BOOTSTRAP APPROACHES FOR LOCALLY STATIONARY PROCESSES

By

Marios Sergides

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
UNIVERSITY OF CYPRUS
NICOSIA CYPRUS
MAY 2008

UNIVERSITY OF CYPRUS

DEPARTMENT OF

DEPARTMENT OF MATHEMATICS AND STATISTICS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "**Bootstrap approaches for locally stationary processes**" by **Marios Sergides** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: May 2008

External Examiner: ______________________

Jens-Peter Kreiss

Research Supervisor: ______________________

Efstathios Paparoditis

Examing Committee: ______________________

Jürgen Franke

____

Konstantinos Fokianos

____

Theofanis Sapatinas

UNIVERSITY OF CYPRUS

Date: **May 2008**

Author: **Marios Sergides**

Title: **Bootstrap approaches for locally stationary processes**

Department: **Department of Mathematics and Statistics**

Degree: **Ph.D.** Convocation: **May** Year: **2008**

Signature of Author

To my wife

Table of Contents

Table of Contents **v**

Abstract in Greek **ix**

Abstract **xi**

Acknowledgements **xiii**

1 Introduction **2**
- 1.1 Locally stationary processes . . . 2
- 1.2 Resampling the local periodogram . . . 8
- 1.3 Testing semi-parametric hypothesis . . . 11
- 1.4 Contribution . . . 12

2 Bootstrapping the local periodogram **14**
- 2.1 Motivation . . . 14
- 2.2 Assumptions . . . 16
- 2.3 The Bootstrap procedure . . . 18
- 2.4 Asymptotic Properties . . . 20
- 2.5 Some remarks on choosing the resampling parameters . . . 26
- 2.6 Numerical Results . . . 28
  - 2.6.1 Simulations . . . 28
  - 2.6.2 A real-data example . . . 34

3 Testing for locally stationary processes **37**
- 3.1 The Testing procedure . . . 37
  - 3.1.1 The set-up . . . 37
  - 3.1.2 The test statistic . . . 38
  - 3.1.3 Asymptotic distribution under the null hypothesis . . . 41
- 3.2 Testing for a time-varying autoregressive structure . . . 42
  - 3.2.1 Consistency . . . 42
  - 3.2.2 Bootstrapping the test statistic . . . 45
- 3.3 Applications . . . 47
  - 3.3.1 Some remarks on choosing the testing parameters . . . 47
  - 3.3.2 Simulations . . . 48

4 Conclusions and further research **54**
- 4.1 Conclusions . . . 54
- 4.2 Further research . . . 54

A Auxiliary results and proofs for Chapter 2 **56**

B Auxiliary results and proofs for Chapter 3 **68**

Bibliography **83**

List of Figures

2.1 Estimated mean, variance, and mean square error of the standard deviation of the first-order sample autocorrelation using the pure tvAR($\cdot$) bootstrap and of the local periodogram bootstrap 32

2.2 Estimated mean, variance, and mean square error of the standard deviation of the first-order sample autocorrelation using the normal approximation and of the local periodogram bootstrap 33

2.3 2048 observations of an earthquake obtained by Shumway and Stoffer [39]. 34

2.4 Plots of the estimated sample local autocorrelation function $\widetilde{\rho}(u,\tau)$ against u for values of $\tau = 1,2,3,4$ of the earthquake data together with 95% pointwise local periodogram bootstrap confidence intervals. 36

3.1 Estimated density of the distribution of the test statistic Q_T under the null hypothesis of a first order tvAR process and its different approximations. The solid lines in (a) and (b) are the estimated exact densities, the dashed lines are the estimated densities corresponding to the median bootstrap approximations while the dotted lines are the densities of the asymptotic Gaussian approximations. 50

List of Tables

2.1 Local Periodogram Bootstrap (LPB) and Local Autoregressive Bootstrap (LARB) estimates of the standard deviation of the first-order local sample autocorrelation. 31

3.1 Rejection frequencies in 500 replications of the tvAR(2) model $X_{t,T} = 0.9\cos(1.5 - \cos(4\pi t/T))X_{t-1,T} - \phi_2 X_{t-2,T} + \varepsilon_t$ for different values of ϕ_2 and of the testing parameters. 52

Abstract in Greek

Η στασιμότητα έχει διαδραματίσει σημαντικό ρόλο στην ανάλυση των χρονοσειρών τις τελευταίες δεκαετίες. Η υπόθεση αυτή, αν και ελκυστική απο θεωρητική πλευρά, επειδή επιτρέπει την ανάπτυξη μιας καλά τεκμηριωμένης στατιστικής συμπερασματολογίας, είναι περιοριστική σε εφαρμογές. Μια πιο ρεαλιστική προσέγγιση στην ανάλυση των χρονοσειρών είναι αυτή που θα επιτρέπει τη δομή εξάρτησης της στοχαστικής ανέλιξης και πιο συγκεκριμένα τις ιδιότητες δευτέρας τάξης να αλλάζουν ομαλά με το χρόνο. Η ανάπτυξη μιας τέτοιας προσέγγισης επιβάλλει την ύπαρξη κάποιων περιορισμών στις αποκλίσεις από τη στασιμότητα για να γίνει δυνατή μια στατιστική συμπερασματολογία. Ένας τρόπος για να μελετήσουμε στατιστική συμπερασματολογία για χρονικά μεταβαλλόμενες ανελίξεις είναι να αυξάνουμε την πληροφορία που έχουμε τοπικά για την ανέλιξη όσο αυξάνεται το μέγεθος της χρονοσειράς. Τοπικά στάσιμες ανελίξεις είναι μη στάσιμες στοχαστικές ανελίξεις των οποίων οι ροπές πρώτης και δεύτερης τάξης αλλάζουν αργά στο χρόνο.

Η παρούσα διατριβή έχει δύο κυρίως σκοπούς:

1) Να αναπτύξει μια μέθοδο αναδειγματοληψίας που να δημιουργεί ψευτοπραγματώσεις του τοπικού περιοδογράμματος μιας τοπικά στάσιμης στοχαστικής ανέλιξης.

2) Να προταθεί ένας έλεγχος της υπόθεσης ότι η χρονικά μεταβαλλόμενη φασματική πυκνότητα έχει μια παραμετρική ή ημιπαραμετρική δομή. Ο έλεγχος μπορεί να εφαρμοστεί και σε χρονικά μεταβαλλόμενες ανελίξεις αυτοπαλινδρόμισης.

Η μέθοδος αναδειγματοληψίας που προτείνεται δημιουργεί ψευδοαντίγραφα του τοπικού περιοδογράμματος και συνδυάζει μία παραμετρική προσέγγιση στο χρόνο με μία απαραμετρική προσέγγιση φάσματος. Εφαρμόζουμε πρώτα τοπικά ένα χρονικά μεταβαλλόμενο μοντέλο αυτοπαλινδρόμισης ώστε να περιγράψουμε τα βασικά χαρακτηριστικά της ανέλιξης. Ένας, τοπικά υπολογισμένος, μη παραμετρικός διορθωτής στο

φασματικό πεδίο χρησιμοποιείται μετά για να βελτιωθεί η παραμετρική αυτοπαλινδρόμιση. Διερευνούμε τις ασυμπτωτικές ιδιότητες της μεθόδου στις οικογένειες των τοπικών φασματικών μέσων και τοπικών στατιστικών πηλίκου. Ασυμπτωτική αποτελεσματικότητα της μεθοδου αναδειγματοληψίας αποδεικνύεται σε δύο περιπτώσεις. Η μια αφορά την περίπτωση όπου η στάσιμη προσέγγιση της ανέλιξης μπορεί να παρασταθεί σαν μια αυτοπαλινδρόμιση άπειρης τάξης και η τάξη του μοντέλου που εφαρμόζεται τείνει στο άπειρο και η άλλη όταν η τάξη του μοντέλου είναι σταθερή. Προσομοιώσεις εξετάζουν τη δυνατότητα της μεθόδου να δίνει καλούς εκτιμητές των ποσοτήτων που μας ενδιαφέρουν σε δείγματα πεπερασμένου μεγέθους. Η ανάλυση ολοκληρώνεται με την παρουσίαση και εφαρμογή της μεθόδου σε πραγματικά δεδομένα.

Ο έλεγχος υποθέσεων που προτείνεται, βασίζεται στην απόσταση L^2 του σταθμισμένου το-πικού περιοδογράμματος από την αναμενόμενη τιμή του κάτω από τη μηδενική υπόθεση. Η στάθμιση γίνεται με την εκτιμώμενη ημιπαραμετρική φασματική πυκνότητα του μοντέλου. Η ασυμπτωτική κατανομή της ελεγχοσυνάρτησης που προτείνεται έχει υπολογιστεί κάτω από τη μηδενική υπόθεση για μια μεγάλη οικογένεια ημιπαραμετρικών μοντέλων τοπικά στάσιμων στοχαστικών ανελίξεων. Σαν ειδική περίπτωση, γίνεται ανάλυση του ελέγχου της ύπαρξης μιας χρονικά μεταβαλλόμενης αυτοπαλινδρόμισης. Για την καλύτερη προσέγγιση της κατανομής της ελεγχοσυνάρτησης κάτω από τη μηδενική υπόθεση προτείνεται μια μέθοδος αναδειγματοληψίας και αποδεικνύεται θεωρητικά ότι αυτή οδηγεί στα σωστά αποτελέσματα. Προσομοιώσεις παρουσιάζουν την αποτελεσματικότητα της μεθοδολογίας αναδειγματοληψίας και την αποδοτικότητα του ελέγχου σε πεπερασμένα δείγματα.

Abstract

Stationarity has played a major role in time series analysis during the last decades. Although this assumption is attractive from a theoretical point of view because it allows for the development of statistical inference procedures with good properties, it seems rather restrictive in applications. A more realistic framework in time series analysis is one which allows for the dependence structure of the underlying stochastic process and more specifically for its second order properties to vary smoothly over time. Developing a useful approach of statistical inference in such a context requires however, that some restrictions have to be imposed on the deviations from stationarity which are allowed. One way to investigate properties of statistical inference procedures for time-varying stochastic processes, is to allow for the amount of local information available to increase to infinity as the sample size increases. Locally stationary processes are non-stationary stochastic processes whose second order structure varies smoothly over time.

The aim of this thesis is twofold:

1) To develop a method to bootstrap the local periodogram of a locally stationary process and

2) To propose a test of the hypothesis that the time varying spectral density of a locally stationary process has a semiparametric structure including that of the time varying autoregressive moving average model.

The bootstrap method proposed generates pseudo local periodogram ordinates by combining a parametric time and non-parametric frequency domain bootstrap approach. We first fit locally a time varying autoregressive model in order to capture the essential characteristics of the underlying process. A locally calculated non-parametric correction in the frequency domain is then used in order to improve upon

the locally parametric autoregressive fit. Some remarks on choosing the resampling parameters are considered. As an application, we investigate the asymptotic properties of the bootstrap method proposed applied to the class of local spectral means and local ratio statistics. Asymptotic normality of these statistics have been proven under two cases, i.e. when the stationary approximation of the process has an infinitive order autoregressive and the order of the fitted model tends to infinitive and when the order of the fitted model is fixed. Some simulations demonstrate the ability of our method to give accurate estimates of the quantities of interest and an application to an earthquake data set is presented.

Concerning the test introduced, it is based on the L^2-distance of a kernel smoothed version of the local periodogram rescaled by the estimated semiparametric, time varying spectral density. The asymptotic distribution of the test statistic proposed is derived under the null hypothesis and it is shown that this distribution is a Gaussian distribution with the nice feature that its parameters do not depend on characteristics or parameters of the underlying process. As an interesting special case, we consider the problem of testing the presence of a time-varying autoregressive structure. A bootstrap procedure to approximate more accurately the distribution of the test statistic under the null hypothesis is proposed and theoretically justified. Remarks on choosing the resampling parameters are considered. Some simulations illustrate that the bootstrap provides a considerably better approximation of the distribution of the test statistic under the null hypothesis than the normal approximation.

Acknowledgements

I would like to thank, first and foremost, my advisor, Efstathios Paparoditis, for his guidance and support during my studies.
I am also grateful to the members of my committee for their insightful comments and suggestions. I would also like to thank the University of Cyprus and especially the Department of Mathematics and Statistics for providing me an excellent environment for learning and research.
Thanks also to my family for supporting and encouraging me in my educational pursuits.

Chapter 1

Introduction

1.1 Locally stationary processes

Most existing models in time series analysis assume that the underlying process is second-order stationary. This assumption is useful in order to have estimators for certain characteristics or parameters of the process with good statistical properties, such as consistency, efficiency or central limit theorems. Weak stationarity refers to the property that the first and second order moment structure of a stochastic process is invariant with respect to time translations. It has been the dominating paradigm in time series analysis for many decades. The theoretical setting of weak stationarity together with an appropriate notion of weak dependence have been proven to be quite effective in time series analysis dealing to the development of a powerful asymptotic theory capable to investigate properties of statistical inference procedures, ref. [36], [4] , [5], [1]. An important theorem for weak stationary processes is the spectral representation theorem, cf. Brockwell and Davis [5], which states that it is possible to write one stationary process as a stochastic integral i.e.

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t} A(\lambda) d\xi(\lambda), \quad t \in \mathbb{N} \tag{1.1.1}$$

with a transfer function $A(\lambda)$ and an orthogonal increment process $\xi(\lambda)$.

However, a more realistic framework in time series analysis is one which assumes that the second order characteristics of the observed process vary over time. Priestley [35] considers stochastic processes with a time varying spectral representation similar to that of stationary process; cf. also [36]. The development of statistical inference procedures for such processes has attracted considerably interest in the literature. Valuable statistical inference requires that the amount of local information available increases to infinity as the sample size increases. In this context, a framework for the development of an asymptotic theory of statistical inference has been provided by Dahlhaus [9] who introduced the class of locally stationary processes. Locally stationary processes are stochastic processes whose spectral structure varies smoothly over time. This concept can be extended/modified in several directions. For instance, Nason et al. [27] adopted the concept of local stationarity but replaced the spectral representation and the Fourier basis involved by a representation with respect to a wavelets basis; see also [30].

The idea of a nonstationary process with time-varying characteristics was made rigorous in Priestley's theory of processes with evolutionary spectra. Priestly investigates processes $\{X_t, t \in \mathbb{N}\}$, where X_t has time varying spectral representation,

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t} A_t(\lambda) d\xi(\lambda), \quad t \in \mathbb{N} \tag{1.1.2}$$

with a time-varying transfer function $A_t(\lambda)$ and an orthogonal increment process $\xi(\lambda)$. This approach, however, do not allow for asymptotic considerations due to the nature of the nonstationarity considered. As a result, important tools like consistency, asymptotic normality, efficiency etc. can not be proved in the theoretical treatment of statistical procedures for such processes.

In order to overcome this problem Dahlhaus [9] introduced processes with time-varying spectral representation, an approach similar to that in nonparametric regression. A simple example is the process

$$X_t = \sigma_t Y_t, \quad t = 0, 1, 2, \ldots \tag{1.1.3}$$

where $Y_t, \quad t = 0, 1, 2, \ldots$ is a zero mean stationary process with unit variance and σ_t is a deterministic positive function of the time parameter t. Here the degree of nonstationarity of the process X_t is measured by its time-varying variance function σ_t. Suppose that we have observations $X_1, X_2, \ldots, X_T$ and that we want to estimate the deterministic function σ_t. For this some regularity assumptions on the function

σ_t have to be imposed. However, to estimate the variance function σ_t only T observations are available and with this approach we do not get any increasing amount of information on the local structure of σ_t as the sample size increases. Thus, asymptotic considerations are difficult to use in the statistical analysis of such processes.

To set up an adequate asymptotic theory for nonstationary processes, Dahlhaus [9] proceeds with a rescaling of the time variable transforming the support of σ_t to be the interval $[0, 1]$. The process is then rewritten as,

$$X_{t,T} = \sigma(\frac{t}{T})Y_t, \tag{1.1.4}$$

where $\{\mathbf{X}_T\}_{T\in\mathbb{N}} = \{X_{t,T}, t = 1, \ldots, T\}_{T\in\mathbb{N}}$, refers now to a triangular array of stochastic processes. In this approach we have two scales of time: the observed time which is the usual scale of time $1, 2, \ldots, T$, and, the rescaled time defined on $[0, 1]$. This rescaling is a standard approach in nonparametric statistics. Now, letting T tending to infinity means in this context that more and more observations on the local structure of the function $\sigma(u)$ are available.

A second example is the time-varying AR(1) processes

$$X_t = a(t)X_{t-1} + \varepsilon_t, \tag{1.1.5}$$

where the ε_t's are assumed to be i.i.d. standard normal random variables. As in nonparametric regression and using a rescaling principle like the one described above, the function $a(t)$ is rescaled to the unit interval and a triangular array of stochastic processes $\{\mathbf{X}_T\}_{T\in\mathbb{N}} = \{X_{t,T}, t = 1, \ldots, T\}_{T\in\mathbb{N}}$ is considered, where

$$X_{t,T} = a(\frac{t}{T})X_{t-1,T} + \varepsilon_t. \tag{1.1.6}$$

Comparing the process (1.1.6) with the spectral representation (1.1.2) one could investigate a process of the form

$$X_{t,T} = \int_{-\pi}^{\pi} e^{i\lambda t} A(\frac{t}{T}, \lambda) d\xi(\lambda). \tag{1.1.7}$$

However, it can be shown that the model (1.1.6) has not exact but only approximative a solution of the form (1.1.7), see [9]. This observation led Dahlhaus [9] to introduce the following, general definition of locally stationary processes.

Definition 1.1.1. *A triangular array of stochastic processes $X_{t,T}$ (t=1,...,T), $T \in \mathbb{N}$ is called locally stationary with transfer function A^0 if there exists a representation*

$$X_{t,T} = \int_{-\pi}^{\pi} e^{i\lambda t} A^0_{t,T}(\lambda) d\xi(\lambda), \tag{1.1.8}$$

where the following holds

(i) *$\xi(\lambda)$ is a stochastic process on $[-\pi, \pi]$ with $\overline{\xi(\lambda)} = \xi(-\lambda)$ and*

$$cum_k\{d\xi(\lambda_1), \ldots, d\xi(\lambda_k)\} = \eta\left(\sum_{j=1}^{k} \lambda_j\right) g_k(\lambda_1, \ldots, \lambda_{k-1}) d\lambda_1 \ldots d\lambda_k$$

where $cum_k\{\cdot\}$ denotes the kth order cumulant, $g_1 = 0$, $g_2(\lambda) = 1$, $|g_k(\lambda_1, \ldots, \lambda_{k-1})| \leq const_k$ for all k and $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ is the period 2π extension of the Dirac delta function.

(ii) *There exists constant K and a 2π-periodic function $A : [0, 1] \times \mathbb{R} \to \mathcal{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$ such that*

$$\sup_{t,\lambda} \left| A^0_{t,T}(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1} \tag{1.1.9}$$

for all T. $A(u, \lambda)$ is assumed to be continuous in u.

Remark: The complicated construction with the two functions $A(\frac{t}{T}, \lambda)$ and $A^0_{t,T}(\lambda)$ is necessary because on the one hand the smoothness assumptions on $A(\frac{t}{T}, \lambda)$ guarantees that the process has a locally stationary behavior and on the other hand $A^0_{t,T}(\lambda)$ ensures that the class of processes considered is rich enough to include interesting applications like the time-varying model (1.1.6).

Another definition for locally stationary processes was given by Dahlhaus and Polonik [16] using an infinitive order, time-varying moving average representation instead of the spectral representation (1.1.8).

Definition 1.1.2. *A triangular array $\{\mathbf{X}_T\}_{T \in N}$ of stochastic processes $\mathbf{X}_T = \{X_{t,T}, t = 1, \ldots, T\}$ is called locally stationary if $X_{t,T}$ fulfills the following conditions.*

(i) *$X_{t,T}$ has the representation*

$$X_{t,T} = \sum_{j=-\infty}^{\infty} \alpha_{t,T}(j) \varepsilon_{t-j} \tag{1.1.10}$$

where ε_t *are independent, identically distributed random variables with* $E\varepsilon_t = 0$, $E\varepsilon_t^2 = 1$ *and* $E{\varepsilon_t}^4 < \infty$. *Let* $\kappa_4 = E{\varepsilon_t}^4 - 3$ *be the fourth cumulant of* ε_t.

(ii) *A sequence* $\{\ell(j), j \in \mathbb{Z}\}$ *satisfying*

$$\sum_{j=-\infty}^{\infty} \frac{|j|}{\ell(j)} < \infty$$

exists, such that

$$\sup_t |\alpha_{t,T}(j)| \leq \frac{K}{\ell(j)}. \tag{1.1.11}$$

(iii) *Functions* $\alpha(\cdot, j) : (0, 1] \to \mathbb{R}$ *exist satisfying*

$$\sup_u |\alpha(u, j)| \leq \frac{K}{\ell(j)}, \tag{1.1.12}$$

$$\sup_{u,v} |\alpha(u, j) - \alpha(v, j)| \leq \frac{K|u - v|}{\ell(j)} \tag{1.1.13}$$

and

$$\sup_t \left|\alpha_{t,T}(j) - \alpha(\frac{t}{T}, j)\right| \leq \frac{K}{T\ell(j)}. \tag{1.1.14}$$

Dahlhaus [11] discusses the equivalence of the spectral and the infinite moving average representations of a locally stationary process.

The concept of local stationarity has been extended, modified in several directions during the last decade. For instance, Nason et al. [27] introduced the class of locally stationary wavelet processes replacing the Fourier basis by a wavelet basis. Locally stationary wavelet processes have the representation

$$X_{t,T} = \sum_j \sum_k w^0_{j,k;T} \psi_{jk}(t) \xi_{jt} \tag{1.1.15}$$

where ξ_{jt} is a random orthonormal increment sequence and $\psi_{jk}(t)$ a discrete, non-decimated family of wavelets. Also they assume that there exist, appropriately defined, functions $W_j(z)$ and constants C_j such that similar to (1.1.14),

$$\sup_k \left|w^0_{j,k;T} - W_j(k/T)\right| \leq C_j/T. \tag{1.1.16}$$

In the following we elaborate on the stationary approximation $\tilde{X}_t(u)$ of a locally stationary process. Let $u \in [0, 1]$ be fixed and define

$$\tilde{X}_t(u) := \int_{-\pi}^{\pi} e^{i\lambda t} A(u,\lambda) d\xi(\lambda), \tag{1.1.17}$$

which has the following moving average representation

$$\tilde{X}_t(u) := \sum_{j=-\infty}^{\infty} \alpha(u,j)\, \varepsilon_{t-j}. \tag{1.1.18}$$

It then follows by simple algebra that

$$|X_{t,T} - \widetilde{X}_t(u)| \leq K\{|t/T - u| + 1/T\} U_t, \tag{1.1.19}$$

where $\{U_t\}$ is the stationary process

$$U_t := \sum_{j=-\infty}^{\infty} \ell^{-1}(j)\, |\varepsilon_{t-j}|\,. \tag{1.1.20}$$

Thus, according to (1.1.19) the stationary process $\widetilde{X}_t(u)$ is an approximation of $X_{t,T}$ in a local neighborhood around $u = t/T$; cf. [17] and [40].

Recall that the spectral density of a stationary process $X_t = \int_{-\pi}^{\pi} e^{i\lambda t} A(\lambda) d\xi(\lambda)$, which satisfies $\sum_h \gamma(h) < \infty$ where $\gamma(\cdot)$ is the covariance function of the process, is defined by $f(\lambda) = (2\pi)^{-1}|A(\lambda)|^2$.

Similarly to this for locally stationary processes we have the following definition of a time-varying spectral density.

Definition 1.1.3. *The function*

$$f(u,\lambda) = \frac{1}{2\pi}|A(u,\lambda)|^2, \tag{1.1.21}$$

where $A(\cdot,\cdot)$ is given in Definition 1.1.1 is the time-varying spectral density of a locally stationary process.

Definition 1.1.4. *The Fourier transform of the time-varying spectral density is the time-varying covariance of lag k, $k \in \mathbb{Z}$, at time u, $u \in (0,1)$, and it is defined by*

$$c(u,k) := \int_{-\pi}^{\pi} f(u,\lambda) e^{i\lambda t} d\lambda. \tag{1.1.22}$$

Note that $f(u,\lambda)$ is the spectral density of the stationary approximation $\{\widetilde{X}_t(u), t \in \mathbb{Z}\}$ and that

$$\int_{-\pi}^{\pi} |f_T(u,\lambda) - f(u,\lambda)|^2 \, d\lambda = o(1), \quad \text{as} \quad T \to \infty \tag{1.1.23}$$

where $f_T(u,\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} Cov(X_{[uT-k/2],T}, X_{[uT+k/2],T}) \exp(-i\lambda k)$; cf. [7] Theorem 2.2. This means that if the process is locally stationary with a smooth function $A(u,\lambda)$ then $f(u,\lambda)$ is uniquely defined by the triangular array.

1.2 Resampling the local periodogram

In inferring properties of stationary processes in the frequency domain one important tool is the periodogram $I_T(\lambda)$ defined by

$$I_T(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X_t e^{-i\lambda t} \right|^2, \tag{1.2.24}$$

where $X_1, X_2, \ldots, X_T$ are observations.

The periodogram ordinates behave for large sample sizes like independent, exponentially distributed random variables and they are asymptotically unbiased but not consistent estimators of the spectral density.

When dealing with locally stationary process one possibility is to consider the periodogram over a segment of length N length of consecutive observations around a time point $[uT]$, $u \in (0,1)$, of the observed series.

Definition 1.2.1. *The local periodogram is defined for every* $\lambda \in [-\pi,\pi]$ *and* $u \in [0,1]$ *by*

$$I_N(u,\lambda) = \frac{1}{2\pi N} \left| \sum_{s=0}^{N-1} X_{[uT]-N/2+s+1,T} e^{-is\lambda} \right|^2, \tag{1.2.25}$$

where $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$ *are observations.*

A common assumption in inferring properties of statistics based on the local periodogram is that the "time window" width N tends to infinity at an appropriate rate as the sample size increases.

Several interesting classes of statistics used in the analysis of locally stationary processes can be expressed as functions of the local periodogram. For instance, spectral means are obtained by averaging over all frequencies the local periodogram multiplied by an appropriately defined complex valued function ϕ in $[-\pi, \pi]$. To be more specific, such a statistic is given for $u \in (0,1)$, by

$$M_T(u, \phi) = \int_{-\pi}^{\pi} \phi(\lambda) I_N(u, \lambda) d\lambda. \tag{1.2.26}$$

An interesting special cases of $M_T(u, \phi)$ which is frequently used in inferring properties of locally stationary processes is obtained if we set $\phi(\lambda) = \exp(i\lambda\tau)$ for some τ, $0 \leq \tau \leq N-1$. For this choice of ϕ the above statistic becomes the sample time varying covariance $\hat{c}(u, \tau)$ given by

$$\begin{aligned}
\hat{c}(u, \tau) &= \int_{-\pi}^{\pi} \exp(i\lambda\tau) I_N(u, \lambda) d\lambda \\
&= \frac{1}{N} \sum_{\substack{k,l=0 \\ k-l=\tau}}^{N-1} X_{[uT]-N/2+k+1,T} X_{[uT]-N/2+l+1,T}.
\end{aligned}$$

Notice that $\hat{c}(u, \tau)$ is an estimator of the time varying covariance

$$c(u, \tau) = \int_{-\pi}^{\pi} \exp(-i\lambda\tau) f(u, \lambda) d\lambda,$$

where $f(u, \lambda)$ denotes the local spectral density of $\{\mathbf{X}_T\}_{T \in \mathbb{N}}$.

Another class of statistics derived from that of local spectral means is that of ratio statistics which are defined for $u \in (0,1)$ by

$$\begin{aligned}
R_T(u, \phi) &= \frac{M_T(u, \phi)}{M_T(u, 1)} \\
&= \frac{\int_{-\pi}^{\pi} \phi(\lambda) I_N(u, \lambda) d\lambda}{\int_{-\pi}^{\pi} I_N(u, \lambda) d\lambda}.
\end{aligned} \tag{1.2.27}$$

An important member of the class (1.2.27) is the time varying sample autocorrelation

$$\hat{\rho}(u, \tau) = \hat{c}(u, \tau) / \hat{c}(u, 0).$$

The asymptotic behavior of statistics like $M_T(u, \phi)$ and $R_T(u, \phi)$ has been investigated by Dahlhaus [9] and Dahlhaus and Giraitis [13]. Under certain smoothness conditions they showed asymptotic normality of appropriately centered and rescaled versions of these statistics. In the second chapter of the thesis we use our bootstrap method to bootstrap the local periodogram of a locally stationary stationary processes, to approximate the asymptotic distribution of statistics like $M_T(u, \phi)$ and $R_T(u, \phi)$.

Bootstrap methods for locally stationary processes have received little attention in the literature. A time domain local block bootstrap procedure for locally stationary processes has been proposed by Paparoditis and Politis [34] and by Dowla et al. [18]. For stationary processes frequency domain bootstrap methods have been considered among others, by Nordgaard [29] and Theiler et al. [41]. Using a similar to (2.1.1) property for the periodogram of a stationary process, Hurvich and Zeger [22] and Franke and Härdle [19], proposed a nonparametric residual-based bootstrap method. Dahlhaus and Janas [12] extended the validity of this bootstrap procedure to the class of the ratio statistics and to Whittle estimators. An alternative idea to bootstrap the periodogram of a stationary process has been proposed by Paparoditis and Politis [33]. An overview of the different methods to bootstrap stationary time series in the frequency domain is given by Paparoditis [32]. A common feature of the aforementioned bootstrap approaches for stationary processes is that the generated bootstrap periodogram ordinates are independent. This restricts the applicability of the corresponding methods to statistics for which the asymptotically negligible dependence of the periodogram does not affect properties of their asymptotic distribution; cf. Dahlhaus and Janas [12] and Paparoditis [32]. More recently, and in order to overcome these problems, Kreiss and Paparoditis [24] proposed a bootstrap method for the periodogram of a stationary process which is based on a combination of a parametric time domain and a nonparametric frequency domain bootstrap and which generates bootstrap periodogram replicates that capture to some extent the dependence structure of the periodogram. The procedure proposed in the first chapter extends to locally stationary process this idea of combining a time domain and a frequency domain approach to bootstrap the periodogram of a stationary process. Furthermore, it justifies theoretically the use of such an approach to approximate the distribution of statistics like (1.2.26) and (1.2.27). The theory developed in this chapter can be also used to establish validity of our bootstrap procedure applied to other classes of statistics than (1.2.26) and (1.2.27) and which are based on the local periodogram $I_N(u,\lambda)$. For instance, frequency domain estimators of the parameters of a locally stationary parametric process and nonparametric estimators of the local spectral density belong to this class.

1.3 Testing semi-parametric hypothesis

Interesting subclasses of locally stationary processes are obtained by parameterizing in a proper way the associated time-varying amplitude function and consequently the underlying time varying spectral density. Such an interesting subclass of locally processes is for instance, that of time-varying, autoregressive moving-average (tvARMA) models. tvARMA models are autoregressive moving-average model which satisfy the following system of difference equations

$$\sum_{j=0}^{p} \phi_j(\frac{t}{T}) X_{t-j,T} = \sum_{k=0}^{q} \beta_k(\frac{t}{T}) \sigma(\frac{t-k}{T}) \varepsilon_{t-k}, \qquad (1.3.28)$$

$\phi_0(\cdot) = 1$, $\beta_0(\cdot) = 1$, ε_t are i.i.d. with $E\varepsilon_t = 0$ and $E\varepsilon_t^2 < \infty$. If all $\phi_j(\cdot)$, $\beta_k(\cdot)$ and $\sigma^2(\cdot)$ are of bounded variation and $\sum\limits_{j=0}^{p} \phi_j(u) z^j \neq 0$ for all u and all $0 < |z| \leq 1 + \delta$ for some $\delta > 0$ then there exists a solution of the form

$$X_{t,T} = \sum_{j=0}^{\infty} a_{t,T}(j) \varepsilon_{t-j}.$$

The time-varying spectral density is given in this case by

$$f(u, \lambda) = \frac{\sigma^2(u)}{2\pi} \frac{\left| \sum\limits_{k=0}^{q} \beta_k(u) e^{i\lambda k} \right|^2}{\left| \sum\limits_{j=0}^{p} \phi_j(u) e^{i\lambda j} \right|^2},$$

cf. Dahlhaus[7]. Notice that if all $\phi_j(\cdot)$, $\beta_k(\cdot)$ and $\sigma^2(\cdot)$ are constant and independent of t we obtain the stationary autoregressive-moving average process $ARMA(p, q)$.

Estimation procedures for locally stationary processes have been considered by many authors under different settings and assumptions. We mention here among others the contributions by Neumann and von Sachs [28], Dahlhaus et al. [14], Chang and Morettin [6] and van Bellegem and Dahlhaus [42]. Forecasting problems for non-stationary time series have been considered by Fryzlewicz et al. [20]. An overview on some of the different developments can be found in Dahlhaus [11]. However, the important problem of testing for the presence of a parametric or semiparametric structure of the underlying locally stationary process, has attracted less attention in the literature. Testing for the presence of such a structure is important because it allows for the use of efficient, i.e., model-based estimation and forecasting procedures. For Gaussian locally stationary processes, Sakiyama and Taniguchi [37] proposed likelihood ratio, Wald and Lagrange multiplier tests of the null hypothesis that the time-varying spectral density depends on a finite dimensional, real-valued parameter vector against a real-valued parametric alternative. However, the class of parametric time-varying spectral densities allowed in this context, is rather restrictive in that it does not include for instance the important case of testing for the presence of a semiparametric tvARMA structure against an unspecified, locally stationary alternative.

In the third chapter of this thesis, we address the important problem of testing whether a locally stationary process belongs to a semiparametric class of time varying processes. The semiparametric class considered under the null is large enough to include several interesting processes. The test statistic developed, evaluates over all frequencies and over an increasing set of time points, a L^2-type distance between the sample local spectral density (local periodogram) and the time-varying spectral density of the fitted semiparametric model postulated under the null. The asymptotic distribution of the test statistic proposed is derived under the null hypothesis and it is shown that this distribution is a Gaussian distribution with the nice feature that its parameters do not depend on characteristics or parameters of the underlying process. As an interesting special case we focus on the problem of testing for the presence of a semiparametric, time-varying autoregressive model. In this context, a bootstrap procedure is proposed to approximate more accurately the distribution of the test statistic under the null hypothesis. Theoretical properties of the bootstrap procedure are discussed and its asymptotic validity is established. It is demonstrated by means of numerical examples that in the testing set-up considered in this chapter, the bootstrap is a very powerful and valuable tool to obtain critical values in finite sample situations.

1.4 Contribution

This thesis is contributing to the literature by proposing a bootstrap method for resampling the local periodogram of a locally stationary process and a test of the hypothesis that the time varying spectral density of a locally stationary process has a semiparametric structure. Properties of the methods proposed have been studied theoretically and investigated by means of simulations. An application to a real-data set is also given.

Chapter 2

Bootstrapping the local periodogram of a locally stationary process.

2.1 Motivation

The aim of this chapter is to develop an alternative, bootstrap-based method to approximate the distribution of statistics like (1.2.26) and (1.2.27). Our method works by generating replicates $I^*_N(u,\lambda)$ of the local periodogram $I_N(u,\lambda)$. To describe heuristically the basic idea underlying our method notice first that under certain assumptions on the underlying process, the local periodogram of a locally stationary process can be approximately written as

$$\begin{aligned} I_N(u,\lambda) &= I_{N,\widetilde{X}}(u,\lambda) + O_p(\frac{N}{T}) \\ &= f(u,\lambda) I_{N,\varepsilon}(u,\lambda) + O_p(\frac{N}{T}) + O_p(\frac{1}{N}) \qquad (2.1.1) \end{aligned}$$

where $I_{N,\widetilde{X}}(u,\lambda)$ is the local periodogram based on observations $\widetilde{X}_1(u), \widetilde{X}_2(u), \ldots, \widetilde{X}_T(u)$ of the process $\{\widetilde{X}_t(u)\}$ which is defined in equation (1.1.10) of Section 1.1. $I_{N,\varepsilon}(u,\lambda) = (2\pi N)^{-1}\Big|\sum_{s=0}^{N-1} \varepsilon_{[uT]-N/2+s+1} e^{-is\lambda}\Big|^2$ is the local periodogram of the i.i.d series $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T$. Proof of the above approximation can be obtained by using Lemma A.0.4(i) and Theorem 10.3.1 of [5].

To proceed, let $f_{tvAR}(t/T,\lambda)$ be the local spectral density of the pth order locally autoregressive process which best fits (in the mean square sense) the locally stationary process $X_{t,T}$. Using (2.1.1) we can write

$$I_N(u,\lambda) \;\approx\; \frac{f(u,\lambda)}{f_{tvAR}(u,\lambda)} f_{tvAR}(u,\lambda) I_{N,\varepsilon}(u,\lambda).$$

Now, since $f_{tvAR}(u,\lambda)I_{N,\varepsilon}(u,\lambda)$ can be considered as the local periodogram of a pth order locally stationary autoregressive process, we end up with the following approximative expression for the local periodogram,

$$I_N(u,\lambda) \;\approx\; g(u,\lambda) I_{N,tvAR}(u,\lambda), \tag{2.1.2}$$

where we have used the notation $I_{N,tvAR}(u,\lambda) = f_{tvAR}(u,\lambda)I_{N,\varepsilon}(u,\lambda)$ and $g(u,\lambda) = f(u,\lambda)/f_{tvAR}(u,\lambda)$.

Expression (2.1.2) motivates the following procedure to generate replicates $I_N^*(u,\lambda)$ of the local periodogram. We first fit in the time domain a pth order time varying autoregressive model and use the fitted model to generate pseudo series $X_{1,T}^+, X_{2,T}^+, \ldots, X_{T,T}^+$. The local periodogram $I_{N,tvAR}^+(u,\lambda)$ of this series is obtained which can be used to mimic the random behavior of $I_{N,tvAR}^+(u,\lambda)$ in (2.1.2). A nonparametric kernel estimator $\hat{g}(u,\lambda)$ of $g(u,\lambda)$ is calculated in the frequency domain by smoothing the rescaled local periodogram $I_N(u,\lambda)/\widehat{f}_{tvAR}(u,\lambda)$, where $\widehat{f}_{tvAR}(u,\lambda)$ is the local spectral density of the fitted autoregressive process. Following (2.1.2) the bootstrapped local periodogram is then obtained as $I_N^*(u,\lambda) = \hat{v}(u,\lambda)I_{N,tvAR}^+(u,\lambda)$. Details of this procedure are given in Section 2.3. Notice that our method to bootstrap the local periodogram is based on a combination of a parametric time domain and a nonparametric frequency domain bootstrap. The parametric time domain bootstrap generating $I_{N,tvAR}^+(u,\lambda)$ is used to capture the essential features and to reproduce (at least to some extent) the dependence structure of the local periodogram $I_N(u,\lambda)$. The nonparametric estimator $\hat{g}(u,\lambda)$ is used to reproduce features of the local periodogram that are not captured by the local parametric autoregressive fit.

Notice that if we set $\hat{g}(u,\lambda) \equiv 1$ the method described above is just a local version of the autoregressive bootstrap. If, additionally, the underlying process has an infinite order autoregressive representation and we allow the order p of the locally fitted autoregressive process to increase to infinity as the sample size increases, then we have a local version of the autoregressive sieve bootstrap. However, our method differs from such an autoregressive bootstrap scheme due to the frequency domain nonparametric correction via the function $\hat{g}(u,\lambda)$. Due to this correction our method is more general than the pure local autoregressive bootstrap in that it leads to asymptotically valid approximations for a larger class of statistics, cf. Section 2.4.

2.2 Assumptions

In this section we impose the assumptions needed in this chapter.

Assumption 2.1 The triangular array $\{\mathbf{X}_T\}_{T\in N}$ of stochastic processes $\mathbf{X}_T = \{X_{t,T}, t = 1,\ldots,T\}$ satisfies Definition 1.1.2.

In the following we consider also the case where the local approximating process $\widetilde{X}_t(u)$ satisfies the following condition.

Assumption 2.2 The process $\{\widetilde{X}_t(u), t \in \mathbb{Z}\}$ has the representation

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

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

Regarding the time window width N we require that the following condition is satisfied.

Assumption 2.3 The window width N satisfies $N \to \infty$ such that $N^{3/2}/T \to 0$ as $N \to \infty$.

Remark: As a careful read of the proofs shows, the condition $N^{3/2}/T \to 0$ is needed in order to make the difference between the centered local spectral mean based on the local periodogram of the observations $X^+_{1,T}, X^+_{2,T}, \ldots, X^+_{T,T}$ and the corresponding centered local spectral mean based on the local periodogram of the observations $\widetilde{X}^+_{1,T}, \widetilde{X}^+_{2,T}, \ldots, \widetilde{X}^+_{T,T}$ coming from the local approximating process, asymptotically negligible. For instance, we get using the notation of Appendix A, that

$$J^+_N(\phi) - E^+(J^+_N(\phi)) = J^+_{N,\widetilde{X}^+}(\phi) - E^+(J^+_{N,\widetilde{X}^+}(\phi)) + O_p(N/T + p^2/N). \qquad (2.2.3)$$

The $O_p(N/T)$ error term on the right hand side above, is due to the variance of the term

$$d_N^{+(2)}(u,\lambda) = \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} \sum_{m=0}^{\infty} (\hat{\alpha}_{[uT]-N/2+s+1,T}(m) - \hat{\alpha}_{[uT],T}(m))\varepsilon_{s-m}^{+} \exp(-i\lambda s),$$

which is of order N^2/T^2; see the proof of Lemma A.0.7 in Appendix A. Now, multiplying equation (2.2.3) by $\sqrt{N}$, leads to the requirement $N^{3/2}/T \to 0$ as $T \to \infty$. We conjecture that the above condition can be relaxed to a less restrictive condition $N/T \to 0$, by using procedures like those used to reduce the bias of the local periodogram, e.g., by using tapered local bootstrap periodogram; see Dahlhaus and Giraitis [13].

Since we are interested in investigating the properties of our bootstrap method applied in order to approximate the distribution of statistics like those defined in (1.2.26) and (1.2.27), we impose the following conditions on the functions ϕ appearing in the corresponding definitions.

Assumption 2.4 $\phi \in \Phi$, where Φ is the set of complex-valued bounded functions equipped with the uniform norm $\|\phi\|_\infty = \sup_x |\phi(x)|$. Furthermore ϕ is periodically extended to $\mathbb{R}$ with period 2π and has a bounded second derivative $\phi^{''}(x) = d^2\phi(x)/d^2x$.

As mentioned in the Introduction, our bootstrap method uses a nonparametric estimator of the function $q(u,\lambda)$ in the frequency domain in order to capture features of the local spectral density not represented by the local parametric autoregressive fit. This is done by smoothing the local periodogram rescaled by the local spectral density of the fitted autoregressive model. To obtain a nonparametric kernel estimator of the function $g(u,\lambda)$ used in our procedure, Assumptions 2.4 and 2.5 below are imposed. They deal with the properties of the smoothing kernel K and the smoothing bandwidth h.

Assumption 2.5 K is a nonnegative kernel function with support $[-\pi,\pi]$. Furthermore, the Fourier transform k of K is symmetric, continuous, bounded and satisfies $k(0) = 2\pi$ and $\int_{-\infty}^{\infty} k^2(x)dx < \infty$.

Assumption 2.6 The smoothing bandwidth $h = h(N)$ satisfies $h \to 0$ such that $Nh \to \infty$ as $N \to \infty$.

2.3 The Bootstrap procedure

Before presenting the bootstrap algorithm in more detail we fix some notation. Recall that in its first step our procedure is based on fitting locally to the time series a pth order autoregressive process. A least squares estimator of the corresponding autoregressive parameters $\beta_1(u), \ldots, \beta_p(u)$ is obtained by minimizing the local quadratic deviation

$$\frac{1}{N-p}\sum_{j=p}^{N}\Big(X_{[uT]-N/2+j,T}-\sum_{i=1}^{p}c_i(u)X_{[uT]-N/2+j-i,T}\Big)^2 \tag{2.3.4}$$

with respect to $c_i(u),\ i=1,2,\ldots,p$. This leads to the estimates $c_1(u)=\hat{\beta}_1(u), \ldots,$ $c_p(u)=\hat{\beta}_p(u)$ where $\hat{\beta}_u(p)'=(\hat{\beta}_1(u),\ldots,\hat{\beta}_p(u))$ satisfies the system of equations

$$\hat{R}_u(p)\hat{\beta}_u(p)=\hat{r}_u(p).$$

Here,

$$\hat{R}_u(p)=\sum_{j=p}^{N-1}X_j(u,p)X_j(u,p)'/(N-p), \quad \hat{r}_u(p)=\sum_{j=p}^{N-1}X_j(u,p)X_{[uT]-N/2+j,T}/(N-p)$$

and

$$X_j(u,p)'=\Big(X_{[uT]-N/2+j,T}, X_{[uT]-N/2+j-1,T}, \ldots, X_{[uT]-N/2+j-p,T}\Big).$$

Let

$$\hat{\sigma}_p^2(u)=\frac{1}{N-p}\sum_{j=p}^{N-1}X^2_{[uT]-N/2+j-p,T}-\hat{\beta}_u(p)'\hat{r}_u(p)$$

be the estimated variance of the errors of the local autoregressive fit. We are now ready to formulate our bootstrap algorithm which consists of the following five steps:

STEP 1: Fit locally a time varying autoregressive model of order p to the observations $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$ and calculate the estimated parameters $\hat{\beta}(t/T,p)'=(\hat{\beta}_1(t/T),\ldots,\hat{\beta}_p(t/T))$ and the error variance $\hat{\sigma_p}^2(t/T)$. Consider the rescaled residuals

$$\widetilde{\varepsilon}_{t,T}=(X_{t,T}-\sum_{i=1}^{p}\hat{\beta}_i(\frac{t}{T})X_{t-i,T})/\hat{\sigma_p}(t/T), \qquad t=p+1,\ldots,T$$

and let

$$\hat{F}_T(x)=\frac{1}{T-p}\sum_{j=p+1}^{T}I_{(-\infty,x]}(\hat{\varepsilon}_{j,T}),$$

where $\hat{\varepsilon}_{t,T} = \widetilde{\varepsilon}_{t,T} - \bar{\varepsilon}$, $\bar{\varepsilon} = \frac{1}{T-p} \sum_{t=p+1}^{T} \widetilde{\varepsilon}_{t,T}$ and $I_A(x)$ is the indicator function of the set $A \subset \mathbb{R}$.

STEP 2: Generate bootstrap observations $X^+_{1,T}, X^+_{2,T}, \ldots, X^+_{T,T}$ using the fitted local autoregressive model, i.e,

$$X^+_{t,T} = \sum_{i=1}^{p} \hat{\beta}_i(\frac{t}{T}) X^+_{t-i,T} + \hat{\sigma}_p(\frac{t}{T}) \cdot \varepsilon^+_t,$$

where $X^+_{j,T} = X_{j,T}$ for $j = 1, 2, \ldots, p$ and ε^+_t are i.i.d random variables with $\varepsilon^+_t \backsim \hat{F}_T$.

STEP 3: Compute the local periodogram $I^+_N(u, \lambda)$ over a segment of length N of the bootstrap pseudo-observations $X^+_{t,T}$, i.e., compute

$$I^+_N(u, \lambda) = \frac{1}{2\pi N} \sum_{s_1=0}^{N-1} \sum_{s_2=0}^{N-1} X^+_{[uT]-N/2+s_1+1,T} X^+_{[uT]-N/2+s_2+1,T} e^{-i\lambda(s_1-s_2)}.$$

STEP 4: Compute the local kernel estimator $\hat{g}(u, \lambda)$ defined by

$$\hat{g}(u, \lambda) = \frac{1}{N} \sum_{j=-M_N}^{M_N} K_h(\lambda - \lambda_j) \frac{I_N(u, \lambda_j)}{\hat{f}_{tvAR}(u, \lambda_j)}$$

where $\{\lambda_j = 2\pi j/N, \ \ j = -M_N, \ldots, M_N\}$, $M_N = [N/2]$ and

$$\hat{f}_{tvAR}(u, \lambda) = \frac{\hat{\sigma_p}^2(u)}{2\pi} \Big| 1 - \sum_{r=1}^{p} \hat{\beta}_r(u) e^{-i\lambda r} \Big|^{-2}.$$

STEP 5: The bootstrapped local periodogram is then defined by

$$I^*_N(u, \lambda) = \hat{g}(u, \lambda) I^+_N(u, \lambda).$$

Recall that if we set $\hat{g}(u, \lambda) \equiv 1$ for all u and λ in STEP 4 of the above bootstrap procedure, then we have a version of the local autoregressive sieve bootstrap. This is so since in this case the bootstrapped periodogram $I^*_N(u, \lambda)$ is given by $I^+_N(u, \lambda)$ which is the local periodogram calculated using the replicates of the autoregressive fit. Thus if the underlying locally stationary process satisfies Assumption 2.2, then the local autoregressive sieve bootstrap procedure applied to the local periodogram can be considered as a special case of our approach to bootstrap the local periodogram.

2.4 Asymptotic Properties

In this section we investigate the asymptotic properties of our bootstrap method, summarized by steps 1-5, applied to the class of local spectral means (1.2.26) and local ratio statistics (1.2.27). Let

$$\tilde{f}(u,\lambda) = \hat{g}(u,\lambda)\hat{f}_{tvAR}(u,\lambda) \tag{2.4.5}$$

which can be considered as a prewhitening type estimator of the local spectral density $f(u,\lambda)$. Proposition A.0.2 and Lemma A.0.5 of Appendix A imply that for every $u \in (0,1)$ and $\lambda \in [-\pi,\pi]$ we have under the assumptions made there, that, as $T \to \infty$,

$$\tilde{f}(u,\lambda) \to f(u,\lambda) \tag{2.4.6}$$

in probability.

To approximate the distribution of the centered spectral mean

$$\sqrt{N}\Big(\int_{-\pi}^{\pi}\phi(\lambda)I_N(u,\lambda)d\lambda \;-\; \int_{-\pi}^{\pi}\phi(\lambda)f(u,\lambda)d\lambda\Big), \tag{2.4.7}$$

our proposal is to use the distribution of the bootstrap statistic

$$\sqrt{N}\Big(\int_{-\pi}^{\pi}\phi(\lambda)I_N^*(u,\lambda)d\lambda \;-\; \int_{-\pi}^{\pi}\phi(\lambda)\widetilde{f}(u,\lambda)d\lambda\Big).$$

Note that if instead of (1.2.26) the discretized version $2\pi N^{-1/2}\sum_{j=-M_N}^{M_N}\phi(\lambda_j)I_N(u,\lambda_j)$ is used, then the corresponding bootstrap statistic will be $2\pi N^{-1/2}\sum_{j=-M_N}^{M_N}\phi(\lambda_j)I_N^*(u,\lambda_j)$.

Our first theorem deals with the case where the underlying process fulfills Assumption 2.1 and the order p of the local approximating process remains fixed as the sample size increases.

Theorem 2.4.1. *Let Assumption 2.1 and Assumptions 2.3 to 2.6 be satisfied. For all fixed $p \in \mathbb{N}$ we have as $T \to \infty$, that*

$$\mathcal{L}\left\{\sqrt{N}\Big(\int_{-\pi}^{\pi}\phi_j(\lambda)I_N^*(u,\lambda)d\lambda - \int_{-\pi}^{\pi}\phi_j(\lambda)\widetilde{f}(u,\lambda)d\lambda\Big)_{j=1,\ldots,m}|X_1,\ldots,X_T\right\} \Rightarrow \{\xi_j\}_{j=1,\ldots,m}$$

in probability for functions $\phi_j(\cdot), j = 1,\ldots,m$ satisfying Assumption 2.4, where $\underline{\xi} = (\xi_1,\xi_2,\ldots,\xi_m)'$ is a Gaussian random vector with mean zero,

$$cov(\xi_i, \xi_j) = 2\pi\Big\{ \int_{-\pi}^{\pi} \phi_i(\lambda)\{\phi_j(\lambda) + \phi_j(-\lambda)\} f^2(u,\lambda) d\lambda + \kappa_4(p) \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_i(\lambda)\phi_j(-\mu) f(u,\lambda) f(u,\mu) d\lambda d\mu \Big\}$$

and $\kappa_4(p) = \int_0^1 \kappa_4(u,p)du - 3$ *with* $\kappa_4(u,p) = E(\widetilde{X}_p(u) - \sum_{j=1}^{p} \beta_{j,p}(u)\widetilde{X}_{p-j}(u))^4/\sigma_p^4(u)$. *Here,* $\beta_{j,p}(u)$, $j = 1, 2, \ldots, p$ *are the coefficients minimizing the mean square error* $E(\widetilde{X}_p(u) - \sum_{j=1}^{p} b_j \widetilde{X}_{p-j}(u))^2$ *and* $\sigma_p^2(u) = E(\widetilde{X}_p(u) - \sum_{j=1}^{p} \beta_{j,p}(u)\widetilde{X}_{p-j}(u))^2$.

The following lemma from Dahlhaus and Giraitis [13] gives the asymptotic distribution of local spectral means.

Lemma 2.4.2. *As* $T \to \infty$,

$$\mathcal{L}\Big\{\sqrt{N}\Big(\int_{-\pi}^{\pi} \phi_j(\lambda) I_N(u,\lambda) d\lambda - \int_{-\pi}^{\pi} \phi_j(\lambda) f(u,\lambda) d\lambda\Big)_{j=1,\ldots,m}\Big\} \Rightarrow \{\xi_j\}_{j=1,\ldots,m}$$

where $\underline{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)'$ *is Gaussian random vector with mean zero and*

$$cov(\xi_i, \xi_j) = 2\pi\Big\{ \int_{-\pi}^{\pi} \phi_i(\lambda)\{\phi_j(\lambda) + \phi_j(-\lambda)\} f^2(u,\lambda) d\lambda + \kappa_4 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_i(\lambda)\phi_j(-\mu) f(u,\lambda) f(u,\mu) d\lambda d\mu \Big\}.$$

Notice that the term κ_4 appearing in the above expression for the covariance of the limiting Gaussian distribution is due to the asymptotically vanishing dependence of the local periodogram ordinates $I_N(u,\lambda_j)$. A comparison of the above distribution to the limiting distribution given in Theorem 4.1 makes it clear that the local periodogram bootstrap manages to reproduce to some extend the effects of the dependence of the local periodogram on this limiting distribution. Furthermore, the closer is $\kappa_4(p)$ to κ_4 the closer is the limiting distribution of the bootstrap statistic to the limiting distribution of the statistic of interest. It is therefore interesting to bound the difference between the two fourth order cumulants. The following proposition gives such a bound for the case where the underlying locally stationary process $\widetilde{X}_t(u)$ is causal.

Proposition 2.4.1. *Assume that $\kappa_4 \neq 0$ (see Defintion 1.1.2) and that the locally stationary process (1.1.10) satisfies $a(u,j) = 0$ for $j < 0$. Then for any $p \in \mathbb{N}$ we have*

$$\Big|\frac{\kappa_4(p)}{\kappa_4} - 1\Big| \;\leq\; 2\int_0^1 \frac{L_2(u,p)}{1+L_2(u,p)}du \tag{2.4.8}$$

where $L_2(u,p) = \sum_{j=1}^{\infty} \ell_p^2(u,j)$ and $\ell_p(u,j) = \alpha(u,j) - \sum_{k=1}^{p} \beta_{j,p}(u)\alpha(u,j-k)$ for $j \geq 1$ and $\beta_{j,p}(u)$ are the coefficients given in Theorem 2.4.1.

To shed some light onto the usefulness of the above bound, consider as an example the case where $\widetilde{X}_t(u)$ is the following simple locally stationary moving average process:

$$\widetilde{X}_t(u) = \sigma(u)Y_t,$$

where $\inf_u \sigma(u) > 0$, $\sup_u \sigma(u) < \infty$ and Y_t is the nonivertible first order moving average process $Y_t = \varepsilon_t + \varepsilon_{t-1}$. Notice that the process $\{\widetilde{X}_t(u)\}$ above, does not satisfy Assumption 2.2. Now, straightforward calculations yield

$$\ell_p(u,j) = \begin{cases} 1 - \beta_{1,p}(u) & \text{for} \quad j = 1 \\ -\beta_{j-1,p}(u) - \beta_{j,p}(u) & \text{for} \quad j \in \{2,\ldots,p\} \\ -\beta_{p,p}(u) & \text{for} \quad j = p+1 \\ 0 & \text{for} \quad j \geq p+2, \end{cases}$$

where for $j = 1, 2, \ldots, p$,

$$\beta_{j,p}(u) = (-1)^{j-1}\Big(1 - \frac{j}{p+1}\Big).$$

Simple algebra yields then

$$\int_0^1 \frac{L_2(u,p)}{1+L_2(u,p)}du = \frac{1}{p+2},$$

which shows that for p sufficiently large the difference between the fourth order cumulants $\kappa_4(p)$ and κ_4, and consequently between the corresponding limiting Gaussian distributions, can be made arbitrary small.

We proceed with our investigation concerning the asymptotic properties of our local bootstrap procedure by considering the case where the approximating stationary process

$\{\widetilde{X}_t(u), t \in Z\}$ has an infinite order autoregressive representation and satisfies Assumption 2.2. In this case we additionally assume that the order of the fitted approximating autoregressive process increases to infinity with the sample size T. We can then establish the following theorem which deals with the properties of our bootstrap method applied to the class of local spectral means.

Theorem 2.4.3. *Let Assumption 2.1 to 2.6 be satisfied. If $p \to \infty$ such that $p^4/N \to 0$ we have as $T \to \infty$, that*

$$\mathcal{L}\left\{\sqrt{N}\Big(\int_{-\pi}^{\pi} \phi_j(\lambda) I_N^*(u,\lambda)d\lambda - \int_{-\pi}^{\pi} \phi_j(\lambda)\widetilde{f}(u,\lambda)d\lambda\Big)_{j=1,\ldots,m} |X_1,\ldots,X_T\right\} \Rightarrow \{\xi_j\}_{j=1,\ldots,m}$$

in probability where $\underline{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)'$ is Gaussian random vector with mean zero and

$$\begin{aligned} cov(\xi_i,\xi_j) &= 2\pi\Big\{\int_{-\pi}^{\pi} \phi_i(\lambda)\{\phi_j(\lambda)+\phi_j(-\lambda)\}f^2(u,\lambda)d\lambda \\ &\quad +\kappa_4 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_i(\lambda)\phi_j(-\mu)f(u,\lambda)f(u,\mu)d\lambda d\mu\Big\} \end{aligned}$$

Applying the so-called δ-method we can extend the validity of the proposed bootstrap method to the class of ratio statistics. Recall that in this case the limiting distribution of the corresponding statistics does not depend on characteristics of the error process and in particular on the fourth order cumulant of ε_t; cf. Dahlhaus and Janas [12].

Theorem 2.4.4. *Let Assumption 2.1 and Assumptions 2.3 to 2.6 be satisfied. For all fixed $p \in Ne$ we have that,*

$$\mathcal{L}\left\{\sqrt{N}\left(\frac{\int_{-\pi}^{\pi} \phi_j(\lambda) I_N^*(u,\lambda)d\lambda}{\int_{-\pi}^{\pi} I_N^*(u,\lambda)d\lambda} - \frac{\int_{-\pi}^{\pi} \phi_j(\lambda)\widetilde{f}(u,\lambda)d\lambda}{\int_{-\pi}^{\pi} \widetilde{f}(u,\lambda)d\lambda}\right)_{j=1,\ldots,m} |X_1,\ldots,X_T\right\} \Rightarrow \{\xi_j\}_{j=1,\ldots,m}$$

in probability as $T \to \infty$, where, $\underline{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)'$ is Gaussian random vector with mean zero,

$$cov(\xi_i,\xi_j) = 2\pi\Big\{\int_{-\pi}^{\pi} \psi_i(\lambda)\{\psi_j(\lambda)+\psi_j(-\lambda)\}f^2(u,\lambda)d\lambda \Big/ \Big(\int_{-\pi}^{\pi} f(u,\lambda)d\lambda\Big)^4\Big\}$$

and $\psi_j(\lambda) = \phi_j(\lambda)\int_{-\pi}^{\pi} f(u,\mu)d\mu - \int_{-\pi}^{\pi} \phi_j(\mu)f(u,\mu)d\mu$.

As we have already mentioned in Section 2.3, if the underlying process satisfies Assumption 2.2 then for $\hat{g}(u,\lambda) \equiv 1$ our approach can be considered as a local version of the so-called autoregressive sieve bootstrap. The following corollary summarizes the performance of this method for the classes of local spectral means and of local ratio statistics.

Corollary 2.4.5. *Let Assumption 2.1 to 2.6 be satisfied and set $\widehat{q}(u,\lambda) \equiv 1$ for all $u \in [0,1]$ and $\lambda \in [-\pi,\pi]$. If $p \to \infty$ such that $p^4/N \to 0$ we have as $T \to \infty$, that*

(i)

$$\mathcal{L}\left\{\sqrt{N}\Big(\int_{-\pi}^{\pi} \phi_j(\lambda) I_N^+(u,\lambda)d\lambda - \int_{-\pi}^{\pi} \phi_j(\lambda)\hat{f}_{tvAR}(u,\lambda)d\lambda\Big)_{j=1,\ldots,m} |X_1,\ldots,X_T\right\} \Rightarrow \{\xi_j\}_{j=1,\ldots,m}$$

in probability where, $\underline{\xi} = (\xi_1,\xi_2,\ldots,\xi_m)'$ is Gaussian random vector with mean zero and

$$\begin{aligned} cov(\xi_i,\xi_j) &= 2\pi\Big\{\int_{-\pi}^{\pi} \phi_i(\lambda)\{\phi_j(\lambda)+\phi_j(-\lambda)\}f^2(u,\lambda)d\lambda \\ &\quad +\kappa_4 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_i(\lambda)\phi_j(-\mu)f(u,\lambda)f(u,\mu)d\lambda d\mu\Big\} \end{aligned}$$

and

(ii)

$$\mathcal{L}\left\{\sqrt{N}\left(\frac{\int_{-\pi}^{\pi} \phi_j(\lambda) I_N^+(u,\lambda)d\lambda}{\int_{-\pi}^{\pi} I_N^+(u,\lambda)d\lambda} - \frac{\int_{-\pi}^{\pi} \phi_j(\lambda)\hat{f}_{tvAR}(u,\lambda)d\lambda}{\int_{-\pi}^{\pi} \hat{f}_{tvAR}(u,\lambda)d\lambda}\right)_{j=1,\ldots,m} |X_1,\ldots,X_T\right\} \Rightarrow \{\xi_j\}_{j=1,\ldots,m}$$

in probability where, $\underline{\xi} = (\xi_1,\xi_2,\ldots,\xi_m)'$ is Gaussian random vector with mean zero,

$$cov(\xi_i,\xi_j) = 2\pi\Big\{\int_{-\pi}^{\pi} \psi_i(\lambda)\{\psi_j(\lambda)+\psi_j(-\lambda)\}f^2(u,\lambda)d\lambda/\Big(\int_{-\pi}^{\pi} f(u,\lambda)d\lambda\Big)^4\Big\}$$

and $\psi_j(\lambda) = \phi_j(\lambda)\int_{-\pi}^{\pi} f(u,\mu)d\mu - \int_{-\pi}^{\pi}\phi_j(\mu)f(u,\mu)d\mu$.

From the above corollary it is clear that if the underlying process satisfies Assumption 2.2, then the local autoregressive sieve bootstrap leads to asymptotically valid approximations of the distribution of the statistics of interest. Recall that in this case the local periodogram bootstrap method using the nonparametric correction function $\hat{g}(\cdot,\cdot)$ works too; cf. Theorem 2.4.2. However, if the underlying process does not have locally an infinite order autoregressive representation , i.e., if it does not satisfy Assumption 2.2 and we apply a pth order autoregressive bootstrap where p is fixed, then it can be shown that in this case,

$$\mathcal{L}\left\{\sqrt{N}\left(\frac{\int\limits_{-\pi}^{\pi}\phi(\lambda)I_N^{+}(u,\lambda)d\lambda}{\int\limits_{-\pi}^{\pi}I_N^{+}(u,\lambda)d\lambda}-\frac{\int\limits_{-\pi}^{\pi}\phi(\lambda)\hat{f}_{tvAR}(u,\lambda)d\lambda}{\int\limits_{-\pi}^{\pi}\hat{f}_{tvAR}(u,\lambda)d\lambda}\right)|X_1,\ldots,X_T\right\}\Rightarrow N(0,\sigma_\phi^2(p))$$

in probability where,

$$\sigma_\phi^2(p)=2\pi\Big\{\int\limits_{-\pi}^{\pi}\psi(\lambda)\{\psi(\lambda)+\psi(-\lambda)\}f_{tvAR}^2(u,\lambda)/\Big(\int\limits_{-\pi}^{\pi}f_{tvAR}(u,\lambda)d\lambda\Big)^4\Big\},$$

$\psi(\lambda)=\phi(\lambda)\int_{-\pi}^{\pi}f(u,\mu)d\mu-\int_{-\pi}^{\pi}\phi(\mu)f(u,\mu)d\mu$ and

$$f_{tvAR}(u,\lambda)=\frac{\sigma_p^2(u)}{2\pi}\Big|1-\sum_{r=1}^{p}\beta_{r,p}(u)e^{-i\lambda r}\Big|^{-2}.$$

This makes clear that even for the class of ratio statistics where the limiting distribution is free from parameters of the error process, the pure autoregressive sieve bootstrap does not work if the approximating process $\{X_t(u),t\in\mathbb{Z}\}$ does not have the infinite order autoregressive representation stated in Assumption 2.2. This is in contrast to the local periodogram bootstrap proposed in this chapter which due to the nonparametric correction in the frequency domain via the estimated function $\hat{g}(u,\lambda)$ leads to asymptotically valid approximations in this case and works therefore for a larger class of stochastic processes and for a larger class of statistics.

We conclude this section by an application of the bootstrap to estimate the distribution of nonparametric local spectral density estimators. An interesting class of such estimators is given by

$$\hat{f}(u,\lambda)=\frac{1}{N}\sum_{j=-M_N}^{M_N}K_b(\lambda-\lambda_j)I_N(u,\lambda_j) \tag{2.4.9}$$

where $K(\cdot)$ is a kernel satisfying Assumption 2.5, $K_b(\cdot)=b^{-1}K(\cdot/b)$ and b is a smoothing bandwidth satisfying $b\to 0$ such that $Nb\to\infty$ as $N\to\infty$. Suppose we are interested in estimating the distribution of

$$\sqrt{Nb}\left(\hat{f}(u,\lambda)-E(\hat{f}(u,\lambda))\right). \tag{2.4.10}$$

For this, the bootstrap analogue

$$\sqrt{Nb}\left(\hat{f}^*(u,\lambda) - E^*(\hat{f}^*(u,\lambda))\right) \tag{2.4.11}$$

can be used, where

$$\hat{f}^*(u,\lambda) = \frac{1}{N}\sum_{j=-M_N}^{M_N} K_b(\lambda-\lambda_j) I_N^*(u,\lambda_j). \tag{2.4.12}$$

It can be shown under certain assumptions that (2.4.10) converges weakly to a Gaussian distribution with mean zero and variance given by $\tau^2(u,\lambda) = (1+\delta(\lambda))f^2(u,\lambda)\int_{-\pi}^{\pi} K^2(w)dw$, where $\delta(\lambda) = 1$ if $\lambda = 0$ or being a multiple of $\pm\pi$ and $\delta(\lambda) = 0$ else; (see the proof of Theorem 2.4.6 below). The following theorem shows asymptotic validity of the proposed bootstrap approximation.

Theorem 2.4.6. *Let Assumptions 2.1, 2.3, 2.5, and 2.6 be satisfied. Suppose that the smoothing bandwidth $b = b(N)$ satisfies $b \to 0$ such that $Nb \to \infty$ as $N \to \infty$. Then for all fixed $p \in Ne$, we have that,*

$$\mathcal{L}\left\{\sqrt{Nb}\left(\hat{f}^*(u,\lambda) - E(\hat{f}^*(u,\lambda))\right)|X_1,\ldots,X_T\right\} \Rightarrow N(0,\tau^2(u,\lambda))$$

in probability, as $T \to \infty$, where $\tau^2(u,\lambda) = (1+\delta(\lambda))f^2(u,\lambda)\int_{-\pi}^{\pi} K^2(w)dw$.

2.5 Some remarks on choosing the resampling parameters

From the previous discussion it is clear that implementation of our method requires the selection of three parameters, that is of the time window width N, of the order of the locally fitted autoregression p and of the frequency domain smoothing bandwidth h.

Concerning the time window width N, we stress here the fact that the selection of its length is not inherit to our local bootstrap procedure but to any statical inference procedure for locally stationary process which is based on the local periodogram. That is, if some procedure to select N exist, which is needed in order to calculate statistics (1.2.26) and (1.2.27), then the same time window N can be used in the local bootstrap procedure applied to infer properties of these statistics.

Nevertheless, the selection N can be investigated theoretically by minimizing over all time points u and over all frequencies λ, an L^2-distance between the estimated locally spectral density $\tilde{f}(u,\lambda) = \hat{g}(u,\lambda)\hat{f}_{tvAR}(u,\lambda)$ and its theoretical counterpart $f(u,\lambda)$. In particular, selection of N can be based on minimization of the leading terms of the time integrated mean square error

$$MISE = \int_0^1 \int_{-\pi}^{\pi} E(\hat{f}(u,\lambda) - f(u,\lambda))^2 d\lambda du$$

where $\hat{f}(u,\lambda) = b^{-1}\int_{-\pi}^{\pi} K((\lambda-\mu)/b) I_N(u,\mu)d\mu$ is a kernel estimator of the local spectral density $f(u,\lambda)$.

From Dahlhaus [8], Theorem 2.2, we get that

$$E(\hat{f}(u,\lambda)) = f(u,\lambda) + \frac{1}{24}\frac{N^2}{T^2}\frac{\partial^2}{\partial u^2}f(u,\lambda) + \frac{1}{2}b^2\int x^2K(x)dx\frac{\partial^2}{\partial\lambda^2}f(u,\lambda) + o(\frac{N^2}{T^2} + \frac{\log N}{N} + b^2)$$

and

$$Var(\hat{f}(u,\lambda)) = (Nb)^{-1}\frac{1}{12}f^2(u,\lambda)\int x^2K(x)dx.$$

Using these results the mean square error equals

$$\begin{aligned} E(\hat{f}(u,\lambda) - f(u,\lambda))^2 &= (Nb)^{-1}\frac{1}{12}f^2(u,\lambda)\int x^2K(x)dx + \frac{1}{576}\frac{N^4}{T^4}(\frac{\partial^2}{\partial u^2}f(u,\lambda))^2 \\ &\quad + \frac{1}{24}\frac{N^2b^2}{T^2}\frac{\partial^2}{\partial u^2}f(u,\lambda)\int x^2K(x)dx\frac{\partial^2}{\partial\lambda^2}f(u,\lambda) + O(b^4). \end{aligned}$$

Notice the above expression of the mean square error is dominated by the first two terms which lead to the following approximation:

$$AMISE = (Nb)^{-1}\frac{1}{12}\int_0^1\int_{-\pi}^{\pi} f^2(u,\lambda)d\lambda du\int x^2K(x)dx + \frac{1}{576}\frac{N^4}{T^4}\int_0^1\int_{-\pi}^{\pi}(\frac{\partial^2}{\partial u^2}f(u,\lambda))^2 d\lambda du.$$

Now for $b = N^{-\delta}$, and $0 < \delta < 1$, the above AMISE is minimized for N_{opt} given by

$$N_{opt} = T^{4/(5-\delta)}\left(\frac{12(1-\delta)\int_0^1\int_{-\pi}^{\pi} f^2(u,\lambda)d\lambda du\int x^2K(x)dx}{\int_0^1\int_{-\pi}^{\pi}(\frac{\partial^2}{\partial u^2}f(u,\lambda))^2 d\lambda du}\right)^{1/(5-\delta)}.$$

Clearly, implementation of this rule to select N in practice requires estimates of the quantities $\int_0^1\int_{-\pi}^{\pi} f^2(u,\lambda)d\lambda du$ and $\int_0^1\int_{-\pi}^{\pi}(\frac{\partial^2}{\partial u^2}f(u,\lambda))^2 d\lambda du$.

The selection of the autoregressive order p and of the smoothing bandwidth h are more inherit to our local bootstrap procedure. Concerning the order p of the fitted autoregressive process a practical rule to determine this order is to use a local version of the AIC-criterion, like for instance the one proposed by Dahlhaus [9]. We mention here however, that as our numerical examples show, due to the nonparametric correction in the frequency domain, the numerical results obtained are less sensitive to the choice of the order p used in the autoregressive fit.

Concerning the choice of the smoothing parameter h one possible approach is to select h using a local version of a cross-validation criterion analogous to the one proposed by Beltrão and Bloomfield [2]. To elaborate on, this approach uses as a starting point a generalization of the Whittle function, given by

$$\sum_{i=1}^{M} \sum_{j=-M_N}^{M_N} \left\{ \log f(u_i, \lambda_j) + \frac{I_N(u_i, \lambda_j)}{f(u_i, \lambda_j)} \right\}. \tag{2.5.13}$$

A leave-one-out estimator for $q(u, \lambda_j)$ is given by

$$\hat{q}_{-j}(u, \lambda_j) = \frac{1}{N} \sum_{j \in N_j} K_h(\lambda_j - \lambda_s) \frac{I_N(u, \lambda_j)}{\hat{f}_{tvAR}(u, \lambda_j)} \tag{2.5.14}$$

where $N_j = \{s : -M_N \leq s \leq M_N \text{ and } j - s \neq \pm j \mod M_N\}$, see Beltrão and Bloomfield [2] and Kreiss and Paparoditis [24]. Notice that $\hat{q}_{-j}$ is a kernel estimator of q obtained by ignoring the jth local periodogram ordinate. Now, substituting $q(u, \lambda_j) f_{tvAR}(u, \lambda_j)$ for $f(u, \lambda_j)$, $\hat{q}_{-j}(u, \lambda_j)$ for $q(u, \lambda_j)$ and $\hat{f}_{tvAR}(u, \lambda_j)$ for $f_{tvAR}(u, \lambda_j)$ in (3.3.11) leads after ignoring the factor $\log f_{tvAR}(u, \lambda_j)$ to the function

$$CV(h) = \sum_{i=1}^{M} \sum_{j=1}^{N} \left\{ \log \hat{q}_{-j}(u_i, \lambda_j) + \frac{I_N(u_i, \lambda_j) / \hat{f}_{tvAR}(u_i, \lambda_j)}{\hat{q}_{-j}(u_i, \lambda_j)} \right\}, \tag{2.5.15}$$

which can be used as a cross-validation-type criterion to select h.

2.6 Numerical Results

2.6.1 Simulations

We investigate the performance of our method in finite sample situations and compare the results of the local periodogram bootstrap procedure to that of the pure local sieve autoregressive bootstrap. For this we consider a case where both methods work asymptotically, i.e, where both methods lead to asymptotically correct approximations. In particular, we consider the problem of estimating the standard deviation of the time varying autocorrelation estimator

$$\hat{\rho}(u,1) = \frac{\hat{c}(u,1)}{\hat{c}(u,0)} = \frac{\int_{-\pi}^{\pi} I_N(u,\lambda)\exp(i\lambda)d\lambda}{\int_{-\pi}^{\pi} I_N(u,\lambda)d\lambda}$$

for different values of u in the interval $(0,1)$. Samples of length $T = 512$ observations from the time varying MA(1)-process

$$X_{t,T} = 1.1\cos(1.5 - \cos(4\pi t/T))\varepsilon_{t-1} + \varepsilon_t \tag{2.6.16}$$

are considered, where the ε_t's are i.i.d with $\varepsilon_t \sim N(0,1)$.

To estimate the exact standard deviation of $\hat{\rho}(u,1)$ calculated over 40 equally spaced points u in the interval $(0,1)$, we generate 5000 samples of the above process. The so obtained estimates of the exact standard deviation are presented in Figure 2.1(a) and Figure 2.1(b) by solid lines. The small crosses in these lines indicate the particular points u in the interval $[0,1]$ for which the estimates of the standard deviation of $\hat{\rho}(u,1)$ have been calculated. To investigate the performance of the bootstrap, we generate 50 different series from the above process and for each of these series we use our bootstrap method by producing 300 bootstrap local periodogram replicates. For the corresponding bootstrap estimates of the standard deviation we calculate the mean, the standard deviation and the mean square error.

We apply both local periodogram bootstrap procedures using two different values of the autoregressive order p. For the nonparametric correction of the local periodogram bootstrap the Bartlett-Priestley kernel given by $K(x) = 3(4\pi)^{-1}(1-(x/\pi)^2)$ for $|x| \leq \pi$ and a bandwidth of $h = 0.2$ has been used. We first fit a time varying autoregressive model of order $p = 1$ and use a window length of $N = 40$ observations. The results for both bootstrap methods for $p = 1$ are presented in Figure 2.1(a), Figure 2.1(c) and Figure 2.1(e). Figure 2.1(a) clearly shows the effect of the nonparametric correction via the function $\hat{g}(u,\lambda)$. The mean estimate using our method captures quite closely the exact standard deviation of $\hat{\rho}(u,1)$, while the mean estimate using the pure autoregressive bootstrap with $p = 1$ is very biased. Although, Figure 2.1(c) shows that the standard deviation of our method is larger than that of the pure autoregressive sieve bootstrap, the mean square error of both procedures shown in Figure 2.1(e) clearly demonstrates that our method performs much better.

In order to see the effects of increasing the autoregressive order p, we fit in a second run to the same set of series a time varying autoregressive model of order $p = 3$. As Figure 2.1(b) shows, in this case the mean estimates using both methods capture the exact standard deviation quite well. Comparing Figure 2.1(c) and Figure 2.1(d) we see that there is an increase in the standard deviation using the pure local autoregressive sieve bootstrap which is due to the increase of the autoregressive order p from $p = 1$ to $p = 3$. For $p = 3$ the behavior of the mean square errors using both methods is very similar with a slight advantage for our method, see Figure 2.1(f). A detailed presentation of the simulation results is given in Table 2.1.

To see the quality of the asymptotic normal approximation and to compare its performance with that of the bootstrap, we calculate the standard deviation of the first order sample autocorrelation using the asymptotic formula for the variance given in Theorem 2.4.4. To estimate the local spectral density $f(u, \lambda)$ involved, we use the non parametric estimator $\hat{f}(u, \lambda)$ with the bandwidth minimizing the mean square error $E(\hat{f}(u, \lambda) - f(u, \lambda))^2$. The results are shown in Figure 2.2.

From this limited simulation study we can clearly conclude that the bootstrap procedure proposed leads to very accurate estimates of the standard deviation of the time varying first order autocorrelation $\hat{\rho}(u, 1)$. These estimates outperform those of the other methods available in the literature.

Table 2.1: Local Periodogram Bootstrap (LPB) and Local Autoregressive Bootstrap (LARB) estimates of the standard deviation of the first-order local sample autocorrelation.

	Est.	LPB			LARB		
	Exact	Mean($\hat{\sigma}_1^*$)	SD($\hat{\sigma}_1^*$)$\times$10	MSE($\hat{\sigma}_1^*$)$\times 10^3$	Mean($\hat{\sigma}_1^*$)	SD($\hat{\sigma}_1^*$)$\times$10	MSE($\hat{\sigma}_1^*$)$\times 10^3$
u	σ_1	p=1, h=0.2			p=1		
0.200	0.122	0.120	0.1405	0.1993	0.151	0.0829	0.8944
0.246	0.121	0.115	0.1640	0.2938	0.147	0.0809	0.7506
0.292	0.120	0.123	0.1564	0.2494	0.150	0.0900	0.9799
0.338	0.149	0.148	0.1775	0.3097	0.164	0.1040	0.3434
0.385	0.163	0.165	0.2726	0.7311	0.176	0.1484	0.3782
0.431	0.121	0.127	0.1844	0.3639	0.152	0.0794	1.0105
0.477	0.115	0.117	0.1626	0.2631	0.147	0.0776	1.1039
0.523	0.113	0.118	0.1544	0.2539	0.148	0.0806	1.2923
0.569	0.120	0.124	0.1343	0.1893	0.153	0.0914	1.1473
0.615	0.164	0.163	0.2416	0.5734	0.175	0.1108	0.2317
0.662	0.149	0.147	0.2444	0.5880	0.167	0.0925	0.4005
0.708	0.121	0.120	0.1352	0.1803	0.150	0.0864	0.9111
0.754	0.121	0.114	0.1684	0.3254	0.146	0.0863	0.7470
0.800	0.124	0.122	0.1951	0.3791	0.151	0.0818	0.7426
		p=3, h=0.2			p=3		
0.200	0.122	0.121	0.1586	0.2473	0.126	0.1673	0.2889
0.246	0.121	0.118	0.1775	0.3178	0.122	0.1771	0.3087
0.292	0.120	0.125	0.1735	0.3208	0.130	0.1839	0.4319
0.338	0.149	0.150	0.1851	0.3377	0.154	0.1997	0.4149
0.385	0.163	0.161	0.2456	0.5943	0.166	0.2539	0.6227
0.431	0.121	0.127	0.1865	0.3718	0.128	0.1831	0.3799
0.477	0.115	0.120	0.1703	0.3105	0.119	0.1668	0.2942
0.523	0.113	0.119	0.1710	0.3237	0.121	0.1780	0.3682
0.569	0.121	0.127	0.1437	0.2439	0.126	0.1448	0.2417
0.615	0.164	0.161	0.2213	0.4938	0.164	0.2361	0.5464
0.662	0.149	0.148	0.2493	0.6099	0.153	0.2553	0.6555
0.708	0.121	0.121	0.1504	0.2218	0.125	0.1769	0.3242
0.754	0.121	0.115	0.1808	0.3499	0.121	0.1953	0.3749
0.800	0.124	0.123	0.1847	0.3356	0.128	0.1957	0.3903

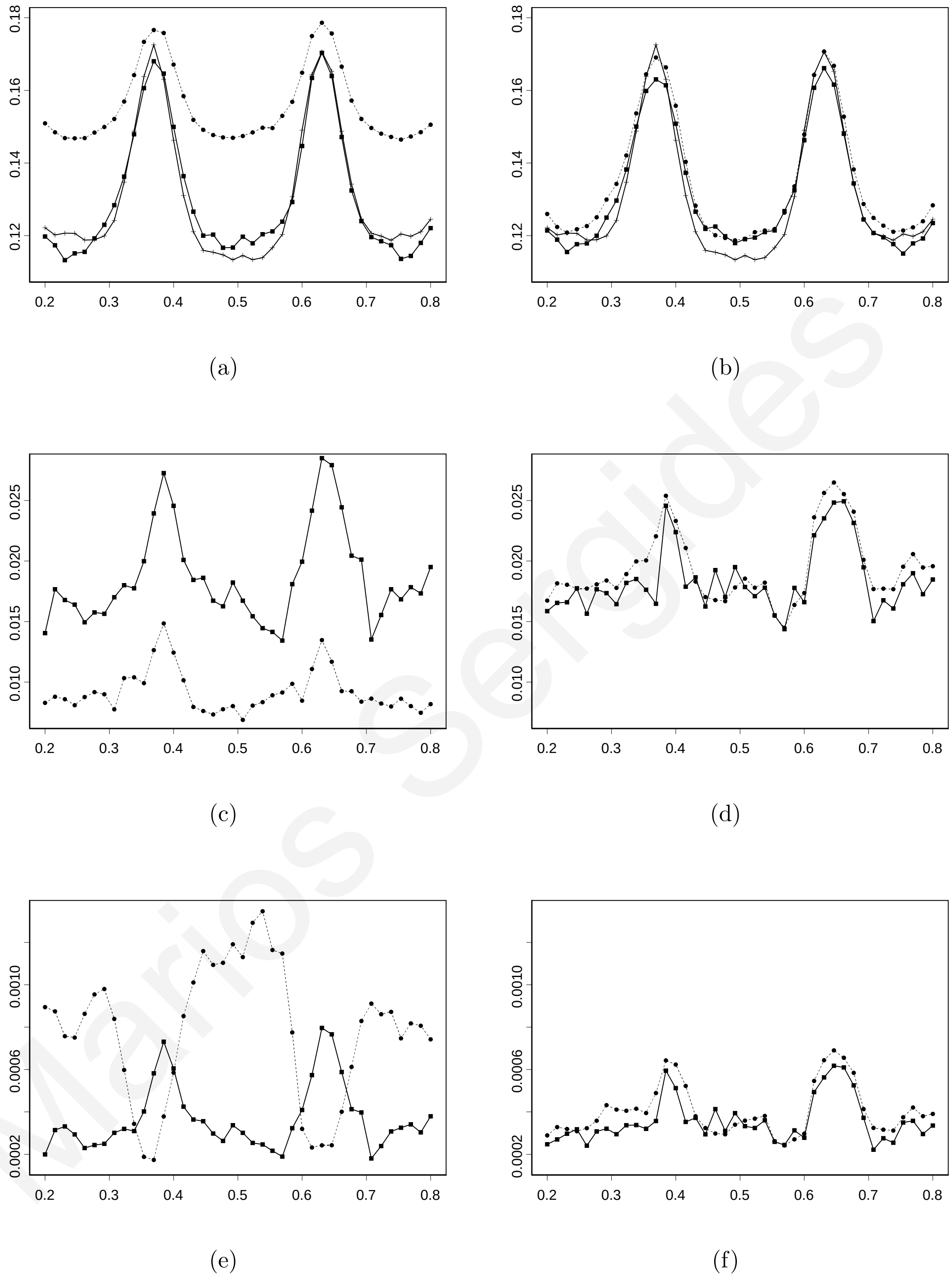


Figure 2.1: Estimated mean, variance, and mean square error of the standard deviation of the first-order sample autocorrelation. On the left a tvAR(1) model is fitted locally to the series while on the right a tvAR(3) model is fitted. The solid line in (a) and (b) is the estimated exact standard deviation. The dashed line with circles and the solid line with squares in (a) and (b) are the mean, in (c) and (d) the variance and in (e) and (f) the mean square error of the standard deviation estimates of the pure tvAR($\cdot$) bootstrap and of the local periodogram bootstrap respectively.

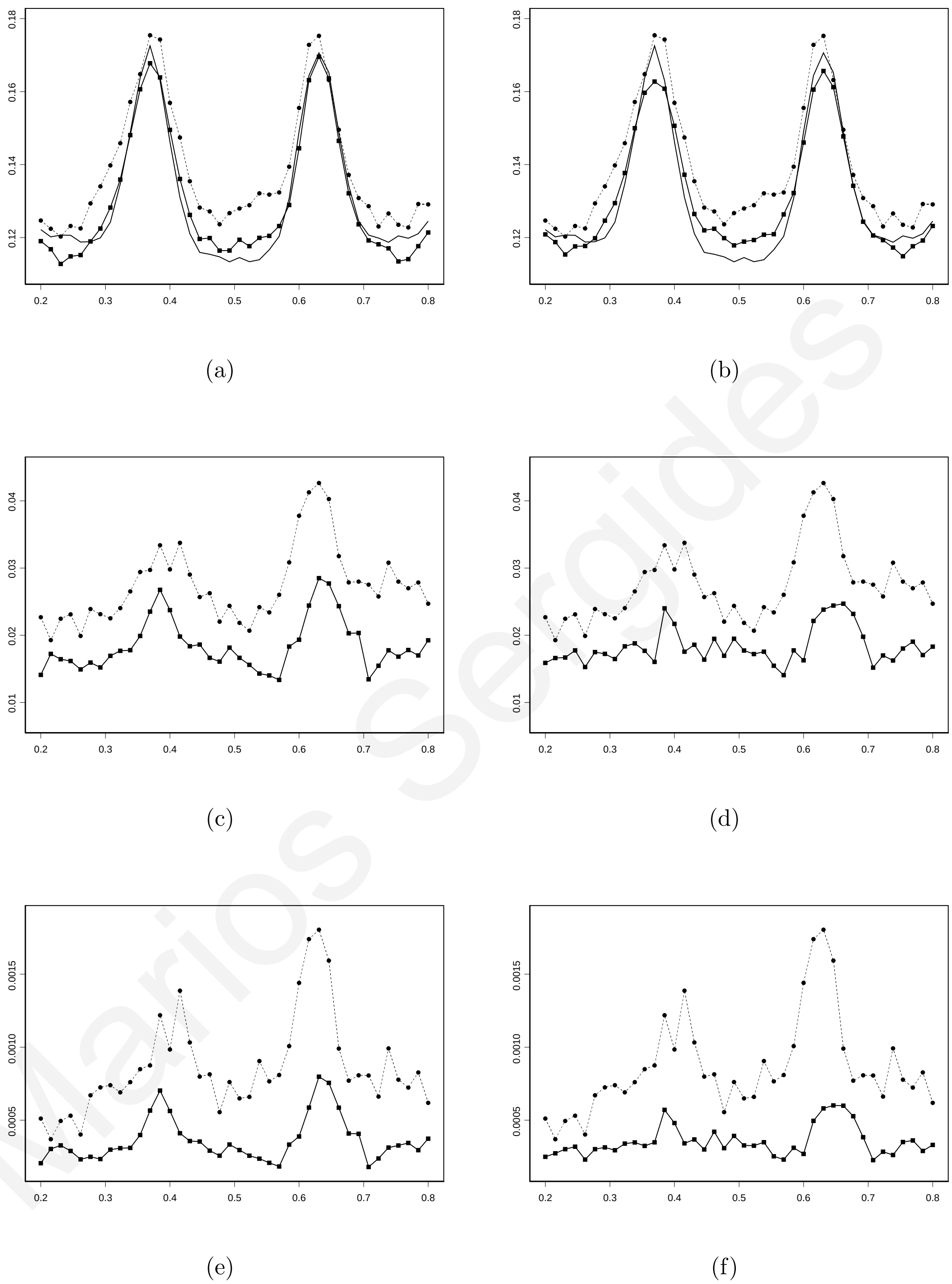


Figure 2.2: Estimated mean, variance, and mean square error of the standard deviation of the first-order sample autocorrelation. On the left a tvAR(1) model is fitted locally to the series while on the right a tvAR(3) model is fitted. The solid line in (a) and (b) is the estimated exact standard deviation. The dashed line with circles and the solid line with squares in (a) and (b) are the mean, in (c) and (d) the variance and in (e) and (f) the mean square error of the standard deviation estimates of the normal approximation and of the local periodogram bootstrap respectively.

2.6.2 A real-data example

In this example, we consider $T = 2048$ observations of the so-called earthquake data obtained from Shumway and Stoffer [39]. The series is recorded at a seismic recording station in Scandinavia where the recording instruments observe earthquakes and mining explosions. It is shown in Figure 2.3.

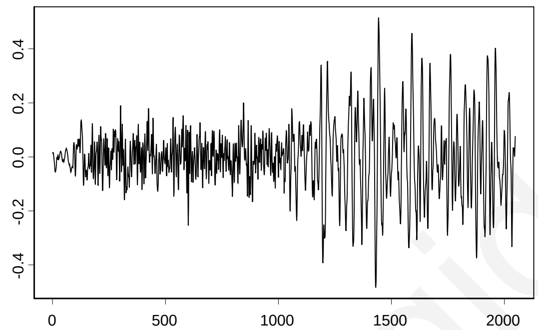


Figure 2.3: 2048 observations of an earthquake obtained by Shumway and Stoffer [39].

For this data set, we are interested in estimating the time-varying autocorrelation function $\widetilde{\rho}(u, \tau)$ for values of τ equal to $\tau = 1, 2, 3, 4$ and in constructing pointwise confidence intervals for the unknown $\rho(u, \tau)$ using the local periodogram bootstrap. To estimate $\rho(u, \tau)$ we now use a modified kernel estimator given by, $\widetilde{\rho}(u, \tau) = \widetilde{c}(u, \tau)/\widetilde{c}(u, 0)$, where

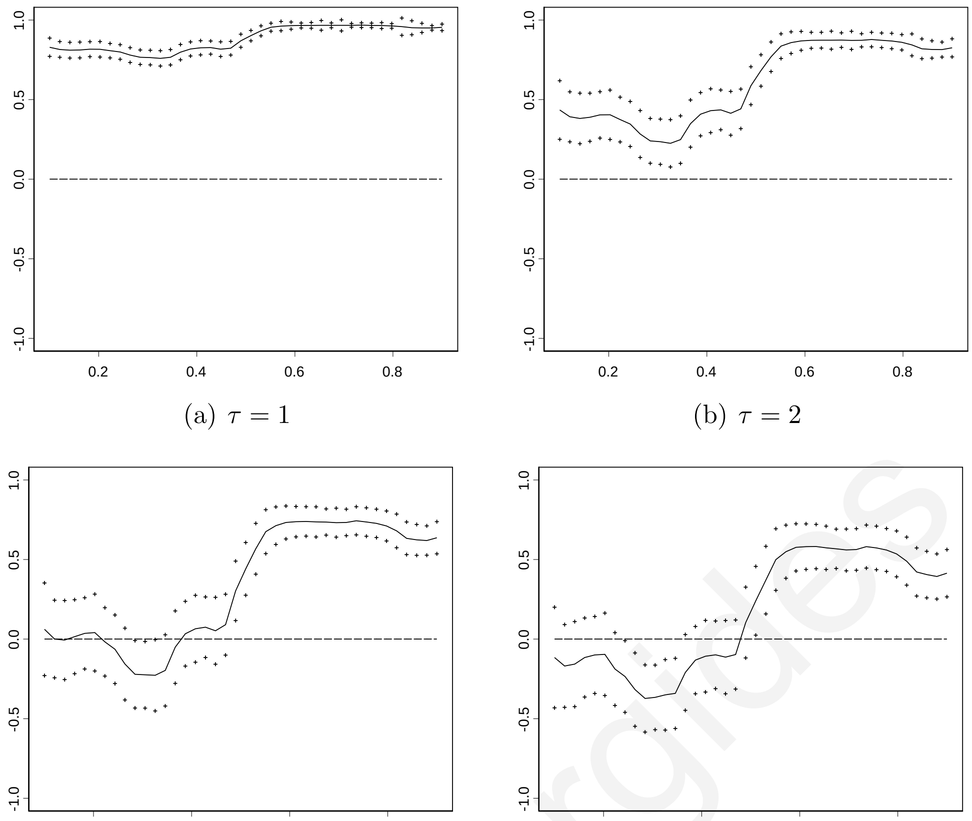
$$\widetilde{c}(u, \tau) = \frac{\sum_{t=1}^{T} K\left(\frac{uT-(t+\tau/2)}{N}\right) X_{t,T} X_{t+\tau,T}}{\sum_{t=1}^{T} K\left(\frac{uT-(t+\tau/2)}{N}\right)},$$

$K(x) = (3/2)(1 - 4x^2)$ for $|x| \leq 0.5$ and $N = 250$ observations; cf. Dahlhaus (2003). The above estimator has been calculated for 40 equally spaced points u in the interval $[0, 1]$. The estimated functions $\widetilde{\rho}(u, \tau)$ are shown in Figure 2.4(a) to Figure 2.4(d) by a solid line.

To estimate the standard deviation of $\widetilde{\rho}(u, \tau)$ we use our local periodogram bootstrap procedure. For this, we fit locally an autoregressive model of order $p = 3$ and use a window length of $N = 250$ observations. The kernel estimates have been calculated using the Bartlett-Priestley's kernel together with the bandwidth $h = 0.1$. For each of the 40 points u_i considered, we calculate the corresponding bootstrap standard deviation using B=1000 bootstrap replications. A simple 95% pointwise bootstrap confidence interval has been then obtained using the limiting distribution of $\widetilde{\rho}(u,\tau)$ and formula

$$[\widetilde{\rho}(u_i,\tau) - 1.96 s^*(u_i,\tau)\ ,\ \widetilde{\rho}(u_i,\tau) + 1.96 s^*(u_i,\tau)],$$

where $s^*(u_i,\tau)$ denotes the bootstrap estimate of the standard deviation of $\widetilde{\rho}(u_i,\tau)$. The estimated functions $\widetilde{\rho}(u_i,\tau)$ together with the so obtained 95% pointwise bootstrap confidence intervals are shown in Figure 2.4. As this figure shows the earthquake leads to a change in the correlation structure of the series. In particular, in the first part of the series the observations are less correlated compared to the second part where the earthquake occurred.



(a) $\tau = 1$ (b) $\tau = 2$

(c) $\tau = 3$ (d) $\tau = 4$

Figure 2.4: Plots of the estimated sample local autocorrelation function $\widetilde{\rho}(u, \tau)$ against u for values of $\tau = 1, 2, 3, 4$ of the earthquake data together with 95% pointwise local periodogram bootstrap confidence intervals.

Chapter 3

Testing semi-parametric hypothesis for locally stationary processes

3.1 The Testing procedure

3.1.1 The set-up

Following Dahlhaus [9] we consider in this section triangular arrays $\{\mathbf{X}_T\}_{T\in\mathbb{N}}$, $\mathbf{X}_T = \{X_{t,T}, t = 1, \ldots, T\}$ of stochastic processes which are locally stationary.

Assumption 3.1.1. *For all $T \in \mathbb{N}$, $\{\mathbf{X}_T\}_{T\in\mathbb{N}}$ is a Gaussian locally stationary process satisfying Definition 1.1.1*

The aim of this chapter is to develop tests of the hypothesis that the time-varying local spectral density $f(u, \lambda)$ has a semiparametric structure. To elaborate on the kind of null and alternative hypothesis considered, let $\mathcal{F}_{\mathcal{LS}}$ be the set of local spectral densities of processes satisfying Assumption 3.1.1 and denote by $\mathcal{F}_{\mathcal{PLS}} \subseteq \mathcal{F}_{\mathcal{LS}}$ a semiparametric model class of local spectral densities, i.e.,

$$\mathcal{F}_{\mathcal{PLS}} = \{f(u,\lambda) = f(u,\lambda;\vartheta(u)), \quad \vartheta(u) = (\vartheta_1(u),\ldots,\vartheta_m(u)), m\in\mathbb{N}, u\in[0,1], \lambda\in\mathbb{R}\},$$

where $\vartheta_i(\cdot) : [0, 1] \to \mathbb{R}$, $i = 1, 2, \ldots, m$, are appropriately defined real-valued functions. We assume that in the set $\mathcal{F}_{\mathcal{PLS}}$, the time-varying local spectral density $f(u, \lambda, \vartheta(u))$ is fully determined by the unknown functions $\vartheta_i(\cdot)$, $i = 1, 2, \ldots, m$, and as we will see in the sequel, we impose some rather mild assumptions on $\vartheta(\cdot)$ allowing for several interesting classes of semiparametric models.

To give one important example which fits in the above set-up, consider the case where $\mathcal{F}_{\mathcal{PLS}}$ is the semiparametric class of local spectral densities possessed by the class of time-varying autoregressive moving-average (tvARMA) models. Recall that a locally stationary process $\{X_{t,T}\}$ satisfying Assumption 2.1 has a tvARMA(p,q) representation if $X_{t,T}$ is generated by the equation

$$X_{t,T} + \sum_{j=1}^{p} a_j(t/T) X_{t-j,T} = \varepsilon_{t,T} + \sum_{j=1}^{q} b_j(t/T)\varepsilon_{t-j,T} \tag{3.1.1}$$

where $a_0(\cdot) \equiv b_0(\cdot) \equiv 1$, the ε_t's are i.i.d. $N(0, \sigma^2(t/T))$ distributed random variables, $a_p(u) \neq 0$ and $b_q(u) \neq 0$. Furthermore if all functions $\alpha_j(\cdot)$ and $\beta_k(\cdot)$ as well as the variance function $\sigma^2(\cdot)$ are of bounded variation and $\sum_{j=0}^{p} \alpha_j(u) z^j \neq 0$ for all $u \in [0,1]$ and all $0 < |z| \leq 1+\delta$ for some $\delta > 0$, then model (3.1.1) belongs to the locally stationary process class; see Dahlhaus [7]. Recall that, model (3.1.1) possesses a time-varying spectral density given by

$$f(u,\lambda;\vartheta(u)) = \frac{\sigma^2(u)}{2\pi} \left| \sum_{j=0}^{q} b_j(u) e^{i\lambda j} \right|^2 \Big/ \left| \sum_{j=0}^{p} a_j(u) e^{i\lambda j} \right|^2,$$

where $\vartheta(u) = (a_1(u), \ldots, a_p(u), b_1(u), \ldots, b_q(u), \sigma^2(u))$.

Based on the above discussion, the testing problem considered in this paper is described by

$$H_0 : f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{PLS}} \quad \text{vs} \quad H_1 : f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{\mathcal{PLS}}. \tag{3.1.2}$$

The specific case where $\vartheta(u)$ is a constant function of the time variable u, that is where $\vartheta(u) = (\vartheta_1, \ldots, \vartheta_m) \in \Theta \subset \mathbb{R}^m$ for all $u \in (0,1)$, is also allowed by (3.1.2). Such a case occurs for instance if one is interested in testing the null hypothesis that the underlying stochastic process is a parametric stationary process against the alternative of a time-varying locally stationary process.

3.1.2 The test statistic

We start our construction of the test statistic by first considering the tapered local periodogram defined for $N < T$, $N \in \mathbb{N}$, by

$$I_N(u,\lambda) = \frac{1}{2\pi H_{2,N}(0)} |d_N(u,\lambda)|^2, \tag{3.1.3}$$

where

$$d_N(u,\lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) X_{[uT]-N/2+s+1} e^{-i\lambda s}$$

and

$$H_{k,N}(\lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) e^{-i\lambda s}.$$

To introduce, the basic statistic used, suppose first for simplicity that the parametric curves $\vartheta(u)$ determining the local spectral density $f(u, \lambda; \vartheta(u))$ under the null hypothesis are known, that is that $\vartheta(u) = \vartheta_0(u)$. Consider then the random variables

$$Y(u, \lambda_j) = \frac{I_N(u, \lambda_j)}{f(u, \lambda_j; \vartheta_0(u))}, \qquad j = -[(N-1)/2], \ldots, [N/2].$$

It is easy to see that if the null hypothesis is true, then

$$E[Y(u, \lambda_j)] \;=\; 1 \;+\; O(N/T + 1/N),$$

for all $u \in [0, 1]$ and $\lambda_j \in (-\pi, \pi]$. Furthermore, if the alternative hypothesis is true, i.e., if $f(u, \lambda_j) \neq f(u, \lambda_j; \vartheta_0(u))$, then

$$E[Y(u, \lambda_j)] = \frac{f(u, \lambda_j)}{f(u, \lambda_j; \vartheta_0(u))} + O(N/T + 1/N),$$

where the function $f(\cdot, \cdot)/f(\cdot, \cdot; \vartheta_0(\cdot))$ is different from the unit function on $[0, 1] \times (-\pi, \pi]$.

Motivated by the above observations the idea used to obtain a test statistic for the null hypothesis that $f(u, \lambda) = f(u, \lambda, \vartheta_0(u))$, is to estimate first non-parametrically the mean function

$$q(u, \lambda) = E[Y(u, \lambda) - 1]$$

and then to evaluate its distance from the zero function using an appropriate L^2-distance measure. To elaborate on, for given $u \in (0, 1)$ and $\lambda \in [0, \pi]$, we use the kernel estimator

$$\hat{q}(u, \lambda) = \frac{1}{N} \sum_j K_b(\lambda - \lambda_j) \left(\frac{I_N(u, \lambda_j)}{f(u, \lambda_j; \vartheta_0(u))} - 1 \right) \qquad (3.1.4)$$

to estimate the unknown mean function $q(u, \lambda)$ non-parametrically. Here $K_b(\cdot) = b^{-1}K(\cdot/b)$ where $K(\cdot)$ is an appropriate defined kernel and b a smoothing bandwidth satisfying certain conditions; see Assumption 3.2.2 below.

To proceed with the construction of the test statistic proposed, we calculate $\hat{q}(u_j, \lambda)$ for different instants of time u_j by using the local periodogram $I_N(u_j, \lambda)$ for segments of observations having midpoints $u_j = t_j/T$, where $t_j := S(j-1) + N/2$ and $j = 1, \ldots, M$. Here the constant S denotes the shift from segment to segment while M refers to the total number of time points in the interval (0,1) considered. Note that by the above construction we have $T = S(M-1) + N$. Now, using a L^2-measure to evaluate the distance of the so estimated mean function $\hat{q}(u_j, \lambda)$ from the zero function and averaging over all time points $u_j = t_j/T$ and over all frequencies λ considered, we end-up with the test statistic

$$Q_{0,T} = \frac{1}{M} \sum_{s=1}^{M} \int_{-\pi}^{\pi} \Big(\hat{q}(u_s, \lambda) \Big)^2 d\lambda. \tag{3.1.5}$$

It can be shown that under some rather standard assumptions to be discussed later and if $M \to \infty$ as $T \to \infty$, then, in probability,

$$Q_{0,T} \;\; \to \begin{cases} 0 & \text{if } H_0 \text{ is true} \\ \int_0^1 \int_{-\pi}^{\pi} \left(\frac{f(u,\lambda)}{f(u,\lambda,\vartheta_0)} - 1 \right)^2 d\lambda du & \text{if } H_1 \text{ is true.} \end{cases}$$

This behavior of $Q_{0,T}$ justifies its use for testing the null hypothesis of interest.

Recall that in order to derive the test statistic (3.1.5) we have assumed that the parameterizing functions $\vartheta(u)$ are known. This corresponds to the case of testing a simple hypothesis, that is a hypothesis where the local spectral density under the null is fully specified. To extend the testing procedure proposed to the more interesting case of testing a composite hypotheses, that is to the case where the functions $\vartheta(u)$ determining the local spectral density are unknown, we replace $\vartheta(\cdot)$ in (3.1.4) by $\sqrt{N}$-consistent estimators. Let $\hat{\vartheta}(\cdot) = (\hat{\vartheta}_1(\cdot), \ldots, \hat{\vartheta}_m(\cdot))'$ be such an estimator of $\vartheta(\cdot) = (\vartheta_1(\cdot), \ldots, \vartheta_m(\cdot))'$. Analogously to (3.1.5), the test statistic used in this case is then given by

$$Q_T = \frac{1}{M} \sum_{s=1}^{M} \int_{-\pi}^{\pi} \left\{ \frac{1}{N} \sum_{j=-M_N}^{M_N} K_b(\lambda - \lambda_j) \left(\frac{I_N(u_s, \lambda_j)}{f(u_s, \lambda_j; \hat{\vartheta}(u_s))} - 1 \right) \right\}^2 d\lambda. \tag{3.1.6}$$

Notice that $f(u_s, \lambda_j; \hat{\vartheta}(u_s))$ appearing in the denominator above, is the semiparametric local spectral density obtained by substituting $\vartheta(\cdot)$ appearing in $f(u, \lambda_j; \vartheta(u))$ by its estimator $\hat{\vartheta}(\cdot)$. The test statistic (3.1.6) is the one proposed in this chapter for testing the pair of hypotheses (3.1.2).

3.1.3 Asymptotic distribution under the null hypothesis

We first establish a basic theorem which deals with the asymptotic distribution of the test statistic (3.1.6) under the null hypothesis in (3.1.2). For this the following set of assumptions is imposed.

Assumption 3.1.2.

(i) *K is a bounded, symmetric, nonnegative kernel function on $(-\infty, \infty)$ with support $[-\pi, \pi]$ such that $(2\pi)^{-1} \int_{-\infty}^{\infty} K(x)dx = 1$.*

(ii) *The window length N satisfies $N \sim T^{\delta}$ for some $1/5 < \delta < 4/5$. Furthermore, $N = [\kappa S]$ where κ is a positive constant independent of N and S.*

(iii) *The smoothing bandwidth b satisfies $b \sim N^{-\lambda}$, where*

$$\max\{0, \frac{9\delta - 7}{\delta}\} < \lambda < \min\{\frac{5\delta - 1}{3\delta}, \frac{1}{2}, \frac{1-\delta}{\delta}\}.$$

(iv) *The taper function h is of bounded variation and vanishes outside the interval [0,1].*

(v) *$\sqrt{N}(\hat{\theta}(u) - \theta(u)) = O_p(1)$ where the $O_p(\cdot)$ term does not depend on u.*

Some remarks concerning the above assumptions are in order. Note that the constant κ appearing in (ii) determines the degree of overlapping between the segments used. We consider the case $\kappa \geq 1$ only, since for $\kappa < 1$ the shift from segment to segment described by S is greater than the segment length N. In the later case, a loss of efficiency is expected due to the fact that some observations are omitted. If $\kappa = 1$ then the observed series is partitioned in nonoverlapping segments of length N while if $\kappa > 1$ then the segments considered overlap. Concerning the rate at which the segment length N is allowed to increase to infinity given in (ii) and the rate at which the bandwidth b is allowed to converge to zero given in (iii), we mention that they are controlled in a way that leads to simple expressions for the mean and for the variance of the limiting distribution of Q_T under H_0. Notice that the range of values of N and of b is large enough allowing for a flexibility in choosing these parameters in practice. Assumption 3.1.2(v) is general enough and allows for different estimators of $\vartheta(u)$; see among others, Dahlhaus and Giraitis [13], Dahlhaus [10], Dahlhaus and Neumann [15] and van Bellegem and Dahlhaus [42] for different proposals.

The following theorem establishes the asymptotic distribution of Q_T when the null hypothesis is true.

Theorem 3.1.1. *Under Assumption 3.1.1 and 3.1.2 and if H_0 is true, then, as $T \to \infty$,*

$$N\sqrt{Mb}(Q_T - \mu_T) \Rightarrow N(0, \tau^2),$$

where

$$\mu_T = \frac{(tap(1))^{1/2}}{Nb}\int_{-\pi}^{\pi} K^2(x)dx + \frac{(tap(1))^{1/2}}{4\pi N}\int_{-\pi}^{\pi}\int_{-2\pi}^{2\pi} K(x)K(x-u)dxdu,$$

$$\tau^2 = tap(\kappa)\frac{2}{\pi}\int_{-2\pi}^{2\pi}\left(\int K(u)K(u+x)du\right)^2 dx$$

and for $s \in \{1, 2, \ldots, m\}$

$$tap(s) = \frac{\sum_{|m|<s}\left(\int_0^{1-|m|/s} h^2(u)h^2(u+|m|/s)du\right)^2}{\left(\int_0^1 h^2(x)dx\right)^4}.$$

According to the above theorem, an attractive feature of the test statistic Q_T, is that its limiting distribution under the null hypothesis does not depend on unknown parameters or characteristics of the underlying locally stationary process $\{X_{t,T}\}$. Furthermore, and based on this theorem, an asymptotically α-level test is obtained by rejecting the null hypothesis if

$$Q_T \geq \mu_T + \frac{\tau}{N\sqrt{Mb}} z_\alpha,$$

where z_α denotes the $100(1-\alpha)\%$ percentile of the standard Gaussian distribution.

3.2 Testing for a time-varying autoregressive structure

3.2.1 Consistency

A special case of the testing problem (3.1.2) and which commonly arises in many situations, is that of testing for the presence of a time-varying autoregressive (tvAR) model. Recall that a locally stationary process satisfying definition 1.1.1 obeys a time-varying autoregressive representation of order p if $X_{t,T}$ is generated by the equation

$$X_{t,T} = \sum_{j=1}^{p} \beta_j(t/T) X_{t-j,T} + \varepsilon_{t,T}, \tag{3.2.7}$$

where the $\varepsilon_{t,T}$'s are i.i.d. $N(0, \sigma^2(t/T))$ random variables, $\beta_p(u) \neq 0$ for all $u \in [0,1]$, the functions $\beta_j(\cdot)$ as well as the variance function $\sigma^2(\cdot)$ are of bounded variation and $\sum_{j=1}^{p} \beta_j(u) z^j \neq 0$ for all $u \in [0,1]$ and all $0 < |z| \leq 1 + \delta$, for some $\delta > 0$. Although the results of this chapter can be easily adapted to cover other special types of semiparametric locally stationary processes i.e. tvARMA(p,q) or tvMA(q), we concentrate on the class of time-varying autoregressive process because these processes provide due to their simplicity, easy implementation and interpretation, a very interesting subclass of semiparametric time varying processes. Now let $\mathcal{F}_{tvAR(p)}$ be the set of local spectral densities of time-varying autoregressive processes of order p. The testing problem considered in this section is then described by the following pair of null and alternative hypothesis

$$H_0: \; f(\cdot,\cdot) \in \mathcal{F}_{tvAR(p)} \quad \text{vs} \quad H_1: \; f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{tvAR(p)}. \tag{3.2.8}$$

Note that the set $\mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{tvAR(p)}$ contains also all locally stationary autoregressive processes with an autoregressive order different from p.

We first discuss a consistency property of our test. For this, suppose that the true spectral density $f(u,\lambda)$ lies in the alternative and measure for $u \in [0,1]$ the distance between $f(u,\lambda)$ and $f(u,\lambda;\vartheta(u))$ by the function

$$\mathcal{L}(u,\vartheta(u)) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log[f(u,\lambda;\vartheta(u))] + \frac{f(u,\lambda)}{f(u,\lambda;\vartheta(u))} \right) d\lambda. \tag{3.2.9}$$

Let $\overline{\vartheta}(u)$ be the value of $\vartheta(u)$ which minimizes $\mathcal{L}(u,\vartheta(u))$ and let $\hat{\vartheta}(u)$ be the estimator of $\overline{\vartheta}(u)$ which is obtained by minimizing the local Whittle likelihood, i.e., $\overline{\vartheta}(u) = \arg\min \mathcal{L}_N(u,\vartheta(u))$, where

$$\mathcal{L}_N\{u,\vartheta(u)\} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log[f(u,\lambda;\vartheta(u))] + \frac{I_N(u,\lambda)}{f(u,\lambda;\vartheta(u))} \right) d\lambda.$$

Notice that

$$\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left(\frac{\log[f(u,\lambda;\vartheta(u))]}{f(u,\lambda)} + \frac{f(u,\lambda)}{f(u,\lambda;\vartheta(u))} - 1 \right) d\lambda du$$

is the asymptotic Kullback-Leibler information divergence between two Gaussian locally stationary processes with time-varying spectral densities $f(u,\lambda;\vartheta(u))$ and $f(u,\lambda)$ respectively; see Theorem 3.4 of Dahlhaus [7]. The curve $\overline{\vartheta}(u)$ obtained by minimizing (3.2.9) is that leading to the best time-varying autoregressive fit, that is to the p-th order autoregressive fit which minimizes the Kullback-Leibler information divergence (3.2.9).

Assumption 3.2.1. *Let $\nabla = (\partial/\partial\vartheta_1, \dots, \partial/\partial\vartheta_m)'$ be the gradient with respect to ϑ.*

(i) $\nabla \mathcal{L}_N(u, \hat{\vartheta}(u)) = 0$, $\nabla \mathcal{L}_N(u, \overline{\vartheta}(u)) = 0$ *for all u and N.*

(ii) The derivatives $\partial^2 A(u,\lambda)/\partial u \partial\lambda$ and $\partial^3 A(u,\lambda)/\partial u^3$ are uniformly bounded in $(u,\lambda) \in [0,1] \times [-\pi,\pi]$.

(iii) The derivatives

$$\frac{\partial^3}{\partial\vartheta_{i_1}\partial\vartheta_{i_2}\partial\vartheta_{i_3}} f^{-1}(u,\lambda;\vartheta(u)), \quad \frac{\partial^3}{\partial\vartheta_{i_1}\partial\vartheta_{i_2}\partial\vartheta_{i_3}} f(u,\lambda;\vartheta(u)), \quad \frac{\partial^2}{\partial\lambda^2}\frac{\partial}{\partial\vartheta_{i_1}} f^{-1}(u,\lambda;\vartheta(u))$$

are bounded for $1 \leq i_1, i_2, i_3 \leq p$ uniformly in $(u,\lambda,\vartheta) \in [0,1] \times [-\pi,\pi] \times \Theta$, where Θ is an open convex subset of $\mathbb{R}^p$.

(iv) $\sup_{0 \leq u \leq 1, \vartheta \in \Theta} ||\nabla^2 \mathcal{L}^{-1}(u, \hat{\vartheta}(u))||_{sp}$ *where $||\cdot||_{sp}$ denotes the spectral norm of a matrix.*

We first state the following result which deals with the limiting properties of Q_T when the alternative hypothesis is true.

Theorem 3.2.1. *Under Assumptions 3.1.1, 3.1.2 and 3.2.1 and if $f(\cdot,\cdot) \in \mathcal{F}_{\mathcal{LS}} \setminus \mathcal{F}_{tvAR(p)}$, then as $T \to \infty$,*

$$Q_T \to D^2 = \int_0^1 \int_{-\pi}^{\pi} \left(\frac{f(u,\lambda)}{f(u,\lambda;\overline{\vartheta}(u))} - 1 \right)^2 d\lambda du,$$

in probability.

Notice that the limit D^2 given above is a L_2-distance measure between the true local spectral density $f(u,\lambda)$ and its best parametric fit $f(u,\lambda;\overline{\vartheta}(u))$. Theorem 3.2.1 implies then that under the assumptions made and if H_1 is true, then $\lim_{T\to\infty} P(N\sqrt{Mb}(Q_T - \mu_T)/\tau \geq z_\alpha) = 1$, that is the test Q_T is consistent against any alternative for which $D^2 > 0$.

3.2.2 Bootstrapping the test statistic

To obtain critical values of the test, Theorem 3.1.1 enables us to approximate the unknown distribution of $N\sqrt{Mb}(Q_T - \mu_T)/\tau$ by that of a standard Gaussian distribution. We experienced, however, that the quality of this approximation is rather poor in finite sample situations and very large to huge sample sizes are required in order for this approximation to be valuable in practice; see Section 3.4 for a numerical illustration of this point. To improve upon the large sample Gaussian approximation of Theorem 3.1.1, we propose here, an alternative, bootstrap-based procedure, which leads in finite sample situations to more accurate estimates of the distribution of Q_T under the null. The procedure proposed works by generating pseudo-observations $X^+_{1,T}, X^+_{2,T}, \ldots, X^+_{T,T}$ using the fitted tvAR(p) process and calculating the test statistic Q_T of interest using the so generated pseudo-observations.

To elaborate on, we first fit locally to the time series the p*th* order time-varying autoregressive process postulated under the null hypothesis. This can be done using local Yule-Walker or local least squares estimators of the autoregressive parameter functions $\beta_1(u), \ldots, \beta_p(u)$. Yule-Walker estimators $\hat{\beta}_u(p)' = (\hat{\beta}_1(u), \ldots, \hat{\beta}_p(u))$, for instance satisfy the system of equations

$$\hat{R}_u(p)\hat{\beta}_u(p) = \hat{r}_u(p),$$

with

$$\hat{R}_u(p) = \hat{c}_N(u, i-j)_{i,j=1,\ldots,p}, \qquad \hat{r}_u(p) = (\hat{c}_N(u,1), \ldots, \hat{c}_N(u,p))'$$

and

$$\hat{c}_N(u,j) = \frac{1}{N}\sum_{\substack{k,l=0\\k-l=\tau}}^{N-1} X_{[uT]-N/2+k+1,T}X_{[uT]-N/2+l+1,T}.$$

Let

$$\hat{\sigma}^2_N(u) = \hat{c}_N(u,0) + \hat{\beta}'_u(p)\hat{r}_u(p)$$

be the corresponding estimator of the variance function $\sigma^2(u)$ of the errors. Properties of the estimators $\hat{\beta}_u(p)$ and $\hat{\sigma}^2_p(u)$ have been investigated by Dahlhaus and Giraitis [13]; see also Section 2.3. Notice that other estimators can be also used provided Assumption 2.1(v) is satisfied; cf. Dahlhaus et al. [14] and van Bellegem and Dahlhaus [42].

The bootstrap algorithm proposed to approximate the distribution of Q_T under the null hypothesis of a tvAR(p) process consists then of the following four Steps:

STEP 1: Fit locally the time-varying autoregressive model of order p to the observations $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$ and calculate the estimated parameters $\hat{\beta}_{t/T}(p)' = (\hat{\beta}_1(t/T), \ldots, \hat{\beta}_p(t/T))$ and $\hat{\sigma_p}^2(t/T)$.

STEP 2: Generate bootstrap observations $X^+_{1,T}, X^+_{2,T}, \ldots, X^+_{T,T}$ using the fitted local autoregressive model, that is,

$$X^+_{t,T} = \sum_{j=1}^{p} \hat{\beta}_j(\frac{t}{T}) X^+_{t-j,T} + \hat{\sigma}_p(\frac{t}{T}) \cdot \varepsilon^+_t,$$

where $X^+_{j,T} = X_{j,T}$ for $j = 1, 2, \ldots, p$ and ε^+_t are i.i.d random variables with $\varepsilon^+_t \backsim N(0,1)$.

STEP 3: Compute the local periodogram $I^+_N(u, \lambda)$ over segments of length N of the bootstrap pseudo-observations $X^+_{t,T}$, i.e., compute

$$I^+_N(u, \lambda) = \frac{1}{2\pi H_{2,N}(0)} |d^+_N(u, \lambda)|^2 \tag{3.2.10}$$

where

$$d^+_N(u, \lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) X^+_{[uT]-N/2+s+1} e^{-i\lambda s}.$$

STEP 4: The bootstrapped test statistic is then defined by

$$Q^+_T = \frac{1}{M} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \left\{ \frac{1}{N} \sum_{j=-M_N}^{M_N} K_b(\lambda - \lambda_j) \left(\frac{I^+_N(u_i, \lambda_j)}{f(u_i, \lambda_j; \hat{\vartheta})} - 1 \right) \right\}^2 d\lambda$$

Notice that we could have in STEP 4 rescaled the local bootstrap periodogram $I^+_N(u, \lambda)$ by $f(u_i, \lambda_j; \hat{\vartheta}^+)$ instead by $f(u_i, \lambda_j; \hat{\vartheta})$, where $\hat{\vartheta}(\cdot)^+$ denotes the estimator of the autoregressive parameter functions $\vartheta(\cdot)$ obtained using the bootstrap pseudo-series $X^+_{1,T}$, $X^+_{2,T}$, $\ldots, X^+_{T,T}$. The specification of Q^+_T used is, however, preferred because besides of being computationally more convenient, it is also justified theoretically by the fact that the limiting distribution of the test statistic Q_T under the null is not affected if the unknown $\vartheta(\cdot)$ is replaced by a $\sqrt{N}$-consistent estimator $\hat{\vartheta}(\cdot)$.

The following theorem shows that the bootstrap procedure proposed leads to an asymptotically valid approximation of the distribution of the test statistic Q_T under the null hypothesis of a tvAR(p) process.

Theorem 3.2.2. *Let Assumptions 3.1.1, 3.1.2 and 3.2.1 be satisfied. Then, conditionally on $X_{1,T}, X_{2,T}, \ldots, X_{T,T}$, we have as $T \to \infty$,*

$$N\sqrt{Mb}(Q_T^+ - \mu_T) \Rightarrow N(0, \tau^2),$$

in probability, where μ_T and τ^2 are defined in Theorem 3.1.1.

3.3 Applications

3.3.1 Some remarks on choosing the testing parameters

From the previous discussion it is clear that implementation of the testing procedure proposed, requires essentially the selection of two parameters: the time window width N and the smoothing bandwidth b. Although a thorough investigation of this problem is beyond the scope of this chapter, we in what follows we give a rather heuristic discussion on how to select this parameters in practice.

Concerning the value of the time window width N, we mention that the selection of this parameter is inherit to any statistical inference procedure for locally stationary process which is based on segments of observations. Choosing N to large will induce a large bias since a large N is associated with a loss of information on the local structure of the underlying process. On the other hand, choosing N to small will lead to an increase of the variance of the estimators involved due to the small number of observations used. Any approach to select N should therefore be guided by the requirement that N should be large enough to allow for reasonable local estimation but not too large to avoid a 'smoothing out' the interesting local characteristics of the process. Based on this observation and depending on the overall size n of the time series at hand, we propose for numerical reasons, to choose N to be some power of 2, where the choices $N = 64$ or $N = 128$ are more convenient in most situations.

Concerning the choice of the smoothing parameter b, one way to proceed is to select this parameter using a local version of a cross-validation criterion like the one proposed by Beltrão and Bloomfield [2]. To elaborate on, notice first that our aim is to obatin a "good" estimate of the function $q(u, \lambda) = f(u, \lambda)/f(u, \lambda; \vartheta)$. To stress the dependence of this function on the estimated parametric curves, we write $q(u, \lambda, \vartheta(u))$ in the sequel. Using as a starting point the function

$$\sum_{i=1}^{M}\sum_{j=1}^{N}\left\{\log q(u_i, \lambda_j; \hat{\vartheta}) + \frac{I_N(u_i, \lambda_j)/f(u_i, \lambda_j, \hat{\vartheta})}{q(u_i, \lambda_j; \hat{\vartheta})}\right\}, \tag{3.3.11}$$

a leave-one-out estimator of $q(u, \lambda_j; \vartheta)$ is given by

$$\hat{q}_{-j}(u, \lambda_j; \hat{\vartheta}) = \frac{1}{N}\sum_{j \in N_j} K_h(\lambda_j - \lambda_s)\frac{I_N(u, \lambda_j)}{f(u, \lambda_j; \hat{\vartheta})} \tag{3.3.12}$$

where $N_j = \{s : -M_N \leq s \leq M_N \text{ and } j - s \neq \pm j \mod M_N\}$. Notice that $\hat{q}_{-j}$ is a kernel estimator of q obtained by ignoring the jth ordinate local periodogram $I_N(u, \lambda_j)$. Now, substituting $\hat{q}_{-j}(u, \lambda_j; \hat{\vartheta})$ for $q(u, \lambda_j; \hat{\vartheta})$ in (3.3.11) leads to the function

$$CV(b) = \sum_{i=1}^{M}\sum_{j=1}^{N}\left\{\log \hat{q}_{-j}(u_i, \lambda_j; \hat{\vartheta}) + \frac{I_N(u_i, \lambda_j)/f(u_i, \lambda_j; \hat{\vartheta})}{\hat{q}_{-j}(u_i, \lambda_j; \hat{\vartheta})}\right\}, \tag{3.3.13}$$

which can be used as a cross-validation-type criterion to select b.

3.3.2 Simulations

Bootstrap Approximations

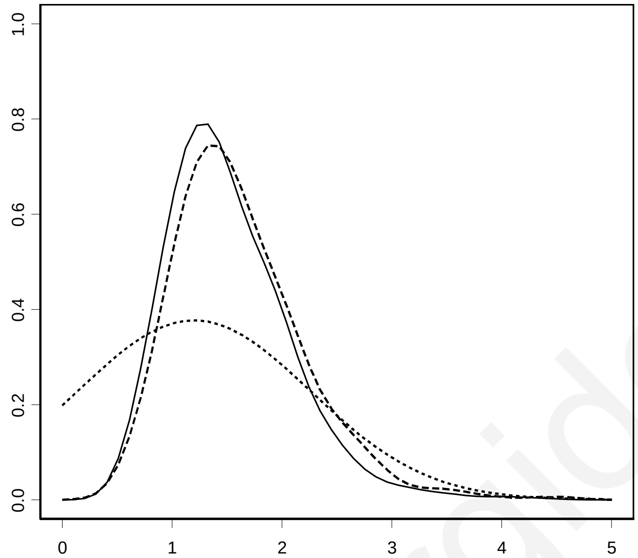
We first, illustrate the advantages of using the bootstrap procedure proposed by comparing its performance in approximating the distribution of Q_T under the null with that of the limiting Gaussian approximation. For this purpose, observations $\{X_{t,T}, t = 1, \ldots, T\}$ from the first order, time-varying autoregressive model

$$X_{t,T} = \phi(\frac{t}{T})X_{t-1,T} + \varepsilon_t \tag{3.3.14}$$

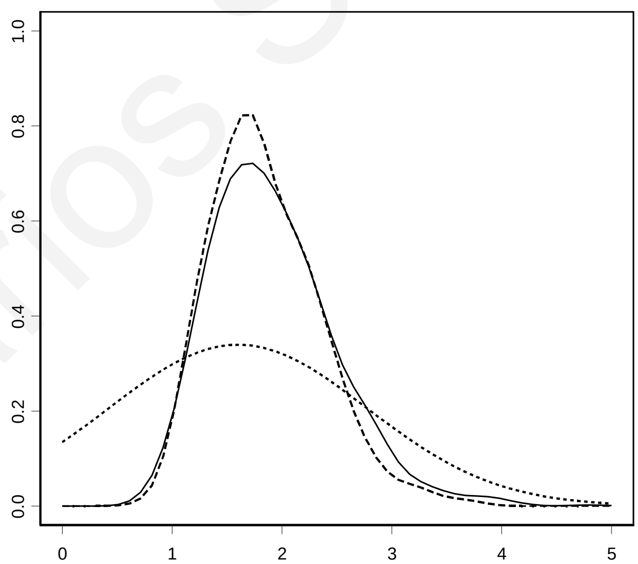
have been generated, where $\phi(t/T) = 0.9\cos(1.5 - \cos(4\pi(t/T)))$ and the ε_t's are i.i.d. random variables with $\varepsilon_t \sim N(0, 1)$. To estimate the exact distribution of the test statistic Q_T we generate 1000 series of length $T = 1024$ and for each of these series we calculated Q_T using the Bartlett-Priestley kernel, $K(x) = 1_{[-\pi,\pi]}(x)3(4\pi)^{-1}(1 - (x/\pi)^2)$ and the bandwidth $b = 0.2$. The window width N has been set equal to $N = 128$ and two different shifts, $S = 128$ and $S = 64$, have been considered. Notice that for $S = 128$ we have $\kappa = 1$, while for $S = 64$, $\kappa = 2$.

To investigate the performance of the bootstrap method, we choose randomly 21 series from the generated 1000 replications of process (3.3.14) and for each of the selected series we apply the bootstrap procedure proposed using 300 bootstrap replications. Based on the bootstrap replications, we estimated for each series the density $\hat{g}^*$ of the corresponding bootstrap approximation of the distribution of Q_T. We also estimated the density of the exact distribution of Q_T based on the 1000 replications of process (3.3.14). The so estimated density is denoted by $\hat{g}$. The density estimates $\hat{g}^*$ and $\hat{g}$ have been obtained using standard SPlus smoothing routines. We then compare the estimated exact density $\hat{g}$ with the Gaussian approximation given in Theorem 3.1.1 and with the median bootstrap approximation. The median bootstrap approximations is that for which $\sum_{x_i} |\hat{g}^*(x_i) - \hat{g}(x_i)|$ takes its median value over the 21 series used. Figure 3.1 shows the estimated densities of the exact, the asymptotic Gaussian and the median bootstrap approximation.

As it is clearly seen from these exhibits, the estimation results based on bootstrap are striking. In particular, the bootstrap performs much better compared to the Gaussian approximation and estimates very accurately the exact distribution of interest.



(a) $\kappa = 1$



(b) $\kappa = 2$

Figure 3.1: Estimated density of the distribution of the test statistic Q_T under the null hypothesis of a first order tvAR process and its different approximations. The solid lines in (a) and (b) are the estimated exact densities, the dashed lines are the estimated densities corresponding to the median bootstrap approximations while the dotted lines are the densities of the asymptotic Gaussian approximations.

Size and power performance of the test

We next investigate the size and power performance of the test in finite sample situations by means of a small simulation study. For this, we consider realizations of length $T = 512$ and $T = 1024$ of the time-varying AR(2) model

$$X_{t,T} = 0.9\cos(1.5 - \cos(4\pi t/T))X_{t-1,T} - \phi_2 X_{t-2,T} + \varepsilon_t \qquad (3.3.15)$$

where the ε_t's are independent, standard Gaussian distributed random variables. The null hypothesis is that the underlying process is a time-varying first order autoregressive process. Different values of the parameter ϕ_2 have been considered corresponding to validity of the null ($\phi_2 = 0$) and of the alternative hypothesis ($\phi_2 \neq 0$). In each case we fit a time-varying AR(1) model using a local least squares estimator and compute the test statistic Q_T using the Bartlett-Priestley kernel and different values of the bandwidth parameter b. We also apply the test proposed for different segment lengths N and shifts S. In all cases the critical values of the test have been obtained using B=300 replications of the bootstrap procedure described in Section 3.2. The results obtained over 500 replications are summarized in Table 3.1.

T=512, N=64	$b = 0.3$			$b = 0.2$		
	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$
$\phi_2 = 0.0$						
$\kappa = 1$	0.008	0.038	0.084	0.008	0.042	0.078
$\kappa = 2$	0.008	0.032	0.070	0.008	0.032	0.068
$\phi_2 = 0.2$						
$\kappa = 1$	0.072	0.204	0.358	0.098	0.244	0.364
$\kappa = 2$	0.200	0.296	0.408	0.174	0.354	0.474
$\phi_2 = 0.25$						
$\kappa = 1$	0.188	0.410	0.582	0.238	0.482	0.612
$\kappa = 2$	0.408	0.564	0.680	0.422	0.640	0.744
$\phi_2 = 0.3$						
$\kappa = 1$	0.392	0.668	0.814	0.474	0.752	0.852
$\kappa = 2$	0.692	0.820	0.898	0.738	0.882	0.936

T=1024, N=128	$b = 0.2$			$b = 0.1$		
	$a = 0.01$	$a = 0.05$	$a = 0.1$	$a = 0.01$	$a = 0.05$	$a = 0.1$
$\phi_2 = 0.0$						
$\kappa = 1$	0.010	0.040	0.080	0.012	0.044	0.082
$\kappa = 2$	0.008	0.044	0.092	0.012	0.048	0.098
$\phi_2 = 0.2$						
$\kappa = 1$	0.272	0.512	0.618	0.226	0.474	0.584
$\kappa = 2$	0.480	0.732	0.800	0.500	0.674	0.776
$\phi_2 = 0.25$						
$\kappa = 1$	0.622	0.830	0.894	0.542	0.798	0.866
$\kappa = 2$	0.824	0.952	0.974	0.834	0.924	0.968
$\phi_2 = 0.3$						
$\kappa = 1$	0.914	0.992	0.996	0.884	0.974	0.992
$\kappa = 2$	0.986	1.000	1.000	0.986	1.000	1.000

Table 3.1: Rejection frequencies in 500 replications of the tvAR(2) model $X_{t,T} = 0.9\cos(1.5 - \cos(4\pi t/T))X_{t-1,T} - \phi_2 X_{t-2,T} + \varepsilon_t$ for different values of ϕ_2 and of the testing parameters.

As Table 3.1 shows, estimating the critical values of the test using the bootstrap procedure proposed, leads to a very good size and power behavior of the test. Notice that for both sample sizes and all combinations of bandwidth values, segment lengths and shifts considered, the empirical size of the test is very close to the nominal level of 5%. Furthermore, under the alternative, the test has power even for small deviations from the null and the power of the test increases rapidly approaching unity as the deviations from the null and/or the sample size become larger.

Chapter 4

Conclusions and further research

4.1 Conclusions

We have proposed a bootstrap method to produce replicates of the local periodogram and applied this method to the important classes of local spectral means and local ratio statistics. We have derived the asymptotic distributions of the bootstrap analogues of these statistics and some simulations have demonstrated the performance of our bootstrap procedure against the local autoregressive sieve bootstrap and the normal approximation. An application to a real-data set is given.

We have also addressed the important problem of testing whether a locally stationary process belongs to a semiparametric class of time varying processes. The asymptotic distribution of the test statistic proposed is derived. As an interesting special case we focus on the problem of testing for the presence of a semiparametric, time-varying autoregressive model and a bootstrap procedure is proposed to approximate the distribution of the test statistic under the null hypothesis. Theoretical properties of the bootstrap procedure are discussed and its asymptotic validity is established. Simulations demonstrated that, in the testing set-up considered, the bootstrap is a very powerful and valuable tool to obtain critical values.

4.2 Further research

The bootstrap approach in the second chapter can be applied to the calculation of pointwise confidence intervals for some parameters of interest. We demonstrated such an application of the bootstrap by considering pointwise confidence intervals for the time varying autocorrelation function $\rho(u,\tau)$, $u \in [0,1]$ and $\tau \in \mathbb{N}$. An interesting problem for future research would be the construction of simultaneous confidence bands for such parameters. Applied to the time varying autocorrelation function, this problem requires, for instance, the investigation of the distribution of statistics like

$$\sup_{u\in[0,1]} |\hat{\rho}(u,\tau) - \rho(u,\tau)|$$

and the development of a bootstrap procedure which is capable to approximate its distribution.

Concerning the testing methodology proposed in Chapter 3, it will be interesting to investigate more closely the power behavior of the test for classes of fixed and of local alternatives. Furthermore, it is interesting to investigate how the testing methodology proposed can be applied to the problem of testing stationarity in time series analysis.

Another interesting problem for future research is how to bootstrap the preperiodogram defined for every $\lambda \in [-\pi, \pi]$ and $u \in [0, 1]$ by

$$J_N(u,\lambda) = \frac{1}{2\pi} \sum_k X_{uT+(k+1)/2,T} X_{uT-(k-1)/2,T} e^{-ik\lambda}, \tag{4.2.1}$$

where the sum over k is for $k \in \mathbb{Z}$ such that $1 \leq [uT-(k-1)/2], [uT+(k+1)/2] \leq T$.

The preperiodogram is an important tool in the analysis of locally stationary processes and several statistics proposed in the literature are based on it. The method of bootstrapping the local periodogram, proposed in the second chapter, can not be directly applied to bootstrap the preperiodogram.

Finally, our bootstrap procedures depend on the choice of different smoothing parameters like the window length N, the smoothing bandwidth b, S and M. In this thesis, we gave only some guidelines on how to choose these parameters in practice. It will be interesting to develop a theory on how to choose these parameters based on some optimality criteria.

Appendix A

Auxiliary results and proofs for Chapter 2

First we define the process $Z_{t,T}(u) = X_t - \widetilde{X}_t(u)$ which satisfies

$$\sum_{s=-\infty}^{\infty} cov(Z_{t,T}(u), Z_{t+s,T}(u)) = O(\frac{1}{T} + |\frac{t}{N} - u|) \tag{A.0.1}$$

Lemma A.0.1. *If $\{X_{t,T}\}$ are locally stationary processes satisfying Assumption 2.1 and*

$$\varepsilon_{p,j}(u) = X_{[uT]-N/2+j,T} - \sum_{i=1}^{p} \beta_i(u) X_{[uT]-N/2+j-i,T},$$

$$\varepsilon_j(u) = \widetilde{X}_{[uT]-N/2+j}(u) - \sum_{i=1}^{\infty} \beta_i(u) \widetilde{X}_{[uT]-N/2+j-i}(u)$$

for $p \in \mathbb{N}$,

$$E\left(\frac{1}{N-p}\sum_{j=p+1}^{N} X_j(u,p)(\varepsilon_{p,j}(u) - \varepsilon_j(u))\right)^2 \leq K\left(\left(\frac{N}{T}\right)^2 \sup_{0\leq u\leq 1}\sum_{i=0}^{p} {\beta_i}^2(u) + \sup_{0\leq u\leq 1}\sum_{k=p+1}^{\infty} {\beta_k}^2(u)\right)$$

where $X_j(u,p)$ is defined in Section 2.3.

Proof : Since

$$\varepsilon_{p,j}(u) - \varepsilon_j(u) = \sum_{i=0}^{p} \beta_i(u) Z_{[uT]-N/2+j-i}(u) + \sum_{i=p+1}^{\infty} \beta_i(u) \widetilde{X}_{[uT]-N/2+j-i}(u) \tag{A.0.2}$$

the result now follows by (A.0.1).

■

Lemma A.0.2. *Let $\{X_{t,T}\}$ be a locally stationary process satisfying Assumption 2.1, and $p \in \mathbb{N}$. Then*

$$E((N-p)^{-1}\sum_{k=p}^{N} X_{[uT]-N/2+k+1-i,T}X_{[uT]-N/2+k+1-j,T}-c(u,i-j))^2 \leq K\left(\frac{N^2}{T^2}+(N-p)^{-1}\right)$$

for every $(i,j) \in \mathbb{N}_p \times \mathbb{N}_p$ where $\mathbb{N}_p = \{1,2,\ldots,p\}$ and $c(u,\tau)$ is the time varying covariance defined in Section 1.1.

Proof : Let $[uT] - N/2 + 1 = \nu$. We then have

$$\begin{aligned}
\frac{1}{N-p}\sum_{k=p}^{N} X_{\nu+k-i,T}X_{\nu+k-j,T} - c(u,i-j) &= \frac{1}{N-p}\sum_{k=p}^{N} \widetilde{X}_{\nu+k-i}(u)\widetilde{X}_{\nu+k-j}(u) - c(u,i-j) \\
&+\frac{1}{N-p}\sum_{k=p}^{N} Z_{\nu+k-i}(u)Z_{\nu+k-j}(u) \\
&+\frac{1}{N-p}\sum_{k=p}^{N} Z_{\nu+k-i}(u)\widetilde{X}_{\nu+k-j}(u) \\
&+\frac{1}{N-p}\sum_{k=p}^{N} Z_{\nu+k-j}(u)\widetilde{X}_{\nu+k-i}(u) \\
&= T_{1,N} + T_{2,N} + T_{3,N} + T_{4,N}
\end{aligned}$$

with an obvious notation for $T_{i,N}$. Since $\widetilde{X}_t(u)$ is a stationary process we have that $E(T_{1,N})^2 = O((N-p)^{-1})$ uniformly in u and the result follows because (A.0.1) gives that

$E(T_{2,N})^2 = O(N^4/T^4)$ and $E(T_{3,N})^2$ and $E(T_{4,N})^2$ are $O(N^2/T^2)$. ■

Before establishing the next lemma we recall some properties of covariance matrices. From Grenander and Szegö [21] we have that if $0 < F_1 < f(\lambda) < F_2 < \infty$ and $\lambda_1 < \lambda_2 < \cdots < \lambda_p$ are the eigenvalues of $R(p) = [\gamma(i-j)]_{i,j=1,\ldots,p}$ with $\gamma(\tau) = \int_{-\pi}^{\pi} f(\lambda)\exp(i\lambda\tau)d\lambda$, then $2\pi F_1 \leq \lambda_1 < \cdots < \lambda_p \leq 2\pi F_2$. Assumption 2.1 implies that for each u the local spectral density $f(u,\lambda)$ is continuous in λ, and there are constants F_1 and F_2 such that $0 < F_1 < f(u,\lambda) < F_2 < \infty$ for all $u \in [0,1]$. Consequently, if we assume that $\lambda_1(u) < \lambda_2(u) < \cdots < \lambda_p(u)$ are the eigenvalues of $R_u(p) = [c(u,i-j)]_{i,j=1,\ldots,p}$, then $2\pi F_1 \leq \lambda_1(u) < \cdots < \lambda_p(u) \leq 2\pi F_2$.

For a matrix A, let $\| A \| = \sup_{\|x\|_2 \leq 1} \| Ax \|_2$, where $\| \, . \, \|_2$ is the Euclidean norm. Thus $\| A \| \leq \| A \|_2$ and if A is positive defined and symmetric $\| A \| = |\lambda_{max}|$ where λ_{max} is the largest of the eigenvalues of the matrix A.

We then have that, uniformly in u,

$$\| R_u(p) \| \leq 2\pi F_2, \qquad \| R_u^{-1}(p) \| \leq 1/(2\pi F_1). \tag{A.0.3}$$

Lemma A.0.3. *Let $\{\tilde{X}_t(u)\}$ satisfy Assumptions 2.2, 2.3 and $p \to \infty$ such that $p^3/N \to 0$. Then for every $u \in [0,1]$*

$$\sqrt{p} \parallel \hat{R}_u^{-1}(p) - R_u^{-1}(p) \parallel \to 0$$

in probability.

Proof : Let

$$\| R_u^{-1}(p) \| = k(u,p), \quad \| \hat{R}_u^{-1}(p) - R_u^{-1}(p) \| = q_T(u,p), \quad \| \hat{R}_u(p) - R_u(p) \| = Q_T(u,p)$$

and note that as in Berk (1974),

$$\begin{aligned} q_T(u,p) &= \| \hat{R}_u^{-1}(p)(\hat{R}_u(p) - R_u(p))R_u^{-1}(p) \| \\ &\leq (\| \hat{R}_u^{-1}(p) - R_u^{-1}(p) \| + \| R_u^{-1}(p) \|) \| \hat{R}_u(p) - R_u(p) \| \| R_u^{-1}(p) \| \\ &= (k(u,p) + q_T(u,p))Q_T(u,p)k(u,p). \end{aligned}$$

Choose T large enough such that $k(u,p)Q_T(u,p) < 1$ holds in probability. We have,

$$q_T(u,p) \leq \frac{k^2(u,p)Q_T(u,p)}{1 - k(u,p)Q_T(u,p)}. \tag{A.0.4}$$

By Lemma A.0.2 we get

$$E(Q_T^2(u,p)) \leq Kp^2\Bigg(N^2T^{-2} + (N-p)^{-1}\Bigg). \tag{A.0.5}$$

which implies that

$$E(\sqrt{p}\ Q_T(u,p))^2 \leq Kp^3\Bigg(N^2T^{-2} + (N-p)^{-1}\Bigg) \to 0$$

as $T \to \infty$, since by assumption $p^3/N \to 0$. Assertions (A.0.3), (A.0.4) and (A.0.5) yields

$$\sqrt{p} \parallel \hat{R}_u^{-1}(p) - R_u^{-1}(p) \parallel = \sqrt{p}\ q_T(u,p) \leq \frac{k^2(u,p)\sqrt{p}\ Q_T(u,p)}{1 - k(u,p)Q_T(u,p)} \to 0$$

in probability. ■

Proposition A.0.1.

(i) Let $\{\widetilde{X}_t(u)\}$ satisfy Assumptions 2.2, 2.3 and $p \to \infty$ such that $p^3/N \to 0$. Then

$$\| \hat{\beta}_u(p) - \beta_u(p) \| = O_p\left(p^{1/2}N^{-1/2}\right)$$

where $\beta_u(p) = (\beta_1(u), \beta_2(u), \cdots, \beta_p(u))'$.

(ii) If Assumptions 2.1 and 2.3 are satisfied and $p \in \mathbb{N}$ fixed, then

$$\| \hat{\beta}_u(p) - \widetilde{\beta}_u(p) \| = O_p\left(N^{-1/2}\right)$$

where $\widetilde{\beta}_u(p) = (\beta_{1,p}(u), \beta_{2,p}(u), \cdots, \beta_{1,p}(u))'$.

Proof : (i)

$$\begin{aligned}
\| \beta_u(p) - \hat{\beta}_u(p) \| &= \| \hat{R}_u^{-1}(p) \sum_{j=p}^{N-1} X_j(u,p)\Big(X_j(u,p)'\beta_u(p) - X_{[uT]-N/2+j,T}\Big)/(N-p) \| \\
&\leq \| \hat{R}_u^{-1}(p) - R_u^{-1}(p) \| \| \sum_{j=p}^{N-1} X_j(u,p)\varepsilon_{p,j}(u)/(N-p) \| \\
&\quad + \| R_u^{-1}(p) \| \| \sum_{j=p}^{N-1} X_j(u,p)(\varepsilon_{p,j}(u) - \varepsilon_j(u))/(N-p) \| \\
&\quad + \| R_u^{-1}(p) \| \| \sum_{j=p}^{N-1} X_j(u,p)\varepsilon_j(u)/(N-p) \|
\end{aligned}$$

where $\varepsilon_{p,j}(u)$ and $\varepsilon_j(u)$ are defined in Lemma A.0.1.

Since for $1 \leq s \leq p$ we have,

$$E\Big(\sum_{j=p}^{N-1} X_{[uT]-N/2+j-s,T}\varepsilon_j(u)\Big)^2 = \sum_{j=p}^{N-1}\sum_{i=0}^{\infty} \alpha^2_{[uT]-N/2+j-s,T}(i) \leq K(N-p).$$

we get

$$E \| \sum_{j=p}^{N-1} X_j(u,p)\varepsilon_j(u)/(N-p) \|^2 \leq Kp(N-p)^{-1} \to 0. \tag{A.0.6}$$

A direct consequence of Lemma A.0.1 is that

$$\sqrt{\frac{N}{p}}E \| \sum_{j=p}^{N-1} X_j(u,p)(\varepsilon_{p,j}(u) - \varepsilon_j(u))/(N-p) \|^2 = O(\frac{p^{1/2}N^{3/2}}{T^2}) + O(\sqrt{Np}\sup_{0\leq u\leq 1}\sum_{k=p+1}^{\infty} {\beta_k}^2(u)). \tag{A.0.7}$$

(A.0.6) and (A.0.7) imply that

$$\sqrt{\frac{N}{p}} \| \sum_{j=p}^{N-1} X_j(u,p)\varepsilon_{p,j}(u)/(N-p) \| \to 0. \tag{A.0.8}$$

The desired result follows then by (A.0.3), (A.0.8) and Lemma A.0.3.

(ii) By Lemma A.0.2 we have

$$\sqrt{N} \parallel \hat{\beta}_u(p) - \widetilde{\beta}_u(p) \parallel \;=\; \sqrt{N} \parallel \hat{R}_u^{-1}(p)\hat{r}_u(p) - R_u^{-1}(p)r_u(p) \parallel = O_p(N(N-p)^{-1}).$$

■

To establish the next proposition we first define,

$$B(u,z) = 1 - \sum_{k=1}^{\infty} \beta_k(u) z^k, \;\; \hat{B}_p(u,z) = 1 - \sum_{k=1}^{p} \hat{\beta}_k(u) z^k \;\text{ and }\; \widetilde{B}_p(u,z) = 1 - \sum_{k=1}^{p} \beta_{k,p}(u) z^k.$$

Proposition A.0.2.

(i) Under Assumptions 2.1 to 2.3 and for $p \to \infty$ such that $p^3/N \to 0$, we have for every $u \in [0,1]$

$$\sup_{\lambda \in [-\pi,\pi]} \left| \hat{f}_{tvAR}(u,\lambda) - f(u,\lambda) \right| \to 0$$

in probability, as $T \to \infty$.

(ii) Under Assumptions 2.1 and 2.3 and for every $p \in \mathbb{N}$ fixed, we have for every $u \in [0,1]$

$$\sup_{\lambda \in [-\pi,\pi]} \left| \hat{f}_{tvAR}(u,\lambda) - f_{tvAR}(u,\lambda) \right| \to 0$$

in probability, as $T \to \infty$.

Proof: (i) It suffices to show that for every $u \in [0,1]$,

$$\hat{\sigma_p}^2(u) \to \alpha(u,0) \quad \text{and} \quad \hat{B}_p(u,e^{i\lambda}) \to B(u,e^{i\lambda}) \tag{A.0.9}$$

in probability, where the last convergence is uniformly in $\lambda \in [-\pi,\pi]$. By Proposition A.0.1(i) and Assumption 2.2 we have

$$\sup_{\lambda \in [-\pi,\pi]} \left| \hat{B}_p(u,e^{i\lambda}) - B(u,e^{i\lambda}) \right| \le \parallel \hat{\beta}_u(p) - \beta_u(p) \parallel + \sum_{j=p+1}^{\infty} |\beta_j(u)| \to 0.$$

To see that $\hat{\sigma_p}^2(u) \to \alpha(u,0)$ in probability note that $E(\widetilde{X}_t(u))^2 = \sum_{k=1}^{\infty} \beta_k(u)c(u,k) + \alpha(u,0)$, which gives

$$E(\widetilde{X}_t(u))^2 - \sum_{k=1}^{\infty} \beta_k(u)c(u,k) = c(u,0) + \beta_u(p)' r_u(p) + \sum_{k=p+1}^{\infty} \beta_k(u)c(u,k) = \alpha(u,0). \tag{A.0.10}$$

Now,

$$
\begin{aligned}
\hat{\sigma_p}^2(u) - \alpha(u,0) &= \frac{1}{N-p}\sum_{j=p}^{N-1} X^2_{[uT]-N/2+j-p,T} - \hat{\beta}_u(p)'\hat{r}_u(p) - \alpha(u,0)\\
&= \frac{1}{N-p}\sum_{j=p}^{N-1} X^2_{[uT]-N/2+j-p,T} + (\hat{\beta}_u(p)-\beta_u(p))'(\hat{r}_u(p)-r_u(p))\\
&\quad +(\hat{\beta}_u(p)-\beta_u(p))'r_u(p) + \hat{\beta}_u(p)'(\hat{r}_u(p)-r_u(p)) + \beta_u(p)'r_u(p) - \alpha(u,0)
\end{aligned}
$$

and a direct application of (A.0.10) gives

$$
\begin{aligned}
\left|\hat{\sigma_p}^2(u) - \alpha(u,0)\right| &\le |\frac{1}{N-p}\sum_{j=p}^{N-1} X^2_{[uT]-N/2+j-p,T} - c(u,0)| + \| (\hat{\beta}_u(p)-\beta_u(p))' \|\\
&\quad \times \| (\hat{r}_u(p)-r_u(p)) \| + \| (\hat{\beta}_u(p)-\beta_u(p))' \|\\
&\quad \times \| (r_u(p)) + \hat{\beta}_u(p)' \| \| (\hat{r}_u(p)-r_u(p)) \| + \sum_{k=p+1}^{\infty} |\beta_k(u)c(u,k)|\\
&= O(N^{-1/2}p^{-1/2}) + O(N^{1/2}p^{1/2}T^{-1})
\end{aligned}
$$

which implies that the right hand side of the above equation converges to zero.

(ii) Proposition A.0.1(ii) and Lemma A.0.2 imply that for every $u \in [0,1]$ $\sup_{\lambda\in[-\pi,\pi]} |\hat{B}_p(u,e^{i\lambda}) - \widetilde{B}_p(u,e^{i\lambda})| \to 0$ and $\hat{\sigma_p}^2(u) \to c(u,0) - \widetilde{\beta}_u(p)'r_u(p) = \sigma_p^2(u)$ in probability. ■

Lemma A.0.4. *Let $\{X_{t,T}\}$ be a locally stationary process satisfying Assumption 2.1. Then*

(i)

$$I_{N,X}(u,\lambda) = I_{N,\widetilde{X}}(u,\lambda) + \widetilde{R}_N(u,\lambda)$$

where $I_{N,\widetilde{X}}(u,\lambda)$ is the local periodogram of the series $\widetilde{X}_1(u),\ldots,\widetilde{X}_T(u)$ and $E(\widetilde{R}_N(u,\lambda))^2 = O(N^2/T^2)$ uniformly in u and λ .

(ii) For $\{\lambda_j = 2\pi j/N, j = 1,\ldots,(M_N-1)\}$ we have that

$$E\Big(I_{N,X}(u,\lambda_j)\Big) = f(u,\lambda_j) + O\Big(\frac{1}{N}+\frac{N}{T}\Big)$$

$$Var\Big(I_{N,X}(u,\lambda_j)\Big) = f^2(u,\lambda_j) + O\Big(\frac{1}{N}+\frac{N}{T}\Big)$$

$$Cov\Big(I_{N,X}(u,\lambda_j), I_{N,X}(u,\lambda_k)\Big) = \frac{1}{N}\kappa_4 f(u,\lambda_j)f(u,\lambda_k) + O\Big(\frac{N}{T}\Big) + o(N^{-1}) \quad for \quad \lambda_k \neq \lambda_j.$$

Proof: (i) Let $[uT] - N/2 + 1 = \nu$. We then have for the local discrete Fourier transform $J_X(u,\lambda)$ of $\{X_{t,T}\}$ that

$$\begin{aligned} J_X(u,\lambda) &= N^{-1/2} \sum_{s=0}^{N-1} X_{\nu+s,T} e^{-i\lambda s} \\ &= N^{-1/2} \sum_{s=0}^{N-1} Z_{\nu+s,T}(u) e^{-i\lambda s} + N^{-1/2} \sum_{s=0}^{N-1} \widetilde{X}_{\nu+s}(u) e^{-i\lambda s} \\ &= J_Z(u,\lambda) + J_{\widetilde{X}}(u,\lambda) \end{aligned}$$

Using that

$$I_{N,X}(u,\lambda) = I_{N,\widetilde{X}}(u,\lambda) + I_{N,Z}(u,\lambda) + J_{\widetilde{X}}(u,\lambda) J_Z(u,-\lambda) + J_Z(u,\lambda) J_{\widetilde{X}}(u,-\lambda),$$

and (A.0.1) which implies that $E(I_{N,Y}(u,\lambda)) = O\left(N^2/T^2\right)$ and $E(I_{N,Y}(u,\lambda))^2 = O\left(N^4/T^4\right)$ the proof for part (i) is completed.

(ii) The assertion for the variance follows using part (i), and the fact that for each u, $Var(I_{N,\widetilde{X}}(u,\lambda_j)) = f^2(u,\lambda_j) + O\left(N^{-1}\right)$ and $|Cov(\widetilde{R}(u,\lambda_j), I_{N,\widetilde{X}}(u,\lambda_j)| = O(NT^{-1})$ by Cauchy's inequality. For the covariance of the local periodogram we use the same arguments as above and that $Cov\Big(I_{N,\widetilde{X}}(u,\lambda_j), I_{N,\widetilde{X}}(u,\lambda_k)\Big) = \frac{1}{N}\kappa_4 f(u,\lambda_j) f(u,\lambda_k) + o(N^{-1})$ for $\lambda_k \neq \lambda_j$. ∎

Lemma A.0.5.

(i) Let Assumptions 2.1, 2.2, 2.3, 2.5 and 2.6 be satisfied. If $p \to \infty$ such that $p^3/N \to 0$ as $T \to \infty$, then for each $u \in [0,1]$
$\hat{g}(u,\lambda) \to 1$ and $\int_{-\pi}^{\pi} |\hat{g}(u,\lambda) - 1|\, d\lambda \to 0$ in probability.

(ii) Let Assumptions 2.1, 2.3, 2.5 and 2.6 be satisfied. For all fixed $p \in \mathbb{N}$ and for each $u \in [0,1]$ we have that
$\hat{g}(u,\lambda) \to f(u,\lambda)/f_{tvAR}(u,\lambda)$ and $\int_{-\pi}^{\pi} |\hat{g}(u,\lambda) - f(u,\lambda)/f_{tvAR}(u,\lambda)|\, d\lambda \to 0$ in probability.

Proof: (i) By Proposition A.0.2(i) we obtain

$$\begin{aligned}\hat{q}(u,\lambda) &= \frac{1}{N}\sum_{j=-M_N}^{M_N} K_h(\lambda-\lambda_j)\frac{I_N(u,\lambda_j)}{\hat{f}_{tvAR}(u,\lambda_j)}\\ &= \frac{1}{f(u,\lambda)}\frac{1}{N}\sum_{j=-M_N}^{M_N} K_h(\lambda-\lambda_j)I_N(u,\lambda_j)+o_p(1)\\ &= 1+o_p(1).\end{aligned}$$

By Lemma A.0.4 we have

$$\begin{aligned}& E\Big(\int_{-\pi}^{\pi}\Big|\frac{1}{N}\sum_{j=-M_N}^{M_N} K_h(\lambda-\lambda_j)(I_N(u,\lambda_j)-EI_N(u,\lambda_j))\Big|\,d\lambda\Big)^2\\ \leq\ & \frac{1}{N^2}\int_{-\pi}^{\pi}\sum_{j=-M_N}^{M_N} K_h^2(\lambda-\lambda_j)Var\left(I_N(u,\lambda_j)\right)d\lambda\\ & +\int_{-\pi}^{\pi}\sum_{\substack{j,k=-M_N\\ j\neq k}}^{M_N} K_h(\lambda-\lambda_j)K_h(\lambda-\lambda_k)Cov(I_N(u,\lambda_j),I_N(u,\lambda_k))d\lambda=o(1).\end{aligned}$$

(ii) Follows by using the same arguments as in (i) and Proposition A.0.2(ii). ■

Before establishing the next lemma we define E^+ and cov^+ the expectation and the covariance function respectively with respect to the measure $\hat{F}$.

Lemma A.0.6. *(i) Let Assumptions 2.1 and 2.3 be satisfied. For all fixed $p\in\mathbb{N}$ we have that $E^+(\varepsilon_1^+)^4-3\to\kappa_4(p)$ in probability.*
(ii) Let Assumptions 2.1 to 2.3 be satisfied. If $p\to\infty$ such that $p^3/N\to 0$ then $E^+(\varepsilon_1^+)^4-3\to\kappa_4$ in probability.

Proof: (i) We have using $\overline{\varepsilon}=O((N-p)^{-1/2})$ and the notation $Y_{t,T}=X_{t,T}-\sum_{i=1}^{p}\hat{\beta}_i(\frac{t}{T})X_{t-i,T}$ that

$$E^+(\hat{\varepsilon}_t^+)^4=\frac{1}{T-p}\sum_{t=p+1}^{T}\frac{Y_{t,T}^4}{\hat{\sigma}_p^4(t/T)}+o_p(1). \tag{A.0.11}$$

Furthermore, let $\widetilde{Y}_t(u)=\widetilde{X}_t(u)-\sum_{i=1}^{p}\beta_{i,p}(u)\widetilde{X}_{t-i}(u)$ and consider the difference

$$\begin{aligned}Y_{t,T}-\widetilde{Y}_t(t/T) &= \left(X_{t,T}-\sum_{i=1}^{p}\hat{\beta}_i(\frac{t}{T})X_{t-i,T}\right)-\left(\widetilde{X}_t(t/T)-\sum_{i=1}^{p}\beta_{i,p}(\frac{t}{T})\widetilde{X}_{t-i}(t/T)\right)\\ &= Z_{t,T}(t/T)+\sum_{i=1}^{p}(\beta_{i,p}(\frac{t}{T})-\hat{\beta}_i(\frac{t}{T}))\widetilde{X}_{t-i}(t/T)+\sum_{i=1}^{p}\hat{\beta}_i(\frac{t}{T})Z_{t-i,T}(t/T)\\ &= O_p(N^{-1/2})\end{aligned} \tag{A.0.12}$$

By (A.0.12) and because $\hat{\sigma}_p(u) \to \sigma_p(u)$ in probability (see proof of Proposition 6.2(ii)), we deduce that

$$\left| \frac{1}{T-p} \sum_{t=p+1}^{T} \left(\frac{Y_{t,T}^4}{\hat{\sigma}_p^4(t/T)} - \frac{\widetilde{Y}_t^4(t/T)}{\sigma_p^4(t/T)} \right) \right| \to 0$$

in probability. Since $E(\widetilde{Y}_t(u))^4 < \infty \;\; \forall u \in [0,1]$, we have that

$$\frac{1}{T-p} \sum_{t=p+1}^{T} E(\widetilde{Y}_t^4(t/T)) \to \int_0^1 E(\widetilde{Y}_t^4(u))du.$$

(ii) This is proved by using standard arguments for the autoregressive sieve bootstrap, see for instance Kreiss [23], Proposition 3.1. ■

The process $X_{t,T}^+$ possesses the following representation,

$$X_{t,T}^+ = \sum_{j=0}^{\infty} \hat{\alpha}_{t,T}(j)\varepsilon_{t-j}^+ \tag{A.0.13}$$

where $\hat{\alpha}_{t,T}(j)$ are obtained by $(\hat{\alpha}_{t,T}(0) = 1)$:

$$\left(1 - \sum_{j=0}^{p} \hat{\beta}_j(\frac{t}{T})z^j\right)^{-1} = 1 + \sum_{j=0}^{\infty} \hat{\alpha}_{t,T}(j)z^j. \tag{A.0.14}$$

Let Assumptions 2.1 to 2.6 be satisfied and let $p \to \infty$ such that $p^3/N \to 0$. By Proposition A.0.1(i) we have that

$$\sum_{j=0}^{\infty} |\hat{\alpha}_{t,T}(j) - \alpha_{t,T}(j)| = O_p(p^2 N^{-1/2}) \tag{A.0.15}$$

Also under Assumption 2.1 and Assumptions 2.3 to 2.6 and letting the order of the fitted approximating process fixed, by Proposition A.0.1(ii) we have that

$$\sum_{j=0}^{\infty} |\hat{\alpha}_{t,T}(j) - \alpha_{j,p}(t/T)| = O_p(N^{-1/2}) \tag{A.0.16}$$

where $\{\alpha_{j,p}(t/T), j \in \mathbb{N}\}$ is defined as $\{\hat{\alpha}_{t,T}(j), j \in \mathbb{N}\}$ if we replace $\hat{\beta}_j(t/T)$ by $\beta_{j,p}(t/T)$ in (A.0.14), see Kreiss(1999) Lemma 8.3.

Lemma A.0.7. *Let Assumption 2.1 and Assumptions 2.3 to 2.6 and keep $p \in \mathbb{N}$ fixed or Assumptions 2.1 to 2.6 and let $p \to \infty$ such that $p^4/N \to 0$ is satisfied. Let*

$$J_N^+(\phi) = \int_{-\pi}^{\pi} \phi(\lambda)\hat{g}(u,\lambda)I_N^+(u,\lambda)d\lambda \quad \textit{and} \quad J_{N,\widetilde{X}^+}^+(\phi) = \int_{-\pi}^{\pi} \phi(\lambda)\hat{g}(u,\lambda)I_{N,\widetilde{X}^+}^+(u,\lambda)d\lambda$$

where $I^{+}_{N,\widetilde{X}^{+}}(u,\lambda)$ *is the periodogram on the segment* $[uT]-N/2+1,...,[uT]+N/2$ *of the bootstrap process* $\widetilde{X}^{+}_t(u)=\sum_{i=0}^{\infty}\hat{\alpha}_{[uT],T}(i)\varepsilon^{+}_{t-i}$. *Then as* $T\to\infty$,

$$J^{+}_N(\phi)-E^{+}(J^{+}_N(\phi))=J^{+}_{N,\widetilde{X}^{+}}(\phi)-E^{+}(J^{+}_{N,\widetilde{X}^{+}}(\phi))+O_p(N/T+p^2/N)$$

Proof:

Let $\nu=[uT]-N/2+1$,

$$d^{+(1)}_N(u,\lambda)=\frac{1}{\sqrt{N}}\sum_{s=0}^{N-1}\sum_{m=0}^{\infty}\hat{\alpha}_{[uT],T}(m)\varepsilon^{+}_{s-m}\exp(-i\lambda s),$$

$$d^{+(2)}_N(u,\lambda)=\frac{1}{\sqrt{N}}\sum_{s=0}^{N-1}\sum_{m=0}^{\infty}(\hat{\alpha}_{\nu+s,T}(m)-\hat{\alpha}_{[uT],T}(m))\varepsilon^{+}_{s-m}\exp(-i\lambda s)$$

Then

$$J^{+}_N(\phi)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(\lambda)\hat{g}(u,\lambda)\sum_{j,l=1}^{2}d^{+(j)}_N(u,\lambda)d^{+(l)}_N(u,-\lambda)d\lambda=:\sum_{j,l=1}^{2}J^{+(j,l)}_N(\phi)$$

Since $J^{+(1,1)}_N(\phi)=J^{+}_{N,\widetilde{X}^{+}}(\phi)$ it suffices to show that $J^{+(j,l)}_N(\phi)-E^{+}J^{+(j,l)}_N(\phi)=O_p(N/T+p^2/N)$ for $j+l>2$. Using (A.0.15) we get that $E^{+}|(d^{+(1)}_N(u,\lambda))|^2=O_p(1)$, $E^{+}|(d^{+(2)}_N(u,\lambda))|^2=O_p(N^2/T^2+p^4/N^2+p^2N^{-1/2}T^{-1})$ and for $j+l>2$

$$E|N^{1/2}(J^{+(j,l)}_N(\phi)-E^{+}J^{+(j,l)}_N(\phi))|^2\le K(\frac{N^3}{T^2}+\frac{p^4}{N})$$

where $0<K<\infty$ is constant. ■

Proof of Proposition 2.4.1: First write the processes $\widetilde{X}_t(u)$ and $\widetilde{Y}_t(u):=\widetilde{X}_t(u)-\sum_{i=1}^{p}\beta_{i,p}(u)\widetilde{X}_{t-i}(u)$ as $\widetilde{X}_t(u)=\Psi(u,B)\varepsilon_t$ where $\Psi(u,z)=\sum_{j=0}^{\infty}\alpha(u,j)z^j$, and $\widetilde{Y}_t(u)=\Phi_p(u,B)\widetilde{X}_t(u)$ where $\Phi_p(u,z)=1-\sum_{j=1}^{p}\beta_{j,p}(u)z^j$ respectively, where B is the backward shift operator defined by $B^jX_t=X_{t-j}$. Observe that

$$\widetilde{Y}_t(u)=\Lambda_p(u,B)\varepsilon_t$$

where

$$\Lambda_p(u,z)=\Phi_p(u,z)\Psi(u,z)=\sum_{j=0}^{\infty}\ell_p(u,j)z^j,$$

$\ell_p(u,0)=1$ and $\ell_p(u,j)=\alpha(u,j)-\sum_{k=1}^{p}\beta_{j,p}(u)\alpha(u,j-k)$ for $j\ge 1$. Straightforward calculations give

$$cum_s(\widetilde{Y}_t(u))=cum_s(\varepsilon_t)\sum_{j=0}^{\infty}\ell_p(u,j)=cum_s(\varepsilon_t)+cum_s(\varepsilon_t)L_s(u,p)$$

where cum_s is the sth ordered cumulant and $L_s(u,p) = \sum_{j=0}^{\infty} \ell_p^s(u,j)$. Recall the definition of $\kappa_4(p)$ in Theorem 4.1 to see that

$$\kappa_4(p) = \int_0^1 \frac{cum_4(\widetilde{Y}_t(u))}{cum_2^2(\widetilde{Y}_t(u))} du$$

and use that $L_4(u,p) \leq L_2^2(u,p)$ to get

$$\begin{aligned} \left|\frac{\kappa_4(p)}{\kappa_4} - 1\right| &= \left|\int_0^1 \frac{(L_2(u,p)-2)L_2(u,p)+L_4(u,p)}{(1+L_2(u,p))^2} du\right| \\ &\leq 2\int_0^1 \frac{L_2(u,p)}{1+L_2(u,p)} du. \end{aligned}$$

■

Proof of Theorem 2.4.1 and Theorem 2.4.2:

Using similar arguments as in Theorem 4.1 of Kreiss and Paparoditis [24] we have that

$$\begin{aligned} & cov^+(J^+_{N,\widetilde{X}^+}(\phi_i), J^+_{N,\widetilde{X}^+}(\phi_j)) = \\ & 2\pi\Big\{ \int_{-\pi}^{\pi} \phi_i(\lambda)\{\phi_j(\lambda)+\phi_j(-\lambda)\}\hat{q}^2(u,\lambda) f^2_{tvAR}(u,\lambda) d\lambda \\ + \quad & \kappa_4^+ \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_i(\lambda)\phi_j(-\mu)\hat{g}(u,\lambda) f_{tvAR}(u,\lambda)\hat{q}(u,\mu) f_{tvAR}(u,\mu) d\lambda d\mu \Big\} \end{aligned}$$

where $J^+_{N,\widetilde{X}^+}(\phi)$ is defined in Lemma A.0.7 and κ_4^+ is the fourth cumulant of the bootstrapped residuals ε_t^+ . Because of Lemmas A.0.5 and A.0.6 and Proposition A.0.2 we can replace $E^+(\varepsilon_1^+)^4$, $\hat{g}(u,\lambda)$ and $\hat{f}_{tvAR}(u,\lambda)$ by their limits and we obtain that, $cov^+(J^+_{N,\widetilde{X}^+}, J^+_{N,\widetilde{X}^+}) \to cov(\xi_i,\xi_j)$ in probability which with Lemma A.0.7 complete the proof of both theorems. ■

Proof of Theorem 2.4.4: Using Lemma A.0.4 (i) we can show that

$$\sqrt{Nb}\left(\hat{f}(u,\lambda) - E(\hat{f}(u,\lambda))\right) = \sqrt{Nb}\left(\widetilde{f}(u,\lambda) - E(\widetilde{f}(u,\lambda))\right) + o_p(1) \qquad \text{(A.0.17)}$$

where $\widetilde{f}(u,\lambda) = N^{-1}\sum_{j=-M_N}^{M_N} K_b(\lambda-\lambda_j) I_{N,\widetilde{X}}(u,\lambda_j)$ and $I_{N,\widetilde{X}}(u,w)$ is the local periodogram of the stationary series $\widetilde{X}_1(u),\ldots,\widetilde{X}_T(u)$. Recall that $\{\widetilde{X}_s(u)\}$ is a stationary process which approximates $X_{t,T}$ in a local neighborhood around $u = t/T$. Now, set

$$\widetilde{f}^*(u,\lambda) = N^{-1} \sum_{j=-M_N}^{M_N} K_b(\lambda - \lambda_j) I^*_{N,\widetilde{X}^*}(u,\lambda_j)$$

where $I^*_{N,\widetilde{X}^*}(u,\lambda) = \hat{g}(u,\lambda) I^+_{N,\widetilde{X}^+}(u,\lambda)$ and $I^+_{N,\widetilde{X}^+}(u,\lambda)$ is the local periodogram of the segment of observations $X^+_{[uT]-N/2+1}(u), ..., X^+_{[uT]+N/2}(u)$ coming from the stationary bootstrap process $\widetilde{X}^+_t(u) = \sum_{i=0}^{\infty} \hat{\alpha}_{[uT],T}(i)\varepsilon^+_{t-i}$. Following the same steps as in Lemma A.0.7 the following result can be established

$$\sqrt{Nb}\left(\hat{f}^*(u,\lambda) - E(\hat{f}^*(u,\lambda))\right) = \sqrt{Nb}\left(\widetilde{f}^*(u,\lambda) - E(\widetilde{f}^*(u,\lambda))\right) + o_p(1). \quad \text{(A.0.18)}$$

The assertion of the theorem follows then by the same arguments as of those used in the proof of Theorem 5.1 of Kreiss and Paparoditis [24], since $\sqrt{Nb}\left(\widetilde{f}(u,\lambda) - E(\widetilde{f}(u,\lambda))\right)$ and $\sqrt{Nb}\left(\widetilde{f}^*(u,\lambda) - E(\widetilde{f}^*(u,\lambda))\right)$ are based on realizations of stationary processes.

Appendix B

Auxiliary results and proofs for Chapter 3

A useful tool for handling taper data, is the periodic extension (with period 2π) of the function $L_T(\alpha) : \mathbb{R} \to \mathbb{R}$, with

$$L_T(\alpha) = \begin{cases} T, & |\alpha| \leq 1/T \\ 1/|\alpha|, & 1/T \leq |\alpha| \leq \pi. \end{cases} \tag{B.0.1}$$

For the proof of the following lemma see Dahlhaus [9].

Lemma B.0.8.

a. $\int_\Pi L_T^k(\alpha) \leq KT^{k-1}$ *for all* $k > 1$.

b. $\int_\Pi L_T(\alpha) \leq K \log(T)$

c. $|\alpha| L_T(\alpha) \leq K$

d. $\int_\Pi L_T(\beta - \alpha) L_S(\alpha + \gamma) \leq K \max\{\log(T), \log(S)\} L_{\min\{T,S\}}(\beta + \gamma)$

For a complex-valued function f define $H_N(f(\cdot), \lambda) := \sum_{s=0}^{N-1} f(s) e^{-i\lambda s}$ and let

$$H_{k,N}(\lambda) = H_N(h^k(\frac{\cdot}{N}), \lambda)$$

and

$$H_N(\lambda) = H_{1,N}(\lambda).$$

Straightforward calculation gives

$$\sum_j H_{k,N}(\alpha - \lambda_j) H_{\ell,N}(\lambda_j - \beta) = 2\pi N H_{k+\ell,N}(\alpha - \beta)$$

where the sum extends over all the Fourier frequencies

$\lambda_j = 2\pi j/N, \quad j = -[(N-1)/2], \ldots, [N/2]$. Under Assumption 3.1.2 (iv) there is a constant C independent of T and λ such that

$$|H_{k,N}(\lambda)| \leq C L_N(\alpha) \tag{B.0.2}$$

and

$$K_b(\lambda) \leq C b L^2_{1/b}(\lambda). \tag{B.0.3}$$

Lemma B.0.9.

(i) Let $N, T \in \mathbb{N}$. Suppose that the data taper h satisfies Assumption 3.1.2 (iv) and $\psi : [0,1] \to \mathbb{R}$ is Lipshitz continuous. Then we have for $0 \leq t \leq N$,

$$H_N\left(\psi\left(\frac{\cdot}{T}\right) h\left(\frac{\cdot}{N}\right), \lambda\right) = \psi\left(\frac{t}{T}\right) H_N(\lambda) + O\left(\frac{N}{T} L_N(\lambda)\right).$$

The same holds, if $\psi(\cdot/T)$ on the left side is replaced by numbers $\psi_{s,T}$ with $\sup_s |\psi_{s,T} - \psi(s/T)| = O(T^{-1})$

(ii) Let $t_j = S(j-1) + N/2$, $u_j = t_j/T$ with N, M, S and T satisfying Assumption 3.1.2 and $\psi : [0,1] \to \mathbb{R}$ be Lipshitz continuous. Then

$$\left|\sum_{j=1}^{M} \psi(u_j) e^{i\lambda S j}\right| \leq K L_M(S\lambda).$$

Proof: The proof is identical to the proof of Lemmas A.5 and A.6 in Dahlhaus [9]. ■

Before proceeding with the next lemma we use for simplicity the notation $f_{\vartheta_0}(u,\lambda)$ for $f(u,\lambda;\vartheta_0(u))$ and $f_{\hat{\vartheta}}(u,\lambda)$ for $f(u,\lambda;\hat{\vartheta}(u))$.

Lemma B.0.10. *Under Assumptions 3.1.1, 3.1.2 and if H_0 is true, then*

$$E(N\sqrt{Mb}\,Q_{0,T}) = \mu_T + o(1)$$

where

$$\mu_T = \frac{M^{1/2}c_{tap}}{b^{1/2}}\int_{-\pi}^{\pi}K^2(x)dx + \frac{M^{1/2}b^{1/2}c_{tap}}{4\pi}\int_{-\pi}^{\pi}\int_{-2\pi}^{2\pi}K(x)K(x-u)dxdu$$

and $c_{tap} = \int_0^1 h^4(x)/(\int_0^1 h^2(x))^2$.

Proof: First note that

$$\begin{aligned} E(N\sqrt{Mb}Q_{0,T}) &= \frac{b^{1/2}}{M^{1/2}N}\sum_{m=1}^{M}\int_{-\pi}^{\pi}\sum_j\sum_s \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_j)f_{\vartheta_0}(u_i,\lambda_s)} \\ &\quad\times cum\left(I_N(u_m,\lambda_j), I_N(u_m,\lambda_s)\right)d\lambda \\ &\quad\quad + O(\sqrt{Mb}N^5/T^4 + \sqrt{Mb}\log^2(N)/N). \end{aligned}$$

Using (3.1.3) and the following property for cumulants

$$cum(Z_1Z_2, Z_3Z_4) = cum(Z_1,Z_3)cum(Z_2,Z_4) + cum(Z_1,Z_4)cum(Z_2,Z_3), \tag{B.0.4}$$

for Z_i Gaussian random variables, we get

$$\begin{aligned} E(N\sqrt{Mb}Q_{0,T}) &= \frac{b^{1/2}}{4\pi^2M^{1/2}NH_{2,N}^2(0)}\sum_{m=1}^{M}\int_{-\pi}^{\pi}\sum_j\sum_s \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_m,\lambda_j)f_{\vartheta_0}(u_m,\lambda_s)}d\lambda \\ &\quad\times\Big(cum\left(d_N(u_m,\lambda_j), d_N(u_m,\lambda_s)\right)cum\left(d_N(u_m,-\lambda_j), d_N(u_m,-\lambda_s)\right) \\ &\quad + cum\left(d_N(u_m,\lambda_j), d_N(u_m,-\lambda_s)\right)cum\left(d_N(u_m,-\lambda_j), d_N(u_m,\lambda_s)\right)\Big) \\ &\quad + o(1) \\ &= \mu_{1,T} + \mu_{2,T} + o(1) \end{aligned}$$

with an obvious notation for $\mu_{i,T}$, $i = 1, 2$.

Recall the definition of $d_N(u,\lambda)$ to see that

$$\begin{aligned} & cum\left(d_N(u_m,\lambda_j), d_N(u_m,\lambda_s)\right)cum\left(d_N(u_m,-\lambda_j), d_N(u_m,-\lambda_s)\right) \\ = & \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_N(A^0_{t_m-N/2+1+\cdot,T}(\mu_1)h(\frac{\cdot}{N}),\lambda_j-\mu_1)H_N(A^0_{t_m-N/2+1+\cdot,T}(-\mu_1)h(\frac{\cdot}{N}),-\lambda_s+\mu_1) \\ & \times H_N(A^0_{t_m-N/2+1+\cdot,T}(\mu_2)h(\frac{\cdot}{N}),-\lambda_j-\mu_2)H_N(A^0_{t_m-N/2+1+\cdot,T}(-\mu_2)h(\frac{\cdot}{N}),\lambda_s+\mu_2)d\mu_1 d\mu_2. \end{aligned}$$

Substituting $A^0_{t_m-N/2+1+\cdot,T}(\mu_2)$ by $A(t/T,\mu_2)$ on the above expression, and using (1.1.9) and the fact that $A(\cdot,\cdot)$ is Lipshitz continuous, we get that the above term is equal to

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\vartheta_0}(u_m,\mu_1)f_{\vartheta_0}(u_m,\mu_2)H_N(\lambda_j-\mu_1)H_N(-\lambda_s+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_s+\mu_2)d\mu_1 d\mu_2 \tag{B.0.5}$$

plus a remainder term R_m depending on the difference $|A^0_{t_m-N/2+1+\cdot,T}(\mu_2)-A(t/T,\mu_2)|$ which satisfies

$$\begin{aligned}
&\left|\frac{b^{1/2}}{M^{1/2}N^3}\sum_{m=1}^{M}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_m,\lambda_j)f_{\vartheta_0}(u_m,\lambda_s)}R_m\right| \\
\leq\ & \frac{N}{T}\frac{b^{1/2}M^{1/2}}{N^3}\log^2(N)b\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}L^2_{1/b}(\lambda_j-\lambda_s)L^2_N(\lambda_j-\lambda_s) \\
=\ & O(\frac{NM^{1/2}\log^2(N)}{b^{1/2}T}).
\end{aligned} \tag{B.0.6}$$

Using the bound (B.0.6) and replacing $f_{\vartheta_0}(u_m,\mu_1)$ and $f_{\vartheta_0}(u_m,\mu_2)$ by $f_{\vartheta_0}(u_m,\lambda_j)$ and $f_{\vartheta_0}(u_m,\lambda_s)$ respectively, we get that the term $\mu_{1,T}$ is equal to

$$\begin{aligned}
&\frac{b^{1/2}M^{1/2}}{NH^2_{2,N}(0)}\int_{-\pi}^{\pi}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)|H_N(\lambda_j-\lambda_s)|^2d\lambda \\
=\ & \frac{M^{1/2}c_{tap}}{b^{1/2}}\int_{-\pi}^{\pi}K^2(x)dx+O(\frac{\log(N)M^{1/2}}{Nb^{3/2}})
\end{aligned}$$

plus a remainder term R_m which depends on the difference $|f_{\vartheta_0}(u_m,\mu_2)-f_{\vartheta_0}(u_m,\lambda_j)|$ and which satisfies

$$\begin{aligned}
&\left|\frac{b^{1/2}}{M^{1/2}N^3}\sum_{m=1}^{M}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)}{f_{\vartheta_0}(u_m,\lambda_j)f_{\vartheta_0}(u_m,\lambda_s)}R_m\right| \\
\leq\ & \frac{b^{1/2}M^{1/2}}{N^3}\int_{-\pi}^{\pi}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)L_N(\lambda_s-\lambda_j)d\lambda \\
=\ & O(\frac{b^{1/2}M^{1/2}\log^2(N)}{N}).
\end{aligned}$$

Using similar arguments we get that the second term $\mu_{2,T}$ is equal to

$$\begin{aligned}
&\frac{b^{1/2}M^{1/2}}{NH^2_{2,N}(0)}\int_{-\pi}^{\pi}\sum_{j=-J_N}^{J_N}\sum_{s=-J_N}^{J_N}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)|H_N(\lambda_j+\lambda_s)|^2d\lambda+o(1) \\
=\ & \frac{M^{1/2}b^{1/2}c_{tap}}{4\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}K(x)K(x-u)dxdu+o(1)
\end{aligned}$$

$\blacksquare$

Lemma B.0.11. *Under Assumptions 3.1.1 and 3.1.2 and if H_0 is true, then*

$$Var(N\sqrt{Mb}Q_{0,T})=\tau^2+o(1)$$

where τ^2 is defined in Theorem 3.1.1.

Proof: First note that

$$
\begin{aligned}
& Var(N\sqrt{Mb}Q_{0,T}) \\
= \; & \frac{b}{MN^2}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l=-J_N}^{J_N}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
& \Big(cum(I_N(u_{m_1},\lambda_j),I_N(u_{m_2},\lambda_k))cum(I_N(u_{m_1},\lambda_s),I_N(u_{m_2},\lambda_l)) \\
& +cum(I_N(u_{m_1},\lambda_j),I_N(u_{m_2},\lambda_l))cum(I_N(u_{m_1},\lambda_s),I_N(u_{m_2},\lambda_k)) \\
& +cum(I_N(u_{m_1},\lambda_j),I_N(u_{m_2},\lambda_k),I_N(u_{m_1},\lambda_s),I_N(u_{m_2},\lambda_l))\Big)d\lambda d\mu + o(1) \\
= \; & V_{1,T}+V_{2,T}+V_{3,T}
\end{aligned}
$$

with an obvious notation for $V_{i,T}\;\; i=1,2,3$. From (3.1.3) and (B.0.4) we get that the terms $V_{1,T}$ and $V_{2,T}$ can be written as the sum of four terms, that is for $j=1,2$ we can write $V_{j,T}=\sum_{i=1}^{4}V_{j,T}^{(i)}$.

The first term in this decomposition equals

$$
\begin{aligned}
V_{1,T}^{(i)} \;=\; & \frac{b}{MN^2}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,l,s}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
& cum\left(d_N(u_{m_1},\lambda_j),d_N(u_{m_2},-\lambda_k)\right)cum\left(d_N(u_{m_1},-\lambda_j),d_N(u_{m_2},\lambda_k)\right) \\
& \times cum\left(d_N(u_{m_1},\lambda_s),d_N(u_{m_2},-\lambda_l)\right)cum\left(d_N(u_{m_1},-\lambda_s),d_N(u_{m_2},\lambda_l)\right)
\end{aligned}
$$

To handle this term notice that using arguments similar to those used in the proof of Lemma B.0.10 we have that the term

$$
\begin{aligned}
& \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_1)h(\frac{\cdot}{N}),\lambda_j-\mu_1) \\
& \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_1)h(\frac{\cdot}{N}),-\lambda_k+\mu_1)H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_2)h(\frac{\cdot}{N}),-\lambda_j-\mu_2) \\
& \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_2)h(\frac{\cdot}{N}),\lambda_k+\mu_2)H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_3)h(\frac{\cdot}{N}),\lambda_s-\mu_3) \\
& \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_3)h(\frac{\cdot}{N}),-\lambda_l+\mu_3)H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\mu_4)h(\frac{\cdot}{N}),-\lambda_s-\mu_4) \\
& \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(-\mu_4)h(\frac{\cdot}{N}),\lambda_l+\mu_4) \\
& \times \exp\left\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\right\}d\mu_1 d\mu_2 d\mu_3 d\mu_4
\end{aligned}
$$

is equal to

$$
\begin{aligned}
& \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}A(u_{m_1},\mu_1)A(u_{m_2},-\mu_1)A(u_{m_1},\mu_2)A(u_{m_2},-\mu_2)A(u_{m_1},\mu_3)A(u_{m_2},-\mu_3) \\
& \times A(u_{m_1},\mu_4)A(u_{m_2},-\mu_4)H_N(\lambda_j-\mu_1)H_N(-\lambda_k+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_k+\mu_2) \\
& \times H_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)H_N(-\lambda_s-\mu_4)H_N(\lambda_l+\mu_4) \\
& \times \exp\left\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\right\}d\mu_1 d\mu_2 d\mu_3 d\mu_4 + R_1(m_1,m_2)
\end{aligned}
$$

where $R_1(m_1, m_2)$ satisfies

$$\frac{b}{MN^2}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}\frac{K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)}$$
$$\times R_1(m_1,m_2)d\lambda d\mu$$
$$\leq \frac{K_1}{H_N^4}\frac{b}{MN^2}\frac{N}{T}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)$$
$$\times L_N(\lambda_j-\mu_1)L_N(-\lambda_k+\mu_1)L_N(-\lambda_j-\mu_2)L_N(\lambda_k+\mu_2)L_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)$$
$$\times L_N(-\lambda_s-\mu_4)L_N(\lambda_l+\mu_4)L_M^2(S(\mu_1+\mu_2+\mu_3+\mu_4))d\mu_1d\mu_2d\mu_3d\mu_4d\lambda d\mu \qquad \text{(B.0.7)}$$

since by Lemma A.6 of Dahlhaus [9] we have

$$\sum_{m=1}^{M}\frac{1}{f_{\vartheta_0}(u_m,\lambda_k)f_{\vartheta_0}(u_m,\lambda_l)}e^{i((\mu_1+\mu_2+\mu_3+\mu_4)(Sm)}=O(L_M(S(\mu_1+\mu_2+\mu_3+\mu_4))).$$

Now using Lemma A.4(e) and Lemma A.4(j) of Dahlhaus [9], expression (B.0.7) can be bounded by

$$\frac{K_1}{H_N^4}\frac{b}{MN^2}\frac{N}{T}\frac{NM}{S}\log^3(N)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l}K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)$$
$$\times L_N^2(\lambda_j-\lambda_k)L_N^2(\lambda_s-\lambda_l)d\lambda d\mu=O(\frac{\log^3(N)N^2}{ST}).$$

Furthermore, replacing $A(u_i,\mu_1)$ by $A(u_i,\lambda_j)$ in the first term of (B.0.7) we get that

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}A(u_{m_1},\mu_1)A(u_{m_2},-\mu_1)A(u_{m_1},\mu_2)A(u_{m_2},-\mu_2)A(u_{m_1},\mu_3)A(u_{m_2},-\mu_3)$$
$$\times A(u_{m_1},\mu_4)A(u_{m_2},-\mu_4)H_N(\lambda_j-\mu_1)H_N(-\lambda_k+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_k+\mu_2)$$
$$\times H_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)H_N(-\lambda_s-\mu_4)H_N(\lambda_l+\mu_4)$$
$$\times\exp\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\}\,d\mu_1d\mu_2d\mu_3d\mu_4$$
$$=\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}A(u_{m_1},\lambda_j)A(u_{m_2},-\mu_1)A(u_{m_1},\mu_2)A(u_{m_2},-\mu_2)A(u_{m_1},\mu_3)A(u_{m_2},-\mu_3)$$
$$\times A(u_{m_1},\mu_4)A(u_{m_2},-\mu_4)H_N(\lambda_j-\mu_1)H_N(-\lambda_k+\mu_1)H_N(-\lambda_j-\mu_2)H_N(\lambda_k+\mu_2)$$
$$\times H_N(\lambda_s-\mu_3)H_N(-\lambda_l+\mu_3)H_N(-\lambda_s-\mu_4)H_N(\lambda_l+\mu_4)$$
$$\times\exp\{i(\mu_1+\mu_2+\mu_3+\mu_4)(t_{m_1}-t_{m_2})\}\,d\mu_1d\mu_2d\mu_3d\mu_4+R_2(m_1,m_2)$$

where the remainder term $R_2(m_1, m_2)$ satisfies

$$\frac{b}{MN^2H_N^4}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}\frac{K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)}$$
$$\times R_2(m_1,m_2)d\lambda d\mu$$
$$\leq \frac{K_1}{H_N^4}\frac{b}{MN^2}\frac{MN}{S}\log^3(N)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,s,l}K_b(\lambda-\lambda_j)K_b(\mu-\lambda_s)K_b(\lambda-\lambda_k)K_b(\mu-\lambda_l)$$
$$\times L_N(\lambda_j-\lambda_k)L_N^2(\lambda_l-\lambda_s)d\lambda d\mu = O(\frac{\log^3(N)}{S}).$$

From (B.0.8) we get that the term $V_{1,T}^{(1)}$ can be expressed as:

$$\begin{aligned}
V_{1,T}^{(1)} &= \frac{b}{16\pi^4MN^2H_N^4b^4}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Bigg[\sum_{j,k}K(\frac{\lambda-\lambda_j}{b})K(\frac{\mu-\lambda_k}{b})\sum_{s_1,s_2,s_3,s_4}h(\frac{s_1}{N})h(\frac{s_2}{N}) \\
&\quad\times h(\frac{s_3}{N})h(\frac{s_4}{N})e^{-i[s_1\lambda_j-s_2\lambda_k-s_3\lambda_j+s_4\lambda_k]}\int_{-\pi}^{\pi}e^{i[\mu_1(s_1-s_2)+\mu_1S(m_1-m_2)]}d\mu_1 \\
&\quad\times\int_{-\pi}^{\pi}e^{i[\mu_2(s_3-s_4)+\mu_2S(m_1-m_2)]}d\mu_2\Bigg]^2 d\lambda d\mu+o(1) \\
&= \frac{b}{N^2H_N^4b^4}\sum_{|m|<N/S}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Bigg[\sum_{j,k}K(\frac{\lambda-\lambda_j}{b})K(\frac{\mu-\lambda_k}{b}) \\
&\quad\times\left|\sum_{s_1=0}^{N-1-Sm}h(\frac{s_1}{N})h(\frac{s_1+Sm}{N})e^{-i[s_1(\lambda_j-\lambda_k)]}\right|^2\Bigg]^2 d\lambda d\mu+o(1),
\end{aligned}$$

which by straightforward calculations yield

$$V_{1,T}^{(1)} = \frac{\sum_{|m|<\kappa}\left(\int_0^{1-m/\kappa}h^2(u)h^2(u+m/\kappa)du\right)^2}{2\pi\left(\int_0^1 h^2(x)\right)^4}\int_{-2\pi}^{2\pi}\left(\int K(u)K(u+x)du\right)^2dx$$
$$+O(\frac{\log^2(N)}{N^2b^4}).$$

The terms $V_{1,T}^{(j)}$, $j=2,3,4$ are handled similarly. In particular, we get

$$V_{1,T}^{(2)} = \frac{\sum_{|m|<\kappa}\left(\int_0^{1-m/\kappa}h^2(u)h^2(u+m/\kappa)du\right)^2}{2\pi\left(\int_0^1 h^2(x)\right)^4}\int_{-2\pi}^{2\pi}\left(\int K(u)K(u-x)du\right)^2dx+o(1),$$

$$\begin{aligned}
V_{1,T}^{(3)} &= \frac{b}{MN^2}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,l,s}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
&\quad\times cum\left(d_N(u_{m_1},\lambda_j),d_N(u_{m_2},\lambda_k)\right)cum\left(d_N(u_{m_1},-\lambda_j),d_N(u_{m_2},-\lambda_k)\right) \\
&\quad\times cum\left(d_N(u_{m_1},\lambda_s),d_N(u_{m_2},-\lambda_l)\right)cum\left(d_N(u_{m_1},-\lambda_s),d_N(u_{m_2},\lambda_l)\right)=O(b)
\end{aligned}$$

and

$$
\begin{aligned}
V_{1,T}^{(4)} &= \frac{b}{MN^2}\sum_{m_1,m_2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,k,l,s}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
&\quad \times cum\,(d_N(u_{m_1},\lambda_j),d_N(u_{m_2},-\lambda_k))\,cum\,(d_N(u_{m_1},-\lambda_j),d_N(u_{m_2},\lambda_k)) \\
&\quad \times cum\,(d_N(u_{m_1},\lambda_s),d_N(u_{m_2},\lambda_l))\,cum\,(d_N(u_{m_1},-\lambda_s),d_N(u_{m_2},-\lambda_l)) \\
&= O(b).
\end{aligned}
$$

The term $V_{2,T}$ has the same structure as the term $V_{1,T}$ and converges, therefore, to the same limit. Finally, for the term $V_{3,T}$ we have

$$
\begin{aligned}
V_{3,T} &= \frac{b}{MN^2}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,s,k,l=-J_N}^{J_N}\frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\
&\quad \times cum(d_N(u_{m_1},\lambda_j)d_N(u_{m_1},-\lambda_j),d_N(u_{m_2},\lambda_k)d_N(u_{m_2},-\lambda_k),d_N(u_{m_1},\lambda_s) \\
&\quad \times d_N(u_{m_1},-\lambda_s),d_N(u_{m_2},\lambda_l)d_N(u_{m_2},-\lambda_l))d\lambda d\mu.
\end{aligned}
$$

To handle this term notice that using the product theorem of cumulants, see Brillinger [4], we have to sum over all indecomposable partitions $P_1,\ldots,P_m$ of the scheme

$$
\begin{array}{cc}
a_1 & b_1 \\
a_2 & b_2 \\
a_3 & b_3 \\
a_4 & b_4
\end{array}
$$

where a_1 stands for the position of $d_N(u_{m_1},\lambda_j)$, b_1 for the position of $d_N(u_{m_1},-\lambda_j)$, etc. Following the notation of Dahlhaus [9], let $P_i=\{c_1,\ldots,c_k\}$, $\overline{P}_i:=\{c_1,\ldots,c_{k-1}\}$,

$\beta_{P_i} := (\beta_{c_1}, \ldots, \beta_{c_{k-1}})$ and $\beta_{c_k} = -\sum_{j=1}^{k-1} \beta_{c_j}$. Also, let m be the size of the corresponding partition and $\beta = (\beta_{\overline{P}_i}, \ldots, \beta_{\overline{P}_m})$. We then get

$$\begin{aligned} V_{3,T} \;=\; & \frac{b}{MN^2H_N^4} \sum_{ip} \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{j,s,k,l} \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\ & \int_{\Pi^{8-m}} H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{a_1})h(\frac{\cdot}{N}), \lambda_j - \beta_{a_1}) \\ & \times H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{b_1})h(\frac{\cdot}{N}), -\lambda_j - \beta_{b_1}) \\ & \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{a_2})h(\frac{\cdot}{N}), \lambda_k - \beta_{a_2}) H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{b_2})h(\frac{\cdot}{N}), -\lambda_k - \beta_{b_2}) \\ & \times H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{a_3})h(\frac{\cdot}{N}), \lambda_s - \beta_{a_3}) H_N(A^0_{t_{m_1}-N/2+1+\cdot,T}(\beta_{b_3})h(\frac{\cdot}{N}), -\lambda_s - \beta_{b_3}) \\ & \times H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{a_4})h(\frac{\cdot}{N}), \lambda_l - \beta_{a_4}) H_N(A^0_{t_{m_2}-N/2+1+\cdot,T}(\beta_{b_4})h(\frac{\cdot}{N}), -\lambda_l - \beta_{b_4}) \\ & \prod_{\nu=1}^{m} g_{|P_\nu|}(\beta_{\overline{P}_\nu}) \exp\{i(t_{m_1}(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}) \\ & \qquad + t_{m_2}(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))\}\, d\beta d\lambda d\mu \end{aligned} \tag{B.0.8}$$

Now replace in (B.0.8) the terms $H_N(A^0_{t_{m_i}-N/2+1+\cdot,T}(\beta)h(\frac{\cdot}{N}), -\lambda_k - \beta)$ by $A(u_{m_i}, \beta)H_N(-\lambda_k - \beta)$ to get

$$\begin{aligned} V_{3,T} \;=\; & \frac{b}{MN^2H_N^4} \sum_{ip} \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{j,s,k,l} \frac{K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l)}{f_{\vartheta_0}(u_{m_1},\lambda_j)f_{\vartheta_0}(u_{m_1},\lambda_s)f_{\vartheta_0}(u_{m_2},\lambda_k)f_{\vartheta_0}(u_{m_2},\lambda_l)} \\ & \int_{\Pi^{8-m}} A(u_{m_1},\beta_{a_1})H_N(\lambda_j-\beta_{a_1})A(u_{m_1},\beta_{b_1})H_N(-\lambda_j-\beta_{b_1})A(u_{m_2},\beta_{a_2})H_N(\lambda_k-\beta_{a_2}) \\ & \times A(u_{m_2},\beta_{b_2})H_N(-\lambda_k-\beta_{b_2})A(u_{m_1},\beta_{a_3})H_N(\lambda_s-\beta_{a_3})A(u_{m_1},\beta_{b_3})H_N(-\lambda_s-\beta_{b_3}) \\ & \times H_N(A(u_{m_2},\beta_{a_4})H_N(\lambda_l-\beta_{a_4})A(u_{m_2},\beta_{b_4})H_N(-\lambda_l-\beta_{b_4}) \prod_{\nu=1}^{m} g_{|P_\nu|}(\beta_{\overline{P}_\nu}) \\ & \times \exp\{i(t_{m_1}(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}) + t_{m_2}(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))\}\, d\beta d\lambda d\mu \\ & + E_T, \end{aligned} \tag{B.0.9}$$

where due to the indecomposability of the partitions considered, the following upper bound is true for the error term E_T

$$\begin{aligned} & \frac{b}{MN^2H_N^4} \frac{N}{T} \sum_{ip} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{j,s,k,l=-J_N}^{J_N} K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)K_b(\mu-\lambda_k)K_b(\mu-\lambda_l) \\ & \int_{\Pi^{8-m}} L_N(\lambda_j-\beta_{a_1})L_N(-\lambda_j-\beta_{b_1})L_N(\lambda_k-\beta_{a_2})L_N(-\lambda_k-\beta_{b_2})L_N(\lambda_s-\beta_{a_3}) \\ & \times L_N(-\lambda_s-\beta_{b_3})L_N(\lambda_l-\beta_{a_4})L_N(-\lambda_l-\beta_{b_4})L_M(S(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3})) \\ & \times L_M(S(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))d\beta d\lambda d\mu \\ \le\; & \frac{b\log^4(N)}{MN^2} \frac{N}{T} \sum_{ip} \int_{\Pi^{8-m}} L_N(-\beta_{a_1}-\beta_{b_1})L_N(-\beta_{a_2}-\beta_{b_2})L_N(-\beta_{a_3}-\beta_{b_3})L_N(-\beta_{a_4}-\beta_{b_4}) \\ & \times L_M(S(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}))L_M(S(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))d\beta. \end{aligned}$$

Therefore, and because $-\beta_{a_i}-\beta_{b_i}\neq 0 \quad \forall i$, $\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}\neq 0$ and $\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}\neq 0$, we get that

$$\begin{aligned}
& \frac{b\log^4(N)}{MN^2}\frac{N}{T}\sum_{ip}\int_{\Pi^{8-m}} L_N(-\beta_{a_1}-\beta_{b_1})L_N(-\beta_{a_2}-\beta_{b_2})L_N(-\beta_{a_3}-\beta_{b_3})L_N(-\beta_{a_4}-\beta_{b_4}) \\
& \times L_M(S(\beta_{a_1}+\beta_{b_1}+\beta_{a_3}+\beta_{b_3}))L_M(S(\beta_{a_2}+\beta_{b_2}+\beta_{a_4}+\beta_{b_4}))d\beta \\
\leq\ & \frac{b\log^4(N)}{N^2M}\frac{N}{T}\frac{N^4}{S^3}\log^3(M)\log^3(S)\to 0.
\end{aligned}$$

Similarly the first term on the right hand side of (B.0.9) is bounded by

$$\frac{b\log^4(N)}{N^2M}\frac{N^4}{S^3}\log^3(M)\log^3(S)\to 0,$$

which shows that $V_{3,T}\to 0$ as $T\to\infty$. ■

Lemma B.0.12. *Under Assumptions 3.1.1 and 3.1.2 and if H_0 is true, we have for every $\ell\geq 3$ that*

$$N^\ell M^{\ell/2}h^{\ell/2}cum_\ell(Q_{0,T})=o(1)$$

Proof: Let $\Pi=(-\pi,\pi]$ and $\mu=(\mu_1,\ldots,\mu_\ell)$. We then have

$$\begin{aligned}
& N^\ell M^{\ell/2}b^{\ell/2}cum_\ell(Q_{0,T}) \\
=\ & N^{-\ell}M^{-\ell/2}b^{\ell/2}\sum_{m_1,\ldots,m_l=1}^{M}\sum_{j_{1,1},\ldots,j_{1,\ell}=-J_N}^{J_N}\sum_{j_{2,1},\ldots,j_{2,\ell}=-J_N}^{J_N}\int_{\Pi^l}\prod_{\nu=1}^{\ell}\frac{K_b(\mu_\nu-\lambda_{j_{1,\nu}})K_b(\mu_\nu-\lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})f_{\vartheta_0}(u_{m_\nu},\lambda_{j_{2,\nu}})} \\
& cum\{\prod_{k=1}^{2}\left(I_N(u_{m_1},\lambda_{j_{k,1}})-f_{\vartheta_0}(u_{m_1},\lambda_{j_{k,1}})\right),\ldots,\prod_{k=1}^{2}\left(I_N(u_{m_\ell},\lambda_{j_{k,\ell}})-f_{\vartheta_0}(u_{m_\ell},\lambda_{j_{k,\ell}})\right) \\
& d\mu_1\ldots d\mu_\ell.
\end{aligned}$$

Using the product theorem for cumulants, we have that

$$\begin{aligned}
& cum\{\prod_{k=1}^{2}\left(I_N(u_{m_1},\lambda_{j_{k,1}})-f_{\vartheta_0}(u_{m_1},\lambda_{j_{k,1}})\right),\ldots,\prod_{k=1}^{2}\left(I_N(u_{m_\ell},\lambda_{j_{k,\ell}})-f_{\vartheta_0}(u_{m_\ell},\lambda_{j_{k,\ell}})\right) \\
=\ & \sum_{i.p.}\prod_{s=1}^{n}cum\{\left(I_N(u_{m_p},\lambda_{j_{q,p}})-f_{\vartheta_0}(u_{m_p},\lambda_{j_{q,p}})\right),(p,q)\in P_s\}
\end{aligned}$$

where the sum is over all indecomposable partitions $\{P_1,\ldots,P_n\}$ of the table

$$\begin{array}{cc}
(1,1) & (1,2) \\
\vdots & \vdots \\
(\ell,1) & (\ell,2).
\end{array}$$

We consider the sum $\sum_{i.p.1}$ over all partitions with $|P_i| > 1$. That is,

$$N^{-\ell} H_{2,N}^{-2\ell}(0) M^{-\ell/2} b^{\ell/2} \sum_{i.p.1} \sum_{m_1,\ldots,m_l=1}^{M} \sum_{j_{1,1},\ldots,j_{1,\ell}=-J_N}^{J_N} \sum_{j_{2,1},\ldots,j_{2,\ell}=-J_N}^{J_N}$$
$$\int_{\Pi^l} \prod_{\nu=1}^{\ell} \frac{K_b(\mu_\nu - \lambda_{j_{1,\nu}}) K_b(\mu_\nu - \lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu}, \lambda_{j_{2,\nu}}) f_{\vartheta_0}(u_{m_\nu}, \lambda_{j_{2,\nu}})} \prod_{s=1}^{n} cum\{d_N(u_{m_p}, \lambda_{j_{q,p}}) d_N(u_{m_p}, -\lambda_{j_{q,p}}), (p,q) \in P_s\}$$
$$d\mu_1 \ldots d\mu_\ell \tag{B.0.10}$$

Using again the product theorem of comulants, we have to sum over all indecomposable partitions $\{Q_{s,1}, \ldots, Q_{s,m}\}$ of the table

$$\begin{array}{cc} a_{p_{s_1},q_{s_1}} & b_{p_{s_1},q_{s_1}} \\ \vdots & \vdots \\ a_{p_{s_{|P_s|}},q_{s_{|P_s|}}} & b_{p_{s_{|P_s|}},q_{s_{|P_s|}}} \end{array}$$

for all sets $P_s = \{(p_{s_1}, q_{s_1}), \ldots, (p_{s_{|P_s|}}, q_{s_{|P_s|}})\}$. Note that $a_{p_{s_r},q_{s_r}}$ and $b_{p_{s_r},q_{s_r}}$ stand for the position of $d_N(u_{m_p^{(r)}}, \lambda_{j_{q,p}^{(r)}})$ and $d_N(u_{m_p^{(r)}}, -\lambda_{j_{q,p}^{(r)}})$ respectively where (r) denotes the position of $d_N(u_{m_p^{(r)}}, -\lambda_{j_{q,p}^{(r)}})$ in a fixed order. For simplicity we use the notation $a_{p_{s_r},q_{s_r}} := a_{s,r}$ and $b_{p_{s_r},q_{s_r}} := b_{s,r}$. Furthermore, if $Q_{s,i} = \{c_{s,1}, \ldots, c_{s,k}\}$ we set $\overline{Q}_{s,i} = \{c_{s,1}, \ldots, c_{s,k-1}\}$, $\beta_{\overline{Q}_{s,i}} := (\beta_{c_{s,1}}, \ldots, \beta_{c_{s,k-1}})$, $\beta_{c_{s,k}} = -\sum_{j=1}^{k-1} \beta_{c_{s,j}}$ and $\beta^{(s)} := (\beta_{\overline{Q}_{s,1}}, \ldots, \beta_{\overline{Q}_{s,m}})$. We then get that (B.0.10) is equal to

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \sum_{i.p.1} \sum_{i.p.^*} \sum_{m_1,\ldots,m_l=1}^{M} \sum_{j_{1,1},\ldots,j_{1,\ell}} \sum_{j_{2,1},\ldots,j_{2,\ell}} \int_{\Pi^l} \prod_{\nu=1}^{\ell} \frac{K_b(\mu_\nu - \lambda_{j_{1,\nu}}) K_b(\mu_\nu - \lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu}, \lambda_{j_{2,\nu}}) f_{\vartheta_0}(u_{m_\nu}, \lambda_{j_{2,\nu}})}$$
$$\times \prod_{s=1}^{n} \int_{\Pi^{2|P_s|-k}} \Bigg\{ \prod_{\substack{r=1 \\ (p,q)\in P_s}}^{|P_s|} H_N(A^0_{t_{m_p}^{(r)}-N/2+1+\cdot,T}(\beta_{a_{s,r}}) h(\frac{\cdot}{N}), \lambda_{j_{q,p}^{(r)}} - \beta_{a_{s,r}})$$
$$\times H_N(A^0_{t_{m_p}^{(r)}-N/2+1+\cdot,T}(\beta_{b_{s,r}}) h(\frac{\cdot}{N}), -\lambda_{j_{q,p}^{(r)}} - \beta_{b_{s,r}}) \Bigg\} \Bigg\{ \prod_{r=1}^{m} g_{|Q_{s,r}|}(\beta_{\overline{|Q_{s,r}|}}) \Bigg\}$$
$$\times \exp\left\{ i \sum_{r=1}^{|P_s|} t_{m_p}^{(r)} (\beta_{a_{s,r}} + \beta_{b_{s,r}}) \right\} d\beta^{(1)} \ldots d\beta^{(n)} d\mu_1 \ldots d\mu_\ell$$

Replace the terms $H_N(A^0_{t_{m_p}^{(r)}-N/2+1+\cdot,T}(\beta) h(\frac{\cdot}{N}), \lambda - \beta)$ by the terms $A(u_{m_p}^{(r)}, \beta) H_N(\lambda -$

β) to get that the above expression is equal to

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \sum_{i.p.1} \sum_{i.p.*} \sum_{m_1,\dots,m_l=1}^{M} \sum_{j_{1,1},\dots,j_{1,\ell}} \sum_{j_{2,1},\dots,j_{2,\ell}} \int_{\Pi^l} \prod_{\nu=1}^{\ell} \frac{K_b(\mu_\nu - \lambda_{j_{1,\nu}}) K_b(\mu_\nu - \lambda_{j_{2,\nu}})}{f_{\vartheta_0}(u_{m_\nu}, \lambda_{j_{2,\nu}}) f_{\vartheta_0}(u_{m_\nu}, \lambda_{j_{2,\nu}})}$$
$$\times \prod_{s=1}^{n} \int_{\Pi^{2|P_s|-k}} \left\{ \prod_{\substack{r=1 \\ (p,q)\in P_s}}^{|P_s|} A(u_{m_p}^{(r)}, \beta_{a_{s,r}}) H_N(\lambda_{j_{q,p}^{(r)}} - \beta_{a_{s,r}}) A(u_{m_p}^{(r)}, \beta_{b_{s,r}}) H_N(-\lambda_{j_{q,p}^{(r)}} - \beta_{b_{s,r}}) \right\}$$
$$\times \left\{ \prod_{r=1}^{m} g_{|Q_{s,r}|}(\beta_{\overline{|Q_{s,r}|}}) \right\} \exp\left\{ i \sum_{r=1}^{|P_s|} t_{m_p}^{(r)}(\beta_{a_{s,r}} + \beta_{b_{s,r}}) \right\} d\beta^{(1)} \dots d\beta^{(n)} d\mu_1 \dots d\mu_\ell$$
$$+E_T \tag{B.0.11}$$

where the error term E_T is bounded by

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \frac{N}{T} \sum_{i.p.1} \sum_{i.p.*} \sum_{j_{1,1},\dots,j_{1,\ell}=-J_N}^{J_N} \sum_{j_{2,1},\dots,j_{2,\ell}=-J_N}^{J_N} \int_{\Pi^l} \prod_{\nu=1}^{\ell} K_b(\mu_\nu - \lambda_{j_{1,\nu}})$$
$$\times K_b(\mu_\nu - \lambda_{j_{2,\nu}}) \prod_{s=1}^{n} \left\{ \prod_{\substack{r=1 \\ (p,q)\in P_s}}^{|P_s|} \int_{\Pi^{2|P_s|-m}} L_N(\lambda_{j_{q,p}^{(r)}} - \beta_{a_{s,r}}) L_N(-\lambda_{j_{q,p}^{(r)}} - \beta_{b_{s,r}}) \right.$$
$$\left. \times L_M(S(\beta_{a_{s,r}} + \beta_{b_{s,r}} + \beta_{a_{x,y}} + \beta_{b_{x,y}})) \right\} d\beta^{(1)} \dots d\beta^{(n)} d\mu_1 \dots d\mu_\ell$$

for some $(x,y) \in \{1,\dots,n\} \times \{1,\dots,|P_x|\}$ with $x \neq s$. Integration over all $\beta_{a_{s,r}}$ and $\beta_{b_{s,r}}$ gives that expression (B.0.11) is bounded by

$$\left(\frac{b}{N^2 H_{2,N}^4(0) M}\right)^{\ell/2} \frac{N}{T} \frac{M^{\ell-1} N}{S} N^{2\ell} \to 0,$$

which completes the proof. ■

Lemma B.0.13. *Under Assumptions 3.1.1 and 3.1.2 and if H_0 is true, we have that*

$$N\sqrt{M b}(Q_T - \mu_T) = N\sqrt{M b}(Q_{0,T} - \mu_t) + o_p(1)$$

Proof:

$$\begin{aligned} Q_T &= Q_{0,T} + \frac{1}{MN^2} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \left\{ \sum_{j=-J_N}^{J_N} K_b(\lambda - \lambda_j) \left(\frac{I_N(u_i,\lambda_j)}{f_{\hat\vartheta}(u_i,\lambda_j)} - \frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} \right) \right\}^2 d\lambda \\ &+ \frac{2b^{1/2}}{NM^{1/2}} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \sum_{j,s} K_b(\lambda - \lambda_j) K_b(\lambda - \lambda_s) \left(\frac{I_N(u_i,\lambda_j)}{f_{\hat\vartheta}(u_i,\lambda_j)} - \frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} \right) \\ &\times \left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)} - 1 \right) d\lambda \\ &= Q_{0,T} + Y_{1,T} + Y_{2,T} \end{aligned}$$

with an obvious notation for $Y_{1,T}$ and $Y_{2,T}$. The term $Y_{1,T}$ is bounded by

$$\begin{aligned} |Y_{1,T}| &\leq \sup_{u,j} \left(\frac{f_{\vartheta_0}(u,\lambda_j) - f_{\hat{\vartheta}}(u,\lambda_j)}{f_{\hat{\vartheta}}(u,\lambda_j)} \right)^2 \frac{b^{1/2}}{M^{1/2}N} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \left\{ \sum_j K_b(\lambda-\lambda_j) \frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} \right\}^2 d\lambda \\ &= O_p\left(\frac{b^{1/2}}{M^{1/2}} \right). \end{aligned}$$

For the second term we have

$$\begin{aligned} Y_{2,T} \;=\; & \frac{b^{1/2}}{NM^{1/2}} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \sum_{j,s} K_b(\lambda-\lambda_j) K_b(\lambda-\lambda_s) \left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)} - 1 \right) \\ & \times \left(\frac{f_{\hat{\vartheta}}(u_i,\lambda_j) - f_{\vartheta_0}(u_i,\lambda_j)}{f_{\hat{\vartheta}}(u_i,\lambda_j)} \right) \left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)} - 1 \right) d\lambda \\ & + \frac{b^{1/2}}{NM^{1/2}} \sum_{i=1}^{M} \int_{-\pi}^{\pi} \sum_{j,s} K_b(\lambda-\lambda_j) K_b(\lambda-\lambda_s) \\ & \times \left(\frac{f_{\hat{\vartheta}}(u_i,\lambda_j) - f_{\vartheta_0}(u_i,\lambda_j)}{f_{\hat{\vartheta}}(u_i,\lambda_j)} \right) \left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)} - 1 \right) d\lambda \\ =\; & W_{1,T} + W_{2,T} \end{aligned}$$

with an obvious notation for $W_{1,T}$ and $W_{2,T}$.

By a standard Taylor series argument, for fixed u, we have that for $\hat{\vartheta}(u) = (\hat{\vartheta}_1(u), \hat{\vartheta}_2(u), \ldots, \hat{\vartheta}_p(u))'$ and $\vartheta_0(u) = (\vartheta_1(u), \vartheta_2(u), \ldots, \vartheta_p(u))'$, $\tilde{\vartheta}_0(u) = (\tilde{\vartheta}_1(u), \tilde{\vartheta}_2(u), \ldots, \tilde{\vartheta}_p(u))'$ with $||\tilde{\vartheta}(u) - \vartheta_0(u)|| \leq ||\hat{\vartheta}(u) - \vartheta_0(u)||$ exists such that

$$\begin{aligned} & \frac{f_{\hat{\vartheta}}(u,\lambda) - f_{\vartheta_0}(u,\lambda)}{f_{\hat{\vartheta}}(u,\lambda)} = \\ & O_p(1) \left\{ \sum_{m=1}^{p} (\hat{\vartheta}_m(u) - \vartheta_m(u)) f_T^{(1)}(\vartheta_m, \lambda) \right. \\ & \left. + \frac{1}{2} \sum_{m=1}^{p} \sum_{l=1}^{p} (\hat{\vartheta}_m(u) - \vartheta_m(u))(\hat{\vartheta}_l(u) - \vartheta_l(u)) f_T^{(2)}(\tilde{\vartheta}_m, \tilde{\vartheta}_l, \lambda) \right\} \end{aligned}$$

where $f_T^{(1)}(\vartheta_m, \lambda)$ and $f_T^{(2)}(\tilde{\vartheta}_m, \tilde{\vartheta}_l, \lambda)$ denote the first and second second partial derivatives of f with respect to ϑ_m and ϑ_l and ϑ_m respectively, and evaluated at ϑ_m and $(\tilde{\vartheta}_m, \tilde{\vartheta}_l)$. Notice that the $O_p(1)$ term appear in (B.0.12) is due to the fact that $|1/f_{\hat{\vartheta}}(u,\lambda)| = O_p(1)$. Using (B.0.12) we get

$$
\begin{aligned}
W_{1,T} &= O_p(1)\frac{b^{1/2}}{NM^{1/2}}\sum_{i=1}^{M}\sum_{m=1}^{p}(\hat{\vartheta}_m(u_i)-\vartheta_m(u_i))\int_{-\pi}^{\pi}\sum_{j,s}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)f^{(1)}(\vartheta_m,\lambda_j) \\
&\quad \times\left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)}-1\right)\left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)}-1\right)d\lambda \\
&\quad +O_p(1)\frac{b^{1/2}}{NM^{1/2}}\sum_{i=1}^{M}\sum_{l=1}^{p}\sum_{m=1}^{p}(\hat{\vartheta}_l(u_i)-\vartheta_l(u_i))(\hat{\vartheta}_m(u_i)-\vartheta_m(u_i))\int_{-\pi}^{\pi}\sum_{j,s}K_b(\lambda-\lambda_j) \\
&\quad \times K_b(\lambda-\lambda_s)f^{(2)}(\tilde{\vartheta}_m,\tilde{\vartheta}_l,\lambda_j)\left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)}-1\right)\left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)}-1\right)d\lambda \\
&= O_p(N^{-1/2})+O_p(b^{1/2}).
\end{aligned}
$$

The $O_p(N^{-1/2})$ term is due to the fact that $\sup_u|\hat{\vartheta}_m(u)-\vartheta_m(u)| = O_p(N^{-1/2})$ and that

$$
\begin{aligned}
&\frac{b^{1/2}}{NM^{1/2}}\sum_{i=1}^{M}\int_{-\pi}^{\pi}\sum_{j,s}K_b(\lambda-\lambda_j)K_b(\lambda-\lambda_s)f^{(1)}(\vartheta_m,\lambda_j)\left(\frac{I_N(u_i,\lambda_j)}{f_{\vartheta_0}(u_i,\lambda_j)}-1\right) \\
&\times\left(\frac{I_N(u_i,\lambda_s)}{f_{\vartheta_0}(u_i,\lambda_s)}-1\right)d\lambda
\end{aligned}
$$

can be handled as $Q_{0,T}$. Similarly we can show that $W_{2,T} = o_p(1)$ which completes the proof. ■

Proof of Theorem 3.1.1: By Lemma $B.0.10$, $B.0.11$ and $B.0.12$ we have that the cumulants of all orders of $Q_{0,T}$ converge to the corresponding cumulants of the limiting Gaussian distribution. The assertion of the theorem follows then by Lemma $B.0.13$. ■

Proof of Theorem 3.2.1 : Follow the same steps as in the proof of Lemma B.0.13 substituting $\overline{\vartheta}$ for ϑ_0 in $f_{\vartheta_0}(u_i,\lambda_j)$ and using the property that under the alternative hypothesis, $\hat{\vartheta}$ is a $\sqrt{N}$-consistent estimator of $\overline{\vartheta}$. ■

Proof of Theorem 3.2.2: First notice that $X^+_{t,T}$ is locally stationary with transfer function $\hat{A}^0$, that is,

$$X^+_{t,T} = \int_{-\pi}^{\pi}\hat{A}^0_{t,T}(\lambda)e^{i\lambda t}d\xi^+(\lambda) \tag{B.0.12}$$

where

(i) $\xi^+(\lambda)$ is a Gaussian stochastic process on $(-\pi,\pi]$ and

$$cov^+\{\xi^+(\lambda_k),\xi^+(\lambda_j)\} = \delta(k,j)d\lambda_k \tag{B.0.13}$$

(ii) There exists a constant K and a function $\hat{A}(u,\lambda)$ on $[0,1]\times(-\pi,\pi)$ such that for all T,

$$\sup_{t,\lambda} |\hat{A}^0_{t,T} - \hat{A}(t/T,\lambda)| \leq K/T$$

(iii) Furthermore,

$$f_{\hat{\vartheta}(u)}(u,\lambda) = \frac{1}{2\pi}|\hat{A}(u,\lambda)|^2 \tag{B.0.14}$$

where $\hat{\vartheta}(u) = (\hat{\beta}_1(u),\ldots,\hat{\beta}_p(u),\hat{\sigma}^2(u))$ and the function $1/f(u,\lambda;\hat{\vartheta})$ is bounded in probability.

Now, following the same steps as in the proof of Lemma $B.0.10$, $B.0.11$ and $B.0.12$, we get that the limits of all cumulants of the bootstrap test statistic $N\sqrt{Mb}(Q_T^+ - \mu_T)$ converge to the cumulants of the limiting Gaussian distribution given in Theorem 3.2.2. $\blacksquare$

Bibliography

[1] T.W. Anderson, *The Statistical Analysis of Time Series.* John Wiley and Sons, New York, 1971.

[2] K. I. Beltrão and P. Bloolfield, *Determining the bandwidth of a kernel spectrum estimate.* Journal of Time Series Analysis, **8**, 21–38, 1987.

[3] K. N. Berk, *Consistent autoregressive spectral estimates.* The Annals of Statistics, **2** , 489–502, 1974.

[4] D. R. Brillinger, *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco, 1981.

[5] P. Brockwell and R. A. Davis, *Time Series: Theory and Methods, 2nd Ed.*, Springer, New York, 1991.

[6] C. Chang and P. Morettin, *Estimation of time varying linear systems.* Statist. Inference Stoch. Proc., **2**, 253–285, 1999.

[7] R. Dahlhaus, *On the Kullback-Leibler information diverenge of locally stationary processes.* Stochastic Processes and their Applications, **62**, 139–168, 1996a.

[8] R. Dahlhaus, *Asymptotic statistical inference for nonstationary process with evolutionary spectra.* In: Athens Conference on Applied Probability and Time Series Analysis, Volume II (Ed. by P.M Robinson and M. Rosenblatt), Springer Verlag, New York, 1996b.

[9] R. Dahlhaus, *Fitting time series models to nonstationary processes.* The Annals of Statistics, **25**, 1–37, 1997.

[10] R. Dahlhaus, *A likelihood approximation for locally stationary processes.* The Annals of Statistics, **28**, 1762–1794, 2000.

[11] R. Dahlhaus, *Curve estimation for locally stationary time series models.* In: Recent Advances and Trends in Nonparametric Statistics (ed. by M. G. Akritas and D. N. Politis) Elservier Science B.V., 451–466, 2003.

[12] R. Dahlhaus and D. Janas, *A frequency domain bootstrap for ratio statistics in time series analysis.* The Annals of Statistics, **24**, 1934–1963, 1996.

[13] R. Dahlhaus and L. Giraitis, *On optimal segment length for parameter estimators for locally stationary time series.* Journal of Time Series Analysis, **19**, 629–655, 1998.

[14] R. Dahlhaus, M.H. Neumann and R. von Sachs, *Nonlinear wavelet estimation of time-varying autoregressive processes.* Bernoulli, **5**, 873–906, 1999

[15] R. Dahlhaus M.H Neumann, *Locally adaptive fitting of semiparametric models to nonstationary time series.* Stochastic Processes and their Applications, **91**, 277–308, 2001.

[16] R. Dahlhaus and W. Polonik, *Nonparametric maximum likelihood estimation and empirical spectral processes for locally stationary processes.* Preprint. Universitat Heidelberg, 2003.

[17] R. Dahlhaus and S. Subba Rao, *Statistical inference of time varying ARCH processes.* The Annals of Statistics, **34**, 1075–1114, 2006.

[18] A. Dowla, E. Paparoditis and D. N. Politis, *Locally stationary processes and the local block bootstrap.* In Recent Advances and Trends in Non Parametric Statistics (ed. by M. G. Akritas and D. N. Politis) Elservier Science B.V., 437–444, 2003.

[19] J. Franke and W. Härdle, *On bootstrapping kernel spectral estimates.* The Annals of Statistics, **20**, 121–145, 1992.

[20] P. Fryzlewicz, S. Van Bellegem and R. von Sachs, *Forecasting non-stationary time series by wavelet process modelling*. Ann. Inst. Statist. Math. **55**, 737-764, 2003.

[21] U. Grenander and G. Szegö,*Toeplitz Forms and Their Applications*, Univ. of California Press, Berkeley, 1958.

[22] C. M. Hurvich and S. L. Zeger, *Frequency domain bootstrap methods for time series.* Technical Report 87–115, Graduate School of Business Administration, New York Univ, 1987.

[23] J. -P. Kreiss, *Residual and wild bootstrap for infinite order autoregression.* Unpublished manuscript, 1999.

[24] J. -P. Kreiss and E. Paparoditis, *Autoregressive-aided periodogram bootstrap for time series*. The Annals of Statistics, **31**, 1923–1955, 2003.

[25] S. G. Mallat, G. Papanicolaou, Z. and Zhang,*Adaptive covariance estimation of locally stationary processes*. The Annals of Statistics, **26**, 1–47, 1998.

[26] G. Mélard and A. Herteleer-De Schutter, *Contributions to evolutionary spectral theory.* Journal of Time Series Analysis, **10**, 41–63, 1989.

[27] G. P. Nayson, R. von Sachs, and G. Kroisandt, *Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum.* J. R. Stat. Soc. Ser. B Stat. Methol., **62**, 271–292, 2000.

[28] M. Neumann, and R. von Sachs, *Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra.* The Annals of Statistics, **25**, 38–76, 1997.

[29] A. Nordgaard, *Resampling stochastic processes using a bootstrap approach.* In: Lecture Notes in Econom. and Math. Systems **376**, 181–185, 1992.

[30] H. Ombao, J. Raz, R. von Sachs, and W. Guo *The SLEX model of a non-stationary random process. Nonparametric approach to time series analysis.* Ann. Inst. Statist. Math. **54**, 171–200, 2002.

[31] H. Ombao, R. von Sachs and W. Guo, *SLEX analysis of multivariate nonstationary time series.* J. Amer. Stat. Assoc. **100**, 519–531, 2005.

[32] E. Paparoditis, *Frequency domain bootstrap for time series.* In: Empirical Process Techniques for Dependent Data (Ed. by H. Dehling, T. Mikosch and M. Sørensen), Birkhäuser, Boston, 365–381, 2002.

[33] E. Paparoditis, and D. N. Politis, *The local bootstrap for the periodogram.* Journal of Time Series Analysis, **20**, 193–222, 1999.

[34] E. Paparoditis and D. N. Politis, *Local block bootstrap.*C.R. Acad. Sci. Paris, Ser. I, **335**, 959–962, 2002.

[35] M. B. Priestley, *Evolutionary spectra and non-stationary processes.* J. R. Stat. Soc. Ser. B, **62**, 204–237, 1965.

[36] M. B. Priestley, *Spectral Analysis and Time Series.* Academic Press, New York, 1981.

[37] S. Sakiyama and M. Taniguchi, *Testing composite hypotheses for locally stationary processes.* Journal of Time Series Analysis, **24**, 483–504, 2003.

[38] M. Sergides and E. Paparoditis, *Bootstrapping the local periodogram of locally stationary processes.* Journal of Time Series Analysis, 2007. to appear.

[39] R. Shumway and D. Stoffer *Time Series and its Applications*, Springer, New York, 2000.

[40] S. Subba Rao, *On some nonstationarity, nonlinear random processes and their stationary approximations.* Preprint, 2004.

[41] J. Theiler, L. S. Paul and D. M. Rubin, *Detecting nonlinearity in data with long coherence times.* In: Times Series Prediction (Ed. by A. Weigend and N. Gershenfeld), Addison-Wesley, Reading, MA, 1994.

[42] S. Van Bellegem and R. Dahlhaus, *Semiparametric estimation by model selection for locally stationary processes.* J. R. Stat. Soc. Ser. B Stat. Methol., **68**, 721–746, 2006.