now $\tau \in (0,1]$. Then $|y^\tau - x^\tau| \leq x^{\tau-1}|y-x|$ for $x > 0$, $y \geq 0$. By using this inequality we get

$$\begin{aligned}
\Big|\frac{1}{n}\sum_{i=1}^{n}\big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big)\Big| &\leq \frac{1}{n}\sum_{i=1}^{n}\big||\hat{e}_{ni}|^\tau - |e_i|^\tau\big| \leq \frac{1}{n}\sum_{i=1}^{n}|e_i|^{\tau-1}\big||\hat{e}_{ni}| - |e_i|\big| \\
&\leq \frac{1}{n}\sum_{i=1}^{n}|e_i|^{\tau-1}|\hat{e}_{ni} - e_i| \leq \|\hat{\rho}_n - \rho\|\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\||e_i|^{\tau-1}.
\end{aligned}$$

If $\tau = 1$, then $|e_1|^{\tau-1} = 1$, so that $\mathrm{E}_F(|e_1|^{\tau-1}) = 1$ for all $F \in M$ in this case, and for $\tau \in (0,1)$ it is $\sup_{F\in M}\mathrm{E}_F(|e_1|^{\tau-1}) < \infty$ by assumption. Moreover, since (3.9) holds it follows from Lemma 5.4 (i) that there is a constant $\widetilde{K} \in (0,\infty)$ with

$$\sup_{F\in M}\mathrm{E}_F(\|\mathbf{X}_{i-1}\|) \leq \widetilde{K} \quad \forall\, i \in \mathbb{N}.$$

Hence,

$$\begin{aligned}
&P_F\Big(\Big|\frac{1}{n}\sum_{i=1}^{n}\big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big)\Big| \geq \epsilon\, b(F)\Big) \\
&\leq P_F\Big(a(\delta)\frac{1}{n^{3/2}}\sum_{i=1}^{n}\|\mathbf{X}_{i-1}\||e_i|^{\tau-1} \geq \epsilon\, b(F)\Big) + P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| \geq a(\delta)\big) \\
&\leq \frac{a(\delta)}{\epsilon\, b(F)}\mathrm{E}_F(|e_1|^{\tau-1})\frac{1}{n^{3/2}}\sum_{i=1}^{n}\mathrm{E}_F(\|\mathbf{X}_{i-1}\|) + \sup_{F\in M}P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| \geq a(\delta)\big) \\
&\leq \frac{a(\delta)}{\epsilon\, b(F)}\frac{\widetilde{K}}{\sqrt{n}}\sup_{F\in M}\mathrm{E}_F(|e_1|^{\tau-1}) + \sup_{F\in M}P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| \geq a(\delta)\big) \\
&< \frac{\delta}{2} + \frac{\delta}{2} = \delta
\end{aligned}$$

for all $F \in M$ and $n \in \mathbb{N}$ such that

$$\sqrt{n} > \frac{1}{b(F)}\max\Big(\frac{2a(\delta)\widetilde{K}\sup_{F\in M}\mathrm{E}_F(|e_1|^{\tau-1})}{\epsilon\,\delta}, \sqrt{n_0(\delta)}\Big).$$

It remains to investigate the case $\tau > 1$. Recall that it has been shown in the proof of Lemma 6.10 that

$$\Big|\frac{1}{n}\sum_{i=1}^{n}\big(|\hat{e}_{ni}|^\tau - |e_i|^\tau\big)\Big| \leq \mathrm{I}_n + \mathrm{II}_n,$$

where

$$\mathrm{I}_n \le \sum_{k=1}^{\lfloor\tau\rfloor} c(k)\|\hat\rho_n-\rho\|^k \frac{1}{n}\sum_{i=1}^n |e_i|^{\tau-k}\cdot\|\mathbf{X}_{i-1}\|^k =: \sum_{k=1}^{\lfloor\tau\rfloor} A_{nk}$$

with $c(k) = \big(\tau(\tau-1)\cdot\ldots\cdot(\tau-k+1)\big)/k!$ for $1 \le k \le \lfloor\tau\rfloor$, and

$$\mathrm{II}_n \begin{cases} = 0 & \text{if } \tau\in\mathbb{N},\\ \le c(\lfloor\tau\rfloor)\cdot\lfloor\tau\rfloor\cdot\|\hat\rho_n-\rho\|^\tau\frac{1}{n}\sum_{i=1}^n\|\mathbf{X}_{i-1}\|^\tau & \text{if } \tau\notin\mathbb{N}.\end{cases}$$

Now by Lyapunov's inequality condition (6.45) implies that $\sup_{F\in M}\int_{\mathbb{R}}|x|^r F(dx) < \infty$ for all $r\in(0,\tau)$. Hence, Lemma 6.9 ensures the existence of positive constants $\widetilde{K}(k)$, $k=1,\ldots,\lfloor\tau\rfloor$, such that

$$\sup_{F\in M}\mathrm{E}_F(\|\mathbf{X}_{i-1}\|^k) \le \widetilde{K}(k) \quad \forall\, i\in\mathbb{N}.$$

Thus we have for every $k=1,\ldots,\lfloor\tau\rfloor$

$$\begin{aligned} P_F\big(A_{nk}\ge\epsilon\, b(F)\big) &\le P_F\Big(\frac{c(k)a(\delta)^k}{n^{1+k/2}}\sum_{i=1}^n|e_i|^{\tau-k}\cdot\|\mathbf{X}_{i-1}\|^k \ge \epsilon\, b(F)\Big) + P_F\big(\sqrt{n}\|\hat\rho_n-\rho\|\ge a(\delta)\big)\\ &\le \frac{c(k)a(\delta)^k}{\epsilon\, b(F)}\,\mathrm{E}_F(|e_1|^{\tau-k})\frac{1}{n^{1+k/2}}\sum_{i=1}^n \mathrm{E}_F(\|\mathbf{X}_{i-1}\|^k)\\ &\quad + \sup_{F\in M} P_F\big(\sqrt{n}\|\hat\rho_n-\rho\|\ge a(\delta)\big)\\ &\le \frac{c(k)a(\delta)^k\widetilde{K}(k)}{\epsilon\, b(F)n^{k/2}}\sup_{F\in M}\int_{\mathbb{R}}|x|^{\tau-k}F(dx) + \sup_{F\in M}P_F\big(\sqrt{n}\|\hat\rho_n-\rho\|\ge a(\delta)\big)\\ &< \frac{\delta}{2}+\frac{\delta}{2} = \delta \end{aligned}$$

for all $F\in M$ and $n\in\mathbb{N}$ such that

$$\sqrt{n} > \frac{1}{b(F)}\max\Big(\frac{2c(k)a(\delta)^k\widetilde{K}(k)\sup_{F\in M}\int_{\mathbb{R}}|x|^{\tau-k}F(dx)}{\epsilon\,\delta}, \sqrt{n_0(\delta)}\Big).$$

It follows from this that there is a positive constant $C(\epsilon,\delta)$ so that

$$P_F\big(\mathrm{I}_n\ge\epsilon\, b(F)\big) < \delta$$

for all $F\in M$ and all natural numbers n with $\sqrt{n} > C(\epsilon,\delta)/b(F)$.

Let us finally study II_n. As already mentioned, $\mathrm{II}_n = 0$ if $\tau\in\mathbb{N}$, so there is nothing to prove in this case. Hence, let $\tau\notin\mathbb{N}$. By Lemma 6.9 there is a $\widetilde{K}(\tau)\in(0,\infty)$ with

$$\sup_{F\in M}\mathrm{E}_F(\|\mathbf{X}_{i-1}\|^\tau) \le \widetilde{K}(\tau) \quad \forall\, i\in\mathbb{N}.$$

Thus

$$\begin{aligned} P_F\big(\mathrm{II}_n\ge\epsilon\, b(F)\big) &\le P_F\Big(\frac{c(\lfloor\tau\rfloor)\lfloor\tau\rfloor a(\delta)^\tau}{n^{1+\tau/2}}\sum_{i=1}^n\|\mathbf{X}_{i-1}\|^\tau\ge\epsilon\, b(F)\Big) + P_F\big(\sqrt{n}\|\hat\rho_n-\rho\|\ge a(\delta)\big)\\ &\le \frac{c(\lfloor\tau\rfloor)\lfloor\tau\rfloor a(\delta)^\tau}{\epsilon\, b(F)n^{\tau/2}}\frac{1}{n}\sum_{i=1}^n\mathrm{E}_F(\|\mathbf{X}_{i-1}\|^\tau) + \sup_{F\in M}P_F\big(\sqrt{n}\|\hat\rho_n-\rho\|\ge a(\delta)\big) \end{aligned}$$

$$\leq \frac{c(\lfloor\tau\rfloor)\lfloor\tau\rfloor a(\delta)^{\tau}\widetilde{K}(\tau)}{\epsilon\, b(F) n^{\tau/2}} + \sup_{F\in M} P_F\big(\sqrt{n}\|\hat{\rho}_n - \rho\| \geq a(\delta)\big)$$

$$< \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

for all $F \in M$ and $n \in \mathbb{N}$ such that

$$\sqrt{n} > \frac{1}{b(F)} \max\Big(\frac{2c(\lfloor\tau\rfloor)\lfloor\tau\rfloor a(\delta)^{\tau}\widetilde{K}(\tau)}{\epsilon\,\delta}, \sqrt{n_0(\delta)}\Big). \qquad \square$$

The next result is an analog of Lemma 4.13 for $\hat{\sigma}_{n,res}$.

Lemma 6.14
Let $\emptyset \neq M \subset \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}$. If $\tau \geq 1/2$, suppose that (4.36) holds, and if $\tau \in (0, 1/2)$ suppose that condition (3.9) is satisfied. For $\tau \in (0,1)$ assume further that $\sup_{F\in M}\int_{\mathbb{R}^*} |x|^{\tau-1} F(dx) < \infty$. Also, let the sequence of estimators $(\hat{\rho}_n)_{n\geq 2}$ for ρ be such that (5.8) is satisfied. Then for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C(\epsilon,\delta) > 0$ such that

$$P_F\big(\big|\hat{\sigma}_{n,res} - 1\big| \geq \epsilon\, b(F)\big) < \delta$$

for every $F \in M$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C(\epsilon,\delta)/b(F)$.

Proof. First, note that we may as well assume that $\epsilon \in (0,1)$. Then $\epsilon\, b(F) \in (0,1)$, as $b(F) \in (0,1]$. Thus, by inequality (4.37) we have

$$P_F\big(\big|\hat{\sigma}_{n,res} - 1\big| \geq \epsilon\, b(F)\big) \leq P_F\big(\big|\hat{\sigma}^{\tau}_{n,res} - 1\big| \geq K_\tau \epsilon\, b(F)\big)$$

$$\leq P_F\Big(\Big|\frac{1}{n}\sum_{i=1}^{n}\big(|\hat{e}_{ni}|^{\tau} - |e_i|^{\tau}\big)\Big| \geq \frac{K_\tau \epsilon\, b(F)}{2\tau}\Big) + P_F\Big(\big|\hat{\sigma}^{\tau}_{n} - 1\big| \geq \frac{K_\tau \epsilon\, b(F)}{2}\Big) =: A_{n,F} + B_{n,F}.$$

Now let $\delta \in (0,1)$. As for every value of τ the conditions of Lemma 6.13 are satisfied, it follows from this very lemma that there is a $C_1(\epsilon,\delta) \in (0,\infty)$ so that $A_{n,F} < \delta/2$ for all $F \in M$ and $n \geq 2$ with $\sqrt{n} > C_1(\epsilon,\delta)/b(F)$. Moreover, we have

$$B_{n,F} \leq \frac{4}{K_\tau^2 \epsilon^2 b(F)^2} \sup_{F\in M} \mathrm{E}_F\big(|\hat{\sigma}^{\tau}_{n} - 1|^2\big)$$

using Markov's inequality, and since it follows from the assumptions that (4.36) holds as well for every $\tau \in (0,\infty)$, part (i) of Lemma 4.12 ensures the existence of a $\widetilde{K} \in (0,\infty)$ such that

$$\sup_{F\in M} \mathrm{E}_F\big(|\hat{\sigma}^{\tau}_{n} - 1|^2\big) \leq \frac{1}{n}\widetilde{K}$$

for every $n \geq 2$. Now set $C_2(\epsilon,\delta) := \big(8\widetilde{K}/(K_\tau^2\epsilon^2\delta)\big)^{1/2}$. Then for every $F \in M$ it is

$$P_F\big(\big|\hat{\sigma}_{n,res} - 1\big| \geq \epsilon\, b(F)\big) < \delta$$

for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > \max\big(C_1(\epsilon,\delta), C_2(\epsilon,\delta)\big)/b(F)$. $\square$

We have now gathered all results necessary to prove an analog of Proposition 4.14 for $b_{n,res}$.

Proposition 6.15
Let $\emptyset \neq M \subset \widetilde{\mathcal{G}}_q^u \setminus \{F_\tau\}$. If the assumptions of Lemma 6.14 are satisfied, then for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C(\epsilon,\delta) > 0$ such that

$$P_F\big(\,\big|b_{n,res}(F) - b(F)\big| \geq \epsilon\, b(F)\,\big) \; < \; \delta$$

for all $F \in M$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C(\epsilon,\delta)/b(F)$.

Proof. First, note that under the above assumptions the set M also satisfies the requirements of Lemma 6.12, whence it follows that $\sup_{F\in M} P_F(\hat{\sigma}_{n,res} = 0) \to 0$ as $n \to \infty$ because the set $\{\hat{\sigma}_{n,res} = 0\}$ is a subset of $\{|\hat{\sigma}_{n,res} - 1| \geq \epsilon\}$ for every $\epsilon \in (0,1)$. Hence, we can and will assume that $\hat{\sigma}_{n,res} > 0$. Now using the results of Lemma 6.12 and Lemma 6.14, the above statement is proven analogously to Proposition 4.14 by simply replacing $\hat{\sigma}_n$ by $\hat{\sigma}_{n,res}$ and b_n by $b_{n,res}$ in the proof. □

Now recall that our aim is to show that the sequences $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ satisfy Wieand's condition (WIII). Before we can proceed with this, we have to adjust the set of possible error distribution functions, because we need the alternative distributions to be sufficiently smooth again.

Hence, let us consider again the set of distribution functions $\mathcal{G}_{q,\gamma,w}$ defined in (6.9). In this subsection, we will always use weight functions w of the form $w(x) = |x|^s$, $x \in \mathbb{R}$, for some $s > 0$. Unlike the weight functions considered in the previous subsection, this w is obviously not bounded away from zero. To stress the dependency of $\|\cdot\|_{w,\infty}$ on s, we will denote it by $\|\cdot\|_{s,\infty}$, i.e.,

$$\|h\|_{s,\infty} = \sup_{x\in\mathbb{R}} |x|^s |h(x)| \in [0,\infty]$$

for any function $h\colon \mathbb{R} \to \mathbb{R}$, and for the same reason we will denote $\mathcal{G}_{q,\gamma,w}$ henceforth by $\mathcal{G}_{q,\gamma,s}$, so that

$$\mathcal{G}_{q,\gamma,s} = \Big\{F\colon F \text{ is a distribution function having a Lebesgue density } f \text{ that satisfies} \\ [f]_\gamma + \|f\|_{s,\infty} + \int_{\mathbb{R}} |x|^q f(x)dx < \infty \text{ and } \int_{\mathbb{R}} x f(x) dx = 0\Big\}$$

for every $q \in (2,\infty)$, $\gamma \in (0,1]$ and $s \in (0,\infty)$. Observe here that for every $F \in \mathcal{G}_{q,\gamma,s}$ its density f is bounded, as it is uniformly continuous because of $[f]_\gamma < \infty$. Moreover, note that if $F \in \mathcal{G}_{q,\gamma,s}$, then the whole scale family $\{F(\cdot/\sigma)\colon \sigma \in (0,\infty)\}$ generated by F is contained in $\mathcal{G}_{q,\gamma,s}$. To verify this, observe that every $F(\cdot/\sigma)$ is centered again with finite q-th moment, and its density $f_\sigma := \sigma^{-1} f(\cdot/\sigma)$ satisfies $[f_\sigma]_\gamma = \sigma^{-(\gamma+1)}[f]_\gamma < \infty$ and $\|f_\sigma\|_{s,\infty} = \sigma^{s-1}\|f\|_{s,\infty} < \infty$.

Next, we will equip $\mathcal{G}_{q,\gamma,s}$ with a suitable metric. On $\mathcal{G}_{q,\gamma,s} \times \mathcal{G}_{q,\gamma,s}$ we consider the function $d_{q,\gamma,s}$ defined by

$$d_{q,\gamma,s}(F,G) := [f-g]_\gamma + \|f-g\|_{s,\infty} + \|f-g\|_\infty + \Big|\int_{\mathbb{R}} |x|^q F(dx) - \int_{\mathbb{R}} |x|^q G(dx)\Big|,$$

where f and g are the densities of F and G. It is easy to see that $d_{q,\gamma,s}$ is a metric on $\mathcal{G}_{q,\gamma,s}$, so that $(\mathcal{G}_{q,\gamma,s}, d_{q,\gamma,s})$ is a metric space.

We will continue by listing some properties of the density f_τ of F_τ, where τ is fixed in $(0,\infty)$.

First, note that $\|f_\tau\|_{s,\infty} < \infty$ for every $s > 0$, because the function $\mathbb{R} \ni x \mapsto |x|^s f_\tau(x) \in [0,\infty)$ is continuous and $\lim_{|x|\to\infty} |x|^s f_\tau(x) = 0$, hence it is bounded.

The next lemma sheds light on the Hölder continuity of f_τ.

Lemma 6.16
If $\tau \geq 1$, then $[f_\tau]_\gamma < \infty$ for every $\gamma \in (0,1]$. If $\tau \in (0,1)$, then $[f_\tau]_\gamma < \infty$ if and only if $\gamma \in (0,\tau]$.

Proof. Recall that for every $\tau > 0$ the density f_τ is differentiable for all $x \in \mathbb{R}^*$ with

$$f'_\tau(x) = f_\tau(x)\tau|x|^{\tau-1} \cdot \big(-\operatorname{sgn}(x)\big).$$

For $\tau > 1$, f_τ is obviously differentiable in 0 as well with $f_\tau'(0) = 0$. Now by the fundamental theorem of calculus it is

$$f_\tau(y) - f_\tau(x) = \int_x^y f_\tau'(u)\,du$$

for every $-\infty < x < y < \infty$ and every $\tau \in (0,\infty)$. Hence, for every $x, y \in \mathbb{R}$ we have

$$|f_\tau(y) - f_\tau(x)| \le \Big| \int_x^y |f_\tau'(u)|\,du \Big| = \int_{-\infty}^{\infty} |f_\tau'(u)| \cdot 1_{[x\wedge y, x\vee y]}(u)\,du. \tag{6.47}$$

Consider now $\tau \ge 1$. Then (6.47) yields

$$|f_\tau(y) - f_\tau(x)| \le \|f_\tau'\|_\infty \cdot |y-x| \quad \forall\, x, y \in \mathbb{R}.$$

Since f_τ' is bounded (for $\tau = 1$, set $f_1'(0) := 0$ for example), this shows that $[f_\tau]_1 < \infty$, i.e., f_τ is Lipschitz continuous for every $\tau \ge 1$. As already mentioned in subsection 6.1, this implies that $[f_\tau]_\gamma < \infty$ for every $\gamma \in (0,1)$ as well.

Next, we examine the case $\tau \in (0,1)$. We will show first that $[f_\tau]_\tau < \infty$, whence $[f_\tau]_\gamma < \infty$ for every $\gamma \in (0,\tau)$ follows. For this, observe that the function $g_1\colon \mathbb{R} \ni x \mapsto |x|^\tau \in [0,\infty)$ is Hölder continuous with exponent τ, because $\big||x|^\tau - |y|^\tau\big| \le |x-y|^\tau$ for all $x, y \in \mathbb{R}$. Moreover, the function $g_2\colon [0,\infty) \ni x \mapsto \exp(-x) \in (0,\infty)$ satisfies $|g_2(x) - g_2(y)| \le |x-y|$ for every $x, y \in [0,\infty)$ by the mean value theorem, using that $|g_2'| \le 1$. Combining all this, we see that for every $x, y \in \mathbb{R}$ the function $g := g_2 \circ g_1$ satisfies

$$|g(x) - g(y)| \le \big|g_1(x) - g_1(y)\big| \le |x-y|^\tau.$$

Thus, we have verified that $[g]_\tau < \infty$. But since $f_\tau = C(\tau)g$ for some positive norming constant $C(\tau)$, the statement follows.
Now let $\gamma \in (\tau, 1]$. Then f_τ is not γ-Hölder continuous, for if it were, there would be a constant $K \in (0,\infty)$ such that

$$\frac{|f_\tau(x) - f_\tau(y)|}{|x-y|^\gamma} \le K \quad \forall\, x \ne y \in \mathbb{R}. \tag{6.48}$$

But for every $x > 0$ we have

$$\frac{|f_\tau(0) - f_\tau(x)|}{|0-x|^\gamma} = \frac{C(\tau)\big(1 - e^{-x^\tau}\big)}{x^\gamma},$$

and by l'Hospital's rule we see that

$$\lim_{x \searrow 0} \frac{1 - e^{-x^\tau}}{x^\gamma} = \lim_{x \searrow 0} \frac{\tau x^{\tau-1} e^{-x^\tau}}{\gamma x^{\gamma-1}} = \frac{\tau}{\gamma} \lim_{x \searrow 0} x^{\tau-\gamma} e^{-x^\tau} = \infty$$

since $\tau - \gamma < 0$, which contradicts (6.48). □

Since F_τ is centered and has moments of all order, it follows that it is an element of the set $\mathcal{G}_{q,\gamma,s}$ for every $q > 2$, $s > 0$ and $\gamma \in (0, \tau \wedge 1]$. The foregoing lemma implies moreover that $F_{\tau^*} \in \mathcal{G}_{q,\gamma,s}$ for all $\tau^* \ge \gamma$.

For the rest of this subsection, we assume that $q = 2\tau$ if $\tau > 1$ and $q > 2$ otherwise. Additionally, we assume that $s > 0$ and $\gamma \in (0, \tau \wedge 1]$.

Suppose now that the random variables $(e_i)_{i\in\mathbb{N}}$ are independent and identically distributed according to an unknown distribution function $F \in \mathcal{G}_{q,\gamma,s}$ and that we want to test

$$\mathrm{H}_0\colon F \in \mathcal{F}_\tau = \big\{F_\tau(\cdot/\sigma)\colon \sigma \in (0,\infty)\big\} \quad \text{versus} \quad \mathrm{H}_1\colon F \in \mathcal{G}_{q,\gamma,s} \setminus \mathcal{F}_\tau. \tag{6.49}$$

In comparison to the initial testing problem (6.17), the set of distribution functions considered under H_1 is reduced here to the subset $\mathcal{G}_{q,\gamma,s} \setminus \mathcal{F}_\tau$ of $\mathcal{G}_q^u \setminus \mathcal{F}_\tau$.

As before, we will test the composite null hypothesis in (6.49) with the asymptotic level α tests based on the Cramér-von Mises statistics $\widehat{\mathcal{W}}_{n,res}$ and $\widehat{\mathcal{V}}_{n,res}$. For the following investigations, we have to specify the sequence of estimators for the autoregressive parameter ρ. Henceforth, we will use the sequence of least squares estimators $(\hat{\rho}_n^{LS})_{n\geq 2}$ to estimate ρ. Recall that this estimator fulfills the assumptions (6.18) and (6.20), where the latter was shown in Remark 6.7.

Proceeding analogously to before, we note that since $\mathcal{G}_{q,\gamma,s} \subset \mathcal{G}_q^u$, the relation $\sim_R$ in (6.37) is also an equivalence relation on $\mathcal{G}_{q,\gamma,s}$, and for every $F \in \mathcal{G}_{q,\gamma,s}$ its equivalence class under this relation is just the scale family generated by it, i.e.,

$$[F]_R := \{G \in \mathcal{G}_{q,\gamma,s} \colon G \sim_R F\} = \big\{F\big(\cdot/c\big) \colon c \in (0,\infty)\big\}.$$

Since the least squares estimator $\hat{\rho}_n^{LS}$ is scale invariant, it follows as before that the mappings

$$F \mapsto P_F \circ \widehat{\mathcal{W}}_{n,res}^{-1} \quad \text{and} \quad F \mapsto P_F \circ \widehat{\mathcal{V}}_{n,res}^{-1}$$

from $\mathcal{G}_{q,\gamma,s}$ into the set of probability measures on $\mathcal{B}^*$ are compatible with $\sim_R$, so that in order to examine the asymptotic behavior of the relative efficiency $N_2(\alpha,\beta,F)/N_1(\alpha,\beta,F)$ of $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ with respect to $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ when the alternative approaches the null hypothesis we have to identify distribution functions deriving from the same scale family again. Because of this, we consider the mappings

$$[F]_R \mapsto P_F \circ \widehat{\mathcal{W}}_{n,res}^{-1} \quad \text{and} \quad [F]_R \mapsto P_F \circ \widehat{\mathcal{V}}_{n,res}^{-1}$$

on the quotient set $\mathcal{G}_{q,\gamma,s}/{\sim_R} := \{[F]_R \colon F \in \mathcal{G}_{q,\gamma,s}\}$. In analogy to the approach of subsection 4.2 we introduce on $\mathcal{G}_{q,\gamma,s}/{\sim_R}$ the following metrics.

Lemma 6.17
For every fixed $c > 0$, set

$$\tilde{d}_{q,\gamma,s,c}([F]_R,[G]_R) := d_{q,\gamma,s}\big(F\big(m_\tau(F)/c\,\cdot\,\big), G\big(m_\tau(G)/c\,\cdot\,\big)\big)$$

for every $[F]_R$, $[G]_R \in \mathcal{G}_{q,\gamma,s}/{\sim_R}$, where $m_\tau(F) = (\tau \int_{\mathbb{R}} |x|^\tau F(dx))^{1/\tau}$ for every $F \in \mathcal{G}_{q,\gamma,s}$. Then

(i) $\tilde{d}_{q,\gamma,s,c}$ is well-defined and a metric on $\mathcal{G}_{q,\gamma,s}/{\sim_R}$,

(ii) for any two constants $c_1, c_2 \in (0,\infty)$ the metrics $\tilde{d}_{q,\gamma,s,c_1}$ and $\tilde{d}_{q,\gamma,s,c_2}$ are uniformly equivalent.

Proof.
(i) To verify that $\tilde{d}_{q,\gamma,s,c}$ is well-defined, observe that for every F_i, $G_i \in \mathcal{G}_{q,\gamma,s}$, $i = 1,2$, with $F_1 \sim_R F_2$, $G_1 \sim_R G_2$ it is

$$F_1(m_\tau(F_1)\,\cdot\,) = F_2(m_\tau(F_2)\,\cdot\,) \quad \text{and} \quad G_1(m_\tau(G_1)\,\cdot\,) = G_2(m_\tau(G_2)\,\cdot\,)$$

by definition of $\sim_R$, cf. (6.37). Because of this,

$$d_{q,\gamma,s}\big(F_1(m_\tau(F_1)/c\,\cdot\,), G_1(m_\tau(G_1)/c\,\cdot\,)\big) = d_{q,\gamma,s}\big(F_2(m_\tau(F_2)/c\,\cdot\,), G_2(m_\tau(G_2)/c\,\cdot\,)\big),$$

so that $\tilde{d}_{q,\gamma,s,c}([F_1]_R,[G_1]_R) = \tilde{d}_{q,\gamma,s,c}([F_2]_R,[G_2]_R)$.

The metric properties of $\tilde{d}_{q,\gamma,s,c}$ follow directly from the respective properties of $d_{q,\gamma,s}$ and the definition of $\sim_R$.

(ii) Let $G, H \in \mathcal{G}_{q,\gamma,s}$, and denote by $\tilde{g}$ and $\tilde{h}$ the densities of $\widetilde{G} := G(m_\tau(G)\,\cdot\,)$ and $\widetilde{H} := H(m_\tau(H)\,\cdot\,)$, respectively. Then

$$\Big[\frac{1}{c_1}\tilde{g}\Big(\frac{\cdot}{c_1}\Big) - \frac{1}{c_1}\tilde{h}\Big(\frac{\cdot}{c_1}\Big)\Big]_\gamma = \frac{1}{c_1^{1+\gamma}}\big[\tilde{g} - \tilde{h}\big]_\gamma = \Big(\frac{c_2}{c_1}\Big)^{1+\gamma}\Big[\frac{1}{c_2}\tilde{g}\Big(\frac{\cdot}{c_2}\Big) - \frac{1}{c_2}\tilde{h}\Big(\frac{\cdot}{c_2}\Big)\Big]_\gamma,$$

$$\Big\|\frac{1}{c_1}\tilde{g}\Big(\frac{\cdot}{c_1}\Big) - \frac{1}{c_1}\tilde{h}\Big(\frac{\cdot}{c_1}\Big)\Big\|_\infty = \frac{1}{c_1}\|\tilde{g} - \tilde{h}\|_\infty = \frac{c_2}{c_1}\Big\|\frac{1}{c_2}\tilde{g}\Big(\frac{\cdot}{c_2}\Big) - \frac{1}{c_2}\tilde{h}\Big(\frac{\cdot}{c_2}\Big)\Big\|_\infty$$

and

$$\Big\|\frac{1}{c_1}\tilde{g}\Big(\frac{\cdot}{c_1}\Big) - \frac{1}{c_1}\tilde{h}\Big(\frac{\cdot}{c_1}\Big)\Big\|_{s,\infty} = c_1^{s-1}\|\tilde{g} - \tilde{h}\|_{s,\infty} = \Big(\frac{c_1}{c_2}\Big)^{s-1}\Big\|\frac{1}{c_2}\tilde{g}\Big(\frac{\cdot}{c_2}\Big) - \frac{1}{c_2}\tilde{h}\Big(\frac{\cdot}{c_2}\Big)\Big\|_{s,\infty}.$$

Moreover, it is

$$\Big|\int_{\mathbb{R}}|x|^q\frac{1}{c_1}\tilde{g}\Big(\frac{x}{c_1}\Big)dx - \int_{\mathbb{R}}|x|^q\frac{1}{c_1}\tilde{h}\Big(\frac{x}{c_1}\Big)dx\Big| = \Big(\frac{c_1}{c_2}\Big)^q\Big|\int_{\mathbb{R}}|x|^q\frac{1}{c_2}\tilde{g}\Big(\frac{x}{c_2}\Big)dx - \int_{\mathbb{R}}|x|^q\frac{1}{c_2}\tilde{h}\Big(\frac{x}{c_2}\Big)dx\Big|.$$

Now set $k := c_1/c_2$ and $S := \{k^{-(1+\gamma)}, k^{-1}, k^{s-1}, k^q\}$. Using the above, it follows that

$$\begin{aligned}\tilde{d}_{q,\gamma,s,c_1}([G]_R,[H]_R) &= d_{q,\gamma,s}\big(\widetilde{G}(\cdot/c_1),\widetilde{H}(\cdot/c_1)\big)\\ &\le \max S\cdot d_{q,\gamma,s}\big(\widetilde{G}(\cdot/c_2),\widetilde{H}(\cdot/c_2)\big) = \max S\cdot \tilde{d}_{q,\gamma,s,c_2}([G]_R,[H]_R)\end{aligned}$$

and

$$\begin{aligned}\tilde{d}_{q,\gamma,s,c_1}([G]_R,[H]_R) &= d_{q,\gamma,s}\big(\widetilde{G}(\cdot/c_1),\widetilde{H}(\cdot/c_1)\big)\\ &\ge \min S\cdot d_{q,\gamma,s}\big(\widetilde{G}(\cdot/c_2),\widetilde{H}(\cdot/c_2)\big) = \min S\cdot \tilde{d}_{q,\gamma,s,c_2}([G]_R,[H]_R),\end{aligned}$$

and since S does neither depend on $[G]_R$ nor $[H]_R$, this concludes the proof. □

The previous lemma shows that any two of the metric spaces $(\mathcal{G}_{q,\gamma,s}/\!\sim_R, \tilde{d}_{q,\gamma,s,c})$, $c > 0$, are uniformly (hence, topologically) isomorphic, so that for our purposes it suffices again to identify these spaces and to work on $(\mathcal{G}_{q,\gamma,s}/\!\sim_R, \tilde{d}_{q,\gamma,s,1})$. Now note that

$$\widetilde{\mathcal{G}}_{q,\gamma,s} := \big\{F(m_\tau(F)\,\cdot\,)\colon F\in\mathcal{G}_{q,\gamma,s}\big\} = \big\{F\in\mathcal{G}_{q,\gamma,s}\colon \tau\textstyle\int_{\mathbb{R}}|x|^\tau F(dx) = 1\big\}$$

is a complete set of equivalence class representatives of $\sim_R$ on $\mathcal{G}_{q,\gamma,s}$, and the well-defined mapping

$$(\mathcal{G}_{q,\gamma,s}/\!\sim_R, \tilde{d}_{q,\gamma,s,1}) \ni [F]_R \mapsto F(m_\tau(F)\,\cdot\,) \in (\widetilde{\mathcal{G}}_{q,\gamma,s}, d_{q,\gamma,s})$$

is easily seen to be an isometric isomorphism. Consequently, the two metric spaces $(\mathcal{G}_{q,\gamma,s}/\!\sim_R, \tilde{d}_{q,\gamma,s,1})$ and $(\widetilde{\mathcal{G}}_{q,\gamma,s}, d_{q,\gamma,s})$ are isometrically isomorphic, and we will not differentiate between them in the following.

For investigating the asymptotic behavior of $N_2(\alpha,\beta,F)/N_1(\alpha,\beta,F)$ we will therefore assume from now on that the unknown distribution function F of the errors $(e_i)_{i\in\mathbb{N}}$ is an element of $\widetilde{\mathcal{G}}_{q,\gamma,s}$, and we will measure the distance of any two distribution functions in $\widetilde{\mathcal{G}}_{q,\gamma,s}$ with the metric $d_{q,\gamma,s}$ if not stated otherwise. Because of this, we consider in the following the testing problem

$$\mathrm{H}_0\colon F = F_\tau \quad \text{versus} \quad \mathrm{H}_1\colon F\in\widetilde{\mathcal{G}}_{q,\gamma,s}\setminus\{F_\tau\}. \tag{6.50}$$

Recall that we have studied the testing problem (6.38) before, where the set of alternatives is larger than in the problem above. But since $\widetilde{\mathcal{G}}_{q,\gamma,s}$ is a subset of $\widetilde{\mathcal{G}}_q^u$, all results previously derived under H_1 still hold true when restricting the alternatives to $\widetilde{\mathcal{G}}_{q,\gamma,s}\setminus\{F_\tau\}$.

Remark 6.18: The foregoing testing problem satisfies condition (2.2). To see this, note again that it follows from Lemma 6.16 that $F_{\tau^*} \in \mathcal{G}_{q,\gamma,s}$ for every $\tau^* \geq \gamma$. Hence, the set $\widetilde{\mathcal{G}}_{q,\gamma,s}$ contains the distinct elements $F_{\tau^*}(m_\tau(F_{\tau^*}) \cdot\ )$, $\tau^* \geq \gamma$, so that the set of alternatives in (6.50) is evidently not empty.
Now fix an $F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \setminus \{F_\tau\}$ and set $F_t := tF + (1-t)F_\tau$ for every $t \in (0,1)$. Again, it is easy to see that $F_t \in \mathcal{G}_q \setminus \{F_\tau\}$ and that its Lebesgue density $f_t := tf + (1-t)f_\tau$ satisfies $[f_t]_\gamma < \infty$ and $\|f_t\|_{s,\infty} < \infty$, with f denoting the density of F. Thus, $F_t \in \mathcal{G}_{q,\gamma,s} \setminus \{F_\tau\}$. Moreover, it is

$$\int_{\mathbb{R}} |x|^\tau F_t(dx) = t \cdot \int_{\mathbb{R}} |x|^\tau F(dx) + (1-t) \cdot \int_{\mathbb{R}} |x|^\tau F_\tau(dx) = \frac{1}{\tau}(t + 1 - t) = \frac{1}{\tau},$$

whence it follows that $F_t \in \widetilde{\mathcal{G}}_{q,\gamma,s} \setminus \{F_\tau\}$ for all $t \in (0,1)$. Now since

$$d_{q,\gamma,s}(F_t, F_\tau) = t \cdot d_{q,\gamma,s}(F, F_\tau) \underset{t \to 0}{\longrightarrow} 0,$$

the claim follows. ♦

Let us briefly mention some properties of the metric space $(\widetilde{\mathcal{G}}_{q,\gamma,s}, d_{q,\gamma,s})$. The first result is an analog of Lemma 6.2.

Lemma 6.19
The identity function

$$id\colon (\widetilde{\mathcal{G}}_{q,\gamma,s}, d_{q,\gamma,s}) \ \ni\ F \ \mapsto\ F \ \in\ (\widetilde{\mathcal{G}}_{q,\gamma,s}, d_q)$$

is continuous, where d_q is the Kantorovich-Wasserstein metric defined in (4.1)*.*

Proof. The proof follows along the same lines as the one of Lemma 6.2, except that the convergence $\|f_n^* - f\|_\infty \to 0$ as $n \to \infty$ is here a direct consequence of $d_{q,\gamma,s}(F_n^*, F) \to 0$ as $n \to \infty$. □

Thus, in combination with Lemma 4.1 the previous lemma yields the following:

Corollary 6.20
(i) For every $r \in [1,q]$ the function

$$(\widetilde{\mathcal{G}}_{q,\gamma,s}, d_{q,\gamma,s}) \ \ni\ F \ \mapsto\ \int_{\mathbb{R}} |x|^r F(dx) \ \in (\mathbb{R}, |\cdot|)$$

is continuous.

(ii) The identity function

$$id\colon (\widetilde{\mathcal{G}}_{q,\gamma,s}, d_{q,\gamma,s}) \ \ni\ F \ \mapsto\ F \ \in (\widetilde{\mathcal{G}}_{q,\gamma,s}, d_K)$$

is continuous, where d_K is the Kolmogorov metric.

We are now able to show that the sequences $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ fulfill condition (WIII).

Theorem 6.21
The sequences $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\geq 2}$ fulfill Wieand's condition (WIII) *with*

$$b\colon \widetilde{\mathcal{G}}_{q,\gamma,s} \setminus \{F_\tau\} \ni F \mapsto \Big(\int_{-\infty}^{\infty} \big(F(x) - F_\tau(x)\big)^2 F_\tau(dx)\Big)^{1/2} \in (0,1].$$

Proof. We will first verify the statement for $(\widehat{\mathcal{W}}_{n,res})_{n\geq 2}$.

Let $K := (\int_{\mathbb{R}} x^2 F_\tau(dx))/2 > 0$. It follows from Corollary 6.20 (i) that there is a $\varrho > 0$ such that

$$\Big|\int_{\mathbb{R}} x^2 F(dx) - \int_{\mathbb{R}} x^2 F_\tau(dx)\Big| < K \quad \text{for all } F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \text{ with } d_{q,\gamma,s}(F, F_\tau) < \varrho. \tag{6.51}$$

Now consider the set

$$M := U_\varrho(F_\tau) \setminus \{F_\tau\} = \big\{F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \colon d_{q,\gamma,s}(F, F_\tau) < \varrho\big\} \setminus \{F_\tau\}.$$

We will show next that this set satisfies conditions (5.12)–(5.14), (3.5), (3.7) and (4.36).

As $d_{q,\gamma,s}(F, F_\tau) < \varrho$ for every $F \in M$, it is $[f - f_\tau]_\gamma < \varrho$, $\|f - f_\tau\|_{s,\infty} < \varrho$ and $\|f - f_\tau\|_\infty < \varrho$, with f denoting the density of F. Hence,

$$[f]_\gamma \leq [f - f_\tau]_\gamma + [f_\tau]_\gamma < \varrho + [f_\tau]_\gamma < \infty,$$

whence it follows that the set $\{[f]_\gamma \colon F \in M\}$ is bounded. Consequently, the family $\{f \colon F \in M\}$ is uniformly equicontinuous, which proves (5.12). Moreover, for every $x \in \mathbb{R}^*$ we have

$$f(x) \leq |f(x) - f_\tau(x)| + f_\tau(x) \leq \frac{1}{|x|^s} \cdot \|f - f_\tau\|_{s,\infty} + f_\tau(x) < \frac{\varrho}{|x|^s} + f_\tau(x),$$

which yields

$$\sup_{F \in M} f(x) \leq \frac{\varrho}{|x|^s} + f_\tau(x) \underset{|x|\to\infty}{\longrightarrow} 0,$$

so that condition (5.13) is also shown. In addition, we see that (5.14) holds because

$$\|f\|_\infty \leq \|f - f_\tau\|_\infty + \|f_\tau\|_\infty < \varrho + \|f_\tau\|_\infty < \infty$$

for every $F \in M$.

Since it follows from (6.51) that $\inf_{F\in M} \int_{\mathbb{R}} x^2 F(dx) \geq K > 0$, the set M also satisfies (3.5).

Now note that we have

$$\Big|\int_{\mathbb{R}} |x|^q F(dx) - \int_{\mathbb{R}} |x|^q F_\tau(dx)\Big| < \varrho \quad \forall\ F \in M,$$

so that

$$\sup_{F \in M} \int_{\mathbb{R}} |x|^q F(dx) < \infty. \tag{6.52}$$

But as $q = 2\tau$ for $\tau > 1$ and $q > 2$ for $\tau \in (0, 1]$, it is $q > 2$ in both cases. Hence,

$$\sup_{F \in M} \int_{\{|x|>c\}} x^2 F(dx) \leq c^{2-q} \sup_{F \in M} \int_{\mathbb{R}} |x|^q F(dx) < \infty$$

for every $c \in (0, \infty)$, which yields (3.7).

Observe next that if $\tau > 1$, then (6.52) is just condition (4.36), since $q = 2\tau$ in this case. If $0 < \tau \leq 1$ ($\Leftrightarrow 0 < 2\tau \leq 2$), then (3.7) implies (4.36), so that the latter condition holds in this case as well.

Furthermore, note that condition (3.9) holds for any value of τ, as it follows from (3.7).

If $\tau \in (0, 1)$, then we also have for every $F \in M$ that

$$\int_{\mathbb{R}^*} |x|^{\tau-1} F(dx) = \int_{\{0<|x|\leq 1\}} |x|^{\tau-1} f(x)dx + \int_{\{|x|>1\}} |x|^{\tau-1} f(x)dx$$

$$\leq \|f\|_\infty 2 \int_0^1 x^{\tau-1} dx + \int_{\mathbb{R}} f(x) dx \leq \frac{2}{\tau} \sup_{F \in M} \|f\|_\infty + 1, \tag{6.53}$$

and the right-hand side of the last inequality is finite by (5.14). Hence, for every $\tau \in (0,1)$ it is $\sup_{F \in M} \int_{\mathbb{R}^*} |x|^{\tau-1} F(dx) < \infty$.

Since it has been shown above that M fulfills the assumptions of Proposition 5.7, it follows that the least squares estimator satisfies $\sqrt{n}(\hat{\rho}_n^{LS} - \rho) = O_P^u(1)$ in M as $n \to \infty$.

Moreover, as M meets the requirements of Lemma 6.12, we have $\sup_{F \in M} P_F(\hat{\sigma}_{n,res} = 0) \to 0$ as $n \to \infty$. Consequently, for every $\delta \in (0,1)$ there is a $C(\delta) > 0$ such that

$$P_F\big(\hat{\sigma}_{n,res} = 0\big) < \delta$$

for every $F \in M$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C(\delta)/b(F) \geq C(\delta)$, so that for the following investigations with respect to the measures P_F, $F \in M$, we can assume again that $\hat{\sigma}_{n,res} > 0$.

Now recall that by (6.39) and (6.40) it is

$$\Big|\frac{\widehat{\mathcal{W}}_{n,res}}{\sqrt{n}} - b(F)\Big| \leq \|F_{n,res} - F\|_\infty + \big|b_{n,res}(F) - b(F)\big|$$

for every $F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \setminus \{F_\tau\}$ and $n \geq 2$. But

$$\|F_{n,res} - F\|_\infty \leq \|F_{n,res} - F_n\|_\infty + \|F_n - F\|_\infty,$$

and it follows as in the proof of Theorem 6.5 that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C_1(\epsilon, \delta) > 0$ with

$$P_F\big(\|F_{n,res} - F_n\|_\infty + \|F_n - F\|_\infty \geq \epsilon\, b(F)\big) < \delta$$

for every $F \in M$ and $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C_1(\epsilon,\delta)/b(F)$.

Let us now take a look at $\big|b_{n,res}(F) - b(F)\big|$. Note that we have verified above that M satisfies the assumptions of Lemma 6.14. Thus, Proposition 6.15 states that for every $\epsilon > 0$ and $\delta \in (0,1)$ there is a $C_2(\epsilon, \delta) > 0$ such that

$$P_F\Big(\big|b_{n,res}(F) - b(F)\big| \geq \epsilon\, b(F)\Big) \ < \ \delta$$

for all $F \in M$ and for all $n \in \mathbb{N}$, $n \geq 2$, with $\sqrt{n} > C_2(\epsilon, \delta)/b(F)$.

Combining these results, this shows that $(\widehat{\mathcal{W}}_{n,res})_{n \geq 2}$ satisfies Wieand's condition (WIII).

It remains to investigate the sequence $(\widehat{\mathcal{V}}_{n,res})_{n \geq 2}$.

By part (i) of Corollary 6.20 there is a $\varrho_1' > 0$ such that

$$\Big|\int_{\mathbb{R}} |x| F(dx) - \int_{\mathbb{R}} |x| F_\tau(dx)\Big| < K' \ \text{ for all } F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \text{ with } d_{q,\gamma,s}(F, F_\tau) < \varrho_1' \tag{6.54}$$

with $K' := (\int_{\mathbb{R}} |x| F_\tau(dx))/2 > 0$. Moreover, it is $0 < F_\tau < 1$ because f_τ is strictly positive. Hence, there are real numbers $x_1 < 0$ and $x_2 > 0$ with $F_\tau(x_1) > 0$ and $F_\tau(x_2) < 1$. It follows now from part (ii) of Corollary 6.20 that for $K'' := \min\big(F_\tau(x_1), 1 - F_\tau(x_2)\big)/2 > 0$ there is a $\varrho_2' > 0$ such that

$$d_K(F, F_\tau) = \|F - F_\tau\|_\infty < K'' \ \text{ for all } F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \text{ with } d_{q,\gamma,s}(F, F_\tau) < \varrho_2'. \tag{6.55}$$

Let us examine the set

$$M' := U_{\varrho'}(F_\tau) \setminus \{F_\tau\} = \big\{F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \colon d_{q,\gamma,s}(F, F_\tau) < \varrho'\big\} \setminus \{F_\tau\}$$

with $\varrho' := \min(\varrho_1', \varrho_2')$.

Note that we can show that M' satisfies conditions (5.12)–(5.14), (3.7) and (4.36) analogously to the verification of the very conditions for the set M before by simply replacing M by M' and ϱ by ϱ'. Moreover, we see as in (6.53) that (5.14) implies $\sup_{F\in M'}\int_{\mathbb{R}^*}|x|^{\tau-1}F(dx)<\infty$ for $\tau\in(0,1)$.

By (6.54) we also have $\int_{\mathbb{R}}|x|F(dx)>K'$ for all $F\in M'$, which shows that (3.8) holds for M'.

We will verify next that M' also satisfies (5.75). But this is easily seen to be true, because

$$F(x_1)>F_\tau(x_1)-K''>0 \quad \text{and} \quad F(x_2)<F_\tau(x_2)+K''<1$$

for every $F\in M'$ by (6.55).

To sum up, we have verified that M' fulfills conditions (3.7), (3.8), (4.36), (5.12)–(5.14) and (5.75). Note that M' also satisfies condition (3.9), since it is implied by (3.7).

As condition (3.5) follows from (3.8), we get from Proposition 5.7 that $\sqrt{n}(\hat\rho_n^{LS}-\rho)=O_P^u(1)$ in M' as $n\to\infty$.

We have shown above that the set M' satisfies the requirements of Lemma 6.12, whence it follows that $\sup_{F\in M'}P_F(\hat\sigma_{n,res}=0)\to 0$ as $n\to\infty$. Just like before, the following investigations with respect to the measures P_F, $F\in M'$, can therefore be carried out on the event $\{\hat\sigma_{n,res}>0\}$.

Now as mentioned in (6.41), it is

$$\Big|\frac{\widehat{\mathcal{V}}_{n,res}}{\sqrt{n}}-b(F)\Big|\le\|\widetilde{F}_{n,res}-F\|_\infty+\big|b_{n,res}(F)-b(F)\big|$$

for every $F\in\widetilde{\mathcal{G}}_{q,\gamma,s}\setminus\{F_\tau\}$ and $n\ge 2$. As in the proof of Theorem 6.5 we can show that for every $\epsilon>0$ and $\delta\in(0,1)$ there is a $C_1(\epsilon,\delta)'>0$ such that

$$P_F\big(\|\widetilde{F}_{n,res}-F\|_\infty\ge\epsilon\, b(F)\big)<\delta$$

for every $F\in M'$ and $n\in\mathbb{N}$, $n\ge 2$, with $\sqrt{n}>C_1(\epsilon,\delta)'/b(F)$. Moreover, observe that we have checked above that the set M' also fulfills the requirements of Lemma 6.14. Consequently, we get from Proposition 6.15 that for every $\epsilon>0$ and $\delta\in(0,1)$ there is a $C_2(\epsilon,\delta)'>0$ with

$$P_F\big(\,\big|b_{n,res}(F)-b(F)\big|\ge\epsilon\, b(F)\,\big)\;<\;\delta$$

for all $F\in M'$ and for all $n\in\mathbb{N}$, $n\ge 2$, with $\sqrt{n}>C_2(\epsilon,\delta)'/b(F)$. A combination of these results evidently implies that (WIII) holds for the sequence $(\widehat{\mathcal{V}}_{n,res})_{n\ge 2}$ as well, and this completes the proof. □

We have now collected all results that are needed to show that the approximate Bahadur ARE of $(\widehat{\mathcal{W}}_{n,res})_{n\ge 2}$ relative to $(\widehat{\mathcal{V}}_{n,res})_{n\ge 2}$ determined in Proposition 6.11 is equal to the limiting (as $\alpha\to 0$) Pitman ARE of these sequences. This follows again from Theorem 2.3 once we have checked that the two sequences of test statistics meet its requirements.
Hence, let us summarize what we have shown. We have verified that the sequences $(\widehat{\mathcal{W}}_{n,res})_{n\ge 2}$ and $(\widehat{\mathcal{V}}_{n,res})_{n\ge 2}$ fulfill (BI), (BII) and (WIII), which means that they satisfy condition (i) of Theorem 2.3. Moreover, we have noted before that the random variables $\widehat{\mathcal{W}}$ in (6.35) and $\widehat{\mathcal{V}}$ in (6.36) have distribution functions that are strictly increasing in their right tails, so that condition (ii) of Theorem 2.3 also holds. Condition (iii) of this theorem is satisfied as well, as is easily seen using $0<b(F)\le d_K(F,F_\tau)$ for every $F\in\widetilde{\mathcal{G}}_{q,\gamma,s}\setminus\{F_\tau\}$ and part (ii) of Corollary 6.20. As condition (iv) of Theorem 2.3 is again trivially satisfied, we have thus verified the following theorem.

Theorem 6.22
Set $T_{1n} = \widehat{\mathcal{W}}_{n,res}$ *and* $T_{2n} = \widehat{\mathcal{V}}_{n,res}$, $n \geq 2$. *Then we have for every* $\beta \in (0,1)$

$$\lim_{\alpha \to 0} \liminf_{\substack{F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \setminus \{F_\tau\}, \\ d_{q,\gamma,s}(F,F_\tau) \to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \lim_{\alpha \to 0} \limsup_{\substack{F \in \widetilde{\mathcal{G}}_{q,\gamma,s} \setminus \{F_\tau\}, \\ d_{q,\gamma,s}(F,F_\tau) \to 0}} \frac{N_2(\alpha,\beta,F)}{N_1(\alpha,\beta,F)} = \frac{\tilde{\lambda}_1^*}{\tilde{\lambda}_1} \ < \ 1.$$

This shows that the limiting Pitman ARE of the sequence $(\widehat{\mathcal{W}}_{n,res})_{n \geq 2}$ with respect to the sequence $(\widehat{\mathcal{V}}_{n,res})_{n \geq 2}$ is also strictly less than one, so that the sequence of tests corresponding to $(\widehat{\mathcal{V}}_{n,res})_{n \geq 2}$ is to be preferred to the one based on $(\widehat{\mathcal{W}}_{n,res})_{n \geq 2}$. The explicit value of the above limiting Pitman ARE for $\tau = 1$ and $\tau = 2$ can be found in Table 2 on page 52.

Appendices

A Auxiliary Results

Lemma A.1
Let $n \geq 2$ and $x_1, \dots, x_n \in \mathbb{R}$ with

$$\min_{1 \leq i \leq n} x_i < 0 < \max_{1 \leq i \leq n} x_i. \tag{A.1}$$

Then there is exactly one $t = t(x_1, \dots, x_n) \in \mathbb{R}$ with

$$\Big(\frac{1}{n} - 1\Big) \frac{1}{\max\limits_{1 \leq i \leq n} x_i} < t < \Big(\frac{1}{n} - 1\Big) \frac{1}{\min\limits_{1 \leq i \leq n} x_i}$$

and

$$\sum_{i=1}^{n} \frac{x_i}{1 + t x_i} = 0.$$

Proof. Set $\underline{x} := \min_{1 \leq i \leq n} x_i$, $\overline{x} := \max_{1 \leq i \leq n} x_i$ and

$$I := \Big(-\frac{1}{\overline{x}}, -\frac{1}{\underline{x}}\Big).$$

First, note that

$$1 + t x_i > 0 \quad \forall\, i = 1, \dots, n,\ t \in I.$$

To verify this, let $i \in \{1, \dots, n\}$ and $t \in I$. If $t \leq 0$, then $x_i \leq \overline{x}$ implies $1 + t x_i \geq 1 + t\overline{x}$, and the right-hand side of the last inequality is positive since $t > -1/\overline{x}$. Similarly, if $t > 0$, then $x_i \geq \underline{x}$ implies that $1 + t x_i \geq 1 + t\underline{x}$, and the right-hand side of the last inequality is positive because of $t < -1/\underline{x}$.

Thus, the function

$$f(t) := \sum_{i=1}^{n} \frac{x_i}{1 + t x_i}, \quad t \in I,$$

is well-defined. Obviously, f is continuously differentiable with

$$f'(t) = -\sum_{i=1}^{n} \frac{x_i^2}{(1 + t x_i)^2}, \quad t \in I,$$

and $f' < 0$ on I because of (A.1), so that f is strictly decreasing. Now define

$$t_n^- := \Big(\frac{1}{n} - 1\Big) \frac{1}{\max\limits_{1 \leq i \leq n} x_i} \quad \text{and} \quad t_n^+ := \Big(\frac{1}{n} - 1\Big) \frac{1}{\min\limits_{1 \leq i \leq n} x_i}.$$

Note that $0 \in (t_n^-, t_n^+) \subset I$. If we can show that $f(t_n^-) > 0$ and $f(t_n^+) < 0$, the statement follows from the continuity and monotonicity of f. To do this, consider for $t \in \mathbb{R}^*$ the function

$$g_t(x) := \frac{x}{1 + tx}, \quad x \in \mathbb{R} \setminus \{-1/t\}.$$

Then g_t is differentiable on $\mathbb{R} \setminus \{-1/t\}$ with derivative

$$g_t'(x) = \frac{1}{(1 + tx)^2} > 0.$$

Hence, g_t is strictly increasing on $(-\infty, -1/t)$ and $(-1/t, \infty)$. By what was shown above, we have $1 + t_n^- x_i > 0$ for all $i = 1, \ldots, n$. Since $t_n^- < 0$, this implies that $x_i < -1/t_n^-$ for $i = 1, \ldots, n$. Thus,

$$f(t_n^-) = \sum_{i=1}^{n} \frac{x_i}{1 + t_n^- x_i} \;\geq\; \frac{\overline{x}}{1 + t_n^- \overline{x}} + (n-1) \cdot \frac{\underline{x}}{1 + t_n^- \underline{x}},$$

since at least one summand in the left sum is equal to $\overline{x}/(1 + t_n^- \overline{x})$ and

$$\frac{x_i}{1 + t_n^- x_i} = g_{t_n^-}(x_i) \;\geq\; g_{t_n^-}(\underline{x}) = \frac{\underline{x}}{1 + t_n^- \underline{x}} \quad \forall\, i = 1, \ldots, n$$

because of $\underline{x} \leq x_i < -1/t_n^-$ and the monotonicity of $g_{t_n^-}$. Now since $(1-n)\underline{x} > 0$ and $n\overline{x} > 0$, we have

$$\begin{aligned} \frac{(1-n)\underline{x}}{n\overline{x} + (1-n)\underline{x}} < 1 \quad &\Rightarrow \quad 0 < 1 + \frac{(n-1)\underline{x}}{n\overline{x} + (1-n)\underline{x}} \\ &\Rightarrow \quad 0 < n\overline{x} + \frac{n\overline{x}(n-1)\underline{x}}{n\overline{x} + (1-n)\underline{x}} = \frac{\overline{x}}{1 + t_n^- \overline{x}} + (n-1) \cdot \frac{\underline{x}}{1 + t_n^- \underline{x}}, \end{aligned}$$

so that $f(t_n^-) > 0$.

Analogously, we have $1 + t_n^+ x_i > 0$ for all $i = 1, \ldots, n$, and it follows from this that $x_i > -1/t_n^+$ because t_n^+ is positive. Then

$$f(t_n^+) = \sum_{i=1}^{n} \frac{x_i}{1 + t_n^+ x_i} \;\leq\; \frac{\underline{x}}{1 + t_n^+ \underline{x}} + (n-1) \cdot \frac{\overline{x}}{1 + t_n^+ \overline{x}},$$

because at least one summand in the left sum is equal to $\underline{x}/(1 + t_n^+ \underline{x})$ and

$$\frac{x_i}{1 + t_n^+ x_i} = g_{t_n^+}(x_i) \;\leq\; g_{t_n^+}(\overline{x}) = \frac{\overline{x}}{1 + t_n^+ \overline{x}} \quad \forall\, i = 1, \ldots, n$$

since $-1/t_n^+ < x_i \leq \overline{x}$ and $g_{t_n^+}$ is monotonically increasing on $(-1/t_n^+, \infty)$. Using $n\underline{x} < 0$ and $(1-n)\overline{x} < 0$, this yields

$$\begin{aligned} \frac{(1-n)\overline{x}}{n\underline{x} + (1-n)\overline{x}} < 1 \quad &\Rightarrow \quad 0 < 1 + \frac{(n-1)\overline{x}}{n\underline{x} + (1-n)\overline{x}} \\ &\Rightarrow \quad 0 > n\underline{x} + \frac{n\underline{x}(n-1)\overline{x}}{n\underline{x} + (1-n)\overline{x}} = \frac{\underline{x}}{1 + t_n^+ \underline{x}} + (n-1) \cdot \frac{\overline{x}}{1 + t_n^+ \overline{x}}, \end{aligned}$$

whence $f(t_n^+) < 0$ follows. □

Let $n \geq 2$. Consider the open set

$$B_n := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon \min_{1 \leq i \leq n} x_i < 0 < \max_{1 \leq i \leq n} x_i\}.$$

For every $(x_1, \ldots, x_n) \in B_n$ let $t(x_1, \ldots, x_n)$ be as in the previous lemma.

Lemma A.2
The function

$$B_n \ni (x_1, \ldots, x_n) \;\mapsto\; t(x_1, \ldots, x_n) \in \mathbb{R}$$

is continuous.

Proof. Let $(\tilde{x}_1, \dots, \tilde{x}_n)$ be an arbitrary point in B_n, and let $\big((x_1^m, \dots, x_n^m)\big)_{m\in\mathbb{N}}$ be a sequence in B_n such that $(x_1^m, \dots, x_n^m) \to (\tilde{x}_1, \dots, \tilde{x}_n)$ as $m \to \infty$. We prove that

$$t(x_1^m, \dots, x_n^m) \underset{m\to\infty}{\longrightarrow} t(\tilde{x}_1, \dots, \tilde{x}_n)$$

by showing that every subsequence of $\big(t(x_1^m, \dots, x_n^m)\big)_{m\in\mathbb{N}}$ has a further subsequence that converges to $t(\tilde{x}_1, \dots, \tilde{x}_n)$.

Let $\big(t(x_1^{m_k}, \dots, x_n^{m_k})\big)_{k\in\mathbb{N}}$ be a subsequence of $\big(t(x_1^m, \dots, x_n^m)\big)_{m\in\mathbb{N}}$. Then $\big((x_1^{m_k}, \dots, x_n^{m_k})\big)_{k\in\mathbb{N}}$ is obviously a subsequence of $\big((x_1^m, \dots, x_n^m)\big)_{m\in\mathbb{N}}$, and therefore

$$(x_1^{m_k}, \dots, x_n^{m_k}) \underset{k\to\infty}{\longrightarrow} (\tilde{x}_1, \dots, \tilde{x}_n).$$

We have shown in Lemma A.1 that t satisfies

$$t^-_{n,m_k} := \Big(\frac{1}{n} - 1\Big)\frac{1}{\max\limits_{1\le i\le n} x_i^{m_k}} < t(x_1^{m_k}, \dots, x_n^{m_k}) < \Big(\frac{1}{n} - 1\Big)\frac{1}{\min\limits_{1\le i\le n} x_i^{m_k}} =: t^+_{n,m_k} \qquad (*)$$

for every $k \in \mathbb{N}$. Because of the continuity of min and max the bounds t^-_{n,m_k} and t^+_{n,m_k} converge in $\mathbb{R}$ as $k \to \infty$, and therefore are bounded themselves. This implies that $\big(t(x_1^{m_k}, \dots, x_n^{m_k})\big)_{k\in\mathbb{N}}$ is a bounded sequence, and therefore it has a convergent subsequence, i.e., there is a sequence $\big((x_1^{m_{k_l}}, \dots, x_n^{m_{k_l}})\big)_{l\in\mathbb{N}}$ and a $c \in \mathbb{R}$ such that

$$t(x_1^{m_{k_l}}, \dots, x_n^{m_{k_l}}) \underset{l\to\infty}{\longrightarrow} c.$$

Because of $(*)$ it is

$$c \;\in\; \Big[\Big(\frac{1}{n} - 1\Big)\frac{1}{\max\limits_{1\le i\le n} \tilde{x}_i}, \Big(\frac{1}{n} - 1\Big)\frac{1}{\min\limits_{1\le i\le n} \tilde{x}_i}\Big] \;\subset\; \Big(-\frac{1}{\max\limits_{1\le i\le n} \tilde{x}_i}, -\frac{1}{\min\limits_{1\le i\le n} \tilde{x}_i}\Big).$$

Now

$$\sum_{i=1}^n \frac{x_i^{m_{k_l}}}{1 + t(x_1^{m_{k_l}}, \dots, x_n^{m_{k_l}})x_i^{m_{k_l}}} \quad \underset{l\to\infty}{\longrightarrow} \quad \sum_{i=1}^n \frac{\tilde{x}_i}{1 + c\,\tilde{x}_i},$$

and by the definition of $t(x_1^{m_{k_l}}, \dots, x_n^{m_{k_l}})$ we know that the left-hand side equals zero for every $l \in \mathbb{N}$. Hence, the limit vanishes as well. But the proof of Lemma A.1 shows that the equation

$$\sum_{i=1}^n \frac{\tilde{x}_i}{1 + t\,\tilde{x}_i} = 0$$

has a unique solution in the interval $\big(-1/\max_{1\le i\le n} \tilde{x}_i, -1/\min_{1\le i\le n} \tilde{x}_i\big)$, namely $t(\tilde{x}_1, \dots, \tilde{x}_n)$. This shows that $t(\tilde{x}_1, \dots, \tilde{x}_n) = c$, and this concludes the proof. □

Proposition A.3
Let M be a nonempty set of distribution functions F such that each F has uniformly continuous Lebesgue density f and

(i) *the family $\{f\colon F \in M\}$ is uniformly equicontinuous,*

(ii) $\sup\limits_{F\in M} f(x) \underset{|x|\to\infty}{\longrightarrow} 0.$

Moreover, let G be a distribution function that is continuously differentiable and strictly increasing. Then

$$\lim_{\delta\downarrow 0} \sup_{F\in M} \sup_{\substack{x,y\in\mathbb{R}\\ |G(x)-G(y)|\le\delta}} |f(x) - f(y)| = 0.$$

Proof. Let $\epsilon > 0$ be arbitrary, but fixed. Since

$$\sup_{F\in M} \sup\{|f(x) - f(y)|\colon x, y \in \mathbb{R}, |G(x) - G(y)| \le \delta\}$$

is non-decreasing in δ, it suffices to show that there is a $\delta_\epsilon > 0$ such that

$$\sup_{F\in M} \sup\{|f(x) - f(y)|\colon x, y \in \mathbb{R}, |G(x) - G(y)| \le \delta_\epsilon\} \le \epsilon. \tag{A.2}$$

Because of (ii) there is an $x_0 > 0$ with $\sup\{f(x)\colon F \in M\} \le \epsilon/2$ for all $x \in \mathbb{R}$ with $|x| \ge x_0$. Moreover, there are $x_1, x_2 \in \mathbb{R}$ such that $G(x) < 1/4$ for all $x \le x_1$ and $G(x) > 3/4$ for $x \ge x_2$. Define

$$c_\epsilon^- := (-x_0) \wedge x_1 \quad \text{and} \quad c_\epsilon^+ := x_0 \vee x_2.$$

Then $G(c_\epsilon^+) - G(c_\epsilon^-) \ge G(x_2) - G(x_1) > 1/2$. Because of (i) there is a $\tilde{\delta}_\epsilon > 0$ such that

$$|f(x) - f(y)| \le \epsilon$$

for all $x, y \in \mathbb{R}$ with $|x - y| \le \tilde{\delta}_\epsilon$ and for all $f = F'$ with $F \in M$. Now set

$$I := \Big[\frac{1}{2}G(c_\epsilon^-), \frac{1}{2} + \frac{1}{2}G(c_\epsilon^+)\Big] \subset (0,1) \quad \text{and} \quad K := \max_{z\in I} |G^{-1'}(z)|,$$

where $G^{-1'}$ is the continuous derivative of G^{-1}, the inverse function of G. Note that K is well-defined and in $(0,\infty)$ because as G is continuously differentiable and strictly increasing, so is G^{-1}.

We will show now that every

$$\delta_\epsilon \in \Big(0, \min\Big\{\frac{1}{2}G(c_\epsilon^-), \frac{1}{2}\big(1 - G(c_\epsilon^+)\big), \tilde{\delta}_\epsilon/K\Big\}\Big) \subset \big(0, 1/2\big)$$

satisfies (A.2). Thus, fix such a δ_ϵ and take $x, y \in \mathbb{R}$ with $|G(x) - G(y)| \le \delta_\epsilon$. We investigate the following cases:

Case A: $x \le c_\epsilon^-$. Then either $y \le c_\epsilon^-$ as well, or $y > c_\epsilon^-$.
In the first case

$$|f(x) - f(y)| \le f(x) + f(y) \le \sup_{F\in M} f(x) + \sup_{F\in M} f(y) \le 2 \cdot \frac{\epsilon}{2} = \epsilon$$

for every $F \in M$.
In the latter case, i.e., if $y > c_\epsilon^-$, y has to be less than or equal to c_ϵ^+, since otherwise

$$\frac{1}{2} < G(c_\epsilon^+) - G(c_\epsilon^-) < G(y) - G(x) \le \delta_\epsilon < \frac{1}{2}.$$

This implies $G(y) \in \big(G(c_\epsilon^-), G(c_\epsilon^+)\big] \subset I$. Additionally,

$$G(x) \ge G(y) - |G(x) - G(y)| \ge G(c_\epsilon^-) - \delta_\epsilon \ge G(c_\epsilon^-) - \frac{1}{2}G(c_\epsilon^-) = \frac{1}{2}G(c_\epsilon^-),$$

and so it follows that $G(x) \in \big[\frac{1}{2}G(c_\epsilon^-), G(c_\epsilon^-)\big] \subset I$. Hence, using the mean value theorem we have

$$|y - x| = \big|G^{-1}\big(G(y)\big) - G^{-1}\big(G(x)\big)\big| \le K \cdot \big|G(y) - G(x)\big| \le K \cdot \delta_\epsilon \le \tilde{\delta}_\epsilon,$$

whence it follows that $|f(x) - f(y)| \le \epsilon$ for every $F \in M$.

Case B: $c_\epsilon^- < x < c_\epsilon^+$. Then

$$G(y) \geq G(x) - |G(x) - G(y)| \geq G(c_\epsilon^-) - \delta_\epsilon \geq \frac{1}{2} G(c_\epsilon^-)$$

and

$$G(y) \leq G(c_\epsilon^+) + \delta_\epsilon \leq \frac{1}{2} + \frac{1}{2} G(c_\epsilon^+),$$

so that $G(y) \in I$. Moreover, $G(x) \in \big(G(c_\epsilon^-), G(c_\epsilon^+)\big) \subset I$. Thus,

$$|y - x| \leq K \cdot \big|G(y) - G(x)\big| \leq K \cdot \delta_\epsilon \leq \tilde{\delta}_\epsilon,$$

which implies $|f(y) - f(x)| \leq \epsilon$ for every $F \in M$.

Case C: $x \geq c_\epsilon^+$. The proof of this case follows analogously to that of Case A. Therefore it is omitted here. □

The next result is an exponential inequality for bounded martingale difference sequences (MDS) and follows from the martingale inequality in Lemma 1 of Häusler [14].

Lemma A.4
Let $\zeta_1, \ldots, \zeta_n$ be a MDS with respect to the filtration $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n$ and $|\zeta_i| \leq K < \infty$ for all $i = 1, \ldots, n$. Then for every $\epsilon, L > 0$

$$\begin{aligned} &P\Big(\Big\{\Big|\sum_{i=1}^n \zeta_i\Big| \geq \epsilon\Big\} \cap \Big\{\sum_{i=1}^n \mathrm{E}(\zeta_i^2 | \mathcal{F}_{i-1}) \leq L\Big\}\Big) \\ &\leq 2 \exp\Big(\frac{\epsilon}{K} - \Big[\frac{\epsilon}{K} + \frac{L}{K^2}\Big] \log\Big(1 + \frac{\epsilon K}{L}\Big)\Big) \qquad (A.3) \\ &\leq 2 \exp\Big(-\frac{\epsilon^2}{2L} + \frac{1}{2} \frac{\epsilon^3 K}{L^2}\Big). \qquad (A.4) \end{aligned}$$

B Eigenvalues of certain Hilbert-Schmidt integral operators

Consider the complex Hilbert space

$$L_2(0,1) := \big\{ f\colon (0,1) \to \mathbb{C} \,\big|\, |f|^2 \text{ is integrable with respect to Lebesgue measure } \lambda \big\}$$

with inner product $\langle f, g\rangle := \int_0^1 f(t)\bar{g}(t)dt$, where the bar denotes the complex conjugate, and induced norm $\|f\|_{L_2} := \sqrt{\langle f, f\rangle}$. As usual, we do not distinguish between two functions $f, g \in L_2(0,1)$ that differ only on a set of zero Lebesgue measure. Let the function $k\colon (0,1)^2 \to \mathbb{C}$ be such that $|k|^2$ is integrable with respect to two-dimensional Lebesgue measure. Then the mapping

$$T_k\colon L_2(0,1) \ni f \mapsto T_k f \in L_2(0,1), \qquad (T_k f)(s) = \int_0^1 k(s,t) f(t)dt, \;\; s \in (0,1)$$

is called the *Hilbert-Schmidt integral operator* with *kernel function* k.

Now assume that $k\colon [0,1]^2 \to \mathbb{R}$ is the continuous covariance function of a non-trivial, real-valued, square-integrable and centered stochastic process $X = (X(t))_{t\in[0,1]}$, i.e.,

$$k\colon [0,1]^2 \ni (s,t) \mapsto \operatorname{cov}(X(s), X(t)) = \mathrm{E}(X(s)X(t)) \in \mathbb{R}. \tag{B.1}$$

Then k is bounded, and therefore its restriction to $(0,1)^2$ is square-integrable with respect to two-dimensional Lebesgue measure. Consequently, the restriction of k is a kernel function. Henceforth we will not distinguish between k as in (B.1) and its restriction to $(0,1)^2$. In the following, we will compare the largest eigenvalues of some Hilbert-Schmidt integral operators with certain covariance kernel functions. It is well known that the operator T_k with k as in (B.1) has at most a countable set of eigenvalues accumulating only at zero, and that all of its eigenvalues are in $[0, \|T_k\|]$, where $\|T_k\|$ is the operator norm of T_k.

Let k denote henceforth the covariance function of the Brownian bridge B°, i.e., $k(s,t) := s \wedge t - s\cdot t$, $s,t \in [0,1]$. Then it is well known that T_k is positive definite and has simple eigenvalues $\lambda_j := 1/(j\pi)^2$ with corresponding eigenfunctions $g_j(\cdot) := \sqrt{2}\sin(j\pi\cdot)$, $j \in \mathbb{N}$.

Example B.1
Let $F \in \mathcal{G}_q$ for some fixed $q \geq 2$, and set

$$\psi_1(s) := \sigma_F^{-1} U_F(F^{-1}(s)) = \sigma_F^{-1} \int_0^s F^{-1}(u)du, \quad s \in [0,1],$$

where $\sigma_F^2 = \int_{\mathbb{R}} x^2 F(dx) \in (0,\infty)$. Then

$$\rho_1(s,t) := k(s,t) - \psi_1(s)\psi_1(t), \quad s,t \in [0,1],$$

is the covariance function of the process $W \circ F^{-1}$, cf. (4.4). Let $(\lambda_j^*)_{j\in\mathbb{N}}$ denote the decreasing sequence of positive eigenvalues of T_{ρ_1} such that every eigenvalue is repeated in the sequence according to its multiplicity. Then by Remark 5.4 in [15] it is $1/(2\pi)^2 = \lambda_2 \leq \lambda_1^* \leq \lambda_1 = 1/\pi^2$, and

$$\lambda_1^* < \lambda_1 \iff 0 \neq \langle \psi_1, g_1\rangle, \tag{B.2}$$

as follows from Theorem 5.2 in [15]. Now note that since F is continuous and centered, neither $F^{-1}(u) \geq 0$ for all $u \in (0,1)$ nor $F^{-1}(u) \leq 0$ for all $u \in (0,1)$ is possible. As F^{-1} is moreover strictly increasing, this implies that there is a $u_0 \in (0,1)$ such that $F^{-1} < 0$ on $(0,u_0)$ and $F^{-1} > 0$ on $(u_0,1)$. Thus, the continuous function $h(\cdot) := U_F(F^{-1}(\cdot))$ is strictly decreasing on $(0,u_0]$, strictly increasing on $(u_0,1)$, and negative on $(0,1)$. Hence,

$$\langle \psi_1, g_1\rangle = \int_0^1 \psi_1(u)g_1(u)du = \frac{\sqrt{2}}{\sigma_F}\int_0^1 h(u)\sin(\pi u)du \neq 0,$$

as the integrand is negative on $(0,1)$. By (B.2), this implies that $\lambda_1^* < \lambda_1 = 1/\pi^2$.

Example B.2
Let F_τ and f_τ be as before the distribution function and the Lebesgue density of the exponential power distribution, $\tau \in (0,\infty)$. Set

$$\psi_2(s) := \tau^{-1/2} F_\tau^{-1}(s) f_\tau(F_\tau^{-1}(s)), \quad s \in [0,1].$$

Then $\rho_2(s,t) := k(s,t) - \psi_2(s)\psi_2(t)$ is the covariance function of the process $Z \circ F_\tau^{-1}$, cf. (4.23). Let $(\tilde{\lambda}_j)_{j\in\mathbb{N}}$ denote the decreasing sequence of positive eigenvalues of the corresponding integral operator T_{ρ_2} in which every positive eigenvalue appears as many times as its multiplicity. Moreover, let ψ_1, ρ_1 and $(\lambda_j^*)_{j\in\mathbb{N}}$ be as in Example B.1 with $F = F_\tau$. The function

$$\rho_{12}(s,t) := k(s,t) - \psi_1(s)\psi_1(t) - \psi_2(s)\psi_2(t), \quad s,t \in [0,1],$$

is then the covariance function of the process $V \circ F_\tau^{-1}$, cf. (4.25). By $(\tilde{\lambda}_j^*)_{j\in\mathbb{N}}$ we will denote the decreasing sequence of positive eigenvalues of $T_{\rho_{12}}$ where again each positive eigenvalue appears as often as its multiplicity.
Now note that both T_{ρ_1} and T_{ρ_2} are injective. For the former operator, this is shown in Proposition 6.1 of [15], and for the latter operator it follows similarly, see Proposition B.1 below. As the kernel ρ_{12} can be written as $\rho_{12}(s,t) = \rho_1(s,t) - \psi_2(s)\psi_2(t) = \rho_2(s,t) - \psi_1(s)\psi_1(t)$, it thus follows from Remark 5.4 in [15] that $\lambda_2^* \le \tilde{\lambda}_1^* \le \lambda_1^*$ as well as $\tilde{\lambda}_2 \le \tilde{\lambda}_1^* \le \tilde{\lambda}_1$, so that $\tilde{\lambda}_1^* \le \min(\lambda_1^*, \tilde{\lambda}_1)$. Hence,

$$\lambda_1^* < \tilde{\lambda}_1 \implies \tilde{\lambda}_1^* < \tilde{\lambda}_1.$$

Let us examine $\tilde{\lambda}_1$. It follows from the symmetry of f_τ that $\langle \psi_2, g_1\rangle = 0$, which implies that $\tilde{\lambda}_1 = \lambda_1 = 1/\pi^2$, see Theorem 5.2 in [15]. Now recall that we have shown in Example B.1 that $\lambda_1^* < \lambda_1 = 1/\pi^2$, so that $\lambda_1^* < \tilde{\lambda}_1$ indeed holds. Hence, it is

$$\tilde{\lambda}_1^* < \tilde{\lambda}_1.$$

Proposition B.1
Let ρ_2 be as in Example B.2. Then the Hilbert-Schmidt integral operator T_{ρ_2} is injective.

Proof. Let $g \in L_2(0,1)$ with $T_{\rho_2} g = 0$. Then

$$0 = (T_{\rho_2} g)(s) = \int_0^1 \rho_2(s,t) g(t)dt = \int_0^1 (s \wedge t - s\cdot t) g(t)dt - \psi_2(s)\int_0^1 \psi_2(t) g(t)dt \tag{B.3}$$
$$= \int_0^s t g(t)dt + s\int_s^1 g(t)dt - s\int_0^1 t g(t)dt - \psi_2(s)\int_0^1 \psi_2(t)g(t)dt \tag{B.4}$$

for λ-almost every $s \in (0,1)$. Since $(0,1) \ni t \mapsto tg(t) \in \mathbb{C}$ and g are integrable on $(0,1)$ with respect to λ, the fundamental theorem of calculus for Lebesgue integrals implies that the functions $(0,1) \ni s \mapsto \int_0^s tg(t)dt$ and $(0,1) \ni s \mapsto \int_s^1 g(t)dt$ are differentiable λ-almost everywhere with derivatives $sg(s)$ and $-g(s)$ respectively. Moreover, note that the function ψ_2 is differentiable for all $s \in (0,1) \setminus \{1/2\}$ with

$$\psi_2'(s) = \frac{1}{\sqrt{\tau}}\big(1 - \tau\big|F_\tau^{-1}(s)\big|^\tau\big).$$

Hence, we get from (B.4) that

$$0 = sg(s) + \int_s^1 g(t)dt - sg(s) - \int_0^1 tg(t)dt - \psi_2'(s)\int_0^1 \psi_2(t)g(t)dt$$
$$= \int_s^1 g(t)dt - \int_0^1 tg(t)dt - \psi_2'(s)\int_0^1 \psi_2(t)g(t)dt \tag{B.5}$$

for λ-almost every $s \in (0,1)$. Now if $\int_0^1 \psi_2(t)g(t)dt = 0$, it follows from (B.3) that

$$0 = \int_0^1 (s \wedge t - s \cdot t)g(t)dt \qquad \lambda\text{-a.e.}$$

But since the integral operator with kernel $k(s,t) = s \wedge t - s \cdot t$ is injective, g has to be zero a.e. Hence, $g \neq 0$ a.e. implies $\int_0^1 \psi_2(t)g(t)dt \neq 0$. Now suppose that $g \neq 0$ a.e. Then by (B.5) we have

$$\psi_2'(s) = \Big(\int_s^1 g(t)dt - \int_0^1 tg(t)dt\Big) \cdot \Big(\int_0^1 \psi_2(t)g(t)dt\Big)^{-1},$$

so that

$$\lim_{s\to 0} \psi_2'(s) = \Big(\int_0^1 g(t)dt - \int_0^1 tg(t)dt\Big) \cdot \Big(\int_0^1 \psi_2(t)g(t)dt\Big)^{-1} > -\infty,$$

which contradicts $\lim_{s\to 0} \psi_2'(s) = -\infty$. It follows from this that $g = 0$ λ-a.e., which means that T_{ρ_2} is injective. $\square$

Acknowledgment

I would like to express my sincere gratitude to Prof. Dr. Erich Häusler for his help and guidance during the preparation of this thesis.

Moreover, I would like to thank my family and friends for their constant support and encouragement.

References

[1] R. R. Bahadur, *Stochastic comparison of tests*, Ann. Math. Statist. 31 (1960) 276–295

[2] R. R. Bahadur, *Rates of convergence of estimates and test statistics*, Ann. Math. Statist. 38 (1967) 303–324

[3] R. Beran, *Tail probabilities of noncentral quadratic forms*, Ann. Statist. 3 (4) (1975) 969–974

[4] P. J. Bickel and D. A. Freedman, *Some asymptotic theory for the bootstrap*, Ann. Statist. 9 (6) (1981) 1196–1217

[5] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, and J. A. Wellner, *Efficient and adaptive estimation for semiparametric models*, Springer-Verlag, New York, 1998, reprint of the 1993 original

[6] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*, Springer Series in Statistics, second edn., Springer-Verlag, New York, 1991

[7] J. Durbin, *Weak convergence of the sample distribution function when parameters are estimated*, Ann. Statist. 1 (1973) 279–290

[8] F. Eicker, *Central limit theorems for families of sequences of random variables.*, Ann. Math. Statist. 34 (1963) 439–446

[9] W. Feller, *An introduction to probability theory and its applications. Vol. II*, John Wiley & Sons Inc., New York, 1966

[10] M. Genz, *Schätzung der Fehlerverteilung in stabilen autoregressiven Prozessen erster Ordnung*, Diploma thesis, Justus-Liebig-Universität Gießen (2001)

[11] M. Genz, *Anwendungen des Empirischen Likelihood-Schätzers der Fehlerverteilung in AR(1)-Prozessen*, Dissertation, Justus-Liebig-Universität Gießen (2004)

[12] M. Genz and E. Haeusler, *Empirical processes with estimated parameters under auxiliary information*, J. Comput. Appl. Math. 186 (1) (2006) 191–216

[13] G. G. Gregory, *On efficiency and optimality of quadratic tests*, Ann. Statist. 8 (1) (1980) 116–131

[14] E. Haeusler, *An exact rate of convergence in the functional central limit theorem for special martingale difference arrays*, Z. Wahrsch. Verw. Gebiete 65 (4) (1984) 523–534

[15] E. S. Hörmann, *The Cramér-von Mises Test for Centered Distributions*, Diploma thesis, Justus-Liebig-Universität Gießen (2007)

[16] W. C. M. Kallenberg and A. J. Koning, *On Wieand's theorem*, Statist. Probab. Lett. 25 (2) (1995) 121–132

[17] S. Kiwitt, E.-R. Nagel, and N. Neumeyer, *Empirical likelihood estimators for the error distribution in nonparametric regression models*, Math. Methods Statist. 17 (3) (2008) 241–260

[18] H. L. Koul, *Weighted empirical processes in dynamic nonlinear models*, vol. 166 of *Lecture Notes in Statistics*, Springer-Verlag, New York, 2002, second edition of *Weighted empiricals and linear models* [Inst. Math. Statist., Hayward, CA, 1992]

[19] W. Linde, *Gaussian measure of large balls in* l_p, Ann. Probab. 19 (3) (1991) 1264–1279

[20] S. Nadarajah, *A generalized normal distribution*, J. Appl. Stat. 32 (7) (2005) 685–694

[21] Y. Nikitin, *Asymptotic Efficiency of Nonparametric Tests*, Cambridge University Press, 1995

[22] A. B. Owen, *Empirical likelihood ratio confidence intervals for a single functional*, Biometrika 75 (2) (1988) 237–249

[23] A. B. Owen, *Empirical likelihood ratio confidence regions*, Ann. Statist. 18 (1) (1990) 90–120

[24] A. B. Owen, *Empirical likelihood for linear models*, Ann. Statist. 19 (4) (1991) 1725–1747

[25] A. B. Owen, *Empirical Likelihood*, Chapman & Hall/CRC, 2001

[26] D. Pollard, *Convergence of stochastic processes*, Springer Series in Statistics, Springer-Verlag, New York, 1984

[27] J. Qin and J. Lawless, *Empirical likelihood and general estimating equations*, Ann. Statist. 22 (1) (1994) 300–325

[28] J. R. Schott, *Matrix analysis for statistics*, Wiley Series in Probability and Statistics: Applied Probability and Statistics, John Wiley & Sons Inc., New York, 1997

[29] M. T. Wells, *The relative efficiency of goodness-of-fit statistics in the simple and composite hypothesis-testing problem*, J. Amer. Statist. Assoc. 85 (410) (1990) 459–463

[30] H. S. Wieand, *A condition under which the Pitman and Bahadur approaches to efficiency coincide*, Ann. Statist. 4 (5) (1976) 1003–1011

[31] B. Zhang, *Estimating a distribution function in the presence of auxiliary information*, Metrika 46 (3) (1997) 221–244