Resampling in the frequency domain of time series to determine critical values for change-point tests*

Claudia Kirch†

2008

Abstract

We study a model with an abrupt change in the mean and dependent errors that form a linear process. Different kinds of statistics are considered, such as maximum-type statistics (particularly different CUSUM procedures) or sum-type statistics. Approximations of the critical values for change-point tests are obtained through permutation methods in the frequency domain. The theoretical results show that the original test statistics and their corresponding frequency permutation counterparts follow the same distributional asymptotics. The main step in the proof is to obtain limit theorems for the corresponding rank statistics and then deduce the permutation asymptotics conditionally on the given data.

Some simulation studies illustrate that the permutation tests usually behave better than the original tests if performance is measured by the α- and β-errors respectively.

Keywords: Permutation principle, change in mean, rank statistic, dependent observations, linear process, frequency domain

AMS Subject Classification 2000: 62G10, 62G09, 62M15

1. Introduction

Recently a number of papers has been published on the use of bootstrapping or permutation methods for obtaining reasonable approximations to the critical values of change-point tests. This approach was first suggested by Antoch and Hušková [1] and later pursued by other authors (cf. Hušková [17] for a recent survey). So far, it has mostly been dealt with independent observations, yet in many situations dependent observations are much more realistic. Kirch and Steinebach [21] considered change-point tests

*This is a preprint of an article that will be published in *Statistics and Decision*.

†Technical University Kaiserslautern, Departement of Mathematics, Erwin-Schrödinger-Str., 67663 Kaiserslautern, ckirch@mathematik.uni-kl.de

for possibly dependent processes under strong invariance. In that situation we were close enough to the independent case for the usual procedure to work.

Here, we are interested in a variation of the change-point problem with a possible change in the mean and i.i.d. errors. Since in many situations the independency is too strong an assumption, Horváth [16] and Antoch et al. [2] investigated the model where the errors are no longer independent but form a linear process. Kirch [20] examined the use of a block permutation method to obtain critical values for that model. This particular bootstrap scheme was used to take account of the dependency. There is, however, one major drawback: Simulations conducted in Kirch [20] show that the performance of the block bootstrap depends strongly on a good choice of the block-length. Automatic procedures are not available. There is only few literature on the optimal choice of the block-length and only for few situations (confer e.g. Hall et al. [14]), yet even then the optimal block-length depends on unknown parameters.
This is why we are interested in yet another bootstrap scheme in this paper. We propose bootstrapping the Fourier coefficients in the frequency domain. This is motivated by the fact that Fourier coefficients are asymptotically independent in a certain sense.

Several authors have proposed bootstrapping in the frequency domain. Franke and Härdle [11], e.g., propose bootstrapping kernel spectral density estimates based on resampling from the periodogram of the original data. The idea behind that approach is that a random vector of the periodograms of finitely many frequencies is approximately independent and exponentially distributed. Later this approach was also pursued for different models, e.g. for ratio statistics such as autocorrelations by Dahlhaus and Janas [9] or in regression models by Hidalgo [15].
In the above papers the estimation problem as a whole was transformed into the frequency domain. As a contrast we backtransform the bootstrapped coefficients and look at the new sequences – back in the time domain – as new pseudo-time series. Then we construct the estimator using empirical distribution functions.

The paper is organized as follows: In a first section we describe the model, frequently used statistics as well as their null asymptotics. Then we explain the details of the resampling procedure we consider. In Section 4 we state the main result, which shows that the resampling procedure gives asymptotically correct critical values. Then we give a short extract of a simulation study in Section 5. To prove the validity of the procedure we first need some results for the corresponding rank statistic, which we state and prove in Section 6. In a final section we can use these to derive the asymptotics of the permutation statistic. We give some auxiliary results about trigonometric functions in an appendix.

2. Model and Null Asymptotic

We consider the following At-Most-One-Change (AMOC) location model

$$X(i) = \mu + d\,1_{\{i>m\}} + e(i), \quad 1 \leqslant i \leqslant n, \tag{2.1}$$

where the errors $\{e(i), 1 \leqslant i \leqslant n\}$ are given by the linear process

$$e(i) = \sum_{j \geqslant 0} w_j\, \epsilon(i-j)$$

and $m = m(n)$, $d = d(n)$ may depend on n. We are interested in testing the null hypothesis of "no change"

$$H_0 : \quad m = n$$

against the alternative of a change in the mean

$$H_1 : \quad 1 \leqslant m < n \text{ and } d \neq 0.$$

Moreover we assume that the innovations $\{\epsilon(i) : -\infty < i < \infty\}$ are i.i.d. random variables with

$$\mathrm{E}\,\epsilon(i) = 0, \quad 0 < \sigma^2 = \mathrm{E}\,\epsilon(i)^2 < \infty, \quad \mathrm{E}\,|\epsilon(i)|^4 < \infty. \tag{2.2}$$

We suppose that the weights $\{w_j : j \geqslant 0\}$ satisfy

$$\sum_{j\geqslant 0} w_j \neq 0, \quad \sum_{j\geqslant 0} \sqrt{j}\,|w_j| < \infty. \tag{2.3}$$

In the case where the errors $\{e(i) : 1 \leqslant i \leqslant n\}$ are a sequence of i.i.d. random variables (with mean zero and finite variance) one often uses statistics based on the partial sums $S_k = \sum_{i=1}^{k}(X_i - \bar{X}_n)$ where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
It turns out that these statistics work also very well in the present setup (confer Antoch et al. [2] and Horváth [16]).
Typical test statistics are

$$T_n^{(1)}(q) = \max_{1\leqslant k<n}\left(\frac{1}{\sqrt{n}\,q(\frac{k}{n})}\,|S_k|\right), \qquad T_n^{(2)}(r) = \frac{1}{n}\sum_{k=1}^{n-1}\frac{1}{r(k/n)}\left(\frac{1}{\sqrt{n}}S_k\right)^2,$$
$$T_n^{(3)}(\epsilon) = \max_{\epsilon n\leqslant k\leqslant(1-\epsilon)n}\left(\sqrt{\frac{n}{k(n-k)}}\,|S_k|\right),$$

where $q(\cdot)$ and $r(\cdot)$ are weight functions defined on $(0,1)$ and $\epsilon > 0$.
We assume that the weight function q belongs to the class

$$Q_{0,1} = \{q \,:\, q \text{ is non-decreasing in a neighborhood of zero, non-increasing in a neighborhood of one and } \inf_{\eta\leqslant t\leqslant 1-\eta} q(t) > 0 \text{ for all } 0 < \eta < 1/2\}.$$

The following integral plays a crucial role for the convergence of statistics based on a weight function q

$$I^*(q,c) = \int_0^1 \frac{1}{t(1-t)}\exp\left\{\frac{-cq^2(t)}{t(1-t)}\right\} dt.$$

For details and further references confer Csörgő and Horváth [7], Chapter 4.
We assume that the weight function r fulfills for all $x \in (0,1)$

$$r(x) > 0 \quad \text{and} \quad \int_0^1 \frac{t(1-t)}{r(t)}\,dt < \infty. \tag{2.4}$$

Details and further references on the sum statistic $T_n^{(2)}(r)$ can be found in Csörgő and Horváth [8], Chapter 2.

The following theorem gives the asymptotic distribution under H_0 for the above statistics.

Theorem 2.1. *Assume that* (2.1) - (2.3) *and* H_0 *holds.*

a) If $q \in Q_{0,1}$ *and* $I^*(q,c) < \infty$ *for some* $c > 0$*, then*

$$\frac{1}{\widehat{\tau}}\, T_n^{(1)}(q) \xrightarrow{\mathcal{D}} \sup_{0<t<1} \frac{|B(t)|}{q(t)} \qquad as\ n \to \infty.$$

b) If r *fulfills condition* (2.4)*, then*

$$\frac{1}{\widehat{\tau}^2}\, T_n^{(2)}(r) \xrightarrow{\mathcal{D}} \int_0^1 \frac{B^2(t)}{r(t)}\, dt \qquad as\ n \to \infty.$$

c) We have for all $\epsilon > 0$ *and* $x \in \mathbb{R}$

$$\frac{T_n^{(3)}(\epsilon)}{\widehat{\tau}} \xrightarrow{\mathcal{D}} \sup_{\epsilon \leqslant t \leqslant 1-\epsilon} \left| \frac{1}{\sqrt{t(1-t)}} B(t) \right| \qquad as\ n \to \infty.$$

Here $\widehat{\tau}$ *is an estimator for* $\tau = \sigma \left|\sum_{s\geqslant 0} w_s\right|$ *with* $\widehat{\tau} - \tau = o_P(1)$ *and* $\{B(t) : 0 \leqslant t \leqslant 1\}$ *is a Brownian bridge.*

Proof. Confer Theorem 2.1 in Antoch et al. [2], where it is assumed $\sum_{j\geqslant 0} j\, |w_j| < \infty$. Yet this assumption is only needed to obtain the Beveridge-Nelson decomposition, which also holds under (2.3), confer Phillips and Solo [24], equation (15) et sqq. ■

Remark 2.1. A popular choice for the weight function $q(\cdot)$ is $q(t) = (t(1-t))^\gamma, 0 \leqslant \gamma < 1/2$. For $\gamma = 1/2$ we no longer have distributional convergence but get an extreme-value behavior. γ close to 0 detects changes in the middle of the observation period better, whereas $\gamma = 1/2$ detects early and late changes more easily. It is also interesting that the statistics for general $q(\cdot)$ and for $\gamma = 1/2$ are asymptotically independent (cf. e.g. Antoch et al. [2]).
Note that the change-point estimators in (3.2) are closely related to this class of test statistics.

The above theorem allows us to construct asymptotic tests where we choose the quantiles of the limit distributions as critical values. However, there are several problems with this approach. First the limit distributions of $T_n^{(1)}(q)$ and $T_n^{(2)}(r)$ are explicitly only known for $q, r \equiv 1$. Secondly the convergence can be rather slow (especially for extreme value statistics). But most importantly the limit distribution is the same no matter what the underlying dependency structure is. Yet small sample simulations (cf. Kirch [18]) show that the exact critical values depend strongly on the dependency structure.
This is why we are interested in alternative methods to obtain critical values for the test, especially the block resampling method, that has been discussed in Kirch [20], and the frequency resampling method that we discuss in this paper.

3. The frequency resampling procedure

Due to limitations of space we focus our attention in the following to the case where n is even. From a practical point of view this is the more important case, since then the

Fast Fourier Transform is much more effective. Yet the proofs for n odd are analogous, for details confer Kirch [18], Chapter 4.

In this section we explain the idea behind the resampling procedure in the frequency domain and give a thorough description of the algorithm.

First Step: Stationarization

First we need to get closer to the actual error sequence which we know forms a linear process. So we use estimators for the change-point and the mean before and after the change and can thus estimate the error sequence:

$$\widetilde{e}(i) := (X(i) - \widehat{\mu}_1)1_{[1,\widehat{m}]}(i) + (X(i) - \widehat{\mu}_2)1_{[\widehat{m}+1,n]}(i), \tag{3.1}$$

where for some arbitrary $0 \leqslant \gamma \leqslant \frac{1}{2}$

$$\begin{aligned} &\widehat{m} = \widehat{m}(\gamma) = \min(\arg\max(|S_k(\gamma)|, k = 1, \ldots, n-1)), \\ &\qquad \text{where } S_k(\gamma) = \left(\frac{n}{k(n-k)}\right)^{\gamma} \sum_{i=1}^{k} (X(i) - \bar{X}_n), \\ &\widehat{\mu}_1 = \widehat{\mu}_1(\gamma) = \frac{1}{\widehat{m}} \sum_{i=1}^{\widehat{m}} X(i), \quad \widehat{\mu}_2 = \widehat{\mu}_2(\gamma) = \frac{1}{n-\widehat{m}} \sum_{i=\widehat{m}+1}^{n} X(i), \\ &\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X(i). \end{aligned} \tag{3.2}$$

Second Step: Fourier Transform and Resampling of Fourier Coefficients

Compute the Fourier coefficients of $\{\widetilde{e}(i) : 1 \leqslant i \leqslant n\}$:

$$\omega(j) := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \widetilde{e}(k) \exp(-2\pi i jk/n).$$

It is known that Fourier coefficients of some time series such as linear processes are – in a certain sense – asymptotically (complex) normally distributed and independent with mean 0 and variance $2\pi f(\lambda_j)$; f is the spectral density of e; $\lambda_j = 2\pi i j/n$ (for details confer e.g. Brillinger [5], Theorem 4.4.1, p. 94). Brockwell and Davis [6], Theorem 10.3.1, give a similar result for the periodograms. This is less general because it specifically deals with linear processes fulfilling (2.2) and (2.3) but the conditions on the weights and moments are less stringent. Our proofs are closely related to theirs.
Let $g(1) := \mathrm{Re}(\omega(1)), g(2) := \mathrm{Im}(\omega(1)), \ldots, g(\tilde{n}-1) := \mathrm{Re}\left(\omega\left(\frac{\tilde{n}}{2}\right)\right), g(\tilde{n}) := \mathrm{Im}\left(\omega\left(\frac{\tilde{n}}{2}\right)\right)$, where $\tilde{n} := n-2$. We bootstrap the centered coefficients $g(i) - \frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} g(j)$ either with or without replacement. Here, we focus on the bootstrap without replacement, since it is the more difficult case. The proof of the validity of the bootstrap with replacement is very similar, for details confer Kirch [18], Section 4.7.
Since $\omega(n-j) = \overline{\omega(j)}$ we set the bootstrap coefficients for $j > n/2$ in the corresponding way. The bootstrap coefficient for $j = n/2$ is set to 0. Note that $\omega(n)$ is the mean of the

sequence (both before and after bootstrapping). Since all the statistics we use center the input sequence, the mean is irrelevant, so we set the bootstrap coefficient for $j = n$ equal to 0.

Remark 3.1. The covariance structure of the original sequence is coded in the variances of the coefficients. Resampling in the above way will destroy that, but correspond to a similar sequence with independent errors and variance $\sigma^2 \sum w_s^2$, where w_s are the weights of the given linear process. This is why we still need an estimator for $\sigma^2 \left(\sum w_s\right)^2$ in order to use the critical values of the permutation statistic for the null hypothesis.

Third Step: Backtransformation

Use inverse Fourier transform to obtain a similar sample as the original one:

$$X_{\mathbf{R}}(l) := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \omega_{\mathbf{R}}(k) \exp(2\pi i l k/n), \tag{3.3}$$

where – as described above –

$$\omega_{\mathbf{R}}(l) = g(R_l) - \bar{g} + i\left(g(R_{\tilde{n}+1-l}) - \bar{g}\right), \quad l = 1, \ldots, \frac{\tilde{n}}{2},$$

$\bar{g} = \frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} g(j)$. Moreover $\omega_{\mathbf{R}}(n) = 0$, $\omega_{\mathbf{R}}(n/2) = 0$ (for n even);
$\omega_{\mathbf{R}}(n-l) = \overline{\omega_{\mathbf{R}}(l)}$, $l = 1, \ldots, \frac{\tilde{n}}{2}$, the conjugated complex of $\omega_{\mathbf{R}}(l)$.
Here, $\mathbf{R} = (R_1, \ldots, R_{\tilde{n}})$ is a random permutation of $(1, \ldots, \tilde{n})$.

Now calculate the value of the chosen statistic for the sample $X_{\mathbf{R}}$. This seems to work in simulations. Our proof, however, only holds true if we just use $N := \frac{n}{\alpha(n)}$ of the n components of $X_{\mathbf{R}}$ where $\alpha(n) \to \infty$, no matter how slowly, and $N \to \infty$. It is also not important which ones we use, although it is reasonable to use successive ones. For $\alpha(n) = 1$ the Noether (and thus the Lindeberg) condition is not fulfilled. For a detailed discussion we refer to Kirch [18], Chapter 4.8.2.

The permutation statistic is then standardized with the exact variance of the corresponding rank statistic, confer also Remark 4.2, i.e.

$$\frac{2}{\tilde{n}} \sum_{l=1}^{\tilde{n}} \left(g(l) - \frac{1}{\tilde{n}} \sum_{k=1}^{\tilde{n}} g(l) \right)^2 .$$

This is essentially the factor needed to standardize the Fourier coefficients $g(\cdot)$. The factor 2 is needed because we use each of these coefficients twice (in $\omega_{\mathbf{R}}(\cdot)$ as well as $\omega_{\mathbf{R}}(n-\cdot)$)

Remark 3.2. It is usually (there are some additional conditions on the weight functions q and r) also possible to use the uncentered Fourier coefficients or keep the middle term, for details confer Kirch [18].

Forth Step: Calculation of the Critical Values

Repeat the second and third step t times and calculate the α-quantile of the statistic based on these t "realizations".

We reject the null hypothesis if the value of the statistic for the original sample (here we have to divide by the asymptotic variance $\sigma^2 \left(\sum w_s\right)^2$ or an appropriate estimate) is larger than the above α-quantile.

4. Main result: Asymptotics of the Permutation Statistics

The main theorem below states that the critical values obtained by the frequency permutation method as proposed in the previous section are asymptotically correct. Precisely it states that the quantiles we obtain from the permutation method are asymptotically the same as the ones corresponding to the distribution of the original statistic under the null hypothesis. Thus, even if our observations follow an alternative, we get a good approximation of the critical values corresponding to the null distribution.
We would like to point out that – under alternatives – this is not necessary for the test to be consistent. In fact under quite general assumptions one can show (cf. e.g Csörgő and Horváth [8]) that the classical CUSUM statistic, for example, converges stochastically to infinity with rate $|d|\, m(n-m)/n$. On the other hand it is possible to show under some mild moment conditions that the critical values obtained by the permutation statistic converge at most logarithmically to infinity, so that the frequency permutation test remains consistent for a much broader class of alternatives than what we consider in this paper.

For notational reasons choose $\alpha(n)$ such that $\frac{n}{\alpha(n)}$ is an integer.

Some simple calculations (for details cf. Kirch [18], Section 4.3) show that the the resampled time series (3.3) is equal to (with a different permutation leading to this representation)

$$X_{\mathbf{R}}(s) = \frac{2}{n}\sum_{l=1}^{n-2} c_s(l)\left(\sum_{j=1}^{n}\widetilde{e}(j)c_j(R_l) - \frac{1}{n-2}\sum_{k=1}^{n-2}\sum_{j=1}^{n}\widetilde{e}(j)c_j(k)\right), \tag{4.1}$$

where $\mathbf{R} = (R_1, \ldots, R_{n-2})$ is a random permutation of $(1, \ldots, n-2)$ and

$$\begin{aligned} c_j &= (0, 0, c_j(1), \ldots, c_j(n-2))^T \\ &= (0, 0, \cos(2\pi j/n), \sin(-2\pi j/n), \cos(2\cdot 2\pi j/n), \ldots, \sin(-(n/2-1)\cdot 2\pi j/n))^T, \end{aligned} \tag{4.2}$$

$j = 1, \ldots, n$, $c_j(R) = (0, 0, c_j(R_1), \ldots, c_j(R_{n-2}))^T$.

In order to prove the validity of the procedure given in Section 3 it thus suffices to investigate

$$Z_n^{\mathbf{X}}(u, \mathbf{R}) = \sqrt{\frac{1}{N}}\sum_{s\leqslant Nu} X_{\mathbf{R}}(\beta(s)), \quad \text{for } u = \frac{\alpha(n)}{n}, \frac{2\alpha(n)}{n}, \ldots, 1.$$

Let $Z_n^{\mathbf{X}}(0, \mathbf{R}) = 0$ and $Z_n^{\mathbf{X}}(t, \mathbf{R})$ be linearly interpolated between $(i-1)/n$ and i/n for $i = 1, \ldots, n$. The function β defines which $\frac{n}{\alpha(n)}$ of the n resampled variables we choose, e.g. $\beta(s) = s$.

Theorem 4.1. *Let (2.2) and (2.3) be fulfilled, $\alpha(n) \to \infty$. Let $m = \lfloor \vartheta n \rfloor$, $0 < \vartheta \leqslant 1$, and $d = d(n)$ bounded. Then*

$$P\left(f[(Z_n^{\mathbf{X}}(id, \mathbf{R}) - id\, Z_n^{\mathbf{X}}(1, \mathbf{R}))/\widehat{\sigma}_n] \le x \,\middle|\, X(1), \ldots, X(n)\right) \xrightarrow{P} P\left(f(B(\cdot)) \leqslant x\right)$$

for all continuous $f : C[0,1] \to \mathbb{R}$ and for all $x \in \mathbb{R}$. Here

$$\widehat{\sigma}_n^2 := \frac{2}{n\tilde{n}} \sum_{l=1}^{\tilde{n}} \left(\sum_{i=1}^{n} \widetilde{e}(i) c_i(l) - \frac{1}{\tilde{n}} \sum_{k=1}^{\tilde{n}} \sum_{j=1}^{n} \widetilde{e}(j) c_j(k) \right)^2 .$$

The following corollary states the limit behavior of our statistics of interest.

Corollary 4.1. *Under the conditions of Theorem 4.1 the following holds:*

a) If $q \in FC_0^1$ and $\int_0^1 \frac{1}{q^2(t)}\,dt < \infty$, then it holds

$$P\left(T_n^{(1f)}(\mathbf{R}, q) \leqslant x \,|\, X(1), \ldots, X(n)\right) \xrightarrow{P} P\left(\sup_{0 \leqslant t \leqslant 1} \frac{|B(t)|}{q(t)} \leqslant x\right),$$

$$\text{where } T_n^{(1f)}(\mathbf{R}, q) := \max_{1 \leqslant k < N} \frac{1}{q\left(\frac{k}{N}\right)} \left| Z_n^{\mathbf{X}}\left(\frac{k}{N}, \mathbf{R}\right) - \frac{k}{N} Z_n^{\mathbf{X}}(1, \mathbf{R}) \right| .$$

b) For $\int_0^1 \frac{(t(1-t))^s}{r(t)}\,dt < \infty$ for some $0 \leqslant s < 1$, it holds

$$P\left(T_n^{(2f)}(\mathbf{R}) \leqslant x \,|\, X(1), \ldots, X(n)\right) \xrightarrow{P} P\left(\int_0^1 \frac{B^2(t)}{r(t)}\,dt \leqslant x\right),$$

$$\text{where } T_n^{(2f)}(\mathbf{R}, r) := \int_0^1 \frac{1}{r(t)} |Z_n^{\mathbf{X}}(t, \mathbf{R}) - t Z_n^{\mathbf{X}}(1, \mathbf{R})|^2\,dt.$$

c) For all $\epsilon > 0$ we get for all $x \in \mathbb{R}$

$$P\left(T_n^{(3f)}(\mathbf{R}) \leqslant x \,|\, X(1), \ldots, X(n)\right) \xrightarrow{P} P\left(\sup_{\epsilon \leqslant t \leqslant 1-\epsilon} \frac{|B(t)|}{(t(1-t))^{1/2}} \leqslant x\right),$$

$$\text{where } T_n^{(3f)}(\mathbf{R}) := \sup_{\epsilon \leqslant t \leqslant 1-\epsilon} \sqrt{\frac{1}{t(1-t)}} |Z_n^{\mathbf{X}}(t, \mathbf{R}) - t Z_n^{\mathbf{X}}(1, \mathbf{R})|.$$

Here $\{B(t) : 0 \leqslant t \leqslant 1\}$ denotes a Brownian bridge.

Remark 4.1. For some variations of the above algorithm such as not centering the Fourier coefficients as well as the proof for a somewhat broader class of alternatives or estimators (for the change-point, the mean before and the mean after the change) confer Kirch [18], Chapter 4. The validity of the bootstrap with replacement is also given there. Furthermore it is shown that the results remain true if the innovation sequence has only existing moments larger than the second one, yet the assumptions on $\alpha(\cdot)$ need to be somewhat strengthened.

Remark 4.2. Theorem A.1 yields after some calculations

$$\frac{2}{n\tilde{n}} \sum_{l=1}^{\tilde{n}} \left(\sum_{j=1}^{n} \widetilde{e}(j) c_j(l) \right)^2$$
$$= \frac{1}{\tilde{n}} \sum_{j=1}^{n} \widetilde{e}(j)^2 - \frac{2}{n\tilde{n}} \left(\sum_{j=1}^{n/2} \widetilde{e}(2j) \right)^2 - \frac{2}{n\tilde{n}} \left(\sum_{j=0}^{n/2-1} \widetilde{e}(2j+1) \right)^2 .$$

The proof of Theorem 4.1 shows that the last two terms converge to 0 in a P-stochastic sense and $\frac{1}{n^{3/2}}\sum_l\sum_j \widetilde{e}(j)c_j(l) = o_P(1)$. Theorem 3.7 of Phillips and Solo [24] and the proof of Theorem 4.1 state that $\frac{1}{\tilde{n}}\sum_{j=1}^n \widetilde{e}^2(j) \to \sigma^2\sum_{s\geqslant 0} w_s^2$ almost surely as $n\to\infty$. Together this shows

$$\frac{2}{n\tilde{n}}\sum_{l=1}^{\tilde{n}}\Big(\sum_{i=1}^{n}\widetilde{e}(i)c_i(l) - \frac{1}{\tilde{n}}\sum_{k=1}^{\tilde{n}}\sum_{j=1}^{n}\widetilde{e}(j)c_j(k)\Big)^2 \xrightarrow{P} \sigma^2\sum_{s\geqslant 0} w_s^2$$

under the same assumptions as in Theorem 4.1.
Thus we standardize the permutation statistic asymptotically with $\sigma\sqrt{\sum w_s^2}$. On the other hand the original statistic is asymptotically standardized with $\tau = \sigma|\sum w_s|$.

5. Simulations

In the previous chapter we have shown that the frequency bootstrap yields asymptotically correct critical values, yet the question remains how well it performs for small samples. Moreover we are specifically interested in how well it performs for small samples sizes in comparison to the asymptotic test.

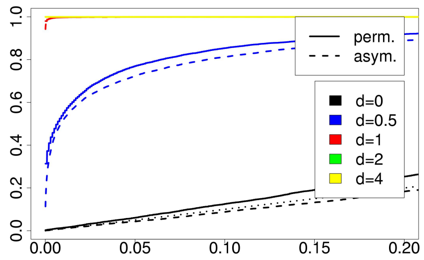
Due to limitation of space and similarity of results we only provide some results for $T_n(q_1)$, $q_1 \equiv 1$. A much more extensive simulation study including tables providing the critical values, the results for $q_2(t) = (t(1-t))^{1/4}$ as well as for the other statistics and double exponentially distributed innovations can be found in Kirch [19], confer also Section 6.2 of Kirch [18].

In the simulation study we use the model of Section 2, where $\{e(i) : i \geqslant 1\}$ forms an AR(1) sequence with autoregressive coefficient $\rho \in \{-0.5, -0.3 : 0.3, 0.5, 0.7\}$ and $\{\epsilon(j) : -\infty < j < \infty\}$ are i.i.d. $N(0,1)$, hence $\tau = \frac{1}{1-\rho}$. Sample size is 80, the change-point is at 40. Then we use the algorithm as described in Section 3 with 60 respectively 70 bootstrap r.v.´s, i.e. $\frac{n}{\alpha(n)} = 60$ respectively $= 70$. Here, we take the first 60 respectively first 70. Moreover we use all 80 r.v.´s, which corresponds to $\alpha(n) = 1$. We use the estimators with $\gamma = 0$.

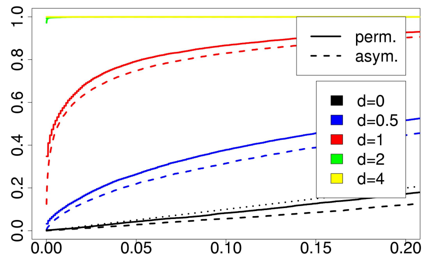
Since we are interested in the performance of the procedure developed here rather than in the performance of some variance estimator, we use the actual variance in the fourth step for both tests.

The goodness of a test is usually determined by two criteria. First and foremost we want the test to hold the nominal level (α-error). This can be very wrong in small samples for tests which are only asymptotically consistent. Secondly the power of the test should be as large as possible.

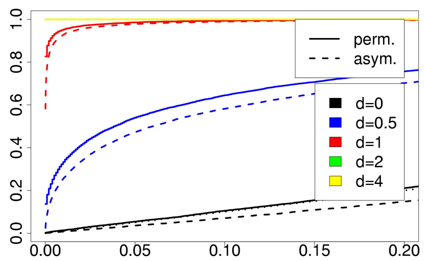
To visualize these two qualities for the asymptotic as well as the frequency permutation test we create size-power curves under the null hypothesis as well as under various alternatives. Size-power-curves are plots of the empirical distribution function of the p-values of the statistic $T_n(q_1)$ for the null hypothesis respectively given alternatives with respect to the distribution used to determine the critical values of the test.
What we get is a plot that shows the empirical size and power (i.e. the empirical α–errors resp. $1-(\beta$–errors)) on the y-axis for the chosen level on the x-axis. So, the graph for



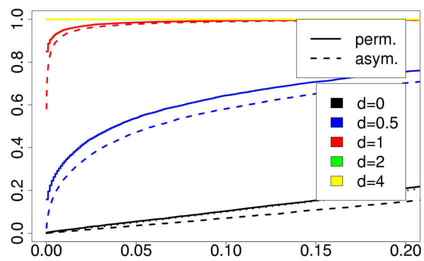
(1) $\rho = -0.3$, $N = 80$



(2) $\rho = 0.3$, $N = 80$



(3) $\rho = 0$, $N = 60$



(4) $\rho = 0$, $N = 80$

Figure 5.1: SPC-plots for $T_n^{(3f)}(\mathbf{R}, q_1)$ for $n = 80$, $m = 40$ and different values for ρ and N

the null hypothesis should be close to the diagonal (which is given by the dotted line) and for the alternatives it should be as steep as possible.

The curves can be found in Figure 5.1. Note that the solid line gives the result for the frequency permutation test and the dashed line for the asymptotic test. From bottom to top lines belong to the null hypothesis ($d = 0$) and different alternatives ($d = 0.5, 1, 2, 4$). The alternatives not visible in the picture have a power of 1. The plots look very similar, no matter what the value of α respectively N is. This remains true for a different selection function β. To illustrate this point we give all two plots for $\rho = 0$ (yet this is true for every other choice of ρ as well). For a selection of other values of ρ we then only give the plot where we use all n values. To create the SPC-plots we use 10 000 time series according to the model (null hypothesis as well as alternatives) and for each of these 1 000 permutations.

The simulations show that the frequency permutation test holds the chosen level usually better than the asymptotic test. Taking different actual levels into account the size of both tests is comparable.

Considering the results given in Section 4 we would also like to see how well the frequency permutation distribution for a given observation sequence actually fits the exact null distribution. For the performance of the test this is not essential, because the value of the statistic of an observation sequence is only compared to the quantiles of the permutation statistic for that same underlying sequence. Still it is interesting to verify the assertions of Section 4 for small samples.

Therefore we create QQ-plots of the permutation distribution of one specific realization against the (empirical) distribution of the statistic under H_0.

To obtain the empirical distribution function we use 1 000 realizations of H_0 respectively 1 000 permutations. Again the plots are very similar for different choices of $\alpha(\cdot)$. The results can be found in Figure 5.2.

We would also like to mention that the frequency permutation distribution is already quite stable for different observation sequences. That means we get essentially the same plot for different observations.

6. Asymptotics of the Corresponding Rank Statistics

The derivation of the permutation asymptotics is based on the corresponding results of rank asymptotics. In this section we analyze the corresponding frequency rank statistics. In the proofs we use simple linear rank statistic results of Hájek et al. [13].

The frequency rank statistics are essentially determined by $Z_n^{\mathbf{x}}(u,\mathbf{R}) := \widetilde{Z}_n^{\mathbf{x}}(u,\mathbf{R}) - \mathrm{E}\,\widetilde{Z}_n^{\mathbf{x}}(u,\mathbf{R})$, where

$$\begin{aligned}\widetilde{Z}_n^{\mathbf{x}}(u,\mathbf{R}) &= \frac{2\sqrt{\alpha(n)}}{n^{3/2}} \sum_{l=1}^{\tilde{n}} \sum_{s \leqslant \frac{n}{\alpha(n)} u} c_{\beta(s)}(l) \sum_{j=1}^{n} x_{j,n} c_j(R_l) \\ &= \sqrt{\frac{\alpha(n)}{n}} \sum_{s \leqslant \frac{n}{\alpha(n)} u} x_{\mathbf{R}}(\beta(s))\end{aligned} \tag{6.1}$$

for $u = j\frac{\alpha(n)}{n}$ and linearly interpolated in between. $\{x_{j,n} : j = 1, \dots, n, n \geqslant 1\}$ are some scores.

We assume that the scores fulfill the following conditions:

$$\begin{aligned}&\frac{2}{n\tilde{n}} \sum_{l=1}^{\tilde{n}} \left(\sum_{i=1}^{n} x_{i,n} c_i(l) - \frac{1}{\tilde{n}} \sum_{k=1}^{\tilde{n}} \sum_{j=1}^{n} x_{j,n} c_j(k) \right)^2 = 1 \\ &\frac{1}{n} \sum_{l=1}^{\tilde{n}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_{i,n} c_i(l) - \frac{1}{\tilde{n}} \sum_{k=1}^{\tilde{n}} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} x_{j,n} c_j(k) \right|^4 = o(\alpha(n)).\end{aligned} \tag{6.2}$$

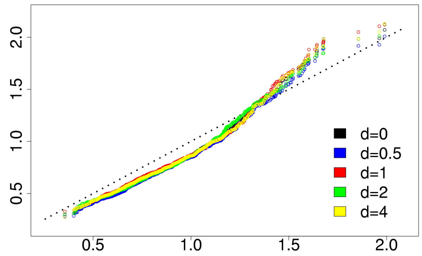
Theorem 6.1. *Under conditions* (6.2) *and if* $\alpha(n) \to \infty$,*then*

$$\{Z_n^{\mathbf{x}}(u,\mathbf{R}) : 0 \leqslant u \leqslant 1\} \xrightarrow{C[0,1]} \{W(u) : 0 \leqslant u \leqslant 1\},$$

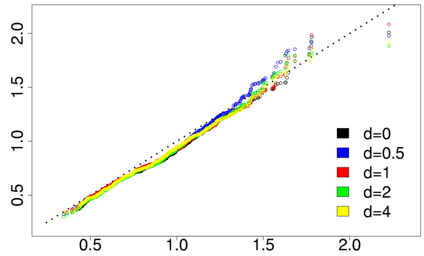
where $\{W(u) : 0 \leqslant u \leqslant 1\}$ *is a standard Wiener process.*

Proof. This follows immediately from Billingsley [4], Theorem 8.1, in regard of Lemmas 6.2 and 6.3 below. ■

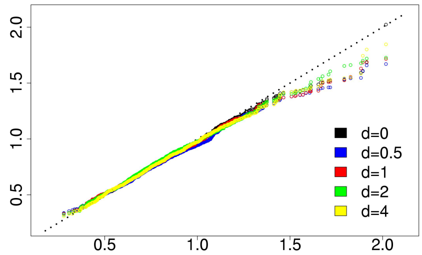
From the above theorem we can now derive the asymptotics for the frequency rank statistics we are interested in.



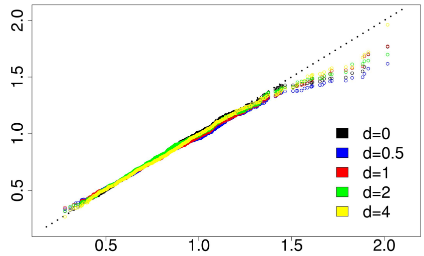
(1) $\rho = -0.5,\ N = 80$



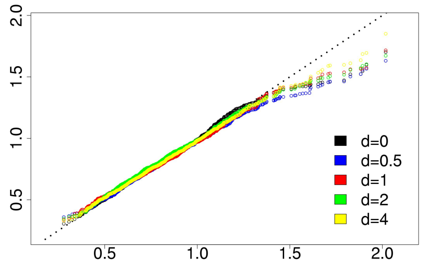
(2) $\rho = -0.3,\ N = 80$



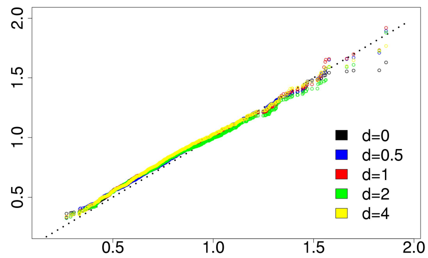
(3) $\rho = 0,\ N = 60$



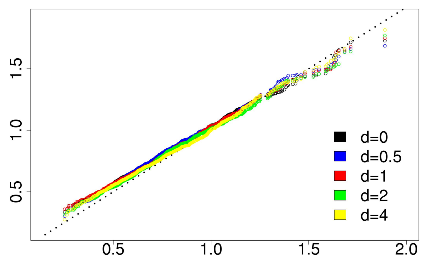
(4) $\rho = 0,\ N = 70$



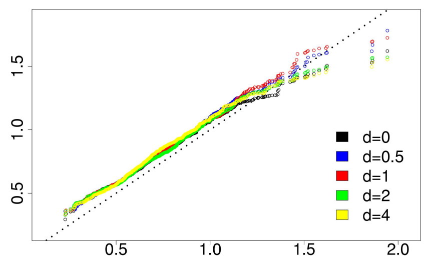
(5) $\rho = 0,\ N = 80$



(6) $\rho = 0.1,\ N = 80$



(7) $\rho = 0.3,\ N = 80$



(8) $\rho = 0.5,\ N = 80$

Figure 5.2: QQ-plots of $T_n^{(3)}(q_1)$ (under H_0) against $T_n^{(3f)}(\mathbf{R}, q_1)$ for $n = 80$, $m = 40$ and different values for ρ and N

Corollary 6.1. *Let conditions* (6.2) *be fulfilled and* $\alpha(n) \to \infty$.

a) If $q \in FC_0^1$ *and* $\int_0^1 \frac{1}{q^2(t)}\,dt < \infty$*, then it holds*

$$T_n^{(1f)}(\mathbf{x}, q) := \max_{1 \leqslant k < N} \frac{1}{q\left(\frac{k}{N}\right)} \left| Z_n^{\mathbf{x}}\left(\frac{k}{N}, \mathbf{R}\right) - \frac{k}{N} Z_n^{\mathbf{x}}(1, \mathbf{R}) \right| \xrightarrow{\mathcal{D}} \sup_{0 \leqslant t \leqslant 1} \frac{|B(t)|}{q(t)}.$$

b) For $\int_0^1 \frac{(t(1-t))^{\kappa}}{r(t)}\,dt < \infty$ *for some* $0 \leqslant \kappa < 1$*, it holds*

$$T_n^{(2f)}(\mathbf{x}, r) := \int_0^1 \frac{1}{r(t)} |Z_n^{\mathbf{x}}(t, \mathbf{R}) - t Z_n^{\mathbf{x}}(1, \mathbf{R})|^2\,dt \xrightarrow{\mathcal{D}} \int_0^1 \frac{B^2(t)}{r(t)}\,dt.$$

c) For all $\epsilon > 0$ *we get*

$$T_n^{(3f)}(\mathbf{x}) := \sup_{\epsilon \leqslant t \leqslant 1-\epsilon} |Z_n^{\mathbf{x}}(t, \mathbf{R}) - t Z_n^{\mathbf{x}}(1, \mathbf{R})| \xrightarrow{\mathcal{D}} \sup_{\epsilon \leqslant t \leqslant 1-\epsilon} \sqrt{\frac{1}{t(1-t)}} |B(t)|.$$

Here $\{B(t) : 0 \leqslant t \leqslant 1\}$ *denotes a Brownian bridge.*

Proof. The result for c) follows immediately from Theorem 6.1, since we can deduce from the Portmanteau theorem (cf. e.g. Billingsley [4], Theorem 2.1) [or the continuous mapping theorem]

$$f(Z_n^{\mathbf{x}}(\cdot, \mathbf{R})) \xrightarrow{\mathcal{D}} f(W(\cdot)),$$

for all continuous $f : C[0,1] \to \mathbb{R}$ where $\{W(t) : 0 \leqslant t \leqslant 1\}$ is a Wiener process. In fact one can deduce from the Portmanteau theorem that this is equivalent to the assertion in Theorem 6.1. Note that $C[0,1]$ is provided with the sup-norm. Thus the transformation in c) is continuous, which gives the assertion.
For the proof of a) note that in the same way we get for any $0 < \eta < \frac{1}{2}$

$$\max_{N\eta \leqslant k \leqslant N - N\eta} \frac{1}{q\left(\frac{k}{N}\right)} \left| Z_n^{\mathbf{x}}\left(\frac{k}{N}, \mathbf{R}\right) - \frac{k}{N} Z_n^{\mathbf{x}}(1, \mathbf{R}) \right| \xrightarrow{\mathcal{D}} \sup_{\eta \leqslant t \leqslant 1-\eta} \frac{|B(t)|}{q(t)}, \tag{6.3}$$

because $\inf_{\eta \leqslant t \leqslant 1-\eta} q(t) > 0$. Furthermore Corollaries 1.2 and 1.3 in Csörgő and Horváth [7] show $\lim_{\eta \to 0} \sup_{0<t<\eta, 1-\eta<t<1} \frac{|B(t)|}{q(t)} = 0$ almost surely. Similarly an application of the Markov inequality together with Theorem 3.1 in Móricz et al. [23] and Theorem 1.1 in Fazekas and Klesov [10] in view of (6.11) give

$$P\left(\max_{1 \leqslant k < \eta N} \frac{1}{q\left(\frac{k}{N}\right)} |Z_n^{\mathbf{x}}(k/N, \mathbf{R})| > x\right) = O(1) \frac{1}{x^4} \frac{1}{N^2} \sum_{k=1}^{\eta N} \frac{k}{q^4\left(\frac{k}{N}\right)}$$
$$= O(1) \frac{1}{x^4} \sup_{0 \leqslant t \leqslant \eta} \frac{t}{q^2(t)} \int_0^{\eta} \frac{1}{q^2(t)}\,dt \to 0 \quad \text{as } \eta \to 0.$$

Now standard arguments yield the assertion. Similar arguments allow us to deduce b) from a). For details confer Kirch [18], Corollary 4.5.1. ■

We will now derive the convergence of the finite-dimensional distributions as well as tightness to prove the above convergence in $C[0,1]$. Our first lemma states that the linearly interpolated part of $Z_n^{\mathbf{x}}(\cdot, \mathbf{R})$ can be neglected to derive the asymptotics.

Lemma 6.1. *Under conditions* (6.2) *it holds as* $n \to \infty$*, uniformly in* $0 \leqslant u \leqslant 1$,

$$Z_n^{\mathbf{x}}(u, \mathbf{R}) - Z_n^{\mathbf{x}}(\lfloor Nu \rfloor /N, \mathbf{R}) = O_P\left(\sqrt{\frac{1}{N}}\right).$$

Proof. Note that $\widetilde{Z}_n^{\mathbf{x}}(u, \mathbf{R}) - \widetilde{Z}_n^{\mathbf{x}}(\lfloor Nu \rfloor /N, \mathbf{R}) = (Nu - \lfloor Nu \rfloor)\frac{1}{\sqrt{N}} x_{\mathbf{R}}(\beta(\lceil Nu \rceil))$, which is a linear rank statistic. Some calculations yield that under (6.2) it holds

$$\operatorname{var}\left(\frac{2}{n}\sum_{l=1}^{\tilde{n}} c_{\beta(\lceil Nu \rceil)}(l) \sum_{j=1}^{n} x_{j,n} c_j(R_l)\right) = O(1).$$

For a closed formula of the variance of a linear rank statistic confer Hájek et al. [13], Theorem 3.3.3. An application of the Markov inequality now yields the assertion, for details confer Kirch [18], Lemma 4.5.1. ■

Now we can state the convergence of the finite-dimensional distributions. This is the only place where we actually need $\alpha(n) \to \infty$, otherwise the Noether condition (6.7) and thus the Lindeberg condition are not fulfilled (which does not necessarily mean that we do not have asymptotic normality).

Lemma 6.2. *For any* $0 \leqslant u_1 < \ldots < u_k \leqslant 1, k \geqslant 1$, *it holds under* (6.2)

$$(Z_n^{\mathbf{x}}(u_1, \mathbf{R}), \ldots, Z_n^{\mathbf{x}}(u_k, \mathbf{R})) \xrightarrow{\mathcal{D}} (W(u_1), \ldots, W(u_k)),$$

if $\alpha(n) \to \infty$*. Here* $\{W(t) : 0 \leqslant t \leqslant 1\}$ *is a Wiener process,*

Proof. By the Cramer-Wold device and Lemma 6.1 it suffices to prove for $\gamma_1, \ldots, \gamma_k \neq 0$ that $\sum_{i=1}^{k} \gamma_i Z_n^{\mathbf{x}}(\lfloor Nu_i \rfloor /N, \mathbf{R}) \xrightarrow{\mathcal{D}} \sum_{i=1}^{k} \gamma_i W(u_i)$ (w.l.o.g. $u_1 \neq 0$). Under (6.2) the left hand side is again a linear rank statistic with (cf. Hájek et al. [13], Theorem 3.3.3)

$$\operatorname{var}\left(\sum_{i=1}^{k} \gamma_i Z_n^{\mathbf{x}}\left(\frac{\lfloor Nu_i \rfloor}{N}, \mathbf{R}\right)\right) \to \sum_{i=1}^{k}\sum_{j=1}^{k} \gamma_i \gamma_j \min(u_i, u_j) = \operatorname{var}\left(\sum_{i=1}^{k} \gamma_i W(u_i)\right),$$

since Lemma A.1 yields after some calculations

$$\frac{2}{nN}\sum_{l=1}^{\tilde{n}}\left(\sum_{i=1}^{k} \gamma_i \sum_{s=1}^{\lfloor Nu_i \rfloor} c_{\beta(s)}(l)\right)^2 \to \sum_{i=1}^{k}\sum_{j=1}^{k} \gamma_i \gamma_j \min(u_i, u_j) \tag{6.4}$$

and Lemma A.3 gives

$$\left|\frac{1}{\sqrt{N}n}\sum_{l=1}^{\tilde{n}}\sum_{i=1}^{k} \gamma_i \sum_{s=1}^{\lfloor Nu_i \rfloor} c_{\beta(s)}(l)\right| = O(1)\frac{\log(N)}{\sqrt{N}} \to 0 \quad (n \to \infty). \tag{6.5}$$

To complete the proof we verify the Lindeberg condition for rank statistics, confer Hájek et al. [13], problems 2 and 3 in Section 6.1. Let $d_n(l) := \sqrt{\frac{2}{nN}} \sum_{i=1}^{k} \gamma_i \sum_{s=1}^{\lfloor Nu_i \rfloor} c_{\beta(s)}(l)$ and $a_n(l) := \sqrt{\frac{2}{n}} \sum_{j=1}^{n} x_{j,n} c_j(l)$. (6.2), (6.4), and (6.5) show that

$$\frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}}(d_n(i) - \bar{d}_n)^2 \sum_{j=1}^{\tilde{n}}(a_n(j) - \bar{a}_n)^2 \to \sum_{i=1}^{k}\sum_{j=1}^{k} \gamma_i \gamma_j \min(u_i, u_j) > 0. \tag{6.6}$$

Moreover

$$\max_{1\leqslant l\leqslant \tilde n} \left(d_n(l)-\bar d_n\right)^2 = O\left(\frac{1}{\alpha(n)}\right) \to 0. \tag{6.7}$$

Thus the Lindeberg condition can be reduced to

$$\frac{1}{\tilde n}\sum_{|a_n(i)-\bar a_n|>\epsilon\sqrt{\alpha(n)}} (a_n(i)-\bar a_n)^2 \to 0 \qquad \text{for any } \epsilon>0.$$

(6.2) shows that the corresponding Lyapunov-type condition (with the 4th moment) is fulfilled, which completes the proof. ■

We finish this section with the proof of the tightness of the process $Z_n^{\mathbf{x}}(\cdot,\mathbf{R})$.

Lemma 6.3. *Under conditions* (6.2) *the sequence of processes* $\{Z_n^{\mathbf{x}}(u,\mathbf{R}) : 0\leqslant u\leqslant 1\}$ *is tight.*

Proof. First $Z_n^{\mathbf{x}}(t,\mathbf{R}) - Z_n^{\mathbf{x}}(u,\mathbf{R}) = \frac{2}{n\sqrt{N}}\sum_{l=1}^{\tilde n} d_n(l)\ (a_n(R_l)-\bar a_n)$ is a linear rank statistic with

$$d_n(l) = \sum_{s=\lceil Nu\rceil+1}^{\lfloor Nt\rfloor} c_{\beta(s)}(l) + (Nt-\lfloor Nt\rfloor)\, c_{\beta(\lceil Nt\rceil)}(l) + (\lceil Nu\rceil - Nu)\, c_{\beta(\lceil Nu\rceil)}(l),$$

$$a_n(l) = \sum_{i=1}^{n} x_{i,n}\, c_i(l).$$

Here $\bar a_n := \frac{1}{\tilde n}\sum_{l=1}^{\tilde n} a_n(l)$ and an equivalent expression for $\bar d_n$. Define

$$z_{2d} := \sum_{l=1}^{\tilde n}\left(d_n(l)-\bar d_n\right)^2 = \sum_{l=1}^{\tilde n} d_n^2(l) - \tilde n(\bar d_n)^2 \leqslant \sum_{l=1}^{\tilde n} d_n^2(l),$$

$$z_{4d} := \sum_{l=1}^{\tilde n}\left(d_n(l)-\bar d_n\right)^4,$$

z_{2a}, z_{4a} analogously. To prove tightness it suffices according to Billingsley [4], Theorem 12.3, to show that for some $C>0$ and all $0\leqslant u\leqslant t\leqslant 1$ it holds $\mathrm{E}[Z_n^{\mathbf{x}}(t,\mathbf{R}) - Z_n^{\mathbf{x}}(u,\mathbf{R})]^4 \leqslant C(t-u)^2$. Hájek et al. [13], problem 25 in Section 3.3, yields

$$\begin{aligned}&\mathrm{E}[Z_n^{\mathbf{x}}(t,\mathbf{R}) - Z_n^{\mathbf{x}}(u,\mathbf{R})]^4\\ &\quad = O\left(\frac{1}{N^2n^6} z_{2a}^2 z_{2d}^2 + \frac{1}{N^2n^5} z_{4a} z_{4d} + \frac{1}{N^2n^6} z_{4a} z_{2d}^2 + \frac{1}{N^2n^6} z_{2a}^2 z_{4d}\right).\end{aligned} \tag{6.8}$$

By assumption (6.2) we have

$$z_{2a} = O(n^2), \qquad z_{4a} = O(n^3). \tag{6.9}$$

Lemma A.3 gives $\bar d_n = O(\log N)$ and anyway $\bar d_n = O(N(t-u))$, which together means $n(\bar d_n)^4 = O(nN^2(t-u)^2\log^2(N)) = O(n^2N^2(t-u)^2)$. Moreover Lemma A.2 shows $\sum_{l=1}^{n} d_n^4(l) = O(nN^3(t-u)^3 + N^4(t-u)^4) = O(n^2N^2(t-u)^2)$. This proves

$$z_{4d} = O(n^2N^2(t-u)^2). \tag{6.10}$$

Lemma A.1 shows $z_{2d} = O(nN(t-u))$, which together with (6.8) to (6.10) yields

$$\mathrm{E}[Z_n^{\mathbf{x}}(t,\mathbf{R}) - Z_n^{\mathbf{x}}(u,\mathbf{R})]^4 \leqslant C(t-u)^2, \tag{6.11}$$

which gives the assertion. ■

7. Proofs of the Main Results

In this section we prove Theorem 4.1 respectively Corollary 4.1. We start with two auxiliary lemmas stating properties of the estimators $\mu_1 = \mu_1(\gamma)$, $\mu_2 = \mu_2(\gamma)$, and $\widehat{m} = \widehat{m}(\gamma)$, $0 \leqslant \gamma \leqslant 1/2$, as in (3.2).

Lemma 7.1. *Let* (2.1) - (2.3) *hold, then under the null hypothesis, i.e.* $m = n, d = 0$, *as well as local alternatives with* $m = \lfloor \vartheta n \rfloor$ *and* $nd^4 = O(1)$, *it holds for* $r = 1, 2$

$$
\begin{aligned}
a)\quad & |\mu_i - \widehat{\mu}_j| = o_P\left(\frac{\sqrt{n}}{\log n}\right), \qquad i, j = 1, 2,\\
b)\quad & \frac{m \wedge \widehat{m}}{n}|\mu_1 - \widehat{\mu}_1|^r = o_P(1), \quad \frac{n - (m \vee \widehat{m})}{n}|\mu_2 - \widehat{\mu}_2|^r 1_{\{\widehat{m}<n\}} = o_P(1),\\
c)\quad & \frac{(m \wedge \widehat{m})^3}{n^2}(\mu_1 - \widehat{\mu}_1)^4 = O_P(1),\\
& \frac{(n - (m \vee \widehat{m}))^3}{n^2}(\mu_2 - \widehat{\mu}_2)^4 1_{\{\widehat{m}<n\}} = O_P(1),\\
d)\quad & (\mu_2 - \widehat{\mu}_1)^4 \frac{(\widehat{m} - m)_+^3}{n^2} = O_P(1), \quad (\mu_1 - \widehat{\mu}_2)^4 \frac{(m - \widehat{m})_+^3}{n^2} 1_{\{\widehat{m}<n\}} = O_P(1),\\
e)\quad & \frac{(\widehat{m} - m)_+}{n}|\mu_2 - \widehat{\mu}_1|^r = o_P(1), \quad \frac{(m - \widehat{m})_+}{n}|\mu_1 - \widehat{\mu}_2|^r 1_{\{\widehat{m}<n\}} = o_P(1),
\end{aligned}
$$

where $a_+ = \max(a, 0)$.

Proof. We will only prove the first equality of c) and d), the rest is analogous. For details confer Kirch [18], Lemma 4.6.1. Note that

$$
\widehat{\mu}_1 - \mu_1 = \frac{1}{\widehat{m}} \sum_{j=1}^{\widehat{m}} (X(j) - \operatorname{E} X(j)) + d\, \frac{(\widehat{m} - m)_+}{\widehat{m}},
$$

$$
\widehat{\mu}_1 - \mu_2 = \frac{1}{\widehat{m}} \sum_{j=1}^{\widehat{m}} (X(j) - \operatorname{E} X(j)) - d\, \frac{m \wedge \widehat{m}}{\widehat{m}}.
$$

First we prove the assertion for the term $\frac{1}{\widehat{m}} \sum_{j=1}^{\widehat{m}} (X(j) - \operatorname{E} X(j)) = \frac{1}{\widehat{m}} \sum_{j=1}^{\widehat{m}} e(j)$. Then we have

$$
\frac{\widehat{m}^3}{n^2} \left| \frac{1}{\widehat{m}} \sum_{j=1}^{\widehat{m}} e(j) \right|^4 \leqslant \frac{1}{n} \max_{1 \leqslant k \leqslant n} \left| \frac{1}{(kn)^{\frac{1}{4}}} \sum_{j=1}^{k} e(j) \right|^4 = O_P\left(\frac{1}{n}\right),
$$

because the Hájek-Renyi inequality for linear processes (cf. e.g. Bai [3], Proposition 1, in addition to Phillips and Solo [24], equation (15) et sqq.) gives:

$$
P\left(\max_{1 \leqslant k \leqslant n} \left| \frac{1}{(kn)^{\frac{1}{4}}} \sum_{j=1}^{k} e(j) \right| \geqslant C \right) = O(1) \frac{1}{C^2 n^{1/2}} \sum_{k=1}^{n} k^{-\frac{1}{2}} = O(1) \frac{1}{C^2}.
$$

For the terms involving d we have

$$
\begin{aligned}
& \frac{\min(m, \widehat{m})^3}{n^2} \left| d\, \frac{(\widehat{m} - m)_+}{\widehat{m}} \right|^4 \leqslant |d|^4 \min\left(\frac{m^3}{n^2}, \frac{(\widehat{m} - m)_+^3}{n^2} \right)\\
& \leqslant |d|^4 n \min(\vartheta, 1 - \vartheta)^3 = O(1)
\end{aligned}
$$

and

$$\frac{(\widehat{m}-m)_+^3}{n^2}\left|d\,\frac{m}{\widehat{m}}\right|^4 \leqslant |d|^4 \min\left(\frac{m^3}{n^2}, \frac{(\widehat{m}-m)_+^3}{n^2}\right) = O(1).$$

This completes the proof. ■

Lemma 7.2. *Let* (2.1) - (2.3) *hold, then under alternatives with* $m = \lfloor \vartheta n \rfloor$, $0 < \vartheta < 1$, *and* $d\sqrt{\frac{n}{\log n}} \to \infty$, *$d$ bounded, it holds*

$$\begin{aligned}
&a) \qquad |\mu_i - \widehat{\mu}_j| = o_P\left(\frac{\sqrt{n}}{\log n}\right), \qquad i,j = 1,2,\\
&b) \qquad n^{\frac{1}{4}}|\mu_j - \widehat{\mu}_j| = O_P(1), \qquad j = 1,2,\\
&c) \qquad (\mu_j - \widehat{\mu}_j)^4\frac{|m-\widehat{m}|^3}{n^2} + |d|^4\frac{|m-\widehat{m}|^3}{n^2} = O_P(1), \qquad j = 1,2,\\
&d) \qquad \frac{|m-\widehat{m}|}{n}|d|^r = o_P(1), \qquad r = 1,2.
\end{aligned}$$

Proof. The Hájek-Renyi inequality for linear processes (cf. e.g. Bai [3], Proposition 1, in addition to Phillips and Solo [24], equation (15) et sqq.) states $(0 \leqslant \gamma \leqslant 1/2)$

$$P\left(n^{\gamma-1}\max_{1\leqslant k\leqslant n}\frac{1}{k^\gamma}\left|\sum_{j=1}^{k}(X(j) - \mathrm{E}\,X(j))\right| \geqslant C\right) = O(1)\frac{n^{2(\gamma-1)}}{C^2}\sum_{k=1}^{n}\frac{1}{k^{2\gamma}}$$
$$= O(1)\frac{1}{C^2}n^{-1}\log n$$

and an analogous expression for $\max_{1\leqslant k<n}\frac{1}{(n-k)^\gamma}\sum_{j=k+1}^{n}(X(j) - \mathrm{E}\,X(j))$ (the Hájek-Renyi inequality in this case can be analogous proven, cf. e.g. Kirch [18], Lemma B.1). This together with equations (3.11) and (3.12) in the proof of Theorem 1.1 in Kokoszka and Leipus [22] gives now

$$|d|\frac{|\widehat{m}-m|}{n} = O_P\left(\sqrt{\frac{\log n}{n}}\right). \tag{7.1}$$

Hence it holds $\frac{\widehat{m}-m}{n} = o_P(1)$, because $d\sqrt{\frac{n}{\log n}} \to \infty$. This gives

$$\frac{n}{\min(\widehat{m}, n-\widehat{m})} = O_P(1), \tag{7.2}$$

since $\frac{n}{\widehat{m}} = \left(\frac{\widehat{m}-m}{n} + \vartheta\right)^{-1} \stackrel{P}{\longrightarrow} \vartheta^{-1} < \infty$ and an analogous argument for $\frac{n}{n-\widehat{m}} \stackrel{P}{\longrightarrow} (1-\vartheta)^{-1}$. As before $\widehat{\mu}_1 - \mu_1 = \frac{1}{\widehat{m}}\sum_{j=1}^{\widehat{m}}(X(j) - \mathrm{E}\,X(j)) + d\,\frac{(\widehat{m}-m)_+}{\widehat{m}}$. For the second term it holds because of (7.1) respectively (7.2)

$$\frac{d(\widehat{m}-m)_+}{\widehat{m}} = O_P\left(\sqrt{\frac{\log n}{n}}\right), \tag{7.3}$$

Moreover by the Hájek-Renyi inequality

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{\widehat{m}}(X(i) - \mathrm{E}\,X(i)) \leqslant \frac{1}{\sqrt{n}}\max_{1\leqslant k<n}\sum_{i=1}^{k}(X(i) - \mathrm{E}\,X(i)) = O_P(1).$$

This together with (7.2) and (7.3) yields

$$\sqrt{\frac{n}{\log n}}|\widehat{\mu}_j - \mu_j| = O_P(1), \tag{7.4}$$

Here the assertion for $j = 2$ follows in the same way. Note that $|\mu_i - \widehat{\mu}_j| \leqslant |\mu_j - \widehat{\mu}_j| + d$. Putting together (7.1) and (7.4) we arrive at the assertion. ■

Lemma 7.3. *Let* (2.2) *and* (2.3) *be fulfilled, n even. Then it holds*

a) $$\frac{1}{n}\sum_{j=1}^{n} e(j)^2 \to \sigma^2 \left(\sum_{s\geqslant 0} w_s\right)^2 > 0 \qquad a.s.,$$

b) $$\frac{1}{n^{3/2}}\sum_{l=1}^{\tilde{n}}\sum_{j=1}^{n} e(j)c_j(l) = o_P(1),$$

c) $$\frac{1}{n}\sum_{l=1}^{\tilde{n}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n} e(j)c_j(l)\right|^4 = O_P(1),$$

d) $$\frac{2}{n}\sum_{j=1}^{n/2} e(2j) = o_P(1), \qquad \frac{2}{n}\sum_{j=1}^{n/2} e(2j-1) = o_P(1).$$

Proof. a) follows immediately from Phillips and Solo [24], Theorem 3.7, and d) from Theorem 3.3, because $e(2\diamond)$ and $e(2\diamond -1)$ can easily be written as sums of two linear processes fulfilling (2.2) and (2.3). The proof for b) and c) goes along the lines of the proof of Theorem 10.3.1 in Brockwell and Davis [6]. We will therefore only sketch the steps, details can be found in Kirch [18], Theorem 4.6.1. We will now prove b). First the decomposition given by Brockwell and Davis yields

$$\frac{1}{n}\sum_{l=1}^{\tilde{n}}\frac{1}{\sqrt{n}}\sum_{j=1}^{n} e(j)c_j(l) = \frac{1}{n}\sum_{j=1}^{n}\epsilon(j)\left[\frac{1}{\sqrt{n}}\sum_{l=1}^{\tilde{n}} c_j(l)d_n(l)\right] + \frac{1}{n}\sum_{l=1}^{\tilde{n}} Y_n(l),$$

where $d_n(\cdot)$ are some constants and $Y_n(\cdot)$ are centered r.v.´s with, as $n \to \infty$,

$$\max_{l=1,\dots,\tilde{n}} |d_n(l)| = O(1), \qquad \operatorname{var} Y_n(l) = O(1/n) \quad \text{uniformly in } l.$$

An application of the Chebyshev inequality and Remark A.1 yields

$$\frac{1}{n}\sum_{j=1}^{n}\epsilon(j)\left[\frac{1}{\sqrt{n}}\sum_{l=1}^{\tilde{n}} c_j(l)d_n(l)\right] = O_P\left(n^{-1/2}\right).$$

Furthermore an application of the Chebyshev and Cauchy-Schwarz inequalities shows

$$\frac{1}{n}\sum_{l=1}^{\tilde{n}} Y_n(l) = O_P\left(n^{-1/2}\right),$$

which completes the proof of b). Concerning c) we get similarly by the decomposition of Brockwell and Davis [6], proof of Theorem 10.3.1,

$$\frac{1}{n}\sum_{l=1}^{\tilde{n}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n} e(j)c_j(l)\right|^4 = O(1)\,\frac{1}{n}\sum_{l=1}^{\tilde{n}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\epsilon(j)c_j(l)\right|^4 + O(1)\,\frac{1}{n}\sum_{l=1}^{\tilde{n}}|Y_n(l)|^4,$$

where $\mathrm{E}\,|Y_n(l)|^4 = O(1/n^2)$ uniformly in l. Thus the Markov inequality shows that the last term is $O_P(1/n^2)$. Assertion c) follows now by the Markov inequality, since

$$P\left(\frac{1}{n^3}\sum_{l=1}^{\tilde{n}}\left|\sum_{j=1}^{n}\epsilon(j)c_j(l)\right|^4 \geqslant C\right) \leqslant \frac{1}{C}\max_{1\leqslant l\leqslant\tilde{n}}\mathrm{E}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\epsilon(j)c_j(l)\right|^4 = O(1)\,\frac{1}{C},$$

where the last estimate follows e.g. by Stout [25], Theorem 3.7.8. ■

Now we are ready to prove Theorem 4.1 respectively Corollary 4.1.

Proof. Theorem 4.1 and Corollary 4.1 The idea of the proof is the following: We use Theorem 6.1 respectively Corollary 6.1 for the scores

$$x_i := \widetilde{e}(i)\left(\frac{2}{n\tilde{n}}\sum_{l=1}^{\tilde{n}}\left(\sum_{i=1}^{n}\widetilde{e}(i)c_i(l) - \frac{1}{\tilde{n}}\sum_{k=1}^{\tilde{n}}\sum_{j=1}^{n}\widetilde{e}(j)c_j(k)\right)^2\right)^{-\frac{1}{2}}.$$

We will verify that these scores fulfill assumptions (6.2) in a P-stochastic sense. The subsequence principle gives then the assertion. Note that the convergence in $C[0,1]$ can be equivalently expressed via a distributional convergence of all real continuous transformations of the process (confer also the proof of Corollary 6.1).

In view of Remark 4.2 and Lemma 7.3 it suffices to show that

$$\frac{1}{n}\sum_{i=1}^{n}|\widetilde{e}(i) - e(i)|^2 = o_P(1), \tag{7.5}$$

$$\frac{1}{n^{3/2}}\sum_{l=1}^{\tilde{n}}\sum_{j=1}^{n}(\widetilde{e}(j) - e(j))c_j(l) = o_P(1), \tag{7.6}$$

$$\frac{1}{n}\sum_{l=1}^{\tilde{n}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}c_j(l)(\widetilde{e}(j) - e(j))\right|^4 = O_P(1), \tag{7.7}$$

$$\frac{1}{n}\sum_{j=1}^{n/2}(\widetilde{e}(2j) - e(2j)) = o_P(1), \tag{7.8}$$

$$\frac{1}{n}\sum_{j=1}^{n/2}(\widetilde{e}(2j-1) - e(2j-1)) = o_P(1). \tag{7.9}$$

Concerning (7.5) note that $(a+b)^2 \leqslant 2a^2 + 2b^2$, which gives $(a-b)^2 \geqslant \frac{1}{2}a^2 - b^2$. This shows in our situation $\frac{1}{n}\sum_{i=1}^{n}|\widetilde{e}(i)|^2 \geqslant \frac{1}{2n}\sum_{i=1}^{n}|e(i)|^2 - \frac{1}{n}\sum_{i=1}^{n}|\widetilde{e}(i) - e(i)|^2$.

Instead of (7.6) we will verify

$$\frac{\log n}{\sqrt{n}}\max_{j=1,\dots,n}|\widetilde{e}(j) - e(j)| = o_P(1), \tag{7.10}$$

which implies (7.6) because of Lemma A.3. We will first prove (7.5) and (7.7) - (7.10) under alternatives with $d\sqrt{\frac{n}{\log n}} \to \infty$. Recall $m = \lfloor \vartheta n \rfloor$ and d is bounded. It holds

$$\begin{aligned}
&\widetilde{e}(i) - e(i) \\
&= (X(i) - \widehat{\mu}_1)1_{[1,\widehat{m}]}(i) + (X(i) - \widehat{\mu}_2)1_{[\widehat{m}+1,n]}(i) - (X(i) - \mu_1)1_{[1,m]}(i) \\
&\quad - (X(i) - \mu_2)1_{[m+1,n]}(i) \\
&= (\mu_1 - \widehat{\mu}_1)1_{[1,m\wedge\widehat{m}]}(i) + (\mu_2 - \widehat{\mu}_2)1_{(m\vee\widehat{m},n]}(i) + (\mu_1 - \widehat{\mu}_2)1_{(\widehat{m},m]}(i) \\
&\quad + (\mu_2 - \widehat{\mu}_1)1_{(m,\widehat{m}]}(i).
\end{aligned}$$

The triangle inequality together with Lemma 7.2 now gives for $r = 1, 2$

$$\begin{aligned}
&\frac{m\wedge\widehat{m}}{n}|\mu_1 - \widehat{\mu}_1|^r + \frac{n-(m\vee\widehat{m})}{n}|\mu_2 - \widehat{\mu}_2|^r + \frac{(m-\widehat{m})_+}{n}|\mu_1 - \widehat{\mu}_2|^r \\
&\quad + \frac{(\widehat{m}-m)_+}{n}|\mu_2 - \widehat{\mu}_1|^r \leqslant o_P(1) + \frac{|m-\widehat{m}|}{n}|d|^r = o_P(1).
\end{aligned}$$

By Lemma A.2

$$\begin{aligned}
&\frac{1}{n^3}\sum_{l=1}^{\tilde{n}}(\mu_1 - \widehat{\mu}_1)^4\left(\sum_{j=1}^{m\wedge\widehat{m}} c_j(l)\right)^4 = O(1)\,\frac{(m\wedge\widehat{m})^3}{n^2}(\mu_1 - \widehat{\mu}_1)^4 \\
&= O(1)\,n(\mu_1 - \widehat{\mu}_1)^4 = O_P(1).
\end{aligned}$$

Analogously we get

$$\frac{1}{n^3}\sum_{l=1}^{\tilde{n}}(\mu_2 - \widehat{\mu}_2)^4\left(\sum_{j=m\vee\widehat{m}}^{n} c_j(l)\right)^4 = O_P(1).$$

Moreover for $\widehat{m} < m \leqslant n$

$$\begin{aligned}
&(\mu_1 - \widehat{\mu}_2)^4\frac{1}{n^3}\sum_{l=1}^{\tilde{n}}\left(\sum_{j=\widehat{m}+1}^{m} c_j(l)\right)^4 = O(1)\,(\mu_1 - \widehat{\mu}_2)^4\frac{(m-\widehat{m})_+^3 n}{n^3} \\
&= O(1)\,(\mu_2 - \widehat{\mu}_2)^4\frac{|m-\widehat{m}|^3}{n^2} + O(1)\,|d|^4\frac{|m-\widehat{m}|^3}{n^2} = O_P(1).
\end{aligned}$$

Analogously for $m < \widehat{m}$

$$(\mu_2 - \widehat{\mu}_1)^4\frac{1}{n^3}\sum_{l=1}^{\tilde{n}}\left(\sum_{j=m+1}^{\widehat{m}} c_j(l)\right)^4 = O_P(1).$$

Under the null hypothesis as well as local alternatives with $|d|^4 n = O(1)$ analogous arguments in addition to Lemma 7.1 give the assertion. Since we can divide each sequence into two sub-sequences (possibly one of which is empty), one fulfilling $|d|^4 n = O(1)$, the other one $|d|\sqrt{n/\log(n)} \to \infty$, this completes the proof. ■

A. Some Properties of Trigonometric Functions

The frequency permutation statistics are based on trigonometric polynomials. In the proofs for the rank statistic asymptotics we make use of the special structure of trigonometric polynomials. We state some results in this appendix.

Lemma A.1. *a) For $j = 1, \ldots, n$ it holds*

$$\sum_{l=1}^{n-2} c_j(l)^2 = \frac{n-2}{2}.$$

b) For $i, j = 1, \ldots, n, i \neq j$ it holds

$$\sum_{l=1}^{n-2} c_j(l)\, c_i(l) = -\frac{1}{2}(1 + (-1)^{i-j}) = \begin{cases} 0 & i-j \text{ odd}, \\ -1 & i-j \text{ even}. \end{cases}$$

Proof. Straightforward calculations give the assertion, for details confer Kirch [18], Theorem 4.4.1. ■

Remark A.1. Note that $\{\sqrt{\frac{2}{n}} c_\diamond(l), l = 1, \ldots, \tilde{n}\}$ forms an ON-System (confer e.g. Brockwell and Davis [6], page 333).

Lemma A.2. *It holds*

$$\sum_{l=1}^{n-2} c_{s_1}(l) c_{s_2}(l) c_{s_3}(l) c_{s_4}(l) = \begin{cases} O(n), & \sum_{j=1}^{4} \delta_\pm^{(j)} s_j = in,\ i \in \mathbb{Z},\ \delta_\pm^{(j)} = \pm 1 \\ & \quad \text{with } \sum_{j=1}^{n} \delta_\pm^{(j)} = 0 \text{ or } \sum_{j=1}^{n} \delta_\pm^{(j)} = 4, \\ 0, & s_1 + s_2 + s_3 + s_4 \text{ odd}, \\ -1, & \text{otherwise}. \end{cases}$$

The condition above means that the sum is only of linear order if s_4 is determined by a finite number of combinations of s_1, s_2, s_3.

Proof. Again direct calculations give the assertion, for details confer Kirch [18], Lemma 4.4.1. ■

Lemma A.3. *If k is not a multiple of n, it holds*

$$\sum_{j=1}^{M} \cos(2\pi kj/n) = O\left(\max\left(\frac{n}{k}, \frac{n}{n-k}\right)\right)$$

$$\text{and} \qquad \sum_{j=1}^{M} \sin(2\pi kj/n) = O\left(\max\left(\frac{n}{k}, \frac{n}{n-k}\right)\right)$$

uniformly in M and anyway $\leqslant M$ uniformly in k. Note

$$\sum_{k=1}^{n} \max\left(\frac{n}{k}, \frac{n}{n-k}\right) \leqslant 2n \sum_{k=1}^{n} \frac{1}{k} = O(n \log n),$$

$$\sum_{k=1}^{n/2} \min\left(M, \frac{n}{k}\right) \leqslant \sum_{k=1}^{n/M} M + \sum_{k=n/M+1}^{n/2} \frac{n}{k} \leqslant n + n \int_{n/M}^{n/2} \frac{1}{x}\, dx = O(n \log M).$$

Proof. An application of equation 1.342 of Gradshteyn and Ryzhik [12] gives the assertion. For details confer Kirch [18], Lemma 4.4.2. ■

Acknowledgment

I would like to express my thanks to my supervisor Prof. Dr. Josef Steinebach for introducing me to the field of change-point analysis. Moreover I am much obliged to Prof. Dr. Marie Hušková from Charles University, Prague, whose deep insights in rank statistics as well as the field of change-point analysis helped solve many problems. Furthermore it was her who suggested the possibility of bootstrapping in the frequency domain. Last but not least my thanks go to Prof. Dr. David Mason from the University of Delaware for pointing out to me that the frequency bootstrap is essentially just another linear rank statistic problem.

References

[1] Antoch, J., and Hušková, M. Permutation tests for change point analysis. *Statist. Probab. Lett.*, 53:37–46, 2001.

[2] Antoch, J., Hušková, M., and Prášková, Z. Effect of dependence on statistics for determination of change. *J. Statist. Plann. Inference*, 60:291–310, 1997.

[3] J. Bai. Least squares estimation of a shift in linear processes. *J. Time Ser. Anal.*, 15:435–472, 1994.

[4] Billingsley, P. *Convergence of Probability Measures.* Wiley, New York, 1968.

[5] Brillinger, D.R. *Time Series, Data Analysis and Theory.* SIAM, San Francisco, 1981.

[6] Brockwell, P.J., and Davis, R.A. *Time Series: Theory and Methods.* Springer, New York, second edition, 1991.

[7] Csörgő, M., and Horváth, L. *Weighted Approximations in Probability and Statistics.* Wiley, Chichester, 1993.

[8] Csörgő, M., and Horváth, L. *Limit Theorems in Change-Point Analysis.* Wiley, Chichester, 1997.

[9] Dahlhaus, R., and Janas, D. A frequence domain bootstrap for ratio statistics in time series analysis. *Ann. Statist.*, 24:1934–1963, 1996.

[10] Fazekas, I., and Klesov, O. A general approach to the strong law of large numbers. *Theory Probab. Appl.*, 45:436–448, 2000.

[11] Franke, J., and Härdle, W. On bootstrapping kernel spectral estimates. *Ann. Statist.*, 20:121–145, 1992.

[12] Gradshteyn, I.S., and Ryzhik, I.M. *Table of Integrals, Series, and Products.* Academic Press, New York, corrected and enlarged edition, 1980.

[13] Hájek, J., Šidák, Z., and Sen, P.K. *Theory of Rank Tests.* Academic Press, San Diego, second edition, 1999.

[14] Hall, P., Horowitz, J. L., and Jing, B. On blocking rules for the bootstrap with dependent data. *Biometrika*, 82:561–74, 1995.

[15] Hidalgo, J. An alternative bootstrap to moving blocks for time series regression models. *J. Econometrics*, 117:369–399, 2003.

[16] Horváth, L. Detection of changes in linear sequences. *Ann. Inst. Statist. Math.*, 49:271–283, 1997.

[17] Hušková, M. Permutation principle and bootstrap in change point analysis. *Fields Inst. Commun.*, 44:273–291, 2004.

[18] Kirch, C. *Resampling Methods for the Change Analysis of Dependent Data*. PhD thesis, Universität zu Köln, Köln, 2006. http://kups.ub.uni-koeln.de/volltexte/2006/1795/.

[19] Kirch, C. Simulation study of the block permutation method as well as frequency permutation method for the change analysis of dependent data. http://www.mi.uni-koeln.de/~jost/simulations.pdf, 2006. University of Cologne, 109 p.

[20] Kirch, C. Block permutation principles for the change analysis of dependent data. *J. Statist. Plann. Inference*, 137:2453–2474, 2007.

[21] Kirch, C., and Steinebach, J. Permutation principles for the change analysis of stochastic processes under strong invariance. *J. Comput. Appl. Math.*, 186:64–88, 2006.

[22] Kokoszka, P., and Leipus, R. Change-point in the mean of dependent observations. *Statist. Probab. Lett.*, 40:385–393, 1998.

[23] Móricz, F., Serfling, R., and Stout, W. Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *Ann. Probab.*, 10:1032–1040, 1982.

[24] Phillips, P.C.B., and Solo, V. Asymptotics for linear processes. *Ann. Statist.*, 20:971–1001, 1992.

[25] Stout, W.F. *Almost Sure Convergence.* Academic Press, New York, 1974.