The moving Fourier transformation of locally stationary processes with application to bootstrap procedures

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)

an der Fakultät für Mathematik
des Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Dipl.-Math. Franziska Lindner geb. Häfner
aus Bayreuth

Tag der mündlichen Prüfung:	16.4.2014
Referent:	JProf. Dr. Claudia Kirch
Korreferent:	Prof. Dr. Rainer Dahlhaus
Korreferent:	Prof. Dr. Norbert Henze

Contents

List of Figures

Danksagung

„Die Arbeit an einer Dissertation ist ein Prozess, der ohne die wohlwollende Unterstützung zahlreicher Personen nicht möglich wäre. Allein dies ist schon Grund, eben diesen Personen Dank zu sagen.“ Lindner [35]

Allen voranstehend möchte ich mich bei meiner Doktormutter Frau Juniorprofessorin Dr. Claudia Kirch für die sehr gute und beispielgebende Betreuung bedanken. Von der Themenstellung bis zur Vollendung hat sie die Entstehung meiner Arbeit sehr aufmerksam und entgegenkommend begleitet und mir in vielen intensiven Gesprächen wertvolle wissenschaftliche Anregungen gegeben. Ebenso danke ich sehr herzlich meinen Korreferenten Herrn Professor Dr. Rainer Dahlhaus von der Universität Heidelberg und Herrn Professor Dr. Norbert Henze, die sehr interessiert meine Arbeit verfolgt und mit bereichernden Fragen unterstützt haben.

In den letzten Jahren habe ich am Institut für Stochastik ein angenehmes und wissenschaftliche Arbeit bereicherndes Klima vorgefunden, für das ich zahlreichen Kollegen danken möchte. Namentlich danken möchte ich Herrn Professor Dr. Günter Last, der mir die Möglichkeit gegeben hat, mit ihm zusammen eine Großveranstaltung zu betreuen. Für die sehr nette und respektvolle Zusammenarbeit möchte ich mich sehr herzlich bedanken. Zu einem wunderbaren Arbeitsklima am Institut tragen ganz entscheident auch Frau Michaela Regelin und Frau Tatjana Dominic bei, die stets mit Rat und Tat zur Seite stehen und dem Institut Zusammenhalt geben. Ein ganz großes Danke geht an diejenigen, die mich Tag für Tag im Büro motiviert haben und mit denen jeder Arbeitstag Spaß gemacht hat: Frau Dipl.-Math. Viola Riess, die tagein, tagaus treu mit mir unser Sportprogramm durchgezogen hat, und in der ich immer jemanden zum Zuhören hatte, Herrn Dipl.-Math. Sebastian Kimmig, meinem gutmütigen und knuffigen Gegenübersitzer, sowie Herrn Dipl.-Math. oec. Markus Schloz (äh, Scholz), der mir stets mit seinen Consulting-Fähigkeiten zur Seite stand. Und natürlich an Herrn PD Dr. Gerd Herzog für die moralische Unterstützung!

Mein äußerster Dank gilt meinen Eltern für ihren beständigen Rückhalt und ihr liebevolles Beiseitestehen in jeder Art und Weise. Ich danke meinem Ehemann Dr. Tobias Lindner MdB (DPRZ) für seine kompetente und liebe Unterstützung. Schließlich danke ich Herrn Dr. Bernhard Klar für seine unendliche Geduld, fortwährende Ermutigung und Liebe.

Karlsruhe, den 7.5.2014

CHAPTER 1

Introduction

Getting acquainted with locally stationary processes:

QUESTION:	What are locally stationary processes and why do we need them?
ANSWER (*Sergides* [49])	Locally stationary processes are nonstationary stochastic processes whose second order structure varies smoothly over time. [We need them to create a] more realistic framework in time series analysis.
ANSWER (*Martin and Flandrin* [37]):	[That's because in reality] the assumption of stationarity fails to be true: the physical character of random signals demands a nonstationary approach such as in acoustics, speech, geophysic, biology, biomedicine fields, etc. However, a spectrum of [a] nonstationary process(...) cannot be defined by simply generalizing the ordinary stationary spectrum.
QUESTION:	Before we go into a thorough discussion on the historical approach to model nonstationarity in general, could you please briefly point out the main ideas the subsequent work is based on?
ANSWER:	Introducing a time varying spectral representation similar to stationary processes and thus allowing to study processes with continuously changing spectral patterns has first been suggested by Priestley [46]. His time dependent spectral functions are called evolutionary spectra, which have a physical interpretation as local energy distributions over frequency. As pointed out by Dahlhaus [6] the approach of Priestley [46] does, however, not allow for meaningful local asymptotic considerations. In order to overcome these difficulties, Dahlhaus [7] suggested to consider a triangular array of data.

The difference of the approach of Dahlhaus [7] to the approach of Priestley [46] is, that Dahlhaus [7] uses double indexed processes and makes use of asymptotic considerations. His concern is to provide a representation which allows for rigorous asymptotic treatment of statistical inference problems, whereas Priestley [46] intended to gain a stochastic representation of the process.

Modelling non-stationarity – historical overview

When dealing with time series in applications, it has already been pointed out that the assumption of stationarity is more than questionable. Modelling time-dependent processes has therefore been dealt with for several decades. There is, of course the possibility to model time-dependent processes in the time domain as done by Hallin [23] and Subba Rao [51]. Subba Rao [51] considered AR-processes with time-varying coefficients represented as expansions of orthogonal polynomials and weighted least squares estimation of the time-varying coefficients. However, he also considered the evolutionary spectral approach developed by Priestley [46]. This concept of an evolutionary spectrum will be discussed later on.

Especially when encountering the field of signal processing and acoustics the assumption of a stationary signal is not convincing. A stationary signal (in continuous time) can be described by the power spectral density

$$f(\omega) = \int_{\tau} \mathrm{Cov}(X_{t+\tau}, X_{t-\tau}) e^{-i2\pi\omega\tau} d\tau, \quad 0 < \omega < \infty$$

(Hlawatsch and Matz [25]). Contrasting stationary processes, non-stationary signals call for time-frequency methods to account for the change of the signal throughout time in order to provide a complete and unique description of the process' second order statistics and spectral properties (cf. Hlawatsch and Matz [25]).
The aim is, thus, to generalize the power spectral density in a way that we get a natural extension with an explicit time-dependence of the classical notion of power spectral density together with most of its "nice" properties (cf. Flandrin [19]). Unfortunately, there is no chance to obtain such a time-dependent spectrum which is unique and well-defined. Whenever choosing a definition, we have to sacrifice one desirable property we would have liked the time-dependent spectrum to have.
There has also been some heated discussion of what conditions are the necessary ones and when a function is allowed to be called a spectrum (cf. Loynes [36] and the discussion of the paper in the appendix).
Priestley [46] reviews the research on the problem of characterizing non-stationary processes via a spectral density: In 1960, Cramér [5] considered the class of non-stationary harmonizable (in the Loève sense) processes. That is, processes with the Cramér representation

$$X_t = \int_{-\infty}^{\infty} e^{i\omega t} dZ(\omega), \qquad -\infty < t < \infty.$$

The increment process $Z(\omega)$, however, is not orthogonal anymore, the increments can be correlated. Cramér then defined the integrated spectrum of such a process by $dF(\omega, \nu) := E(|dZ(\omega)dZ^*(\nu)|)$. A major drawback of this approach, as pointed out by Priestley [47], is the difficulty of interpreting this two-dimensional spectral density function.
Another definition has been given by Hatanaka and Suzuki (unpublished). They define the spectral density function of non-stationary processes as the limit of of the expected value of the periodogram as sample size tends to infinity. Both, Cramér and Hatanaka and Suzuki intended to characterize the behaviour of the non-stationary process over the whole parameter space with the help of a single function.
When concentrating on looking for a local description of the spectrum of a non-stationary process one inevitably comes to Page [43] who was the first to be toying with the idea of a changing spectrum. He defines the instantaneous power spectrum of a non-stationary process.

Instantaneous power spectrum (→ Page [43]) Considering the energy of a signal to be distributed over time and frequency, the density of the energy in the time-frequency plane is called $\rho(t, f)$. For some fixed t, this is called the instantaneous power spectrum at time t.

> Motivation (cf. Priestley [46] and Page [43]):
>
> $\int_0^T \int_{-\infty}^{\infty} \rho(t, f) df dt$ is the total energy of the signal output up to time T. To get the increase in total power from time T to $T + \Delta T$ one differentiates with respect to t. Now differentiating with respect to t yields
>
> $$\int_{-\infty}^{\infty} \rho(T, f) df,$$
>
> which is the instantaneous power of the signal at time T. Thus, $\rho(T, f)$ describes the difference between the energy on the interval $(0, T)$ to the interval $(0, T + \delta T)$ and is called the instantaneous power spectrum at time T. Approximately,
>
> $$\begin{aligned} \rho(T, f) &\approx \frac{1}{\Delta T} E\left(\left| \int_0^{T+\Delta T} X_t e^{-ift} dt \right|^2 - \left| \int_0^T X_t e^{-ift} dt \right|^2 \right) \\ &\approx \frac{1}{\Delta T} \left(\int_0^{T+\Delta T} \rho(t, f) dt - \int_0^T \rho(t, f) dt \right). \end{aligned}$$

Integration of the instantaneous power spectrum over time yields the conventional spectrum (cf. Page [43]). His definition of the conventional spectrum is the same as of Hatanaka and Suzuki.

The reason for Priestley [46] to resort to evolutionary spectra is that he is unhappy with the physical interpretation of Page's instantaneous power spectrum. Priestley considered it far more important to study the spectral content of the process within the interval $(T, T+\delta T)$ than studying the difference between the spectral contents of the intervals $(0, T)$ and $(0, T+\delta T)$. His evolutionary spectrum at time T can roughly be understood as

$$f(t,\omega) \approx E\left(\left|\int_T^{T+\Delta T} X_t e^{-ift} dt\right|^2\right).$$

For the interpretation of the definitions, see the discussions at the end of Priestley [46], pp. 234,235.

Evolutionary spectrum (→ Priestley [46],[47]) Priestley's [47] concept is to generalize the representation of a stationary process as

$$X_t = \int_{-\infty}^{\infty} e^{i\omega t} dZ(\omega), \quad -\infty < t < \infty,$$

with $dZ(\omega)$ remaining an orthogonal process. Not giving up on the increments, i.e. the random amplitudes, being uncorrelated ensures easy interpretation, which has not been provided by Cramér [5].
In order maintain this uncorrelatedness of the increments, Priestley [46] restricts attention to the class of processes for which there exists a family $\mathcal{F}$ of functions $\{\phi_t(\omega)\}$ defined on the real line, indexed by t, and a measure $\mu(\omega)$ on the real line, such that for each $-\infty < s, t < \infty$ the covariance function can be written as

$$\text{Cov}(X_t, X_s) = \int_{-\infty}^{\infty} \phi_s(\omega)\overline{\phi_t(\omega)} d\mu(\omega).$$

Referring to Parzen [45], Priestley [46] points out that for the parameter space being a bounded interval $(0 \le t \le T)$ it is always possible to obtain this kind of representation. Given $\phi_t(\omega)$ is quadratic integrable for each t, X_t admits a representation of the form

$$X_t = \int_{-\infty}^{\infty} \phi_t(\omega) dZ(\omega),$$

where dZ is an orthogonal process with $E|Z(\omega)|^2 = d\mu(\omega)$. (Note: $\mu(\omega)$ here mirrors the role of $F(\omega)$ in the stationary case.) Depending on which family of functions is chosen for ϕ_t one gets a wide variety of different representations of the process. This again is a result of Parzen [45] and has been taken up by Priestley [46].
By choosing $\phi_t(\omega) = e^{i\omega t}$ we get the stationary case. Aiming to consider non-stationary processes, we ought to choose another family of functions. Priestley [46] picked out oscillatory functions (as to preserve the physical concept of frequency):

Definition of an oscillatory function:
→*Priestley [46]*

> The function of t, $\phi_t(\omega)$, will be said to be an oscillatory function if, for some (necessarily unique) $\theta(\omega)$ it may be written in the form $\phi_t(\omega) = A_t(\omega)e^{i\theta(\omega)t}$, where $A_t(\omega)$ is of the form
>
> $$A_t(\omega) = \int_{-\infty}^{\infty} e^{it\theta} dH_\omega(\theta),$$
>
> with $|dH_\omega(\theta)|$ having an absolute maximum at $\theta = 0$.

Note: With $A_t(\omega) = 1$ and $\theta(\omega) = \omega$ the class of oscillatory processes certainly includes all second-order stationary processes.
An oscillatory process whose second-oder characteristics change "slowly" over time, is considered by Priestley [46] to be a semi-stationary process. (Of course, in Priestley [46] the term slowly is defined mathematically.)
For a non-stationary process X_t represented by

$$X_t = \int_{-\infty}^{\infty} A_t(\omega)e^{i\omega t} dZ(\omega),$$

with an orthogonal increment process $dZ(\omega)$, we can interpret $A_t(\omega)dZ(\omega)$ as random amplitudes and consider X_t to be the limit of a sum of many sine and cosine waves with different frequencies and amplitudes $A_t(\omega)dZ(\omega)$. Hence, the power that is contributed by frequency ω is

$$|A_t(\omega)|^2 dF(\omega) = |A_t(\omega)|^2 |dZ(\omega)dZ^*(\omega)|.$$

The evolutionary power spectrum by Priestley [46] is then defined to be

$$f_t(\omega) := |A_t(\omega)|^2 dF(\omega).$$

F is the spectral distribution function of the corresponding stationary process

$$X_t = \int_{-\infty}^{\infty} e^{i\omega t} dZ(\omega).$$

The evolutionary spectrum has the same physical interpretation as the spectrum of a stationary process (cf. Bruscato and Toloi [4]), namely, it describes a distribution of power over frequency, but whereas the latter is determined by the behaviour of the process for all time t, the former represents specifically the spectral content of the process in the neighbourhood of each time instant t.
Unfortunately, this evolutionary spectrum is by no means unique and depends on the family $\mathcal{F}$ considered. Moreover, as pointed out by Dahlhaus [7], the approach of [46] does not allow for rigorous local asymptotic considerations. In order to overcome these difficulties, Dahlhaus [7] suggested to consider a triangular array of data. Subba Rao [51] not only considered estimation in the time domain, but he also used

the evolutionary spectral approach developed by Priestley to estimate the time-varying parameters of the time-dependent AR-processes. Subba Rao's modelling of non-stationary time series with time-dependent AR-models has been resumed by Grenier [22] and Kitagawa and Gersch [29]. The latter restricted the time-varying coefficients by introducing smoothness priors, that is setting up stochastically perturbed difference equations for the coefficients. By doing so, they also create a local time-varying structural model, which does not have global structural time-varying properties.

Evolutive spectrum (→ Tjostheim [53] and Mélard [38]) The evolutive spectrum has independently been proposed by Tjostheim [53] and Mélard [38]. It is defined for discrete time processes and is a special case of Priestley's evolutionary spectra with respect to some special family $\mathcal{F}$. This is explained in more detail in Flandrin [19].

Wigner-Ville spectrum (→ Martin and Flandrin [37]) Another popular definition of a spectrum of nonstationary processes is the Wigner-Ville spectrum (cf. Bruscato and Toloi [4]):
When generalizing the classical ordinary spectrum for stationary time series "under natural conditions" Martin and Flandrin [37] find the Wigner-Ville spectrum to be the only time-varying spectrum to sufficiently comply with those conditions, such as satisfying the linear time-frequency dualism and reducing to the ordinary spectral density if the process is stationary. The major drawback of the Wigner-Ville spectrum is the sacrifice of the non-negativity, which does no longer allow for the physical interpretation of local energy over time. The Wigner-Ville spectrum is (uniquely) defined as the expected value of the Wigner-Ville distribution:

$$f_{WV}(t,\omega) := E[W_x(t,\omega)] = E\left[\int_{-\infty}^{\infty} X_{t+\frac{\tau}{2}}\overline{X_{t-\frac{\tau}{2}}}e^{-i\omega\tau}d\tau\right], \quad -\infty < t < \infty.$$

For the discrete case

$$f_{WV}(t,\omega) := 2\sum_{\tau=-\infty}^{\infty}\gamma(t+\tau,t-\tau)e^{-2i\omega\tau}$$

defines the discrete Wigner-Ville spectrum. We can see, that this is a representation similar to the one for stationary processes.

Time varying spectral density The most recent amendment to the techniques of modelling non-stationary time series has been made by Dahlhaus [8]. He introduces the class of locally stationary processes and along with it, the concept of a time varying spectral density, which is the spectral density of the stationary approximations at different points in time. For this time-varying spectral density, he picks up the idea of Priestley [46] of locally describing the spectral density, but, as pointed out before, he introduced double-indexed processes allowing for asymptotic considerations. Also contrasting the evolutionary spectrum, the time varying spectral density of a locally stationary process is unique and equals the limit of the Wigner-Ville spectrum of this process (see Theorem 2.1).

Aims of this work

Looking at the long list of approaches to consider deviations from stationarity, one can see the great relevance of the topic – and also the difficulties coming along with it, among them the problem of generalizing the stationary model maintaining the possibility of asymptotic theory and the difficulty of generalizing the concept of a spectrum to the non-stationary case – not to speak of estimating it. This thesis, based on the concept of locally stationary time series introduced by Dahlhaus [8], aims to develop a modification of the Fourier transform which enables us to transfer the local structure of the data from the time domain to the frequency domain, yet preserving the convenient property of the resulting Fourier coefficients being at least uncorrelated in the frequency domain. This is then the basis for the application of bootstrap techniques. Of course, some appropriate inverse transformation should be constructed to allow for the bootstrapped coefficients to be converted back to time domain data, again, without losing structural information. The first main goal is thus to generalize the TFT bootstrap by Kirch and Politis [28] to locally stationary time series. This, of course, first implies to find a suitable estimator of the time varying spectral density as well as proving its consistency. It also requires to prove that the TFT bootstrap for locally stationary time series yields the correct covariance structure of the bootstrap observations.
The second objective is to validate that the new way of Fourier transforming is applicable to other state-of-the-art bootstraps. There exist extensions to the wild hybrid bootstrap (Kreiss and Paparoditis [33]) as well as the autoregressive periodogram bootstrap (Kreiss and Paparoditis [31]) using the local periodogram. We intend to generalize these procedures to stationary time series using the periodogram resulting from our new transform and compare the performance of our obtained procedures to the extensions already in existence.
The third aspect is a practical one as it is intended to implement the new transform as well as the new version of the TFT bootstrap and the two other bootstrap procedures. We will moreover deal with the question whether there is any way of reducing the complexity of the algorithm. As test statistics can often be written as spectral means, we also aim to structurally investigate those spectral means being based on the newly introduced periodogram. Naturally, it is also intended to survey those statistics using simulations. The practical part even goes to such lengths as to introduce uniform confidence bands for the autocorrelation and to examine them thoroughly with respect to different error distributions.

Agenda

We start with an introduction to locally stationary processes proposed by Dahlhaus [8]. This concept inspired us to extend the ordinary Fourier transform to a moving version. In Chapter 3, the derivation and construction of the transform is presented and a corresponding transformation to return to the time domain is introduced. The moving Fourier transform as well as the resulting moving periodogram have been thoroughly investigated in Chapter 4 and 5 concerning their asymptotic properties. We have even gone further taking account of moving spectral means and their asymptotic characteristics. Chapter 6 includes some philosophical aspects on possible modifications of the moving Fourier transform. Application of the moving transform to bootstrapping has been discussed in Chapter 7. Chapter 7 also exposes the need for an appropriate estimator for the time varying spectral density. The construction of an estimator as well as the proof of adequateness has been done in Chapter 8. Finally, Chapter 9 looks at the bootstrap data emerging from the moving wild TFT bootstrap and discovers that the autocovariance structure is mimicked well. We have stochastic convergence to the correct autocovariance function, uniformly in lag h, when h is smaller than the window width used for the transformation. Chapter 10 is dedicated to the investigation whether an analogon of the Fast Fourier transform can be constructed to reduce numerical complexity. The final Chapter 11 presents a simulation study including the moving Fourier transform and the bootstrap procedures developed in Chapter 7. We construct simultaneous confidence bands for the autocorrelation function as well as for the autocovariance function of locally stationary data and investigate their performance with respect to different bootstrap procedures and different data generating processes.

CHAPTER 2

Locally stationary processes

2.1 The concept of local stationarity

2.1.1 Asymptotic theory

If $X_1, \dots, X_T$ are the observations at hand, letting T tend to infinity which means extending the process into the future, does, in case of a non-stationary process, not yield any more information. Thus, asymptotic considerations have to be adequately adapted in the sense that letting T tend to infinity does indeed reveal more information on the process. Exemplarily, the process

$$X_t = g^*(t)X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad t = 1, \dots, T,$$

with some function $g^* : \{1, 2, \dots, T\} \to \mathbb{R}$ is considered. Currently we have information on the unknown function $g^*(t)$ on the grid $\{1, 2, \dots, T\}$. Dahlhaus [6],[7] and [8] sets down the asymptotic theory not by assuming the function g^* to be observed for a longer period of time on an extending grid with constant grid width, but to be observed on a finer and finer grid on the same interval. This is done by rescaling the unknown function g^* to the interval $(0, 1]$ in the way that the rescaled function g now reads $g^*(t) = g\left(\frac{t}{T}\right)$ and thus

$$X_{t,T} = g\left(\frac{t}{T}\right) X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad t = 1, \dots, T. \tag{2.1}$$

We can see that the larger T grows the finer the grid on which we observe the function g gets, but the domain of the rescaled function g remains to be the interval $(0, 1]$. This means that more and more information on the function g is available as T tends to infinity. Still, Dahlhaus [6] indicates to exercise caution when interpreting the asymptotics. The big difference to stationary time series is that the approach using rescaling is purely an abstraction for judging statistical inference. As a consequence it makes for example no sense to ask for a real data example that fulfills the rescaling property introduced in Equation (2.1).

2.1.2 Definition of locally stationary processes

Dahlhaus [8] introduced the time-varying spectral representation of locally stationary processes in analogy to stationary processes. Easier to work with, however, is the equivalent time varying MA(∞)-representation of locally stationary time series as given by Dahlhaus [10], Eq. (11).

Following Sergides [49], all forthcoming calculations are based on the definition, that a triangular array $\{X_T\}_{T\in\mathbb{N}} = \{X_{t,T}, t = 1, \dots, T\}_{T\in\mathbb{N}}$ is called locally stationary, if the processes have a tvMA(∞)-representation with time varying coefficients $a_{t,T}(j)$ (fulfilling certain smoothness conditions stated below):

$$X_{t,T} = \sum_{j=-\infty}^{\infty} a_{t,T}(j)\varepsilon_{t-j},$$

with $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ being independent, identically distributed random variables with zero mean and variance 1. Hence, we do only consider centered time series and focus on changes in the autocovariance structure. The exact definition used is

---

Definition 2.1 (tvMA(∞) representation of locally stationary processes). *→ Dahlhaus [10], Ass. 2.1 , Dahlhaus and Polonik [15], Ass. 2.1, Sergides [49], Ass. 1*

A sequence of stochastic processes $X_{t,T}$, $t = 1, \dots, T$, is called locally stationary if there exists a representation

$$X_{t,T} = \sum_{j=-\infty}^{\infty} a_{t,T}(j)\varepsilon_{t-j}, \tag{2.2}$$

where the following holds

(a) $\varepsilon_t \overset{iid}{\sim} (0,1)$ *with finite fourth moment* $E\varepsilon_t^4 < \infty$,

(b) $\sup_t |a_{t,T}(j)| \leq \frac{K}{l(j)}$, *and*

let $\{l(j)\}$ *be a positive sequence with* $l(j) := \begin{cases} 1, & |j| \leq 1 \\ |j| log^{1+\kappa}|j|, & |j| > 1 \end{cases}$

for some $\kappa > 0$.

(c) There exist functions $a(\cdot, j) : (0,1] \to \mathbb{R}$, $j \in \mathbb{Z}$, *with*

(i) $\sup_t \left|a_{t,T}(j) - a(\frac{t}{T}, j)\right| \leq \frac{K}{Tl(j)}$.

(ii) $|a(u,j) - a(v,j)| \leq \frac{K|u-v|}{l(j)}$

(iii) $\sup_u \left|\frac{\partial^i a(u,j)}{\partial u^i}\right| \leq \frac{K}{l(j)}$, $i = 0, 1, 2, 3.$

Remark 2.1
Throughout this thesis we use K and C as generic positive constants not depending on any other quantities if not stated otherwise.

Remark 2.2

1. *The rather complicated construction using the coefficients $a_{t,T}(j)$ and $a(u,j)$ is justified in Dahlhaus [10], p.454 and Dahlhaus and Polonik [14], Remark 2.12 (i). The function $a(\cdot,j)$ is needed for rescaling and to impose necessary smoothness conditions in the time direction, while the additional use of $a_{t,T}(j)$ makes the class rich enough to cover interesting cases, such as tvAR models.*
2. *Despite the fact that Definition 2.1 appears to admit only homoscedastic innovations, Dahlhaus and Polonik [14], Remark 2.12 (ii) state that a time varying scaling factor of the innovations may be included in the coefficients $a_{t,T}(j)$.*

Remark 2.3
For Lemma 5.1 we use a slightly different assumption to Definition 2.1(b): Let $\{l(j)\}$ be a positive sequence with

$$\sum_{k=-\infty}^{\infty} \sum_{j>k} \frac{1}{l(j)} < \infty.$$

2.1.3 Stationary approximation

Taking up the wording of Sergides [49] that a locally stationary process is a stochastic process whose second order structure varies slowly over time, it feels intuitive to consider this process stationary within a local neighbourhood of some point in time. We are now going to formally clarify what is meant by 'changing slowly'.
Based on Sergides [49], Dahlhaus and Subba Rao [16] and Subba Rao [50] we define, for some $u \in (0,1)$, the stationary process $\tilde{X}_t(u)$ by

$$\tilde{X}_t(u) := \sum_{j=-\infty}^{\infty} a(u,j)\varepsilon_{t-j}, \tag{2.3}$$

where $a(\cdot,j)$ are the functions used in the definition of a locally stationary process and the errors are those of the locally stationary process

$$X_{t,T} = \sum_{j=-\infty}^{\infty} \alpha_{t,T}(j)\varepsilon_{t-j}.$$

Now, comparing these two processes yields (cf. Sergides [49], Equation (1.1.19))

$$\left|X_{t,T} - \tilde{X}_t(u)\right| \le K \left(\left| \frac{t}{T} - u \right| + \frac{1}{T} \right) \sum_{j=-\infty}^{\infty} \frac{|\varepsilon_{t-j}|}{l(j)},$$

which implies

$$X_{t,T} = \tilde{X}_t(u) + O_P \left(\left| \frac{t}{T} - u \right| + \frac{1}{T} \right). \tag{2.4}$$

$\tilde{X}_t(u)$ is a stationary approximation to $X_{t,T}$ in some local neighbourhood of u (Note: u is the time parameter in rescaled time). That is, if $\frac{t}{T}$ is close to u – meaning we are only looking at $X_{t,T}$ in some local neighbourhood of u – $X_{t,T}$ and $\tilde{X}_t(u)$ are very close, and $X_{t,T}$ is 'basically' stationary.
As it can be seen above, the degree of approximation depends on the rescaling factor T and the deviation $\left|\frac{t}{T} - u\right|$ (cf. Dahlhaus and Subba Rao [16], p.4).

To study the behaviour of $\{X_{t,T}\}$, we will follow Sergides [49] and use the process $Z_{t,T}(u) := X_{t,T} - \tilde{X}_t(u)$, which then has the tvMA(∞)-representation

$$Z_{t,T}(u) = \sum_{j=-\infty}^{\infty} \left(a_{t,T}(j) - a(u,j)\right) \varepsilon_{t-j}, \tag{2.5}$$

with ε_t, $a_{t,T}(j)$ and $a(\cdot, j)$ from Definition 2.1.

2.1.4 Time varying spectral density and covariance

In order to work theoretically with the concept of locally stationary time series, it needs to be clarified what is meant by the spectral density or the covariance of a locally stationary process and how these functions relate to their stationary counterparts.
First, the concepts time varying spectral density and time varying covariance are introduced.

Definition 2.2 (time varying spectral density and covariance).
→ Dahlhaus and Polonik [15]

Let $X_{t,T}$ be a locally stationary process.

(a) The time varying spectral density of a locally stationary process is given by

$$f(u,\lambda) = \frac{1}{2\pi} |A(u,\lambda)|^2, \tag{2.6}$$

with $A(u,\lambda) := \sum_{j=-\infty}^{\infty} a(u,j) e^{-i\lambda j}$.

(b) The Fourier transform of the time-varying spectral density (at rescaled time u)

$$c(u,h) := \int_{-\pi}^{\pi} f(u,\lambda)e^{i\lambda h}d\lambda = \sum_{j=-\infty}^{\infty} a(u,h+j)a(u,j) \tag{2.7}$$

denotes the time varying covariance of lag h, $h \in \mathbb{Z}$ (at rescaled time u).

As Sergides [49] notes, both the time varying spectral density and the time varying covariance are the corresponding functions of the stationary approximation $\tilde{X}_t(u)$ of $X_{t,T}$ at time u. However, as $\tilde{X}_t(u)$ does not equal but only serves as an approximation of $X_{t,T}$ at any time other than u, the actual autocovariance function of $X_{t,T}$ will only for $\frac{t}{T} = u$ equal the corresponding time varying covariance. Referring to Dahlhaus [10], Equation (17), we have the following coherence between the time varying covariance function and the covariance function of the locally stationary process.

$$\operatorname{Cov}(X_{\lfloor uT \rfloor,T}, X_{\lfloor uT \rfloor+h,T}) \quad = \quad c(u,h) + O\left(\frac{1}{T}\right) \tag{2.8}$$

uniformly in u and h.

Remark 2.4
In the following, the time-varying covariances $c(u,h)$ are assumed to be absolutely summable for every $u \in [0,1]$.

Without asymptotics, one can only identify a finite number of covariances within any approximative stationary environment of $X_{t,T}$ and thus, as Dahlhaus [6] mentions, the spectral density is not uniquely determined. Just like in the case of stationary processes this problem can be solved by employing the asymptotics introduced by Dahlhaus [8] as in each approximately stationary environment more and more data becomes available. Due to that asymptotic approach Dahlhaus [7] is now able to obtain a uniqueness property of the time varying spectral density. To be more specific on this uniqueness we need to consider the Wigner-Ville spectrum, which has already been introduced in the previous section:
We define, for fixed T, $\lambda \in [-\pi,\pi]$ and $u \in [0,1]$ the Wigner-Ville spectrum of a locally stationary process $\{X_{t,T}\}$ as

$$f_T(u,\lambda) := \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \operatorname{Cov}(X_{\lfloor uT-\frac{s}{2} \rfloor,T}, X_{\lfloor uT+\frac{s}{2} \rfloor,T})e^{-i\lambda s}. \tag{2.9}$$

The Wigner-Ville spectrum is a real-valued function of time and frequency. This is, as discussed before, one possibility of defining a time dependent spectrum.
Dahlhaus [7] proved that the time varying spectral density $f(u,\lambda)$ is uniquely determined and equals the limit of the Wigner-Ville spectrum.

Theorem 2.1 (L^2-convergence of $f_T(u,\lambda)$ to $f(u,\lambda)$).
→ *Dahlhaus [7], Theorem 2.2*

Let $X_{t,T}$ be a locally stationary process and $A(u,\lambda)$ uniformly Lipschitz continuous in both components with index $\alpha > \frac{1}{2}$.

We then have for all $u \in (0,1)$:

$$\int_{-\pi}^{\pi} |f_T(u,\lambda) - f(u,\lambda)|^2 d\lambda = o(1).$$

Remark 2.5
Continuous differentiability of A with respect to u and λ is sufficient for $A(u,\lambda)$ being uniformly Lipschitz-continuous (as in Haug [24], Definition 2.7) with $\alpha > \frac{1}{2}$.

Despite the fact that the spectral representation of a non-stationary process is not unique (see Section 1), the above theorem points out that if there exists a tvMA(∞)-representation as in Definition (2.1) of a locally stationary process with $a(u,\lambda)$ (and therefore $A(u,\lambda)$) sufficiently smooth, the time varying spectral density $f(u,\lambda)$ is asymptotically unique. It is determined by the whole triangular array and equals the limit of the Wigner-Ville spectrum, cf. Dahlhaus [7], p.143.

2.1.5 Dependence structure of a locally stationary process

Let $X_{t,T}$ be a locally stationary process as in Definition 2.1. For stationary time series with absolutely summable autocovariance function γ we have $|\gamma(h)| \to 0$ as $|h| \to \infty$. So it does seem only natural that locally stationary processes, as generalizations of stationary processes, do also have a decaying covariance structure as $|h| \to \infty$. From Dahlhaus [15], proof of Proposition 5.4, we obtain

$$c_T\left(\frac{t}{T}, h\right) := \mathrm{Cov}(X_{t,T}, X_{t+h,T}) = \sum_{j=-\infty}^{\infty} a_{t,T}(j) a_{t+h,T}(j+h). \tag{2.10}$$

They then prove (Equation (51)) that the above relation yields

$$\sup_t \left| c_T\left(\frac{t}{T}, h\right)\right| \leq \sum_{j=-\infty}^{\infty} \frac{K}{l(j)l(j+h)} \leq \frac{K}{l(h)},$$

with $\sup_t |a_{t,T}(j)| \leq \frac{K}{l(j)}$ (from Definition 2.1 (b)). The last inequality results from the fact that

$$\sup_{j\in\mathbb{Z}} \frac{1}{l(j+h)} = \frac{1}{l(h)}.$$

Considering the definition of l in Definition 2.1 (b), we can see that $\frac{K}{l(h)}$ converges to zero for $|h| \to \infty$. Thus, the following Lemma results:

Lemma 2.1.
$\to$ *Dahlhaus [15]*
The time varying covariance $c_T\left(\frac{t}{T}, h\right)$ *of a locally stationary process* $\{X_{t,T}\}$ *at time* $t = 1, \ldots, T$ *converges to zero for lags* $|h| \to \infty$:

$$c_T\left(\frac{t}{T}, h\right) = o(1).$$

The notation c_T is borrowed from Neumann and von Sachs [40].

CHAPTER 3

Adapting the Fourier transformation

A locally stationary process $\{X_{t,T}\}$ can, as the name suggests, be locally (i.e. in a small environment U) approximated by a stationary process, see Equation (2.4) for a formal description. In order to preserve the changing nature of a locally stationary time series for the frequency domain it may therefore seem only natural to apply the Fourier transformation to each environment. The local moving Fourier transformation is introduced as an intuitive and numerically cheap procedure to meet these needs.

3.1 Prerequisites

Concerning the sample size T and the segments' length $2m + 1$, we require the following conditions to hold:

- $m \to \infty$ (for $T \to \infty$).
- $\frac{m^{\frac{3}{2}}}{T} \to 0$ (for $T \to \infty$) i.e. the sample size increases considerably faster than the window size.

For the sake of simplicity, we introduce the following concepts for $j \in \mathbb{Z}$:

$$mod(j) := \begin{cases} m, & \text{if } m \text{ is a factor of } j \in \mathbb{Z}, \\ j \mod (m), & j > 0 \ \wedge m \,|\, j, \\ m - [(-j) \mod (m)], & j < 0 \ \wedge m \nmid j. \end{cases} \tag{3.1}$$

$$div(j) := \left\lceil \frac{j}{m} \right\rceil. \tag{3.2}$$

$$\text{Then,} \quad j = mod(j) + (div(j) - 1)m. \tag{3.3}$$

3.2 The local moving Fourier transform

In the following, no distinction is made between the actual process and the realization of the process. We heuristically describe how the local moving Fourier transform is developed. The formal definition can be found in Definition 3.2.
In order to simplify notation, we will from here on assume to not only have $\{X_{t,T}\}_{t\in[1,T]}$ – our time series of length T we wish to transform – available, but also additionally a sufficient (depending on the window size $2m+1$) number of preceding and succeeding observations, i.e.

$$X_{-m+1,T}, \dots, X_{1,T}, \dots, X_{T,T}, \dots, X_{T+m,T}.$$

The reason is, as already stated above, the advantage of keeping notation simple enough to be able to fully focus on the new way of transforming $\{X_{t,T}\}_{t\in[1,T]}$.

As a locally stationary process can locally (in an environment getting larger at some sufficiently slower pace than T) be considered stationary, we can look at

$$\{X_{t,T}\}_{t\in[t_0-C\cdot m, t_0+C\cdot m]}, \tag{3.4}$$

$C > 0$, $t_0 \in [1, T]$, as an approximately stationary time series. This especially holds true for the sequence

$$X_{t_0-m,T}, \dots, X_{t_0+m,T}. \tag{3.5}$$

Without taking into account that there are more observations than those $2m+1$we now apply the usual Fourier transform to the stationary sequence (3.5).

$$\begin{aligned} \mathcal{F}(X_{t_0-m,T}, \dots, X_{t_0+m,T}; \lambda_k) &:= \frac{1}{\sqrt{2m+1}} \sum_{l=0}^{2m} X_{l+t_0-m,T} e^{-il\lambda_k} \\ &= \frac{1}{\sqrt{2m+1}} \sum_{l=t_0-m}^{t_0+m} X_{l,T} e^{-il\lambda_k} e^{i(t_0-m)\lambda_k}, \end{aligned} \tag{3.6}$$

with $1 \le k \le 2m$, and $\lambda_k := \frac{2\pi k}{2m+1}$ denoting the Fourier frequencies.

We now concentrate on the shifted stretch

$$X_{t_0-m+1,T}, \dots, X_{t_0+m+1,T} \tag{3.7}$$

and calculate $\mathcal{F}(X_{t_0-m+1,T}, \dots, X_{t_0+m+1,T}; \lambda_k)$, $k = 1, \dots, m$.
The motive for shifting and doing another Fourier transform of a slightly different stretch becomes more obvious when noticing that the observations (3.7) also fall within (3.4). Heuristics then indicate that the Fourier coefficients of (3.5) and (3.7) should also possess similar statistical properties. Accordingly, Fourier coefficients stemming from adjacent stretches may be interchanged without major changes to

statistical inference.
We use these heuristics to come up with the following construction: Instead of calculating all $2m+1$ Fourier coefficients for every single stretch, we calculate just one Fourier coefficient per stretch and then move on to the next stretch. That is, the centre of the stretch considered is no longer a fixed t_0 as in the transformation (3.6), but depends on the index k of the Fourier frequency λ_k considered:

$$\mathcal{F}(X_{k-m,T},\ldots,X_{k+m,T};\lambda_k) = \frac{1}{\sqrt{2m+1}}\sum_{l=k-m}^{k+m} X_{l,T}e^{-il\lambda_k}e^{i(k-m)\lambda_k}, \quad 1 \le k \le m, \tag{3.8}$$

with $\lambda_k := \frac{2\pi k}{2m+1}$ denoting Fourier frequencies.
Exemplarily, we calculate, say $\mathcal{F}(X_{1-m,T},\ldots,X_{1+m,T};\lambda_1)$ for the first stretch and then move on to the adjacent stretch and calculate $\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2)$. Consequently, as the Fourier coefficients of adjacent stretches are interchangeable, the Fourier coefficients $\mathcal{F}(X_{1-m,T},\ldots,X_{1+m,T};\lambda_1)$, $\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2)$ are, from a statistical point of view, as good as $\mathcal{F}(X_{1-m,T},\ldots,X_{1+m,T};\lambda_1)$, $\mathcal{F}(X_{1-m,T},\ldots,X_{1+m,T};\lambda_2)$. We then move on to the next stretch, from which we calculate $\mathcal{F}(X_{3-m},\ldots,X_{3+m};\lambda_3)$. Again, $\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2)$, $\mathcal{F}(X_{3-m,T},\ldots,X_{3+m,T};\lambda_3)$ should, concerning statistical properties, be as good as $\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2)$, $\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_3)$.
So intuitively, instead of

$$\mathcal{F}(X_{1-m,T},\ldots,X_{1+m,T};\lambda_1),\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2),\mathcal{F}(X_{3-m,T},\ldots,X_{3+m,T};\lambda_3),$$

we can also use

$$\mathcal{F}(X_{j-m,T},\ldots,X_{j+m,T};\lambda_1),\mathcal{F}(X_{j-m,T},\ldots,X_{j+m,T};\lambda_2),\mathcal{F}(X_{j-m,T},\ldots,X_{j+m,T};\lambda_3),$$

$j = 1,2,3$, basically without any change in statistical characteristics. Shifting the time window of length $2m+1$, $m-1$ times (each time generating an additional Fourier coefficient stemming from the actual stretch) we finally obtain m Fourier coefficients

$$\mathcal{F}(X_{1-m,T},\ldots,X_{1+m,T};\lambda_1),\mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2),\ldots,\mathcal{F}(X_{0,T},\ldots,X_{2m,T};\lambda_m) \tag{3.9}$$

This is still not a transformation of a time series of length T, but captures the basic idea!

Note that the observations used for those Fourier coefficients are $X_{1-m,T},\ldots,X_{2m,T}$, with $X_{0,T},\ldots,X_{1+m}$ being part of each of the $m-1$ Fourier transforms. The set (3.9) therefore consists of Fourier coefficients of basically $X_{0,T},\ldots,X_{m+1}$, as these are the most influential observations on the coefficients. When intending to refer to the set (3.9) as local moving Fourier coefficients at some time k, it is thus apparent that we should speak of (3.9) as the local moving Fourier coefficients at time $k = \lfloor \frac{m}{2} \rfloor + 1$. To ease the understanding of the following definition, we will consider the next two points in time and look at the local moving Fourier coefficients at the time $k = \lfloor \frac{m}{2} \rfloor + 2$ and $k = \lfloor \frac{m}{2} \rfloor + 3$:

$$\mathcal{F}(X_{1,T},\ldots,X_{2m+1,T};\lambda_1), \mathcal{F}(X_{2-m,T},\ldots,X_{2+m,T};\lambda_2), \quad \ldots, \quad \mathcal{F}(X_{0,T},\ldots,X_{2m,T};\lambda_m)$$
$$\mathcal{F}(X_{1,T},\ldots,X_{2m+1,T};\lambda_1), \quad \mathcal{F}(X_{2,T},\ldots,X_{2m+2,T};\lambda_2), \quad \ldots, \quad \mathcal{F}(X_{0,T},\ldots,X_{2m,T};\lambda_m).$$

We can see that by moving on in time, the coefficients, starting at frequency λ_1 get replaced by more recent coefficients at the same frequency. This scheme continues until we get to time $k = \lfloor \frac{m}{2} \rfloor + m + 1$:

$$\mathcal{F}(X_{1,T},\ldots,X_{2m+1,T};\lambda_1), \quad \mathcal{F}(X_{2,T},\ldots,X_{2m+2,T};\lambda_2), \quad \ldots, \quad \mathcal{F}(X_{m,T},\ldots,X_{3m,T};\lambda_m),$$

and then starts anew, substituting $\mathcal{F}(X_{1,T},\ldots,X_{2m+1,T};\lambda_1)$ by the more recent coefficient $\mathcal{F}(X_{m+1,T},\ldots,X_{3m+1,T};\lambda_1)$.
The formal definition of the local moving Fourier coefficients at time k, $MF_k(\lambda_j)$, is as follows:

Definition 3.1 (Local moving Fourier coefficients).
The local moving Fourier coefficients at time k for frequencies λ_l, $l = 1,\ldots,m$, are given by

$$MF_k(\lambda_l) \quad := \quad \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} X_{l+\left(div\left(k-\lfloor \frac{m}{2} \rfloor\right)-\mathbb{1}_{\{l \geq mod(k-\lfloor \frac{m}{2} \rfloor)\}}\right)m-m+t,T} e^{-it\lambda_l}.$$

Furthermore,

$$\begin{aligned} MF_k(\lambda_{2m+1-j}) &:= \overline{MF_k(\lambda_j)}, \; j = 0,\ldots,m, \\ \textit{and } MF_k(\lambda_0) &:= 0. \end{aligned} \tag{3.10}$$

The operators mod and div are defined according to (3.1) and (3.2).

The reason for defining $MF_k(\lambda_{2m+1-j}) := \overline{MF_k(\lambda_j)}$, $j = 0,\ldots,m$, and $MF_k(\lambda_0) := 0$ is given in Remark 3.4.

Remark 3.1
The local moving Fourier coefficients at time $k + \lfloor \frac{m}{2} \rfloor$ and frequency λ_l, $l = 1,\ldots,m$ are depending on

$$X_{l+0-m+\left[div(k)-\mathbb{1}_{\{l \geq mod(k)\}}\right]m}, X_{l+1-m+\left[div(k)-\mathbb{1}_{\{l \geq mod(k)\}}\right]m}, \ldots, X_{l+m+\left[div(k)-\mathbb{1}_{\{l \geq mod(k)\}}\right]m}.$$

Removing the indicator function and using $k = mod(k) + [div(k) - 1]m$, this is for $l < mod(k)$

$$X_{l+k-mod(k)}, X_{l+k-mod(k)+1}, \ldots, X_{l+k-mod(k)+2m}$$

and for $l \geq mod(k)$

$$X_{l+k-mod(k)-m}, X_{l+k-mod(k)-m+1}, \ldots, X_{k-mod(k)+m}.$$

That is, the set

$$\{MF_{k+\lfloor \frac{m}{2} \rfloor}(\lambda_l)\}_{l=1,\ldots,m}$$

incorporates the observations $X_{k-m}, \ldots, X_{k+2m-1}$. Of those $3m$ observations,

$$X_{k-1}, \ldots, X_{k+m}$$

occur in all of the local moving Fourier coefficients. In other words, the set of local moving Fourier coefficients basically describes the time series in an environment of time $k + \lfloor \frac{m}{2} \rfloor$.

We now extend our construction in order to finally be able to fully transform $\{X_{t,T}\}_{t\in[1,T]}$. This is done by starting with $X_{1-m,T}, \ldots, X_{1+m,T}$ and shifting the time window of length $2m+1$ not just $m-1$ times, but $T-1$ times (each time generating an additional Fourier coefficient stemming from the actual stretch). By doing so, we finally obtain T Fourier coefficients. Some attention, however, has to be paid to the frequencies, as we only calculate the coefficients for frequencies $\lambda_{mod(k)}$, which guarantees the index to remain between 1 and m (see also Remark 3.4).

Definition 3.2 (Moving Fourier transform).

Let $X_{t,T}$ be a locally stationary process as in Definition 2.1. The moving Fourier coefficients c_k $(1 \leq k \leq T)$ of $X_{t,T}$ are then defined by

$$\begin{aligned} c_k := \mathcal{F}^{div(k)-1}\left(\lambda_{mod(k)}\right) &:= \mathcal{F}(X_{k-m,T}, \ldots, X_{k+m,T}; \lambda_{mod(k)}) \\ &= \frac{1}{\sqrt{2m+1}} \sum_{l=k-m}^{k+m} X_{l,T} e^{-il\lambda_{mod(k)}} e^{i(k-m)\lambda_{mod(k)}}, \end{aligned} \tag{3.11}$$

with $\lambda_{mod(k)} := \frac{2\pi\, mod(k)}{2m+1}$ denoting the Fourier frequencies and the operator mod according to (3.1).

Following the algorithm (3.11), hence, yields the moving Fourier coefficients

$$c_1, \ldots, c_T,$$

which code the time series $X_{1,T}, \ldots, X_{T,T}$. Due to the continuous shifting, local structural information is preserved.

Remark 3.2
We will speak of <u>*moving Fourier coefficients*</u> *when referring to* $c_1, \ldots, c_T$*, while we use the term* <u>*local moving Fourier coefficients at time k*</u> *to indicate that we are locally, at one point in time, looking at* m *of the moving Fourier coefficients and rearranging them according to their frequencies. Hence, local moving Fourier coefficients refers to the set* $MF_k(\lambda_1), \ldots, MF_k(\lambda_m)$ *of* m *rearranged moving Fourier coefficients at some point in time* k*. They relate to each other by*

$$MF_k(\lambda_l) = c_{l+\left[div\left(k-\lfloor\frac{m}{2}\rfloor\right)-\mathbb{1}_{\{l\geq mod\left(k-\lfloor\frac{m}{2}\rfloor\right)\}}\right]m} \tag{3.12}$$

Remark 3.3
The additional notation $\mathcal{F}^{div(k)-1}\left(\lambda_{mod(k)}\right)$ *instead of* c_k *in Definition 3.2 is introduced to ease the understanding of the concept of the moving Fourier transform. When constructing local moving Fourier coefficients we combine moving Fourier coefficients located around the point in time considered and do some rearranging. We might therefore encounter a set of coefficients which consists of the moving Fourier coefficients, say,* $c_{m+1}, \ldots, c_{m+17}, c_{18}, \ldots, c_m$*. Due to the moving, some coefficients are 'older' than others. Sorting with respect to the currentness of the coefficients yields* $c_{18}, \ldots, c_{m+17}$*. The notation with the calligraphic* F *is chosen to prominently display via the superscript where the discontinuity concerning the up-to-dateness of the coefficients is. In the example, we would write*

$$\mathcal{F}^0\left(\lambda_1\right), \mathcal{F}^0\left(\lambda_2\right), \ldots, \mathcal{F}^0\left(\lambda_{17}\right), \mathcal{F}^{-1}\left(\lambda_{18}\right), \ldots, \mathcal{F}^{-1}\left(\lambda_m\right).$$

The notation is used in Theorem 5.4.

Remark 3.4
We have restricted the range of k *to* $\{1, \ldots, m\}$*. The reason we imply this restriction is as follows:*
The spectral density of a stationary process (of length $2m+1$*) is uniquely specified by values within the interval* $[0, \pi]$*. This means that in order to extract all information on the spectral density, only the Fourier coefficients corresponding to the frequencies* $\lambda_0, \ldots, \lambda_m$ *are needed. The remaining Fourier coefficients (in the stationary case of a time series of length* $2m+1$*) follow using symmetry arguments and the conjugated complexes of the already calculated coefficients. In detail: Suppose we are given a time series of length* $2m+1$ *and have calculated* $\mathcal{F}(X_{1-m,T}, \ldots, X_{1+m,T}; \lambda_1)$*,* $\mathcal{F}(X_{2-m,T}, \ldots, X_{2+m,T}; \lambda_2)$*, …,* $\mathcal{F}(X_{0,T}, \ldots, X_{2m,T}; \lambda_m)$ *We may now write*

$$\mathcal{F}(X_{j-m,T}, \ldots, X_{j+m,T}; \lambda_{2m+1-j}) = \overline{\mathcal{F}(X_{j-m,T}, \ldots, X_{j+m,T}; \lambda_j)}, \quad \text{for } j = 0, \ldots, m.$$

$\mathcal{F}(X_{-m,T}, \ldots, X_{m,T}; \lambda_0)$ *carries information on the mean. As we start out with a time series with mean zero, we may set these to zero in order for the back transformed time series to be centred as well. This is the reason why we only gather Fourier coefficients for frequencies* $\lambda_1, \ldots, \lambda_m$ *from the given time series* $\{X_{t,T}\}$*.*

Remark 3.5
The assumption that we have data

$$X_{-m+1,T}, \dots, X_{T+m,T}$$

available, i.e. a time series of length $T+2m$ instead of just a time series of length T, can easily be abandoned by slightly changing the scheme of transformation, employing the ordinary Fourier transform (cf. (3.6)) for the first and last stretch and retrieving not one, but m Fourier coefficients.

However, the question is not only how to transform the data to the frequency domain, but how to obtain (bootstrapped) time series data from the moving Fourier coefficients. This is where the special definition of the local moving Fourier coefficients are of great importance. The procedure will be explained in Section 3.3.

Of course, the definition of local moving Fourier coefficients implies that there is also a moving periodogram.

Definition 3.3 (Local moving periodogram).
Consider a locally stationary process $X_{t,T}$ according to Definition 2.1 and its local moving Fourier coefficients at time k as in Definition 3.1. The local moving periodogram $MI_k : [0, 2\pi] \to \mathbb{R}$ at time k is then defined by

$$MI_k(\lambda_j) := |MF_k(\lambda_j)|^2, \tag{3.13}$$

with $\lambda_j := \frac{2\pi j}{N}$, $j = 1, \dots, m$, denoting the Fourier frequencies and $k = 1, \dots, T$.

The local moving periodogram can be periodically extended.

Remark 3.6
The intention of introducing the new term moving periodogram is to create a sequence of local periodograms which 'move' through the time series. At each point in time k, however, the local moving periodogram equals the local periodogram $I_{2m+1,X}\left(\frac{k}{T}, \cdot\right)$ used by Sergides [49]. The local periodogram is defined by

$$I_{2m+1,X}(u, \lambda) := \frac{1}{2\pi(2m+1)} \left| \sum_{l=0}^{2m} X_{l-m+\lfloor uT \rfloor} e^{-i\lambda l} \right|^2 \tag{3.14}$$

and we have $2\pi I_{2m+1,X}\left(u, \lambda_{mod(\lfloor uT \rfloor)}\right) = |c_{\lfloor uT \rfloor}|^2$.

3.3 Moving inverse Fourier transform

3.3.1 Principle of construction

In the following section, we will construct a transformation from the frequency domain to the time domain.

We start witha sequence of length $T+2\left\lfloor \frac{m}{2}\right\rfloor$ of arbitrary complex random variables in the frequency domain:

$$c_{1-\lfloor \frac{m}{2}\rfloor}, \ldots, c_1, \ldots, c_{T+\lfloor \frac{m}{2}\rfloor}.$$

Now, for each point in time k, we can select $c_{k-\lfloor \frac{m}{2}\rfloor}, \ldots, c_{k+\lceil \frac{m}{2}\rceil-1}$ and rearrange them in the manner of Remark 3.2, Equation 3.12. Having done so, we call the elements of the set

$$MF_k^{(c)}(\lambda_1), \ldots, MF_k^{(c)}(\lambda_m).$$

(We chose the notation like this, because the construction mirrors the construction of the local moving Fourier coefficients MF from the moving Fourier coefficients c.) Analogously to Definition 3.1, the definition of the local moving Fourier coefficients, we use the complex conjugated values for the missing frequencies:

$$0, MF_k^{(c)}(\lambda_1), \ldots, MF_k^{(c)}(\lambda_m), \overline{MF_k^{(c)}(\lambda_m)}, \overline{MF_k^{(c)}(\lambda_{m-1})}, \ldots, \overline{MF_k^{(c)}(\lambda_1)}.$$

We now apply the ordinary inverse Fourier transformation of length $2m+1$ at time k to this data.
For all other points in time we proceed analogously.

Summing up, the idea underlying this transformation includes shifting a window of length m along the given sequence $c_{1-\lfloor \frac{m}{2}\rfloor}, \ldots, c_1, \ldots, c_{T+\lfloor \frac{m}{2}\rfloor}$, rearranging the elements resulting from each shift to create the $MF^{(c)}$'s and applying the ordinary inverse Fourier transform of length $2m+1$ at the corresponding time to each of the sets, which results in T elements in the time domain.
The formal definition of the new transformation is as follows:

Definition 3.4 (Moving inverse Fourier transform).

Let $c_{1-\lfloor \frac{m}{2}\rfloor}, \ldots, c_{T+\lfloor \frac{m}{2}\rfloor}$ be elements in the frequency domain. The transformation yielding a sample in the time domain is called moving inverse Fourier transform and is defined by

$$\begin{aligned} X_{t,T}^{back} &:= \mathcal{F}^{-1}\left(MF_t^{(c)}(\lambda_1), MF_t^{(c)}(\lambda_2), \ldots, MF_t^{(c)}(\lambda_m); t\right) \\ &:= 0 \cdot e^{it\lambda_0} + \frac{1}{\sqrt{2m+1}} \sum_{l=1}^{m} c_{l+\left(div\left(t-\lfloor \frac{m}{2}\rfloor\right)-\mathbb{1}_{\{l\geq mod(t-\lfloor \frac{m}{2}\rfloor)\}}\right)m} e^{i\lambda_l t} \qquad (3.15) \\ &+ \frac{1}{\sqrt{2m+1}} \sum_{l=1}^{m} \overline{c}_{l+\left(div\left(t-\lfloor \frac{m}{2}\rfloor\right)-\mathbb{1}_{\{l\geq mod(t-\lfloor \frac{m}{2}\rfloor)\}}\right)m} e^{-i\lambda_l t} \qquad (3.16) \end{aligned}$$

with $\lambda_k := \frac{2\pi k}{N}$, $k=0,\ldots,m$, denoting the Fourier frequencies and $t=1,\ldots,T$.

---

k	$0, MF_k^{(c)}(\lambda_1), MF_k^{(c)}(\lambda_2), \dots, MF_k^{(c)}(\lambda_m)$	$MF_k^{(c)}(\lambda_{m+1}), \dots, MF_k^{(c)}(\lambda_{2m})$	$X_{k,T}^{back}$
1	$0, c_1, \dots, c_{\lceil \frac{m}{2} \rceil}, c_{1-\lfloor \frac{m}{2} \rfloor}, c_{2-\lfloor \frac{m}{2} \rfloor}, c_{3-\lfloor \frac{m}{2} \rfloor} \cdots, c_0,$	$\overline{c_0}, \dots, \overline{c_{1-\lfloor \frac{m}{2} \rfloor}}, \overline{c_{\lceil \frac{m}{2} \rceil}}, \dots, \overline{c_1}$	$\Rightarrow X_{1,T}^{back}$
2	$0, c_1, \dots, c_{\lceil \frac{m}{2} \rceil}, c_{\lceil \frac{m}{2} \rceil+1}, c_{2-\lfloor \frac{m}{2} \rfloor}, c_{3-\lfloor \frac{m}{2} \rfloor} \cdots, c_0,$	$\overline{c_0}, \dots, \overline{c_{2-\lfloor \frac{m}{2} \rfloor}}, \overline{c_{\lceil \frac{m}{2} \rceil+1}}, \dots, \overline{c_1}$	$\Rightarrow X_{2,T}^{back}$
3	$0, c_1, \dots, c_{\lceil \frac{m}{2} \rceil}, c_{\lceil \frac{m}{2} \rceil+1}, c_{\lceil \frac{m}{2} \rceil+2}, c_{3-\lfloor \frac{m}{2} \rfloor}, \dots, c_0,$	$\overline{c_0}, \dots, \overline{c_{3-\lfloor \frac{m}{2} \rfloor}}, \overline{c_{\lceil \frac{m}{2} \rceil+2}}, \dots, \overline{c_1}$	$\Rightarrow X_{3,T}^{back}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$1+\lfloor \frac{m}{2} \rfloor$	$0, c_1, \dots, c_m,$	$\overline{c_m}, \dots, \overline{c_3}, \overline{c_2}, \overline{c_1}$	$\Rightarrow X_{1+\lfloor \frac{m}{2} \rfloor,T}^{back}$
$2+\lfloor \frac{m}{2} \rfloor$	$0, c_{m+1}, c_2, \dots, c_m,$	$\overline{c_m}, \dots, \overline{c_3}, \overline{c_2}, \overline{c_{m+1}}$	$\Rightarrow X_{2+\lfloor \frac{m}{2} \rfloor,T}^{back}$
$3+\lfloor \frac{m}{2} \rfloor$	$0, c_{m+1}, c_{m+2}, c_3, \dots, c_m,$	$\overline{c_m}, \dots, \overline{c_3}, \overline{c_{m+2}}, \overline{c_{m+1}}$	$\Rightarrow X_{3+\lfloor \frac{m}{2} \rfloor,T}^{back}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m+1+\lfloor \frac{m}{2} \rfloor$	$0, c_{m+1}, \dots, c_{2m},$	$\overline{c_{2m}}, \dots, \overline{c_{m+1}}$	$\Rightarrow X_{m+1+\lfloor \frac{m}{2} \rfloor,T}^{back}$
$m+2+\lfloor \frac{m}{2} \rfloor$	$0, c_{2m+1}, c_{m+2}, \dots, c_{2m},$	$\overline{c_{2m}}, \dots, \overline{c_{m+2}}, \overline{c_{2m+1}}$	$\Rightarrow X_{m+2+\lfloor \frac{m}{2} \rfloor,T}^{back}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T-\lceil \frac{m}{2} \rceil+1$	$0, c_{T-m+1}, c_{T-m+2}, c_{T-m+3}, \dots, c_T,$	$\overline{c_T}, \dots, \overline{c_{T-m+3}}, \overline{c_{T-m+2}}, \overline{c_{T-m+1}}$	$\Rightarrow X_{T-\lceil \frac{m}{2} \rceil+1,T}^{back}$
$T-\lceil \frac{m}{2} \rceil+2$	$0, c_{T+1}, c_{T-m+2}, c_{T-m+3}, \dots, c_T,$	$\overline{c_T}, \dots, \overline{c_{T-m+3}}, \overline{c_{T-m+2}}, \overline{c_{T+1}}$	$\Rightarrow X_{T-\lceil \frac{m}{2} \rceil+2,T}^{back}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
T	$0, c_{T+1} \dots, c_{T+\lfloor \frac{m}{2} \rfloor}, c_{T-\lceil \frac{m}{2} \rceil}, \dots, c_T,$	$\overline{c_T}, \dots, \overline{c_{T-\lceil \frac{m}{2} \rceil}}, \overline{c_{T+\lfloor \frac{m}{2} \rfloor}}, \dots, \overline{c_{T+1}}$	$\Rightarrow X_{T,T}^{back}$

Figure 3.1: Illustrating the moving inverse Fourier transform (Definition 3.4)

3.3.2 Inverse - quote unquote

The definition of the moving inverse Fourier transform enables us to obtain samples back in the time domain. We now have the possibility, having applied the moving Fourier transformation to time domain data, to go back to the time domain. The resulting observations will not be the original $X_{t,T}$ one starts out with, due to the shifting performed, but some time series with similar characteristics.
We are interested in how many (and which) of the original observations of the time series $\{X_{t,T}\}$ are used to construct one observation $X_{t,T}^{back}$?

Definition 3.4 then yields, using the coefficients c_k obtained as in Definition 3.2 and the coefficients MF_k as in Definition 3.1:

$$X_{t,T}^{back} \quad = \quad \mathcal{F}^{-1}\left(MF_t(\lambda_1), MF_t(\lambda_2), \ldots, MF_t(\lambda_m); t\right).$$

$X_{t,T}^{back}$ is, thus, constructed using the local moving Fourier coefficients at time t. According to Remark 3.1, the set $\{MF_t(\lambda_l)\}_{l=1,\ldots,m}$ incorporates the observations $X_{t-\lfloor\frac{m}{2}\rfloor-m}, \ldots, X_{t+\lceil\frac{m}{2}\rceil+m-1}$.
Therefore, in order to construct $X_{t,T}^{back}$, we need a stretch of observations of length $3m$, namely the stretch $X_{t-\lfloor\frac{m}{2}\rfloor-m}, \ldots, X_{t+\lceil\frac{m}{2}\rceil+m-1}$.

Of those $3m$ observations,

$$X_{t-\lfloor\frac{m}{2}\rfloor-1}, \ldots, X_{t+\lceil\frac{m}{2}\rceil}$$

occur in all of the local moving Fourier coefficients. That is, those $m+2$ observations have the main influence on $X_{t,T}^{back}$.

CHAPTER 4

Basic properties of the moving Fourier coefficients

The following section is devoted to determining distributional characteristics of the local moving Fourier coefficients. As our Definition 2.1 of locally stationary processes assumes errors with mean zero, we may directly state that the expected value of the moving Fourier transform equals zero.

All results also hold for $\lambda_j, j = 1, \ldots, 2m+1$, due to symmetry. The case of $j = 1, \ldots, m$ is considered w.l.o.g. for the sake of readability.

Lemma 4.1. *Let $X_{t,T}$ be a locally stationary process as in Definition 2.1. It holds that*

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E\left(MF_{\lfloor uT \rfloor}(\lambda_l)\right) = 0. \tag{4.1}$$

According to Definition 3.1, with $\zeta_{k,l} := div\left(k - \lfloor \frac{m}{2} \rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(k - \lfloor \frac{m}{2} \rfloor\right)\right\}}$,

$$MF_k^{\varepsilon}(\lambda_l) := \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \varepsilon_{l+\zeta_{k,l}m-m+t} e^{-it\lambda_l} \tag{4.2}$$

denotes the local moving Fourier coefficient at time k of the innovations at frequency λ_l. Analogously, we define for the stationary approximation at time k

$$MF_k^{\tilde{X}}(\lambda_l) := \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \tilde{X}_{l+\zeta_{k,l}m-m+t}\left(\frac{k}{T}\right) e^{-it\lambda_l}. \tag{4.3}$$

For the asymptotic considerations we will make use of the rescaling as introduced by Dahlhaus [8]. However, not only the relationship between the local moving Fourier

coefficients but also between the moving periodograms is relevant for further proofs throughout this work. If squared, the local moving Fourier coefficient at time $\lfloor uT \rfloor$ and frequency λ_l of the innovations yields the value of the moving periodogram of the innovations at that frequency: $|MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)|^2 = MI^{\varepsilon}_{\lfloor uT \rfloor,m}(\lambda_l)$.
The following approach is taken:
In a first step, the local moving Fourier transform of a stationary time series is considered and linked to the local moving Fourier transform of the innovations. This is done in Lemma 4.2. To prove this we extend the proof of Theorem 10.3.1 in Brockwell and Davis [3]. Lemma 4.3 then is a generalization to locally stationary time series.
In order to gain analogous results for the moving periodograms, Lemma 4.2 is used as a basis to prove Theorem 4.2. The culminating result is Theorem 4.3, which links the moving periodogram of a locally stationary time series to the moving periodogram of the innovations.
In the beginning some technical requirements are proved.

4.1 Technical basics

Proposition 4.1
Let $\tilde{X}_t(u)$ denote the stationary approximation of $X_{t,T}$ at time $\lfloor uT \rfloor$.
Let further $l = 1, \ldots, m$. With the Definitions given by (4.2) and (4.3),

(a)

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)|^2 = 1. \tag{4.4}$$

(b)

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)|^4 < \infty. \tag{4.5}$$

(c)

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)|^4 < \infty, \tag{4.6}$$

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF_{\lfloor uT \rfloor}(\lambda_l)|^4 < \infty. \tag{4.7}$$

Proof. Let $\zeta_{\lfloor uT \rfloor,l} = div\left(\lfloor uT \rfloor - \lfloor \frac{m}{2} \rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(\lfloor uT \rfloor - \lfloor \frac{m}{2} \rfloor\right)\right\}}$.

(a) Because $E\varepsilon_k\varepsilon_l = \delta_{k,l}$ it holds

$$\begin{aligned}
& \sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)|^2 \\
= & \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{2m+1} E\left(\sum_{t_1,t_2=0}^{2m} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_2} e^{-i(t_1-t_2)\lambda_l} \right) \\
= & \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{2m+1} \sum_{t_1,t_2=0}^{2m} E(\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_2}) e^{-i(t_1-t_2)\lambda_l} \\
= & E(\varepsilon_1^2) = 1.
\end{aligned}$$

(b) Since

$$\begin{aligned}
& E(\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_2} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_3} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_4}) \\
= & \begin{cases} E(\varepsilon_1^4), & \text{if } t_1 = t_2 = t_3 = t_4, \\ 1, & \text{if } \exists i_1, i_2, j_1, j_2 : t_{i_1} = t_{i_2} \neq t_{j_1} = t_{j_2}, \\ 0, & \text{else}, \end{cases} \qquad (4.8)
\end{aligned}$$

we get

$$\begin{aligned}
& \sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)|^4 \\
= & \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{(2m+1)^2} \Bigg| \sum_{t_1,t_2,t_3,t_4=0}^{2m} E(\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_2} \\
& \cdot \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_3} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_4}) e^{i(t_2-t_1+t_3-t_4)\lambda_l} \Bigg| \\
\leq & \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{(2m+1)^2} \left(K \left| \sum_{t_1,t_2=0}^{2m} 1 \right| + K \left| \sum_{t_1=0}^{2m} E(\varepsilon_1^4) \right| \right) < \infty.
\end{aligned}$$

(c) The same case differentiation as in (4.8) needs to be done, however, note that instead of just having to consider t_i being equal or not, we need to be concerned whether indices $t_i - j_i$ are equal, as we face

$$
\begin{aligned}
& \sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|MF^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)|^4 \\
= \; & \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{(2m+1)^2} \Bigg| \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \\
& a\left(\frac{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_1-j_1}{T}, j_1\right) a\left(\frac{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_2-j_2}{T}, j_2\right) \\
& a\left(\frac{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_3-j_3}{T}, j_3\right) a\left(\frac{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_4-j_4}{T}, j_4\right) \\
& E(\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_1-j_1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_2-j_2} \\
& \cdot \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_3-j_3} \varepsilon_{l+\zeta_{\lfloor uT \rfloor,l}m-m+t_4-j_4}) e^{i(t_2-t_1+t_3-t_4)\lambda_l} \Bigg| .
\end{aligned}
$$

Still, since

$$
\sup_{u\in[0,1]} |a(u, j_1)| \le \frac{K}{l(j_1)},
$$

by Definition 2.1, we get for the case of all indices being equal, that is the case of $t_1 - j_1 = t_2 - j_2 = t_3 - j_3 = t_4 - j_4$, an upper bound of the above expression of

$$
\begin{aligned}
& E(\varepsilon_1^4) \frac{K}{(2m+1)^2} \\
& \cdot \Bigg| \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \left|\frac{1}{l(j_1)}\right| \left|\frac{1}{l(j_2)}\right| \cdot \left|\frac{1}{l(j_3)}\right| \cdot \left|\frac{1}{l(j_4)}\right| \mathbb{1}_{\{t_1-j_1=t_2-j_2=t_3-j_3=t_4-j_4\}} \\
\le \; & \frac{K(2m+1)E(\varepsilon_1^4)}{(2m+1)^2} = O\left(\frac{1}{m}\right).
\end{aligned}
$$

The other possibility is any two indices being equal. There are three cases:

$$
\begin{aligned}
t_1 - j_1 = t_2 - j_2 &\neq t_3 - j_3 = t_4 - j_4, \\
t_1 - j_1 = t_3 - j_3 &\neq t_2 - j_2 = t_4 - j_4, \\
t_1 - j_1 = t_4 - j_4 &\neq t_2 - j_2 = t_3 - j_3.
\end{aligned}
$$

Exemplarily, we will consider $t_1 - j_1 = t_4 - j_4 \neq t_2 - j_2 = t_3 - j_3$. Hence, the upper bound in that case is

$$\begin{aligned} & \frac{K}{(2m+1)^2} \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \left|\frac{1}{l(j_1)}\right| \cdot \left|\frac{1}{l(j_2)}\right| \\ & \cdot \left|\frac{1}{l(j_3)}\right| \cdot \left|\frac{1}{l(j_4)}\right| \mathbb{1}_{\{t_1-j_1=t_4-j_4\neq t_2-j_2=t_3-j_3\}} \\ = \; & O(1) . \end{aligned}$$

Finally, we now get

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)|^4 = O(1) .$$

For $\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF_{\lfloor uT \rfloor}(\lambda_l)|^4$ we get the same result, as we merely need to substitute all functions $a\left(\frac{t}{T}, j\right)$ by $a_{t,T}(j)$. For the new coefficients the same bounds apply. See Definition 2.1. □

.

Proposition 4.2
Let $\tilde{X}_t(u)$ denote the stationary approximation of $X_{t,T}$ at time $\lfloor uT \rfloor$.
Let further $l = 1,\ldots,m$ and $\zeta_{\lfloor uT \rfloor,l} = div\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)\right\}}$ and $A(u,\lambda) := \sum_{j=-\infty}^{\infty} a(u,j)e^{-i\lambda j}$, $A_{t,T}(\lambda) := \sum_{j=-\infty}^{\infty} a_{t,T}(j)e^{-i\lambda j}$.
Then

(a)

$$\sup_{u\in[0,1]} \sum_{j\in\mathbb{Z}} \left|a_{\lfloor uT \rfloor,T}(j)\right| < \infty. \tag{4.9}$$

(b)

$$\sup_{x\in[0,1]} \sup_{l=1,\ldots,m} |A(x,\lambda_l)| < \infty. \tag{4.10}$$

(c) For $z \in \mathbb{R}_{>0}$

$$\sup_{m\in\mathbb{N}} \sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \left| A\left(\frac{l+\zeta_{\lfloor uT \rfloor,l}m - m}{T}, \lambda_l\right) - A(u,\lambda_l)\right|^z = O\left(\frac{m^z}{T^z}\right) . \tag{4.11}$$

(d) For $z \in \mathbb{R}_{>0}$

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \left| A\left(\frac{l+\zeta_{\lfloor uT \rfloor,l}m - m}{T}, \lambda_l\right) - A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)\right|^z = O\left(\frac{1}{T^z}\right) . \tag{4.12}$$

For $z \in \mathbb{R}_{>0}$

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \left| A\left(u, \lambda_l\right) - A_{l+\zeta_{\lfloor uT \rfloor, l}m-m,T}(\lambda_l) \right|^z = O\left(\frac{m^z}{T^z}\right). \tag{4.13}$$

(e) Let $j = 1, \ldots, 2m$.

$$\sup_{|u-u'|\leq \frac{Cm}{T}} \left| |A(u,\lambda_j)|^2 - |A(u',\lambda_j)|^2 \right| = O\left(\frac{m^2}{T^2} + \frac{m}{T}\right) = O\left(\frac{m}{T}\right). \tag{4.14}$$

(f) Let $j = 1, \ldots, 2m$.

$$\sup_{|u-u'|\leq \frac{Cm}{T}} \left| |A(u,\lambda_j)|^4 - |A(u',\lambda_j)|^4 \right| = O\left(\frac{m^4}{T^4} + \frac{m^3}{T^3} + \frac{m^2}{T^2} + \frac{m}{T}\right) = O\left(\frac{m}{T}\right). \tag{4.15}$$

Proof. (a) With Definition 2.1,

$$\begin{aligned}
\sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} \left| a_{\lfloor uT \rfloor,T}(j) \right| &\leq \sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} \left| a_{\lfloor uT \rfloor,T}(j) - a(u,j) \right| + \sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} |a(u,j)| \\
&\leq K \sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} \left| \frac{\lfloor uT \rfloor}{T} - u \right| \frac{1}{l(j)} + \sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} \frac{1}{l(j)} < \infty.
\end{aligned}$$

(b)

$$\begin{aligned}
\sup_{x\in[0,1]} \sup_{l=1,\ldots,m} \left| A\left(x, \lambda_l\right) \right| &\leq \sup_{x\in[0,1]} \sup_{l=1,\ldots,m} \sum_{j=-\infty}^{\infty} \left| a\left(x, j\right) \right| \\
&\leq \sum_{j=-\infty}^{\infty} \frac{K}{l(j)} < \infty,
\end{aligned}$$

cf. Definition 2.1(b).

(c) Since $\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\left|\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T}-u\right|\leq C\frac{m}{T}$, it holds

$$\begin{aligned}
&\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\left|A\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},\lambda_l\right)-A(u,\lambda_l)\right| \\
\leq\ &\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\sum_{j=-\infty}^{\infty}\left|a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},j\right)-a(u,j)\right| \\
\leq\ &\sum_{j=-\infty}^{\infty}\frac{Km}{Tl(j)}=O\left(\frac{m}{T}\right),
\end{aligned}$$

cf. Definition 2.1(b) and (c). Consequently, (4.11).

(d) Since

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\left|\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T}-u\right|\leq K\frac{m}{T},$$

it holds that

$$\begin{aligned}
&\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\left|A\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},\lambda_l\right)-A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\right| \\
\leq\ &\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\sum_{j=-\infty}^{\infty}\left|a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},j\right)-a_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(j)\right| \\
\leq\ &\sum_{j=-\infty}^{\infty}\frac{K}{Tl(j)}=O\left(\frac{1}{T}\right),
\end{aligned}$$

cf. Definition 2.1(b) and (c). Consequently, (4.12).

(e) Results from parts (c) and (d).

(f) Let $|u-u'|\leq\frac{Cm}{T}$

$$\begin{aligned}
|A(u,\lambda_j)|^2 &= \left|\sum_{k=-\infty}^{\infty}a(u,k)e^{-i\lambda_j k}\right|^2 \\
&= \sum_{k_1,k_2=-\infty}^{\infty}a(u,k_1)a(u,k_2)e^{-i\lambda_j(k_1-k_2)}. \\
&= \sum_{k_1,k_2=-\infty}^{\infty}[a(u,k_1)-a(u',k_1)+a(u',k_1)][a(u,k_2)-a(u',k_2)+a(u',k_2)] \\
&\quad\cdot e^{-i\lambda_j(k_1-k_2)}.
\end{aligned}$$

Maintaining the difference $a(u,k_i)-a(u',k_i)$, extracting yields 4 summands with the last one being $a(u', k_1)a(u', k_2)$. Substracting $|A(u', \lambda_j)|^2$ therefore merely gets rid of this one summand. The remaining terms can be bounded by either $\frac{m}{Tl(k_i)}$, if we have a difference $a(u, k_i) - a(u', k_i)$, or $\frac{K}{l(k_i)}$ in the case of $a(u, k_i)$ (cf. Definition 2.1). Hence,

$$\begin{aligned}
&\sup_{|u-u'|\leq \frac{Cm}{T}} \left||A(u, \lambda_j)|^2 - |A(u', \lambda_j)|^2\right| \\
\leq\ & K \sum_{k_1, k_2=-\infty}^{\infty} \left(\frac{m^2}{T^2 l(k_1) l(k_2)} + \frac{2m}{T l(k_1) l(k_2)} \right) \\
=\ & O\left(\frac{m^2}{T^2} + \frac{m}{T} \right).
\end{aligned}$$

(g)

$$\begin{aligned}
& |A(u, \lambda_j)|^4 \\
=\ & \left| \sum_{k=-\infty}^{\infty} a(u, k) e^{-i\lambda_j k} \right|^4 \\
=\ & \sum_{k_1, k_2, k_3, k_4=-\infty}^{\infty} a(u, k_1) a(u, k_2) a(u, k_3) a(u, k_4) e^{-i\lambda_j (k_1 - k_2 + k_3 - k_4)}. \\
=\ & \sum_{k_1, k_2, k_3, k_4=-\infty}^{\infty} [a(u, k_1) - a(u', k_1) + a(u', k_1)][a(u, k_2) - a(u', k_2) + a(u', k_2)] \\
& \cdot [a(u, k_3) - a(u', k_3) + a(u', k_3)][a(u, k_4) - a(u', k_4) + a(u', k_4)] \\
& \cdot e^{-i\lambda_j (k_1 - k_2 + k_3 - k_4)}.
\end{aligned}$$

Maintaining the difference $a(u, k_i) - a(u', k_i)$, extracting yields 16 summands with the last one being $a(u', k_1)a(u', k_2)a(u', k_3)a(u', k_4)$. Substracting $|A(u', \lambda_j)|^4$ therefore merely gets rid of this one summand. The remaining terms can be bounded by either $\frac{m}{Tl(k_i)}$, if we have a difference $a(u, k_i) - a(u', k_i)$, or $\frac{K}{l(k_i)}$ in the case of $a(u, k_i)$ (cf. Definition 2.1). Hence,

$$\begin{aligned}
&\sup_{|u-u'|\leq \frac{Cm}{T}} \left||A(u, \lambda_j)|^4 - |A(u', \lambda_j)|^4\right| \\
\leq\ & K \sum_{k_1, k_2, k_3, k_4=-\infty}^{\infty} \left(\frac{m^4}{T^4 l(k_1) l(k_2) l(k_3) l(k_4)} + \frac{4m^3}{T^3 l(k_1) l(k_2) l(k_3) l(k_4)} \right. \\
& \left. + \frac{6m^2}{T^2 l(k_1) l(k_2) l(k_3) l(k_4)} + \frac{4m}{T l(k_1) l(k_2) l(k_3) l(k_4)} \right) \\
=\ & O\left(\frac{m^4}{T^4} + \frac{m^3}{T^3} + \frac{m^2}{T^2} + \frac{m}{T} \right).
\end{aligned}$$

□

Proposition 4.3
For a sequence $\varepsilon_1, \varepsilon_2, \dots$ of independent identically distributed centred random variables with variance $0 < \sigma^2 < \infty$ and existing fourth moment, the following inequality holds

$$E\left(\sum_{j=1}^{n} \varepsilon_j\right)^4 \leq nE(\varepsilon_1^4) + 3n^2\sigma^4.$$

See also Exercise 10.14 in Brockwell and Davis [3].

Proof.

$$\begin{aligned} E\left(\sum_{j=1}^{n} \varepsilon_j\right)^4 &= \operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j\right)^2 + \left(E\left(\sum_{j=1}^{n} \varepsilon_j\right)^2\right)^2 \\ &=: A_1 + A_2. \end{aligned}$$

$\boxed{A_2}$

$$A_2 = \left(E\left(\sum_{j=1}^{n} \varepsilon_j\right)^2\right)^2 = \left(\operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j\right) + \left(E\left(\sum_{j=1}^{n} \varepsilon_j\right)\right)^2\right)^2.$$

Employing that the random variables are independent identically distributed and centred, we get

$$A_2 = \left(\operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j\right)\right)^2 = \left(\sum_{j=1}^{n} \operatorname{Var}(\varepsilon_j)\right)^2 = n^2\sigma^4.$$

$\boxed{A_1}$

$$\begin{aligned} A_1 &= \operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j\right)^2 = \operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j^2 + \sum_{i \neq j=1}^{n} \varepsilon_i \varepsilon_j\right) \\ &= \operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j^2\right) + \operatorname{Var}\left(\sum_{i \neq j=1}^{n} \varepsilon_i \varepsilon_j\right) + 2\operatorname{Cov}\left(\sum_{j=1}^{n} \varepsilon_j^2, \sum_{k \neq l=1}^{n} \varepsilon_k \varepsilon_l\right). \end{aligned} \tag{4.16}$$

$$\begin{aligned} \operatorname{Var}\left(\sum_{j=1}^{n} \varepsilon_j^2\right) &= \sum_{j=1}^{n} \operatorname{Var}(\varepsilon_j^2) = n \operatorname{Var}(\varepsilon_1^2) = n\left(E(\varepsilon_1^4) - (E(\varepsilon_1^2))^2\right) \\ &= nE(\varepsilon_1^4) - n\sigma^4. \end{aligned} \tag{4.17}$$

$$\begin{aligned}
\operatorname{Var}\left(\sum_{i\neq j=1}^{n} \varepsilon_i\varepsilon_j\right) &= E\left(\sum_{i\neq j=1}^{n} \varepsilon_i\varepsilon_j - E\left(\sum_{i\neq j=1}^{n} \varepsilon_i\varepsilon_j\right)\right)^2 \\
&= E\left(\sum_{i\neq j=1}^{n} \varepsilon_i\varepsilon_j\right)^2 = E\left(\sum_{i\neq j=1}^{n}\sum_{k\neq l=1}^{n} \varepsilon_i\varepsilon_j\varepsilon_k\varepsilon_l\right) \\
&= \sum_{i\neq j=1}^{n}\sum_{k\neq l=1}^{n} E(\varepsilon_i\varepsilon_j\varepsilon_k\varepsilon_l)\cdot \mathbb{1}_{\{(i=k\wedge j=l)\vee(i=l\wedge j=k)\}}
\end{aligned} \tag{4.18}$$

$$= 2\sum_{i\neq j=1}^{n} E(\varepsilon_i^2\varepsilon_j^2) = 2n(n-1)\sigma^4. \tag{4.19}$$

$$\begin{aligned}
\operatorname{Cov}\left(\sum_{j=1}^{n}\varepsilon_j^2, \sum_{k\neq l=1}^{n}\varepsilon_k\varepsilon_l\right) &= \sum_{j=1}^{n}\sum_{k\neq l=1}^{n} \operatorname{Cov}(\varepsilon_j^2, \varepsilon_k\varepsilon_l) \\
&= \sum_{j=1}^{n}\sum_{k\neq l=1}^{n} (E(\varepsilon_j^2\varepsilon_k\varepsilon_l) - E(\varepsilon_j)^2E(\varepsilon_k)E(\varepsilon_l)) \\
&= \sum_{j=1}^{n}\sum_{k\neq l=1}^{n} E(\varepsilon_j^2\varepsilon_k\varepsilon_l) = \sum_{j=1}^{n}\sum_{k\neq l=1}^{n} (E(\varepsilon_j^2\varepsilon_k\varepsilon_l)\mathbb{1}_{\{k=j\vee l=j\}} \\
&= \sum_{j=1}^{n}\sum_{l=1,\, l\neq j}^{n} E(\varepsilon_j^3)E(\varepsilon_l) + \sum_{j=1}^{n}\sum_{k=1,\, k\neq j}^{n} E(\varepsilon_j^3)E(\varepsilon_k) = 0.
\end{aligned} \tag{4.20}$$

Now, with (4.17), (4.19) and (4.20) Equation (4.16) simplifies to

$$A_1 = nE(\varepsilon_1^4) - n\sigma^4 + 2n(n-1)\sigma^4 + 0 = nE(\varepsilon_1^4) + 2n^2\sigma^4 - 2n\sigma^4.$$

With the knowledge about A_1 and A_2 one finally obtains

$$E\left(\sum_{j=1}^{n}\varepsilon_j\right)^4 = nE(\varepsilon_1^4) + 3n^2\sigma^4 - 2n\sigma^4 \leq nE(\varepsilon_1^4) + 3n^2\sigma^4.$$

. □

The following Theorem is kept in the notation of Brockwell and Davis [3], Theorem 10.3 and is an additional result to their theorem.

Theorem 4.1. *Let* $\{Z_t\} \sim IID(0, \sigma^2)$ *and* $\sum_{j\in\mathbb{Z}} |\psi_j|\sqrt{|j|} < \infty$. *Let further, for* $\lambda_j = \frac{2\pi j}{n}$, $j = 1, \ldots, n$

$$\begin{aligned} Y(\lambda_j) &= \frac{1}{\sqrt{n}} \sum_{s=-\infty}^{\infty} \psi_s e^{-i\lambda_j s} U_{n,s}, \\ \textit{with } U_{n,s} &= \sum_{t=1-s}^{n-s} Z_t e^{-i\lambda_j t} - \sum_{t=1}^{n} Z_t e^{-i\lambda_j t}. \end{aligned}$$

Then

$$E\left(\sum_{j=1}^{n} |Y(\lambda_j)|^2\right)^2 = O(1),$$

for $n \to \infty$.

Proof. $Y(\lambda_j) = \frac{1}{\sqrt{n}} \sum_{s=-\infty}^{\infty} \psi_s e^{-i\lambda_j s} U_{n,s}$, with $U_{n,s} = \sum_{t=1-s}^{n-s} Z_t e^{-i\lambda_j t} - \sum_{t=1}^{n} Z_t e^{-i\lambda_j t}$. $U_{n,s}$ is a sum of $2|s|$ independent random variables, if $|s| < n$. If $|s| \geq n$ it is a sum of $2n$ independent random variables.
We now intend to prove that

$$E\left(\sum_{j=1}^{n} |Y(\lambda_j)|^2\right)^2 = O(1),$$

for $n \to \infty$.
Define

$$T(s) := \begin{cases} \{1-s, \ldots, 0\} \cup \{n-s+1, \ldots, n\}, & 0 < s < n, \\ \{1, \ldots, -s\} \cup \{n+1, \ldots, n-s\}, & -n < s < 0, \\ \{1-s, \ldots, n-s\} \cup \{1, \ldots, n\}, & |s| \geq n. \end{cases}$$

This set indicates, which Z_t contribute to $U_{n,s}$. Note that

$$\#T(s) = \min\{2|s|, 2n\}.$$

In the following, we will also be concerned with the cardinality of the intersections of the sets $T(s_i) \cap T(s_j)$, $i, j = 1, 2, 3, 4$, which is

$$\#T(s_i) \cap T(s_j) \leq \min\{\min\{2|s_i|, 2n\}, \min\{2|s_j|, 2n\}\} = O(\min\{n, |s_i|, |s_j|\}).$$

Further note that

$$\min\{n, |s_i|, |s_j|\} \leq \sqrt{|s_i|}\sqrt{|s_j|}.$$

We then set $\smallint_{n,s,t} := \pm 1$, depending on which Z_t's contribute to $U_{n,s}$.

$$
\begin{aligned}
&\left(\sum_{j=1}^{n} |Y(\lambda_j)|^2\right)^2 \\
= \; & \frac{1}{n^2} \sum_{j_1=1}^{n} \sum_{s_1,s_2=-\infty}^{\infty} \psi_{s_1}\overline{\psi_{s_2}} e^{-i\lambda_{j_1}(s_1-s_2)} \sum_{t_1 \in T(s_1)} \smallint_{n,s_1,t_1} Z_{t_1} e^{-i\lambda_{j_1} t_1} \sum_{t_2 \in T(s_2)} \smallint_{n,s_2,t_2} Z_{t_2} e^{+i\lambda_{j_1} t_2} \\
& \cdot \sum_{j_2=1}^{n} \sum_{s_3,s_4=-\infty}^{\infty} \overline{\psi_{s_3}}\psi_{s_4} e^{i\lambda_{j_2}(s_3-s_4)} \sum_{t_3 \in T(s_3)} \smallint_{n,s_3,t_3} Z_{t_3} e^{+i\lambda_{j_2} t_3} \sum_{t_4 \in T(s_4)} \smallint_{n,s_4,t_4} Z_{t_4} e^{-i\lambda_{j_2} t_4} \\
= \; & \sum_{k_1,k_2 \in \mathbb{Z}} \sum_{s_1,s_2,s_3,s_4=-\infty}^{\infty} \psi_{s_1}\overline{\psi_{s_2}\psi_{s_3}}\psi_{s_4} \sum_{t_1 \in T(s_1)} \smallint_{n,s_1,t_1} Z_{t_1} \sum_{t_2 \in T(s_2)} \smallint_{n,s_2,t_2} Z_{t_2} \\
& \cdot \sum_{t_3 \in T(s_3)} \smallint_{n,s_3,t_3} Z_{t_3} \sum_{t_4 \in T(s_4)} \smallint_{n,s_4,t_4} Z_{t_4} \mathbb{1}_{\{t_1-t_2+(s_1-s_2)=k_1 n\}} \mathbb{1}_{\{t_3-t_4+(s_3-s_4)=k_2 n\}}.
\end{aligned}
$$

As we are interested in the value of the expectation of this expression, we only need to look at the cases when $E(Z_{t_1}Z_{t_2}Z_{t_3}Z_{t_4}) \neq 0$. As the fourth moment of the random variables exists, we may bound it, as well as lower moments by some arbitrary constant $C \geq 0$. Formally, we get

$$
\begin{aligned}
&\left| E\left(\sum_{j=1}^{n} |Y(\lambda_j)|^2\right)^2 \right| \\
\leq \; & C \sum_{s_1,s_2,s_3,s_4=-\infty}^{\infty} \psi_{s_1}\overline{\psi_{s_2}\psi_{s_3}}\psi_{s_4} \sum_{k_1,k_2 \in \mathbb{Z}} \sum_{t_i \in T(s_i),\, i=1,2,3,4} \Big[\mathbb{1}_{\{t_1=t_2\}}\mathbb{1}_{\{t_3=t_4\}} + \mathbb{1}_{\{t_1=t_3\}}\mathbb{1}_{\{t_2=t_4\}} \\
& + \mathbb{1}_{\{t_1=t_4\}}\mathbb{1}_{\{t_2=t_3\}}\Big] \cdot \mathbb{1}_{\{t_1-t_2+(s_1-s_2)=k_1 n\}} \mathbb{1}_{\{t_3-t_4+(s_3-s_4)=k_2 n\}}.
\end{aligned}
$$

We hence encounter the following situations

- Case 1:

$$
\begin{aligned}
(I) \quad & t_1 = t_2 \\
(II) \quad & t_3 = t_4
\end{aligned}
$$

as well as

$$
\begin{aligned}
(III) \quad & t_1 - t_2 + (s_1 - s_2) = k_1 n \\
(IV) \quad & t_3 - t_4 + (s_3 - s_4) = k_2 n
\end{aligned}
$$

With (I) and (II) we get

$$
\begin{aligned}
(III') \quad & k_1 = \frac{s_1 - s_2}{n}, \\
(IV') \quad & k_2 = \frac{s_3 - s_4}{n}.
\end{aligned}
$$

k_1 and k_2 are therefore uniquely determined by s_2, s_2, s_3, s_4 and n and can therefore be eliminated. Note that k_1 and k_2 need to be integers. To be exact, we would then have to write

$$k_1 = \frac{s_1 - s_2}{n} \cap \mathbb{Z}, \; k_2 = \frac{s_3 - s_4}{n} \cap \mathbb{Z}.$$

Sufficient for an upper bound, however, is to use the whole range of the s_i, $i = 1, 2, 3, 4$.

$$\begin{aligned}
& \sum_{s_1,s_2,s_3,s_4} \psi_{s_1}\overline{\psi_{s_2}\psi_{s_3}}\psi_{s_4} \sum_{t_1 \in T(s_1)\cap T(s_2)} 1 \sum_{t_3 \in T(s_3)\cap T(s_4)} 1 \\
\leq \;& K \sum_{s_1,s_2,s_3,s_4} \psi_{s_1}\overline{\psi_{s_2}\psi_{s_3}}\psi_{s_4} \min\{n, |s_1|, |s_2|\} \min\{n, |s_3|, |s_4|\} \\
\leq \;& K \sum_{s_1,s_2} \psi_{s_1}\overline{\psi_{s_2}}\sqrt{|s_1|}\sqrt{|s_2|} \sum_{s_3,s_4} \overline{\psi_{s_3}}\psi_{s_4}\sqrt{|s_3|}\sqrt{|s_4|} \\
= \;& O(1).
\end{aligned}$$

- Case 2:

$$\begin{aligned}
(I) \quad & t_1 = t_3 \\
(II) \quad & t_2 = t_4
\end{aligned}$$

as well as

$$\begin{aligned}
(III) \quad & t_1 - t_2 + (s_1 - s_2) = k_1 n \\
(IV) \quad & t_1 - t_2 + (s_3 - s_4) = k_2 n
\end{aligned}$$

With (I) and (II) we get

$$\begin{aligned}
(III') \quad k_1 &= \frac{t_1 - t_2 + s_1 - s_2}{n}, \\
(IV') \quad k_2 &= \frac{t_1 - t_2 + s_3 - s_4}{n}.
\end{aligned}$$

k_1 and k_2 are therefore uniquely determined by $s_2, s_2, s_3, s_4, t_1, t_2$ and n and can therefore be eliminated. Note that k_1 and k_2 need to be integers. To be exact, we would then have to write

$$k_1 = \frac{t_1 - t_2 + s_1 - s_2}{n} \cap \mathbb{Z}, \; k_2 = \frac{t_1 - t_2 + s_3 - s_4}{n} \cap \mathbb{Z}.$$

Now,

$$\begin{aligned}
&\sum_{s_1,s_2,s_3,s_4} \psi_{s_1}\overline{\psi_{s_2}\psi_{s_3}}\psi_{s_4} \sum_{t_1\in T(s_1)\cap T(s_3)} 1 \sum_{t_3\in T(s_2)\cap T(s_4)} 1 \\
\leq\ & K \sum_{s_1,s_2,s_3,s_4} \psi_{s_1}\overline{\psi_{s_2}\psi_{s_3}}\psi_{s_4} \min\{n,|s_1|,|s_3|\} \min\{n,|s_2|,|s_4|\} \\
\leq\ & K \sum_{s_1,s_2} \psi_{s_1}\overline{\psi_{s_2}}\sqrt{|s_1|}\sqrt{|s_2|} \sum_{s_3,s_4} \overline{\psi_{s_3}}\psi_{s_4} \sum \sqrt{|s_3|}\sqrt{|s_4|} \\
=\ & O(1).
\end{aligned}$$

- Case 3: Analogously to Case 2.

. □

Remark 4.1
The condition $\sum_{j\in\mathbb{Z}} |\psi_j|\sqrt{|j|} < \infty$ is not very strong. Dahlhaus and Giraitis [12] use in Corollary 4.1, which is the asymptotic normality of the rescaled spectral mean of the Fourier coefficients of a stationary time series the assumption that $\sum_{j\in\mathbb{Z}} |\psi_j| j^2 < \infty$. As we need the result of the above Theorem to consider the rescaled spectral mean later on, we are on the safe side starting off with a condition not as strong as their final condition.
Moreover, Grenander and Rosenblatt [21] also use this assumption in Theorem 6, when they intend to generalize their results from iid white noise to stationary time series.

4.2 Linking the locally stationary case to the i.i.d. case

Lemma 4.2 (Relationship between $MF^{\tilde{X}}_{\lfloor uT\rfloor}$ and $MF^{\varepsilon}_{\lfloor uT\rfloor}$).
Let $\tilde{X}_t(u)$ denote the stationary approximation of $X_{t,T}$ at time $\lfloor uT\rfloor$.
Let further $A(u,\lambda) := \sum_{j=-\infty}^{\infty} a(u,j)e^{-i\lambda j}$. Then for $l = 1,\dots,m$

$$\begin{aligned}
&MF^{\tilde{X}}_{\lfloor uT\rfloor}(\lambda_l) = A\left(u,\lambda_l\right) MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l) \quad + \quad R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l), \\
&\textit{with} \sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \quad\to\quad 0, \ \textit{as } m\to\infty,
\end{aligned} \tag{4.21}$$

and

$$ER^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l) = 0. \tag{4.22}$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty$,

$$\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 = O\left(\frac{1}{m^2}\right), \tag{4.23}$$

as well as

$$E\left(\frac{1}{\sqrt{2m+1}}\sum_{l=1}^{2m+1}|R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^2\right)^2 \to 0 \text{ as } m\to\infty. \tag{4.24}$$

Proof. With $\zeta_{\lfloor uT\rfloor,l} = div\left(\lfloor uT\rfloor - \lfloor\frac{m}{2}\rfloor\right) - \mathbb{1}_{\left\{l\geq mod\left(\lfloor uT\rfloor - \lfloor\frac{m}{2}\rfloor\right)\right\}}$ it follows that

$$\begin{aligned} MF^{\tilde X}_{\lfloor uT\rfloor}(\lambda_l) &= \frac{1}{\sqrt{2m+1}}\sum_{t=0}^{2m}\tilde X_t\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t}{T},j\right)e^{-it\lambda_l} \\ &= \frac{1}{\sqrt{2m+1}}\sum_{t=0}^{2m}\sum_{j=-\infty}^{\infty}a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t}{T},j\right)\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t-j}e^{-it\lambda_l}. \end{aligned}$$

In order to proceed as in Brockwell and Davis [3], we need to free the coefficients $a(\cdot,j)$ of their dependence on time. We do this by splitting:

$$\begin{aligned} & MF^{\tilde X}_{\lfloor uT\rfloor}(\lambda_l) \\ = & \frac{1}{\sqrt{2m+1}}\sum_{t=0}^{2m}\sum_{j=-\infty}^{\infty}\left[a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t}{T},j\right)\right. \\ & \left.-a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},j\right)\right]\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t-j}e^{-it\lambda_l} \\ + & \frac{1}{\sqrt{2m+1}}\sum_{t=0}^{2m}\sum_{j=-\infty}^{\infty}a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},j\right)\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t-j}e^{-it\lambda_l} \\ =: & Y^{\tilde X,(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + \frac{1}{\sqrt{2m+1}}\sum_{j=-\infty}^{\infty}a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},j\right)e^{-ij\lambda_l}\sum_{t=-j}^{2m-j}\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t}e^{-it\lambda_l} \\ = & Y^{\tilde X,(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + A\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},\lambda_l\right)\frac{1}{\sqrt{2m+1}}\sum_{t=0}^{2m}\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t}e^{-it\lambda_l} \\ + & \frac{1}{\sqrt{2m+1}}\sum_{j=-\infty}^{\infty}a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},j\right)e^{-ij\lambda_l}U_{\lfloor uT\rfloor,m,j}(\lambda_l) \\ =: & Y^{\tilde X,(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + A\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T},\lambda_l\right)MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l) + Y^{\tilde X,(2)}_{\lfloor uT\rfloor,m}(\lambda_l), \end{aligned}$$

where

$$U_{\lfloor uT\rfloor,m,j}(\lambda_l) \;=\; \left(\sum_{t=-j}^{2m-j} \varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t} e^{-it\lambda_l} - \sum_{t=0}^{2m} \varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t} e^{-it\lambda_l}\right).$$

It is evident, that the expectancy of both terms $Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l)$ and $Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l)$ equals zero, implying (4.22) for $R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l) := Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l)$.
It now remains to show that

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \to 0 \text{ as } m\to\infty, \tag{4.25}$$

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \to 0 \text{ as } m\to\infty, \tag{4.26}$$

with adequate rates in the case of the stronger assumption.
To show the L_4-convergence (4.25) first note that

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 = O\left(\frac{m^6}{T^4(2m+1)^2}\right) = O\left(\frac{m^4}{T^4}\right),$$

with the proof basically analogous to the proof of (4.6), however it is slightly more demanding:

$$\begin{aligned}
&\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \\
=\;& \sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \frac{1}{(2m+1)^2} \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \\
&\left(a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_1}{T}, j_1\right) - a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T}, j_1\right)\right) \\
&\cdot\left(a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_2}{T}, j_2\right) - a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T}, j_2\right)\right) \\
&\cdot\left(a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_3}{T}, j_3\right) - a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T}, j_3\right)\right) \\
&\cdot\left(a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_4}{T}, j_4\right) - a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m-m}{T}, j_4\right)\right) \\
&\cdot E(\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_1-j_1}\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_2-j_2}\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_3-j_3}\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_4-j_4}) e^{i(t_2-t_1+t_3-t_4)\lambda_l}.
\end{aligned}$$

Just like in the proof of Proposition 4.1, one needs to distinguish between the different values the expected value can take and split the above expression accordingly. Since for $v\in[0,1]$

$$\sup_{u\in[0,1]} |a(u+v,j_1) - a(u,j_1)| \le \frac{K|v|}{l(j_1)},$$

we get for the case of all indices of the innovations being equal an upper bound of

$$\begin{aligned} & E(\varepsilon_1^4)\frac{K}{(2m+1)^2} \\ & \cdot \left| \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \left|\frac{t_1}{Tl(j_1)}\right| \left|\frac{t_2}{Tl(j_2)}\right| \cdot \left|\frac{t_3}{Tl(j_3)}\right| \cdot \left|\frac{t_4}{Tl(j_4)}\right| \mathbb{1}_{\{t_1-j_1=t_2-j_2=t_3-j_3=t_4-j_4\}} \right. \\ \leq \;& \frac{K(2m+1)^5\, E(\varepsilon_1^4)}{(2m+1)^2 T^4} = O\left(\frac{m^3}{T^4}\right). \end{aligned}$$

The other possibility is any two indices being equal. There are three cases:

$$\begin{aligned} t_1 - j_1 = t_2 - j_2 &\neq t_3 - j_3 = t_4 - j_4, \\ t_1 - j_1 = t_3 - j_3 &\neq t_2 - j_2 = t_4 - j_4, \\ t_1 - j_1 = t_4 - j_4 &\neq t_2 - j_2 = t_3 - j_3. \end{aligned}$$

Exemplarily, we will consider $t_1 - j_1 = t_4 - j_4 \neq t_2 - j_2 = t_3 - j_3$. For this case, an upper bound is

$$\begin{aligned} & \frac{K}{(2m+1)^2} \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \left|\frac{t_1}{Tl(j_1)}\right| \cdot \left|\frac{t_2}{Tl(j_2)}\right| \\ & \cdot \left|\frac{t_3}{Tl(j_3)}\right| \cdot \left|\frac{t_4}{Tl(j_4)}\right| \mathbb{1}_{\{t_1-j_1=t_4-j_4\neq t_2-j_2=t_3-j_3\}} \\ \leq \;& \frac{K(2m+1)^4}{(2m+1)^2T^4} \sum_{t_1,t_2=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \frac{1}{l(j_1)l(j_2)l(j_3)l(j_4)} \\ \leq \;& \frac{K(2m+1)^6}{(2m+1)^2T^4} = O\left(\frac{m^4}{T^4}\right). \end{aligned}$$

Finally, collecting the results on the upper bounds in the different cases, we get

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|Y_{\lfloor uT\rfloor,m}^{\tilde{X},(1)}(\lambda_l)|^4 = O\left(\frac{m^4}{T^4}\right). \tag{4.27}$$

Now for the second error term $Y_{\lfloor uT\rfloor,m}^{\tilde{X},(2)}(\lambda_l)$, the Minkowski inequality yields

$$\begin{aligned} & \sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|Y_{\lfloor uT\rfloor,m}^{\tilde{X},(2)}(\lambda_l)|^4 \\ \leq \;& \sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \frac{1}{(2m+1)^2} \left(\sum_{j=-\infty}^{\infty} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| (E|U_{\lfloor uT\rfloor,m,j}(\lambda_l)|^4)^{\frac{1}{4}} \right)^4 . \end{aligned} \tag{4.28}$$

Note that for $n \in \mathbb{N}$, $E\left(\sum_{j=1}^n e_j\right)^4 \leq nE(e_1^4) + 3n^2$ (cf. Proposition 4.3).

As stated before, for $|j| < 2m+1$, $U_{\lfloor uT\rfloor,m,j}(\lambda_l)$ is the sum of $2|j|$ independent and centred random variables. For $|j| \geq 2m+1$, $U_{\lfloor uT\rfloor,m,j}(\lambda_l)$ is the sum of $2(2m+1) = 4m+2$ independent and centred random variables.

$$\begin{aligned}\sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|U_{\lfloor uT\rfloor,m,j}(\lambda_l)|^4 &\leq 2\min(|j|,2m+1)E(e_1^4) + 12(\min(|j|,2m+1))^2 \\ &= O\left(\min(|j|^2,(2m+1)^2\right).\end{aligned}$$

Let $\mu_0 \in \mathbb{N}$ arbitrary but fixed. It follows that

$$\begin{aligned}&\frac{1}{\sqrt{2m+1}} \sum_{j=-\infty}^{\infty} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| \sqrt[4]{(\min(|j|^2,(2m+1)^2)} \\ \leq\ &\frac{1}{\sqrt{2m+1}} \sum_{|j|\leq\mu_0} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| \cdot \sqrt{|j|} \\ &+ \sum_{|j|>\mu_0} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right|.\end{aligned}$$

Concerning the function a: Note that m is only present in the first component of the function a. We may therefore exploit the assumption made by Definition 2.1(c):

$$\sup_{m,u,l} \left| \sum_{j=-\infty}^{\infty} a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| \leq \sup_{u\in[0,1]} \left| \sum_{j=-\infty}^{\infty} a(u,j)\right| < \infty. \tag{4.29}$$

We note that now

$$\begin{aligned}&\sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{\sqrt{2m+1}} \sum_{j=-\infty}^{\infty} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| \sqrt[4]{(\min(|j|^2,(2m+1)^2)} \\ \leq\ &\frac{K\sqrt{\mu_0}}{\sqrt{2m+1}} + \sup_{x\in[0,1]} \sum_{|j|>\mu_0} |a(x,j)|, \end{aligned}\tag{4.30}$$

with

$$\begin{aligned}&\limsup_{m\to\infty} \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{\sqrt{2m+1}} \sum_{j=-\infty}^{\infty} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| \sqrt[4]{(\min(|j|^2,(2m+1)^2)} \\ \leq\ &\sup_{x\in[0,1]} \sum_{|j|>\mu_0} |a(x,j)|\end{aligned}$$

As μ_0 is arbitrary and due to the uniform absolute summability of $a(u,j)$ (Definition 2.1(c)), the upper bound of (4.28) converges to zero for $T\to\infty$ (and thus $m\to\infty$).

If additionally

$$\sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} |a(u,j)|\sqrt{|j|} < \infty,$$

we get with the same argument as in (4.29) that

$$\sup_{m,u,l} \sum_{j=-\infty}^{\infty} \left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| \sqrt{|j|} < \infty.$$

This then results in a rate of convergence, which we did not get in (4.30):

$$\begin{aligned}
&\sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \\
\leq\ & C\cdot \sup_{u\in[0,1]} \sup_{l=1,\dots,m} \frac{1}{(2m+1)^2}\left(\sum_{j=-\infty}^{\infty}\left| a\left(\frac{l+\zeta_{\lfloor uT\rfloor,l}m - m}{T}, j\right)\right| |j|^{\frac{1}{2}}\right)^4 \\
=\ & O\left(\frac{1}{m^2}\right).
\end{aligned}$$

All in all, we get for $R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l) := Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l)$

$$\begin{aligned}
\sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \ &\leq\ 2^4\left(\sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4\right. \\
&\qquad \left. + \sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4\right) \\
&=\ O\left(\frac{1}{m^2}\right),
\end{aligned}$$

as the assumptions given in Section 3.1 imply $\frac{m^6}{T^4} = o(1)$.

Now, to finally prove (4.24), we note again that

$$R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l) := Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + Y^{\tilde{X},(2)}_{\lfloor uT\rfloor,m}(\lambda_l).$$

By (4.27)

$$\sup_{u\in[0,1]} \sup_{l=1,\dots,m} E|Y^{\tilde{X},(1)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 = O\left(\frac{m^4}{T^4}\right).$$

Hence,

$$E\left(\frac{1}{\sqrt{2m+1}}\sum_{j=1}^{2m+1}|Y_{\lfloor uT\rfloor,m}^{\tilde{X},(1)}(\lambda_l)|^2\right)^2 = O\left(\frac{m^5}{T^4}\right) = o(1),$$

by the assumptions given in Section 3.1.
From Theorem 4.1 we get

$$E\left(\sum_{j=1}^{2m+1}|Y_{\lfloor uT\rfloor,m}^{\tilde{X},(2)}(\lambda_l)|^2\right)^2 = O(1).$$

(4.24) now follows by Cauchy-Schwarz. □

Theorem 4.2 (Relationship between $MI_{\lfloor uT\rfloor,m}^{\tilde{X}}$ and $MI_{\lfloor uT\rfloor,m}^{\varepsilon}$).
In the situation of Lemma 4.2

$$MI_{\lfloor uT\rfloor,m}^{\tilde{X}}(\lambda_l) = |A(u,\lambda_l)|^2\, MI_{\lfloor uT\rfloor,m}^{\varepsilon}(\lambda_l) + R_{\lfloor uT\rfloor,m}(\lambda_l), \tag{4.31}$$

$$\textit{with} \sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R_{\lfloor uT\rfloor,m}(\lambda_l)|^2 \to 0 \textit{ for } m\to\infty. \tag{4.32}$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty$

$$\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R_{\lfloor uT\rfloor,m}(\lambda_l)|^2 = O\left(\frac{1}{m}\right). \tag{4.33}$$

Proof. Now, extending the result of Lemma 4.2 to a relationship between the moving periodograms, one merely needs to consider the remainder $R_{\lfloor uT\rfloor,m}(\lambda_l)$, which is of the form

$$\begin{aligned} R_{\lfloor uT\rfloor,m}(\lambda) &= A(u,\lambda_l)\, MF_{\lfloor uT\rfloor}^{\varepsilon}(\lambda_j)\overline{R_{\lfloor uT\rfloor,m}^{(1)}(\lambda_l)} \\ &+ \overline{A(u,\lambda_l) MF_{\lfloor uT\rfloor}^{\varepsilon}(\lambda_j)}R_{\lfloor uT\rfloor,m}^{(1)}(\lambda_l) + |R_{\lfloor uT\rfloor,m}^{(1)}(\lambda_l)|^2. \end{aligned}$$

Therefore,

$$\begin{aligned} &\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R_{\lfloor uT\rfloor,m}(\lambda_l)|^2 \\ \le\ & 2^2\Bigg(\sup_{u\in[0,1]}\sup_{l=1,\ldots,m}|A(u,\lambda_l)|^2\cdot\left|MF_{\lfloor uT\rfloor}^{\varepsilon}(\lambda_j)\right|^2\cdot\left|R_{\lfloor uT\rfloor,m}^{(1)}(\lambda_l)\right|^2 \\ &+ \sup_{u\in[0,1]}\sup_{l=1,\ldots,m}\left|R_{\lfloor uT\rfloor,m}^{(1)}(\lambda_l)\right|^4\Bigg) \end{aligned} \tag{4.34}$$

According to Lemma 4.2, $\sup_{u\in[0,1]}\sup_{l=1,\dots,m}\left|R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l)\right|^4 = O\left(\frac{1}{m^2}\right)$, if the stronger assumption applies. Otherwise this term simply tends to zero. Concerning the other summand, the application of the Cauchy-Schwarz inequality readily yields the result when additionally employing Propositions 4.1 and 4.2. □

Lemma 4.3 (Relationship between $M\!F_{\lfloor uT\rfloor}$ and $M\!F^{\varepsilon}_{\lfloor uT\rfloor}$).
Under the same assumptions as in Lemma 4.2 and with $A_{t,T}(\lambda_l) := \sum_{j=-\infty}^{\infty} a_{t,T}(j)e^{-i\lambda_l j}$, *the result (4.21) extends to*

$$M\!F_{\lfloor uT\rfloor}(\lambda_l) = A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)M\!F^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l) + R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l), \tag{4.35}$$

with $R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l) := R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l) + \tilde{R}_{\lfloor uT\rfloor,m}(\lambda_l)$ *and* $R^{(1)}_{\lfloor uT\rfloor,m}(\lambda_l)$ *as in Lemma 4.2. Then*

$$ER^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l) = 0 \tag{4.36}$$

and

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} E|\tilde{R}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 = O\left(\frac{m^4}{T^4}\right). \tag{4.37}$$

In particular, we get

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} E|R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 \to 0 \text{ as } m\to\infty. \tag{4.38}$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|} < \infty$, *then even*

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} E|R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)|^4 = O\left(\frac{1}{m^2}\right), \tag{4.39}$$

as well as

$$E\left(\frac{1}{\sqrt{2m+1}}\sum_{l=1}^{2m+1}|R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)|^2\right)^2 \to 0 \text{ as } m\to\infty. \tag{4.40}$$

Proof. We first split $M\!F_{\lfloor uT \rfloor}(\lambda_l)$ to be able to apply Lemma 4.2:

$$\begin{aligned} M\!F_{\lfloor uT \rfloor}(\lambda_l) &= M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l) + (M\!F_{\lfloor uT \rfloor}(\lambda_l) - M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)) \\ &= A\,(u,\lambda_l)\, M\!F^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) + R^{(1)}_{\lfloor uT \rfloor,m}(\lambda_l) + (M\!F_{\lfloor uT \rfloor}(\lambda_l) - M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)) \\ &= A_{l+\zeta_{\lfloor uT \rfloor,l} m-m,T}(\lambda_l) M\!F^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) + R^{(1)}_{\lfloor uT \rfloor,m}(\lambda_l) \\ &+ \Big(A\,(u,\lambda_l) - A_{l+\zeta_{\lfloor uT \rfloor,l} m-m,T}(\lambda_l)\Big)\, M\!F^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) + (M\!F_{\lfloor uT \rfloor}(\lambda_l) - M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)). \end{aligned} \tag{4.41}$$

As already defined in the above Lemma, $\tilde{R}$ is the additional remainder we obtain when bridging the gap from the stationary approximation to the actual locally stationary time series. Inspecting (4.41) yields the exact structure of this remainder:

$$\tilde{R}_{\lfloor uT \rfloor,m}(\lambda_l) = \Big(A\,(u,\lambda_l) - A_{l+\zeta_{\lfloor uT \rfloor,l} m-m,T}(\lambda_l)\Big)\, M\!F^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) + (M\!F_{\lfloor uT \rfloor}(\lambda_l) - M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)).$$

Here, we can easily see that $E\tilde{R}_{\lfloor uT \rfloor,m}(\lambda_l) = 0$. With (4.22), we have $ER^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l) = 0$.

Propositions 4.1 and 4.2 immediately result in

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \Big|A\,(u,\lambda_l) - A_{l+\zeta_{\lfloor uT \rfloor,l} m-m,T}(\lambda_l)\Big|^4 E\left|M\!F^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)\right|^4 = O\left(\frac{m^4}{T^4}\right). \tag{4.42}$$

Concerning the difference between the two Fourier transforms $M\!F_{\lfloor uT \rfloor}(\lambda_l)$ and $M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)$, we may use the procedure as chosen in the proof of Proposition 4.1, since

$$\sup_{u\in[0,1]} \left|a\,(u,j_1) - a_{\lfloor uT \rfloor,T}(j_1)\right| \le \frac{K}{Tl(j_1)}.$$

Hence, we continue with

$$\begin{aligned} &\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|M\!F_{\lfloor uT \rfloor}(\lambda_l) - M\!F^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)|^4 \\ \le\; &\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \frac{1}{(2m+1)^2} \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \\ &\Bigg|\Big[a_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_1,T}(j_1) - a\,(u,j_1)\Big] \cdot \Big[a_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_2,T}(j_2) - a\,(u,j_2)\Big] \\ &\cdot\Big[a_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_3,T}(j_3) - a\,(u,j_3)\Big] \cdot \Big[a_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_4,T}(j_4) - a\,(u,j_4)\Big] \\ &\cdot E(\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_1-j_1}\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_2-j_2}\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_3-j_3}\varepsilon_{l+\zeta_{\lfloor uT \rfloor,l} m-m+t_4-j_4}) \\ &\cdot e^{i(t_2-t_1+t_3-t_4)\lambda_l}\Bigg|. \end{aligned} \tag{4.43}$$

Considering that $\tilde{X}$ is the stationary approximation at time u, we need to pay a little attention, though, as

$$\begin{aligned}
& MF_{\lfloor uT \rfloor}(\lambda_l) - MF^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l) \\
= \; & \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \sum_{j=-\infty}^{\infty} \left[\left(a_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t, T}(j) - a(u,j) \right) \right] \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t - j} e^{-it\lambda_l} \\
= \; & \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \sum_{j=-\infty}^{\infty} \left[\left(a_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t, T}(j) - a\left(\frac{l+\zeta_{\lfloor uT \rfloor, l} m - m + t}{T}, j \right) \right) \right. \\
& \left. + \left(a\left(\frac{l+\zeta_{\lfloor uT \rfloor, l} m - m + t}{T}, j \right) - a(u,j) \right) \right] \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t - j} e^{-it\lambda_l}.
\end{aligned}$$

With Definition 2.1((c)),

$$\sup_{m,u,l} \left| a_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t, T}(j) - a\left(\frac{l+\zeta_{\lfloor uT \rfloor, l} m - m + t}{T} \right) \right| \leq \sup_{s} \left| a_{s,T}(j) - a\left(\frac{s}{T}, j \right) \right| \leq \frac{K}{Tl(j)}.$$

The second summand an be bounded (see Definition 2.1) by

$$\left| a\left(\frac{l+\zeta_{\lfloor uT \rfloor, l} m - m + t}{T}, j \right) - a(u,j) \right| \leq \frac{\left| \frac{l+\zeta_{\lfloor uT \rfloor, l} m - m + t}{T} - u \right|}{l(j)} \leq \frac{K}{Tl(j)}.$$

The next step is now to consider all cases for which the expected value of the errors for $0 \leq t \leq 2m$, see (4.43), is not zero. That is either, in the notation of (4.43), that $\{t_1 - j_1 = t_2 - j_2 = t_3 - j_3 = t_4 - j_4\}$ or that one of the three cases

$$\begin{aligned}
t_1 - j_1 = t_2 - j_2 \; & \neq \; t_3 - j_3 = t_4 - j_4, \\
t_1 - j_1 = t_3 - j_3 \; & \neq \; t_2 - j_2 = t_4 - j_4, \\
t_1 - j_1 = t_4 - j_4 \; & \neq \; t_2 - j_2 = t_3 - j_3
\end{aligned}$$

holds true, because we have defined $\zeta_{\lfloor uT \rfloor, l} := div\left(\lfloor uT \rfloor - \lfloor \frac{m}{2} \rfloor \right) - \mathbb{1}_{\{l \geq mod\left(\lfloor uT \rfloor - \lfloor \frac{m}{2} \rfloor \right)\}}$.

So, again, we split the sums in $\sup_{u \in [0,1]} \sup_{l=1,\ldots,m} E|MF_{\lfloor uT \rfloor}(\lambda_l) - MF^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)|^4$ accordingly and bound them one by one. For equal indices of the errors, we get the upper bound

$$\begin{aligned}
& E(\varepsilon_1^4) \frac{K}{(2m+1)^2} \cdot \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \left(\left| \frac{m}{Tl(j_1)} \right| \right. \\
& \left. \cdot \left| \frac{m}{Tl(j_2)} \right| \cdot \left| \frac{m}{Tl(j_3)} \right| \cdot \left| \frac{m}{Tl(j_4)} \right| \right) \mathbb{1}_{\{t_1 - j_1 = t_2 - j_2 = t_3 - j_3 = t_4 - j_4\}} \\
= \; & O\left(\frac{m^3}{T^4} \right),
\end{aligned}$$

Secondly, we exemplarily consider $t_1 - j_1 = t_4 - j_4 \neq t_2 - j_2 = t_3 - j_3$ and get a bound of

$$\frac{1}{(2m+1)^2} \sum_{t_1,t_2,t_3,t_4=0}^{2m} \sum_{j_1,j_2,j_3,j_4=-\infty}^{\infty} \left(\left| \frac{m}{Tl(j_1)} \right| \cdot \left| \frac{m}{Tl(j_2)} \right| \right.$$
$$\left. \cdot \left| \frac{m}{Tl(j_3)} \right| \cdot \left| \frac{m}{Tl(j_4)} \right| \right) \mathbb{1}_{\{t_1-j_1=t_4-j_4\neq t_2-j_2=t_3-j_3\}}$$
$$= O\left(\frac{m^4}{T^4}\right).$$

Finally, collecting the results on the upper bounds, we get

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|MF_{\lfloor uT \rfloor}(\lambda_l) - MF^{\tilde{X}}_{\lfloor uT \rfloor}(\lambda_l)|^4 = O\left(\frac{m^4}{T^4}\right) = O\left(\frac{1}{m^2}\right). \tag{4.44}$$

At last, the proof of (4.37) is completed by putting together the results of (4.42) and (4.44).
The result (4.39) is then a consequence of the result (4.23) of Lemma 4.2.
Now, to prove (4.24), recall that

$$R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l) := R^{(1)}_{\lfloor uT \rfloor,m}(\lambda_l) + \tilde{R}_{\lfloor uT \rfloor,m}(\lambda_l).$$

According to Lemma 4.2 Equation (4.24)

$$E\left(\frac{1}{\sqrt{2m+1}} \sum_{j=1}^{2m+1} |R^{(1)}_{\lfloor uT \rfloor,m}(\lambda_l)|^2\right)^2 \to 0, \text{ for } m \to \infty.$$

We further know from the above proof that

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} |\tilde{R}_{\lfloor uT \rfloor,m}(\lambda_l)|^4 = O\left(\frac{m^4}{T^4}\right).$$

Hence,

$$E\left(\frac{1}{\sqrt{2m+1}} \sum_{j=1}^{2m+1} |\tilde{R}_{\lfloor uT \rfloor,m}(\lambda_l)|^2\right)^2 = O\left(\frac{m^5}{T^4}\right).$$

(4.40) now follows by Cauchy-Schwarz. □

Theorem 4.3 (Relationship between $MI_{\lfloor uT \rfloor,m}$ and $MI^{\varepsilon}_{\lfloor uT \rfloor,m}$).
Under the same assumptions as in Theorem 4.2 and $A_{t,T}(\lambda_l) := \sum_{j=-\infty}^{\infty} a_{t,T}(j)e^{-i\lambda_l j}$, the result (4.31) extends to

$$MI_{\lfloor uT \rfloor,m}(\lambda_l) = |A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)|^2 MI^{\varepsilon}_{\lfloor uT \rfloor,m}(\lambda_l) + R'_{\lfloor uT \rfloor,m}(\lambda_l), \tag{4.45}$$

with

$$\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R'_{\lfloor uT\rfloor,m}(\lambda_l)|^2 \to 0 \textit{ for } m\to\infty. \tag{4.46}$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty$,

$$\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R'_{\lfloor uT\rfloor,m}(\lambda_l)|^2 = O\left(\frac{1}{m}\right), \tag{4.47}$$

as well as

$$E\left(\frac{1}{\sqrt{2m+1}}\sum_{l=1}^{2m+1}|R'_{\lfloor uT\rfloor,m}(\lambda_l)|\right)^2 \to 0 \textit{ as } m\to\infty. \tag{4.48}$$

Proof. This Theorem is an immediate consequence of Lemma 4.3, Propositions 4.1 and 4.2, as well as the application of the Cauchy-Schwarz inequality. The proof is in analogy to the proof of Theorem 4.2. □

CHAPTER 5

Distributional properties of the moving Fourier coefficients

5.1 Variance

Theorem 5.1.

(a)

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} | \mathit{Var}(\mathit{MF}_{\lfloor uT \rfloor}(\lambda_l)) - 2\pi f(u,\lambda_l)| \to 0.$$

(b) If additionally $\sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} |a(u,j)|\sqrt{|j|} < \infty$,

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} | \mathit{Var}(\mathit{MF}_{\lfloor uT \rfloor}(\lambda_l)) - 2\pi f(u,\lambda_l)| = O\left(\frac{1}{\sqrt{m}}\right).$$

Proof. Lemma 4.3 provides the following relation

$$\mathit{MF}_{\lfloor uT \rfloor}(\lambda_l) = A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)\mathit{MF}^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) + R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l),$$

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l)|^4 \to 0.$$

If additionally $\sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} |a(u,j)|\sqrt{|j|} < \infty$,

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E|R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l)|^4 = O\left(\frac{1}{m^2}\right).$$

If we apply Lemma 4.3, we now get

$$\begin{aligned} \mathrm{Var}(MF_{\lfloor uT \rfloor}(\lambda_l)) &= A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)\overline{A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)}\,\mathrm{Var}(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)) \\ &\quad + A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)\,\mathrm{Cov}(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l), R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l)) \\ &\quad + \overline{A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l)}\,\mathrm{Cov}(R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l)), MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)) + \mathrm{Var}(R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l)) \\ &=: A_1 + A_2 + \overline{A_2} + A_3. \end{aligned}$$

$\boxed{A_1}$

By Proposition 4.1 and Lemma 4.1 it holds

$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \mathrm{Var}(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)) = 1.$

With Proposition 4.2 (b) and (c),

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \left| A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l) - A(u,\lambda_l) \right| = O\left(\frac{m}{T}\right).$$

Hence, with Definition 2.2,

$$\begin{aligned} A_1 &= |A(u,\lambda_l)|^2\,\mathrm{Var}(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l)) + O\left(\frac{m}{T}\right) \\ &= 2\pi f(u,\lambda_l) + O\left(\frac{m}{T}\right). \end{aligned}$$

$\boxed{A_2}$ From Proposition 4.2 it follows that

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} \left| A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l) \right| < \infty.$$

Furthermore, Proposition 4.1 tells us that

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} E\left| MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) \right|^2 = 1.$$

Finally, the use of Lemma 4.3 and the application of the Cauchy-Schwarz inequality yield

$$\sup_{u\in[0,1]} \sup_{l=1,\ldots,m} A_{l+\zeta_{\lfloor uT \rfloor,l}m-m,T}(\lambda_l) E\left(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) \cdot R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_{-l}) \right) = o(1). \tag{5.1}$$

If the stronger assumptions apply, we have the rate $\frac{1}{\sqrt{m}}$.

Note, that (4.36) tells us that $E(R^{(2)}_{\lfloor uT \rfloor,m}(\lambda_l)) = 0$. Due to this and Lemma 4.1, relation (5.1) entails

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\,\mathrm{Cov}(M\!F^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l), R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)) = o(1)$$
$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\,\mathrm{Cov}(R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l), M\!F^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)) = o(1).$$

If the stronger assumptions apply, we have the rate $\frac{1}{\sqrt{m}}$.

$\boxed{A_3}$ Lemma 4.3 and the application of the Cauchy-Schwarz inequality yields

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} E\left(R^{(2)}_{\lfloor uT\rfloor}(\lambda_l)\cdot R^{(2)}_{\lfloor uT\rfloor}(\lambda_{-l})\right) = o(1). \tag{5.2}$$

The rate given the stronger set of assumptions is $\frac{1}{m}$.
Relation (5.2) now entails

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} \mathrm{Var}\left(R^{(2)}_{\lfloor uT\rfloor}(\lambda_l)\right) = o(1).$$

With the set of stronger assumptions, we can now even get

$$\sup_{u\in[0,1]}\sup_{l=1,\dots,m} |\mathrm{Var}(M\!F_{\lfloor uT\rfloor}(\lambda_l)) - 2\pi f(u,\lambda_l)| = O\left(\frac{m}{T}+\frac{1}{\sqrt{m}}\right) = O\left(\frac{1}{\sqrt{m}}\right).$$

. □

5.2 Covariance structure

For the next theorems, we define sets $\mathcal{A}_1(a_m,u)$ and $\mathcal{A}_2$ in order to rule out λ_j,λ_l converging to the same frequency from different sides, as we do not get asymptotic uncorrelated moving Fourier coefficients in these cases. This phenomenon is not indigenous to our moving procedure. There are also procedures like the block bootstrap – procedures which basically mimick the dependence structure but exhibit some minor exceptions at certain cut off points.
Accordingly, two main questions have to be answered: What are the situations in the moving case where asymptotic uncorrelation is not fulfilled? And why do the situations occur in our case?
Ad 1: Looking at the suggestive Figure 3.1, one can see some 'break' concerning the indices. For $k=\lfloor\frac{m}{2}\rfloor+2$ we have the sequence $c_{m+1},c_2,c_3,\dots,c_m$ for the moving Fourier transform at frequencies $\lambda_1,\dots,\lambda_m$. At frequency $\lambda_{mod(\lfloor\frac{m}{2}\rfloor+2-\lfloor\frac{m}{2}\rfloor)}=\lambda_2$, the break occurs. That is, for λ_j and λ_k converging to λ_2 from different sides, we encounter a situation where the actual frequencies get closer and closer whereas the indices of the coefficients do not.

Ad 2: The reason for these situations – getting local moving Fourier coefficients, which are not asymptotically uncorrelated – is due our very good up-to-dateness of the procedure. Meaning: We use only m moving Fourier coefficients and then assume that we have already extracted all information about the second order structure of the time series. See Remark 3.4 for the explicit argument. Usually, one would have calculated all $2m+1$ coefficients. We call that the long version of our transformation. We, however, double the coefficients (see Section 3.3) in an adequate way when going back into time domain. So, having generated the m^{th} coefficient, one would start duplexing the newest information when using the long version of the moving procedure. At that point, our procedure, however, starts updating the oldest information, that is the oldest coefficient. In the easier understandable non-moving case this problem corresponds to: Take $2m+1$ real random variables and use the Fourier transform. This yields m complex coefficients with 'new' information, that is $2m$ real coefficients with new information. So there is a one-to-one relation between the information carried by the original data and the information carried by the transformed data.
If one took $2m+1$ real random variables and got $2m+1$ complex coefficients out of them this would be $4m+2$ real coefficients, all carrying new information and thus the information in the frequency domain is double the amount of information in the time domain, with no doubling of information. Not being in the land of milk and honey we need to pay for this overflowing information – by hiccups in the dependence structure, compared to the long version.

Theorem 5.2. *For $u \in [0,1]$, $j \neq l = 1, \ldots, m$, let*

$$\begin{aligned} \zeta_{\lfloor uT \rfloor, l} &:= div\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(\lfloor uT \rfloor - \lfloor \frac{m}{2} \rfloor\right)\right\}}, \\ l' &:= l + \zeta_{\lfloor uT \rfloor, l}\, m, \\ j' &:= j + \zeta_{\lfloor uT \rfloor, j}\, m. \end{aligned}$$

Then,

$$\sup_{1 \leq l \neq j \leq m} \max\left(\frac{m}{|l'-j'|}, |l-j|\right) \left|E(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l) MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_j))\right| = O(1).$$

Remark 5.1
The difference $l'-j'$ equals either $l-j$, $l-j+m$ or $l-j-m$, depending on the time $\lfloor uT \rfloor$ we are currently at, as the composition of the local moving Fourier coefficients of old and new moving Fourier coefficients changes throughout time.

Proof. Lemma 4.1 allows us to write

$$\begin{aligned} Cov\left(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l), MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)\right) &= E\left(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)\right) \\ &= \frac{1}{2m+1}\sum_{t_1,t_2=0}^{2m} E(\varepsilon_{l'-m+t_1}\varepsilon_{j'-m+t_2}) \\ &\quad \cdot e^{-i\lambda_l t_1}e^{i\lambda_j t_2}. \end{aligned}$$

Since $E(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l), MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)) = 0$ iff $l+\zeta_{\lfloor uT\rfloor,l}m - m + t_1 = j + \zeta_{\lfloor uT\rfloor,j}m - m + t_2$, we get

$$\begin{aligned} &\sup_{1\le l\neq j\le m} \max\left(\frac{m}{|l'-j'|}, |l-j|\right)\left|E(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j))\right| \\ = &\sup_{1\le l\neq j\le m} \max\left(\frac{m}{|l'-j'|}, |l-j|\right)\Bigg|\frac{1}{2m+1}\sum_{t_1,t_2=0}^{2m} E(\varepsilon_{l+\zeta_{\lfloor uT\rfloor,l}m-m+t_1}\varepsilon_{j+\zeta_{\lfloor uT\rfloor,j}m-m+t_2}) \\ &\qquad \cdot e^{-i\lambda_l t_1}e^{i\lambda_j t_2}\cdot \mathbb{1}_{\{l'+t_1=j'+t_2\}}\Bigg| \\ = &\sup_{1\le l\neq j\le m} C\max\left(\frac{1}{|l'-j'|}, \frac{|l-j|}{m}\right)\left|\sum_{t_1=\max\{0,(j'-l')\}}^{\min\{2m,2m+(j'-l')\}} e^{-i\lambda_{l-j}t_1 + i\lambda_j(l'-j')}\right|. \end{aligned}$$

We define

$$g_1(u,l,j) = g_1 := \max\{0,(j'-l')\},\ldots,\min\{2m, 2m+(j'-l')\} =: g_2 = g_2(u,l,j).$$

Then,

$$\begin{aligned} &\sup_{1\le l\neq j\le m} \max\left(\frac{1}{|l'-j'|}, \frac{|l-j|}{m}\right)\left|e^{i\lambda_j(l'-j')}\sum_{t_1=g_1}^{g_2} e^{-i\lambda_{l-j}t_1}\right| \\ = &\sup_{1\le l\neq j\le m} \max\left(\frac{1}{|l'-j'|}, \frac{|l-j|}{m}\right)\left|\sum_{t_1=1}^{g_2-g_1+1} e^{-i\lambda_{l-j}t_1}\right|. \end{aligned}$$

Application of Lemma A.4 in Kirch [27] yields that uniformly in l, j.

$$\left|\sum_{t_1=1}^{g_2-g_1+1}\cos\left(\frac{2\pi(l-j)}{2m+1}t_1\right)\right| = O\left(\min\left(\frac{m}{|l-j|}, |g_2-g_1|\right)\right).$$

Analogously for the sine. Consequently,

$$\left|\sum_{t_1=1}^{g_2-g_1+1} e^{-i\lambda_{l-j}t_1}\right| = O\left(\min\left(\frac{m}{|l-j|}, |g_2-g_1|\right)\right). \tag{5.3}$$

On the other hand

$$\left|\sum_{t_1=1}^{g_2-g_1+1} e^{-i\lambda_{l-j}t_1}\right| = \left|\sum_{t_1=0}^{2m} e^{-i\lambda_{l-j}t_1} - \sum_{t_1=g_2-g_1+2}^{2m+1} e^{-i\lambda_{l-j}t_1}\right| = \left|\sum_{t_1=g_2-g_1+1}^{2m} e^{-i\lambda_{l-j}t_1}\right|$$

using that $\sum_{k=0}^{2m} e^{-ik\lambda} = 0$ for $\lambda \neq 2\pi\mathbb{Z}$.
Again, with Lemma A.4 in Kirch [27],

$$\left|\sum_{t_1=1}^{g_2-g_1+1} e^{-i\lambda_{l-j}t_1}\right| = O\left(\min\left(\frac{m}{|l-j|}, |2m-(g_2-g_1)|\right)\right). \tag{5.4}$$

With (5.4) we get

$$\begin{aligned}
\left|\sum_{t_1=1}^{g_2-g_1+1} e^{-i\lambda_{l-j}t_1}\right| &= O\left(\min\left(\frac{m}{|l-j|}, |g_2-g_1|, |2m-(g_2-g_1)|\right)\right) \\
&= O\left(\min\left(\frac{m}{|l-j|}, |2m-|l'-j'||, |l'-j'|\right)\right) \\
&= O\left(\min\left(\frac{m}{|l-j|}, |l'-j'|\right)\right) \\
&= O\left(\left(\max\left(\frac{|l-j|}{m}, \frac{1}{|l'-j'|}\right)\right)^{-1}\right).
\end{aligned}$$

All in all, we now have that

$$\sup_{1\leq l\neq j\leq m} \max\left(\frac{m}{|l'-j'|}, |l-j|\right) \left|E(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l), MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j))\right| = O(1).$$

. □

Let a_m be a sequence with $a_m \to \infty$ with $a_m/m \to 0$.
The following set $\mathcal{A}_1(a_m, u)$ includes all indices of Fourier frequencies who are either sufficiently far apart or, if they are close to each other, one needs to ensure that the indices do not relate to coefficients which are very different concerning up-to-dateness.
Let a_m be a sequence with $a_m \to \infty$, $\frac{a_m}{m} \to 0$. Then define

$$\mathcal{A}_1(a_m, u) := \{(l,j) \in \{1,\ldots,m\}^2 | (l \neq j) \wedge [|l'-j'| \leq a_m \vee |l-j| \geq a_m]\}, \tag{5.5}$$

with

$$\begin{aligned}
\zeta_{\lfloor uT\rfloor,l} &:= div\left(\lfloor uT\rfloor - \left\lfloor\frac{m}{2}\right\rfloor\right) - \mathbb{1}_{\left\{l\geq mod\left(\lfloor uT\rfloor - \left\lfloor\frac{m}{2}\right\rfloor\right)\right\}}, \\
l' &:= l - \zeta_{\lfloor uT\rfloor,l}m, \\
j' &:= j - \zeta_{\lfloor uT\rfloor,j}m.
\end{aligned}$$

Denote by $\mathcal{A}_1^{cu}(a_m, u) := \{l, j | l \neq j\} \setminus \mathcal{A}_1(a_m, u)$. It holds (see Remark 5.2) that

$$|\mathcal{A}_1^{cu}(a_m, u)| = a_m^2.$$

Theorem 5.3. *Let a_m be a sequence with $a_m \to \infty$, $\frac{a_m}{m} \to 0$ and $\mathcal{A}_1(a_m, u)$ as in (5.5) .*

(a)

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} Cov(MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_l), MF^{\varepsilon}_{\lfloor uT \rfloor}(\lambda_j)) = O\left(\frac{1}{a_m} + \frac{a_m}{m}\right). \tag{5.6}$$

(b)

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} Cov(MF_{\lfloor uT \rfloor}(\lambda_l), MF_{\lfloor uT \rfloor}(\lambda_j)) \to 0. \tag{5.7}$$

(c) If additionally $\sup_{u\in[0,1]} \sum_{j=-\infty}^{\infty} |a(u, j)|\sqrt{|j|} < \infty$,

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} Cov(MF_{\lfloor uT \rfloor}(\lambda_l), MF_{\lfloor uT \rfloor}(\lambda_j)) = O\left(\frac{1}{a_m} + \frac{a_m}{m} + \frac{1}{\sqrt{m}}\right). \tag{5.8}$$

(d)

$$\sup_{1\leq l\neq j\leq m} Cov(MF_{\lfloor uT \rfloor}(\lambda_l), MF_{\lfloor uT \rfloor}(\lambda_j)) = O(1). \tag{5.9}$$

Proof. Part (a) is an immediate corollary of Theorem 5.2, substituting the set $\{1, \ldots, m\}$ by the special set $\mathcal{A}_1(a_m, u)$.

Concerning parts (b) to (d):

Analogously to the previous proof of Theorem 5.1, we now utilize Lemma 4.3

$$\begin{aligned}
& \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} Cov\left(MF_{\lfloor uT\rfloor}(\lambda_l), MF_{\lfloor uT\rfloor}(\lambda_j))\right) \\
= & \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} E\left(MF_{\lfloor uT\rfloor}(\lambda_l)\cdot MF_{\lfloor uT\rfloor}(\lambda_j))\right) \\
\leq & \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(-\lambda_j) \\
& \cdot E\left(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)\cdot MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)\right) \\
& + \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)E\left(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)\cdot R^{(2)}_{\lfloor uT\rfloor,m}(-\lambda_j)\right) \\
& + \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(-\lambda_j)E\left(R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)\cdot MF^{\varepsilon}_{\lfloor uT\rfloor}(-\lambda_j)\right) \\
& + \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} E\left(R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)\cdot R^{(2)}_{\lfloor uT\rfloor,m}(-\lambda_j)\right) \\
=: & \; B_1 + B_2 + \overline{B_2} + B_3.
\end{aligned}$$

Incorporating the result for the errors given by part (a), we get for the term B_1:

$\boxed{B_1}$ With Proposition 4.2 (a) and (c), we get

$$\sup_{1\leq l\neq j\leq m}\left|A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\right| < \infty. \tag{5.10}$$

Hence,

$$\begin{aligned}
\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} |B_1| \leq & \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} |A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(-\lambda_j)| \\
& \cdot \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} E\left(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l), MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)\right) \\
= & \; O\left(\frac{1}{a_m}+\frac{a_m}{m}\right).
\end{aligned}$$

$\boxed{B_2}$ From (5.10) it follows that

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)}\left|A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\right| < \infty.$$

Furthermore, Proposition 4.1 tells us that

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} E\left|MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)\right|^2 = 1.$$

Finally, the use of Lemma 4.3 and the application of the Cauchy-Schwarz inequality yield

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)E\left(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l)\cdot R^{(2)}_{\lfloor uT\rfloor,m}(-\lambda_j)\right) = o(1). \quad (5.11)$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty$ we have the rate $\frac{1}{\sqrt{m}}$.

Due to $\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} E(R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)) = 0$ (4.36) and Lemma 4.1, relation (5.11) entails

$$\begin{aligned}\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\operatorname{Cov}(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l),R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_j)) &= o(1)\\ \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\operatorname{Cov}(R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l),MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)) &= o(1).\end{aligned}$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty$

$$\begin{aligned}\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\operatorname{Cov}(MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_l),R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_j)) &= O\left(\frac{1}{\sqrt{m}}\right)\\ \sup_{(l,j)\in\mathcal{A}_1(a_m,u)} A_{l+\zeta_{\lfloor uT\rfloor,l}m-m,T}(\lambda_l)\operatorname{Cov}(R^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l),MF^{\varepsilon}_{\lfloor uT\rfloor}(\lambda_j)) &= O\left(\frac{1}{\sqrt{m}}\right).\end{aligned}$$

$\boxed{B_3}$ Lemma 4.3 and the application of the Cauchy-Schwarz inequality yields

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} E\left(R^{(2)}_{\lfloor uT\rfloor}(\lambda_l)\cdot R^{(2)}_{\lfloor uT\rfloor}(-\lambda_j)\right) = o(1). \quad (5.12)$$

The rate given the stronger set of assumptions is $\frac{1}{m}$.
Due to (4.36) relation (5.12) entails

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} \operatorname{Cov}\left(R^{(2)}_{\lfloor uT\rfloor}(\lambda_l),R^{(2)}_{\lfloor uT\rfloor}(-\lambda_j)\right) = o(1).$$

If additionally $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty$

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} \operatorname{Cov}\left(R^{(2)}_{\lfloor uT\rfloor}(\lambda_l),R^{(2)}_{\lfloor uT\rfloor}(-\lambda_j)\right) = O\left(\frac{1}{m}\right).$$

. □

Remark 5.2
Let x be defined as

$$x \;:=\; mod\left(\lfloor uT\rfloor - \left\lfloor\frac{m}{2}\right\rfloor\right). \quad (5.13)$$

The set $\mathcal{A}_1(a_m, u)$ can then also be written as

$$\mathcal{A}_1(a_m, u) := \{(l, j) \in \{1, \ldots, m\}^2 | (l \neq j) \wedge [(\max\{|l - x|, |j - x|\} \geq a_m) \\ \veebar (x \leq l, j \leq x + a_m) \veebar (x - a_m \leq l, j < x)]\}, \tag{5.14}$$

which makes the interpretation of the set more obvious. First, we illustrate the meaning of x: When considering a set of local moving coefficients (see Definition 3.1), we have, depending, what time they refer to, a Fourier frequency λ_x whose corresponding cofficient is freshest. The coefficient corresponding to the next Fourier frequency λ_{x+1} is the oldest one, due to construction. The set $\mathcal{A}_1(a_m, u)$, hence, contains all pairs of non-equal indices of Fourier frequencies which are either both smaller or both larger than x or which are, if there is one smaller and one larger, sufficiently far away from x.
The complementary set to $\mathcal{A}(a_m, u)$ includes to following pairs of indices (l, j): Let w.l.o.g. $l \leq j$. If one of the to indices is further away from x than a_m, we are no longer in $\mathcal{A}^c(a_m, u)$. Thus, the complementary set comprises only of pairs of indices $(l, j) \in \{1, \ldots, m\}^2$ with $x - a_m < l < x$ and $x \leq j < x + a_m$, and, not to forget, all pairs (l, l).
Now, concerning the cardinality of $\mathcal{A}^c(a_m, u)$: We have m possibilities to choose l and set $j = l$. Further, there are a_m possibilities, to choose l such that $x - a_m < l < x$ and another a_m to choose for each of the l's an index j, such that $x \leq j < x + a_m$. Hence, we have

$$|\mathcal{A}^c(a_m, u)| = m + a_m^2.$$

If we are only looking at indices (l, j), with $l \neq j$, that is the set $\mathcal{A}_1^{cu}(a_m, u) := \{(l, j) \in \{1, \ldots, m\}^2 | (l \neq j) \wedge \overline{[|l' - j'| \leq a_m \vee |l - j| \geq a_m]}\}$. We have

$$|\mathcal{A}_1^{cu}(a_m, u)| = a_m^2.$$

When not considering moving Fourier coefficients centred around the same point at time $\lfloor uT \rfloor$, but centred around $\lfloor uT \rfloor$ and $\lfloor uT \rfloor + s$, $s = 1, \ldots, cm$, we are still left with the problem we have already faced when formulating Theorem 5.3. In that case we have worked around it with the help of $\mathcal{A}_1(a_m, u)$. In the situation of the second point in time being s apart, the work around, however, is slightly changed and slightly more tricky due to the additional variable s.
For the next result we will therefore first evoke some intuition of how we construct the set $\mathcal{A}_2$ used in Theorem 5.4.
Consider the points in time $\lfloor uT \rfloor$ and $\lfloor uT \rfloor + s$ for some $s = 1, \ldots, cm$. We use Fourier frequencies with indices up to $\pm \left\lfloor \frac{m}{2} \right\rfloor$ around each point. For the sake of notation needed for the proof of Theorem 5.4 we set

$$U(s) \ := \ \left\{ \lfloor uT \rfloor + s - \left\lfloor \frac{m}{2} \right\rfloor, \ldots, \lfloor uT \rfloor + s + \left\lfloor \frac{m}{2} \right\rfloor \right\}. \tag{5.15}$$

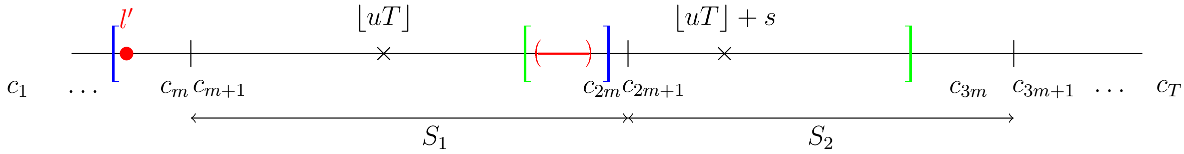


Figure 5.1: Illustrative sketch: The set $U(0)$ – see Equation (5.15) – is marked blue, the set $V(s)$ is marked green.

The coefficients in set $U(0)$ correspond to frequencies $\lambda_{mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)}, \dots, \lambda_{mod\left(\lfloor uT \rfloor + \left\lfloor \frac{m}{2} \right\rfloor\right)}$. Vice versa, when near the time $\lfloor uT \rfloor$, the corresponding moving Fourier coefficient to frequency λ_l, $l = 1, \dots, m$, is, see (3.12),

$$c_{l+\zeta_{\lfloor uT \rfloor, l} m},$$

with $\zeta_{\lfloor uT \rfloor, l} = div\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)\right\}}$. We call the above index of the moving Fourier coefficient l'.
Analogously,

$$j' := j + \zeta_{\lfloor uT \rfloor + s, j} m.$$

l' and j' can now take values from 1 to T, depending on the location of $\lfloor uT \rfloor$ and the value of s.
Now consider the following Figure 5.1 with some exemplary $\lfloor uT \rfloor$ and $\lfloor uT \rfloor + s$. No problem arises, as long as j' and l' remain in the same set, say S_k, with $div(j') = div(l')$, $k \in \mathbb{N}_0$. As soon as they sit in different stretches, the possibility of the problem arises that $|l' - j'| \to \infty$, while $|mod(l) - mod(j)| = |l - j|$ does not. The problem is banned as soon as we require that $|mod(j') - mod(l')| \geq a_m$. That is, for some exemplary l', j' must not be within the red area $l' + m \pm a_m$. In words, we need to ensure that we do not use moving Fourier coefficients corresponding to frequencies that are very near to each other, unless the coefficients themselves are located very near to each other (with a distance of the indices of less than m). By phrasing very near to each other, we mean that one can not find a sequence $a_m \to \infty$, $\frac{a_m}{m} \to 0$, with $|l - j| \geq a_m$.

We can now choose a set of indices $\mathcal{A}_2$, for which the moving Fourier coefficients at the corresponding frequencies are asymptotically uncorrelated.

Theorem 5.4. *Let a_m be a sequence with $a_m \to \infty$, $\frac{a_m}{m} \to 0$ and define*

$$A_2(s) := \{l' \in U(0), j' \in U(s) | [div(l') = div(j')] \vee [|mod(l') - mod(j')| \geq a_m]\},$$

with $U(0)$ and $U(s)$ as in (5.15).
Let $\mathcal{F}_\varepsilon^{-1}(\lambda_{1-\lceil \frac{m}{2} \rceil}), \ldots, \mathcal{F}_\varepsilon^{-1}(\lambda_m), \mathcal{F}_\varepsilon^{0}(\lambda_1), \ldots, \mathcal{F}_\varepsilon^{0}(\lambda_m), \ldots, \mathcal{F}_\varepsilon^{\frac{T}{m}}(\lambda_{\lfloor \frac{m}{2} \rfloor})$ denote the moving Fourier transforms (Definition 3.2 and Remark 3.3) of the innovations. Then

$$\sup_{s\geq 0} \sup_{\varphi \neq \psi \in \mathcal{A}_2(s)} Cov(\mathcal{F}_\varepsilon^{div(\varphi)-1}(\lambda_{mod(\varphi)}), \mathcal{F}_\varepsilon^{div(\psi)-1}(\lambda_{mod(\psi)})) = O\left(\frac{1}{a_m} + \frac{a_m}{m}\right). \quad (5.16)$$

Proof. Due to Lemma 4.1, it is sufficient to consider only

$$\sup_{s\geq 0} \sup_{(\varphi,\psi) \in \mathcal{A}_2(s)} E\left(\mathcal{F}_\varepsilon^{div(\varphi)-1}(\lambda_{mod(\varphi)}) \cdot \mathcal{F}_\varepsilon^{div(\psi)-1}(\lambda_{mod(\psi)})\right).$$

Hence,

$$\begin{aligned} &\sup_{(\varphi,\psi) \in \mathcal{A}_2(s)} E\left(\mathcal{F}_\varepsilon^{div(\varphi)-1}(\lambda_{mod(\varphi)}) \cdot \mathcal{F}_\varepsilon^{div(\psi)-1}(\lambda_{mod(\psi)})\right) \\ = &\sup_{(\varphi,\psi) \in \mathcal{A}_2(s)} \frac{1}{2m+1} \sum_{t_1,t_2=0}^{2m} E(\varepsilon_{\varphi-m+t_1}\varepsilon_{\psi-m+t_2}) e^{-i\lambda_{mod(\varphi)}t_1} e^{i\lambda_{mod(\psi)}t_2}. \end{aligned}$$

The expectancy $E(\varepsilon_{\varphi-m+t_1}\varepsilon_{\psi-m+t_2})$ equals not zero only if $\varphi + t_1 = \psi + t_2$.
The remaining proof is completely analogous to the proof of Theorem 5.2 and we end up with

$$\begin{aligned} &\sup_{s\geq 0} \sup_{(\varphi,\psi) \in \mathcal{A}_2(s)} \frac{1}{2m+1} \sum_{t_1,t_2=0}^{2m} E(\varepsilon_{\varphi-m+t_1}\varepsilon_{\psi-m+t_2}) e^{-i\lambda_{mod(\varphi)}t_1} e^{i\lambda_{mod(\psi)}t_2} \\ = &\; O\left(\sup_{s\geq 0} \sup_{(\varphi,\psi) \in \mathcal{A}_2(s)} \min\left(\frac{m}{|mod(\varphi) - mod(\psi)|}, |\varphi - \psi|, |2m - \varphi - \psi|\right)\right) \\ = &\; O\left(\frac{1}{a_m} + \frac{a_m}{m}\right). \end{aligned}$$

□

Remark 5.3
Theorem 5.4 can also be formulated using the notation of the local moving Fourier coefficients. Doing so, one needs another characterisation of the set $\mathcal{A}_2(s)$:

$$\mathcal{A}_2(s) := \mathcal{A}_{21}(s) \cup \mathcal{A}_{22}(s) \cup \mathcal{A}_{23}.$$

These sets $\mathcal{A}_{21}, \mathcal{A}_{22}, \mathcal{A}_{23}$ *are specified below. However, as the whole definition of* $\mathcal{A}_2(s)$ *is a bit nasty looking, we decided on the neater notation using the moving Fourier coefficients.*

For the alternative definition of $\mathcal{A}_2(s)$*, we are concerned with two points in time,* $\lfloor uT \rfloor$ *and* $\lfloor uT \rfloor + s$*. Define* x *as in (5.13) and*

$$y_s \ := \ mod\left(\lfloor uT \rfloor + s - \left\lfloor \frac{m}{2} \right\rfloor\right), \quad s \in \{-2m, \ldots, 2m\}.$$

Then, for $l, j \in 1, \ldots, m$*, define the conditions*

(c1) $(l < x) \wedge [(j < y_s) \vee ((j \geq y_s) \wedge (|l-j| \geq a_m))]$

(c2) $(l \geq x) \wedge [(j \geq y_s) \wedge ((j < y_s) \wedge (|l-j| \geq a_m))]$

(c3) $(l < x) \wedge [(j \geq y_s) \wedge ((j < y_s) \wedge (|l-j| \geq a_m))]$

(c4) $(l \geq x) \wedge (|j-l| \geq a_m)$

(c5) $|l-j| \geq a_m$.

Now define three sets of indices, depending on s: If $s = 1, \ldots, m-x$

$$\mathcal{A}_{21}(s) := \{(l,j) \in \{1, \ldots, m\}^2 | (l \neq j) \wedge [\,(c1) \veebar (c2)]\}. \tag{5.17}$$

If $s = m-x+1, \ldots, 2m-x$

$$\mathcal{A}_{22}(s) := \{(l,j) \in \{1, \ldots, m\}^2 | (l \neq j) \wedge [\,(c3) \veebar (c4)]\}. \tag{5.18}$$

If $s = 2m-x+1, \ldots, 2m$

$$\mathcal{A}_{23}(s) := \{(l,j) \in \{1, \ldots, m\}^2 | (l \neq j) \wedge (c5)\}. \tag{5.19}$$

Linking these sets, we get the set of indices which corresponds to the set $\mathcal{A}_2(s)$ *when using the notation of local moving Fourier coefficients.*

Lemma 5.1. *Let* $k, l = 1, \ldots, T$ *and denote the Fourier transforms by* $c_1, \ldots, c_T$ *as in Definition 3.2. Further, use the assumption on the function* l *discussed in Remark 2.3.*
Then

$$\sup_{|k-l| \geq 3m} \sup_{k,l} |\, Cov(c_k, c_l)| = o(1). \tag{5.20}$$

The proof is based on the concept of weak dependence developed by Doukhan and Louhichi [18]:

Definition 5.1 (($\vartheta, \mathcal{F}, \psi$)-weak dependence).
→ Doukhan and Louhichi [18], Definition 1

The sequence $(X_n)_{n\in\mathbb{N}}$ of r.v.s is called $(\vartheta, \mathcal{F}, \psi)$-weak dependent, if there exists a class $\mathcal{F}$ of real-valued functions, a sequence $\vartheta = (\vartheta_r)_{r\in\mathbb{N}}$ decreasing to zero at infinity, and a function ψ with arguments $(h, k, u, v) \in \mathcal{F}^2 \times \mathbb{N}^2$ such that for any u-tuple $(i_1, \ldots, i_u)$ and any v-tuple $(j_1, \ldots, j_v)$ with $i_1 \leq \ldots \leq i_u < i_u + r \leq j_1 \leq \ldots \leq j_v$, one has

$$| \, Cov(h(X_{i_1}, \ldots, X_{i_u}), k(X_{j_1}, \ldots, X_{j_v}))| \leq \psi(h, k, u, v)\vartheta_r,$$

for all functions h,k $\in \mathcal{F}$ that are defined on $\mathbb{R}^u$ and $\mathbb{R}^v$, respectively.

In the setting of Doukhan and Louhichi [18] we have $\mathcal{L} :=$ {set of bounded Lipschitz functions $\mathbb{R}^u \to \mathbb{R}$, for some $u \in \mathbb{N}$}. Further,

$$Lip(h) := \sup_{x\neq y} \frac{|h(x) - h(y)|}{||x - y||}$$

denotes the Lipschitz modulus of a function $h : \mathbb{R}^u \to \mathbb{R}$, where $\mathbb{R}^u$ is equipped with its l^1-norm. Furthermore, $\mathcal{L}_1 := \{h \in \mathcal{L}; \|h\|_\infty \leq 1\}$.
We cite subsection 4.3.4 in Nze and Doukhan [41], where the $(\vartheta, \mathcal{L}, \psi)$-weak dependence of two-sided sequences is stated. Accordingly, the infinite moving average

$$X_{t,T} = \sum_{j=-\infty}^{\infty} a_{t,T}(j)e_{t-j},$$

with $\vartheta_r = 2 \cdot E|\varepsilon_0| \cdot \sum_{|j|\geq m} \frac{1}{l(j)}$ and $\psi(h, k, u, v) = (uLip(h) + vLip(k))$ is $(\vartheta, \mathcal{L}, \psi)$-weak dependent.

Proposition 5.1
A locally stationary process as in Definition 2.1 is $(\vartheta, \mathcal{L}_1, \psi)$-weak dependent, with a sequence $\vartheta = (\vartheta_{2m})_{2m\in\mathbb{N}}$ decreasing to zero at infinity, and a function ψ with arguments $(h, k, u, v) \in {\mathcal{L}_1}^2 \times \mathbb{N}^2$ such that for any u-tuple $(i_1, \ldots, i_u)$ and any v-tuple $(j_1, \ldots, j_v)$ with $i_1 \leq \ldots \leq i_u < i_u + 2m \leq j_1 \leq \ldots \leq j_v$

$$| \, Cov(h(X_{i_1,T}, \ldots, X_{i_u,T}), k(X_{j_1,T}, \ldots, X_{j_v,T}))| \leq \psi(h, k, u, v)\vartheta_{2m},$$

with $\psi(g, h, u, v) := (u + v) \cdot \max\{Lip(h), Lip(k)\}$ and

$$\vartheta_{2m} := E|\varepsilon_0| \cdot \left(\sum_{|j|\geq m} \frac{1}{l(j)} \right). \tag{5.21}$$

Proof. To prove the $(\vartheta, \mathcal{L}_1, \psi)$-weak dependence of a locally stationary process as in Definition 2.1, we follow Nze and Doukhan [41] p.1007 and split $X_{t,T}$ in an m-dependent process $Y_{t,T}$ and some process $R_{t,T}$, being asymptotically negligible with respect to the L^1-norm.

$$Y_{t,T} := \sum_{j=-\lfloor \frac{m}{2} \rfloor}^{\lfloor \frac{m}{2} \rfloor} a_{t,T}(j)\varepsilon_{t-j} \qquad R_{t,T} := \sum_{|j|>\lfloor \frac{m}{2} \rfloor} a_{t,T}(j)\varepsilon_{t-j}.$$

Consequently,

$$\begin{aligned} & |\operatorname{Cov}(h(X_{i_1,T},\dots,X_{i_u,T}),k(X_{j_1,T},\dots,X_{j_v,T}))| \\ \leq\ & |\operatorname{Cov}(h(X_{i_1,T},\dots,X_{i_u,T})-h(Y_{i_1,T},\dots,Y_{i_u,T}),k(X_{j_1,T},\dots,X_{j_v,T}))| \\ & +|\operatorname{Cov}(h(Y_{i_1,T},\dots,Y_{i_u,T}),k(Y_{j_1,T},\dots,Y_{j_v,T}))| \\ & +|\operatorname{Cov}(h(Y_{i_1,T},\dots,Y_{i_u,T}),k(X_{j_1,T},\dots,X_{j_v,T})-k(Y_{j_1,T},\dots,Y_{j_v,T}))| \end{aligned}$$

Due to the m-dependence of the process $Y_{t,T}$,

$$|\operatorname{Cov}(h(Y_{i_1,T},\dots,Y_{i_u,T}),k(Y_{j_1,T},\dots,Y_{j_v,T}))| = 0.$$

We may therefore continue with

$$\begin{aligned} & |\operatorname{Cov}(h(X_{i_1,T},\dots,X_{i_u,T}),k(X_{j_1,T},\dots,X_{j_v,T}))| \\ \leq\ & |\operatorname{Cov}(h(X_{i_1,T},\dots,X_{i_u,T})-h(Y_{i_1,T},\dots,Y_{i_u,T}),k(X_{j_1,T},\dots,X_{j_v,T}))| \\ & +|\operatorname{Cov}(h(Y_{i_1,T},\dots,Y_{i_u,T}),k(X_{j_1,T},\dots,X_{j_v,T})-k(Y_{j_1,T},\dots,Y_{j_v,T}))| \\ <\ & 2Lip(h)\|k\|_\infty\|(X_{i_1,T},\dots,X_{i_u,T})-(Y_{i_1,T},\dots,Y_{i_u,T})\|_1 \\ & +2Lip(k)\|h\|_\infty\|(X_{j_1,T},\dots,X_{j_v,T})-(Y_{j_1,T},\dots,Y_{j_v,T})\|_1 \\ \leq\ & 2Lip(h)\|(R_{i_1,T},\dots,R_{i_u,T})\|_1+2Lip(k)\|(R_{j_1,T},\dots,R_{j_v,T})\|_1 \end{aligned}$$

due to $\|h\|_\infty \leq 1$ as well as $\|k\|_\infty \leq 1$.

For $t = 1,\dots,T$,

$$E|R_{t,T}| \ \leq \sum_{|j|\geq\lfloor \frac{m}{2} \rfloor} |a_{t,T}(j)|E|\varepsilon_{t-j}| \leq E|\varepsilon_0| \left(\sum_{|j|\geq\lfloor \frac{m}{2} \rfloor} |a_{t,T}(j)|\right) \leq C\left(\sum_{|j|\geq\lfloor \frac{m}{2} \rfloor} \frac{1}{l(j)}\right).$$

This result also holds for every $u \in [0,1]$:

$$\sup_{u\in[0,1]}\sup_{t\in B_u} E|R_{t,T}| \ \leq\ C\left(\sum_{|j|\geq\lfloor \frac{m}{2} \rfloor} \frac{1}{l(j)}\right).$$

Thus, $X_{t,T}$ is weak dependent with $\psi(g,h,u,v) = (u+v) \cdot \max\{Lip(h), Lip(k)\}$ and

$$\vartheta_m := E|\varepsilon_0| \cdot \left(\sum_{|j| \geq \lfloor \frac{m}{2} \rfloor} \frac{1}{l(j)} \right).$$

$\vartheta_m \to 0$ due to the the behaviour of $l(j)$ as given in Definition 2.1. □

Proof of Lemma 5.1. Let $|k-l| \geq 3m$. Then $|k-m+j-(l-m+i)| \geq m$, for $j, i = 0, \ldots, 2m$.
Consider the special case of $u = v = 2m$, $r \geq m$, $f_1(x) := \sum_{j=0}^{2m} x_{k-m+j} e^{-i\lambda_{mod(k)} j}$ and $f_2(x) := \sum_{j=0}^{2m} x_{l-m+j} e^{-i\lambda_{mod(l)} j}$. Both functions f_1 and f_2 have a Lipschitz modulus of 1, as

$$\frac{\left| \sum_{j=0}^{2m} (x_{k-m+j} - y_{k-m+j}) e^{-i\lambda_{mod(k)} j} \right|}{\sum_{j=0}^{2m} |x_{k-m+j} - y_{k-m+j}|} \leq \frac{\sum_{j=0}^{2m} |x_{k-m+j} - y_{k-m+j}|}{\sum_{j=0}^{2m} |x_{k-m+j} - y_{k-m+j}|} = 1.$$

We begin with

$$\begin{aligned}
&\sup_{|k-l| \geq 3m} \sup_{k,l} |\operatorname{Cov}(c_k, c_l)| \\
= &\sup_{|k-l| \geq 3m} \sup_{k,l} \left| \operatorname{Cov}\left(\frac{1}{\sqrt{2m+1}} \sum_{j=0}^{2m} X_{k-m+j,T} e^{-i\lambda_{mod(k)} j}, \frac{1}{\sqrt{2m+1}} \sum_{j=0}^{2m} X_{l-m+j,T} e^{-i\lambda_{mod(l)} j} \right) \right| \\
= &\sup_{|k-l| \geq 3m} \sup_{k,l} \frac{1}{2m+1} \left| \operatorname{Cov}\left(\sum_{j=0}^{2m} X_{k-m+j,T} e^{-i\lambda_{mod(k)} j}, \sum_{j=0}^{2m} X_{l-m+j,T} e^{-i\lambda_{mod(l)} j} \right) \right| \\
= &\sup_{|k-l| \geq 3m} \sup_{k,l} \frac{1}{2m+1} |\operatorname{Cov}(f_1(X_{k-m,T}, \ldots, X_{k+m,T}), f_2(X_{l-m,T}, \ldots, X_{l+m,T}))| \\
\leq &\sup_{|k-l| \geq 3m} \sup_{k,l} \frac{1}{2m+1} \psi(f_1, f_2, 2m, 2m) \cdot \vartheta_{2m},
\end{aligned}$$

$E|\varepsilon_0| \leq C$ (cf. Definition 2.1). Furthermore, $Lip(f_1) = Lip(f_2) = 1$, which then results in

$$\sup_{|k-l| \geq 3m} \sup_{k,l} \frac{1}{2m+1} \psi(f_1, f_2, 2m, 2m) \cdot \vartheta_{2m} \leq 2 \cdot K \cdot \frac{2m}{2m+1} \sum_{|j| \geq m} \frac{1}{l(j)}.$$

Hence, (5.20) follows. □

Remark 5.4
There is also an alternative, more obvious proof of Lemma 5.1:

Proof. First note that

$$\operatorname{Cov}(c_k, c_l) \quad = \quad \sum_{j_1,j_2=-\infty}^{\infty} a_{k,T}(j_1)a_{l,T}(j_2) \sum_{t_1,t_2=0}^{2m} E\left(\varepsilon_{k-j_1+t_1}\varepsilon_{l-j_2+t_2}\right) e^{-i\lambda_{mod(k)}t_1+i\lambda_{mod(l)}t_2},$$

with $E\left(\varepsilon_{k-j_1+t_1}\varepsilon_{l-j_2+t_2}\right) = \delta_{\{t_1=t_2+l-k+j_1-j_2\}}$. However, as t_1 is restricted to the range from 0 to $2m$, the expectancy can only be non-zero for $|j_1 - j_2| > m$, if $|k-l| \geq 3m$. Hence,

$$\sup_{|k-l|\geq 3m} \sup_{k,l} |\operatorname{Cov}(c_k, c_l)| \leq C \sup_{|k-l|\geq 3m} \sup_{k,l} \sum_{|j_1-j_2|>m} |a_{k,T}(j_1)| \cdot |a_{l,T}(j_2)| \to 0. \quad \square$$

The drawback of this simple proof, however, is that it can not easily extended to periodograms, as the arguments require a lineare structure.

The following Lemma is an extension of Theorem 5.3 (a) to moving periodograms.

Lemma 5.2. *For $u \in [0,1]$, $j \neq k = 1, \ldots, m$, let*

$$\begin{aligned}
\zeta_{\lfloor uT \rfloor, k} &:= div\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)\right\}}, \\
l' &:= l + \zeta_{\lfloor uT \rfloor, l} m \\
j' &:= j + \zeta_{\lfloor uT \rfloor, j} m.
\end{aligned}$$

Then,

$$\sup_{l \neq j = 1, \ldots, m} \left[\min\left(m, \max\left(|l-j|^2, \frac{m^2}{|l'-j'|^2}\right)\right) \left| E(MI^{\varepsilon}_{\lfloor uT \rfloor, m}(\lambda_j) MI^{\varepsilon}_{\lfloor uT \rfloor, m}(\lambda_l)) - 1 \right| \right] = O(1);$$

$$\sup_{j=1,\ldots,m} Var\, MI^{\varepsilon}_{\lfloor uT \rfloor, m}(\lambda_j) = 2 + O\left(\frac{1}{m}\right).$$

Proof.

$$\begin{aligned}
& E(MI^{\varepsilon}_{\lfloor uT \rfloor, m}(\lambda_l) MI^{\varepsilon}_{\lfloor uT \rfloor, m}(\lambda_j)) \\
= \;& \frac{1}{(2m+1)^2} \sum_{t_1, t_2, t_3, t_4 = 0}^{2m} E\big(\varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t_1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t_2} e^{-i(t_1 - t_2)\lambda_l} \\
& \varepsilon_{j+\zeta_{\lfloor uT \rfloor, j} m - m + t_3} \varepsilon_{j+\zeta_{\lfloor uT \rfloor, j} m - m + t_4} e^{-i(t_4 - t_3)\lambda_j}\big) \\
= \;& \frac{1}{(2m+1)^2} \sum_{t_1, t_2, t_3, t_4 = 0}^{2m} e^{-i(t_1 - t_2)\lambda_l} e^{-i(t_4 - t_3)\lambda_j} \\
& \cdot E\big(\varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t_1} \varepsilon_{l' - m + t_2} \varepsilon_{j' - m + t_3} \varepsilon_{j+\zeta_{\lfloor uT \rfloor, j} m - m + t_4}\big).
\end{aligned}$$

We now have

$$\frac{1}{(2m+1)^2} \sum_{t_1, t_2 = l' - m'}^{l'+m} \sum_{t_3, t_4 = j' - m}^{j'+m} E(\varepsilon_{t_1} \varepsilon_{t_2} \varepsilon_{t_3} \varepsilon_{t_4}) e^{-i(t_1 - t_2)\lambda_l} e^{i(t_3 - t_4)\lambda_j}$$

For further calculations we need the following case differentiation with respect to the indices $t_1, \ldots, t_4$.

- $\boxed{t_1 = t_2 = t_3 = t_4}$

  In this case, we can get the upper bound

$$\frac{1}{2m+1} E(\varepsilon_1^4) + \frac{\Sigma(l,j) E(\varepsilon_1^4)}{(2m+1)^2} \tag{5.22}$$

  The exact value value of $\Sigma(l,j)$ is:

$$\Sigma(l,j) := \min\{2m + l', 2m + j'\} - \max\{l', j'\} + 1 = 2m - |l' - j'| + 1.$$

- $\boxed{t_1 = t_2 \neq t_3 = t_4}$
  yields

$$\sigma^2 - \frac{E\varepsilon_1^4}{(2m+1)^2} \sum_{t_1=t_2=t_3=t_4} 1 \quad = \quad 1 - \frac{\Sigma(l,j)E(\varepsilon_1^4)}{(2m+1)^2}.$$

- $\boxed{t_1 = t_3 \neq t_2 = t_4}$

W.l.o.g. $l' > j'$

$$\begin{aligned}
& \frac{1}{(2m+1)^2} \sum_{t_1,t_2=j'-m}^{j'+m} \sum_{t_3,t_4=l'-m}^{l'+m} E(\varepsilon_{t_1}\varepsilon_{t_2}\varepsilon_{t_3}\varepsilon_{t_4}) e^{-i(t_1-t_2)\lambda_l} e^{i(t_3-t_4)\lambda_j} \mathbb{1}_{\{t_1=t_3\neq t_2=t_4\}} \\
= \quad & \frac{1}{(2m+1)^2} \sum_{t_1,t_2=l'}^{2m+j'} e^{-i(t_1-t_2)\lambda_l} e^{i(t_1-t_2)\lambda_j} - \frac{E\varepsilon_1^4}{(2m+1)^2} \sum_{t_1=t_2=t_3=t_4} 1 \\
= \quad & \frac{1}{(2m+1)^2} \left| \sum_{t_1=l'}^{2m+j'} e^{-it_1\lambda_{l-j}} \right|^2 - \frac{\Sigma(l,j)E(\varepsilon_1^4)}{(2m+1)^2}.
\end{aligned}$$

Considering the term

$$\frac{1}{(2m+1)^2} \left| \sum_{t_1=l'}^{2m+j'} e^{-it_1\lambda_{l-j}} \right|^2,$$

which equals 1 for $l = j$, we write

$$\begin{aligned}
\sum_{t_1=l'}^{2m+j'} e^{-it_1\lambda_{l-j}} \quad = \quad & \sum_{t_1=0}^{2m+j'-l'} e^{-i(t_1+l')\lambda_{l-j}} + \sum_{t_1=2m+j'-l'+1}^{2m} e^{-i(t_1+l')\lambda_{l-j}} \\
& - \sum_{t_1=2m+j'-l'+1}^{2m} e^{-i(t_1+l')\lambda_{l-j}} \\
= \quad & - \sum_{t_1=2m+j'-l'+1}^{2m} e^{-i(t_1+l')\lambda_{l-j}} \\
= \quad & - \sum_{t_1=1}^{l'-j'} e^{-i(t_1+2m+l')\lambda_{l-j}}.
\end{aligned}$$

From this, we define

$$T := T(l', j', l, j) \quad := \quad \sum_{t_1=1}^{l'-j'} e^{-i(t_1+2m+j')\lambda_{l-j}},$$

to which we apply Lemma A.4 in Kirch [27]. This yields

$$T = O\left(\min\left(\frac{2m+1}{|l-j|}, |l'-j'|\right)\right).$$

Analogously for the sine term. Hence,

$$\begin{aligned}\frac{1}{(2m+1)^2}|T|^2 &\leq C\min\left(\frac{1}{|l-j|^2}, \frac{|l'-j'|^2}{m^2}, \frac{|l'-j'|}{m|l-j|}\right) + \frac{1}{m}\\ &= C\min\left(\frac{1}{|l-j|^2}, \frac{|l'-j'|^2}{m^2}\right) + \frac{1}{m}.\end{aligned} \tag{5.23}$$

- $\boxed{t_1 = t_4 \neq t_2 = t_3}$ This case is analogous to the case of $t_1 = t_3 \neq t_2 = t_4$ and yields the same result, which includes $1 + O\left(\frac{1}{m}\right)$ for $l = j$. So note that all in all

$$\sup_{j=1,\ldots,m} \operatorname{Var} MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j) = 2 + O\left(\frac{1}{m}\right).$$

Continuing for $l \neq j$ we can write

$$\operatorname{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)\right) = \Sigma(l,j)\cdot\frac{E(\varepsilon_1^4)-3}{(2m+1)^2} + O\left(\min\left(\frac{1}{|l-j|^2}, \frac{|l'-j'|^2}{m^2}\right)\right). \tag{5.24}$$

All in all,

$$\begin{aligned}|E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)) - 1| &\leq C_1\min\left(\frac{1}{|l-j|^2}, \frac{|l'-j'|^2}{m^2}\right) + C_2\frac{1}{m}.\\ &\leq C_3\max\left(\min\left(\frac{1}{|l-j|^2}, \frac{|l'-j'|^2}{m^2}\right), \frac{1}{m}\right).\end{aligned}$$

□

.

Lemma 5.3. *Let a_m be a sequence with $a_m \to \infty$, $\frac{a_m}{m} \to 0$ and $\mathcal{A}_1(a_m, u)$ as in (5.5).*
Then,

(a)

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} \left|E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)) - 1\right| = O\left(\frac{1}{m} + \frac{1}{a_m^2} + \frac{a_m^2}{m^2}\right)$$

(b)

$$\sup_{l\neq j=1,\dots,m} E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)) \;=\; O(1).$$

Proof. The set $\mathcal{A}_1(a_m,u)$ is defined as

$$\mathcal{A}_1(a_m,u) := \{(l,j)\in\{1,\dots,m\}^2|(j\neq l)\wedge[|l'-j'|\leq a_m \vee |l-j|\geq a_m]\}.$$ □

(a) In any of the two cases $|l'-j'|\leq a_m$ or $|l-j|\geq a_m$, the result follows from Lemma 5.2.

(b) Follows immediately from Lemma 5.3.

We introduce the following notation. Let

$$x\succ y :\Leftrightarrow \frac{x}{y}=o(1),$$
$$x\succcurlyeq y :\Leftrightarrow \frac{x}{y}=O(1).$$

Remark 5.5
If additionally to the assumptions of Lemma 5.3 $a_m \preccurlyeq \sqrt{m}$,

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} \left|E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j))-1\right| = O\left(\frac{1}{a_m^2}\right).$$

Corollary 5.1. *In the situation of Lemma 5.2 with $a_m\to\infty$ and $a_m\preccurlyeq\sqrt{m}$,*

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} Cov(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)) \;=\; O\left(\frac{1}{a_m^2}\right).$$

Proof. Under the assumptions made, Lemma 5.3 states that

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} \left|E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j))-1\right| \;=\; O\left(\frac{1}{a_m^2}\right),$$

As $\sup_{u,l} E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l))=1$, see Proposition 4.1, the result follows. □

Lemma 5.4.

$$\frac{1}{2m+1}\sum_{j,l=0,\ldots,2m}\left|Cov(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j))\right| = O(1).$$

Proof. We split the set of indices and then apply Corollary 5.1:

$$\begin{aligned}
&\frac{1}{2m+1}\sum_{j\neq l=0}^{2m}\left|E(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j))-1\right| \\
=\;& \frac{2}{2m+1}\left(\sum_{(l,j)\in\mathcal{A}_1(a_m,u)}\left|\mathrm{Cov}(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j))\right|\right. \\
&\left.+\sum_{(l,j)\notin\mathcal{A}_1(a_m,u)}\left|\mathrm{Cov}(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j))\right|\right) \\
=\;& O\left(\frac{m}{a_m^2}+\frac{a_m^2}{m}\right).
\end{aligned}$$

With $a_m = \sqrt{m}$ the result follows. Note that the case $l = j$ can be included into the sum. ☐

Lemma 5.5. *Let* $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|} < \infty$. *Then,*

$$\frac{1}{2m+1}\sum_{j,k=0}^{2m}\left| Cov(MI_{\lfloor uT\rfloor,m}(\lambda_j), MI_{\lfloor uT\rfloor,m}(\lambda_k)) - |A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(\lambda_j)|^2 \right.$$
$$\left.\cdot|A_{k+\zeta_{\lfloor uT\rfloor,k}m-m,T}(\lambda_k)|^2\cdot Cov\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_k)\right)\right| = o(1).$$

Proof. Theorem 4.3 enables us to express the moving periodogram of a locally stationary time series with the help of the moving periodogram of iid random variables plus some remainder with vanishing second moment. The properties of the remainder are formally stated in Equations (4.47) and (4.48).

We may therefore substitute

$$\frac{1}{2m+1}\sum_{j,k=0}^{2m}\Big|\operatorname{Cov}(MI_{\lfloor uT\rfloor,m}(\lambda_j),MI_{\lfloor uT\rfloor,m}(\lambda_k))-|A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(\lambda_j)|^2$$
$$\cdot|A_{k+\zeta_{\lfloor uT\rfloor,k}m-m,T}(\lambda_k)|^2\operatorname{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j),MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_k)\right)\Big|$$
$$\leq \frac{1}{2m+1}\sum_{j,k=0}^{2m}|A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(\lambda_j)|^2\left|\operatorname{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j),R'_{\lfloor uT\rfloor,m}(\lambda_k)\right)\right|$$
$$+\frac{1}{2m+1}\sum_{j,k=0}^{2m}|A_{k+\zeta_{\lfloor uT\rfloor,k}m-m,T}(\lambda_k)|^2\left|\operatorname{Cov}\left(R'_{\lfloor uT\rfloor,m}(\lambda_j),MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_k)\right)\right|$$
$$+\frac{1}{2m+1}\sum_{j,k=0}^{2m}\left|\operatorname{Cov}\left(R'_{\lfloor uT\rfloor,m}(\lambda_j),R'_{\lfloor uT\rfloor,m}(\lambda_k)\right)\right|,$$

with $A_{t,T}(\lambda_l):=\sum_{j=-\infty}^{\infty}a_{t,T}(j)e^{-i\lambda_l j}$.

According to Proposition 4.2 $\sup_{j=1,\dots,m}|A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(\lambda_j)|^2$ is bounded. Moreover, note that $ER^{(2)}_{\lfloor uT\rfloor,m}(\lambda_l)=0$ (4.36) and consequently $ER^{'}_{\lfloor uT\rfloor,m}(\lambda_l)=0$.

$$\frac{1}{2m+1}\sum_{j,k=0}^{2m}|A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(\lambda_j)|^2\left|E\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)R'_{\lfloor uT\rfloor,m}(\lambda_k)\right)\right|$$
$$\leq E\left(\frac{1}{\sqrt{2m+1}}\sum_{j=0}^{2m}|A_{j+\zeta_{\lfloor uT\rfloor,j}m-m,T}(\lambda_j)|^2\left|MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right|\right.$$
$$\left.\cdot\frac{1}{\sqrt{2m+1}}\sum_{k=0}^{2m}\left|R'_{\lfloor uT\rfloor,m}(\lambda_k)\right|\right)$$
$$\leq CE\left(\frac{1}{\sqrt{2m+1}}\sum_{j=0}^{2m}\left|MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right|\frac{1}{\sqrt{2m+1}}\sum_{k=0}^{2m}\left|R'_{\lfloor uT\rfloor,m}(\lambda_k)\right|\right)$$
$$\leq C\sqrt{\frac{1}{2m+1}E\left(\sum_{j=0}^{2m}\left|MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right|\right)^2}\cdot\sqrt{\frac{1}{2m+1}E\left(\sum_{k=0}^{2m}\left|R'_{\lfloor uT\rfloor,m}(\lambda_k)\right|\right)^2}$$

with the Cauchy-Schwarz inequality. Similarly,

$$\frac{1}{2m+1}\sum_{j,k=0}^{2m}\left|E\left(R'_{\lfloor uT\rfloor,m}(\lambda_j)R'_{\lfloor uT\rfloor,m}(\lambda_k)\right)\right| \leq E\left(\frac{1}{\sqrt{2m+1}}\sum_{k=0}^{2m}\left|R'_{\lfloor uT\rfloor,m}(\lambda_k)\right|\right)^2.$$

We know from Lemma 5.4 that

$$\begin{aligned}
\operatorname{Var}\left(\frac{1}{\sqrt{2m+1}}\sum_{j=0}^{2m}\left|MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right|\right) &= \operatorname{Var}\left(\frac{1}{\sqrt{2m+1}}\sum_{j=0}^{2m}MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right)\\
&= \frac{1}{2m+1}\sum_{j,k=0}^{2m}\operatorname{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_k)\right)\\
&= O(1).
\end{aligned}$$

With Proposition 4.1,

$$E\left(\frac{1}{\sqrt{2m+1}}\sum_{j=0}^{2m}\left|MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right|\right)^2 = O(1).$$

And with Theorem 4.3, Equation (4.48), that is

$$E\left(\frac{1}{\sqrt{2m+1}}\sum_{k=0}^{2m}\left|R'_{\lfloor uT\rfloor,m}(\lambda_k)\right|\right)^2 = o(1),$$

the result follows. □

5.3 Spectral means with moving periodograms: asymptotic characteristics

We define the spectral mean MT using the local moving periodogram:

$$MT(u) := \frac{1}{2m+1}\sum_{j=0}^{2m}\varphi(\lambda_j)|MF_{\lfloor uT\rfloor}(\lambda_j)|^2 \tag{5.25}$$

$$T(u) := \frac{1}{2m+1}\sum_{j=0}^{2m}\varphi(\lambda_j)|\mathcal{F}(X_{\lfloor uT\rfloor-m+1,T},\ldots,X_{\lfloor uT\rfloor+m,T};\lambda_j)|^2 \tag{5.26}$$

$T(u)$ denotes the spectral mean statistic as used by Sergides [49], employing the local periodogram. Here, φ is chosen as in Sergides [49], Assumption 4. φ is a complex-valued bounded function. Moreover, it is periodically extended to $\mathbb{R}$ with period 2π and has a bounded second derivative. For $2m \geq j > m$, we require $\varphi(\lambda_j) = \varphi(\lambda_{2m-j})$.

Proposition 5.2

$$\frac{1}{m}\sum_{l,j=0,l>j}^{2m}\min\left(\frac{1}{|l-j|^2},\frac{\left|l-j+\mathbb{1}_{\{l\geq mod\left(\lfloor uT\rfloor-\lfloor\frac{m}{2}\rfloor\right)>j\}}m\right|^2}{m^2}\right)=O\left(m^{-\frac{1}{3}}\right).$$

Proof.

$$\begin{aligned}
&\frac{1}{m}\sum_{l>j}\min\left(\frac{1}{|l-j|^2},\frac{\left|l-j+\mathbb{1}_{\{l\geq mod\left(\lfloor uT\rfloor-\lfloor\frac{m}{2}\rfloor\right)>j\}}m\right|^2}{m^2}\right)\\
\leq\ &\frac{1}{m}\sum_{l>j}\min\left(\frac{1}{|l-j|^2},\frac{|l-j|^2}{m^2}\right)+\frac{1}{m}\sum_{l\geq j+m^{\frac{1}{3}}}\frac{1}{|l-j|^2}\\
&+\frac{C}{m}\sum_{j=mod\left(\lfloor uT\rfloor-\lfloor\frac{m}{2}\rfloor\right)-m^{\frac{1}{3}}}^{mod\left(\lfloor uT\rfloor-\lfloor\frac{m}{2}\rfloor\right)-1}\ \sum_{l=mod\left(\lfloor uT\rfloor-\lfloor\frac{m}{2}\rfloor\right)}^{j+m^{\frac{1}{3}}}1
\end{aligned}$$

The splitting is chosen according to the different possible values of the indicator function: The first case is assuming that $\mathbb{1}_{\{l\geq mod\left(\lfloor uT\rfloor-\lfloor\frac{m}{2}\rfloor\right)>j\}}=0$, the second one assumes that the indicator function equals 1, but $l>j+m^{\frac{1}{3}}$, and the third case assumes the indictor being 1 and all values of l and j not included in the second case.

With Kreiss and Neuhaus [30], Equation (A.11),

$$\begin{aligned}
&\frac{1}{m}\sum_{l>j}\min\left(\frac{1}{|l-j|^2},\frac{|l-j|^2}{m^2}\right)=\frac{1}{m}\sum_{1\leq|h|\leq O(m)}(2m-|h|)\min\left(\frac{1}{h^2},\frac{h^2}{m^2}\right)\\
\leq\ &O(1)\left|\sum_{|h|\leq\sqrt{m}}\frac{h^2}{m^2}+\sum_{|h|>\sqrt{m}}\frac{1}{h^2}\right|=O\left(\frac{1}{\sqrt{m}}\right).
\end{aligned}$$

Analogously,

$$\frac{1}{m}\sum_{l\geq j+m^{\frac{1}{3}}}\frac{1}{|l-j|^2}=\frac{1}{m}\sum_{O(m)\geq|h|\geq m^{\frac{1}{3}}}\frac{2m-|h|}{h^2}=\sum_{|h|\geq m^{\frac{1}{3}}}\frac{1}{h^2}=O\left(m^{-\frac{1}{3}}\right).$$

And finally,

$$\begin{aligned}
&\frac{1}{m} \sum_{j=mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)-m^{\frac{1}{3}}}^{mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)-1} \ \sum_{l=mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)}^{j+m^{\frac{1}{3}}} 1 \\
=& \frac{1}{m} \sum_{j=mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)-m^{\frac{1}{3}}}^{mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)-1} \left(j + m^{\frac{1}{3}} - mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right) + 1\right) \\
=& \frac{1}{m} \sum_{k=1}^{m^{\frac{1}{3}}} k = O\left(\frac{m^{\frac{2}{3}}}{m}\right) = O\left(m^{-\frac{1}{3}}\right). \qquad \square
\end{aligned}$$

Theorem 5.5. *With MT as in (5.25), for every $u \in [0,1]$ it holds that*

$$\begin{aligned}
& Var\left(\sqrt{2m+1}\,[MT(u) - E(MT(u))]\right) = \\
=& \frac{8\pi^2}{2m+1} \sum_{j=0}^{2m} |\varphi(\lambda_j)|^2 f^2(u,\lambda_j) + (E(\varepsilon_1^4) - 3)\left(\frac{8\pi^2}{(2m+1)^3} \cdot \right. \\
& \left. \cdot \sum_{j \neq l = 0}^{2m} \varphi(\lambda_j)\overline{\varphi(\lambda_l)} f(u,\lambda_j) f(u,\lambda_l) |\Sigma(l,j)| \right) + o(1).
\end{aligned}$$

with

$$\begin{aligned}
\Sigma(l,j) &:= 2m - |l' - j'| + 1, \\
\zeta_{\lfloor uT \rfloor,k} &:= div\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right) - \mathbb{1}_{\left\{l \geq mod\left(\lfloor uT \rfloor - \left\lfloor \frac{m}{2} \right\rfloor\right)\right\}}, \\
l' &:= l + \zeta_{\lfloor uT \rfloor,l}\, m, \\
j' &:= j + \zeta_{\lfloor uT \rfloor,j}\, m.
\end{aligned}$$

Remark 5.6
$\Sigma(l,j)$ is bounded from above by $2m+1$. In the case of local periodograms, as seen in Sergides [49], it equals $2m+1$, as we do not do any shifting.

Proof.

With Lemma 5.5, Definition 2.2 and φ bounded

$$\begin{aligned}
& \frac{1}{2m+1}\sum_{j\neq l}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}\,\mathrm{Cov}(MI_{\lfloor uT\rfloor,m}(\lambda_j),MI_{\lfloor uT\rfloor,m}(\lambda_l)) \\
= & \frac{4\pi^2}{2m+1}\sum_{j\neq l}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}f(u,\lambda_j)f(u,\lambda_l)\,\mathrm{Cov}(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j),MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l))+o(1).
\end{aligned}$$

From this,

$$\begin{aligned}
& \mathrm{Var}\left(\sqrt{2m+1}\,[MT(u)-E(MT(u))]\right) \\
= & \frac{1}{2m+1}\sum_{j=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_j)}\,\mathrm{Var}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j)\right) \\
& +\frac{1}{2m+1}\sum_{j\neq l=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}\,\mathrm{Cov}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j),MI_{\lfloor uT\rfloor,m}(\lambda_l)\right) \\
= & \frac{1}{2m+1}\sum_{j=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_j)}\,\mathrm{Var}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j)\right) \\
& +\frac{2}{2m+1}\sum_{j\neq l}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}\,\mathrm{Cov}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j),MI_{\lfloor uT\rfloor,m}(\lambda_l)\right) \\
= & \frac{1}{2m+1}\sum_{j=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_j)}\,\mathrm{Var}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j)\right) \\
& +\frac{8\pi^2}{2m+1}\sum_{j\neq l}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}f(u,\lambda_j)f(u,\lambda_l)\,\mathrm{Cov}(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j),MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)) \\
& +o(1) \\
=: & A_1+A_2+o(1),
\end{aligned}$$

$\boxed{A_1}$

By slightly modifying Theorem 4.3, we get with Propositions 4.1 and 4.2

$$MI_{\lfloor uT\rfloor,m}(\lambda_j) \;=\; |A\,(u,\lambda_j)|^2\,MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)+R''_{\lfloor uT\rfloor,}(\lambda_j).$$

$R''_{\lfloor uT\rfloor,m}(\lambda_j)$ fulfills (4.46) and, under the additional assumption of

$$\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty}|a(u,j)|\sqrt{|j|}<\infty,$$

(4.47) holds.

With the above modification of Theorem 4.3

$$\begin{aligned}\operatorname{Var}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j)\right) &= \left|A\left(u,\lambda_j\right)\right|^4 \operatorname{Var}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right) + \operatorname{Var}\left(R'_{\lfloor uT\rfloor}(\lambda_j)\right) \\ &\quad +2\left|A\left(u,\lambda_j\right)\right|^2 \operatorname{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), R'_{\lfloor uT\rfloor}(\lambda_j)\right).\end{aligned}$$

Now,

$$\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} \operatorname{Var}\left(R'_{\lfloor uT\rfloor}(\lambda_j)\right) \leq \sup_{u\in[0,1]}\sup_{l=1,\ldots,m} E|R'_{\lfloor uT\rfloor}(\lambda_j)|^2 = o(1)$$

Under the assumption that $\sup_{u\in[0,1]}\sum_{j=-\infty}^{\infty} a(u,j)\sqrt{|j|} < \infty$, we get $O\left(\frac{1}{m}\right)$.

Hence, with Cauchy-Schwarz and Propositions 4.1 and 4.2, we also have that

$$\sup_{u\in[0,1]}\sup_{l=1,\ldots,m} \left|A\left(u,\lambda_j\right)\right|^2 \operatorname{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), R'_{\lfloor uT\rfloor}(\lambda_j)\right) = o(1).$$

Again, with the additional assumption, this yields a rate of $O\left(\frac{1}{\sqrt{m}}\right)$.
From Lemma 5.2 we have that

$$\sup_{j=1,\ldots,m} \operatorname{Var}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)\right) = 2 + O\left(\frac{1}{m}\right).$$

Thus,

$$\frac{1}{2m+1}\sum_{j=0}^{2m} \varphi(\lambda_j)\overline{\varphi(\lambda_j)} \operatorname{Var}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j)\right) = \frac{2}{2m+1}\sum_{j=0}^{2m} |\varphi(\lambda_j)|^2 \left|A\left(u,\lambda_j\right)\right|^4 + o(1).$$

With the definition of the time-varying spectral density, $f(u,\lambda)$, Definition 2.2,

$$\operatorname{Var}(MI_{\lfloor uT\rfloor,m}(\lambda_j)) = 8\pi^2 f^2(u,\lambda_j) + o(1). \tag{5.27}$$

With the additional assumption we get

$$\operatorname{Var}(MI_{\lfloor uT\rfloor,m}(\lambda_j)) = 8\pi^2 f^2(u,\lambda_j) + O\left(\frac{1}{\sqrt{m}}\right). \tag{5.28}$$

And, hence,

$$\begin{aligned}A_1 &= \frac{1}{2m+1}\sum_{j=0}^{2m} \varphi(\lambda_j)\overline{\varphi(\lambda_j)} \operatorname{Var}\left(MI_{\lfloor uT\rfloor,m}(\lambda_j)\right) \\ &= \frac{8\pi^2}{2m+1}\sum_{j=0}^{2m} |\varphi(\lambda_j)|^2 f^2(u,\lambda_j) + o(1).\end{aligned}$$

$\boxed{A_2}$

From the proof in Lemma 5.2, we obtain Equation (5.24), that is

$$\mathrm{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)\right) = \Sigma(l,j)\cdot\frac{E(\varepsilon_1^4)-3}{(2m+1)^2} + O\left(\min\left(\frac{1}{|l-j|^2}, \frac{|l'-j'|^2}{m^2}\right)\right).$$

Hence, by Proposition 5.2,

$$\begin{aligned} A_2 &= \frac{8\pi^2}{2m+1}\sum_{j\neq l=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}f(u,\lambda_j)f(u,\lambda_l)\,\mathrm{Cov}\left(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l)\right) \\ &= (E(\varepsilon_1^4)-3)\left(\frac{1}{(2m+1)^3}\sum_{j\neq l=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}f(u,\lambda_j)f(u,\lambda_l)\Sigma(l,j)\right) + O\left(m^{-\frac{1}{3}}\right). \end{aligned}$$

and therefore the result follows. □

CHAPTER 6

Alternative Fourier transformations

In Chapter 5 we have looked at the covariance of local moving coefficients. In the easiest case, the coefficients referred to and were based on a sequence of i.i.d. random variables ε_t. We will restrict ourselves in the following to this case in order to understandably convey our point. The results can, of course, analogously to the procedure in Chapter 4 be extended to the case of stationary, as well as locally stationary time series.
We look at the statement made by Theorem 5.3:
In the situation of Lemma 5.2 with $a_m \to \infty$ and $a_m/\sqrt{m} \to 0$,

$$\sup_{(l,j)\in\mathcal{A}_1(a_m,u)} \operatorname{Cov}(MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_l), MI^{\varepsilon}_{\lfloor uT\rfloor,m}(\lambda_j)) = O\left(\frac{1}{a_m^2}\right).$$

Here, one might ask oneself if the restriction to the set $\mathcal{A}_1(a_m,u)$ introduced to maintain the correct covariance structure is indeed necessary or whether it can be circumvented by slightly changing the transform in some way. For the definition of the set $\mathcal{A}_1(a_m,u)$ see Equation (5.5).

Note that in the following sections, we will refer to the coefficients

$$MF^{\varepsilon}_k(\lambda_l) := \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \varepsilon_{l+\left(div\left(k-\lfloor \frac{m}{2}\rfloor\right)-\mathbb{1}_{\{l\geq mod(k-\lfloor \frac{m}{2}\rfloor)\}}\right)m-m+t,T}\, e^{-it\lambda_l},$$

(Definition 3.1) as the original local moving Fourier coefficients in contrast to the below-mentioned alterations $MF^{\varepsilon,(1)}_k(\lambda_l)$, $MF^{\varepsilon,(2)}_k(\lambda_l)$, and $MF^{\varepsilon,(3)}_k(\lambda_l)$.

Further, we define two alternative operators to *mod* and *div* for $j \in \mathbb{Z}$:

$$MOD(j) := \begin{cases} 2m+1, & \text{if } 2m+1 \text{ is a factor of } j \in \mathbb{Z}, \\ j \mod (2m+1), & j > 0 \ \wedge \ (2m+1) \mid j, \\ 2m+1-[(-j) \mod (2m+1)], & j < 0 \ \wedge \ (2m+1) \nmid j. \end{cases} \tag{6.1}$$

$$DIV(j) := \left\lceil \frac{j}{2m+1} \right\rceil. \tag{6.2}$$

Then, $\quad j = MOD(j) + (DIV(j) - 1)(2m+1)$.

Those operators arise from the definition of *mod* and *div* by substituting m by $2m+1$.

6.1 Local moving Fourier transform (Alt 1) - adding some rearranging

Motivated by the procedure chosen for the transformation from frequency to time domain (Definition 3.4), one might also want to consider the rearranging of the input data in this case. We have already done this in the frequency domain where it seems very intuitive, as one would want to link the Fourier coefficient c_l (referring to frequency λ_l to the exponential function $e^{i\lambda_l t}$. Simply speaking, one would like to have

$$\varepsilon_t = \sum_{l=0}^{2m} c_l e^{i\lambda_l t},$$

just like we do for the ordinary global Fourier transform. Note, however, that we do move through the sequence $c_1, c_2, \ldots, c_T$ when performing the inverse moving Fourier transform. Hence, we would, in the next step, link c_l with $e^{-i\lambda_{l-1} t}$, as the set of Fourier coefficients we started with would be not $c_1, \ldots, c_m$, but $c_2, \ldots, c_{m+1}$. Now c_2 would be multiplied by $e^{i\lambda_1 \cdot}$, which does not make any sense, as it corresponds to frequency λ_1. Hence, in the frequency domain the reason why we do the rearranging, which is thoroughly described in Chapter 3, is obvious.
We now asked ourselves the question whether we may either get rid of or reduce the cardinality of the set $\mathcal{A}_1(a_m, u)$ by doing the same rearranging in the time domain, ensuring that some random variable ε_t will always, when occuring in any selected stretch of data be linked with $e^{-i\lambda_l t}$ (incorporating the same index t and not incorporating just some index used for summing up the $2m+1$ elements). Without rearranging, the random variable ε_{16}, for example, which is used in $2m+1$ transformations, is always linked to a different value of the index of summation. By rearranging, we ensure that the random variable ε_{16} is always in position 16 of our set of length $2m+1$. That's the idea so far, now comes the theory.

What happens formally, when adding rearranging to the original local moving Fourier transform? Starting out with the time series $\varepsilon_0, \ldots, \varepsilon_T$, we intend to always link ε_0 to the first position in the set, ε_1 to the second position, up to ε_{2m}, which is linked to the position $2m+1$. ε_{2m+1} shall then be again linked to the first position in the sample and so on. We can see that this way the corresponding positions result from using the actual indices $MOD(2m+1)$.
Now consider the exemplary set of

$$\varepsilon_4, \varepsilon_5, \ldots, \varepsilon_{2m+1}, \varepsilon_{2m+2}, \varepsilon_{2m+3}, \varepsilon_{2m+4}.$$

The first step is to find out, what the inherent position of the last element is: $MOD(2m+4)$. This equals 3. Now, the set is rearranged to

$$\varepsilon_{2m+2}, \varepsilon_{2m+3}, \varepsilon_{2m+4}, \varepsilon_4, \varepsilon_5, \ldots, \varepsilon_{2m+1},$$

placing ε_{2m+4} in position 3. The number $MOD(2m+4)$ will be referred to as the splice of the stretch of random variables considered.
Assuming we want to calculate the Fourier coefficient at frequency $1 \leq l \leq m$, the splice in the sequence of indices can be written as

$$\xi_l := MOD(l + \zeta_{\lfloor uT \rfloor, l} m + m)$$

That is

$$\begin{aligned} MF^{\varepsilon,(1)}_{\lfloor uT \rfloor}(\lambda_l) \quad := \quad & \frac{1}{\sqrt{2m+1}} \sum_{t=1}^{\xi_l} \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + 2m + t + 1} e^{-i\lambda_l t} \\ & + \frac{1}{\sqrt{2m+1}} \sum_{t=\xi_l+1}^{2m+1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t} e^{-i\lambda_l t}. \end{aligned}$$

Now, with similar arguments as in the proof of Theorem 5.2,

$$\begin{aligned} & E(MF^{\varepsilon,(1)}_{\lfloor uT \rfloor}(\lambda_l) MF^{\varepsilon,(1)}_{\lfloor uT \rfloor}(\lambda_j)) \\ = \quad & E\left(\left(\frac{1}{\sqrt{2m+1}} \sum_{t=1}^{\xi_l} \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + 2m + t} e^{-i\lambda_l t} + \frac{1}{\sqrt{2m+1}} \sum_{t=\xi_l+1}^{2m+1} \varepsilon_{l+\zeta_{\lfloor uT \rfloor, l} m - m + t} e^{-i\lambda_l t}\right)\right. \\ & \left.\cdot \left(\frac{1}{\sqrt{2m+1}} \sum_{t=1}^{\xi_j} \varepsilon_{j+\zeta_{\lfloor uT \rfloor, j} m - m + 2m + t} e^{i\lambda_j t} + \frac{1}{\sqrt{2m+1}} \sum_{t=\xi_j+1}^{2m+1} \varepsilon_{j+\zeta_{\lfloor uT \rfloor, j} m - m + t} e^{i\lambda_j t}\right)\right) \\ = \quad & \frac{O(1)}{m} \left(\sum_{t_1=1}^{\min(\xi_l, \xi_j)} e^{-i(\lambda_l - \lambda_j) t_1} + \sum_{t_2=\max(\xi_l, \xi_j)+1}^{2m+1} e^{-i(\lambda_l - \lambda_j) t_2} \right) \\ = \quad & \frac{O(1)}{m} \left(- \sum_{t=\min(\xi_l, \xi_j)+1}^{\max(\xi_l, \xi_j)} e^{-i(\lambda_l - \lambda_j) t} \right) = O\left(\min\left(\frac{|\xi_l - \xi_j|}{m}, \frac{1}{l-j}\right)\right). \end{aligned}$$

Here, we can see that for $|l-j|$ small, that is for Fourier frequencies λ_l and λ_j close to each other, we cannot guarantee $|\xi_l - \xi_j|$ to be of an order less than m. An example would be $l + \zeta_{\lfloor uT \rfloor, l} m = m+3$ and $j + \zeta_{\lfloor uT \rfloor, j} m = 2$, $l = 2, j = 3$.

6.2 Local moving Fourier transform (Alt 2) - formally circumventing the stumbling blocks

To tackle the question, whether the introduction of the set $\mathcal{A}_1(a_m, u)$ can be circumvented, one might also want to look critically at the fact that we artificially create some kind of break in the sequence of Fourier coefficients. What we are currently doing is as follows: We calculate the moving Fourier coefficients by shifting along the time series and for each stretch of data we calculate the Fourier transform for one single frequency λ_l and then move on to the next stretch, with $l = 1, \ldots, m$. Having reached λ_m, we start anew with $l = 1$.
The following alternative transformation differs from our original one by the fact that we generate Fourier coefficients corresponding to the whole set of Fourier frequencies $\{\lambda_1, \ldots, \lambda_{2m+1}\}$ before continuing with λ_1. In the original procedure we generated coefficients for frequencies $\lambda_1, \ldots, \lambda_m$ and then started again with λ_1. To adapt the original method, we need to use the two operators MOD and DIV.

The crucial point, however, is that for the procedure to work, we need to get rid of all Fourier coefficients belonging to frequencies $\lambda_{m+1}, \ldots, \lambda_{2m+1}$. To illustrate the concept: Having generated $c_1, \ldots, c_T$, we throw away $c_{m+1}, \ldots, c_{2m+1}$, as well as $c_{3m+2}, \ldots, c_{4m+2}$ and so on, that is in the end we have a gapped stretch of coefficients, because every m times we have thrown away a stretch of length m. If we didn't, the same problem as in Chapter 5 and the resulting need for the set $\mathcal{A}_1(a_m, u)$ would emerge.
Hence,

$$MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_l) := \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \varepsilon_{l + \left[DIV(\lfloor uT \rfloor - m) - \mathbb{1}_{\{l \geq MOD(\lfloor uT \rfloor - m)\}}\right](2m+1) - (2m+1) + t} e^{-i\lambda_l t_1},$$

with $l = 1, \ldots, m$. Define

$$\tilde{\zeta}_{\lfloor uT \rfloor, l} := \left[DIV\left(\lfloor uT \rfloor - m\right) - \mathbb{1}_{\{l \geq MOD(\lfloor uT \rfloor - m)\}}\right].$$

We now propose that

$$\sup_{l \neq j = 1, \ldots, m} E(MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_l) MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_j)) = o(1).$$

The formal proof that this works can be seen in the following:

$$\begin{aligned}
&E(MF^{\varepsilon,(2)}_{\lfloor uT\rfloor}(\lambda_l)MF^{\varepsilon,(2)}_{\lfloor uT\rfloor}(\lambda_j)) \\
&= \frac{1}{2m+1}\sum_{t_1,t_2=0}^{2m} E(\varepsilon_{l+\tilde{\zeta}_{\lfloor uT\rfloor,l}(2m+1)-(2m+1)+t_1}\varepsilon_{j+\tilde{\zeta}_{\lfloor uT\rfloor,j}(2m+1)-(2m+1)+t_2})e^{-i\lambda_l t_1}e^{i\lambda_j t_2}.
\end{aligned}$$

This equals not zero only if $l+\tilde{\zeta}_{\lfloor uT\rfloor,l}(2m+1)-t_1 = j+\tilde{\zeta}_{\lfloor uT\rfloor,j}(2m+1)+t_2$, that is $l'+t_1 = j'+t_2$.
Note, that $MOD(l') = l$ and $MOD(j') = j$.

We thus choose to substitute t_2 by $(l'-j')+t_1$ and, hence, have to correct the range of t_1 to

$$g_1(u,l,j) = g_1 := \max\{0,(j'-l')\},\ldots,\min\{2m,2m-(l'-j')\} =: g_2 = g_2(u,l,j).$$

Let w.l.o.g. $j' > l'$.

$$\begin{aligned}
&E(MF^{\varepsilon,(2)}_{\lfloor uT\rfloor}(\lambda_l)MF^{\varepsilon,(2)}_{\lfloor uT\rfloor}(\lambda_j)) \\
&= \frac{1}{2m+1}e^{-i\lambda_j g_1}\left[\sum_{t_1=g_1}^{2m} e^{i\lambda_{l-j}t_1}\right] \\
&= \frac{1}{2m+1}e^{-i\lambda_j g_1}\left[\sum_{t_1=g_1}^{2m}\left(\cos\left(\frac{2\pi(l-j)}{2m+1}t_1\right) - i\sin\left(\frac{2\pi(l-j)}{2m+1}t_1\right)\right)\right].
\end{aligned}$$

Application of Lemma A.4 in Kirch [27] yields that

$$\sum_{t_1=0}^{2m-j'+l'}\cos\left(\frac{2\pi(l-j)}{2m+1}t_1\right) = O\left(\min\left(\frac{2m+1}{|l-j|},|2m+1-j'+l'|\right)\right).$$

Analogously for the sine term.

$$\begin{aligned}
\sum_{t_1=0}^{2m-j'+l'} e^{i\lambda_{l-j}t_1} &= \sum_{t_1=0}^{2m} e^{i\lambda_{l-j}t_1} - \sum_{t_1=2m-j'+l'+1}^{2m} e^{i\lambda_{l-j}t_1} \\
&= \sum_{t_1=0}^{j'-l'-1} e^{-i\lambda_{l-j}(t_1+2m-j'+l'+1)} = e^{-i\lambda_{l-j}(l'-j')}\sum_{t_1=0}^{j'-l'-1} e^{i\lambda_{l-j}t_1}.
\end{aligned}$$

For this sum, we get, again by Lemma A.4 in Kirch [27],

$$\sum_{t_1=0}^{j'-l'-1} e^{i\lambda_{l-j}t_1} = O\left(\min\left(\frac{2m+1}{|l-j|},|l'-j'|\right)\right).$$

Putting the two results together, one gets

$$E(MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_l)MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_j)) = O\left(\min\left(\frac{1}{|l-j|}, \frac{|l'-j'|}{2m+1}, \frac{|2m+1-j'+l'|}{2m+1}\right)\right)$$

We know that $|l-j|$ is between 0 and $m-1$. The difference $\tilde{\zeta}_{\lfloor uT \rfloor,j} - \tilde{\zeta}_{\lfloor uT \rfloor,l}$ is either ± 1 or 0, depending on the position of l and j with regard to $MOD(\lfloor uT \rfloor - m)$. We have assumed that $j' > l'$. Hence, the case that the difference equals -1 can not occur.

Now consider the case $\tilde{\zeta}_{\lfloor uT \rfloor,j} - \tilde{\zeta}_{\lfloor uT \rfloor,l} = 0$. In this case, $j' - l' = j - l$. Hence,

$$E(MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_l)MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_j)) = O\left(\min\left(\frac{1}{|l-j|}, \frac{|l-j|}{2m+1}\right)\right) = O\left(\frac{1}{\sqrt{m}}\right).$$

On the other hand, if $\tilde{\zeta}_{\lfloor uT \rfloor,j} - \tilde{\zeta}_{\lfloor uT \rfloor,l} = 1$, $j' - l' = j - l + 2m + 1$. In this case, again,

$$E(MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_l)MF^{\varepsilon,(2)}_{\lfloor uT \rfloor}(\lambda_j)) = O\left(\min\left(\frac{1}{|l-j|}, \frac{|l-j|}{2m+1}\right)\right) = O\left(\frac{1}{\sqrt{m}}\right).$$

Thus, the phenomenon which occurred in Chapter 3 can not occur here. And we do not need to make any exceptions to values of j and l.

The reason why we do not get a problem here, is that from a stretch of T real random variables we have created $T/2$ complex random variables, that is T real random variables. In our original transformation, we have used T real random variables to create T complex variables, that is $2T$ real random variables. By doing so, we certainly have to pay a price and this price is this dependence coming in – dependence of coefficients which belong to frequencies with index not in $\mathcal{A}_1(a_m, u)$. This price, however, is not too high to pay as the cardinality of the set of indices not in $\mathcal{A}_1(a_m, u)$ is of an order less than m, which makes it negotiable when speaking of spectral means, ratios etc. Using the first amendment to the list of alternative transformations introduced above, we, however, have to pay the price of actually wasting information on the time series or, putting it in other words, being too slow with collecting information. Which is not important for stationary time series, but very well important for locally stationary time series. In reality, structural changes can happen quite fast, and if one had the choice between a method which uses a stretch of data double the size for the same information avoiding a negligible additional dependency, one would most certainly go for the information which is denser in time.

We have finally extracted the conceptual problem of why we get this restriction to $\mathcal{A}_1(a_m, u)$: By applying the original local moving Fourier transform we gain double the information which is present. Which ought to cost something.

The altered local moving Fourier transform has been the first try to construct a transform in a way that we get a 1-1 relation between the information contained in the time domain and the information contained in the frequency domain.

A further approach is taken in the following section.

6.3 Local moving Fourier transform (Alt 3) - customized to fit the needs

A further amendment to the list of transformations is to use the original transform with the adaption that for the first m times we obtain not the complex Fourier coefficient, but only the real part. The second m times we generate the corresponding imaginary parts. Then the m real parts and m imaginary parts are stuck together as m complex Fourier coefficients to frequencies $\lambda_1, \dots, \lambda_m$. We then continue with the transformation of our time series. The following m stretches serve as data to obtain real parts and the corresponding imaginary parts are generated by the next but one set of m transforms. All in all, having moved through the time series we end up with $T/2$ real parts and $T/2$ imaginary parts, that is $T/2$ moving Fourier coefficients. One could think of this method as a more adapted method than Alt 2, better capturing the aspect of locally changing time series as the information at all times is incorporated in the sample and hence, the change is mirrored more closely. Still, of course, this method suffers from the same flaw as Alt 2, to obtain m Fourier coefficients, we need a number of observations in the time domain which would have generated $2m$ coefficients using the original method.
For $l = 1, \dots, m$, define

$$MF^{\varepsilon,(3)}_{\lfloor uT \rfloor}(\lambda_l) := \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \varepsilon_{l+\left[DIV(\lfloor uT \rfloor - m) - \mathbb{1}_{\{l \geq MOD(\lfloor uT \rfloor - m)\}}\right](2m+1)-(2m+1)+t} \cos(\lambda_l t_1).$$

For $l = m+1, \dots, 2m+1$, define

$$MF^{\varepsilon,(3)}_{\lfloor uT \rfloor}(\lambda_l) := \frac{1}{\sqrt{2m+1}} \sum_{t=0}^{2m} \varepsilon_{l+\left[DIV(\lfloor uT \rfloor - m) - \mathbb{1}_{\{l \geq MOD(\lfloor uT \rfloor - m)\}}\right](2m+1)-(2m+1)+t} \sin(-\lambda_l t_1)$$

Now, for $l = 1, \dots, m$ and $j = m+1, \dots, 2m+1$

$$\begin{aligned} & E(MF^{\varepsilon,(3)}_{\lfloor uT \rfloor}(\lambda_l), MF^{\varepsilon,(3)}_{\lfloor uT \rfloor}(\lambda_j)) \\ = \; & \frac{1}{2m+1} \sum_{t_1,t_2=0}^{2m} E(\varepsilon_{l+\tilde{\zeta}_{\lfloor uT \rfloor,l}(2m+1)-(2m+1)+t_1} \varepsilon_{j+\tilde{\zeta}_{\lfloor uT \rfloor,j}(2m+1)-(2m+1)+t_2}) \\ & \cdot \cos\left(\frac{2\pi l t_1}{2m+1}\right) \sin\left(-\frac{2\pi j t_2}{2m+1}\right). \end{aligned}$$

This equals not zero only if $l + \tilde{\zeta}_{\lfloor uT \rfloor,l}(2m+1) + t_1 = j + \tilde{\zeta}_{\lfloor uT \rfloor,j}(2m+1) + t_2$, that is $l' + t_1 = j' + t_2$. Note, that $MOD(l') = l$ and $MOD(j') = j$.
We thus choose to substitute t_2 by $(l' - j') + t_1$ and, hence, have to correct the range

of t_1 to

$$g_1(u,l,j) = g_1 := \max\{0,(j'-l')\},\dots,\min\{2m, 2m-(l'-j')\} =: g_2 = g_2(u,l,j).$$

Let w.l.o.g. $j' > l'$.

$$\begin{aligned} & E(MF^{\varepsilon,(3)}_{\lfloor uT \rfloor}(\lambda_l), MF^{\varepsilon,(3)}_{\lfloor uT \rfloor}(\lambda_j)) \\ = & \frac{1}{2m+1} \sum_{t_1=(j'-l')}^{2m} \cos\left(\frac{2\pi l t_1}{2m+1}\right) \sin\left(-\frac{2\pi j(t_1+l'-j')}{2m+1}\right). \end{aligned}$$

Simulations have lead to the conclusion that the above sum converges to 0 for all possible combinations of j and l.

6.4 Summary

Concerning Alt 1, we can say: Rearranging unifies the procedures applied for back and forth transform, but does not have any effect on the set $\mathcal{A}_1(a_m, u)$. Looking closer at why the need for this exception arises, one mathematically finds out that if $|j'-l'| \to \infty$ always implied $|l-j| \to \infty$, we would be done. This has been achieved by developing Alt 2. Looking closer at this transformation we have now been able to detect the kernel of the brute. The need for set $\mathcal{A}_1(a_m, u)$ was due to the fact that we overindulged in information. We had m random variables at hand, and created out of them $2m$ random variables. These $2m$ random variables can not possibly all be uncorrelated, each carrying different information, as this information could not all have been stored in the m variables we started with. Getting this bonus of double the random variables with our transform, we need to pay the price of some of them not being uncorrelated. As long as this 'some' is of less than order m, though, all is well. Both, Alt 2 and Alt 3 suffer from the problem of needing to use wider stretches of input data, whereas the local moving Fourier transform uses stretches half as wide, resulting in a more local procedure. Alt 2 and 3 both, of course, get rid of the set $\mathcal{A}_1(a_m, u)$, with Alt 3 definitely being superior to Alt 2. Alt 2 grabs some information, waits some time without getting information, then again grabs another piece of information and so on. By doing so, Alt 2 will miss out on the gradual change in information. Alt 3, however, meets the criterion of constantly updating its information while moving through time, as does the local moving Fourier transform.

Remark 6.1
As we have noticed, rearranging the data stretches does not effect the second order structure of the Fourier coefficients. This leaves room for the conjecture that there is no change of the distributional properties of Fourier coefficients in the stationary case when using shifted data. That is that $\mathcal{F}(X_1, X_2, \dots, X_n; \lambda_1)$ is as far as

distributional characteristics are concerned, equal to $\mathcal{F}(X_n, X_1, \ldots, X_{n-1}; \lambda_1)$.

CHAPTER 7

Application of the moving Fourier transformation

A locally stationary process $\{X_{t,T}\}$ describes a time series with slow changes. In Chapter 3 we have developed a method to transfer the changing information contained in the time series to the frequency domain using the moving Fourier transformation (Definition 3.2). Further, we have also found a way to convert these coefficients back to some time series with the same structural characteristics as the original one.
Now, seeing that the local moving Fourier coefficients at time t (Definition 3.1), which are basically a set of $2m+1$ specially created Fourier coefficients assigned to some time t, as well as the corresponding periodogram ordinates exhibit an asymptotically decreasing covariance, one is reminded of the ordinary Fourier coefficients which are asymptotically iid.
Taking up this discovery that local moving Fourier coefficients asymptotically behave similarly to Fourier coefficients, we extend bootstrap methods in the frequency domain from the stationary to the locally stationary setting.

7.1 Bootstrap methods in the frequency domain

7.1.1 Wild bootstrap

→ *Kirch and Politis [28]*
We apply the standard bootstrap method of wild bootstrap as described in Kirch and Politis [28] to the moving Fourier coefficients (Definition 3.2). In order to perform the wild bootstrap, we need an estimator of the time varying spectral density (Definition 2.2), meeting the requirement

$$\max_{k\in\{\lfloor uT\rfloor-\lceil\frac{m}{2}\rceil+1,\dots,\lfloor uT\rfloor+\lfloor\frac{m}{2}\rfloor\}}\left|\hat{f}\left(\frac{k}{T},\lambda_{mod(k)}\right)-f\left(u,\lambda_{mod(k)}\right)\right|=o_P(1).$$

See Chapter 8 for existence and construction of such an estimator.
We now proceed as follows
Step 1:
Split each c_k into real and imaginary part $c_k := x_k + iy_k$.
Step 2:
Let $G_k, G_{k+T},\ k = 1, \dots, T$, be independent identically standard normal random variables. Generate the bootstrap samples c_k^* according to

$$\begin{aligned} x_k^* &:= \sqrt{\pi \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right)} G_k, \\ y_k^* &:= \sqrt{\pi \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right)} G_{k+T}, \\ c_k^* &:= x_k^* + iy_k^*. \end{aligned}$$

7.1.2 Residual based bootstrap

→ *Kirch and Politis [28]*

The initial requirement is just like in the case of the wild bootstrap: In order to be able to apply the standard bootstrap method of residual based bootstrap as described in Kirch and Politis [28] to the moving Fourier coefficients (Definition 3.2) we need an estimator of the time varying spectral density (Definition 2.2) with

$$\max_{k \in \{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil + 1, \dots, \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor\}} \left| \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right) - f\left(u, \lambda_{mod(k)}\right) \right| = o_P(1).$$

Step 1:
Split each c_k into real and imaginary part $c_k := x_k + iy_k$.

Step 2:
Estimate residuals of real and imaginary part and put them in a vector $\{\tilde{s}_j\}_{1 \le j \le 2T}$

$$\tilde{s}_k := \frac{x_k}{\sqrt{\pi \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right)}}, \quad \tilde{s}_{T+k} := \frac{y_k}{\sqrt{\pi \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right)}}.$$

Step 3:
Standardization yields

$$s_k := \frac{\tilde{s}_k - \frac{1}{T}\sum_{l=1}^{2T} \tilde{s}_l}{\frac{1}{2T}\sum_{t=1}^{2T}\left(\tilde{s}_t - \frac{1}{2T}\sum_{l=1}^{2T} \tilde{s}_l\right)^2}.$$

Step 4:
Ordinary iid resampling with replacement in order to get $s_1^*, \ldots, s_{2T}^*$.

Step 5:
Define bootstrap Fourier coefficients

$$\begin{aligned} x_k^* &:= \sqrt{\pi \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right)} s_k^*, \\ y_k^* &:= \sqrt{\pi \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right)} s_{T+k}^*, \\ c_k^* &:= x_k^* + i y_k^*. \end{aligned}$$

7.1.3 Local bootstrap

→ *Kirch and Politis [28], Paparoditis and Politis [44]*

Step 1:
Select a symmetric, nonnegative kernel $K(\cdot)$ with $\int K(t)dt = 1$. Special assumptions on the kernel K are made in Chapter 8.
Moreover, one needs to select a bandwidth h, fulfilling $h \to 0$, but $mh \to \infty$.
Step 2:
Define iid random variables $J_{1,T}, \ldots, J_{2T,T}$ on $\mathbb{Z}$, with

$$p_{s,T} = P(J_{j,T} = s) = \frac{K(2\pi s/((2m+1)h))}{\sum_{l=-\infty}^{\infty} K(2\pi l/((2m+1)h))}$$

Independent of these, define $2T$ iid $Bern(1/2)$-distributed random variables $B_1, \ldots, B_{2T}$.
Step 3:
The bootstrap is performed as follows:

$$\begin{aligned} \tilde{x}_k^* &:= \begin{cases} x_{k+J_{k,T}}, & \text{if } B_k = 0, \\ y_{k+J_{k,T}}, & \text{if } B_k = 1, \end{cases} \\ \tilde{y}_k^* &:= \begin{cases} y_{k+J_{T+k,T}}, & \text{if } B_{T+k} = 0, \\ x_{k+J_{T+k,T}}, & \text{if } B_{T+k} = 1. \end{cases} \end{aligned}$$

This construction exploits the fact that for a smooth spectral density, the distribution of the moving Fourier coefficients in a small environment should approximately be the same.

The final bootstrap coefficients are then obtained after centering with the weighted mean of the original series and are thus given by

$$\begin{aligned} x_k^* &:= \tilde{x}_k^* - \frac{1}{2}\sum_{s\in\mathbb{Z}} p_{s,T}(x_{k+s}+y_{k+s}), \\ y_k^* &:= \tilde{y}_k^* - \frac{1}{2}\sum_{s\in\mathbb{Z}} p_{s,T}(x_{k+s}+y_{k+s}), \\ c_k^* &:= x_k^* + iy_k^*. \end{aligned}$$

7.2 Bootstrapping time domain data

7.2.1 Moving TFT-bootstrap

As soon as one is provided with some suitable method of transforming back and forth from time to frequency domain the most natural thought when intending to use bootstrapping is to do so in the frequency domain as it allows for iid bootstrap methods. For the first time, this has been done by Kirch and Politis [28] using the ordinary Fourier transformation of length T. With the new method of the moving Fourier transformation at hand, we can now extend the concept to locally stationary processes, perform local iid bootstrap methods in the frequency domain and return to the time domain.
Thus, the moving TFT-bootstrap can essentially be viewed as a three step procedure.

Step 1:
The observed time series is transformed using the so called moving Fourier transform (3.11):

$$c_k = \mathcal{F}(X_{k-m,T},\ldots,X_{k+m,T};\lambda_{mod(k)}) = \frac{1}{\sqrt{2m+1}}\sum_{l=k-m}^{k+m} X_{l,T}e^{-il\lambda_{mod(k)}}e^{i(k-m)\lambda_{mod(k)}},$$

with $\lambda_{mod(k)} := \frac{2\pi\, mod(k)}{2m+1}$ denoting the Fourier frequencies and the operator *mod* according to (3.1).
We now face the T moving Fourier coefficients $c_1,\ldots,c_T$.

Step 2:
In a second step, the resulting moving Fourier coefficients are bootstrapped by a localized standard method of choice, such as the wild, the local or the residual bootstrap. See Section 7.1. This results in

$$c_1^*,\ldots,c_T^*.$$

Step 3:
The moving bootstrap coefficients gained are then transformed back using a moving version of the inverse Fourier transform (3.15).

$$X^*_{t,T} = \frac{1}{\sqrt{2m+1}} \sum_{l=1}^{m} c^*_{l+\left(div\left(t-\lfloor\frac{m}{2}\rfloor\right)-\mathbb{1}_{\{l\geq mod(t-\lfloor\frac{m}{2}\rfloor)\}}\right)m} e^{i\lambda_l t} + \frac{1}{\sqrt{2m+1}} \sum_{l=1}^{m} \overline{c^*}_{l+\left(div\left(t-\lfloor\frac{m}{2}\rfloor\right)-\mathbb{1}_{\{l\geq mod(t-\lfloor\frac{m}{2}\rfloor)\}}\right)m} e^{-i\lambda_l t}$$

with $\lambda_k := \frac{2\pi k}{N}$, $k = 0, \dots, m$, denoting the Fourier frequencies and $t = 1, \dots, T$. This finally yields a bootstrap replicate $X^*_{1,T}, X^*_{2,T}, \dots, X^*_{T,T}$ of the original time series in the time domain.

Basically all bootstrap methods involving the frequency domain which are used for stationary time series can be adapted to the locally stationary situation using the moving Fourier transformation and the moving periodogram, as defined in Definition 3.3. There are also other ways of localizing bootstrap procedures using periodograms, for example the use of the local periodogram as done by Sergides [49] and Kreiss and Paparoditis [32]. We will now modify these two procedures using our moving periodogram and compare their performance.

7.2.2 Moving autoregressive-aided periodogram bootstrap

The local autoregressive-aided periodogram bootstrap by Sergides [49] combines a parametric bootstrap in the time domain with a local nonparametric correction in the frequency domain. It is an extension of the autoregressive-aided periodogram bootstrap by Kreiss and Paparoditis [31] to locally stationary time series and essentially works as follows: The part concerned with the parametric bootstrap is based on locally fitting an AR(p)-model to the data, calculating the residuals and generating bootstrap errors from the empirical distribution function of the residuals. The bootstrap observations then result from using the estimated AR(p)-coefficients and the bootstrap errors. Up to now, we just have, as Sergides [49] point out, a local version of the autoregressive bootstrap. In order to loosen the restriction of an underlying AR(p)-process, a nonparametric correction is added to the bootstrap AR(p)-periodogram. It serves the purpose to correct the bootstrap periodogram of the time varying AR(p)-process for structure of the data that can not be explained by some autoregressive model. The correction is a smoothed version of the local periodogram divided by an estimate of the local spectral AR(p)-density.
Using the autoregressive-aided periodogram bootstrap by Kreiss and Paparoditis [31] as a fundament, Sergides [49] has created a local bootstrap method. We provide a further adaption of the autoregressive-aided periodogram bootstrap of Kreiss and Paparoditis [31] to locally stationary time series. The parametric bootstrap

is mainly mimicked, except for the fact that we draw the bootstrap errors locally using iid resampling. To calculate the bootstrap periodogram, however, we use the moving periodogram.
The next section displays our algorithm for the moving autoregressive-aided periodogram bootstrap. The alterations to Sergides [49] are in the use of the moving periodogram, the remaining parts, however, are borrowed from Sergides [49]. For the sake of simplicity, we restrict ourselves w.l.o.g. to an AR(1) model.
We therefore assume
Let $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$ be a locally stationary time series as in Definition 2.1. In order to keep the algorithm as simple as possible, we will assume that there is a sufficient number of preceding and succeeding observations available. Namely, $X_{1-3m,T}, \ldots, X_{0,T}$ and $X_{T+1,T}, \ldots, X_{T+3m,T}$. If applied to a real set of data, we need to slightly adapt the procedure by settling for a blockwise approach in the beginning and in the end of the time series.

We also assume that (Sergides [49], Assumption 2.2) the stationary approximation of $X_{t,T}$ at time $u \in [0,1]$ has the AR(∞)-representation

$$\tilde{X}_t(u) = \sum_{k=1}^{\infty} \beta_k(u)\tilde{X}_{t-1}(u) + a(u,0)\varepsilon_t,$$

where $1 + \sum_{k=1}^{\infty} a(u,k)z^k = (1 - \sum_{k=1}^{\infty} \beta_k(u)z^k)^{-1}$, $\sum_{k=1}^{\infty} k|\beta_k(u)| < \infty$ and $1 - \sum_{k=1}^{\infty} \beta_k(u)z^k \neq 0$ for all complex z with $|z| \leq 1$.

Step 1: Local fit of $AR(1)$-model
For every point in time $1 \leq t \leq T$ we fit an autoregressive model of order 1 to the data $X_{t-m,T}, \ldots, X_{t+m,T}$ and calculate the estimated parameter $\hat{a}^{(t)} := \hat{\beta}_1\left(\frac{t}{T}\right)$. This leaves us with the estimated coefficients $\hat{a}^{(1)}$, $\hat{a}^{(2)}, \ldots, \hat{a}^{(T)}$ and the estimated standard deviations of the errors $\hat{\sigma}^{(1)}, \ldots, \hat{\sigma}^{(T)}$.
(The exact formula to calculate the standard deviation using Yule-Walker estimators on the stationary approximations is given by Sergides [49], Section 2.3, page 14.)

Step 2: Estimation of the centered and rescaled errors $\hat{\varepsilon}_{1,T}, \ldots, \hat{\varepsilon}_{T,T}$
Consider the rescaled residuals

$$\tilde{\varepsilon}_{t,T} := \frac{1}{\hat{\sigma}^{(t)}}\left(X_{t,T} - \hat{a}^{(t)}X_{t-1,T}\right), \qquad t = 2, \ldots, T.$$

These rescaled residuals are then centered by $\hat{\varepsilon}_{t,T} := \tilde{\varepsilon}_{t,T} - \frac{1}{T}\sum_{\tau=1}^{T} \tilde{\varepsilon}_{\tau,T}$, so we finally get $\hat{\varepsilon}_{1,T}, \ldots, \hat{\varepsilon}_{T,T}$.

Step 3: Generation of the bootstrap errors $\varepsilon^+_{1,T}, \ldots, \varepsilon^+_{T,T}$
For every $t \in \{1, \ldots, T\}$ consider the stretch $\hat{\varepsilon}_{t-m,T}, \ldots, \hat{\varepsilon}_{t+m,T}$ with equal probability assigned to each residual, and draw one residual. This sample is named $\varepsilon^+_{t,T}$.

Step 4: Generation of the bootstrap observations $X^+_{1,T}, \ldots, X^+_{T,T}$
Having created all bootstrap errors $\varepsilon^+_{1,T}, \ldots, \varepsilon^+_{T,T}$, we can now calculate the bootstrap observations by using the locally fitted $AR(1)$- models (cf. Step 1). We set $X^+_{1,T} := X_{1,T}$ and

$$X^+_{t,T} := \hat{a}^{(t)} X^+_{t-1,T} + \hat{\sigma}^{(t)} \varepsilon^+_{t,T},\ t = 2, \ldots, T.$$

Step 5: Calculation of moving periodogram.
Application of the moving Fourier transform (as in Definition 3.2) to the bootstrap observations $X^+_{1,T}, \ldots, X^+_{T,T}$ yields

$$c^+_1, c^+_2, \ldots, c^+_T.$$

Using the local moving coefficients as in Definition 3.1 at each time t results in T sets

$$MF^+_t(\lambda_1), \ldots, MF^+_t(\lambda_m)$$

The moving periodogram $MI_{t,m}(\lambda_j)$ is defined in Definition 3.3 and, thus, analogously

$$MI^+_{t,m}(\lambda_j) := \left| MF^+_t(\lambda_j) \right|^2.$$

Step 6: Local correction
Computation of the local kernel estimator. The assumptions on the kernel are given by (K)(i)-(v) in Chapter 8.

$$\hat{q}\left(\frac{t}{T}, \lambda\right) := \frac{1}{2m+1} \sum_{j=-m}^{2m} K_h(\lambda - \lambda_j) \frac{MI^+_{t,m}(\lambda_j)}{\hat{f}^{(t)}_{AR}\left(\frac{t}{T}, \lambda_j\right)},$$

where

$$\hat{f}^{(t)}_{AR}\left(\frac{t}{T}, \lambda_j\right) := \frac{(\hat{\sigma}^{(t)})^2}{2\pi} \cdot \frac{1}{|1 - \hat{a}^{(t)} e^{-i\lambda_j}|^2}.$$

Step 7: Construction of moving bootstrap periodogram
The moving bootstrap periodogram is then given by

$$MI^*_{t,m}(\lambda_j) := \hat{q}\left(\frac{t}{T}, \lambda_j\right) \cdot MI^+_{t,m}(\lambda_j),$$

$j = 1, \ldots, m$, and $t = 1, \ldots, T$.

7.2.3 Moving wild hybrid bootstrap

The hybrid bootstrap by Kreiss and Paparoditis [32] combines a wild bootstrap in the time domain with a nonparametric approach in the frequency domain. It is an extension of the wild hybrid bootstrap by Kreiss and Paparoditis [33] to locally stationary time series.

It uses two major ideas: Firstly, that the observations $X_{t,T}$ can approximately be written as

$$X_{t,T} \approx \frac{1}{T} \sum_{j=0}^{T-1} \sqrt{f\left(\frac{t}{T}, \lambda_j\right)} J_\varepsilon(\lambda_j) e^{it\lambda_j},$$

with $J_\varepsilon(\lambda_j) := \frac{1}{T} \sum_{l=0}^{T-1} \varepsilon_l e^{-il\lambda_j}$. This is heuristically deduced from the relation $|J_X(\lambda_j)|^2 \approx f(\lambda_j)|J_\varepsilon(\lambda_j)|^2$ in Brockwell and Davis [3], Theorem 10.3.1, and has, for a time-independent density, already been used by Kreiss and Paparoditis [33].

In the moving version this approximating expression is slightly changed to

$$X_{t,T} \approx \frac{1}{2m+1} \sum_{j=0}^{2m} \sqrt{f\left(\frac{t}{T}, \lambda_j\right)} \mathcal{F}_\varepsilon^{div(t)-1}(\lambda_j) e^{it\lambda_j},$$

incorporating the moving Fourier transform instead of the original Fourier transform of the errors, see also Remark 3.3.

The second nip is, as already done in Kreiss and Paparoditis [33], to estimate the fourth order cumulant of the innovations by using the relation

$$\begin{aligned} \mathrm{Cov}({X_t}^2(u), {X_{t+k}}^2(u)) &= \kappa_4 \sum_{j=-\infty}^{\infty} {\psi_j}^2(u) {\psi_{j+k}}^2(u) \\ &+ 2 \cdot \mathrm{Cov}^2(X_t(u), X_{t+k}(u)), \ u \in [0,1]. \end{aligned} \tag{7.1}$$

Here, $X_t(u) = \sum_{j=-\infty}^{\infty} \psi_j(u) \varepsilon_{t-j}$ is the stationary approximation of $X_{t,T}$ at time $\lfloor uT \rfloor$. Equation (7.1) then yields,

$$\kappa_4(u) := \frac{\sum_{k=-\infty}^{\infty} (c_2(u,k) - 2c^2(u,k))}{c^2(u,0)}, \ u \in [0,1],$$

with $c_2(u,k)$ being the autocovariance function of the squared stationary approximation $X_t^2(u)$ at time $\lfloor uT \rfloor$. Contrasting Kreiss and Paparoditis [32], we refrain from integrating over time in a next step in order to avoid evening out changes in the fourth order structure, but to be able to mimick them.

The next paragraph describes the bootstrap algorithm for the moving hybrid bootstrap. Note that just like for the previous procedure in Section 7.2.2, we will also assume that sufficient observations preceding time $t = 1$ and succeeding time $t = T$ are available in order to straighten out notation.

Step 1: Estimating the local fourth order cumulant $\hat{\kappa}_4\left(\frac{t}{T}\right)$ at time t
We follow Kreiss and Paparoditis [32] defining an estimator of the fourth order cumulant. However, we do allow κ_4 to locally vary and account for that variation by local estimation:

$$\hat{\kappa}_4\left(\frac{k}{T}\right) := \frac{\hat{G}_1\left(\frac{k}{T}\right) - \hat{G}_2\left(\frac{k}{T}\right)}{\hat{G}_3\left(\frac{k}{T}\right)}.$$

The functions $\hat{G}_1, \hat{G}_2$ and $\hat{G}_3$ are defined in the following.

$$\hat{G}_1\left(\frac{k}{T}\right) := \sum_{j=-m}^{2m} K_h(0-\lambda_j)\cdot \ MI_{(2),k,m}(\lambda_j),$$

where

$$MI_{(2),k,m}(\lambda_j) := \frac{1}{(2m+1)}\left|\sum_{l=0}^{2m}\left(X^2_{j+\zeta_{k,j}m-m+l,T} - \frac{1}{2m+1}\sum_{r=0}^{2m} X^2_{j+\zeta_{k,j}m-m+r,T}\right) e^{-i\lambda_j l}\right|^2$$

denotes the local moving periodogram of the squared and locally centered time series $X^2_{t,T}$. The smoothing kernel $K_h(\cdot)$ ought to fulfill the assumptions given by (K)(i)-(v) in Chapter 8.

$\hat{G}_1\left(\frac{k}{T}\right)$ is an estimator for (a multiple of) the spectral density of the squared time series $X^2_{t,T}$ at time $1 \leq k \leq T$ and frequency zero. The second estimator $\hat{G}_2\left(\frac{k}{T}\right)$ estimates the sum of the squared autocovariances of the stationary approximation of $X_{t,T}$ at time k:

$$\hat{G}_2\left(\frac{k}{T}\right) := \sum_{l=0}^{2m}\left(MI_{k,m}(\lambda_l)\right)^2.$$

And, at last, $\hat{G}_3\left(\frac{k}{T}\right)$ is an estimator for the squared autocovariance function of the stationary approximation of $X_{t,T}$ at time k and lag zero:

$$\hat{G}_3\left(\frac{k}{T}\right) := \left(\sum_{l=0}^{2m} MI_{k,m}(\lambda_l)\right)^2.$$

Having calculating $\hat{\kappa}_4\left(\frac{1}{T}\right), \ldots, \hat{\kappa}_4\left(\frac{T}{T}\right)$, we aim (see Kreiss and Paparoditis [32]) to get estimates for the fourth moment of the errors, by setting

$$\tilde{\kappa}_4^t := \hat{\kappa}_4\left(\frac{t}{T}\right) + 3,$$

and thus obtain local estimates $\tilde{\kappa}_4^1, \tilde{\kappa}_4^2, \ldots, \tilde{\kappa}_4^T$ of the fourth order moment structure of the errors ε_t.

Step 2:
Knowing about the first second and fourth order moment structure of the errors, one can now generate bootstrap residuals according to the following sampling rule: Generate a sample $\varepsilon_1^*, \ldots, \varepsilon_T^*$ of length T of iid random variables meeting

$$\begin{aligned} P(\varepsilon_t^* = \sqrt{\tilde{\kappa}_4^t}) &= P(\varepsilon_t^* = -\sqrt{\tilde{\kappa}_4^t}) = \frac{1}{2\tilde{\kappa}_4^t}, \\ P(\varepsilon_t^* = 0) &= 1 - \frac{1}{\tilde{\kappa}_4^t}, \end{aligned}$$

for $1 \leq t \leq T$.

Step 3:
Calculation of the moving Fourier transform of the bootstrap errors $\varepsilon_1^*, \ldots, \varepsilon_T^*$, resulting in

$$c_1^{\varepsilon^*}, c_2^{\varepsilon^*}, \ldots, c_T^{\varepsilon^*}$$

Step 4: The local moving Fourier coefficients at each time $t = 1, \ldots, T$, are then given by the T sets

$$MF_t^{\varepsilon^*}(\lambda_1), \ldots, MF_t^{\varepsilon^*}(\lambda_m)$$

Step 5: Generation of the bootstrap observations by

$$X_{t,T}^* := \frac{1}{\sqrt{2m+1}} \sum_{j=0}^{m} \sqrt{\hat{f}\left(\frac{t}{T}, \lambda_j\right)} \left(MF_t^{\varepsilon^*}(\lambda_j) e^{it\lambda_j} + \overline{MF_t^{\varepsilon^*}(\lambda_j)} e^{-it\lambda_j} \right),$$

where $\hat{f}$ is an estimator of the spectral density, fulfilling

$$\max_{k \in \{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil + 1, \ldots, \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor\}} \left| \hat{f}\left(\frac{k}{T}, \lambda_{mod(k)}\right) - f\left(u, \lambda_{mod(k)}\right) \right| = o_P(1).$$

See Chapter 8 for existence and construction of such an estimator.

CHAPTER 8

Moving spectral density estimation

8.1 Approaches to estimate the time-varying spectral density

There are basically two fundamental approaches in literature, how to estimate the time varying spectral density of a locally stationary process.

Localized periodogram

The first is calculating the classical periodogram only locally over a segment of length $N << T$. This estimator, called "localized periodogram", has been introduced by von Sachs and Schneider [54]. It uses a stretch of length N of tapered data with some midpoint $\lfloor uT \rfloor$ to obtain an estimate for the spectral density at this point in time $\lfloor uT \rfloor$.

$$I_N(u,\lambda) = \frac{1}{H_{2,N}} \left| \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) X_{\lfloor uT - \frac{N}{2} + s + 1 \rfloor} e^{-i2\pi\lambda s} \right|^2,$$

with $h : [0,1] \to [0,1]$ being a sufficiently smooth tapering function and $H_{2,N}$ being the appropriate norming factor as in Dahlhaus [8], Section 3.
This is what Sergides [49] basically calls the tapered local periodogram. For $h \equiv 1$ it is the local periodogram. He is not doing any tapering, though and uses another notation of the Fourier transformation. In order to get an estimator for the spectral density at all times, von Sachs and Schneider [54] calculate $I_N(u,\lambda)$ on possibly overlapping segments of $X_{t,T}$ of length N. Denote the shift from segment to segment by S, $1 \le S \le N$. The resulting number of segments is called M. Hence, $I_N(u,\lambda)$ is evaluated at M timepoints $u_i = \frac{t_i}{T}$, where $t_i = S \cdot i + \frac{N}{2}$, $0 \le i \le M-1$.

The drawback of this procedure is for one the computational cost, which is the cost of M times a Fourier transformation of length N, i.e. $O(NM\log(N))$. For another it is the additional parameter N. Dahlhaus and Neumann [13] nicely and understandably pose the problem, which is twofold: First, this parameter delivers a cut-off point, from which on covariances of higher lags than k are excluded from the estimation, which induces a bias in time domain, if N is small.
For the second aspect one needs to bear in mind the so-called uncertainty principle, which says (in the more general case of evolutionary spectra):

> UNCERTAINTY PRINCIPLE
> → Priestley [46], p. 217
>
> In determining evolutionary spectra, one cannot obtain simultaneously a high degree of resolution in both the time domain and the frequency domain.

Now, when using the estimator

$$\hat{f}(u,\lambda) = \frac{1}{b_f}\int K_f\left(\frac{\lambda-\mu}{b_f}\right) I_N(u,\mu)d\mu,$$

with K_f being a symmetric kernel with $\int K_f(x)dx = 1$ and b_f the bandwidth in frequency direction (cf. Dahlhaus [11]), there is already included some smoothing in the time domain, which is not obvious at first glance. That is, as part of the localization of the classical periodogram made by von Sachs and Scheider [54] was obtained by summation over certain time points in segments of chosen length N. Dahlhaus [11] provides in Equation (83) the exact kernel estimate in the time domain, which is implicitly contained and possesses a bandwidth of $b_t = \frac{N}{T}$. Thus, inherently a lower bound for the resolution in the time domain is fixed. This lower bound for the resolution in the time domain immediately results in an upper bound for the resolution in frequency domain, due to the uncertainty principle (cf. Neumann and von Sachs [40]).

Dahlhaus and Neumann [13] draw the following conclusion: Local periodograms therefore lack the possibility to control for the whole amount of smoothing explicitly – in an additional smoothing step. A possible remedy can be to control the smoothing in time domain purely by the choice of N and perform the second smoothing step for smoothing only in the frequency domain. Also, when using a higher degree of smoothing in the second step, for example, a kernel with a bandwidth $b_t >> \frac{N}{T}$, the use of the local periodogram is reasonable.

Preperiodogram

The second approach taken to estimate the time-dependent spectral density is the use of the so-called preperiodogram, which does not incorporate any implicit smoothing. The preperiodogram for a locally stationary time series $\{X_{t,T}\}$ at frequency $\lambda \in [0,\pi]$ has been introduced by Neumann and von Sachs [40]:

Definition 8.1 (Preperiodogram).
→ *Neumann and von Sachs [40], Equation (3.7)*

$$I_{t,T}(\lambda) = \frac{1}{2\pi} \sum_{s:1\le \lfloor t-\frac{s}{2}\rfloor, \lfloor t+\frac{s}{2}\rfloor \le T} X_{\lfloor t-\frac{s}{2}\rfloor,T} X_{\lfloor t+\frac{s}{2}\rfloor,T} e^{-i\lambda s}.$$

Neumann and von Sachs [40] point out that the preperiodogram can serve as a preliminary estimate of the spectral density, which is even more fluctuating than the classical periodogram. Asymptotically, its expected value equals the evolutionary spectrum (introduced by Priestley [47], see Section 1). For fixed length T, its expected value equals the Wigner-Ville spectrum (Martin and Flandrin [37], see Section 1).
There is a nice relation of the preperiodogram to the classical periodogram over the whole stretch of data, which is shown by Dahlhaus [9]. It eases the interpretation of the preperiodogram: The classical periodogram is the average of the preperiodogram over time, that is $I_T(\lambda) = \frac{1}{T}\sum_{t=1}^{T} I_{t,T}(\lambda)$. The preperiodogram uses only the product $X_{\lfloor t-\frac{k}{2}\rfloor,T} X_{\lfloor t+\frac{k}{2}\rfloor,T}$ to estimate the covariance at time t, while the periodogram is the Fourier transformation of the covariance estimator of lag k over the whole segment (see Neumann and von Sachs [40], Section 2.1).

We base our estimator on the first approach. However, instead of smoothing the local periodogram in frequency direction, we do so with our moving periodogram as in Definition 3.3. The difference to the local periodogram is explained in the subsequent Remark 3.6.
We look for an estimator for the time varying spectral density which is still close to the true spectral density at time t, even when we estimate at a time slightly earlier or later than t. In formulae: For every $u \in [0,1]$,

$$\sup_{k\in\{\lfloor uT\rfloor-\lceil\frac{m}{2}\rceil+1,\dots,\lfloor uT\rfloor+\lfloor\frac{m}{2}\rfloor\}} \left|\hat{f}\,(k) - f\left(u, \lambda_{mod(k)}\right)\right| = o_P(1). \tag{8.1}$$

Before we introduce our estimator, we state the assumptions made on the time varying spectral density.

8.2 Prerequisites

(F) Assumptions on the time varying spectral density

(i) f is uniformly Lipschitz continuous in both arguments.

(ii) f is uniformly bounded from above and below: $\exists c$, C with $0 < c \le |f(u,\lambda)| \le C$ for all $u \in [0,1]$, $\lambda \in [0, 2\pi]$.

Remark 8.1
The existence of an upper bound in (F)(ii) follows from Definition 2.1 and Definition 2.2, since

$$\begin{aligned} |f(u,\lambda)| &= \frac{1}{2\pi}\left|\sum_{j=-\infty}^{\infty} a(u,j)e^{-i\lambda j}\right|^2 \leq \frac{1}{2\pi}\sum_{j,k=-\infty}^{\infty} |a(u,j)|\,|a(u,k)| \\ &\leq C\sum_{j=-\infty}^{\infty}\frac{1}{l(j)}\sum_{k=-\infty}^{\infty}\frac{1}{l(k)} \leq C. \end{aligned}$$

Definition 8.2 (Uniformly Lipschitz continuous).
→ Haug [24] Definition 2.7
A function $g: D \subset \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is uniformly Lipschitz continuous of order α in both components (with Lipschitz constants M_1 and M_2), if for all $u, v \in D$

$$|g(u,\mu) - g(v,\mu)| \leq M_1|u-v|^\alpha \quad \forall\, \mu \in \mathbb{R},$$

and for all $\lambda, \mu \in \mathbb{R}$

$$|g(u,\lambda) - g(u,\mu)| \leq M_2|\lambda-\mu|^\alpha \quad \forall\, u \in D.$$

Based on Sergides [49], we use a local kernel density estimator to estimate f. Nevertheless, modifications are needed to adapt to our way of locally Fourier transforming a time series. The kernel K ought to be chosen according to the following criteria

(K) Assumptions on the kernel

(i) K is a nonnegative, symmetric function with compact support.

(ii) $\int K(x)dx = 1,\ |K(x)| \leq const.$,

$$\frac{2\pi}{(2m+1)h}\sum_{j\in\mathbb{Z}} K\left(\frac{2\pi j}{(2m+1)h}\right) = \int K(x)dx + o(1) = 1 + o(1).$$

(iii) K is uniformly Lipschitz continuous.

(iv) $h \to 0$ $(T \to \infty)$ and $hm^{\frac{1}{4}} \to \infty$.

(v) $|K_h(x)| = O\left(\frac{1}{h}\right)$, with $K_h(\cdot) := \frac{1}{h}K\left(\frac{\cdot}{h}\right)$.

8.3 Definition of the estimator: The smoothed moving periodogram

Definition 3.3 provides the concept of the moving periodogram and it is (see Remark 3.6) compared to the concept of the local periodogram as used by Sergides [49]. The new terminology introduced in the chapter's heading also underlines the difference in concept and intended use of the moving periodogram: moving spectral density estimation.
We do not intend to locally estimate the spectral density at one point and then do the same estimation again and again in neighbouring points in time like Sergides [49]: He uses the local periodogram on window of width N to define a local spectral density estimator by

$$\hat{f}(u, \lambda) := \frac{1}{N} \sum_{j=-\lfloor \frac{N}{2} \rfloor}^{\lfloor \frac{N}{2} \rfloor} K_h(\lambda - \lambda_j) I_{N,X}(u, \lambda_j). \tag{8.2}$$

The local (unscaled and untapered) periodogram is given by

$$I_{N,X}(u, \lambda_j) = \frac{1}{N} \left| \sum_{t=0}^{2m} X_{\lfloor uT \rfloor - m + t} e^{-i\lambda_j t} \right|. \tag{8.3}$$

Note, that Sergides' original results are all obtained for the local periodogram rescaled by the factor $\frac{1}{2\pi}$. When referring to his results, however, we will always refer to the unscaled version (8.3).

Contrasting the definition of the local estimator of the spectral density and the local periodogram, we define the smoothed moving periodogram, which is effectively just a function of one argument – of time.

Definition 8.3 (Smoothed moving periodogram).

Consider a locally stationary process $X_{t,T}$ according to Definition 2.1 and a function K fulfilling (K)(i) – (K)(v).
The smoothed moving periodogram $\hat{f} : \{1, \ldots, T\} \to \mathbb{R}$ is then defined by

$$\hat{f}(k) := \frac{1}{m} \sum_{t=-m}^{2m} K_h \left(\lambda_{mod(k)} - \lambda_t \right) MI_{k,m}(\lambda_t), \tag{8.4}$$

with $MI_{k,m}(\lambda_t)$ being the local moving periodogram as in Definition 3.3.

8.4 Locally uniform consistency of the estimator

The following Theorem mathematically formalizes the aim indicated in (8.1).

Theorem 8.1 (Locally uniform convergence).
Let $X_{t,T}$ be a locally stationary time series as in Definition 2.1 with time varying spectral density f meeting (F)(i) and (ii). Further assume that (K)(i)-(v) hold. Then, for every $u \in [0,1]$, the estimator $\hat{f}$ as in Equation (8.4) fulfills condition (8.1), that is

$$\sup_{k \in \{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil + 1, \ldots, \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor\}} \left| \hat{f}\,(k) - f\left(u, \lambda_{mod(k)}\right) \right| = o_P(1).$$

Preliminary work

We have seen in the definition of $\hat{f}$ that the moving spectral density estimator is a function of only one variable k which tells us to use frequency $\lambda_{mod(k)}$ at time k. We have not yet discussed, though, how to specify the point in time we need to consider when intending to estimate the moving spectral density at a certain frequency ω. This problem is addressed in the following:

For the sake of simplicity, define $B_u := \left\{ \lfloor uT \rfloor - \lceil \frac{m}{2} \rceil + 1, \ldots, \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor \right\}$. With

$$s'(\omega) := \min \left\{ l \in \{1, \ldots, m\} \middle| \lambda_l - \frac{\pi}{2m+1} < \omega \leq \lambda_l + \frac{\pi}{2m+1} \right\} \tag{8.5}$$

the frequencies $\lambda_{s'(\omega)}, \lambda_{s'(\vartheta)}$ are the Fourier frequencies closest (in absolute value) to $0 < \omega,\, \vartheta < \pi$ (cf. Brockwell and Davis [3], Definition 10.3.1).

Note that as pointed out in Remark 3.2, the relation between the moving and the local moving Fourier coefficients is as follows:

$$MF_k(\lambda_l) = c_{l + \left[div\left(k - \lfloor \frac{m}{2} \rfloor\right) - \mathbb{1}_{\{l \geq mod\left(k - \lfloor \frac{m}{2} \rfloor\right)\}} \right] m} \tag{8.6}$$

$$c_{k - \lfloor \frac{m}{2} \rfloor} = MF_k\left(\lambda_{mod\left(k - \lfloor \frac{m}{2} \rfloor\right)} \right). \tag{8.7}$$

Lemma 8.1. *Using the assumptions of Theorem 8.1 and with Theorem 5.5,*

$$\frac{1}{2\pi(2m+1)}\sum_{j=1}^{2m}\frac{MI_{\lfloor uT\rfloor,m}(\lambda_j)}{f(u,\lambda_j)} \xrightarrow{P} 1.$$

Proof. Note that with $\varphi(\lambda_j)=\frac{1}{f(u,\lambda_j)}$,

$$\frac{1}{(2m+1)}\sum_{j=1}^{2m}\frac{1}{f(u,\lambda_j)}MI_{\lfloor uT\rfloor,m}(\lambda_j) = MT\left(u\right).$$

in the notation of (5.25). We can thus use the results of Theorem 5.5 and have that for every $u\in[0,1]$ it holds that

$$\begin{aligned}
&\operatorname{Var}\left(\sqrt{2m+1}\left[MT(u)-E(MT(u))\right]\right) = \\
= \quad & \frac{4\pi^2}{2m+1}\sum_{j=0}^{2m}|\varphi(\lambda_j)|^2 f^2(u,\lambda_j) + (E(\varepsilon_1)^4-3)\left(\frac{8\pi^2}{(2m+1)^3}\cdot\right.\\
&\left.\cdot\sum_{j\neq l=0}^{2m}\varphi(\lambda_j)\overline{\varphi(\lambda_l)}f(u,\lambda_j)f(u,\lambda_l)\Sigma(l,j)\right)+o(1).
\end{aligned}$$

with

$$\begin{aligned}
\Sigma(l,j) &:= \max\{2m-l',2m-j',2m\}-\min\{0,-l',-j'\}+1,\\
\zeta_{\lfloor uT\rfloor,l} &:= div\left(\lfloor uT\rfloor-\left\lfloor\frac{m}{2}\right\rfloor\right)-\mathbb{1}_{\left\{l\geq mod\left(\lfloor uT\rfloor-\left\lfloor\frac{m}{2}\right\rfloor\right)\right\}},\\
l' &:= l+\zeta_{\lfloor uT\rfloor,l}m,\\
j' &:= j+\zeta_{\lfloor uT\rfloor,j}m,
\end{aligned}$$

with $\Sigma(l,j)$ bounded from above by $2m+1$.

$$E\left(MT(u)\right) \;=\; 2\pi+O\left(\frac{1}{\sqrt{m}}\right),$$

by Theorem 5.1. Hence, an application of the Markov inequality yields

$$\frac{1}{2m+1}\sum_{j=1}^{2m}\frac{1}{f(u,\lambda_j)}MI_{\lfloor uT\rfloor,m}(\lambda_j)=2\pi+o_P(1).$$

□

Proof of Theorem 8.1. The spectral density estimator is given by

$$\hat{f}\,(k) \;=\; \frac{1}{m}\sum_{t=-m}^{2m} K_h\left(\lambda_{mod(k)} - \lambda_t\right) MI_{k,m}\,(\lambda_t)\,,$$

see Definition 8.3. The set of local moving periodograms incorporated can also be written in the notation introduced in Remark 3.3 – as a set of moving Fourier coefficients. We estimate the spectral density at $k \in \{\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil, \ldots, \lfloor uT\rfloor + \lfloor \frac{m}{2}\rfloor\}$. Note that, depending on the time u, the set of moving Fourier coefficients includes a "jump" in the superscripts, which occurs at frequency $\lambda_{mod\left(\lfloor uT\rfloor - \lfloor \frac{m}{2}\rfloor\right)}$:

Now, for some fixed time u, the corresponding set of moving Fourier coefficients is

$$\mathcal{F}^{div(\lfloor uT\rfloor)}(\lambda_1), \ldots, \mathcal{F}^{div(\lfloor uT\rfloor)}(\lambda_{mod\left(\lfloor uT\rfloor - \lfloor \frac{m}{2}\rfloor\right)-1}),$$
$$\mathcal{F}^{div(\lfloor uT\rfloor)-1}(\lambda_{mod\left(\lfloor uT\rfloor - \lfloor \frac{m}{2}\rfloor\right)}), \ldots, \mathcal{F}^{div(\lfloor uT\rfloor)-1}(\lambda_m).$$

If we restrict the range of k to $m\left(div\left(\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil\right) - 1\right) \le k \le m\, div\left(\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil\right)$, we ensure that – for small h and due to the compact support of the kernel – only moving Fourier coefficients "after the jump" are used for estimation, which then enables us to reformulate the spectral density estimator:

$$\hat{f}(k) \;=\; \frac{1}{m}\sum_{t=-m}^{2m} K_h\left(\lambda_{mod(k)} - \lambda_t\right) |\mathcal{F}^{div\left(\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil\right)-1}(\lambda_t)|^2.$$

With $z := div\left(\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil\right) - 1$,

$$\hat{f}(k) \;=\; \frac{1}{m}\sum_{t=-m}^{2m} K_h\left(\lambda_{mod(k)} - \lambda_t\right) |\mathcal{F}^{z}(\lambda_t)|^2. \tag{8.8}$$

For $m\, div\left(\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil\right) + 1 \le k \le m\,\left(div\left(\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil\right) + 1\right)$, we ensure that only moving Fourier coefficients "before the jump" are used, hence,

$$\hat{f}(k) \;=\; \frac{1}{m}\sum_{t=-m}^{2m} K_h\left(\lambda_{mod(k)} - \lambda_t\right) |\mathcal{F}^{z+1}(\lambda_t)|^2. \tag{8.9}$$

Using restricted values of k allows for a handier representation of the spectral density estimator. In order to be able to use these representations (8.8) and (8.9), we need to perform at the beginning of the proof a split of the term

$$\sup_{\lfloor uT\rfloor - \lceil \frac{m}{2}\rceil < k \le \lfloor uT\rfloor + \lfloor \frac{m}{2}\rfloor} \left|\hat{f}\,(k) - f\left(u, \lambda_{mod(k)}\right)\right|$$

concerning the range of k:

$$
\begin{aligned}
& \sup_{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil < k \leq \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor} \left| \hat{f}(k) - f\left(u, \lambda_{mod(k)}\right) \right| \\
= \; & \max \Bigg(\sup_{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil < k \leq m \, div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right)} \left| \hat{f}(k) - f\left(u, \lambda_{mod(k)}\right) \right| \\
& + \sup_{m \, div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right) + 1 \leq k \leq \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor} \left| \hat{f}(k) - f\left(u, \lambda_{mod(k)}\right) \right| \Bigg) \\
\leq \; & \sup_{m\left(div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right) - 1\right) \leq k \leq m \, div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right)} \left| \hat{f}(k) - f\left(u, \lambda_{mod(k)}\right) \right| \\
& + \sup_{m \, div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right) + 1 \leq k \leq m\left(div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right) + 1\right)} \left| \hat{f}(k) - f\left(u, \lambda_{mod(k)}\right) \right| \\
:= \; & A + B. \qquad (8.10)
\end{aligned}
$$

The treatment of A and B is basically analogous.
In the following we use the same idea as in the proof of Theorem A1 in Franke and Härdle [20], as well as

$$
\boxed{\alpha_m := \frac{h}{m^{\frac{1}{4}}}, \; \mu_m := \left\lfloor \frac{1}{\alpha_m} \right\rfloor}
$$

Part I: $A = o_P(1)$

With $z := div\left(\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil\right) - 1$,

$$
\begin{aligned}
A \; = \; & \sup_{l=1,\ldots,m} \left| \frac{1}{m} \sum_{t=-m}^{2m} K_h(\lambda_l - \lambda_t) \left|\mathcal{F}^z(\lambda_t)\right|^2 - f(u, \lambda_l) \right| \\
\leq \; & \sup_{j \leq \mu_m} \left| \frac{1}{m} \sum_{t=-m}^{2m} K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) \left|\mathcal{F}^z(\lambda_t)\right|^2 - E\left(\frac{1}{m} \sum_{t=-m}^{2m} K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) \left|\mathcal{F}^z(\lambda_t)\right|^2 \right) \right| \\
& + \sup_{j \leq \mu_m} \left| E\left(\frac{1}{m} \sum_{t=-m}^{2m} K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) \left|\mathcal{F}^z(\lambda_t)\right|^2 \right) - f(u, \lambda_j) \right| \\
& + \sup_{|l-j| \leq \frac{\alpha_m(2m+1)}{\pi}} \left| \frac{1}{m} \sum_{t=-m}^{2m} \left[K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) - K_h\left(\lambda_{s'\left(\frac{\pi l}{\mu_m}\right)} - \lambda_t\right) \right] \mathcal{F}^z(\lambda_t)|^2 \right| \\
& + \sup_{|l-j| \leq \frac{\alpha_m(2m+1)}{\pi}} \left| f(u, \lambda_l) - f(u, \lambda_j) \right| \\
=: \; & A_1 + A_2 + A_3 + A_4. \qquad (8.11)
\end{aligned}
$$

$\boxed{A_4}$

The first thing one notices is $A_4 = o(1)$ as $|l-j| \leq \frac{\alpha_m(2m+1)}{\pi} \Leftrightarrow |\lambda_j - \lambda_l| \leq 2\alpha_m$, f is uniformly Lipschitz continuous in both arguments, see assumption (F)(i), and $\alpha_m \to 0$.

$\boxed{A_1}$

Note that from the assumptions (K) on the kernel function $\frac{1}{m}\sum_{j\in\mathbb{Z}} K_h(\lambda_j) = O(1)$, as well as $K_h^2(\cdot) \leq \frac{1}{h}K_h(\cdot)$.

We are interested in

$$\sup_{j\neq l} \operatorname{Cov}\left(|\mathcal{F}^z(\lambda_j)|^2, |\mathcal{F}^z(\lambda_l)|^2\right).$$

Being in the situation A, we always have $|l'-j'| = |l-j|$, due to the superscripts being the very same z for both arguments. That is, the set $\mathcal{A}_1(u, a_m)$ equals $\{1,\ldots,m\}^2$. Hence, with Lemma 5.3, we get

$$\sup_{j\neq l} \operatorname{Cov}\left(|\mathcal{F}^{z,\varepsilon}(\lambda_j)|^2, |\mathcal{F}^{z,\varepsilon}(\lambda_l)|^2\right) = O\left(\frac{1}{m}\right),$$

and, with the Cauchy-Schwarz inequality and Theorem 4.3,

$$\sup_{j\neq l} \operatorname{Cov}\left(|\mathcal{F}^{z}(\lambda_j)|^2, |\mathcal{F}^{z}(\lambda_l)|^2\right) = O\left(\frac{1}{\sqrt{m}}\right). \tag{8.12}$$

We need this result (8.12) when considering

$$\begin{aligned}
P(A_1 > \varepsilon) \quad &\leq \quad \sum_{j=1}^{\mu_m} \frac{1}{m^2\varepsilon^2} \operatorname{Var}\left(\sum_{t=-m}^{2m} K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) |\mathcal{F}^z(\lambda_t)|^2\right) \\
&= \quad \sum_{j=1}^{\mu_m} \frac{1}{m^2\varepsilon^2} \sum_{t=-m}^{2m} K_h^2\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) \operatorname{Var}\left(|\mathcal{F}^z(\lambda_t)|^2\right) \\
&\quad + \sum_{j=1}^{\mu_m} \frac{1}{m^2\varepsilon^2} \sum_{t\neq\tau=-m}^{2m} K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_t\right) K_h\left(\lambda_{s'\left(\frac{\pi j}{\mu_m}\right)} - \lambda_\tau\right) \\
&\quad \cdot \operatorname{Cov}\left(|\mathcal{F}^z(\lambda_t)|^2, |\mathcal{F}^z(\lambda_\tau)|^2\right) \\
&=: \quad A_{11} + A_{12}.
\end{aligned}$$

Now, A_{12} is with the above arguments of order $O\left(\frac{\mu_m}{\sqrt{m}}\right) = O\left(\frac{1}{m^{\frac{1}{4}}h}\right) = o(1)$, as $hm^{\frac{1}{4}} \to \infty$. With (5.28), $A_{11} = O\left(\frac{1}{m^{\frac{3}{4}}h}\right) = o(1)$, as $hm^{\frac{1}{4}} \to \infty$.

$\boxed{A_3}$

Making use of the kernel being uniformly Lipschitz continuous (see assumptions (K)),

$$\sup_{|l-j|\le\frac{\alpha_m(2m+1)}{\pi}} \left|\frac{1}{m}\sum_{t=-m}^{2m}\left[K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)-K_h\left(\lambda_{s'(\frac{\pi l}{\mu_m})}-\lambda_t\right)\right]|\mathcal{F}^z(\lambda_t)|^2\right|$$
$$\le \frac{1}{m}\sum_{t=-m}^{2m}|\mathcal{F}^z(\lambda_t)|^2 O\left(\frac{\alpha_m}{h^2}\right).$$

Note that due to Lemma 8.1, $\frac{1}{m}\sum_{t=-m}^{2m}|\mathcal{F}^z(\lambda_t)|^2 = O_P(1)$, and therefore, as $hm^{\frac{1}{4}}\to\infty$,

$$A_3 = O_P\left(\frac{1}{hm^{\frac{1}{4}}}\right) = o_P(1).$$

$\boxed{A_2}$

With Theorem 5.1 and, again, $\frac{1}{m}\sum_{j\in\mathbb{Z}}K_h(\lambda_j)=O(1)$,

$$\begin{aligned}
&\sup_{j\le\mu_m}\left|E\left(\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)|\mathcal{F}^z(\lambda_t)|^2\right)-f(u,\lambda_j)\right| \\
=\ &\sup_{j\le\mu_m}\left|\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)\right. \\
&\left.\cdot\left(f\left(\frac{s'\left(\frac{\pi j}{\mu_m}\right)+m\left(div\left(\lfloor uT\rfloor-\lceil\frac{m}{2}\rceil\right)-1\right)}{T},\lambda_t\right)-f(u,\lambda_t)\right)\right| \\
&+\sup_{j\le\mu_m}\left|\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)f(u,\lambda_t)-f(u,\lambda_j)\right|+o(1) \\
:=\ &A_{21}+A_{22}+o(1).
\end{aligned}$$

Using the uniform Lipschitz continuity of f yields $A_{21}=o(1)$.

$$\begin{aligned}
A_{22} &\le \sup_{j\le\mu_m}\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)|f(u,\lambda_t)-f(u,\lambda_j)|+o(1) \\
&\le \sup_{j\le\mu_m}\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)|\lambda_t-\lambda_j|+o(1) \\
&\le \sup_{j\le\mu_m}\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)\mathbb{1}_{\{|\lambda_t-\lambda_j|\le Ch\}}|\lambda_t-\lambda_j|+o(1) \\
&\le Ch\frac{1}{m}\sum_{t=-m}^{2m}K_h\left(\lambda_{s'(\frac{\pi j}{\mu_m})}-\lambda_t\right)+o(1)=o(1).
\end{aligned}$$

□

Part II: $B = o_P(1)$

The treatment of B is analogous to A, except for a minor alteration in the splitting up (cf. (8.11)) in four analogue terms to A_i, $i = 1, 2, 3, 4$. Instead of using the superscript z, we have to use the superscript $z+1$. The following proof is, after the change in the superscript, again, completely analogous to part I for the analogue terms to A_1, A_3 and A_4. When considering the analogon to A_2, we merely have to bear in mind that instead of looking at time $s'\left(\frac{\pi j}{\mu_m}\right)+m\left(div\left(\lfloor uT\rfloor - \lceil\frac{m}{2}\rceil\right)-1\right)$, as we did in A_{21}, the point of time concerned is $s'\left(\frac{\pi j}{\mu_m}\right)+m\,div\left(\lfloor uT\rfloor - \lceil\frac{m}{2}\rceil\right)$, which does not make any difference to the behaviour of the analogue term to A_2, as the two times are only m apart.

CHAPTER 9

Covariance structure of the bootstrap sample

This chapter is devoted to proving that the moving TFT-Bootstrap maintains the second order structure of the original process.
W.l.o.g. only lags $h \geq 0$ are considered. A distinction is made between some fixed integer h and h increasing with m in the way that $\frac{h}{m} \to \alpha$ for $m \to \infty$ and $0 < \alpha < 1$. In the case of $\alpha > 1$ or $\frac{h}{m} \to \infty$ the bootstrap observations $X^*_{t,T}$ and $X^*_{t+h,T}$ are independent, due to the m-dependence of the bootstrap scheme.

(B) Assumptions on the local bootstrap Fourier coefficients

(i) $E^*(c^*_k) = 0, \quad \forall k = 1, \ldots, T.$

(ii) Independence of c^*_l and c^*_k $(k \neq l)$. for any $k, l = 1, \ldots, T$.

(iii)

$$\begin{aligned}
\sup_{k \in \{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil + 1, \ldots, \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor\}} \left| \mathrm{Var}^*(Re(c^*_k)) - \pi f(u, \lambda_k) \right| &= o_P(1). \\
\sup_{k \in \{\lfloor uT \rfloor - \lceil \frac{m}{2} \rceil + 1, \ldots, \lfloor uT \rfloor + \lfloor \frac{m}{2} \rfloor\}} \left| \mathrm{Var}^*(Im(c^*_k)) - \pi f(u, \lambda_k) \right| &= o_P(1).
\end{aligned}$$

(iv) Independence of $\mathrm{Re}(c^*_k)$ and $\mathrm{Im}(c^*_k)$.

Remark 9.1
Due to Theorem 8.1, assumption (iii) is fulfilled for the wild bootstrap. (B)(i),(ii),(iv) are true due to construction of the bootstrap replicates.

Theorem 9.1 (Second order structure of bootstrap replicate).
Let $X_{t,T}$ be a locally stationary time series with time varying spectral density f meeting (F)(i) and (ii). Further assume that (K)(i)-(iv) as well as (B)(i)-(iv) hold. $X^*_{t,T}$ is the bootstrap time series created according to the scheme in Section 7.2.1 using the wild bootstrap.
Then

$$\sup_{|h|\leq m}\left|Cov^*(X^*_{\lfloor uT\rfloor,T},X^*_{\lfloor uT\rfloor+h,T})-c(u,h)\right|=o_P(1).$$

Proof. To simplify notation we set $t:=\lfloor uT\rfloor-\left\lfloor\frac{m}{2}\right\rfloor$. The final result is then adjusted by shifting. Note that

$$\mathrm{E}^*(c^*_k\overline{c^*_l})=\mathrm{E}^*(Re(c^*_k)Re(\overline{c^*}_l)+iRe(c^*_k)Im(\overline{c^*}_l)+iIm(c^*_k)Re(\overline{c^*}_l)+Im(c^*_k)Im(\overline{c^*}_l)).$$

Due to Assumption (B)(ii),

$$\mathrm{E}^*(Re(c^*_k)Im(\overline{c^*}_l))=\mathrm{E}^*(Re(c^*_k))\,\mathrm{E}(Im(\overline{c^*}_l))=0,$$

for $k\neq l$. Analogously, $\mathrm{E}^*(Im(c^*_k)Re(\overline{c^*}_l))=0$, for $k\neq l$.

$$\begin{aligned}
&(2m+1)\,\mathrm{Cov}^*(X^*_{t+\lfloor\frac{m}{2}\rfloor,T},X^*_{t+\lfloor\frac{m}{2}\rfloor+h,T})\\
=\;&\mathrm{E}^*\Bigg[\Bigg(\sum_{l=1}^m c^*_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}e^{i\lambda_l\left(t+\lfloor\frac{m}{2}\rfloor\right)}+\sum_{l=1}^m \overline{c^*}_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}e^{-i\lambda_l\left(t+\lfloor\frac{m}{2}\rfloor\right)}\Bigg)\\
&\cdot\Bigg(\sum_{l=1}^m \overline{c^*}_{l+\left(div(t+h)-\mathbb{1}_{\{l\geq mod(t+h)\}}\right)m}e^{-i\lambda_l\left(t+h+\lfloor\frac{m}{2}\rfloor\right)}\\
&+\sum_{l=1}^m c^*_{l+\left(div(t+h)-\mathbb{1}_{\{l\geq mod(t+h)\}}\right)m}e^{i\lambda_l\left(t+h\lfloor\frac{m}{2}\rfloor\right)}\Bigg)\Bigg]\\
=\;&\Bigg(\sum_{l,k=1}^m \mathrm{E}^*\left(c^*_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}\overline{c^*}_{k+\left(div(t+h)-\mathbb{1}_{\{k\geq mod(t+h)\}}\right)m}\right)e^{i\lambda_l\left(t+\lfloor\frac{m}{2}\rfloor\right)}e^{-i\lambda_k\left(t+h+\lfloor\frac{m}{2}\rfloor\right)}\\
&\sum_{l,k=1}^m \mathrm{E}^*\left(c^*_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}c^*_{k+\left(div(t+h)-\mathbb{1}_{\{k\geq mod(t+h)\}}\right)m}\right)e^{i\lambda_l\left(t+\lfloor\frac{m}{2}\rfloor\right)}e^{i\lambda_k\left(t+h+\lfloor\frac{m}{2}\rfloor\right)}\\
&\sum_{l,k=1}^m \mathrm{E}^*\left(\overline{c^*}_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}c^*_{k+\left(div(t+h)-\mathbb{1}_{\{k\geq mod(t+h)\}}\right)m}\right)e^{-i\lambda_l\left(t+\lfloor\frac{m}{2}\rfloor\right)}e^{i\lambda_k\left(t+h+\lfloor\frac{m}{2}\rfloor\right)}\\
&\sum_{l,k=1}^m \mathrm{E}^*\left(\overline{c^*}_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}\overline{c^*}_{k+\left(div(t+h)-\mathbb{1}_{\{k\geq mod(t+h)\}}\right)m}\right)e^{-i\lambda_l\left(t+\lfloor\frac{m}{2}\rfloor\right)}e^{-i\lambda_k\left(t+h+\lfloor\frac{m}{2}\rfloor\right)}\Bigg)
\end{aligned}$$

As the bootstrap coefficients are assumed to be independent for different indices according to (B)(ii), we need to clarify when the indices of $c^*_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}$ and

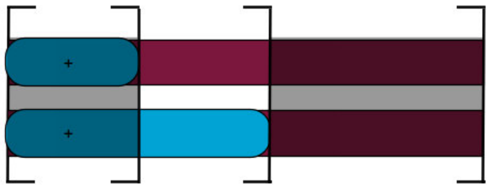
$c^*_{k+\left(div(t+h)-\mathbb{1}_{\{k\geq mod(t+h)\}}\right)m}$ are equal.

Let w.l.o.g. $h \geq 0$. We are hence only concerned with $h = 0, \ldots, m$, as for larger h, the bootstrap covariance equals zero.
Accordingly, we always have $t \leq t+h$.

The following situations can occur:

Case 1: $\{div(t) = div(t+h)\} =: A$

As $h \leq m$, we can definitely say that $mod(t) \leq mod(t+h)$. In this case, the following figure exemplarily states the situation. The shaded area marks the intervals in which the indices of $c^*_{l+\left(div(t)-\mathbb{1}_{\{l\geq mod(t)\}}\right)m}$ and $c^*_{k+\left(div(t+h)-\mathbb{1}_{\{l\geq mod(t+h)\}}\right)m}$ are equal.

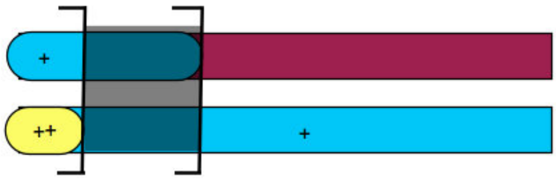


We get matches for $l = 1 \ldots mod(t)-1$ and $l = mod(t)+h, \ldots, m$

We can also write the condition of case 1 in a different way. $A = \{t+h \leq div(t)m\}$.

Case 2: $\{div(t) = div(t+h)-1\} =: B$

In this case, again, as $h \leq m$, we know for sure, that $mod(t+h) \leq mod(t)$.



We get matches for $l = mod(t+h), \ldots, mod(t)-1 = mod(t)+h-m, \ldots, mod(t)-1$

We can also write the condition of case 2 in a different way: $B = \{div(t)m < t+h \leq (div(t)+1)m\}$

With $|h| \leq m$, there is no possibility of getting into the situation of $|div(t) - div(t+h)| \geq 2$.

The Fourier coefficients are constructed using the estimated time varying spectral density. Now, using the result of Chapter 8 concerning the spectral density estimator and the assumption (B)(iii), we may write:

$$\begin{aligned}
&\mathrm{Cov}^*(X^*_{t+\lfloor\frac{m}{2}\rfloor,T}, X^*_{t+\lfloor\frac{m}{2}\rfloor+h,T}) \\
= \quad & \frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=1}^{mod(t)-1}(f(u,\lambda_l)+o_P(1))e^{-i\lambda_l h}+\sum_{l=mod(t)+h}^{m}(f(u,\lambda_l)+o_P(1))e^{-i\lambda_l h}\right) \\
&+\frac{2\pi}{2m+1}\mathbb{1}_B\sum_{l=mod(t)+h-m}^{mod(t)-1}(f(u,\lambda_l)+o_P(1))e^{-i\lambda_l h} \\
&+\frac{2\pi}{2m+1}\mathbb{1}_A\Big(\sum_{l=1}^{mod(t)-1}(f(u,\lambda_l)+o_P(1))e^{+i\lambda_l h}+\sum_{l=mod(t)+h}^{m}(f(u,\lambda_l)+o_P(1))e^{+i\lambda_l h}\Big) \\
&+\frac{2\pi}{2m+1}\mathbb{1}_B\sum_{l=mod(t)+h-m}^{mod(t)-1}(f(u,\lambda_l)+o_P(1))e^{+i\lambda_l h} \\
+ \quad & \frac{1}{2m+1}\mathbb{1}_A\Big(\sum_{l=1}^{mod(t)-1}o_P(1)e^{-i\lambda_l\left(h+2t+2\lfloor\frac{m}{2}\rfloor\right)}+\sum_{l=t+h-m}^{m}o_P(1)e^{-i\lambda_l h}\Big) \\
&+\frac{1}{2m+1}\mathbb{1}_B\sum_{l=mod(t)+h-m}^{mod(t)-1}o_P(1)e^{-i\lambda_l\left(h+2t+2\lfloor\frac{m}{2}\rfloor\right)} \\
&+\frac{1}{2m+1}\mathbb{1}_A\Big(\sum_{l=1}^{mod(t)-1}o_P(1)e^{+i\lambda_l\left(h+2t+2\lfloor\frac{m}{2}\rfloor\right)}+\sum_{l=t+h-m}^{m}o_P(1)e^{+i\lambda_l\left(h+2t+2\lfloor\frac{m}{2}\rfloor\right)}\Big) \\
&+\frac{2\pi}{2m+1}\mathbb{1}_B\sum_{l=mod(t)+h-m}^{mod(t)-1}o_P(1)e^{+i\lambda_l h}.
\end{aligned}$$

Treating all sums involving the term $o_P(1)$ as $o_P(1)$ is possible, as these sums can at most have $2m+1$ summands and we then get the uniform asymptotic behaviour of $\frac{2m+1}{m}\cdot o_P(1)=o_P(1)$. We can thus simplify to

$$\begin{aligned}
&\frac{2\pi}{2m+1}\mathrm{Cov}^*(X^*_{t+\lfloor\frac{m}{2}\rfloor,T}, X^*_{t+\lfloor\frac{m}{2}\rfloor+h,T}) \\
= \quad & \frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=1}^{mod(t)-1}f(u,\lambda_l)e^{-i\lambda_l h}+\sum_{l=mod(t)+h}^{m}f(u,\lambda_l)e^{-i\lambda_l h}\right) \\
&+\frac{2\pi}{2m+1}\mathbb{1}_A\Big(\sum_{l=1}^{mod(t)-1}f(u,\lambda_l)e^{+i\lambda_l h}+\sum_{l=mod(t)+h}^{m}f(u,\lambda_l)e^{+i\lambda_l h}\Big) \\
&+\frac{2\pi}{2m+1}\mathbb{1}_B\left(\sum_{l=mod(t)+h-m}^{mod(t)-1}f(u,\lambda_l)e^{-i\lambda_l h}+\sum_{l=mod(t)+h-m}^{mod(t)-1}f(u,\lambda_l)e^{+i\lambda_l h}\right)+o_P(1)
\end{aligned}$$

$$= \frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=1}^{mod(t)-1} f(u,\lambda_l)e^{-i\lambda_l h} + \sum_{l=mod(t)+h}^{m} f(u,\lambda_l)e^{-i\lambda_l h}\right)$$
$$+\frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=2m+2-mod(t)}^{2m} f(u,\lambda_{2m+1-l})e^{-i\lambda_l h} + \sum_{l=m+1}^{2m+1-mod(t)-h} f(u,\lambda_{2m+1-l})e^{-i\lambda_l h}\right)$$
$$+\frac{2\pi}{2m+1}\mathbb{1}_B\left(\sum_{l=mod(t)+h-m}^{mod(t)-1} f(u,\lambda_l)e^{-i\lambda_l h} + \sum_{l=2m+2-mod(t)}^{3m+1-h-mod(t)} f(u,\lambda_{2m+1-l})e^{-i\lambda_l h}\right) + o_P(1),$$

noting symmetry of the spectral density we substitute $f(u,\lambda_l) = f(u,\lambda_{2m+1-l})$.

We readily get that $\mathrm{Cov}^*(X^*_{t+\lfloor \frac{m}{2}\rfloor,T}, X^*_{t+\lfloor \frac{m}{2}\rfloor+h,T}) = O_P(\frac{m-h}{m})$, as the number of summands is part A as well as in part B is equal to $m-h$.

We can now continue with completing the fragments in part A as well as in part B to a sum from 1 to $2m+1$:

$$\mathrm{Cov}^*(X^*_{t+\lfloor \frac{m}{2}\rfloor,T}, X^*_{t+\lfloor \frac{m}{2}\rfloor+h,T})$$
$$= \frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=1}^{2m+1} f(u,\lambda_l)e^{-i\lambda_l h} - \sum_{l=mod(t)}^{mod(t)+h-1} f(u,\lambda_l)e^{-i\lambda_l h} - f(u,\lambda_{2m+1})e^{-i\lambda_{2m+1}h}\right.$$
$$\left. - \sum_{l=2m+2-mod(t)-h}^{2m+1-mod(t)} f(u,\lambda_l)e^{-i\lambda_l h}\right)$$
$$+\frac{2\pi}{2m+1}\mathbb{1}_B\left(\sum_{l=1}^{2m+1} f(u,\lambda_l)e^{-i\lambda_l h} - \sum_{l=1}^{mod(t)+h-m-1} f(u,\lambda_l)e^{-i\lambda_l h}\right.$$
$$\left. - \sum_{l=mod(t)}^{2m+1-mod(t)} f(u,\lambda_l)e^{-i\lambda_l h} - \sum_{l=3m+2-h-mod(t)}^{2m+1} f(u,\lambda_{2m+1-l})e^{-i\lambda_l h}\right) + o_P(1)$$
$$= \frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=1}^{2m+1} f(u,\lambda_l)e^{-i\lambda_l h} + \frac{2\pi}{2m+1}\mathbb{1}_B\sum_{l=1}^{2m+1} f(u,\lambda_l)e^{-i\lambda_l h} + o_P(1)\right.$$
$$-\frac{2\pi}{2m+1}\mathbb{1}_A\left(\sum_{l=mod(t)}^{mod(t)+h-1} f(u,\lambda_l)e^{-i\lambda_l h} + f(u,\lambda_{2m+1})e^{-i\lambda_{2m+1}h}\right.$$
$$\left. + \sum_{l=2m+1-mod(t)-h}^{2m+1-mod(t)} f(u,\lambda_l)e^{-i\lambda_l h}\right) - \frac{2\pi}{2m+1}\mathbb{1}_B\left(+ \sum_{l=1}^{mod(t)+h-m-1} f(u,\lambda_l)e^{-i\lambda_l h}\right.$$
$$\left. + \sum_{l=mod(t)}^{2m+1-mod(t)} f(u,\lambda_l)e^{-i\lambda_l h} + \sum_{l=3m+2-h-mod(t)}^{2m+1} f(u,\lambda_{2m+1-l})e^{-i\lambda_l h}\right)$$
$$=: A_1 + B_1 + o_P(1) + A_2 + B_2.$$

Note that with Definition 2.2 we get the following relation between the time-varying spectral density at time u and frequency λ_l and the time-varying autocovariance function.

$$\begin{aligned} f(u,\lambda_l) &= \frac{1}{2\pi}|A(u,\lambda_l)|^2 = \frac{1}{2\pi}\sum_{j,k=-\infty}^{\infty} a(u,j)a(u,k)e^{-i(\lambda_l(j-k))} \\ &= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} a(u,j)a(u,j-n)e^{-i\lambda_l n} = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} c(u,n)e^{-i\lambda_l n}. \end{aligned}$$

Hence,

$$\begin{aligned} &\frac{2\pi}{2m+1}\Big(\mathbb{1}_A\sum_{l=1}^{2m+1} f(u,\lambda_l)e^{-i\lambda_l h} + \mathbb{1}_B\sum_{l=1}^{2m+1} f(u,\lambda_l)e^{-i\lambda_l h}\Big) \\ = & \frac{1}{2m+1}\mathbb{1}_A\Big(\sum_{n=-\infty}^{\infty} c(u,n)\sum_{l=1}^{2m+1} e^{-i\lambda_l(h+n)}\Big) \\ + & \frac{1}{2m+1}\mathbb{1}_B\Big(\sum_{n=-\infty}^{\infty} c(u,n)\sum_{l=1}^{2m+1} e^{-i\lambda_l(h+n)}\Big). \end{aligned}$$

The first two sums equal zero except for the case when $h+n=\mathbb{Z}\cdot(2m+1)$. In this case, $\sum_{l=1}^{2m+1} e^{-i\lambda_l(h+n)} = 2m+1$. Therefore, due to the absolute summability of the autocovariance functions (cf. Remark 2.4),

$$\begin{aligned} A_1 + B_1 &= \mathbb{1}_A\Big(c(u,h) + \sum_{|k|\geq 1} c(u,h+k(2m+1))\Big) \\ &\quad + \mathbb{1}_B\Big(c(u,h) + \sum_{|k|\geq 1} c(u,h+k(2m+1))\Big) \\ &= c(u,h) + \sum_{|k|\geq 1} c(u,h+k(2m+1)). \end{aligned}$$

The last sum can, as $|h| \leq m$, be bounded by

$$\sum_{|k|\geq 1} c(u,h+k(2m+1)) \leq \sum_{|l|\geq m} c(u,l) = o(1).$$

With this knowledge, we may now write

$$\begin{aligned} &\sup_{|h|\leq m}\left|\mathrm{Cov}^*(X^*_{\lfloor uT\rfloor,T}, X^*_{\lfloor uT\rfloor+h,T}) - c(u,h)\right| \\ = &\sup_{|h|\leq m}|A_2 + B_2| + o_P(1). \end{aligned}$$

As the term A_2 is only non-zero for the indictor $\mathbb{1}_A$ being equal to 1, we do not need to consider all $0 \leq h \leq m$, but only those that result in $div(t) = div(t+h)$, $t := \lfloor uT \rfloor$. We have noted before that we can also write the set A as $\{t + h \leq div(t)m\}$. This condition can be reformulated as $h \leq div(t)m - t$ and, with the definition of div and mod, we get $h \leq m - mod(t)$. We therefore need to consider only $0 \leq h \leq m - mod(t)$ when looking at A_2. Similarly, we only need to consider $m \geq h > m - mod(t)$ when looking at B_2.

To treat $\sup_{|h| \leq m} |A_2 + B_2|$, firstly, consider

$$\begin{aligned}\sup_{0 \leq h \leq m - mod(t)} |A_2| &= \frac{1}{m} \sup_{0 \leq h \leq m - mod(t)} \left| \sum_{l=mod(t)}^{mod(t)+h-1} f(u, \lambda_l) e^{-i\lambda_l h} \right. \\ &\quad \left. + \sum_{l=2m+1-mod(t)-h}^{2m+1-mod(t)} f(u, \lambda_l) e^{-i\lambda_l h} \right| + O\left(\frac{1}{m}\right)\end{aligned}$$

Hence,

$$\begin{aligned}\sup_{0 \leq h \leq m - mod(t)} |A_2| &= O\left(\frac{1}{m}\right) + \sup_{0 < h \leq m - mod(t)} \frac{1}{m} \sum_{n=-\infty}^{\infty} |c(u,n)| \left(\left| \sum_{l=0}^{h-1} e^{-i\lambda_{l+mod(t)}(n+h)} \right| \right. \\ &\quad \left. + \left| \sum_{l=0}^{h} e^{-i\lambda_{l+2m+1-mod(t)-h}(n+h)} \right| \right) \\ &= O\left(\frac{1}{m}\right) + \sup_{0 < h \leq m - mod(t)} \frac{1}{m} \sum_{n=-\infty}^{\infty} |c(u,n)| \, O\left(\min\left(\frac{2m+1}{|n+h|}, |h| \right) \right)\end{aligned}$$

by Lemma A.4 in Kirch [28]. We continue with

$$\begin{aligned}&\sup_{0 < h \leq m - mod(t)} \frac{1}{m} \sum_{n=-\infty}^{\infty} |c(u,n)| \, O\left(\min\left(\frac{2m+1}{|n+h|}, |h| \right) \right) \\ &= \sup_{0 < h \leq m - mod(t)} \sum_{n=-\infty}^{\infty} |c(u,n)| \, O\left(\min\left(\frac{1}{|n+h|}, \frac{|h|}{m} \right) \right) \\ &= \sup_{0 < h \leq m - mod(t)} \sum_{|n| < \sqrt{h}} |c(u,n)| \, O\left(\min\left(\frac{1}{h}, \frac{|h|}{m} \right) \right) \\ &\quad + \sup_{0 < h \leq m - mod(t)} \sum_{|n| \geq \sqrt{h}} |c(u,n)| \, O\left(\min\left(1, \frac{|h|}{m} \right) \right)\end{aligned}$$

$$
\begin{aligned}
&\leq O\left(\sup_{h\leq m}\min\left(\frac{1}{h},\frac{|h|}{m}\right)\right)+O\left(\sup_{h\leq m}\min\left(\frac{|h|}{m},\sum_{|n|\geq\sqrt{h}}|c(u,n)|\right)\right)\\
&\leq O\left(\frac{1}{\sqrt{m}}\right)+O\left(\max\left(\frac{1}{\sqrt{m}},\sum_{|n|\geq\sqrt{m}}|c(u,n)|\right)\right)\\
&= o(1).
\end{aligned}
$$

Secondly, we look at

$$
\begin{aligned}
\sup_{m-mod(t)<h\leq m}|B_2| \leq& \sup_{m-mod(t)<h\leq m}\frac{1}{m}\left(\left|\sum_{l=1}^{mod(t)+h-m-1} f(u,\lambda_l)e^{-i\lambda_l h}\right|\right.\\
&+\left|\sum_{l=1}^{2m+2-2mod(t)} f(u,\lambda_{l+mod(t)-1})e^{-i\lambda_{l+mod(t)-1}h}\right|\\
&\left.+\left|\sum_{l=1}^{-m+h+mod(t)} f(u,\lambda_{l+3m+1-h-mod(t)})e^{-i\lambda_{l+3m+1-h-mod(t)}h}\right|\right)\\
\leq& \sup_{m-mod(t)<h\leq m}\frac{1}{m}\sum_{n=-\infty}^{\infty}|c(u,n)|\left(\left|\sum_{l=1}^{mod(t)+h-m-1} e^{-i\lambda_l(h+n)}\right|\right.\\
&\left.+\left|\sum_{l=1}^{2m+2-2mod(t)} e^{-i\lambda_{l+mod(t)-1}(h+n)}\right|+\left|\sum_{l=1}^{-m+h+mod(t)} e^{-i\lambda_{l+3m+1-h-mod(t)}(h+n)}\right|\right)
\end{aligned}
$$

With Lemma A.4 in Kirch [28] we can continue analogously as for the term A_2:

$$
\begin{aligned}
&\sup_{m-mod(t)<h\leq m}\frac{1}{m}\sum_{n=-\infty}^{\infty}|c(u,n)|\left(O\left(\min\left(\frac{2m+1}{|n+h|},|mod(t)+h-m-1|\right)\right)\right.\\
&\left.+O\left(\min\left(\frac{2m+1}{|n+h|},|2m+2-2mod(t)|\right)\right)\right)\\
=& \sup_{m-mod(t)<h\leq m}\left[\sum_{|n|<\sqrt{h}}|c(u,n)|\left(O\left(\min\left(\frac{1}{h},\frac{|mod(t)+h-m-1|}{m}\right)\right)\right.\right.\\
&\left.+O\left(\min\left(\frac{1}{h},\frac{|2m+2-2mod(t)|}{m}\right)\right)\right)\\
&+\sum_{|n|\geq\sqrt{h}}|c(u,n)|\left(O\left(\min\left(1,\frac{|mod(t)+h-m-1|}{m}\right)\right)\right.\\
&\left.\left.+O\left(\min\left(1,\frac{|2m+2-2mod(t)|}{m}\right)\right)\right)\right]
\end{aligned}
$$

As $mod(t) \leq m$ and $mod(t) + h - m \geq 0$, we can bound $|mod(t) + h - m|$ by h. Moreover, as also $h > m - mod(t)$, we may bound $|2m + 2 - 2mod(t)|$ by $2h$ as well. Hence,

$$\begin{aligned}\sup_{m-mod(t)<h\leq m} |B_2| &\leq O\left(\sup_{h\leq m} \min\left(\frac{1}{h}, \frac{h}{m}\right)\right) + O\left(\sup_{h\leq m} \min\left(\frac{1}{h}, \frac{2h}{m}\right)\right) \\ &\quad + O\left(\sup_{h\leq m} \min\left(\frac{h}{m}, \sum_{|n|\geq\sqrt{h}} |c(u,n)|\right)\right) \\ &\leq O\left(\frac{1}{\sqrt{m}}\right) + O\left(\max\left(\frac{1}{\sqrt{m}}, \sum_{|n|\geq\sqrt{m}} |c(u,n)|\right)\right)\end{aligned}$$

Thus,

$$\sup_{0\leq h\leq m} |A_2 + B_2| = o(1).$$

So all in all,

$$\sup_{0\leq h\leq m} \text{Cov}^*(X^*_{t+\lfloor \frac{m}{2} \rfloor,T}, X^*_{t+\lfloor \frac{m}{2} \rfloor+h,T}) = o_P(1). \qquad \square$$

CHAPTER 10

Deficiency of the Adapted Fast Fourier Transformation (AFFT)

As the moving Fourier transformation of T values $X_1, \ldots, X_T$ with window width $N := 2m + 1$ is of order $O(mT)$, one might think of exploiting the benefits of developing an algorithm in the style of the ordinary fast Fourier transform to reduce computing time. This is, unfortunately, not possible without compromises. The first section shortly displays the algorithm of the fast Fourier transform. In the second part, we adapt the fast Fourier transform to fit our needs. And in the next step, we then give the reasons of why we can't possibly achieve any improvement in computing time. Finally, an algorithm is suggested which – to a previously chosen extent – compromises on 'locality' for the benefit of speed.

10.1 The fast Fourier transform

This Section follows closely Chapter 3.7 in Schwarz and Koeckler [48].

Assume we have N values $X_0, \ldots, X_{N-1}$. For reasons of simplicity, we assume that $N = 2^q$, which covers the most popular algorithm. There are also algorithms for N being a power of other bases, for example Boor [17], Brigham [2] and Winograd [55]. Nowadays, software practically allows for any N, the amount of complexity depending on the prime factorization of N.
We employ the standard notation and use $\omega_N^j := e^{-\frac{2\pi ij}{N}} = e^{-i\lambda_j}$, $j = 0, \ldots, N-1$, to refer to the j-th unit root. The Fourier coefficient at frequency $\lambda_k = \frac{2\pi k}{N}$ is then given by

$$c_k := \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} X_j e^{-ij\lambda_k}.$$

For the fast Fourier transform one needs to distinguish between odd and even indices.

Let, for the first step to become more illustrating $m := \frac{N}{2} = 2^{q-1}$.

$k = 2l,\ l = 0, \ldots, m-1$

As $\omega_N^{2l(m+j)} = \omega_N^{2lj}\omega_N^{2lm} = \omega_N^{2lj}$ and $\omega_m = \omega_N^2$, we get

$$\begin{aligned} c_{2l} &= \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} X_j e^{-ij\lambda_{2l}} = \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} X_j \omega_N^{2lj} = \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1}(X_j + X_{m+j})\omega_N^{2lj} \\ &= \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1}(X_j + X_{m+j})(\omega_N^2)^{lj} = \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1} z_j\omega_m^{lj}. \end{aligned} \tag{10.1}$$

We have now reduced the Fourier transform of the $N = 2m$ values $X_1, \ldots, X_N$ to a Fourier transform of the m auxiliary variables

$$z_{j,e} := X_j + X_{m+j},\ j = 0, \ldots, m-1.$$

$k = 2l+1,\ l = 0, \ldots, m-1$

As $\omega_N^{(2l+1)(m+j)} = \omega_N^{(2l+1)j}\omega_N^{(2l+1)m} = -\omega_N^{(2l+1)j}$ and $\omega_m = \omega_N^2$, we get

$$\begin{aligned} c_{2l+1} &= \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} X_j\omega_N^{(2l+1)j} = \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1}\left(X_j\omega_N^{(2l+1)j} + X_{m+j}\omega_N^{(2l+1)(m+j)}\right) \\ &= \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1}\left(X_j - X_{m+j}\right)\omega_N^{(2l+1)j} = \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1}\left((X_j - X_{m+j})\omega_N^j\right)\omega_N^{2lj} \\ &= \frac{1}{\sqrt{N}}\sum_{j=0}^{m-1} z_{j+m}\omega_m^{lj}. \end{aligned} \tag{10.2}$$

Again, we have reduced the Fourier transform of the $N = 2m$ values $X_1, \ldots, X_N$ to a Fourier transform of m auxiliary variables

$$z_{j+m,o} := (X_j - X_{m+j})\omega_N^j,\ j = 0, \ldots, m-1.$$

This act of reducing the Fourier transform of $2m$ values to a Fourier transform of m values effectively costs m complex multiplications (for calculating the z_{j+m}).
Note, that for each $k = 2l$, we have the same auxiliary variables z_j, $j = 0, \ldots, m-1$. Analogously, for each $k = 2l+1$, we have the same auxiliary variables z_{j+m}, $j = , \ldots, m-1$. This is actually the key to why we lose complexity - the auxiliary variables remaining unchanged in each group.

The next step is then to reduce both of the new formulae (10.1) and (10.2) to Fourier transforms of $\frac{m}{2}$ values, which costs $2 \cdot \frac{m}{2} = \frac{N}{2}$ complex multiplications. That is, a Fourier transform of order $N = 2^q$ can in q steps be reduced to N Fourier transforms of order 1, which are the desired coefficients. Each of the steps requires $\frac{N}{2}$ complex multiplications and thus, the total complexity is

$$\frac{N}{2} \cdot q = O(N \log_2 N).$$

Compared to the straight calculation of N Fourier coefficients, each of them resulting from a Fourier transform of N values, we have reduced the complexity from $O(N^2)$ to $O(N \log_2 N)$ by using this special algorithm, the Fast Fourier Transformation.

10.2 Procedure

As the moving Fourier transformation performs a Fourier transform of length $2m$ on each stretch of data, one might think of employing the FFT-algorithms to speed things up. However, one has to note that we don't actually calculate all frequencies for each stretch, but calculate one frequency only and then shift to the next stretch. The point of matter, therefore, is whether the shift still allows for a sufficient 'reuse' of the auxiliary variables z_j and $z_{j+\frac{N}{2^q}}$ in each group $j = 0, \dots, \frac{N}{2^q} - 1$, $q = 1, \dots, p$. The parameter q denotes the reduction step we are currently at. To find out, whether shifting causes the algorithm to lose its computational advantage (which it has compared to the ordinary DFT) compared to the straight calculation of the moving Fourier transform, we need to write it down first.
Note that the width of the chosen window for our transformation has been $2m + 1$ in the previous chapters. For reasons of simplicity, as pointed out before, we select a window width of $N := 2m = 2^q$, $q \in \mathcal{N}$. Further, we assume $m|T$.
The procedure of the adapted fast Fourier transform is analogous to the ordinary case. We reduce the transform of N values to 2 transforms of $m := \frac{N}{2}$ values each (distinguishing odd and even indices, as usual).
Note that we want to look at the complexity of transforming a time series of length T, which normally only yields $T - m$ moving Fourier coefficients. As m is so much smaller than T, we can very well consider the coefficients c_j, $1 \leq j \leq T$, when being interested in complexity only.
This results in $\frac{T}{2}$ odd and $\frac{T}{2}$ even indices.

$\boxed{k = 2l,\ l = 0, \dots, m-1}$

Analogously to the stationary case,

$$\begin{aligned} c_{2l} &= \frac{1}{\sqrt{N}} \sum_{j=2l-m+1}^{m+2l} X_j \omega_N^{2lj} = \frac{1}{\sqrt{N}} \sum_{j=2l-m+1}^{2l} (X_j + X_{j+m}) \omega_m^{lj} \\ &= \frac{1}{\sqrt{N}} \sum_{j=2l-m+1}^{2l} z_{j,e}\, \omega_m^{lj}, \end{aligned}$$

with $z_{k,e} := X_k + X_{k+m}$.

Analogously,

$\boxed{k = 2l+1,\ l = 0,\dots,m-1}$

$$\begin{aligned} c_{2l+1} &= \frac{1}{\sqrt{N}} \sum_{j=2l-m+2}^{m+2l+1} X_j \omega_N^{(2l+1)j} = \frac{1}{\sqrt{N}} \sum_{j=2l-m+2}^{2l+1} \left[(X_j - X_{j+m}) \omega_N^j \right] \omega_m^{lj} \\ &= \frac{1}{\sqrt{N}} \sum_{j=2l-m+2}^{2l+1} z_{j,o} \omega_m^{lj}, \end{aligned}$$

with $z_{j,o} := (X_j - X_{j+m})\omega_N^j$.

10.3 Complexity and the reason there is no 'fast' transform

As k ranges from 1 to T, T Fourier coefficients need to be calculated. Let p designate the 'splitting' step we look at. The first splitting step ($p = 1$) is performed in the previous section for odd as well as even indices of the Fourier coefficients. With the definitions $z_{k,e} := X_k + X_{k+m}$ and $z_{k,o} := (X_k - X_{k+m})\omega_N^k$, we can then write each of the Fourier coefficients as a sum of length m of either z_o's (if the index k is odd) or z_e's (if the index k is even).
Hence, a coefficient, for example, c_1 with a sum of only z_o's is followed by a coefficient, c_2 with a sum of only z_e's and so on.
We now continue the construction principle of the fast Fourier transform – the splitting of sums and reducing to Fourier transforms of lower order. For the sake of simplicity, we consider only the case of k being odd. The even case works analogously.
As we continue splitting, each sum of length m of z_o's is split and then rearranged to a sum of length $\frac{m}{2} = \frac{N}{4}$ of either $z_{o,e}$'s or $z_{o,o}$'s, where

$$z_{k,o,e} := z_{k,o} + z_{k+\frac{m}{2},o}, \qquad z_{k,o,o} := (z_{k,o} - z_{k+\frac{m}{2},o})\omega_{\frac{m}{2}}^k,$$

depending on whether $\frac{(k-1)}{2}$ is even (first case) or odd (second case). Figure 10.1 symbolizes the possible combinations of evens and odds for the first 4 steps:

	$p=1$	$p=2$	$p=3$	$p=4$
c_1	o	oe	oee	oeee
c_2	e	eo	eoe	eoee
	o	oo	ooe	ooee
·	e	ee	eeo	eeoe
·	o	oe	oeo	oeoe
·	e	eo	eoo	eooe
	o	oo	ooo	oooe
	e	ee	eee	eeeo
	o	oe	oee	oeeo
	e	eo	eoe	eoeo
	o	oo	ooe	ooeo
	e	ee	eeo	eeoo
	o	oe	oeo	oeoo
	e	eo	eoo	eooo
	o	oo	ooo	oooo
	e	ee	eee	eeeo
	o	oe	oee	oeee
	e	eo	eoe	eoee
	o	oo	ooe	ooee
	e	ee	eeo	eeoe
	o	oe	oeo	oeoe
	e	eo	eoo	eooe
	o	oo	ooo	oooe
	e	ee	eee	eeeo
	⋮	⋮	⋮	⋮
	o	oo	ooe	ooeo
	e	ee	eeo	eeoo
	o	oe	oeo	oeoo
	e	eo	eoo	eooo
	o	oo	ooo	oooo
c_T	e	ee	eee	eeee

Figure 10.1: Suggestive graphic of the indices used for $z_{k,\cdot}$, $k = 1, \ldots, T$.

The more splitting steps we perform, the longer it takes for a combination to repeat itself. In the second splitting step, for example, the combination *oe* is repeated every 2^2 times, whereas, in the third step, the combination *oee* is repeated every 2^3 times. The number of different outcomes is 2^p.

Let's say we are currently at step p and interested in all Fourier coefficients that can be split in a way that only $z_{e,e,e,\ldots e}$'s remain ($p \times e$). In this group, there is the Fourier coefficient c_m, as well as $c_{m+2^p}, c_{m+2\cdot 2^p}, c_{m+3\cdot 2^p}, \ldots$. All in all, there should be $\frac{T}{2^p}$ Fourier coefficients in this group.

Now, to determine the number of calculations to be done to obtain all Fourier coefficients for this group in step p, we start with the easiest case ($p = 1$):

Each of the m calculations of z_e's in the even case costs one complex addition, whereas each of the m calculations of a single z_o in the odd case costs – at first glance – one complex addition and one complex multiplication.

In Figure 10.2 all summands involved in the calculation of two subsequent odd-indexed Fourier coefficients c_{2l_0+1} and c_{2l_0+3} (l_0 fixed) are explicitly listed.
Given $z_{2l_0-m+2,\,o}, \ldots, z_{2l_0+1,\,o}$, we can see from Figure 10.2 that we only need to invest 2 further complex multiplications (and complex additions) in order to obtain all the z_o's needed to construct c_{2l_0+3}. For the very first odd coefficient, however, we need to calculate all z_o's, which costs $\frac{N}{2}$ complex multiplications (and the same amount of complex additions).

The number of complex multiplications needed to calculate all z_o's totals

$$\frac{N}{2} + 2 \cdot \left(\frac{T}{2}\right) = \frac{N}{2} + T$$

Same applies of course, for the complex additions. That was for the first step $p = 1$.

Now, for the p-th step, the idea remains the same, but we have already noticed, that the Fourier coefficients yielding the same output (i.e. sums of z's with the same index) – figuratively spoken (in view of Figure 10.1)– 'move further and further apart' with every splitting step. As we don't use the same data for each transformation, but shift, the number of common elements of two Fourier coefficients yielding the same output gets less and less. Accounting for this the number of z's to be calculated additionally in each further step is no more only 2 (as in step $p = 1$), but depends on p and is equal to 2^p – which is the reason why we fail to create a faster algorithm.

For our convenience, we will refer to Fourier coefficients, which – splitted p-times – in a way that they can be expressed with sums containing only $z_{e,e,e,\ldots e}$'s ($p \times e$) as Fourier coefficients of group $eeeee....e$.
The p-th letter is an e, indicating that the z's are created solely by one complex addition (without complex multiplication)

$$
\begin{aligned}
\sqrt{N}c_{2l_0+1} \;=\; & \underbrace{\left[(X_{2l_0-m+2}-X_{2l_0+2})\omega_N^{(2l_0-m+2)}\right]}_{=z_{2l_0-m+2,\,o}}\omega_m^{l_0(2l_0-m+2)} + \underbrace{\left[(X_{2l_0-m+3}-X_{2l_0+3})\omega_N^{(2l_0-m+3)}\right]}_{=z_{2l_0-m+3,\,o}}\omega_m^{l_0(2l_0-m+3)} \\
&+ \left[(X_{2l_0-m+4}-X_{2l_0+4})\omega_N^{(2l_0-m+4)}\right] + \left[(X_{2l_0-m+5}-X_{2l_0+5})\omega_N^{(2l_0-m+5)}\right] \\
&+ \left[(X_{2l_0-m+6}-X_{2l_0+6})\omega_N^{(2l_0-m+6)}\right] + \left[(X_{2l_0-m+7}-X_{2l_0+7})\omega_N^{(2l_0-m+7)}\right] \\
&\vdots \\
&+ \underbrace{\left[(X_{2l_0}-X_{2l_0+m})\omega_N^{(2l_0)}\right]}_{=z_{2l_0,\,o}}\omega_m^{l_0(2l_0)} + \underbrace{\left[(X_{2l_0+1}-X_{2l_0+m+1})\omega_N^{(2l_0+1)}\right]}_{=z_{2l_0+1,\,o}}\omega_m^{l_0(2l_0+1)} \\
\sqrt{N}c_{2l_0+3} \;=\; & \left[(X_{2l_0-m+4}-X_{2l_0+4})\omega_N^{(2l_0-m+4)}\right]\omega_m^{(l_0+1)(2l_0-m+4)} + \left[(X_{2l_0-m+5}-X_{2l_0+5})\omega_N^{(2l_0-m+5)}\right]\omega_m^{(l_0+1)(2l_0-m+5)} \\
&+ \left[(X_{2l_0-m+6}-X_{2l_0+6})\omega_N^{(2l_0-m+6)}\right]\omega_m^{(l_0+1)(2l_0-m+6)} + \left[(X_{2l_0-m+7}-X_{2l_0+7})\omega_N^{(2l_0-m+7)}\right]\omega_m^{(l_0+1)(2l_0-m+7)} \\
&\vdots \\
&+ \left[(X_{2l_0}-X_{2l_0+m})\omega_N^{(2l_0)}\right]\omega_m^{(l_0+1)(2l_0)} + \left[(X_{2l_0+1}-X_{2l_0+m+1})\omega_N^{(2l_0+1)}\right]\omega_m^{(l_0+1)(2l_0+1)} \\
&+ \left[(X_{2l_0+2}-X_{2l_0+2+m})\omega_N^{(2l_0+2)}\right]\omega_m^{(l_0+1)(2l_0+2)} + \left[(X_{2l_0+3}-X_{2l_0+m+3})\omega_N^{(2l_0+3)}\right]\omega_m^{(l_0+1)(2l_0+3)}
\end{aligned}
$$

Figure 10.2: Comparison of summands used for calculation of two succeeding odd indexed moving Fourier coefficients

If we now exemplarily consider the number of complex multiplications needed to construct all elements of group *oooo...o*, we end up with

$$\frac{N}{2^p} + 2^p \cdot \left(\frac{T}{2^p}\right)$$

- $\frac{N}{2^p} \mathrel{\widehat{=}}$ number of initially to be calculated z's, when no element of group *oooo...o* has been calculated yet.
- $2^p \mathrel{\widehat{=}}$ number of z's that have yet to be calculated, given we have already calculated an element of group *oooo...o*
- $\frac{T}{2^p} \mathrel{\widehat{=}}$ number of Fourier coefficients in group *oooo...o*

complex multiplications and the same number of additions in the p-th step.

Putting the results together, we obtain the following:

- T Fourier coefficients (i.e. sums of length $N = 2^q$) need to be calculated.
- We split each sum p times, in order to be finally left with only one summand – the Fourier coefficient. So $q = \log_2 N$ is the number of splitting steps.
- We have 2^p different *oe*-index-combinations after the $p-$th step.
- For the p-th step, the costs for one group total

  $$\left(\frac{N}{2^p} + 2^p \cdot \left(\frac{T}{2^p}\right)\right) \qquad \text{(multiplication only)}$$

  As we split q times, we have to accept a cost of

  $$\sum_{p=1}^{q} 2^p \left(\frac{N}{2^p} + 2^p \cdot \left(\frac{T-N}{2^p}\right)\right) = O(NT).$$

  So it doesn't actually help to exploit the benefit of 'reusable' (in the sense of: already calculated) elements. If we would have been able to detect an advantage at this point, we would also have had to numerically take into consideration, that the number of common elements (of Fourier coefficients in one group) decreases steadily and thus, at some step p, we don't have any overlapping anymore and have to calculate all $\frac{N}{2^p}$ z's for each Fourier coefficient in each group.

  Though we didn't get a computational advantage, we will write down the general formula for the sake of completeness.

If the last step with overlapping elements is $p = x$, the formula yielding the correct cost is

$$\sum_{p=1}^{x} 2^p \left(\frac{N}{2^p} + 2^p \cdot \left(\frac{T-N}{2^p} \right) \right) + \sum_{p=x+1}^{q} 2^p \left(\frac{N}{2^p} \cdot \left(\frac{T-N}{2^p} \right) \right) = O(NT).$$

Referring back to Section 10.1, we have now seen that the reason we fail to get a numerical advantage to the straight calculation of the moving Fourier transform by applying fast Fourier techniques, is the shifting. The ordinary Fourier transform, after having calculated the summands for the first two Fourier coefficients in one group, completely reuses the summands z for the remaining coefficients in this step and there is no exponentially growing amount of summands to be additionally calculated as in the previous procedure of the adapted fast Fourier transform.

10.4 A compromise between speed and locality

As we have figured out the problem that occurs when intending to adapt the idea of the fast Fourier transform, one might want to try out an alternative, which benefits from the reduction of complexity by creating reusable summands, but at the same time doesn't give up on the aspect of shifting.

Basic idea

Let again $1 \leq j \leq T$. The proposed algorithm of the moving Fourier transform implies that, after the calculation of one Fourier coefficient based on N data values, for example $X_1, \ldots, X_N$, the 'window' shifts and the next Fourier coefficient is calculated based on only almost the same data $X_2, \ldots, X_{N+1}$. So what we are doing is that we are, in a sense, shifting the catchment area of the Fourier transform by 1 unit after the calculation of each coefficient.
The idea of the compromising algorithm is not to shift by 1 unit, but by L units – and also not after the calculation of every single Fourier coefficient, but the shift ought to occur after having calculated L Fourier coefficients on the basis of the same data values.

Formulae

Choose $L := L(N)$ in a way that $\frac{L(N)}{N} \to 0$ and $L(N) \to \infty$ for $N \to \infty$. The calculation of the Fourier transform in this adapted way is suggestively displayed in Figure 10.3.
Note: If $\frac{T}{L}$ is not an integer, the last block is shorter and complexity somewhat smaller.

Now what we have is basically $\frac{T}{L}$ times a Fourier transform of N values.

Case 1: $L \geq 2^p$ As long as $L \geq 2^p$, which means that each possible outcome of the p-th step is covered at least once, we can refer to the ordinary Fourier transform of N values. For simplicity let $L := 2^x$ and we get, for $p \leq x$

$$\frac{T}{L}\sum_{p=1}^{x} 2^p \frac{N}{2^p} = O\left(\frac{TN\log_2 L}{L}\right).$$

Case 2: $L < 2^p$ However, as soon as L is no longer covering all outcomes, which means as soon as $p > x$, we can no longer resort to the idea of $\frac{T}{L}$ separate Fourier transforms, but we have to revisit the previous procedure:

- 2^p different groups
- $\frac{T}{2^p}$ elements in each group
- Taking into consideration the shift on indices: Having calculated one element of the group, we need, for any other element of the group, $\left\lfloor \frac{2^p}{L} \right\rfloor \cdot L$ operations.

$$\sum_{p=x+1}^{q} 2^p \left(\frac{N}{2^p} + \left\lfloor \frac{2^p}{L} \right\rfloor \cdot L \cdot \frac{T}{2^p} - 1\right)$$

So the complexity is now

$$\begin{aligned} &\frac{T}{L}\left(\sum_{p=1}^{x} 2^p \frac{N}{2^p}\right) + \sum_{p=x+1}^{q}\left(2^p \left(\frac{N}{2^p} + \left\lfloor \frac{2^p}{L} \right\rfloor \cdot L \cdot \frac{T}{2^p} - 1\right)\right) \\ = \; & O\left(\frac{TN\log_2 L}{L} + (\log_2 N - \log_2 L)N + \frac{NT}{L}\right). \end{aligned}$$

This encloses the cases of

- $L = 1$ (maximal locality, high complexity) – the moving Fourier transform ($\rightarrow O(NT)$), and
- $L = N$ (minimal locality, low complexity) – the original Fourier transform of $\frac{T}{N}$ blocks of length N ($\rightarrow O(T\log_2 N)$).

Therefore, this method achieves a reduction in complexity for $L > 1$ – to the cost of locality.

$1st$ block of length L :

$$c_1 = \frac{1}{\sqrt{N}} \sum_{k=1}^{N+1} X_{k-\frac{N}{2}} \omega_N^{1\cdot k}, \quad c_2 = \frac{1}{\sqrt{N}} \sum_{k=1}^{N+1} X_{k-\frac{N}{2}} \omega_N^{2\cdot k}, \ldots, \quad c_L = \frac{1}{\sqrt{N}} \sum_{k=1}^{N+1} X_{k-\frac{N}{2}} \omega_N^{L\cdot k}$$

$2nd$ block of length L :

$$c_{L+1} = \frac{1}{\sqrt{N}} \sum_{k=1+L}^{N+1+L} X_{k-\frac{N}{2}} \omega_N^{(L+1)k}, \quad c_{L+2} = \frac{1}{\sqrt{N}} \sum_{k=1+L}^{N+1+L} X_{k-\frac{N}{2}} \omega_N^{(L+2)k}, \quad \ldots, \quad c_{2L} = \frac{1}{\sqrt{N}} \sum_{k=1+L}^{N+1+L} X_{k-\frac{N}{2}} \omega_N^{(2L)k}$$

$$\vdots$$

$\frac{T}{L}th$ block of length L :

$$c_{T-L+1} = \frac{1}{\sqrt{N}} \sum_{k=T-N+1}^{T} X_{k-\frac{N}{2}} \omega_N^{(T-L+1)k}, \quad c_{T-L+2} = \frac{1}{\sqrt{N}} \sum_{k=T-N+1}^{T} X_{k-\frac{N}{2}} \omega_N^{(T-L+2)k},$$

$$\ldots, \quad c_T = \frac{1}{\sqrt{N}} \sum_{k=T-N+1}^{T} X_{k-\frac{N}{2}} \omega_N^{(T)k}$$

Figure 10.3: Construction of moving Fourier coefficients with shift L

CHAPTER 11

Simultaneous confidence bands for the autocorrelation

When estimating statistical quantities, one does also want, as Neumann and Polzahl [39] put it, to give a visual impression of the adequacy and variability of the estimation. This can be done by the presentation of confidence intervals for the values of interest. When estimating functions, however, it does not suffice to provide pointwise confidence intervals for each function value, as the main focus is most of the time on the overall shape of the curve and not the reliabilities of single values. Visualizations of such pointwise confidence bands will most likely also lead to a wrong interpretation by the user of statistical evaluations. It is therefore of great interest to provide uniform or simultaneous confidence bands when estimating functions to offer an easy and intuitive understanding of the preciseness of the estimation.

Why would one want to study autocorrelation functions? Autocorrelation of a time series means that values yet to come depend on past values. Autocorrelation sometimes eases predictions, indicating some persistance in systems meaning that some states perservere for additional time-units as the system is quite inertial. Some examplary time series can be seen in the field of hydrometerology: Garen and Pagano [42] analyze April – September streamflow volume data from 141 unregulated basins in the western United States for trends in persistence. Decadal time-scale changes in lag-1-year autocorrelation (persistence) were observed. The 1930s – 50s was a period of low variability and high persistence, the 1950s – 70s was a period of low variability and antipersistence, and the period after 1980 was highly variable and highly persistent. In particular, regions from California and Nevada to southern Idaho, Utah, and Colorado have recently experienced an unprecedented sequence of consecutive wet years along with multiyear extreme droughts.

Paying attention to autocorrelation functions is not only of major interest against physical background, but also in economic settings: Autocorrelation in stock returns is used as one important measure of securities market pricing. To monitor the autocorrelation of stock returns closely is important, as it may be a sign of genuine

pricing inefficiency (see Anderson [1]).
Building on the example of autocorrelation – estimating the autocorrelation function and constructing confidence bands for the first order autocorrelation – we point out the practical relevance of estimators grasping the time-varying structures occurring in time series. We advertise our method of the moving Fourier transform, the moving periodogram and correspondingly adapted bootstrap procedures to meet these needs.

11.1 Design of simultaneous confidence bands for the autocorrelation

Confidence bands for a time varying autocorrelation function have hardly been studied in literature.
Sergides [49] constructs pointwise confidence bands for the time varying autocorrelation function of a tvMA(1)-process. These pointwise intervals are of variable width and are calculated by adding (and substracting) the bootstrap estimate of the standard deviation times the theoretical quantiles of the standard normal distribution to the estimated autocorrelation function.
Kreiss and Paparoditis [32] do also construct pointwise confidence bands by using the same approach as Sergides [49] employing their hybrid bootstrap method. However, they do not provide any further simulation study but merely apply their bootstrap method to give a numerical example.
As pointed out before, it would be far more reasonable and also more intuitive from a practitioner's point of view to provide simultaneous confidence bands for estimated functions. This is a problem, which is, up to now, mainly addressed in nonparametric regression, where simultaneous confidence bands are constructed for the regression function (see Sun and Loader [52] and Neumann and Polzehl [39]).

There are two basic approaches to construct simultaneous confidence bands: either with fixed or with variable width. For the situation of nonparametric regression, an easy to understand description of how to proceed in either case is given by Neumann and Polzehl [39]. There is also a fairly good manual of how to construct simultaneous confidence bands with variable width using bootstrapping in Lenhoff et al. [34].
We now describe two methods of constructing simultaneous confidence bands for the autocorrelation $\rho(u,h)$, $u \in [0,1]$, of the locally stationary time series $\{X_{t,T}\}$. W.l.o.g. we will restrict ourselves to $h = 1$, that is the 1-lag autocorrelation.
Firstly, the construction of a confidence band of variable width is considered. We aim to use bootstrapping in order to mimic the behaviour of the process

$$\left\{ \frac{\rho(u,1) - \hat{\rho}(u,1)}{\hat{\sigma}_\rho(u)} \right\}_{u \in [0,1]} .$$

This process specifies the maximal weighted deviation of the autocorrelation from the estimate $\hat{\rho}(u,1)$, for all $u \in [0,1]$.

Step 1: Choose a suitable lattice $L[0,1]$ on the interval $[0,1]$. For example, $\{t/T, t = 1,\ldots,T\}$. For every fixed $u \in L[0,1]$ calculate an estimate $\hat{\rho}(u,1)$ of the autocorrelation function.

Step 2: Generate B bootstrap time series by employing a moving bootstrap. For each time series $\left\{X_{t,T}^{*,b}\right\}$, $b = 1,\ldots,B$, estimate the autocorrelation $\hat{\rho}^b(u,1)$ for every $u \in L[0,1]$.

Step 3: Use all B time series to estimate the standard deviation $\hat{\sigma}_{\hat{\rho}}(u)$ of $\hat{\rho}$ for every $u \in L[0,1]$.

Step 4: Choose $C_{boot} > 0$ such that

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left\{\max_{u\in L[0,1]} \frac{\hat{\rho}(u,1)-\hat{\rho}^b(u,1)}{\hat{\sigma}_\rho(u)} \leq C_{boot}\right\} \geq 1-\alpha,$$

for some prescribed α, $0<\alpha<1$.

The simultaneous $\alpha \cdot 100\%$ confidence band for $\rho(u,1)$, $0 \leq u \leq 1$, is then given by

$$CB_{variable} := [\hat{\rho}(u,1) - C_{boot}\cdot\hat{\sigma}_\rho(u),\ \hat{\rho}(u,1) + C_{boot}\cdot\hat{\sigma}_\rho(u)].$$

In order to maintain uniform size of the confidence band, one simply omits the third step of the above algorithm and adapts the fourth step. Doing so, we hence mimic the process of the maximal deviation of the autocorrelation from the estimate $\hat{\rho}(u,1)$, for all $u \in [0,1]$.

$$\{\rho(u,1) - \hat{\rho}(u,1)\}_{u\in[0,1]}\,.$$

Step 1: Create a suitable lattice $L[0,1]$ on the interval $[0,1]$. For every fixed $u \in L[0,1]$ calculate an estimate $\hat{\rho}(u,1)$ of the autocorrelation function.

Step 2: Generate B bootstrap time series by employing a moving bootstrap. For each time series $\left\{X_{t,T}^{*,b}\right\}$, $b = 1,\ldots,B$, estimate the autocorrelation $\hat{\rho}^b(u,1)$ for every $u \in L[0,1]$.

Step 4': Choose $C'_{boot} > 0$ such that

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left\{\max_{u\in L[0,1]} \{\rho(u,1)-\hat{\rho}^b(u,1)\} \leq C_{boot}\right\} \geq 1-\alpha,$$

for some prescribed α, $0<\alpha<1$.

The simultaneous $\alpha \cdot 100\%$ confidence band for $\rho(u,1)$, $0 \leq u \leq 1$, is then given by

$$CB_{fixed} := [\hat{\rho}(u,1) - C'_{boot}, \hat{\rho}(u,1) + C'_{boot}].$$

In our simulations, the moving bootstrap method, as referred to in Step 2, will either be the moving TFT-bootstrap, the moving autoregressive aided periodogram bootstrap or the moving wild hybrid bootstrap.

11.2 Simulation study

The study is structured as follows: At first, we will simulate different types of locally stationary processes. Those processes vary with respect to the model structure and the distribution of the white noise. We will consider a tvAR(1)-process with linearly changing coefficients, as well as a tvMA(1)-process as used by Sergides [49]. Concerning the white noise, we will study standard normal errors, standardized χ^2- as well as standardized exponentially distributed errors. Definition 2.1 merely prescribes that $E\varepsilon_t = 0$, $E\varepsilon_t^2 = 1$, as well as $E\varepsilon_t^4 < \infty$ which is fulfilled after appropriate centering and rescaling of the errors.

DGP 1 (time-varying AR(1)-process)

$$X_{t,T} = a_{t,T} \cdot X_{t-1,T} + \varepsilon_t,$$

with $a_{t,T} = \left(1 - \frac{t}{T}\right) \cdot (-0.6) + \frac{t}{T} \cdot 0.6$ and ε_t independent and identically distributed for all $t = 1, \ldots, T$.

DGP 2 (time-varying MA(1)-process)

$$X_{t,T} = 1.1 \cdot \cos\left(1.5 - \cos\left(\frac{4\pi\, i}{T}\right)\right) \cdot \varepsilon_{t-1} + \varepsilon_t,$$

with ε_t independent and identically distributed for all $t = 1, \ldots, T$.

The following arrangements will be considered:

		Error distribution		
		$\mathcal{N}(0,1)$	$5 \cdot Exp(5) - 1$	$\frac{\chi_3^2 - 3}{\sqrt{6}}$
Model	tvAR(1)	DGP1a	DGP1b	DGP1c
	tvMA(1)	DGP2a	DGP2b	DGP2c

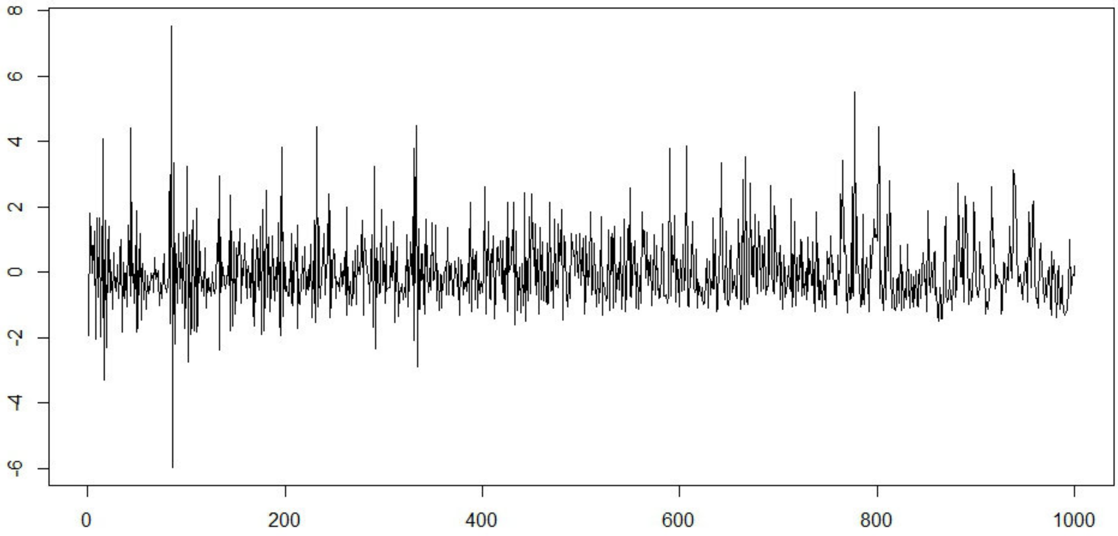


Figure 11.1: A realization of DGP 1c

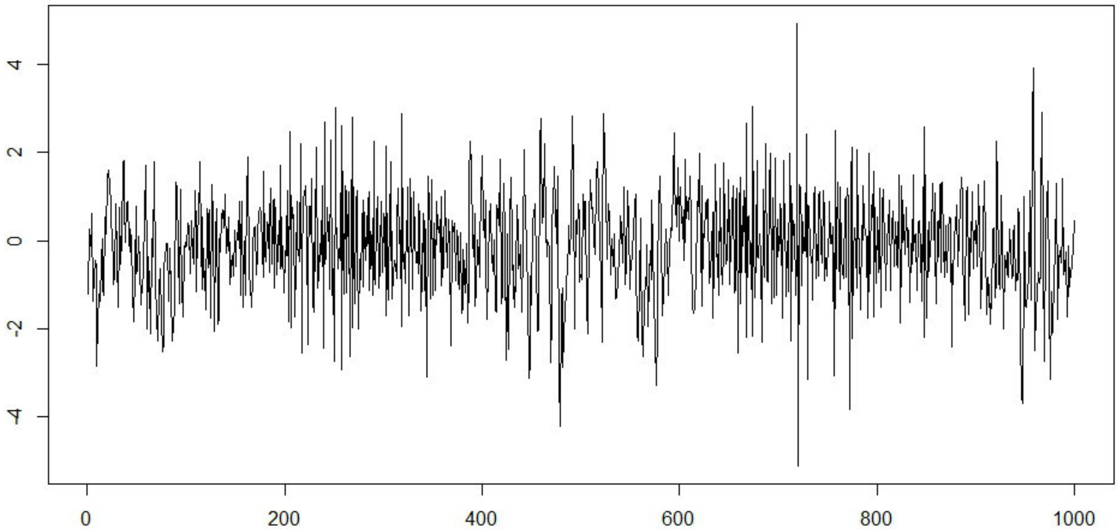


Figure 11.2: A realization of DGP 2a

To account for the boundary effects which occur as we don't use the moving versions of the bootstrap for the very first and the very last $N = 201$ observations, we only evaluate the simulations in between $t = 200$ and $t = 800$, that is at 601 points in time. This is in agreement with Sergides [49]. The following graphics, though, display the whole range of $t = 1$ to $t = 1000$. One can clearly see – for example in Figure 11.3 – the effect of the blockwise bootstrap in the beginning and at the end. We have constructed the confidence bands to a coverage of 95%. In order to verify whether the confidence bands actually meet the intended coverage probability, we calculate the empirical coverage probability using $R = 200$ repetitions. As we work with simultaneous confidence bands, the question is how to characterize a curve to lie in between two other curves. Should we require all points of the curve to be in between, do we allow for some percentage of points that can be outside?

The first check is whether the empirical coverage is dependent on the data. This is done by looking at different seeds. The resulting empirical coverage probability of the theoretical autocorrelation function at lag 1 of a time-varying AR(1)-process with standard normal iid errors (DGP 1a) is given in the following table. The theoretical curve $\rho(1)$ is considered to lie within the confidence band if all values between $t = 200$ and $t = 800$ are within the confidence band.

		Width		Type of bootstrap		
		variable	fixed	mTFT	mAAPB	mH
Seed	1:200	x		0.960	0.99	1.000
	1:200		x	0.970	0.995	1.000
	201:400	x		0.980	0.995	1.000
	201:400		x	0.985	0.995	1.000

We can see that for different seeds, the moving version of the TFT-bootstrap is the most volatile. We should as a rule of thumb consider random deviations of ± 0.01 before drawing conclusions.

Having mentioned the difficult question of clarifying when the theoretical curve $\rho(1)$ is considered to lie within the confidence band, we consider different numbers of points we allow to deviate. The first criterion is that all 601 points do have to lie within the bounds, secondly, only 590 of the 601 points need to be in the band. The most is a miss by 100 points, which is 17% of the curve.

Empirical coverage probability based on DGP1a

		Width		Type of bootstrap		
		variable	fixed	mTFT	mAAPB	mH
$\geq$ bound	601	x		0.96	0.99	1.00
	601		x	0.97	1.00	1.00
	590	x		0.97	0.99	1.00
	590		x	0.98	1.00	1.00
	560	x		0.98	1.00	1.00
	560		x	0.99	1.00	1.00
	500	x		1.00	1.00	1.00
	500		x	1.00	1.00	1.00

We now perform the autoregressive-aided periodogram bootstrap as given by Sergides [49], still with a window width of $N = 201$, but calculating all N Fourier coefficients at each point in time u. Thus, for each u bootstrap replica of $I_N(u, \lambda_j)$ for each $j = 1, \ldots, N$ are produced. From those we obtain $\rho(u, 1)$ for every $u \in [0, 1]$. However, instead of constructing pointwise confidence bands as done by Sergides [49], we now construct uniform confidence bands proceeding as in Section 11.1. The algorithm of obtaining the bootstrap replicates has a complexity of $O(N^2T)$ and thus takes up much more time than construction of simultaneous confidence bands using the moving Fourier transform, which is only of order $O(NT)$. Given the computational ressources available, we have performed the autoregressive-aided periodogram bootstrap $B = 200$ times for each of the 56 trials in order to calculate the empirical coverage probability. For the standard normal iid errors, both, the fixed and the variable confidence band exhibit an empirical coverage of 100%.
A visual comparison of the moving version of the autoregressive-aided periodogram bootstrap and the original version can be found in Figure 11.3.

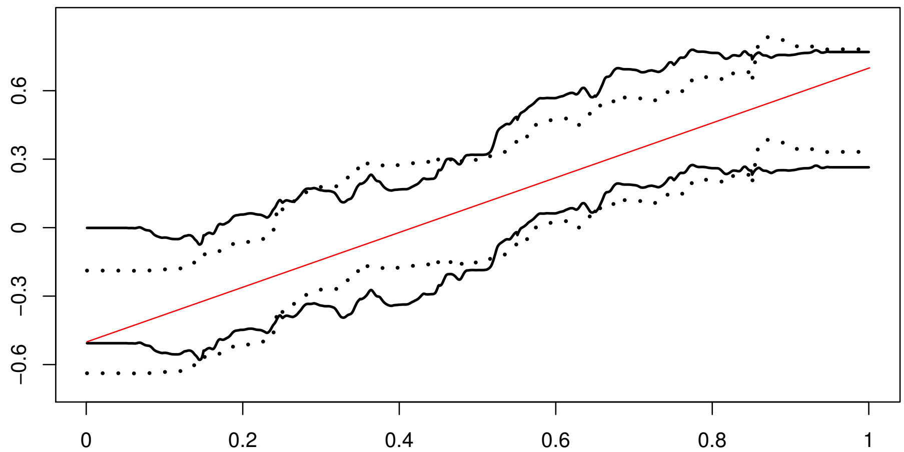


Figure 11.3: Confidence band of the mAAPB (solid) and the AAPB (dotted)

The resulting confidence bands have a mean width of 0.484 with a standard deviation of 0.023 compared to the mean width of 0.485 with standard deviation of 0.022 of the moving version. Both with an empirical coverage of 100%.

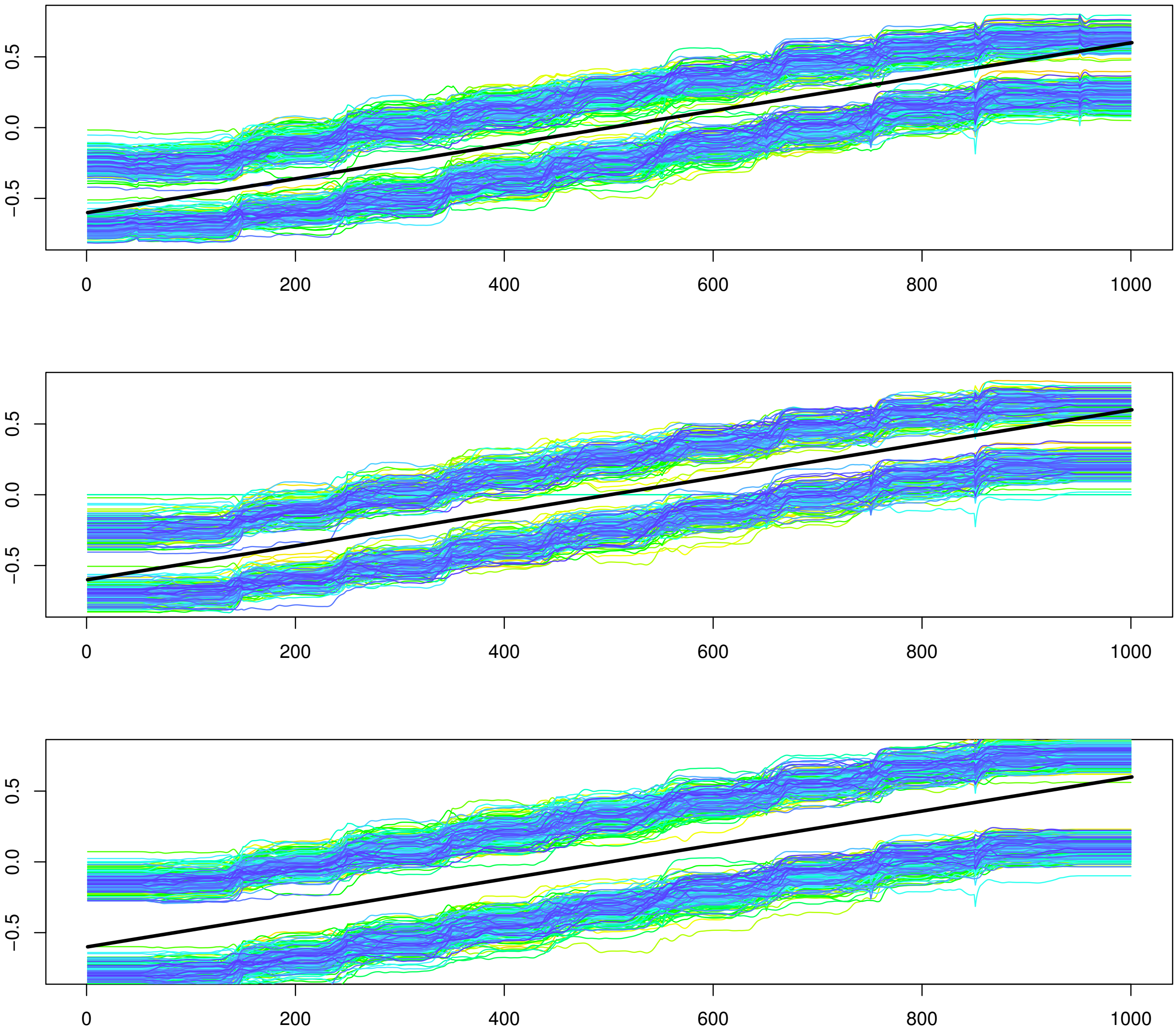


Figure 11.4: Confidence bands (variable width) of moving version of (a) TFT bootstrap, (b) AAP bootstrap and (c) wild hybrid bootstrap for DGP1a for different realizations

Empirical coverage probability based on DGP1 with different error distributions

			Width		Type of bootstrap		
			variable	fixed	mTFT	mAAPB	mH
DGP1b	$\geq$ bound	601	x		0.95	0.99	1.00
		601		x	0.97	0.99	1.00
		590	x		0.96	0.99	1.00
		590		x	0.98	0.99	1.00
		560	x		0.99	0.83	1.00
		560		x	1.00	1.00	1.00
		500	x		1.00	1.00	1.00
		500		x	1.00	1.00	1.00
DGP1c	$>$bound	601	x		0.97	0.99	1.00
		601		x	0.98	0.99	1.00
		590	x		0.99	0.99	1.00
		590		x	0.99	0.99	1.00
		560	x		1.00	1.00	1.00
		560		x	1.00	1.00	1.00
		500	x		1.00	1.00	1.00
		500		x	1.00	1.00	1.00

Empirical coverage probability based on DGP2a

		Width		Type of bootstrap		
		variable	fixed	mTFT	mAAPB	mH
$\geq$ bound	601	x		0.02	0.86	0.41
	601		x	0.04	0.72	0.56
	590	x		0.04	0.91	0.49
	590		x	0.07	0.80	0.66
	560	x		0.19	0.97	0.80
	560		x	0.26	0.97	0.88
	500	x		0.58	1.00	0.97
	500		x	0.78	1.00	1.00

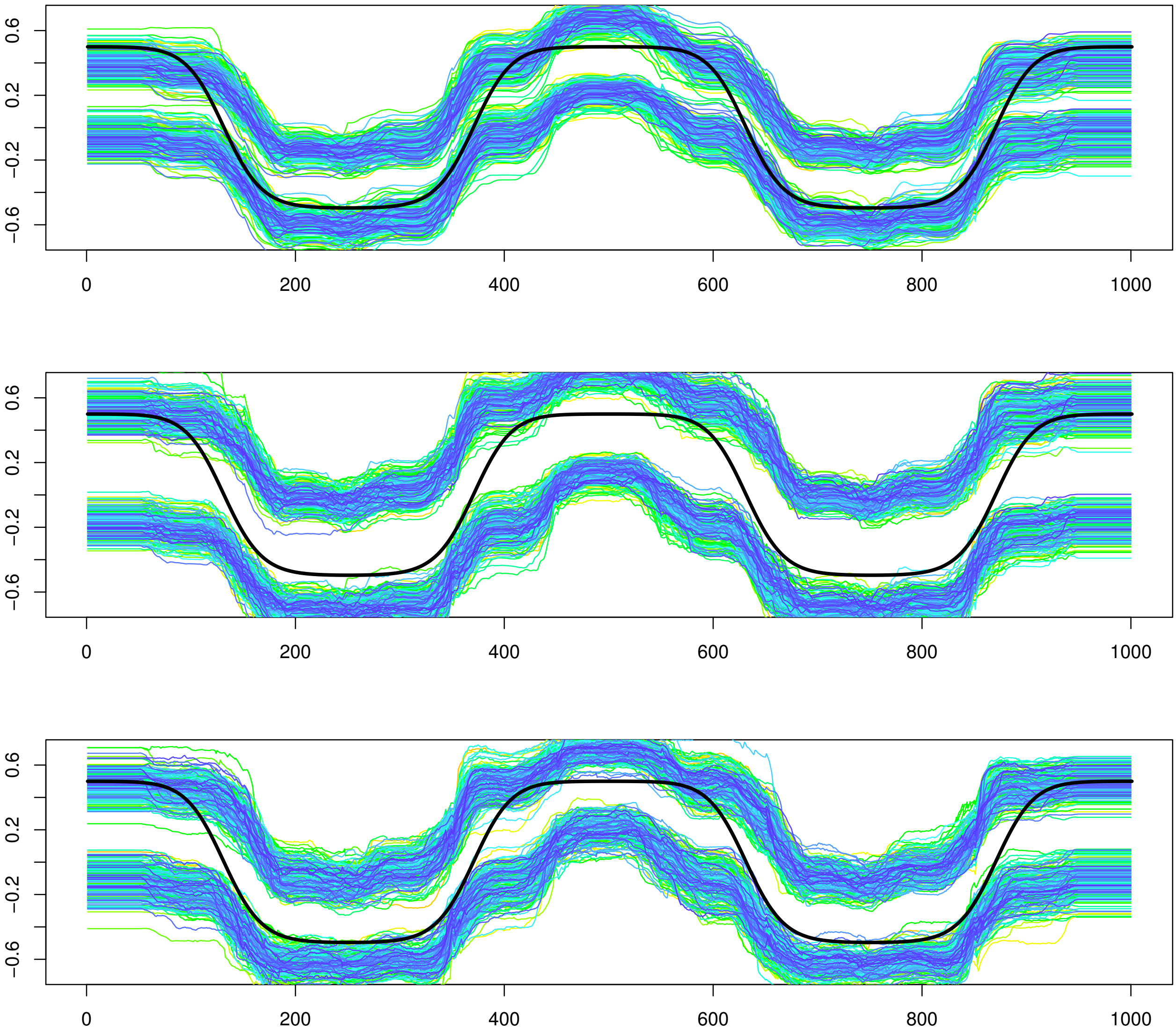


Figure 11.5: Confidence band of moving version of (a) TFT bootstrap, (b) AAP bootstrap and (c) wild hybrid bootstrap

Empirical coverage probability based on DGP2 with different error distributions

			Width		Type of bootstrap		
			variable	fixed	mTFT	mAAPB	mH
DGP2b	$\geq$ bound	601	x		0.01	0.90	0.23
		601		x	0.01	0.78	0.39
		590	x		0.02	0.94	0.32
		590		x	0.03	0.85	0.49
		560	x		0.07	0.99	0.71
		560		x	0.08	0.97	0.83
		500	x		0.18	1.00	0.95
		500		x	0.27	1.00	0.98
DGP2c	$\geq$ bound	601	x		0.02	0.91	0.32
		601		x	0.03	0.76	0.49
		590	x		0.03	0.95	0.43
		590		x	0.06	0.82	0.58
		560	x		0.16	0.99	0.75
		560		x	0.26	0.98	0.82
		500	x		0.61	1.00	0.95
		500		x	0.76	1.00	0.99

One might also want to check, how the pointwise asymptotic confidence bands perform (see Figure 11.6). From a simple look at one realization of the asymptotic 95%-confidence band compared to the true autocorrelation $\rho(u, 1)$ one can readily tell that this band will not be likely to get anywhere near an empirical coverage of 95%.

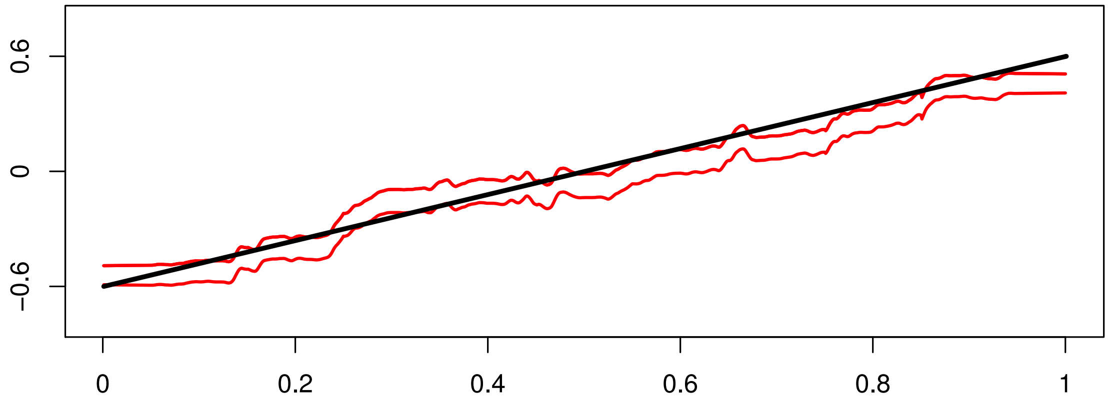


Figure 11.6: Pointwise asymptotic 95%-confidence band for DGP1a

A exemplary visual comparison of the variable and fixed width confidence band in the case of a χ^3_2-distribution of the errors constructed using the moving version of the TFT-bootstrap shows that there is not much difference between the two ways, though the table hints a higher empirical coverage rate of the method using a fixed width.

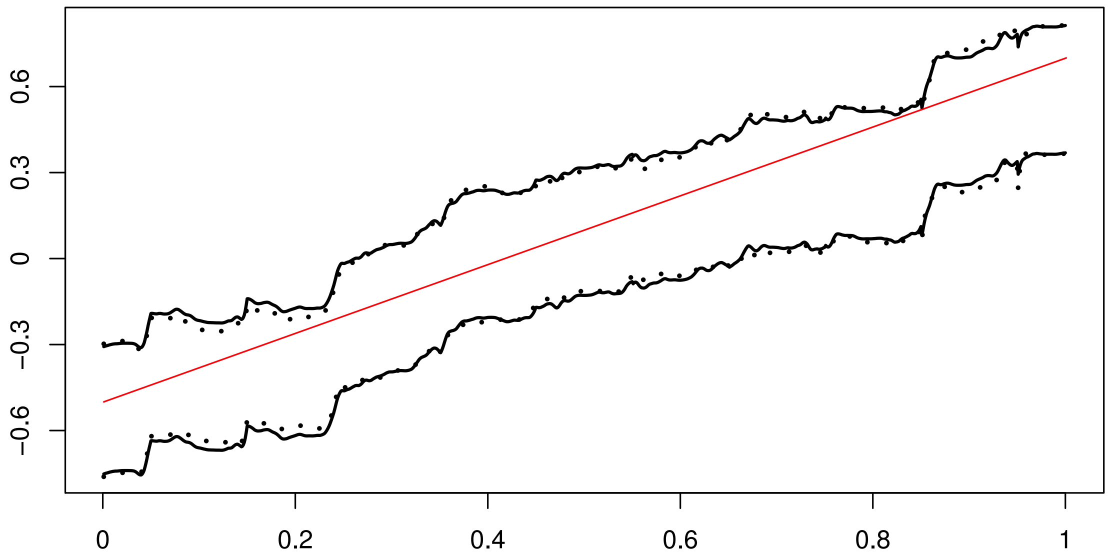


Figure 11.7: DGP 1c: Comparison of the simultaneous confidence bands of fixed (solid) and variable (dotted) width obtained via the moving TFT bootstrap

We have seen that the wild hybrid bootstrap exhibits a coverage of 100% for DGP1a. This is also reflected in the width of the confidence bands. We will now compare the confidence bands with fixed width and present the average width and the standard deviation of the widths of each procedure in the case of DGP1a and DGP2a.
Width of confidence bands

(width)	DGP1a		DGP2a	
	mean	std	mean	std
mTFT	0.4659180	0.01825032	0.4569330	0.01950276
mAAPB	0.4837308	0.02309174	0.7110781	0.04509634
mHB	0.6538740	0.06472582	0.6432225	0.03435888

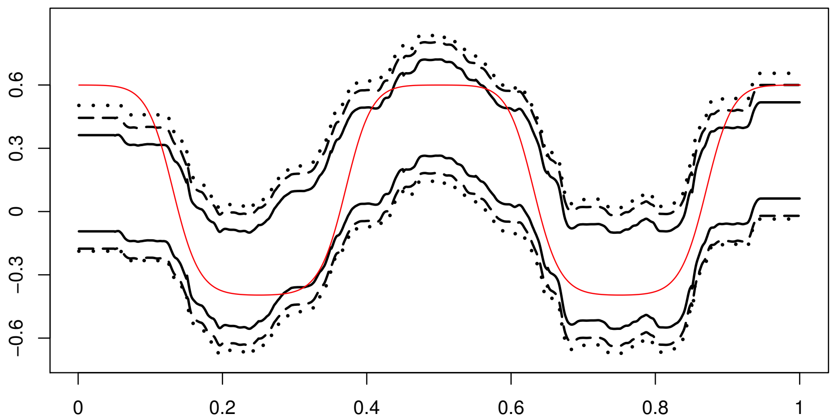


Figure 11.8: DGP 2a: Comparison of the simultaneous confidence bands of fixed width obtained via the moving TFT bootstrap (solid), the moving AAP bootstrap (dotted) and the moving wild hybrid bootstrap (dashed)

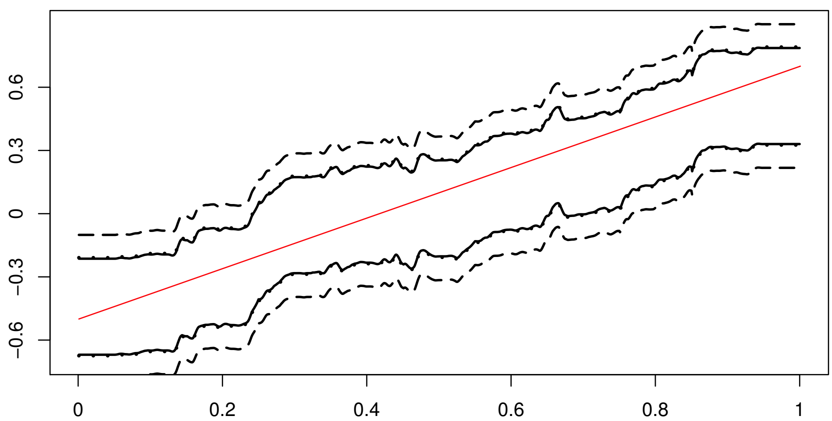
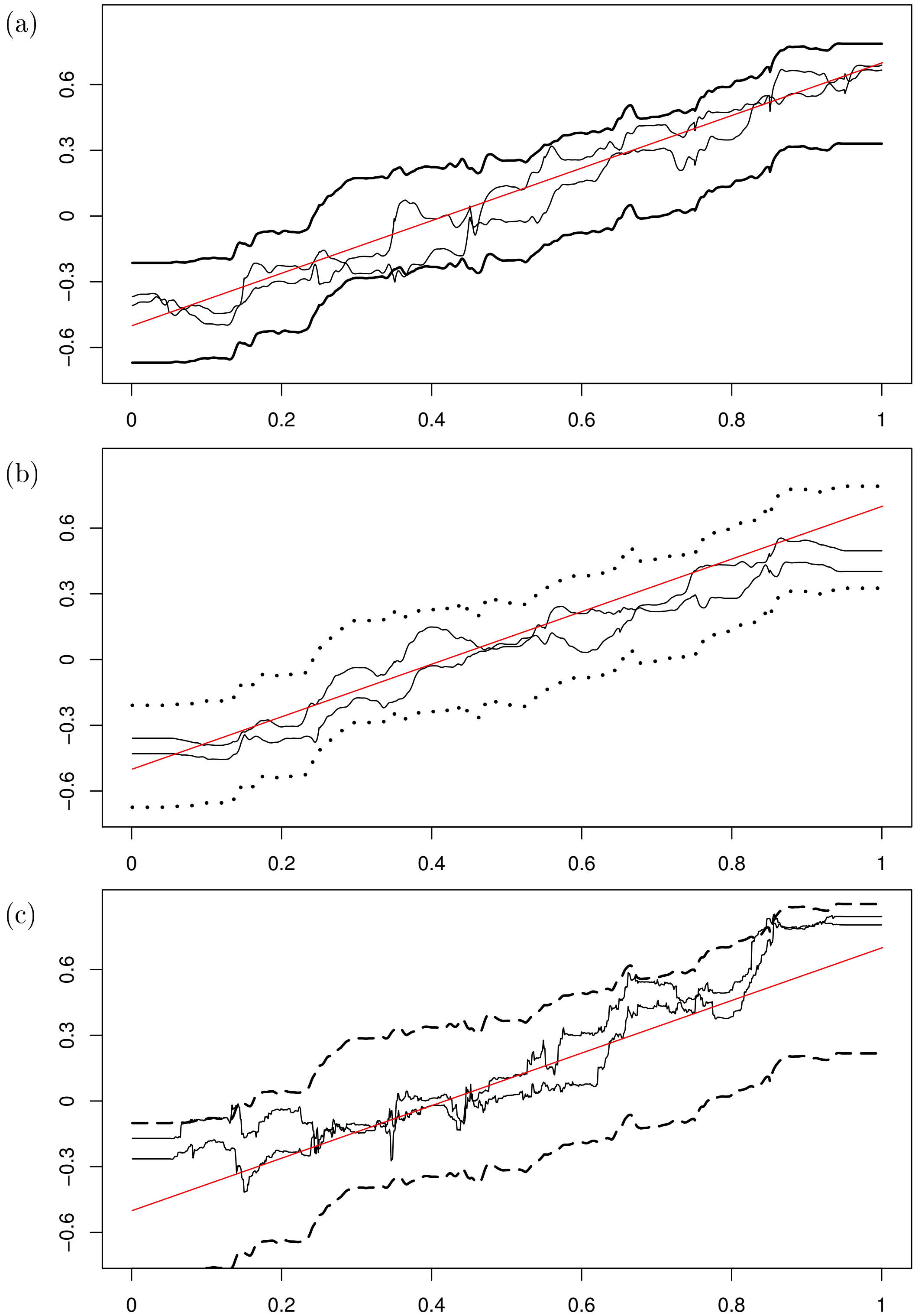


Figure 11.9: DGP 1a: Comparison of the simultaneous confidence bands of fixed width obtained via the moving TFT bootstrap (solid), the moving AAP bootstrap (dotted) and the moving wild hybrid bootstrap (dashed)

Figure 11.10: DGP 1a: Simultaneous 95% confidence bands of fixed width with two realizations each using (a) the moving TFT bootstrap (solid), (b) the moving AAP bootstrap (dotted) and (c) the moving wild hybrid bootstrap (dashed)



We conclude with a look at the autocovariance function. In Chapter 5, we have seen that the variance of the spectral mean incorporating the moving periodogram still depends on the fourth order cumulant. The moving TFT bootstrap has not been designed to bootstrap the fourth order cumulant of the data. However, we were interested in how well the bootstrap still works deviating from the standard normal distribution of the errors. In the following study we have used fixed width confidence bands of the autocovariance function of lag 1 (i.e. of the spectral mean with weight function $\varphi(\lambda) = e^{i\lambda}$). First, using standard normally distributed errors and, second, using standardized exponentially distributed errors. In the first case, we get an excess curtosis of zero, in the second of 6. The empirical coverage of the bands has been surprisingly good in the case of the high excess kurtosis.

	DGP1a	DGP1b
≥ 601	0.88	0.75
≥ 590	0.92	0.82
≥ 560	0.97	0.92
≥ 500	0.99	0.96

Exemplary confidence bands can be seen in the following figures.

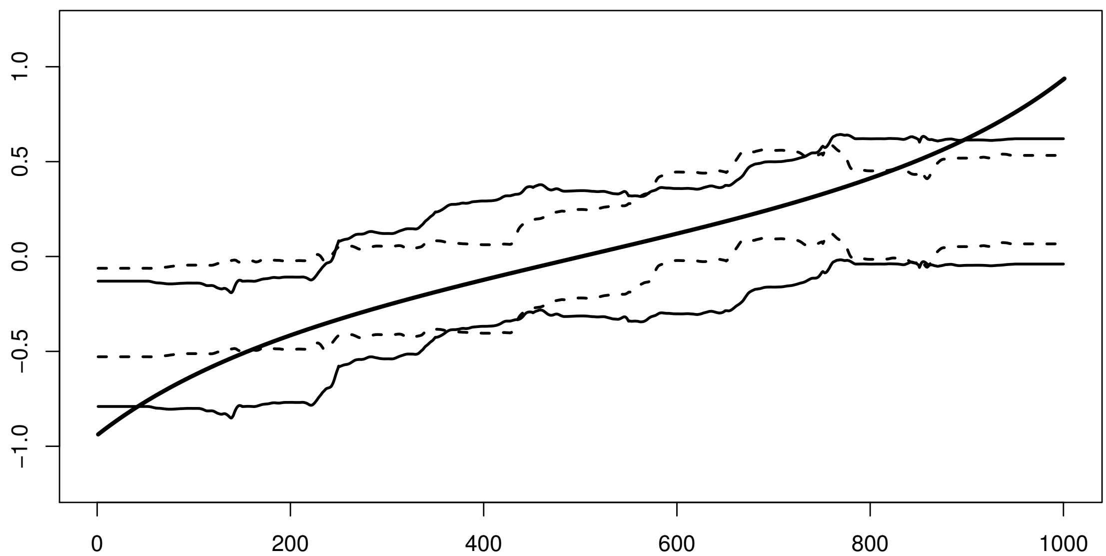


Figure 11.11: Exemplary simultaneous 95% confidence bands of fixed width with the moving TFT bootstrap for DGP1a (solid) and DGP1b (dashed)

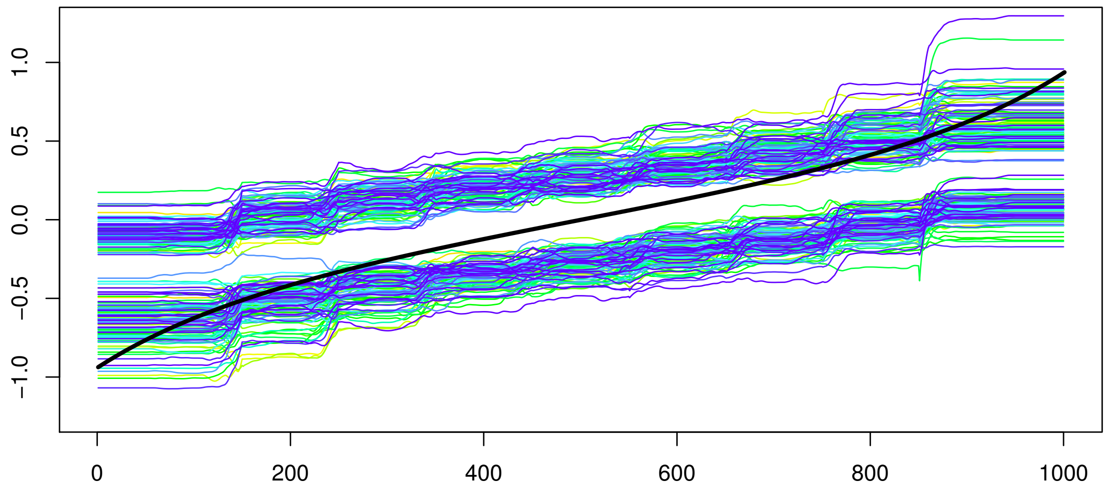


Figure 11.12: Bootstrap versions of simultaneous 95% confidence bands of fixed width with the moving TFT bootstrap for DGP1b

We eye the assumption that $\kappa_4 := E(\varepsilon_t)^4 - 3$ is unchanged throughout time. We don't see any reason for κ_4 to remain unchanged while the variance of the time series changes. Assuming that there is a change in κ_4, would the moving spectral density estimation help to mimick this changing κ_4? Or would it fail, just like the procedure in Kreiss and Paparoditis [32]. They estimate a single value of κ_4 using all data.
It is quite hard to consistently estimate the fourth moment. So what we did was to estimate the time varying autocovariance at lag 1, which is a spectral mean. The asymptotic covariance structure, both in our case (see Theorem 5.5) as well as in the situation when using the local periodogram as an estimator (Lemma 2.4.2 in Sergides [49]) is dependent on κ_4. We now estimate the autocovariance function of iid data with a fourth moment of $m_1 := 3$ for $t = 1, \ldots, 499$ and then switch to iid data with a fourth moment of 18 for $t = 500, \ldots, 1000$. For constructing the data, the construction made by Kreiss and Paparoditis [31] is used: E.g. for $t = 1, \ldots, 499$, $P(\varepsilon_t = \sqrt{m_1}) = P(\varepsilon_t = -\sqrt{m_1}) = \frac{1}{2m_1}$ and $P(\varepsilon_t = 0) = 1 - \frac{1}{m_1}$.
Being interested in how well the change in κ_4 is mimicked, we need to look at the variance of the estimated autocovariances.
When estimating the autocovariance using the moving Fourier transform for 200 different but identically distributed sets of iid random variables, we get 200 values at each time t. The estimated variance at each time can be seen in Figure 11.13. For illustrative reasons the average variance of the first 300 observations as well as of the last 300 observations is marked. It can clearly be seen that the variability of the estimation changes as time passes.

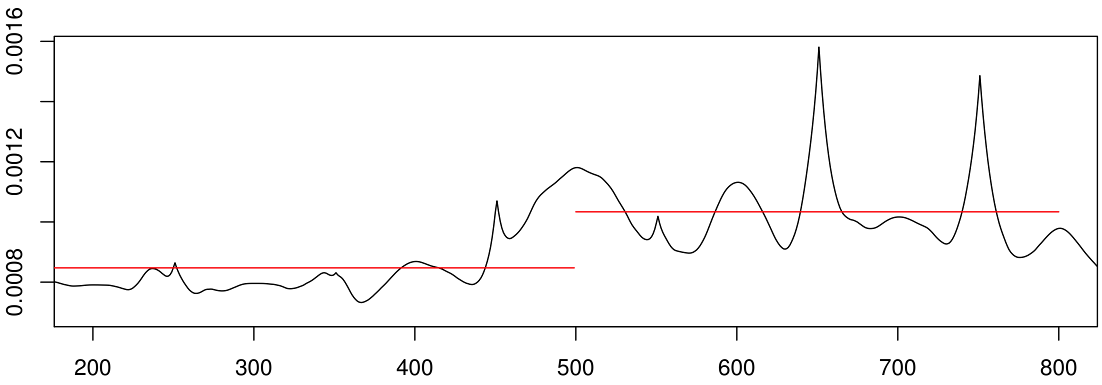


Figure 11.13: Sample variance of 200 realizations of the estimated autocovariance of lag 1 using the moving periodogram

Changes of the fourth moment are completely omitted by Kreiss and Parparoditis [32]. Their bootstrap procedure is therefore only applicable in the restricted setting of constant fourth moment of the innovations. Still, the bootstrap is able to copy the information on the fourth moment. We therefore propose the moving hybrid bootstrap as presented in Section 7.2.3 in order to be able to cover changes in the fourth moment.

CHAPTER 12

Conclusion and outlook

The new aspect of this thesis is the idea of shifting a local Fourier transform along a time series. It refines in an elegant and efficient way the common idea of applying a local Fourier transform to the data: At each time t, only one of the Fourier coefficients is updated. For slow changes, which are characteristical for locally stationary time series, this is an effective way to mirror structural changes. The very last section of Chapter 11 exemplarily shows that even changes in fourth moments can be traced. This interesting aspect of how changing fourth order cumulants can be monitored will certainly be of future interest. We contribute by proposing the moving wild hybrid bootstrap (see Section 7.2.3).

Using the moving Fourier transform, we have been able to develop a well-behaved and numerically cheap estimator for the time varying spectrum, which is locally uniform consistent, which means that the spectral density estimator at some time k in the neighbourhood of t converges to the true spectral density at time t, uniformly in k. This is the local equivalent to the condition required in the stationary setting. We may therefore extend all procedures involving spectral density estimation in the stationary setting to the locally stationary setting. This has explicitly been done for three bootstrap procedures in Chapter 8. We now have two methods, the moving TFT-bootstrap, as well as the moving wild hybrid bootstrap, to generate bootstrap observations of locally stationary data not just in the frequency domain, but also in the time domain. The moving autoregressive aided periodogram bootstrap only generates replicates in the frequency domain. Adapting the extension made by Jentsch and Kreiss [26] to the moving case, however, one could also obtain a moving autoregressive aided Fourier coefficient bootstrap which is able to generate bootstrap observations in the time domain.

The maintaining of the correct covariance structure of the bootstrap data has been proved in Chapter 10 exemplarily for the TFT-bootstrap. We have also investigated whether there is a possibility of being more efficient concerning the calculation of

the moving Fourier transform by introducing some kind of adapted fast Fourier transform. However, it turned out that when intending to reduce complexity, the transform needs to spend a longer period of time without shifting to the next stretch of data. It would therefore be desirable to investigate further methods of speeding up computations.
Referring to the aspect of only transforming a small set of data at a time, the question arises whether the choice of the window width can somehow be locally adapted to the degree of structural change. The question of an optimal choice of the window width has also not been answered yet. Concerning the spectral density estimation, future work will include the examination of the choice of kernel and bandwidth.
In Chapter 5 we have turned our attention to spectral means and provided asymptotic expectation and variance of those statistics. The next step will now be to explicitly prove the asymptotic normality, as explorative simulations have hinted that normal distribution is most likely. In a further step, one should look at ratio statistics and their properties. The simulation part of this thesis has already turned to solve this problem by bootstrapping. The local autocorrelation function of locally stationary processes is studied and we provide uniform bootstrap confidence bands, comparing different bootstrap approaches.

Bibliography

[1] ANDERSON, R. Time-varying risk premia and stock return autocorrelation. Tech. rep., UC Berkeley: Coleman Fung Risk Management Research Center, 2013.

[2] BRIGHAM, E. *The fast Fourier transform and its applications.* Prentica-Hall Inc., 1988.

[3] BROCKWELL, P. J., AND DAVIS, R. A. *Time Series:Theory and Methods.* Springer, 2006.

[4] BRUSCATO, A., AND TOLOI, C. M. Spectral analysis of non-stationary processes using the Fourier transform. *Brasilian Journal of Probability and Statistics 18* (2004), 69–102.

[5] CRAMÉR, H. On some classes of nonstationary processes. In *Fourth Berkeley Symposium Math. Statist. and Prob.* (1960), vol. 2, University of California Press, pp. 57–78.

[6] DAHLHAUS, R. Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In *Athens Conference on Applied Probability and Time Series Analysis*, P. Robinson and M. Rosenblatt, Eds., vol. II. Springer Verlag, 1996, pp. 145–159.

[7] DAHLHAUS, R. On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications 62*, 1 (1996), 139–168.

[8] DAHLHAUS, R. Fitting time series models to nonstationary processes. *The Annals of Statistics 25*, 1 (1997), 1–37.

[9] DAHLHAUS, R. A likelihood approximation for locally stationary processes. *The Annals of Statistics 28*, 6 (2000), 1762–1794.

[10] Dahlhaus, R. Curve estimation for locally stationary time series models. In *Recent Advances and Trends in Nonparametric Statistics*, M. G. Akritas and D. N. Politis, Eds. Elsevier, 2003.

[11] Dahlhaus, R. Locally stationary processes. *ArXiv e-prints* (2011). Provided by the SAO/NASA Astrophysics Data System.

[12] Dahlhaus, R., and Giraitis, L. On the optimal segment length for parameter estimates for locally ststionary time series. *Journal of Time Series Analysis 19*, 6 (1998), 629–655.

[13] Dahlhaus, R., and Neumann, M. Locally adaptive fitting of semiparametric models to nonstationary time series. *Stochastic Processes and their Applications 91* (2001), 277–308.

[14] Dahlhaus, R., and Polonik, W. Empirical spectral processes for locally stationary time series. Working Paper, 2007.

[15] Dahlhaus, R., and Polonik, W. Empirical spectral processes for locally stationary time series. *Bernoulli 15*, 1 (2009), 1–39.

[16] Dahlhaus, R., and Subba Rao, S. Statistical inference for time-varying ARCH-processes. *The Annals of Statistics 34*, 3 (2006), 1075–1114.

[17] de Boor, C. FFT as nested muiltiplication, with a twist. *Journal on Scientific and Statistical Computing 1* (1980), 173–178.

[18] Doukhan, P., and Louhichi, S. A new weak dependence condition and application to moment inequalities. *Stochastic Processes and their Applications*, 84 (1999), 313–342.

[19] Flandrin, P. Time dependent spectra for nonstationary stochastic processes. In *Time and Frequency Representations of Signals and Systems*, G. Longo and B. Picinbono, Eds. Springer Verlag, New York, 1989, pp. 69–124.

[20] Franke, J., and Härdle, W. On bootstrapping kernel density estimates. *The Annals of Statistics 20*, 1 (1992), 121–145.

[21] Grenander, U., and Murray, R. Statistial spectral analysis of time series arising from stationary stochastic processes. *The Annals of Mathematical Statistics 24*, 4 (1953), 537–558.

[22] Grenier, Y. Time dependent ARMA modelling of nonstationary signals. *IEEE Trans. Acoust. Speech Signal Process. 31* (1983), 899–911.

[23] Hallin, M. Non-stationary q-dependent processes and time-varying moving-average models: Invertibility properties and the forecasting problem. *Advances in Applied Probability 18*, 1 (1986), 170–210.

[24] HAUG, S. Untersuchung lokal stationärer Zeitreihen mittels Wavelets. Master's thesis, Technische Universitaet Muenchen, 2002.

[25] HLAWATSCH, F., AND MATZ, G. Time-frequency signal processing: A statistical perspective. *Proc. CSSP-98* (Nov 1998), 207–219.

[26] JENTSCH, C., AND KREISS, J.-P. The multiple hybrid bootstrap - resampling multivariate linear processes. *Journal of Multivariate Analysis 101* (2010), 2320–2345.

[27] KIRCH, C. Resampling time series in the freqeuency domain of time series to determine critical values for change-point tests. *Statistics and Decisions 25*, 3 (2007), 237–261.

[28] KIRCH, C., AND POLITIS, D. TFT-bootstrap: Resampling time series in the freqeuency domain to obtain replicates in the time domain. *The Annals of Statistics 39*, 3 (2011), 1427–1470.

[29] KITAGAWA, G., AND GERSCH, W. A smoothness priors time-varying ar coefficient modeling of nonstationary covariance time series. *IEEE Transactions on Automatic Control 30*, 1 (1985), 48–56.

[30] KREISS, J., AND NEUHAUS, G. *Einführung in die Zeitreihenanalyse*. Springer, 2006.

[31] KREISS, J.-P., AND PAPARODITIS, E. Autoregressive aided periodogram bootstrap for time series. *The Annals of Statistics 31* (2003), 1923–1955.

[32] KREISS, J.-P., AND PAPARODITIS, E. Bootstrapping locally stationary processes. *Journal of the Royal Statistical Society* (2012).

[33] KREISS, J.-P., AND PAPARODITIS, E. The hybrid wild bootstrap for time series. *Journal of the American Statistical Association 107* (2012), 1073–1084.

[34] LENHOFF, M. W., SANTNER, T. J., OTIS, J. C., PETERSON, M. G., WILLIAMS, B. J., AND BACKUS, S. I. Bootstrap prediction and confidence bands: a superior statistical method for the analysis of gait data. *Gait and Posture 9* (1999), 10–17.

[35] LINDNER, T. *Zur Manipulierbarkeit der Allokation öffentlicher Güter : Theoretische Analyse und Simulationsergebnisse*. PhD thesis, 2011.

[36] LOYNES, R. On the concept of the spectrum for non-stationary processes. *Journal of the Royal Statistical Society 30*, 1 (1968), 1–30.

[37] MARTIN, W., AND FLANDRIN, P. Wigner-Ville spectral analsysis of nonstationary processes. *IEEE Transactions on Acoustics, Speech and Signal Processing 33* (1985), 1461–1470.

[38] Mélard, G. Propriétés du spectre évolutif d'un processus non-stationnaire. *Ann. Inst. H. Poincaré 14* (1978), 411–424.

[39] Neumann, M. H., and Polzehl, J. Simulaneous bootstrap confidence bands in nonparametric regression. *Nonparametric Statistics 9* (1998), 307–333.

[40] Neumann, M. H., and von Sachs, R. Wavelet theresholding in anisotronic function classes and application to adaptive estimation. *The Annals of Statistics 25*, 1 (1997), 38–76.

[41] Nze, P. A., and Doukhan, P. Weak dependence: Models and applications to econometrics. *Econometric Theory*, 20 (2004), 995–1045.

[42] Pagano, T., and Garen, D. A recent increase in western u.s. streamflow variability and persistence. *Journal of Hydrometerology 6* (2005), 173—-179.

[43] Page, C. H. Instantaneous power spectra. *Journal of Applied Physics 23*, 1 (1952), 103–106.

[44] Paparoditis, E., and Politis, D. N. The local bootstrap for periodogram statistics. *Journal of Time Series Analysis 20*, 2 (1999), 193–222.

[45] Parzen, E. Statistical inference on time-series by hilbert space methods, I. Tech. Rep. 23, Stanford, U.S.A., 1959.

[46] Priestley, M. B. Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society 27*, 2 (1965), 204–237.

[47] Priestley, M. B. Power spectral analysis of non-stationary random processes. *Journal of Sound and Vibration 6*, 1 (1967), 86–97.

[48] Schwarz, H. R., and Köckler, N. *Numerische Mathematik*, 6 ed. B.G. Teubner Verlag, 2006.

[49] Sergides, M. *Bootstrap Approaches for locally stationary processes.* PhD thesis, University of Cyprus, May 2008.

[50] Subba Rao, S. On some nonstationary, nonlinear random processes and their stationary approximations. *Advances in Applied Probability 38*, 4 (2006), 1155–1172.

[51] Subba Rao, T. The fitting of non-stationary time-series models with time-dependent parameters. *Journal of the Royal Statistical Society 32*, 2 (1970), 312–322.

[52] Sun, J., and Loader, C. R. Confidence bands for linear regression and smoothing. *The Annals of Statistics 22*, 3 (1994), 1328–1345.

[53] Tjostheim, D. Spectral generating operators for non-stationary processes. *Advances in Applied Probability 8*, 4 (1976), 831–846.

[54] von Sachs, R., and Schneider, K. Wavelet smoothing of evolutionary spectra by nonlinear thresholding. *Applied and Computational Harmonic Analysis 3* (1996), 268–282.

[55] Winograd, S. On computing the discrete Fourier transform. *Mathematics of Computation 32*, 141 (1978), 175–199.