A NOVEL APPROACH TO THE ANALYSIS OF NONLINEAR TIME SERIES WITH APPLICATIONS TO FINANCIAL DATA

A Dissertation

by

JUN BUM LEE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Major Subject: Statistics

A NOVEL APPROACH TO THE ANALYSIS OF NONLINEAR TIME SERIES WITH APPLICATIONS TO FINANCIAL DATA

A Dissertation

by

JUN BUM LEE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHLOSOPHY

Approved by:

Chair of Committee,	Suhasini Subba Rao
Committee Members,	Mohsen Pourahmadi
	Soumendra N. Lahiri
	Hwagyun Kim
Head of Department,	Simon J. Sheather

May 2012

Major Subject: Statistics

ABSTRACT

A Novel Approach to the Analysis of Nonlinear Time Series with Applications to Financial Data. (May 2012)

Jun Bum Lee, B.S., Seoul National University;

M.S., Texas A&M University

Chair of Advisory Committee: Dr. Suhasini Subba Rao

The spectral analysis method is an important tool in time series analysis and the spectral density plays a crucial role on the spectral analysis. However, one of limitations of the spectral density is that the spectral density reflects only the covariance structure among several dependence measures in the time series data. To overcome this restriction, we define two spectral densities, the quantile spectral density and the association spectral density. The quantile spectral density can model the pairwise dependence structure and provide identification of nonlinear time series and the association spectral density allows detecting periodicities on different parts of the domain of the time series. We propose the estimators for the quantile spectral density and the association spectral density and derive their sampling properties including asymptotic normality. Furthermore, we use the quantile spectral density to develop a goodness-of-fit tests for time series and explain how this test can be used for comparing the sequential dependence structure of two time series. The asymptotic sampling properties of the test statistic is derived under the null and alternative hypothesis, and a bootstrap procedure is suggested to obtain finite sample approximation. The method is illustrated with simulations and some real data examples. Besides the exploration of the new spectral densities, we consider general quadratic forms of α-mixing time series and derive asymptotic normality of these forms under the relatively weak assumptions.

To my parents

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Suhasini Subba Rao, for her valuable teaching and excellent mentorship through my studies. I always have been amazed by the patience and understanding she has displayed with me over the past few years. Without her guidance and persistent help, this dissertation would have been impossible.

I wish to thank my committee members, Dr. Mohsen Pourahmadi, Dr. Soumendra N. Lahiri, and Dr. Hwagyun Kim, for their willingness to serve on my committee and their helpful comments and suggestions.

I am indebted to my friends, Soojin, Jaehong and Chulhwan. They helped me get thorugh the final stage of this study.

Lastly, and most importantly, I wish to thank my family for all their love and encouragement.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

One objective of time series analysis is to capture the dependence structure of data, and there are two approaches to this. One approach is time domain analysis and one another is frequency domain analysis. Another dichotomy can be applied to time series model itself, linear time series and nonlinear time series.

Because of the easiness of their usage and interpretation, the linear time series model has been more popular than the nonlinear models and most widely used linear model framework is autoregressive moving average(ARMA) model after Box and Jenkins (1970). Due to its nature, the dependence structure of the linear time series is often confined to linear order and autocovariance function(ACF) plays a important role in time domain analysis approach. It describes dependence structure of the linear time series fairly well, and if the innovation in the linear model follows Gaussian distribution, ACF solely can capture the whole dependence structure. The counter part of ACF in the frequency domain is spectral density which is Fourier transformation of autocovariance function. It can be used for detecting periodicities and estimating parameters in linear model. Despite the clear advantage of this simplicity, there are several disadvantages in using the autocovariance and spectral density as tools for describing dependence structure. The autocovariance function only measures the average linear interaction between elements of a time series, so it often fails to provide useful when nature of dependence structure is beyond linear as in most nonlinear time

---

This dissertation follows the style of *Biometrics*.

series model.

Nonlinear time series model naturally arises from observation that there are some features which can not be captured by the linear model. For example, one of stylized facts which are common to a wide set of financial data is absence of autocorrelations and slow decay of autocorrelation in absolute returns. The one set of time series model satisfying this characteristic is ARCH and GARCH model proposed by Engle (1982) and Bollerslev (1986), and they have been widely used in volatility modeling. However, with any statistical tools based on autocovariance function and the classical spectral density, we can not distinguish these models from white noise and this might lead us to the false conclusion of independence in data. Also, heavy-tailedness, one of other stylized facts of the financial time series, often invalidates the usage of classical spectral density, hence many applications of it are based on the finite moment assumption.

Recently several methods have been proposed to overcome these limitations of the classical spectral density. Hong (1999) introduces the generalized spectral density, which is the Fourier transform of the empirical characteristic function of a time series. Li (2008) proposes Laplace spectrum and Laplace periodogram to obtain more robustness in spectral density estimators. His idea is based on that the usual periodogram is the least square coefficient estimator in the regression between time series data and harmonic functions and it suffers from outliers due to the least square(LS) method. To alleviate this problem, the least absolute deviation(LAD) method is used and the LAD estimator in the regression is defined as Laplace periodogram. He shows how the Laplace spectrum is related to spectral density of $\{I(X_t \leq 0)\}_t$ called zero-crossing spectrum and it could be used for detecting the periodicity in $\{I(X_t \leq 0)\}_t$. Hagemann (2011) widens this approach by considering the spectral density of $\{I(X_t \leq q_u)\}$ where q_u is uth-quantile of $\{X_t\}$ and Dette, Hallin, Kley, and Volgushav (2011) also investigates the cross-spectral density of $\{I(X_t \le q_u), I(X_t \le q_v)\}$. There is a similarity in these works considering spectral densities of certain transforms, Hong (1999) for the empirical characteristic function transform and Hagemann (2011) and Dette et al. (2011) for the empirical distribution transform. The use of empirical distribution function has the advantage of the empirical characteristic function for its easy interpretation. In this work, we also introduce a spectral density of the empirical distribution called the quantile spectral density. In contrast to L_1 estimating method in Dette et al. (2011), we propose the L_2 estimator with an analytic form, thus can easily be used in both goodness-of-fit test.

Goodness-of-fit tests are usually done by checking the assumptions imposed in a statistical model. In many time series models, the independent innovation is commonly assumed, and this assumption is verified based on the sample autocorrelation of the residuals from the fitted model. Box and Pierce (1970) proposes this method in ARMA model and its modification was done by Ljung and Box (1978), and Milhøj (1981), Velilla (1994) and Anderson (1997) provide the frequency domain counterparts of these methods. Hong (1996) shows that a test could be more powerful by giving different weight on sample autocorrelations at different lags.

Since these methods only focus on the autocorrelation, they often fail to detect dependence in general form. The more general form of dependence measure is serial dependence, and it dates back to Hoeffding (1948) whose method is used for testing independence in two random variables. This method is based on the fact that if X and Y are independent, then $P(X \le x, Y \le y) - P(X \le x)P(Y \le y) = 0$ for any x, y. For bivariate random sample $\{(X_t, Y_t)\}_{t=1}^T$ from (X, Y), it measures the difference between bivariate empirical distribution and the product of two marginal empirical distributions. Blum, Kiefer, and Rosenblatt (1961) extends this concept for more than 2 random variables case. Applying this measurement for time series data was addressed by Skaug and Tjøstheim (1993) and Hong (1998) in the time domain and by Hong (2000) in frequency domain. The dependence measure they used is $P(X_t \leq x, X_{t+r} \leq y) - P(X_t \leq x)P(X_{t+r} \leq y)$ for the stationary time series $\{X_t\}$.

While Skaug and Tjøstheim (1993) and Hong (1998) use the empirical distribution function to test sequential dependence, Hallin and Puri (1992) propose a method based on ranks and Pinkse (1998) uses the empirical characteristic function for it. Hong (1999) takes this empirical characteristic function approach further defining the generalized spectral density, which is the Fourier transform of the characteristic function of pair-wise dependent data. The goodness-of-fit tests based on the generalized spectral densities of the estimated residuals are presented in Hong (1999) and Hong and Lee (2003). However, sometimes the residuals cannot be or are not easy to estimate. For example, it is possible to estimate the residuals of an ARCH($X_t = Z_t\sigma_t$), possible but difficult with a GARCH and usually impossible for many models of the type $X_t = g(X_{t-1}, \epsilon_t)$. To circumvent this difficulty, we propose a new goodness-of-fit test based on the quantile spectral density. It directly measures the difference of serial dependence structures between the time series data and the fitted model.

In Chapter II, we present the quantile spectral density which captures serial dependence in time series data without requiring linearity and certain moment assumption. We propose the estimator for it and derive its sampling properties including asymptotic normality. A goodness-of-fit test using the quantile spectral density is developed and some simulation results and real data example are given.

In Chapter III, we introduce the association spectral density and its estimator. The asymptotic properites of the estimator are derived.

Chapter IV contains the asymptotic normality of general quadratic forms of nonstationary, α-mixing time series, which we encounter in Chapter II and III.

CHAPTER II

THE QUANTILE SPECTRAL DENSITY AND COMPARISON BASED TESTS FOR NONLINEAR TIME SERIES

1. Introduction

The analysis of most time series is based on a set of assumptions, which in practice need to be tested. This is usually done through a goodness of fit test. The majority of goodness of fit tests for time series are based on fitting the conjectured model to the data, estimating the residuals of the model and testing for lack of correlation, normally with a Ljung-Box type test (see for example, Anderson (1993), Hong (1996), Chen and Deo (2004), and Hallin and Puri (1992) for a robust tests based on ranks). If one restricts the class of models to just linear time series models, then such tests can correctly identify the model. However, problems can arise, if one widens the class of models and allow for nonlinear time series. For example, if the time series were to satisfy an ARCH process, then it will be uncorrelated, but it is not independent. Moreover, the squares will satisfy an autoregressive representation, with errors which are martingale differences. Therefore, correlation based test for nonlinear time series models may not identify the model.

Neumann and Paparoditis (2008) propose a goodness of fit test for Markov time series models based on the one step ahead transition distribution. But this test is specifically for Markov models. An alternative approach is to generalise the notion of correlation to measuring the general dependence between pairs of random variables in a time series. This notion is usually called serial dependence, and dates back to Hoeffding (1948). Skaug and Tjøstheim (1993) and Hong (2000) use this definition to test for serial independence of a time series. Hong (1998) takes these notions further, and as far as we are aware is the first paper to generalise the spectral density to sequential dependence. He does this by defining the generalised spectral density, which is the Fourier transform of the characteristic function of pair-wise dependent data. He uses this device in Hong (1998) and Hong and Lee (2003) to test for goodness of fit of a time series model, mainly through the analysis of the estimated residuals. However, sometimes the residuals cannot be or are not easy to estimate. For example, it is possible to estimate the residuals of an ARCH ($X_t = Z_t\sigma_t$), possible but difficult with a GARCH and usually impossible for many models of the type $X_t = g(X_{t-1}, \varepsilon_t)$.

In this chapter, we use the notion of serial dependence to test for goodness of fit, but without estimating the residuals. Instead our test is based on comparisons. In Section 2.1 we motivate our test by considering the Microsoft daily log return data and comparing it with the GARCH$(1,1)$ model, which is one of the standard models fitted to such data sets. We show that though the GARCH model seems to model well some of the stylised facts of this data, ie. the uncorrelatedness, and positive correlation in the absolute and squares, if one made a deeper analysis and compared the correlation of other transformations such as $\mathrm{cov}(I(X_t \leq x), I(X_{t+r} \leq y))$ (where I denotes the indicator function), there is large difference between the data and GARCH model. This motivates us to define the *quantile* autocovariance function and the *quantile* spectral density. The quantile spectral density can be considered as a measure of serial dependence of a time series. In Sections 2.2 and 2.3 we propose a method for estimating the quantile spectral density, and use the quantile spectral density as the basis of a test based on the quadratic distance which compares the quantile spectral density estimator with the spectral density estimator under the null hypothesis. The asymptotic sampling properties of the quantile spectral density estimator are derived in Section 3.1. Recently there have been several articles defining and estimating the spectral density of sequential dependence. Li (2008), Hagemann (2011) and Dette et al. (2011) define spectral density functions similar to the quantile spectral density, however these authors, estimate the periodogram and the quantile spectral density using L_1 methods. In contrast, we use L_2 methods based on the usual definition of the periodogram, this is because it has an analytic form and can easily be used in a goodness of fit and other tests. It is interesting, and rather surprising, to note that the L_1 estimator proposed in Dette et al. (2011) and our estimator of the quantile spectral density have similar asymptotic properties. In Section 3.2 we derive the asymptotic sampling properties of the test statistic. The advantage of our approach is that it can easily be extended to test other quantities, for example with a small adaption it can be used to test for equality of serial dependence of two time series, this is considered in Section 4. In Section 5 we propose a bootstrap method for estimating the finite sampling distribution of the test statistic under the null. The proofs can be found in Section 7 and some technical details are given in the appendix.

2. The quantile spectral density and the test statistic

2.1. Motivation

To motivate our approach, we analyze the Microsoft daily log returns (MSFT) between March 1986 - June 2003, which we denote as $\{X_t\}$. One argument for fitting GARCH types models to financial data is their ability to model the so called 'stylised facts' seen in such data sets. We now demonstrate why this is the case for the MSFT (see Zivot (2009)). Using the maximim likelihood, the GARCH model which best fits the log differences of the MSFT is $X_t = \mu + \varepsilon_t$, $\epsilon_t = \sigma_t Z_t$, $\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + b\sigma_{t-1}^2$ ($\{Z_t\}$ are independent, identically distributed standard normal random variables), where $\mu = 1.56 \times 10^{-3}$, $a_0 = 1.03 \times 10^{-5}$, $a_1 = 0.06$ and $b = 0.925$. In Figure

1 we give the sample autocorrelation plots of $\{X_t\}$ and $\{|X_t|\}$, together with the autocorrelation plots of the corresponding GARCH$(1,1)$ model. Comparing the two plots, it appears that the GARCH$(1,1)$ captures the 'stylised facts' in the Microsoft data, such as the near zero autocorrelation of the observations and the persistant positive autocorrelations of the absolute and squares of the log returns. However, if we want to check the suitability of the GARCH model for modelling the general pair-wise dependence structure, that is the joint distribution of (X_s, X_t) for all s and t (often called *sequential dependence*), then we need to look beyond the covariance of $\{X_t\}$ and $\{|X_t|\}$. To make a more general comparison we transform the data into indicator variables $\{I(X_t \leq x)\}$ and check the correlation structure of the indicator variables over various x. For example, define the multivariate vector time series $\underline{Y}_t = (I(X_t \leq q_{0.1}), I(X_t < q_{0.5}), I(X_t \leq q_{0.9}))$, where q_α denotes the estimated α-percentile of X_t.

Plots of the cross-covariances of $\underline{Y}_t$ and the corresponding GARCH model (with Gaussian innovations) are given in Figure 2. In Figure 2, there are clear differences in the dependence structure of the data and the GARCH model. The 10th, 50th and 90th percentiles correspond to large negative, zero and large positive values of X_t (big negative change, no change and large positive changes in the returns). In order to do the analysis, we will use the following observations. By using that $\mathrm{cov}(I(X_0 \leq x), I(X_r \leq y)) = P(X_0 \leq x, X_r \leq y) - P(X_0 \leq x)P(X_r \leq y)$, for all $x, y \in \mathbb{R}$ we have

$$\begin{aligned} \mathrm{cov}(I(X_0 \leq x), I(X_r \leq y)) &= \mathrm{cov}(I(X_0 > x), I(X_r > y)) \\ &= -\mathrm{cov}(I(X_0 \leq x), I(X_r > y)). \end{aligned}$$

From Figure 2 we observe:

- The ACF of $I(X_t \leq q_{0.5})$ of the GARCH is zero. This is due to the symmetry

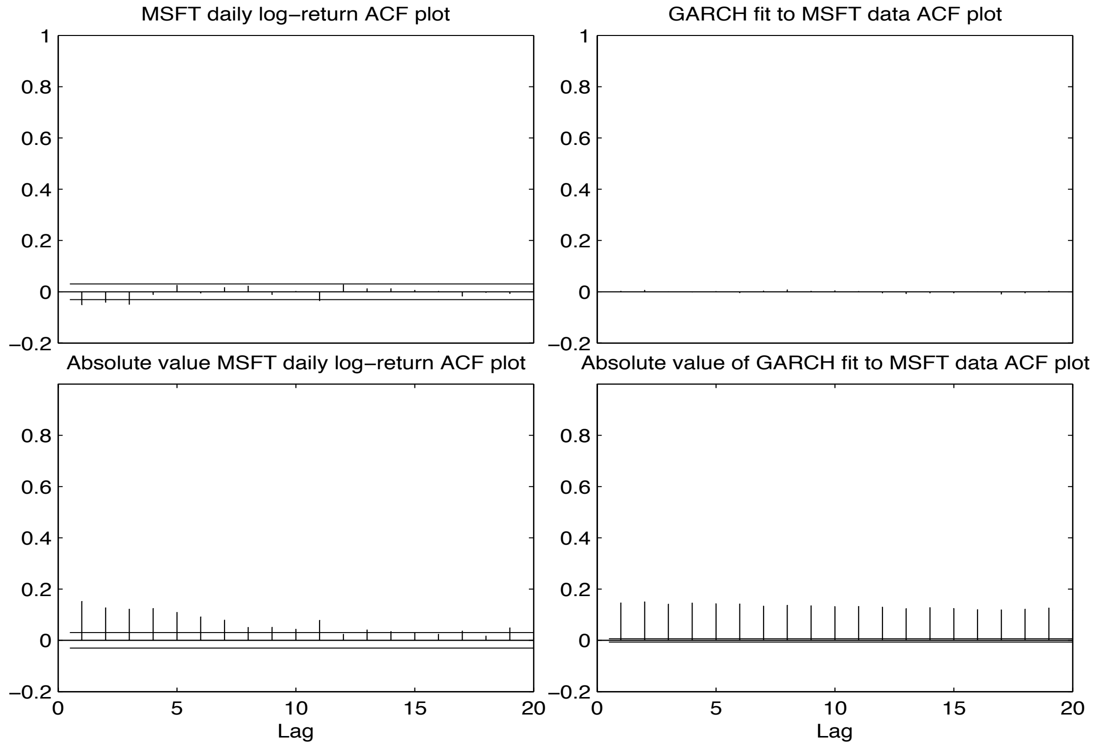

Fig. 1. The ACF plots of $\{X_t\}$ and $\{|X_t|\}$ of the MSFT and the corresponding GARCH model

of the GARCH process, given the event$X_0 \leq 0$, we have equal chance $X_r > 0$ and $X_r < 0$ (ie. $\text{cov}(I(X_0 \leq 0), I(X_r \leq 0)) = -\text{cov}(I(X_0 \leq 0), I(X_r > 0)))$. This means that $\text{cov}(I(X_0 \leq 0), I(X_r \leq 0)) = 0$. On the other hand, for the MSFT data we see that there is a clear positive correlation in the sample autocorrelation of $\{I(X_t < 0)\}$. One interpretation for the MSFT data, is that a decrease in consecutive values, is likely to lead to future decreases.

- The cross correlation of the GARCH of $I(X_t < q_{0.1})I(X_t < q_{0.9})$ is symmetric about zero, this means that $\text{cov}(I(X_0 < q_{0.1}), I(X_r < q_{0.9})) = \text{cov}(I(X_0 < q_{0.1}), I(X_{-r} < q_{0.9}))$. On the other hand, the corresponding sample cross-correlations of the MSFT is not symmetric. Thus the GARCH process is time

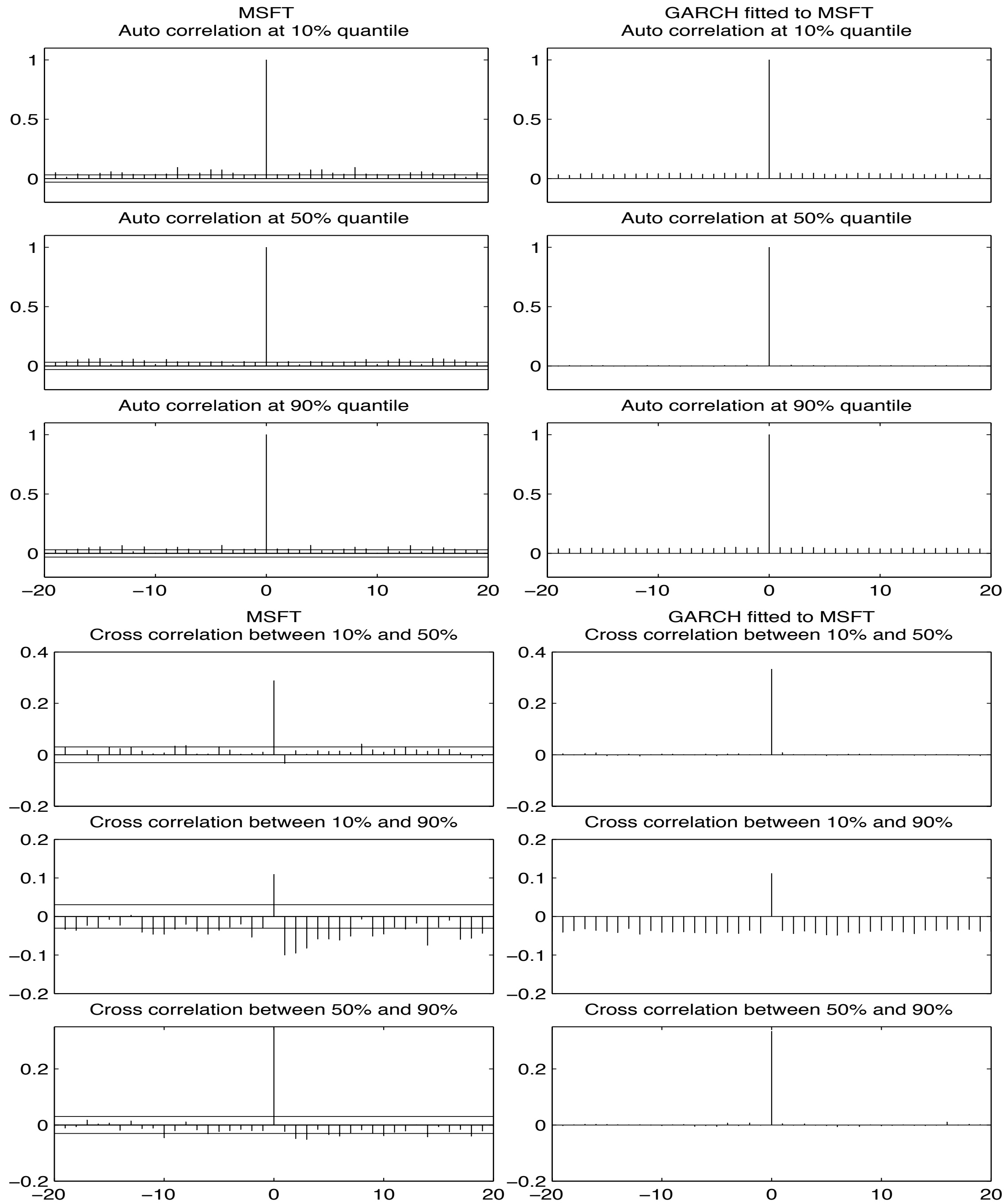

Fig. 2. The quantile covariance of the MSFT and the corresponding GARCH

reversible, whereas it appears that the MSFT data may not be.

The cross and autocovariances in Figure 2 are a graphical representation of the serial dependence structure of the time series. These plots suggest that for MSFT time series the GARCH model may not be the most appropriate model, especially if validity is based on modelling the serial dependence structure. In the sections below we will test this.

2.2. The quantile spectral density function

We now formalise the discussion above. Let us suppose that $\{X_t\}$ is a strictly stationary time series. It is obvious that the cross covariance of the indicator functions $\{I(X_t \leq x), I(X_t \leq y)\}$ is

$$C_r(x,y) := \mathrm{cov}(I(X_0 \leq x), I(X_r \leq y)) = P(X_0 \leq x, X_r \leq y) - P(X_0 \leq x)P(X_r \leq y).$$

Skaug and Tjøstheim (1993) and Hong (2000) use a similar quantity to test for serial independence of a time series (and this definition dates back to Hoeffding (1948)). We will call $C_r(\cdot)$ the *quantile covariance.* If $\{X_t\}$ is an α-mixing time series with mixing rate $s > 1$ (s is defined in Assumption IV.1, below) it can be shown that $\sup_{x,y} \sum_r |\mathrm{cov}(I(X_0 \leq x), I(X_r \leq y))| < \infty$, thus for all $x, y \in \mathbb{R}$, it's Fourier transform

$$G(x,y;\omega) \;\; = \;\; \frac{1}{2\pi} \sum_r C_r(x,y;\omega) \exp(ir\omega),$$

is well defined. Since $G(x,y;\omega)$ can be considered as the cross-spectral density of $\{I(X_t \leq x), I(X_t \leq y)\}$, we call $G(\cdot)$ the *quantile spectral density.*

2.2.1. Properties of the quantile spectral density

The quantile spectral density carries all the information about the serial dependence structure of the time series. For example (i) if $\{X_t\}$ is serially independent, then G does not depend on ω and $G(x, y; \omega) \propto C_0(x, y)$, (ii) if for all r, the distribution function of (X_0, X_r) is identical to the distribution function of (X_0, X_{-r}), then $G(\cdot)$ will be real and (iii) for any given x and y, G gives information about any periodicities that may exists at a given threshold. In addition, $G(\cdot)$ captures the covariance structure of any transformation of $\{X_t\}$. For example, consider the transformation $\{h(X_t)\}$, then it is straightforward to show that the spectral density of the time series $\{h(X_t)\}$ is

$$f_h(\omega) = \frac{1}{2\pi} \sum_r \mathrm{cov}(h(X_0), h(X_r)) \exp(ir\omega) = \int \int h(x)h(y)G(dx, dy; \omega).$$

Of course, $G(x, y; \omega)$ only captures the serial dependency, and may miss higher order structure. Only in the case that $\{X_t\}$ is Markovian, does $G(x, y; \omega)$ capture the entire joint distribution of $\{X_t\}$.

Remark II.1 *The quantile spectral density is closely related to the generalised spectral density introduced in Hong (1998). He defines the generalised spectral density as* $h(x, y; \omega) = \sum_r cov(\exp(ixX_0), \exp(iyX_r)) \exp(ir\omega)$. *Essentially, this is the Fourier transform of the characteristic function of pairwise distributions minus their marginals, therefore the relationship between the quantile spectral density and the generalised spectral density is analogous to that between the distribution function and the characteristic function of a random variable. Hong (1998, 2003) uses the generalised spectral density as a tool in various tests goodness of fit tests, which are mainly based on the residual. On the other hand, the goodness of fit test that we propose, is based on checking for similarity between the estimated quantile spectral density and the pro-*

posed spectral density.

Remark II.2 (The Copula spectral density) *A closely related quantity to the quantile spectral density is the copula spectral density, which is defined as*

$$G_{\mathcal{C}}(u_1, u_2; \omega) \;=\; \frac{1}{2\pi} \sum_{r} \mathcal{C}_r(u_1, u_2) \exp(ir\omega), \tag{2.1}$$

where $\mathcal{C}_r(u_1, u_2) = cov(I(F(X_0) \leq u_1), I(F(X_r) \leq u_2)) = \mathbb{E}(I(F(X_0) \leq u_1)I(F(X_r) \leq u_2)) - u_1u_2$, *and* $F(\cdot)$ *is marginal distribution function of* $\{X_t\}$. *Note that by definition* $u_1, u_2 \in [0,1]$. *Thus, unlike the quantile spectral density, the copula spectral density is invariant to any monotonic transformation of* $\{X_t\}$, *for example mean and variance shifts. By considering the ranks of* $\{X_t\}$, *the methods detailed in the section below can also be used to estimate* $G_{\mathcal{C}}$. *Alternatively, Dette et al. (2011) have recently proposed* L_1*-methods for estimating* $G_{\mathcal{C}}$, *and the asymptotic sampling properties have been derived for this estimator.*

In Figures 3, 4 and 5 we plot the quantile spectral density for the autoregressive ($X_t = 0.9X_{t-1} + Z_t$), ARCH ($X_t = \sigma_t Z_t$ with $\sigma_t^2 = 1/1.9 + 0.9X_{t-1}^2$) and the squared ARCH, with independent, identically distributed (iid) Gaussian innovations Z_t. The diagonals are of $G(x, x; \omega)$, the lower triangle contains the real part of $G(x, y; \omega)$ and the upper triangle the imaginary part of $G(x, y; \omega)$. We observe that the AR and ARCH quantile spectral densities are very different. The AR has a similar shape for all x, whereas for the ARCH, it is flat (like the spectral density of uncorrelated data) at about the 50% percentile, but moves away from flatness at the extremes. Furthermore, recalling that the AR and ARCH squared have the same spectral density (if the moments of the ARCH squared exists), there is a large difference between the quantile spectral density of the AR and the ARCH squared.

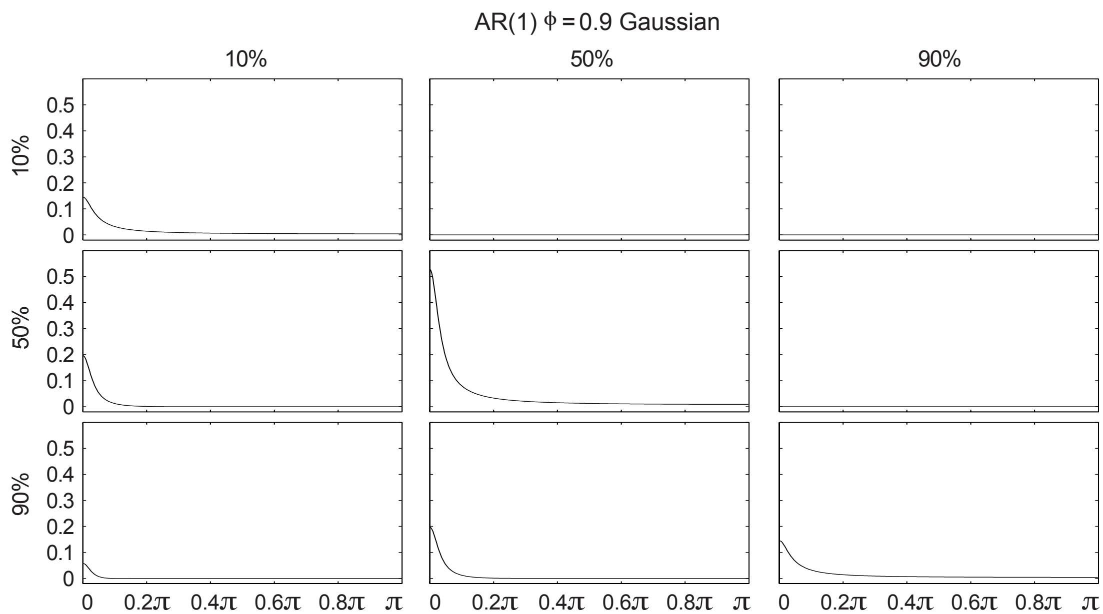

Fig. 3. The quantile spectral density of $X_t = 0.9X_{t-1} + Z_t$

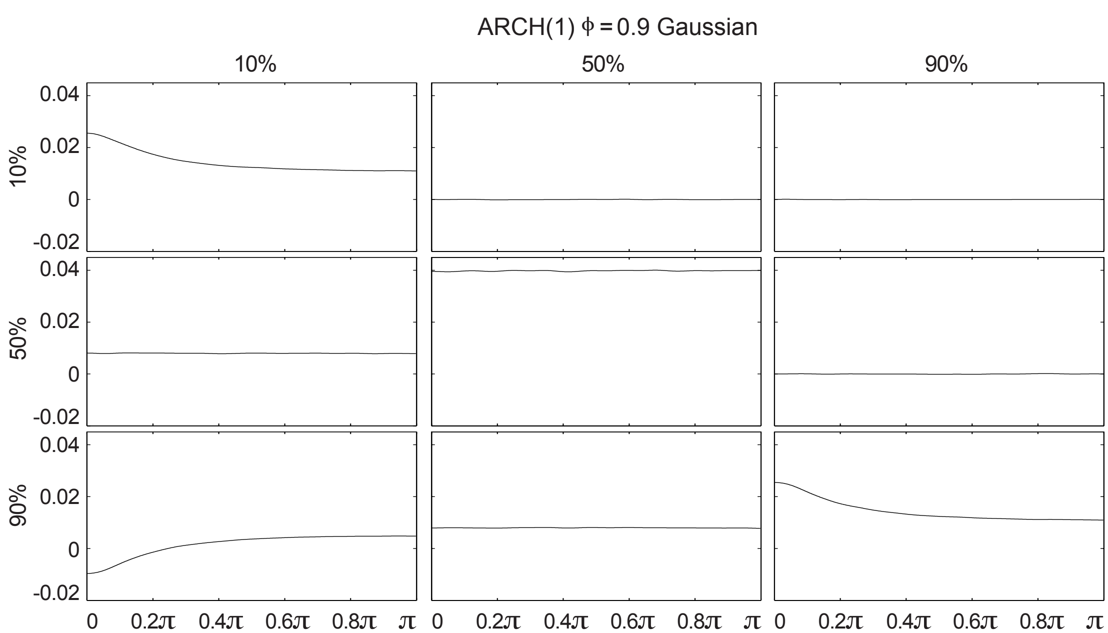

Fig. 4. The quantile spectral density of $X_t = \sigma_t Z_t$, where $\sigma_t^2 = 1/1.9 + 0.9X_{t-1}^2$

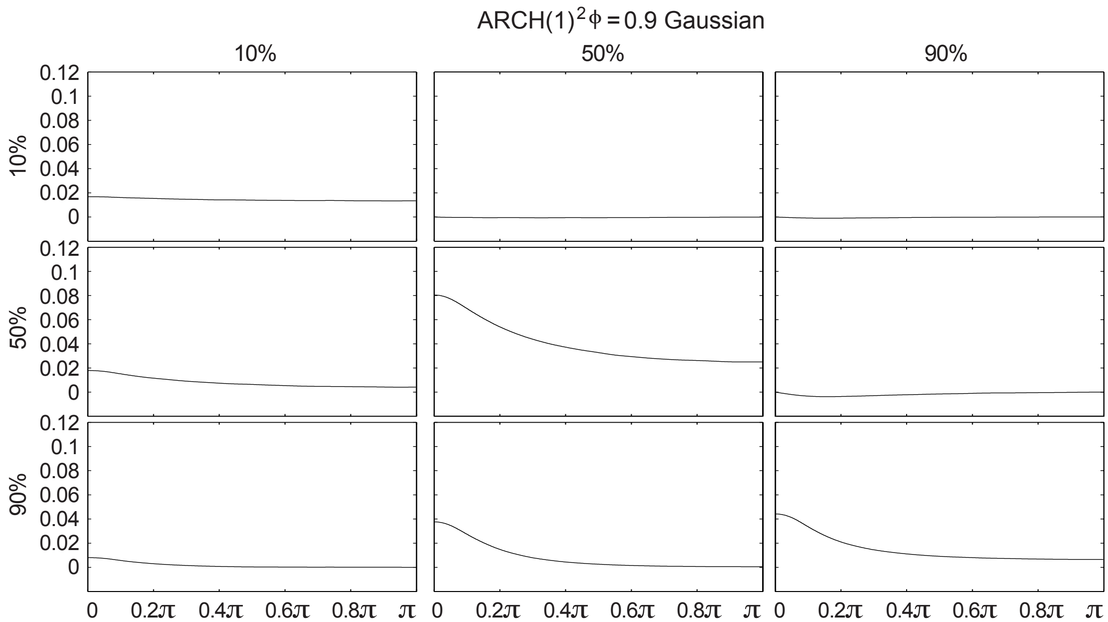

Fig. 5. The quantile spectral density of $X_t^2 = \sigma_t^2 Z_t^2$, where $\sigma_t^2 = 1/1.9 + 0.9X_{t-1}^2$

2.2.2. Estimating the quantile spectral density

The quantile spectral density $G(x, y; \omega)$ can be considered as the cross spectral density of the bivariate time series $\{I(X_t \leq x), I(Y_t \leq y)\}$. Therefore, our estimator of $G(x, y; \omega)$ is motivated by the classical cross spectral. To do this we define the class of lag windows we shall use.

Definition II.1 *The lag window takes the form*

$$\lambda(u) = \Big(\sum_{j=-r}^{r} a_r \exp(i2\pi ru) - \sum_{j=1}^{r} b_j |u|^j\Big) I_{[-1,1]}(u).$$

This class of lag windows is quite large, and includes the truncated window, the Bartlett window and general Tukey window (see, for example, Priestley (1981) Section 6.2.3 for properties of these lag windows).

To obtain an estimator of G, we define the centralised, transformed variable $Z_t(x) = I(X_t \le x) - \widehat{F}_T(x)$ (where $\widehat{F}_T(x) = \frac{1}{T}\sum_t I(X_t \le x)$). We estimate the quantile covariance $C_r(x,y) = P(X_0 \le x, X_r \le y) - P(X_0 \le x)P(X_r \le y)$ with $\widehat{C}_r(x,y) = \frac{1}{T}\sum_{t=1}^{T-|r|} Z_t(x)Z_{t+r}(y)$, and use as an estimator of G

$$\begin{aligned}\widehat{G}_T(x,y;\omega_k) &= \frac{1}{2\pi}\sum_r \lambda_M(r)\widehat{C}_r(x,y)\exp(ir\omega_k) \\ &= \sum_s K_M(\omega_k-\omega_s)J_T(x;\omega_s)\overline{J_T(y;\omega_s)},\end{aligned} \tag{2.2}$$

where

$$\begin{aligned}\lambda_M(r) &= \lambda(r/M) \\ K_M(\omega) &= \frac{1}{T}\sum_r \lambda_M(r)\exp(ir\omega) \\ J_T(x;\omega) &= \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} Z_t(x)\exp(it\omega).\end{aligned}$$

2.3. The test statistic

The proposed test is based on the fit of the estimated quantile spectral density to the conjectured quantile spectral density. More precisely, we test $H_0 : G(x,y;\omega) = G_0(x,y;\omega)$ against $H_A : G(x,y;\omega) \neq G_0(x,y;\omega)$, where G is the quantile spectral density of $\{X_t\}$, $G_0(x,y;\omega) = \frac{1}{2\pi}\sum_r C_{0,r}(x,y)\exp(ir\omega)$ and $C_{0,r}(x,y) = F_{0,r}(x,y) - F_0(x)F_0(y)$. Thus under the null the marginal distribution is $F_0(\cdot)$ and the joint distribution is $F_{0,r}(\cdot)$. We use the quadratic distance to measure the distance between the estimated quantile spectral density and the conjectured spectral density, and define the test statistic as

$$\begin{aligned}
\mathcal{Q}_T &= \frac{1}{T}\sum_{k=1}^{T}\int |\widehat{G}_T(x,y;\omega_k) - \frac{1}{2\pi}\sum_r \lambda_M(r)C_{0,r}(x,y)\exp(ir\omega_k)|^2 dF_0(x)dF_0(y) \\
&= \frac{1}{T}\sum_{k=1}^{T}\int |\widehat{G}_T(x,y;\omega_k) - \sum_{s=1}^{T} K_M(\omega_k-\omega_s)G_0(\omega_s)|^2 dF_0(x)dF_0(y) \\
&= \frac{1}{2\pi}\sum_r \lambda_M(r)^2 \int\int \left|\hat{C}_r(x,y) - C_{0,r}(x,y)\right|^2 dF_0(x)dF_0(y), \qquad (2.3)
\end{aligned}$$

where the above immediately follows from Parseval's theorem. The choice of lag window will have an influence on the type of alternatives the test can detect. For example, the truncated window ($\lambda(u) = I_{[-1,1]}(u)$) gives equal weights to all the quantile covariances, whereas the Bartlett window ($\lambda(u) = (1-|u|)I_{[-1,1]}(u)$) gives more weight to the lower order lags. Therefore the tests ability to detect the alternative will depend on which order of the quantile covariance deviates most from the null, and the weight the lag window places on these. We derive the asymptotic distribution of $\mathcal{Q}_T$ in Section 3.2.

Remark II.3 *The test can be adapted to be invariant to monotonic transformations (such as shifts of mean and variance). This can be done by replacing the quantile spectral density with the copula spectral density $G_{\mathcal{C}}(\cdot)$ defined in (2.1). In this case the null is $H_0 : G_{\mathcal{C}}(x,y;\omega) = G_{\mathcal{C},0}(x,y;\omega) = \frac{1}{2\pi}\sum_r \mathcal{C}_{0,r}(u_1,u_2;\omega)\exp(ir\omega)$ against $H_A : G_{\mathcal{C}}(x,y;\omega) \neq G_{\mathcal{C},0}(x,y;\omega)$. The test statistic in this case is*

$$\mathcal{Q}_{T,\mathcal{C}} = \frac{1}{T}\sum_{k=1}^{T}\int |\widehat{G}_{T,\mathcal{C}}(u_1,u_2;\omega_k) - \frac{1}{2\pi}\sum_r \lambda_M(r)\mathcal{C}_{0,r}(u_1,u_2)\exp(ir\omega_k)|^2 du_1du_2,$$

where we estimate $\widehat{G}_{T,\mathcal{C}}(u_1,u_2;\omega_k)$ in the same way as we have estimated $\widehat{G}_T$ in (2.2) but replace $\{X_t\}_t$ with $\{\hat{F}_T(X_t)\}_t$. The distribution of $\mathcal{Q}_{T,\mathcal{C}}$ is beyond the scope of the current paper.

3. Sampling properties

In this section we derive the sampling properties of the quantile spectral density $\widehat{G}_T$ and the test statistic $\mathcal{Q}_T$. We will use the α-mixing assumptions below.

Assumption II.1 *Let us suppose that $\{X_t\}$ is a strictly stationary α-mixing time series such that*

$$\sup_{\substack{A\in\sigma(X_r,X_{r+1},\ldots)\\ B\in\sigma(X_0,X_{-1},\ldots)}} |P(A\cap B)-P(A)P(B)| \leq \alpha(r),$$

where $\alpha(r)$ are the mixing coefficients which satisfy $\alpha(r)\leq K|r|^{-s}$ for some $s>2$.

3.1. Sampling properties of $\widehat{G}_T$

In the following lemma we derive the limiting distribution of $\widehat{G}_T$, this will allow us to construct point wise confidence intervals for G.

Theorem II.1 *Suppose Assumption IV.1 holds. Then*

$$\mathbb{E}(\widehat{G}_T(x,y;\omega)) = G(x,y;\omega) + O(\frac{1}{M^{s-1}}),$$

and for $0<\omega_k<\pi$ we have

$$V_T(x,y;\omega_k)^{-1/2}\begin{pmatrix}\Re\widehat{G}_T(x,y;\omega_k)-\Re\mathbb{E}(\widehat{G}_T(x,y;\omega_k))\\ \Im\widehat{G}_T(x,y;\omega_k)-\Im\mathbb{E}(\widehat{G}_T(x,y;\omega_k))\end{pmatrix} \stackrel{D}{\to} \mathcal{N}\big(0,I_2\big)$$

$$V_T(x,x;\omega_k)^{-1/2}\Big(\widehat{G}_T(x,x;\omega_k)-\mathbb{E}(\widehat{G}_T(x,x;\omega_k))\Big) \stackrel{D}{\to} \mathcal{N}(0,1),$$

where $M\to\infty$ and $M/T\to 0$ as $T\to\infty$,

$$V_T(x,y;\omega_k) = \sum_{s=1}^{T} K_M(\omega_k-\omega_s)^2\begin{pmatrix}A(x,y;\omega_s) & C(x,y;\omega_s)\\ C(x,y;\omega_s) & B(x,y;\omega_s)\end{pmatrix} = O(\frac{M}{T}),$$

and

$$
\begin{aligned}
A(x, y; \omega_s) &= \frac{1}{2}\Big(G(x, x; \omega_s)G(y, y; \omega_s) + \Re G(x, y; \omega_s)^2 - \Im G(x, y; \omega_s)^2\Big) \\
B(x, y; \omega_s) &= \frac{1}{2}\Big(G(x, x; \omega_s)G(y, y; \omega_s) + \Im G(x, y; \omega_s)^2 - \Re G(x, y; \omega_s)^2\Big) \\
C(x, y; \omega_s) &= \Re G(x, y; \omega_s)\Im G(x, y; \omega_s).
\end{aligned}
$$

Thus, if $\frac{M}{T} >> \frac{1}{M^{2(s-1)}}$, in other words the variance of $\widehat{G}_T$ dominates the bias, then we can use the above result to construct confidence intervals for G.

3.2. Sampling properties of test statistic under the null hypothesis

We now derive the limiting distribution of the test statistic under the null hypothesis. Let

$$
\begin{aligned}
E_T &= \frac{1}{T}\int\int W_M(\omega - \theta)^2 G(x, x; \theta)G(y, y; \theta)dF_0(x)dF_0(y)d\theta d\omega \\
V_T &= \frac{4}{T^2}\int\int \Delta_M(\theta_1 - \theta_2)^2 \prod_{i=1}^{2} G(x_1, x_2; \theta_i)G(y_1, y_2; \theta_i)d\theta_i dF_0(x_i)dF_0(y_i),
\end{aligned}
$$

where

$$
\begin{aligned}
W_M(\theta) = \frac{T}{2\pi}K_M(\theta) &= \frac{1}{2\pi}\sum_r \lambda_M(r)\exp(ir\theta) \\
\Delta_M(\theta_1 - \theta_2) &= \int W_M(\omega - \theta_1)W_M(\omega - \theta_2)d\omega. \qquad (2.4)
\end{aligned}
$$

Lemma II.1 *Suppose that Assumption IV.1 holds and $G(\cdot)$ is the quantile spectral density of $\{X_t\}$. Then under the null hypothesis we have*

$$
\mathbb{E}\big(\mathcal{Q}_T\big) = E_T + O(\frac{1}{T}) = O(\frac{M}{T}) \text{ and } var\big(\mathcal{Q}_T\big) = V_T + O(\frac{1}{T}) = O(\frac{M}{T^2}).
$$

Using the above we obtain the limiting distribution under the null.

Theorem II.2 *Suppose that Assumption IV.1 holds. Then under the null hypothesis* *we have*

$$V_T^{-1/2}\big(\mathcal{Q}_T - E_T\big) \xrightarrow{D} \mathcal{N}(0,1)$$

as $M \to \infty$ *and* $M/T \to 0$ *as* $T \to \infty$.

Using estimates of $\widehat{G}_T(\cdot)$, E_T and V_T can both be estimated. Thus by using the above result, we reject the null at the $\alpha\%$ level if $V_T^{-1/2}\big(\mathcal{Q}_T - E_T\big) > z_{1-\alpha}$ (where $z_{1-\alpha}$ denotes the $1-\alpha$ quantile of a standard normal distribution).

3.3. Behavior of the test statistic under the alternative hypothesis

We now examine the behavior of the test statistic under the alternative

$$H_A : G(x,y;\omega) = G_1(x,y;\omega) = \frac{1}{2\pi}\sum_r \big(F_{1,r}(x,y) - F_1(x)F_1(y)\big)\exp(ir\omega).$$

To obtain the limiting distribution we decompose the test statistic $\mathcal{Q}_T$ as

$$\mathcal{Q}_T = \mathcal{Q}_{T,1} + \mathcal{Q}_{T,2} + \mathcal{Q}_{T,3}$$

where

$$\begin{aligned}
\mathcal{Q}_{T,1} &= \frac{1}{T}\sum_{k=1}^{T}\int \big|\widehat{G}_T(x,y;\omega_k) - \mathbb{E}(\widehat{G}_T(x,y;\omega_k))\big|^2 dF_0(x)dF_0(y) \\
\mathcal{Q}_{T,2} &= \frac{2}{T}\Re\sum_{k=1}^{T}\int \big(\hat{G}_T(x,y;\omega_k) - \mathbb{E}(\hat{G}_T(x,y;\omega_k))\big) \\
&\quad \times\big(\mathbb{E}(\hat{G}_T(x,y;\omega_k)) - \tilde{G}(x,y;\omega_k)\big)dF_0(x)dF_0(y) \\
\mathcal{Q}_{T,3} &= \frac{1}{T}\sum_{k=1}^{T}\int \big|\mathbb{E}(\hat{G}_T(x,y;\omega_k)) - \tilde{G}(x,y;\omega_k)\big|^2 dF_0(x)dF_0(y),
\end{aligned}$$

and

$$\tilde{G}(x,y;\omega_k) = \frac{1}{2\pi}\sum_r \lambda_M(r)C_{0,r}(x,y)\exp(ir\omega) = \sum_s K_M(\omega_k - \omega_s)G_0(x,y;\omega_s).$$

From the decomposition of $\mathcal{Q}_T$, we observe that there are two stochastic terms $\mathcal{Q}_{T,1}$ and $\mathcal{Q}_{T,2}$, and a deterministic term $\mathcal{Q}_{T,3}$. By using Lemma II.1, it can be shown that $\mathcal{Q}_{T,1} = O_p(\frac{M^{1/2}}{T} + \frac{M}{T})$. On the other hand, we show in the proof of the theorem below that $\mathcal{Q}_{T,2}$, is of lower order than $\mathcal{Q}_{T,1}$ and, thus, determines the distribution of $\mathcal{Q}_T$. To understand the role that $\mathcal{Q}_{T,3}$ plays in the test, we replace $\tilde{G}(x,y;\omega)$ and $\mathbb{E}(\widehat{G}_T(x,y;\omega))$ with G_0 and G_1 respectively and obtain

$$\mathcal{Q}_{T,3} \;=\; \frac{1}{T}\sum_{k=1}^{T}\int \big|G_1(x,y;\omega_k) - G_0(x,y;\omega_k)\big|^2 dF_0(x)dF_0(y) + O(\frac{1}{M^{s-1}}).$$

Thus $\mathcal{Q}_{T,3}$ measures the deviation of the alternative from the null hypothesis, and shifts the mean of the test statistic.

Theorem II.3 *Suppose that Assumption IV.1 holds, and for all* r, $\sup_{x,y}|C_{r,0}(x,y)| \leq K|r|^{-(2+\delta)}$, *for some* $\delta > 0$. *Under the alternative hypothesis we have*

$$\sqrt{T}\mathcal{Q}_{T,2} \overset{D}{\to} \mathcal{N}(0, V_{T,2}), \tag{2.5}$$

and

$$\sqrt{T}\big(\mathcal{Q}_T - \mathcal{Q}_{T,3}\big) \overset{D}{\to} \mathcal{N}(0, V_{T,2}), \tag{2.6}$$

where $M \to \infty$ *and* $\sqrt{M}/T \to 0$ *as* $T \to \infty$,

$$\begin{aligned} V_{T,2} \;&= \frac{8}{T}\Re\int\int \Lambda_T(x_1,y_1;\omega)\overline{\Lambda_T(x_2,y_2;\omega)} \\ &\Big\{G_1(x_1,x_2;\omega)G_1(y_1,y_2;\omega) + G_1(x_1,y_2;\omega)G_1(y_1,x_2;\omega)\Big\}d\omega\prod_{i=1}^{2}dF_0(x_i)dF_0(y_i) \\ &+\frac{8}{T}\Re\int\int \Lambda_T(x_1,y_1;\omega_1)\overline{\Lambda_T(x_2,y_2;\omega_2)} \\ &G_{(x_1,y_1,x_2,y_2)}(\omega_1,-\omega_1,\omega_2)\prod_{i=1}^{2}dF_0(x_i)dF_0(y_i)d\omega_i, \end{aligned}$$

$\Lambda_T(x,y;\omega_s) = \frac{1}{2\pi}\sum_r \lambda_M(r)^2(\frac{T-|r|}{T})[C_{1,r}(x,y) - C_{0,r}(x,y)]\exp(ir\omega_k)$ *and* $G_{(x_1,y_1,x_2,y_2)}$ *is the cross tri-spectral density of* $\{(I(X_t \leq x_1), I(X_t \leq y_1), I(X_t \leq x_2), I(X_t \leq y_2))\}_t$.

The theorem above tells us that the mean of the test statistic is shifted the further the alternative is from the null. Interestingly, we observe from the definition of $\Lambda_T(\cdot)$, that the variance also depends on the difference between the null and alternative, which increases as the difference increase. However, for a fixed alternative the above result tells us that the power converges to 100% as the sample size grow.

4. Testing for equality of serial dependence of two time series

The above test statistic can easily be adapted to test other hypothesis. In this section, we consider one such example, and consider testing for equality of serial dependence between two time series. Let us suppose that $\{U_t\}$ and $\{V_t\}$ are two stationary time series, and we wish to test whether they have the same sequential dependence structure. Using the same motivation as that for the the goodness of fit test described above we define the test statistic

$$\mathcal{P}_T = \frac{1}{T}\sum_{k=1}^{T}\int |\widehat{G}_{1,T}(x,y;\omega_k) - \widehat{G}_{2,T}(x,y;\omega_k)|^2 dF(x)dF(y),$$

where $\widehat{G}_{1,T}$ and $\widehat{G}_{2,T}$ are the quantile spectral density estimators based on $\{U_t\}$ and $\{V_t\}$ respectively and F is any distribution function. In order to obtain the limiting distribution under the null hypothesis we have $H_0 : G_1(x,y;\omega) = G_2(x,y;\omega)$ and the alternative $H_A : G_1(x,y;\omega) \neq G_2(x,y;\omega)$ we expand $\mathcal{P}_T$

$$\mathcal{P}_T := \mathcal{Q}_{1,1,T} + \mathcal{Q}_{2,2,T} - \mathcal{Q}_{1,2,T} - \mathcal{Q}_{2,1,T} + 2\mathcal{L}_{1,T} + 2\mathcal{L}_{2,T} + \mathcal{D},$$

where

$$
\begin{aligned}
\mathcal{Q}_{i,j,T} &= \frac{1}{T}\sum_{k=1}^{T}\int \big(\widehat{G}_{i,T}(x,y;\omega_k) - \mathbb{E}(\widehat{G}_{i,T}(x,y;\omega_k))\big) \\
&\quad \overline{\big(\widehat{G}_{j,T}(x,y;\omega_k) - \mathbb{E}(\widehat{G}_{j,T}(x,y;\omega_k))\big)} dF(x)dF(y), \\
\mathcal{L}_{i,T} &= \Re\frac{1}{T}\sum_{k=1}^{T}\int \big(\widehat{G}_{i,T}(x,y;\omega_k) - \mathbb{E}(\widehat{G}_{i,T}(x,y;\omega_k))\big) \\
&\quad \big(\mathbb{E}(\widehat{G}_{1,T}(x,y;\omega)) - \mathbb{E}(\widehat{G}_{2,T}(x,y;\omega))\big) dF(x)dF(y)
\end{aligned}
$$

and

$$
\mathcal{D} = \int\int\int |\mathbb{E}(\widehat{G}_{1,T}(x,y;\omega)) - \mathbb{E}(\widehat{G}_{2,T}(x,y;\omega))|^2 dF(x)dF(y)d\omega.
$$

Therefore, using the above expansion under the null hypothesis we have

$$
\mathcal{P}_T := \mathcal{Q}_{1,1,T} + \mathcal{Q}_{2,2,T} - \mathcal{Q}_{1,2,T} - \mathcal{Q}_{2,1,T},
$$

where the moments are $\mathbb{E}\big(\mathcal{P}_T\big) = E_{T,3} + O(\frac{1}{T}) = O(\frac{M}{T})$ and $\mathrm{var}\big(\mathcal{P}_T\big) = V_{T,3} + O(\frac{1}{T}) = O(\frac{M}{T^2})$, with

$$
\begin{aligned}
E_{T,3} &= \frac{1}{T}\int\int W_M(\omega-\theta)^2\big(\sum_{i=1}^{2} G_i(x,x;\theta)G_i(y,y;\theta)\big) dF(x)dF(y)d\theta d\omega \\
V_{T,3} &= \frac{4}{T^2}\sum_{i=1}^{2}\int\int \Delta_M(\theta_1-\theta_2)^2 \prod_{j=1}^{2} G_i(x_1,y_2;\theta_i)G_j(y_1,x_2;\theta_j)d\theta_j dF(x_j)dF(y_j).
\end{aligned}
$$

By using identical arguments as those used in the proof of Theorem II.2, under the null hypothesis we have

$$
V_{T,3}^{-1/2}\big(\mathcal{P}_T - E_{T,3}\big) \xrightarrow{D} \mathcal{N}(0,1).
$$

Using the above result, we test for equality of sequential dependence, that is we reject the null hypothesis at the α-level if $|V_{T,3}^{-1/2}(\mathcal{P}_T - E_{T,3})| > z_{1-\alpha}$.

The limiting distribution of the alternative can be derived using the same methods as those used to derive the limiting distribution of $\mathcal{Q}_T$ under its alternative. It can be shown that

$$\mathcal{P}_T - \mathcal{D} := \underbrace{2\mathcal{L}_{1,T} + 2\mathcal{L}_{2,T}}_{O_p(\frac{1}{\sqrt{T}})} + O_p(\frac{M^{1/2}}{T}),$$

where $2\mathcal{L}_{1,T} + 2\mathcal{L}_{2,T}$ can be approximated by a quadratic form. Using this quadratic approximation, asymptotic normality of the above can be shown. Thus under a fixed alternative the power grows to 100% as $T \to \infty$.

Remark II.4 *We can easily adapt our method to test that the distributions of* (X_0, X_r) *and* (X_{-r}, X_0) *are identical for all* r *(ie.* $F_r(x,y) = F_{-r}(x,y)$*). This means that the imaginary part of the quantile spectral density* $G(\cdot)$ *is zero over all* x, y *and* ω*. In this case, we can use the test statistic*

$$\mathcal{R}_T = \frac{1}{T}\sum_r \left|\Im \widehat{G}_T(x, y; \omega)\right|^2 dF(x)dF(y),$$

where F *is some distribution. We can use identical methods to those above to derive the distribution under the null. Dette et al. (2011) also consider time reversibility and their impact on the quantile spectral density.*

5. Bootstrap approximation

The asymptotic normality result that we use to obtain the p-value of the test statistic $\mathcal{Q}_T$ is only an approximation. For small samples, the normality approximation may not be particularly good, mainly because $\mathcal{Q}_T$ is a positive random variable, whose distribution will be skewed. This may well lead to more false positive than we can control for in our type I error.

To correct for this, we propose estimating the finite sample distribution of $\mathcal{Q}_T$ using a frequency domain bootstrap procedure. In a multivariate time series, the periodogram matrix at the fundamental frequencies asymptotically follows a Wishart distribution, moreover for our purposes they are close enough to be independent such that we don't loose too much information by treating them as independent (observe that the asymptotic variance of the test statistic $\mathcal{Q}_T$ is only in terms of the pair-wise distributions and does not contain any higher order dependencies). Thus motivated by the frequency domain bootstrap methods proposed in Hurvich and Zeger (1987) and Franke and Härdle (1992) for univariate data and Berkowitz and Diebold (1998) and Dette and Paparoditis (2009) for multivariate data, we propose the following bootstrap scheme to obtain an estimate of the finite sample distribution under the null hypothesis.

Let $x_1 < \cdots < x_q$ be a finite discretization of the real line and note that we approximate $\mathcal{Q}_T$ with the discretization

$$\begin{aligned}\mathcal{Q}_T &= \frac{2\pi}{T}\sum_{k=1}^{T}\sum_{i_1,i_2=2}^{q}|\hat{G}(x_{i_1},x_{i_2};\omega_k) - \sum_s K_M(\omega_k-\omega_s)G_0(x_{i,1},x_{i_2};\omega_k)|^2 \times \\ &\quad (F_0(x_{i_1}) - F_0(x_{i_1-1}))(F_0(x_{i_2}) - F_0(x_{i_2-1})).\end{aligned}$$

We observe that under the null hypothesis that $\mathbf{G_Z}(\omega)$ will be the spectral density matrix of the q-dimensional multivariate time series $\mathbf{Z}_t = (\tilde{Z}_t(x_1), \cdots, \tilde{Z}_t(x_q))$ where $\tilde{Z}_t(x) = I(X_t \leq x) - F(x)$ and $\mathbf{G_Z}(\omega)_{i_1,i_2} = G_0(x_{i_1}, x_{i_2}; \omega)$. Thus we use the transformation of X_t into a high dimensional multivariate time series to construct the the bootstrap distribution.

The steps of the frequency domain boostrap for the test statistic $\mathcal{Q}_T$ are as follows.

Step 1: Generate T independent matrices $\mathbf{I}^*_{\mathbf{Z}}(\omega_k) = \mathbf{G}_{\mathbf{Z}}(\omega_k)^{1/2} W^*_k \mathbf{G}_{\mathbf{Z}}(\omega_k)^{1/2}$, where

$$W^*_k \sim \begin{cases} W^C_q(1, I_q) & 1 \leq k < T/2 \\ W^R_q(1, I_q) & k \in \{0, T/2\} \\ \overline{W^*_{T-k}} & T/2 < k \leq T \end{cases},$$

where W^C and W^R denote the complex and real Wishart distributions.

Step 2: Construct the bootstrap quantile spectral density matrix estimators with $\hat{\mathbf{G}}^*_{\mathbf{Z}}(\omega_k) = \sum_s K_M(\omega_k - \omega_s)\mathbf{I}^*_{\mathbf{Z}}(\omega_s)$ for $k = 1, \ldots, T$.

Step 3: Obtain the bootstrap test statistic

$$\begin{aligned} \mathcal{Q}^*_T &= \frac{2\pi}{T} \sum_{k=1}^{T} \sum_{i_1, i_2 = 2}^{q} |\hat{G}^*(x_{i_1}, x_{i_2}; \omega_k) - G^M_0(x_{i_1}, x_{i_2}; \omega_k)|^2 \times \\ & \quad (F_0(x_{i_1}) - F_0(x_{i_1 - 1}))(F_0(x_{i_2}) - F_0(x_{i_2 - 1})) \end{aligned}$$

where $G^M_0(x, y; \omega) = \frac{1}{2\pi} \sum_k \lambda(\frac{k}{M}) C_{0,r}(x, y) \exp(ir\omega)$.

Step 4: Approximate the distribution of $\mathcal{Q}_T$ under the null by using the empirical distribution of the bootstrap sample $\{\mathcal{Q}^*_T\}$.

Step 5: Based on the bootstrap distribution estimate the p-value of $\mathcal{Q}_T$.

We illustrate our procedure in Figure 6, for this example we use the quantile spectral density G_0, based on an ARCH(1) ($X_t = Z_t \sigma_t$ and $\sigma^2_t = a_0 + a_1 X^2_{t-1}$), where $a_0 = 1/1.9$, $a_1 = 0.9$, Z_t are iid standard normal random variables and $T = 500$. A plot of the normal approximation, the density of $\mathcal{Q}_T$ (which is estimated and based on 500 replications) and the bootstrap estimator of the density (along with their rejection regions) is given in Figure 6. We observe that the skew in the finite sample distribution means that the normal distribution is under estimating the location of the rejection region. However, the bootstrap approximation appears to capture the finite sample distribution (and this the rejection region) quite well.

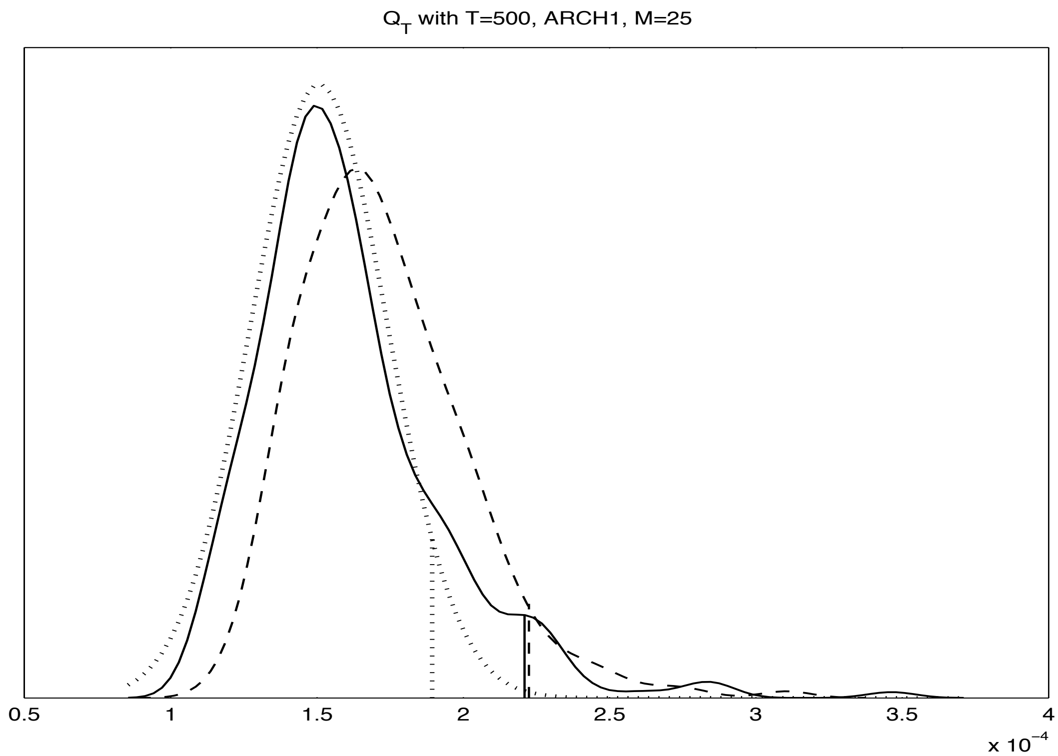

Fig. 6. The fine line is the standard normal (with the 5% rejection line), the thick solid line is the finite sample density of the test statistic (with 5% rejection region) and the thick dashed line is the bootstrap approximation (with 5% rejection region).

6. Simulations and real data examples

6.1. Simulations

In this section we conduct a simulation study. In order to determine the effectiveness of the test we will use two different models that have the same first and second order structure (thus a test based on the covariance structure would not be able to distinguish between them). In particular, we will consider the AR(1) model $X_t = \mu + aX_{t-1} + \varepsilon_t$ and the squares of the ARCH(1) model $Y_t = a_0 + aY_{t-1} + (Z_t^2 - 1)(a_0 + aY_{t-1})$, where $\{\varepsilon_t\}$ and $\{Z_t\}$ are iid zero mean Gaussian random variables with $\mathrm{var}(Z_t) = 1$ and μ and $\mathrm{var}(\varepsilon_t)$ chosen such that X_t and Y_t have the same mean and covariance structure. Note that in the simulation we only consider $a \leq 0.55$, so that the spectral density of the squared ARCH exists. For each model we did 1000 replications and the tests was done at both the $\alpha = 0.1$ and $\alpha = 0.05$ level.

In our simulations we used the Bartlett window, compared the test for various M and used both the normal approximation and the proposed bootstrap procedure. The results for H_0 : AR(1) against the alternative H_A : ARCH(1) (various a, fixing $a_0 = 0.4$) are given in Table 1 and 2. The results for H_0 : ARCH(1) against H_A : AR(1) are given in Table 3 and 4. We use the sample sizes $T = 100$ and 500.

As expected under the null hypothesis the null hypothesis tends to over reject, whereas the bootstrap gives a better approximation of the significance level. There appears to be very little difference in the behavior under the null for various values of a and between the AR and the ARCH. Under the alternative, the power seems to be quite high even for quite small samples. The only model where the power is not close to 100% is when $a = 0.3$, sample size $T = 100$, the null is an AR(1) and the alternative is an ARCH(1). This can be explained by the fact that for small values of a, both the AR and the ARCH models are relatively close to independent observations, thus making it relatively difficult to reject the null.

Table 1. $H_0 : AR(1)$ vs $H_A : ARCH(1)$ $T = 100$

$T = 100$		$\alpha = 0.1$				$\alpha = 0.05$			
		Bootstrap		Normal		Bootstrap		Normal	
a	M	H_0	H_A	H_0	H_A	H_0	H_A	H_0	H_A
0.3	11	0.052	1	0.076	1	0.021	0.972	0.054	1
	16	0.04	0.869	0.062	0.971	0.011	0.262	0.04	0.854
	21	0.048	0.386	0.064	0.561	0.021	0.106	0.043	0.348
	25	0.021	0.071	0.048	0.229	0.014	0.016	0.029	0.12
0.4	11	0.048	1	0.082	1	0.02	1	0.055	1
	16	0.043	1	0.059	1	0.013	0.939	0.041	1
	21	0.046	0.932	0.066	0.997	0.011	0.416	0.046	0.929
	25	0.036	0.582	0.055	0.832	0.01	0.124	0.037	0.598
0.5	11	0.046	1	0.073	1	0.015	1	0.052	1
	16	0.049	1	0.078	1	0.027	1	0.045	1
	21	0.046	1	0.06	1	0.015	0.985	0.037	1
	25	0.047	1	0.062	1	0.015	0.397	0.043	1
0.55	11	0.041	1	0.096	1	0.018	1	0.057	1
	16	0.045	1	0.066	1	0.017	1	0.046	1
	21	0.065	1	0.06	1	0.034	1	0.034	1
	25	0.045	1	0.051	1	0.024	1	0.032	1

Table 2. $H_0 : AR(1)$ vs $H_A : ARCH(1)$ $T = 500$

$T = 500$		$\alpha = 0.1$				$\alpha = 0.05$			
		Bootstrap		Normal		Bootstrap		Normal	
a	M	H_0	H_A	H_0	H_A	H_0	H_A	H_0	H_A
0.3	14	0.053	1	0.098	1	0.024	1	0.063	1
	21	0.064	1	0.082	1	0.023	1	0.052	1
	28	0.06	1	0.093	1	0.024	1	0.062	1
	35	0.07	1	0.086	1	0.033	1	0.062	1
0.4	14	0.043	1	0.092	1	0.014	1	0.064	1
	21	0.058	1	0.092	1	0.015	1	0.056	1
	28	0.066	1	0.094	1	0.03	1	0.061	1
	35	0.073	1	0.087	1	0.032	1	0.052	1
0.5	14	0.031	1	0.105	1	0.018	1	0.072	1
	21	0.059	1	0.079	1	0.03	1	0.05	1
	28	0.076	1	0.111	1	0.046	1	0.069	1
	35	0.053	1	0.086	1	0.022	1	0.055	1
0.55	14	0.038	1	0.107	1	0.014	1	0.077	1
	21	0.056	1	0.108	1	0.021	1	0.067	1
	28	0.071	1	0.103	1	0.032	1	0.06	1
	35	0.051	1	0.089	1	0.026	1	0.06	1

Table 3. $H_0 : ARCH(1)$ vs $H_A : AR(1)$ $T = 100$

$T = 100$		$\alpha = 0.1$				$\alpha = 0.05$			
		Bootstrap		Normal		Bootstrap		Normal	
a	M	H_0	H_A	H_0	H_A	H_0	H_A	H_0	H_A
0.3	11	0.039	0.994	0.08	0.997	0.022	0.984	0.051	0.995
	16	0.043	0.978	0.086	0.991	0.009	0.925	0.055	0.983
	21	0.045	0.98	0.07	0.99	0.016	0.934	0.051	0.983
	25	0.026	0.939	0.059	0.976	0.011	0.895	0.045	0.965
0.4	11	0.046	1	0.086	1	0.012	0.999	0.053	1
	16	0.049	0.993	0.092	0.999	0.014	0.988	0.062	0.996
	21	0.03	0.994	0.07	0.997	0.017	0.983	0.046	0.997
	25	0.038	0.994	0.083	0.997	0.024	0.982	0.059	0.994
0.5	11	0.054	1	0.107	1	0.024	1	0.067	1
	16	0.063	1	0.098	1	0.03	1	0.066	1
	21	0.051	1	0.083	1	0.022	1	0.061	1
	25	0.028	0.997	0.06	0.998	0.012	0.995	0.043	0.998
0.55	11	0.074	1	0.113	1	0.03	1	0.081	1
	16	0.056	1	0.087	1	0.02	1	0.054	1
	21	0.065	1	0.08	1	0.038	1	0.057	1
	25	0.067	1	0.088	1	0.03	1	0.065	1

Table 4. $H_0 : ARCH(1)$ vs $H_A : AR(1)$ $T = 500$

$T = 500$		$\alpha = 0.1$				$\alpha = 0.05$			
		Bootstrap		Normal		Bootstrap		Normal	
a	M	H_0	H_A	H_0	H_A	H_0	H_A	H_0	H_A
0.3	14	0.072	1	0.09	1	0.025	1	0.059	1
	21	0.062	1	0.094	1	0.032	1	0.059	1
	28	0.067	1	0.097	1	0.024	1	0.062	1
	35	0.076	1	0.101	1	0.026	1	0.073	1
0.4	14	0.045	1	0.097	1	0.022	1	0.059	1
	21	0.075	1	0.105	1	0.03	1	0.077	1
	28	0.06	1	0.111	1	0.024	1	0.07	1
	35	0.085	1	0.12	1	0.041	1	0.086	1
0.5	14	0.053	1	0.129	1	0.032	1	0.079	1
	21	0.1	1	0.121	1	0.054	1	0.082	1
	28	0.111	1	0.124	1	0.071	1	0.085	1
	35	0.066	1	0.117	1	0.029	1	0.075	1
0.55	14	0.099	1	0.143	1	0.047	1	0.104	1
	21	0.074	1	0.119	1	0.042	1	0.083	1
	28	0.078	1	0.11	1	0.037	1	0.072	1
	35	0.082	1	0.119	1	0.037	1	0.085	1

6.2. Real Data

In this section we consider the the Microsoft daily return data discussed in Section 2.1 and the Intel monthly log return data (from January 1973 - December 2003), this was considered in Tsay (2005). In the analysis below we will test whether the GARCH and ARCH models are appropriate for the Microsoft and Intel data, respectively. We will be using the Bartlett window.

A plot of the estimated $\widehat{G}_T$ together with the piece-wise confidence intervals (obtained using the results in Theorem II.1) and the corresponding quantile spectral density of the GARCH(1,1) is given in Figure 7 for the Microsoft data. It is clear from the plot that the GARCH$(1,1)$ model with coefficients evaluated using the maximum likelihood estimator is clearly not the appropriate model for fitting to this data. This observation is further confirmed by the results of out test. Using various values of M ranging from $30-70$, the p-value corresponding to $\mathcal{Q}_T$ is zero both using the normal approximation and also the Bootstrap approximation. Therefore, from our analysis it seems that the GARCH$(1,1)$ is not a suitable model for modelling the Microsoft daily returns from 1986-2003. Studying the the quantile spectral density plots we see why this the reason.

We now consider the second data set, the Intel monthly log returns from 1973 - 2003. Tsay (2005) propose fitting an ARCH(1) model to this data, and maximum likelihood yields the estimators $\mu = 0.0166$, $a_0 = 0.0125$ and $a_1 = 0.363$, where $X_t = \mu + \varepsilon_t$, $\varepsilon_t = \sigma_t Z_t$ and $\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2$. A plot of the estimated $\widehat{G}_T$ with the piece-wise confidence intervals together the quantile spectral density of the ARCH(1) model is given in Figure 8. We observe that the quantile spectral density of the ARCH model lies in the confidence intervals for almost all frequencies. These observations are confirmed by the proposed goodness of fit test. A summary of the results for

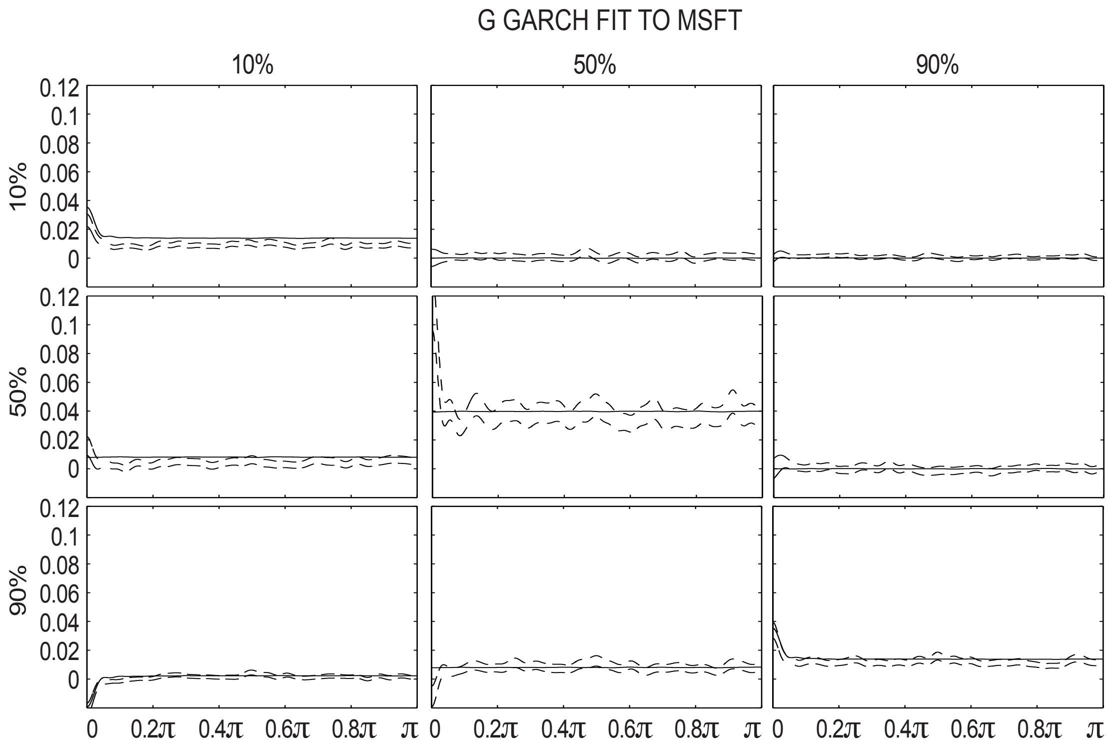

Fig. 7. The quantile spectral density of the fitted GARCH(1, 1) model using Microsoft data with the confidence intervals

various M, using both the normal approximation and the bootstrap method is given in Table 5.

Table 5. The p-values for the Intel Data and various values of M

M	15	20	25	30
Normal p-value	0.0905	0.1279	0.1807	0.2643
Bootstrap p-value	0.3880	0.4320	0.4020	0.4780

The p-values for the normal approximation tend to the smaller than the p-values of the bootstrap method, this is probably due to the skew in the finite sample distribution which results in smaller p-values. However, both the normal approximation

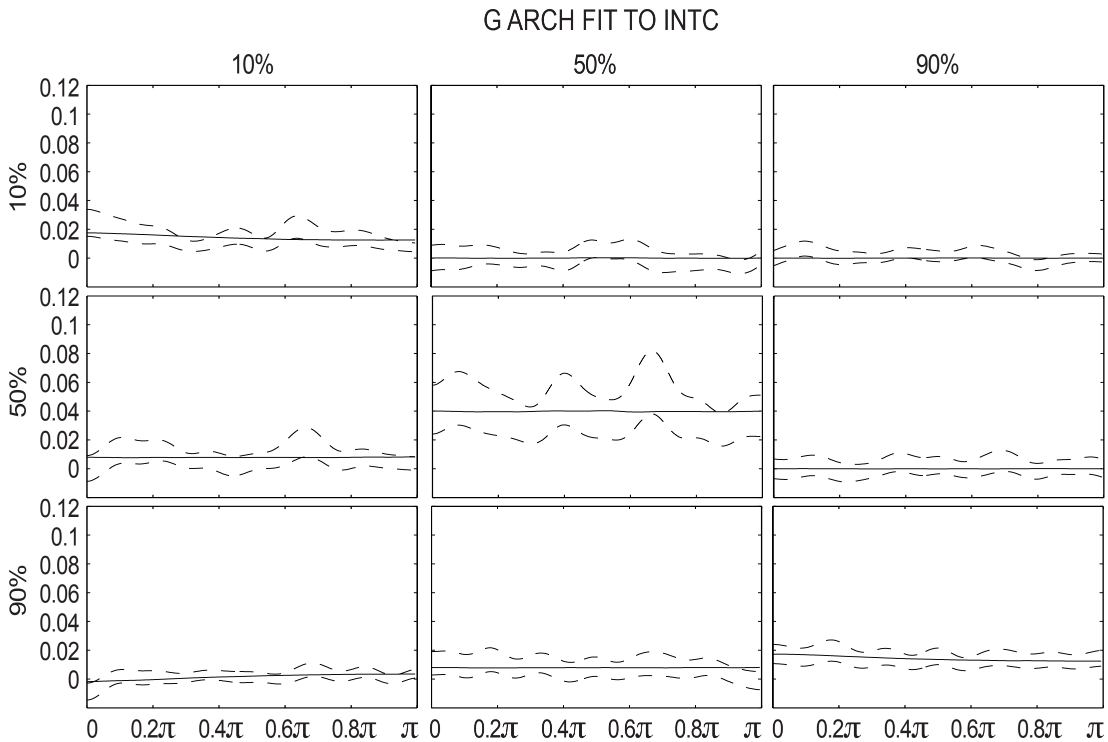

Fig. 8. The quantile spectral density of the fitted ARCH(1) from Intel data with the confidence intervals

and the bootstrap give relatively large p-values for all values of M. Therefore there is not enough evidence to reject the null. This backs the claims in Tsay (2005) that the ARCH(1) may be an appropriate model for the the Intel data.

7. Proofs

7.1. Proof of Theorem II.1

To obtain the sampling properties of $\hat{G}_T(\cdot)$ and $\mathcal{Q}_T$ (under both the null and alternative), we first replace the empirical distribution function $\hat{F}_T(x)$, with the true distribution and show that the error is negligible. Define the zero mean, transformed variable $\tilde{Z}_t(x) = I(X_t \leq x) - F(x)$, where $F(\cdot)$ denotes the marginal distribution of $\{X_t\}$. In addition define $\tilde{C}_r(x,y) = \frac{1}{T}\sum_t \tilde{Z}_t(x)\tilde{Z}_{t+r}(y)$,

$$\tilde{G}_T(x,y;\omega_k) \quad = \quad \frac{1}{2\pi}\sum_r \lambda_M(r)\tilde{C}_r(x,y)\exp(ir\omega_k) = \sum_s K_M(\omega_k - \omega_s)\tilde{J}_T(x;\omega_s)\overline{\tilde{J}_T(y;\omega_s)},$$

$$\tilde{\mathcal{Q}}_T \quad = \quad \frac{1}{T}\sum_{k=1}^{T}\int |\tilde{G}_T(x,y;\omega_k) - \sum_r \lambda_M(r)C_{0,r}(x,y)\exp(ir\omega_k)|^2 dF_0(x)dF_0(y).$$

where $\tilde{J}_T(x;\omega) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} Z_t(x)\exp(it\omega)$.

In the proofs below we shall use the notation $\|X\|_r = (\mathbb{E}(|X|^r))^{1/r}$. We first show that replacing $\hat{F}_T(x)$ with $F(x)$ does not affect the asymptotic sampling properties of $G_T(\cdot)$ and $\mathcal{Q}_T$.

Lemma II.2 *Suppose Assumption IV.1 holds. Then we have*

$$(\mathbb{E}|\widehat{G}_T(x,y;\omega) - \tilde{G}_T(x,y;\omega)\big|^2)^{1/2} = O(\frac{M}{T}) \tag{2.7}$$

and

$$(\mathbb{E}|\mathcal{Q}_T - \tilde{\mathcal{Q}}_T\big|^2)^{1/2} = O(\frac{1}{T}). \tag{2.8}$$

PROOF. We first observe that

$$\begin{aligned} & J_T(x;\omega_k)\overline{J_T(y;\omega_k)} - \tilde{J}_T(x;\omega_k)\overline{\tilde{J}_T(y;\omega_k)} \\ = & \left\{ \begin{array}{cc} 0 & \omega_k \neq 0,\pi \\ T(\hat{F}_T(x) - F(x))(\hat{F}_T(y) - F(y)) & \text{otherwise} \end{array} \right. . \end{aligned}$$

Substituting the above into $\hat{G}_T(\omega_s) - \tilde{G}_T(\omega_s)$ gives

$$\hat{G}_T(\omega_s) - \tilde{G}_T(\omega_s) = TK_M(\omega_s)(\hat{F}_T(x) - F(x))(\hat{F}_T(y) - F(y)). \tag{2.9}$$

Using $K_M(\cdot) = O(\frac{M}{T})$ and $||\hat{F}_T(x) - F(x)||_2 = O(\frac{1}{T})$ in (2.9), we obtain the desired result for (2.7). To prove (2.8) note that

$$
\begin{aligned}
& \mathcal{Q}_T - \tilde{\mathcal{Q}}_T \\
= & \int \frac{1}{T}\sum_{s=1}^{T}\big(\hat{G}_T(x,y;\omega_s) - \tilde{G}_T(x,y;\omega_s)\big)\overline{\big(\hat{G}_T(x,y;\omega_s) + \tilde{G}_T(x,y;\omega_s)\big)}dF_0(x)dF_0(y) \\
+ & \Re\bigg(\int \frac{2}{T}\sum_{s=1}^{T}\big(\hat{G}_T(x,y;\omega_s) - \tilde{G}_T(x,y;\omega_s)\big)G(x,y;\omega_s)dF_0(x)dF_0(y)\bigg).
\end{aligned}
$$

Thus substituting (2.9) into the above gives

$$
\begin{aligned}
& \mathcal{Q}_T - \tilde{\mathcal{Q}}_T \\
= & \int (\widehat{F}_T(x) - F(x))(\widehat{F}_T(y) - F(y)) \times \\
& \bigg(\sum_{s=1}^{T} K_M(\omega_s)\big(\hat{G}_T(x,y;\omega_s) + \tilde{G}_T(x,y;\omega_s)\big)\bigg)dF_0(x)dF_0(y) \\
+ & 2\int (\widehat{F}_T(x) - F(x))(\widehat{F}_T(y) - F(y))\Re\bigg(\sum_{s=1}^{T} K_M(\omega_s)G(x,y;\omega_s)\bigg)dF_0(x)dF_0(y).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
& \big\|\mathcal{Q}_T - \tilde{\mathcal{Q}}_T\big\|_2 \\
\leq & \int \big\|\widehat{F}_T(x) - F(x)\big\|_8 \big\|\widehat{F}_T(y) - F(y)\big\|_8 \times \\
& \bigg(\sum_{s=1}^{T}\big(\big|K_M(\omega_s)\big| \cdot \big(\big\|\hat{G}_T(x,y;\omega_s)\big\|_8 + \|\tilde{G}_T(x,y;\omega_s)\big\|_8\big)\big)\bigg)dF_0(x)dF_0(y) \\
+ & 2\int \big\|\widehat{F}_T(x) - F(x)\big\|_4 \big\|\widehat{F}_T(y) - F(y)\big\|_4 \times \\
& \bigg(\sum_{s=1}^{T}\big|K_M(\omega_s)\big| \cdot \big|G(x,y;\omega_s)\big|\bigg)dF_0(x)dF_0(y).
\end{aligned}
$$

For all $r \geq 2$, we have $\|\widehat{F}_T(x) - F(x)\|_r = O(\frac{1}{\sqrt{T}})$, substituting this into the above gives $\big\|\mathcal{Q}_T - \tilde{\mathcal{Q}}_T\big\|_2 = O(\frac{1}{T})$, and the desired result. □

PROOF of Theorem II.1 To show asymptotic normality of $\widehat{G}_T(\cdot)$, we first replace $\widehat{G}_T$ with $\tilde{G}_T$, by (2.7) the replacement error is $O_p(\frac{M}{T})$. Thus $\widehat{G}_T$ and $\tilde{G}_T$ have the same asymptotic distribution and we can show how asymptotic normality of $\widehat{G}_T$ by considering $\tilde{G}_T(\cdot)$ instead. To show asymptotic normality of $\tilde{G}_T$ we use identical methods to those in Chapter IV, since $\{I(X_t < x)\}$ are bounded random variables, we can use Ibragimov's covariance bounds for bounded random variables. To obtain the limiting variance we note that under Assumption IV.1, since $s > 2$, we have that $\sum_r |r| \cdot |\mathrm{cov}(I(X_0 \leq x), I(X_r \leq y))| < \infty$ and $\sum_{r_1,r_2,r_2}(1 + |r_j|)|\mathrm{cum}(I(X_0 \leq x_0), I(X_{r_1} \leq x_1), I(X_{r_2} \leq x_2), I(X_{r_3} \leq x_3))| < \infty$. Thus, the assumptions in Brillinger (1981), Theorem 3.4.3 are satisfied, which allows us to obtain the stated limiting variance. $\square$

7.2. Proof of Theorem II.2

We use the following lemma to obtain a bound for the variance of $\mathcal{Q}_T$.

7.2.1. Proof of Lemma II.1

Lemma II.3 *Let the lag window be defined as in Definition II.1 and suppose $h_1(\cdot)$ and $h_2(\cdot)$ are bounded functions. Then we have*

$$L_1 = \int h_1(u_1)h_2(u_2)\Delta_M(u_1 - u_2)^2 du_1 du_2 = O(M) \tag{2.10}$$

and

$$L_2 = \int h_1(u_1)h_2(u_2)\Delta_M(u_1 + u_2)\Delta_M(u_1 - u_2) du_1 du_2 = O(1) \tag{2.11}$$

where $\Delta_M(\cdot)$ is defined in (2.4).

PROOF. To simplify notation we prove the result for the truncated lag window $\lambda(u) = I_{[-1,1]}(u)$, but a similar result can also be proven for lag windows which satisfy Definition II.1. In the proof we use the following two identities

$$\sum_{t=0}^{T} e^{it\omega} = e^{\frac{iT\omega}{2}} \frac{\sin(\frac{T+1}{2}\omega)}{\sin(\omega/2)} \quad \text{and} \quad \Big(\int \big|\frac{\sin(\frac{M+1}{2}(u))}{\sin((u)/2)}\big|^{p} du\Big)^{1/p} = O(M^{1-p^{-1}}). \quad (2.12)$$

We start by expanding Δ_M and using the above, to give

$$\begin{aligned}
\Delta_M(\theta_1-\theta_2) &= \int \sum_{j_1,j_2=-M}^{M} \lambda_M(j_1)\lambda_M(j_2)\exp(ij_1(\omega-\theta_1))\exp(ij_2(\omega-\theta_2))d\omega \\
&= \sum_{j} \lambda_M(j)\lambda_M(-j)\exp(ij(\theta_1-\theta_2)) \\
&= \frac{\sin((M+1)(\theta_1-\theta_2)/2)}{\sin((\theta_1-\theta_2)/2)} 2\Re e^{\frac{iM(\theta_1-\theta_2)}{2}}. \qquad (2.13)
\end{aligned}$$

Substituting the above and (2.12) into (2.10) gives

$$\begin{aligned}
|L_1| &= \Big|\int\int h_1(u_1)h_2(u_2)\Delta_M(u_1-u_2)^2 du_1 du_2\Big| \\
&\leq \sup_{u,i} |h_i(u)|^2 \int\int \Big|\frac{\sin(\frac{M+1}{2}(u_1-u_2))}{\sin((u_1-u_2)/2)}\Big|^2 du_1 du_2 \\
&= O(M).
\end{aligned}$$

This proves (2.10). To prove (2.11) we observe that by a change of variables ($v_1 = u_1 - u_2$ and $v_2 = u_1 + u_2$) we have

$$\begin{aligned}
|L_2| &\leq C\int |\Delta_M(u_1+u_2)|\cdot|\Delta_M(u_1-u_2)| du_1 du_2 \\
&\leq C\Big(\int |\Delta_M(u)| du\Big)^2.
\end{aligned}$$

Now by substituting (2.13) and (2.12) into the above gives $L_2 = O(1)$. Thus we have obtained the desired result. $\square$

PROOF of Lemma II.1 We first evaluate the expectation of $\mathcal{Q}_T$. By using Lemma II.2 we have

$$\begin{aligned}
& \mathbb{E}(\mathcal{Q}_T) \\
= & \frac{1}{T}\sum_{s=1}^{T}\int \sum_{k_1,k_2=1}^{T} K_M(\omega_s-\omega_{k_1})K_M(\omega_s-\omega_{k_2}) \times \\
& \mathrm{cov}\big(\tilde{J}_T(x;\omega_{k_1})\overline{\tilde{J}_T(y;\omega_{k_1})}, \tilde{J}_T(x;\omega_{k_2})\overline{\tilde{J}_T(y;\omega_{k_2})}\big) + O(\frac{1}{T}) \\
= & I_1 + I_2 + I_3 + O(\frac{1}{T}),
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{T}\int \sum_{s,k_1,k_2=1}^{T} \Big(K_M(\omega_s-\omega_{k_1})K_M(\omega_s-\omega_{k_2}) \times \\
& \mathrm{cov}\big(\tilde{J}_T(x;\omega_{k_1}), \tilde{J}_T(x;\omega_{k_2})\big)\mathrm{cov}\big(\overline{\tilde{J}_T(y;\omega_{k_1})}, \overline{\tilde{J}_T(y;\omega_{k_2})}\big)\Big) dF_0(x)dF_0(y) \\
I_2 &= \frac{1}{T}\int \sum_{s,k_1,k_2=1}^{T} \Big(K_M(\omega_s-\omega_{k_1})K_M(\omega_s-\omega_{k_2}) \times \\
& \mathrm{cov}\big(\tilde{J}_T(x;\omega_{k_1}), \overline{\tilde{J}_T(y;\omega_{k_2})}\big)\mathrm{cov}\big(\overline{\tilde{J}_T(y;\omega_{k_1})}, \tilde{J}_T(x;\omega_{k_2})\big)\Big) dF_0(x)dF_0(y) \\
I_3 &= \frac{1}{T}\int \sum_{s,k_1,k_2=1}^{T} \Big(K_M(\omega_s-\omega_{k_1})K_M(\omega_s-\omega_{k_2}) \times \\
& \mathrm{cum}\big(\tilde{J}_T(x;\omega_{k_1}), \overline{\tilde{J}_T(y;\omega_{k_1})}, \tilde{J}_T(x;\omega_{k_2}), \overline{\tilde{J}_T(y;\omega_{k_2})}\big)\Big) dF_0(x)dF_0(y).
\end{aligned}$$

Under Assumption IV.1, we have that $\sum_r |r|\cdot|\mathrm{cov}(I(X_0\le x), I(X_r\le y))| < \infty$ and $\sum_{r_1,r_2,r_2}(1+|r_j|)|\mathrm{cum}(I(X_0 \le x_0), I(X_{r_1} \le x_1), I(X_{r_2} \le x_2), I(X_{r_3} \le x_3))| < \infty$.

Therefore we can apply Brillinger (1981), Theorem 3.4.3 to obtain

$$\begin{aligned}
I_1 &= \frac{1}{T}\sum_{s=1}^{T}\int\sum_{k=1}^{T}K_M(\omega_s-\omega_k)^2 G(x,x;\omega_k)G(y,y;\omega_k)dF_0(x)dF_0(y)+O(\frac{1}{T}) \\
&= O(\frac{M}{T}) \\
I_2 &= \frac{1}{T}\sum_{s=1}^{T}\int\sum_{k=1}^{T}K_M(\omega_s-\omega_k)K_M(\omega_s+\omega_k)|G(x,y;\omega_k)|^2 dF_0(x)dF_0(y)+O(\frac{1}{T}) \\
&= O(\frac{1}{T}) \\
I_3 &= \frac{1}{T^2}\int\sum_{r}\lambda_M(r)^2\sum_{t_1,t_2=1}^{T}\mathrm{cum}\big(Z_{t_1}(x),Z_{t_1+r}(y),Z_{t_2}(x),Z_{t_2+r}(y)\big)dF_0(x)dF_0(y) \\
&= O(\frac{1}{T}).
\end{aligned}$$

This gives us an asymptotic expression for the expectation. We now obtain an expression for the variance. Replacing $Z_t(\cdot)$ with $\tilde{Z}_t(\cdot)$ gives

$$\begin{aligned}
&\mathrm{var}(\mathcal{Q}_T) = \\
&\frac{1}{T^2}\sum_{s_1,s_2=1}^{T}\int\Bigg(\sum_{k_1,k_2,k_3,k_4}K_M(\omega_{s_1}-\omega_{k_1})K_M(\omega_{s_1}-\omega_{k_2})K_M(\omega_{s_2}-\omega_{k_3})K_M(\omega_{s_2}-\omega_{k_4}) \\
&\times \mathrm{cov}\big((J_{k_1,x_1}\overline{J}_{k_1,y_1}-\mathbb{E}(J_{k_1,x_1}\overline{J}_{k_1,y_1}))(J_{k_2,x_1}\overline{J}_{k_2,y_1}-\mathbb{E}(J_{k_2,x_1}\overline{J}_{k_2,y_1})), \\
&(J_{k_3,x_2}\overline{J}_{k_3,y_2}-\mathbb{E}(J_{k_3,x_2}\overline{J}_{k_3,y_2}))(J_{k_4,x_2}\overline{J}_{k_4,y_2}-\mathbb{E}(J_{k_4,x_2}\overline{J}_{k_4,y_2}))\big)\Bigg) \\
&dF_0(x_1)dF_0(y_1)dF_0(x_2)dF_0(y_2)+O(\frac{1}{T}) \\
&= II_1+II_2+II_3+O(\frac{1}{T})
\end{aligned}$$

where $J_{k,x}=\tilde{J}_T(x;\omega_k)$,

$$
\begin{aligned}
II_1 &= \frac{1}{T^2}\sum_{s_1,s_2=1}^{T}\int\sum_{k_1,k_2,k_3,k_4}\mathrm{cum}(J_{k_1,x_1}\bar{J}_{k_1,y_1},\bar{J}_{k_3,x_2}J_{k_3,y_2})\mathrm{cum}(J_{k_2,x_1}\bar{J}_{k_2,y_1},\overline{J}_{k_4,x_2}J_{k_4,y_2}) \\
&\quad \prod_{i=1}^{2}K_M(\omega_{s_1}-\omega_{k_i})\prod_{i=3}^{4}K_M(\omega_{s_2}-\omega_{k_i})dF_0(x_1)dF_0(y_1)dF_0(x_2)dF_0(y_2) \\
II_2 &= \frac{1}{T^2}\sum_{s_1,s_2=1}^{T}\int\sum_{k_1,k_2,k_3,k_4}\mathrm{cum}(J_{k_1,x_1}\bar{J}_{k_1,y_1},\overline{J}_{k_4,x_2}J_{k_4,y_2})\mathrm{cum}(J_{k_2,x_1}\bar{J}_{k_2,y_1},\bar{J}_{k_3,x_2}J_{k_3,y_2}) \\
&\quad \prod_{i=1}^{2}K_M(\omega_{s_1}-\omega_{k_i})\prod_{i=3}^{4}K_M(\omega_{s_2}-\omega_{k_i})dF_0(x_1)dF_0(y_1)dF_0(x_2)dF_0(y_2) \\
II_3 &= \frac{1}{T^2}\sum_{s_1,s_2=1}^{T}\int\sum_{k_1,k_2,k_3,k_4}\mathrm{cum}(J_{k_1,x_1}\bar{J}_{k_1,y_1},J_{k_2,x_1}\bar{J}_{k_2,y_1},\bar{J}_{k_3,x_2}J_{k_3,y_2},\overline{J}_{k_4,x_2}J_{k_4,y_2}) \\
&\quad \prod_{i=1}^{2}K_M(\omega_{s_1}-\omega_{k_i})\prod_{i=3}^{4}K_M(\omega_{s_2}-\omega_{k_i})dF_0(x_1)dF_0(y_1)dF_0(x_2)dF_0(y_2).
\end{aligned}
$$

To obtain an expression for the variance we start by expanding II_1

$$\begin{aligned}
II_1 \quad &= \frac{1}{T^2}\sum_{s_1,s_2}\int \sum_{k_1,k_2,k_3,k_4}\prod_{i=1}^{2}K_M(\omega_{s_1}-\omega_{k_i})\prod_{i=3}^{4}K_M(\omega_{s_2}-\omega_{k_i}) \\
&\times \Big(\mathrm{cov}(J_{k_1,x_1},J_{k_3,x_2})\mathrm{cov}(\overline{J}_{k_1,y_1},\overline{J}_{k_3,y_2})\mathrm{cov}(J_{k_2,x_1},J_{k_4,x_2})\mathrm{cov}(\overline{J}_{k_2,y_1},\overline{J}_{k_4,y_2}) \\
&+ \mathrm{cov}(J_{k_1,x_1},J_{k_3,x_2})\mathrm{cov}(\overline{J}_{k_1,y_1},\overline{J}_{k_3,y_2})\mathrm{cov}(J_{k_2,x_1},\overline{J}_{k_4,y_2})\mathrm{cov}(\overline{J}_{k_2,y_1},J_{k_4,x_2}) \\
&+ \mathrm{cov}(J_{k_1,x_1},J_{k_3,x_2})\mathrm{cov}(\overline{J}_{k_1,y_1},\overline{J}_{k_3,y_2})\mathrm{cum}(J_{k_2,x_1},\overline{J}_{k_2,y_1},J_{k_4,x_2},\overline{J}_{k_4,y_2}) \\
&+ \mathrm{cov}(J_{k_1,x_1},\overline{J}_{k_3,y_2})\mathrm{cov}(\overline{J}_{k_1,y_1},J_{k_3,x_2})\mathrm{cov}(J_{k_2,x_1},J_{k_4,x_2})\mathrm{cov}(\overline{J}_{k_2,y_1},\overline{J}_{k_4,y_2}) \\
&+ \mathrm{cov}(J_{k_1,x_1},\overline{J}_{k_3,y_2})\mathrm{cov}(\overline{J}_{k_1,y_1},J_{k_3,x_2})\mathrm{cov}(J_{k_2,x_1},\overline{J}_{k_4,y_2})\mathrm{cov}(\overline{J}_{k_2,y_1},J_{k_4,x_2}) \\
&+ \mathrm{cov}(J_{k_1,x_1},\overline{J}_{k_3,y_2})\mathrm{cov}(\overline{J}_{k_1,y_1},J_{k_3,x_2})\mathrm{cum}(J_{k_2,x_1},\overline{J}_{k_2,y_1},J_{k_4,x_2},\overline{J}_{k_4,y_2}) \\
&+ \mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1},J_{k_3,x_2},\overline{J}_{k_3,y_2})\mathrm{cov}(J_{k_2,x_1},J_{k_4,x_2})\mathrm{cov}(\overline{J}_{k_2,y_1},\overline{J}_{k_4,y_2}) \\
&+ \mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1},J_{k_3,x_2},\overline{J}_{k_3,y_2})\mathrm{cov}(J_{k_2,x_1},\overline{J}_{k_4,y_2})\mathrm{cov}(\overline{J}_{k_2,y_1},J_{k_4,x_2}) \\
&+ \mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1},J_{k_3,x_2},\overline{J}_{k_3,y_2})\mathrm{cum}(J_{k_2,x_1},\overline{J}_{k_2,y_1},J_{k_4,x_2},\overline{J}_{k_4,y_2})\Big)\prod_{j=1}^{2}dF_0(x_j)dF_0(y_j) \\
&:= \sum_{j=1}^{9} II_{1,j}.
\end{aligned}$$

We use Brillinger (1981), Theorem 3.4.3 to obtain the following expression for $II_{1,1}$

$$\begin{aligned}
II_{1,1} = & \frac{1}{T^2}\sum_{s_1,s_2}\int\Big(\sum_{k_1,k_2=1}^{T}(\prod_{i=1}^{2}K_M(\omega_{s_i}-\omega_{k_1})K_M(\omega_{s_i}-\omega_{k_2}))\\
& \mathrm{cov}(J_{k_1,x_1},J_{k_1,x_2})\mathrm{cov}(\overline{J}_{k_1,y_1},\overline{J}_{k_1,y_2})\mathrm{cov}(J_{k_2,x_1},J_{k_2,x_2})\mathrm{cov}(\overline{J}_{k_2,y_1},\overline{J}_{k_2,y_2})\Big)\\
& \prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2})\\
= & \frac{1}{T^2}\int\int\Big(\int W_M(\omega_{s_1}-\theta_1)W_M(\omega_{s_1}-\theta_2)d\omega_{s_1}\Big)\times\\
& \Big(\int W_M(\omega_{s_2}-\theta_1)W_M(\omega_{s_2}-\theta_2)d\omega_{s_2}\Big)\prod_{i=1}^{2}G(x_1,x_2;\theta_i)G(y_1,y_2;-\theta_i)d\theta_i\\
& \prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2})\\
= & \frac{1}{T^2}\int\int\Delta_M(\theta_1-\theta_2)^2\prod_{i=1}^{2}G(x_1,x_2;\theta_i)G(y_1,y_2;-\theta_i)d\theta_i\prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2}).
\end{aligned}$$

Therefore by using (2.10) we have $II_{1,1}=O(\frac{M}{T^2})$. We now consider $II_{1,2}$, by using a similar method we have

$$\begin{aligned}
II_{1,2} &= \frac{1}{T^2}\int W_M(\omega_{s_1}-\theta_1)W_M(\omega_{s_1}-\theta_2)W_M(\omega_{s_2}-\theta_1)W_M(\omega_{s_2}+\theta_2)\\
&\quad G(x_1,x_2,\theta_1)G(y_1,y_2,-\theta_1)G(x_1,y_2,\theta_2)G(y_1,x_2,-\theta_2)d\theta_1 d\theta_2 d\omega_{s_1}d\omega_{s_2}\\
&\quad \prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2})\\
&= \frac{1}{T^2}\int\Delta_M(\theta_1-\theta_2)\Delta_M(\theta_1+\theta_2)G(x_1,x_2,\theta_1)G(y_1,y_2,-\theta_1)\\
&\quad G(x_1,y_2,\theta_2)G(y_1,x_2,-\theta_2)d\theta_1 d\theta_2\prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2}).
\end{aligned}$$

By using (2.11) the above integral is $O(1)$, and altogether $II_{1,2}=O(\frac{1}{T^2})$. Using a similar argument, one can show that $II_{1,3}$, $II_{1,4}$ are smaller than $O(\frac{M}{T^2})$, so negligible.

For $II_{1,5}$, we use that

$$\mathrm{cov}(J_{k_1,x}, \overline{J}_{k_2,y}) = \begin{cases} G(x,y,\omega_{k_1}) & k_1+k_2=T \\ O(\frac{1}{T}) & otherwise \end{cases}$$

which follows from Brillinger (1981), Theorem 3.4.3. This leads to

$$\begin{aligned} II_{1,5} &= \frac{1}{T^2}\sum_{s_1,s_2}\int\Bigg(\sum_{k_1,k_2=1}^{T}(\prod_{i=1}^{2}K_M(\omega_{s_1}-\omega_{k_i})K_M(\omega_{s_2}+\omega_{k_i})) \\ &\quad \mathrm{cov}(J_{k_1,x_1}, J_{k_1,y_2})\mathrm{cov}(\overline{J}_{k_1,y_1}, \overline{J}_{k_1,x_2})\mathrm{cov}(J_{k_2,x_1}, J_{k_2,y_2})\mathrm{cov}(\overline{J}_{k_2,y_1}, \overline{J}_{k_2,x_2})\Bigg) \\ &\quad \prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2}) \\ &= \frac{1}{T^2}\int\int\left(\int W_M(\omega_{s_1}-\theta_1)W_M(\omega_{s_1}-\theta_2)d\omega_{s_1}\right)\times \\ &\quad \left(\int W_M(\omega_{s_2}+\theta_1)W_M(\omega_{s_2}+\theta_2)d\omega_{s_2}\right)\prod_{i=1}^{2}G(x_1,y_2;\theta_i)G(y_1,x_2;-\theta_i)d\theta_i \\ &\quad \prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2}) \\ &= II_{1,1} \end{aligned}$$

because of $\Delta(\theta) = \Delta(-\theta)$ and interchangeability of integrals about (x_1, x_2, y_1, y_2). With a similar method, one can show that $II_{1,6}\ldots, II_{1,9}$ are all dominated by $II_{1,1}$ and $II_{1,5}$ Altogether this gives

$$II_1 = \frac{2}{T^2}\int \Delta_M(\theta_1-\theta_2)^2\prod_{i=1}^{2}G(x_1,x_2;\theta_i)G(y_1,y_2;\theta_i)d\theta_i\prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2}).$$

Using the identical argument with the above, we can show that

$$II_2 = \frac{2}{T^2}\int \Delta_M(\theta_1-\theta_2)^2\prod_{i=1}^{2}G(x_1,x_2;\theta_i)G(y_1,y_2;\theta_i)d\theta_i\prod_{j=1}^{2}dF_0(x_j)dF_0(y_j)+O(\frac{1}{T^2}).$$

To bound II_3 we recall that

$$II_3 = \frac{1}{T^2}\sum_{s_1,s_2=1}^{T}\int\sum_{k_1,k_2,k_3,k_4} K_M(\omega_{s_1}-\omega_{k_1})K_M(\omega_{s_1}-\omega_{k_2})K_M(\omega_{s_2}-\omega_{k_3})K_M(\omega_{s_2}-\omega_{k_4})$$
$$\mathrm{cum}\big(J_{k_1,x_1}\overline{J}_{k_1,y_1}, J_{k_2,x_1}\overline{J}_{k_2,y_1}, J_{k_3,x_2}\overline{J}_{k_3,y_2}, J_{k_4,x_2}\overline{J}_{k_4,y_2}\big)\prod_{j=1}^{2} dF_0(x_j)dF_0(y_j) + O(\frac{1}{T}).$$

The above cumulant is computed as the sum of the products of cumulants in decomposable partitions by Theorem 2.3.2 in Brillinger (1981). We used the mathematica routine by Andrews and Stafford (1998) to find all decomposable partitions. Further information about the indecomposable partitioning could be found in Andrews and Stafford (1998), Stafford (1994) and Smith and Field (2001). This together with Brillinger (1981), Theorem 3.4.3 gives us $II_3 = O(\frac{M}{T^3})$. The detail is given in the Appendix. Combining the expressions for II_1, II_2 and II_3 gives us the expression for the variance and completes the proof. □

7.2.2. Proof of Theorem II.2

Now we show that $\mathcal{Q}_T$ can be approximated by the sum of martingale differences, this will allow us to the the martingale central limit theorem to prove Theorem II.2. We first define the martingale difference decomposition of $\tilde{Z}_t(x) = \sum_{j=0}^{\infty} M_j^{(x)}(t-j)$, where $M_j^{(x)}(t-j) = \mathbb{E}(\tilde{Z}_t(x)|\mathcal{F}_{t-j}) - \mathbb{E}(\tilde{Z}_t(x)|\mathcal{F}_{t-j-1})$, where for $t > 0$ we have $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots, X_1)$ and for $t \leq 0$ we let $\mathcal{F}_t = \sigma(1)$, and $M_j(s) = 0$ for $j \geq s$. Using the above notation we define the random variable

$$\begin{aligned}\mathcal{S}_T &= \frac{1}{T^2}\int\sum_{j_1,\ldots,j_4=0}^{\infty}\sum_{t_1,r,t_2\in\mathcal{A}} \lambda_M(r)^2 M_{j_1}^{(x)}(t_1-j_1)M_{j_2}^{(y)}(t_1+r-j_2) \\ &\quad \times M_{j_3}^{(x)}(t_2-j_3)M_{j_4}^{(y)}(t_2+r-j_4)dF_0(x)dF_0(y), \qquad (2.14)\end{aligned}$$

where $\mathcal{A} = \{(t_1-j_1, t_1+r-j_2, t_2-j_3, t_2+r-j_4) \text{ are all different }\}$.

Theorem II.4 *Suppose Assumption IV.1 holds, $\mathcal{S}_T$ is defined as in (2.14) and the null hypothesis is true. Then we have*

$$\mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T) \;=\; \mathcal{S}_T + O_p\Big(\frac{1}{T} + \frac{M^{1/2}}{T^{3/2}}\Big).$$

and for all $r \geq 2$

$$\big\|\mathcal{S}_T\big\|_r = O\Big(\frac{M^{1/2}}{T}\Big).$$

PROOF. We use a combination of iterative martingales and Burkholder's inequality for martingale differences. First we note that for $r \geq 2$ we have

$$\begin{aligned}
\|M_j^{(x)}(t-j)\|_r &= \|\mathbb{E}(\tilde{Z}_t(x)|\mathcal{F}_{t-j}) - \mathbb{E}(\tilde{Z}_t(x)|\mathcal{F}_{t-j-1})\|_2 \\
&\leq 2\|\mathbb{E}(\tilde{Z}_t(x)|\mathcal{F}_{t-j})\|_r \leq C\alpha(j),
\end{aligned} \qquad (2.15)$$

where $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots, X_1)$, which follows from Ibragimov's inequality. Substituting the representation $\tilde{Z}_t(x) = \sum_{j=1}^{\infty} M_j^{(x)}(t-j)$ into $\mathcal{Q}_T$ gives

$$\begin{aligned}
& \mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T) \\
= \;& \frac{1}{T^2}\int \sum_{j_1,\ldots,j_4=0}^{\infty} \sum_{r=-M}^{M} \lambda_M(r)^2 \\
& \sum_{t_1,t_2} \Big(\overline{M_{j_1}^{(x)}(t_1-j_1)M_{j_2}^{(y)}(t_1+r-j_2)} \times \overline{M_{j_3}^{(x)}(t_2-j_3)M_{j_4}^{(y)}(t_2+r-j_4)} \\
- \;& \mathbb{E}\big(\overline{M_{j_1}^{(x)}(t_1-j_1)M_{j_2}^{(y)}(t_1+r-j_2)} \times \overline{M_{j_3}^{(x)}(t_2-j_3)M_{j_4}^{(y)}(t_2+r-j_4)}\big)\Big) \\
& dF_0(x)dF_0(y),
\end{aligned}$$

where $\overline{X}$ denotes the centralised random variable $X - \mathbb{E}(X)$ (note that $M_j(s) = 0$ for $s \leq 0$). We now partition the above sum into several cases, where we treat $j_1, \ldots, j_4$ as free and condition on t_1, t_2 and r:

(i) $\mathcal{A} = \{(t_1, t_2, r)$ such that $(t_1 - j_1, t_1 + r - j_2, t_2 - j_3, t_2 + r - j_4)$ are all different

}.

(ii) $\mathcal{B} = \{(t_1, t_2, r)$ such that $(t_1 - j_1 = t_1 + r - j_2)$ and $(t_2 - j_3 = t_2 + r - j_4)\}$.

(iii) $\mathcal{C} = \{(t_1, t_2, r)$ such that $(t_1 - j_1) = (t_2 - j_3)$ or $(t_2 + r - j_4)$ and $(t_1 + r - j_2) \neq (t_1 - j_1)\}$.

(iv) $\mathcal{D} = \{(t_1, t_2, r)$ such that $(t_1 + r - j_2) = (t_2 - j_3)$ or $(t_2 + r - j_4)$ and $(t_1 - j_1) \neq (t_1 + r - j_2)\}$.

(v) $\mathcal{E} = \{(t_1, t_2, r)$ such that $(t_2 - j_3) = (t_1 - j_1)$ or $(t_1 + r - j_2)$ and $(t_2 + r - j_4 \neq t_2 - j_3)\}$.

(iv) $\mathcal{F} = \{(t_1, t_2, r)$ such that $(t_2 + r - j_4) = (t_1 - j_2)$ or $(t_1 + r - j_2)$ and $(t_2 - j_3) \neq (t_2 + r - j_4)\}$.

Thus

$$\mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T) = \int \Big(I_{\mathcal{A}} + I_{\mathcal{B}} + I_{\mathcal{C}} + I_{\mathcal{D}} + I_{\mathcal{E}} + I_{\mathcal{F}} \Big) dF_0(x) dF_0(y),$$

where

$$\begin{aligned}
I_{\mathcal{A}} &= \frac{1}{T^2} \sum_{j_1,\ldots,j_4=0}^{\infty} \sum_{r,t_1,t_2 \in \mathcal{A}} \lambda_M(r)^2 \\
&\quad M_{j_1}^{(x)}(t_1 - j_1) M_{j_2}^{(y)}(t_1 + r - j_2) M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 + r - j_4), \\
I_{\mathcal{B}} &= \frac{1}{T^2} \sum_{j_1,\ldots,j_4=0}^{\infty} \sum_{r,t_1,t_2 \in \mathcal{B}} \lambda_M(j_1 - j_2)^2 \\
&\quad \Big(\overline{M_{j_1}^{(x)}(t_1 - j_1) M_{j_2}^{(y)}(t_1 - j_1)} \times \overline{M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 - j_3)} - \\
&\quad \mathbb{E}\big(\overline{M_{j_1}^{(x)}(t_1 - j_1) M_{j_2}^{(y)}(t_1 - j_1)} \times \overline{M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 - j_3)} \big) \Big)
\end{aligned}$$

for $I_{\mathcal{C}}, \ldots, I_{\mathcal{F}}$ are defined similarly.

We first bound $I_{\mathcal{A}}$. We partition $\mathcal{A}$ into 24 cases by the order of $(t_1 - j_1, t_1 + r - j_2, t_2 - j_3, t_2 + r - j_4)$. The first is $\mathcal{A}_1 = \{(t_1, t_2, r)$ such that $t_1 - j_1 > t_1 + r - j_2 > t_2 - j_3 > t_2 + r - j_4\}$ which gives

$$\begin{aligned} I_{\mathcal{A},1} &= \frac{1}{T^2} \sum_{j_1,\ldots,j_4=0}^{\infty} \sum_{t_1,t_2 \in \mathcal{A}_1} \lambda_M(r)^2 \\ & M_{j_1}^{(x)}(t_1 - j_1) M_{j_2}^{(y)}(t_1 + r - j_2) M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 + r - j_4). \end{aligned}$$

The other 23 cases are defined similarly such that we have $I_{\mathcal{A}} = \sum_{j=1}^{24} I_{\mathcal{A}_j}$. We start by bounding $I_{\mathcal{A},1}$. Since $t_1 - j_1 > t_1 + r - j_2 > t_2 - j_3 > t_2 + r - j_4$, it is easy to see that $M_{j_1}^{(x)}(t_1 - j_1) \sum_{r < j_2 - j_1} \lambda_M^2(r) M_{j_2}^{(y)}(t_1 + r - j_2) \sum_{t_2 < t_1 - j_1 + j_3} M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 + r - j_4)$ is a martingale over t_1, $M_{j_2}^{(y)}(t_1 + r - j_2) \sum_{t_2 < t_1 - j_1 + j_3} M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 + r - j_4)$ is a martingale over r and $\{M_{j_3}^{(x)}(t_2 - j_3) M_{j_4}^{(y)}(t_2 + r - j_4)\}$ is a martingale over t_2. Thus by using Burkhölder's inequality together with Hölder's inequality three times, for any $q \geq 2$ we have

$$\begin{aligned} \|I_{\mathcal{A},1}\|_q &= \frac{1}{T^2} \sum_{j_1,\ldots,j_4=0}^{\infty} \Big(\sum_{r,t_1,t_2} \lambda_M(r)^2 \|M_{j_1}^{(x)}(t_1 - j_1)\|_{4q}^2 \|M_{j_2}^{(y)}(t_1 + r - j_2)\|_{4q}^2 \\ & \|M_{j_3}^{(x)}(t_2 - j_3)\|_{4q}^2 \|M_{j_4}^{(y)}(t_2 + r - j_4)\|_{4q}^2 \Big)^{1/2}. \end{aligned}$$

Thus by using (2.15) we have that $\|I_{\mathcal{A},1}\|_q = O(\frac{M^{1/2}}{T})$ and by the same argument we have $I_{\mathcal{A},j} = O(\frac{M^{1/2}}{T})$ (for $2 \leq j \leq 24$). Therefore, altogether this gives $\|I_{\mathcal{A}}\|_q = O(\frac{M^{1/2}}{T})$. We now bound $I_{\mathcal{B}}$. We first define the random variable

$$\begin{aligned} & A_{j_1,j_2;i}^{(x,y)}(t_1 - j_1 - i) = \\ & \mathbb{E}(M_{j_1}^{(x)}(t_1 - j_1) M_{j_2}^{(y)}(t_1 - j_1) | \mathcal{F}_{t_1 - j_1 - i}) - \mathbb{E}(M_{j_1}^{(x)}(t_1 - j_1) M_{j_2}^{(y)}(t_1 - j_1) | \mathcal{F}_{t_1 - j_1 - i}). \end{aligned}$$

To bound $\|A^{(x,y)}_{j_1,j_2;i}(t_1-j_1-i)\|_q$, we repeatedly use Ibragimov's inequality and (2.15) to give

$$\begin{aligned}\|A^{(x,y)}_{j_1,j_2;i}(t_1-j_1-i)\|_q &\leq 2\|\mathbb{E}(\overline{M^{(x)}_{j_1}(t_1-j_1)M^{(y)}_{j_2}(t_1-j_1)}|\mathcal{F}_{t_1-j_1-i})\| \\ &\leq C\alpha(i)\|M^{(x)}_{j_1}(t_1-j_1)M^{(y)}_{j_2}(t_1-j_1)\|_q \\ &\leq C\alpha(i)\alpha(j_1)\alpha(j_2). \qquad (2.16)\end{aligned}$$

This gives the representation

$$\overline{M^{(x)}_{j_1}(t_1-j_1)M^{(y)}_{j_2}(t_1-j_1)} = \sum_i A^{(x,y)}_{j_1,j_2;i}(t_1-j_1-i).$$

Substituting the above representation into $I_{\mathcal{B}}$ gives

$$\begin{aligned}I_{\mathcal{B}} &= \frac{1}{T^2}\sum_{\substack{j_1,\ldots,j_4,\\ i_1,i_2=0}}^{\infty}\sum_{t_1,t_2\in\mathcal{B}}\lambda_M(j_1-j_2)^2\big[A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2) \\ &\quad -\mathbb{E}(A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2))\big] \\ &:= I_{\mathcal{B},1}+I_{\mathcal{B},2}+I_{\mathcal{B},3},\end{aligned}$$

where

$$\begin{aligned}I_{\mathcal{B},1} &:= \frac{1}{T^2}\sum_{\substack{j_1,\ldots,j_4,\\ i_1,i_2=0}}^{\infty}\sum_{\substack{t_1-j_1-i_1>\\ t_2-j_3-i_2}}\lambda_M(j_1-j_2)^2A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2) \\ I_{\mathcal{B},2} &:= \frac{1}{T^2}\sum_{\substack{j_1,\ldots,j_4,\\ i_1,i_2=0}}^{\infty}\sum_{\substack{t_1-j_1-i_1<\\ t_2-j_3-i_2}}\lambda_M(j_1-j_2)^2A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2) \\ I_{\mathcal{B},3} &:= \frac{1}{T^2}\sum_{\substack{j_1,\ldots,j_4,\\ i_1,i_2=0}}^{\infty}\sum_{\substack{t_1-j_1-i_1=\\ t_2-j_3-i_2}}\lambda_M(j_1-j_2)^2\overline{A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2)}.\end{aligned}$$

Using similar techniques to those used to bound $\|I_{\mathcal{A},1}\|_q$, Burkhölder's and Hölder's inequalities twice on $\|I_{\mathcal{B},1}\|_q$, together with (2.16), we obtain the bound $\|I_{\mathcal{B},1}\|_q = O(\frac{1}{T})$. A similar argument can be used for $\|I_{\mathcal{B},2}\|_q = O(\frac{1}{T})$. To bound $\|I_{\mathcal{B},3}\|_q$, we need to decompose

$$A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2)-\mathbb{E}(A^{(x,y)}_{j_1,j_2;i_1}(t_1-j_1-i)A^{(x,y)}_{j_3,j_4;i_2}(t_2-j_3-i_2)),$$

into the sum of martingale differences, using this martingale decomposition we can use the same argument as those used above to obtain $\|I_{\mathcal{B},3}\| = O(\frac{1}{T^{3/2}})$. Therefore, altogether we have $\|I_{\mathcal{B}}\|_q = O(\frac{1}{T})$. Now by using similar arguments and repeated decompositions into martingale differences we can show that $\|I_{\mathcal{C}}\|_q,\ldots,\|I_{\mathcal{F}}\|_q = O(\frac{M^{1/2}}{T^{3/2}})$. Thus we have shown that $I_{\mathcal{A}}$ is the dominating term in $\mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T)$. Since $\mathcal{S}_T = \int I_{\mathcal{A}} dF_0(x)dF_0(y)$ we have obtained the desired result. $\square$

To prove the asymptotic normality of $\mathcal{Q}_T$ under the null hypothesis, we use the martingale central limit theorem on $\mathcal{S}_T$ in (2.14). To do this, we use the same decompositions of $I_{\mathcal{A}}$, as that used in the proof of Theorem II.4. We set $\mathcal{S}_{T,i} := I_{\mathcal{A},i}$, recalling that

$$\begin{aligned}\mathcal{S}_{T,i} &= \frac{1}{T^2}\int\sum_{j_1,\ldots,j_4=0}^{\infty}\sum_{r,t_1,t_2\in\mathcal{A},i}\lambda_M(r)^2\\ &\quad M^{(x)}_{j_1}(t_1-j_1)M^{(y)}_{j_2}(t_1+r-j_2)M^{(x)}_{j_3}(t_2-j_3)M^{(y)}_{j_4}(t_2+r-j_4)dF_0(x)dF_0(y),\end{aligned}$$

where $\mathcal{A}_i$ is some ordering of $\{t_1-j_1, t_1+r-j_2, t_2-j_3, t_2+r-j_4\}$. We show that $\mathcal{S}_{T,i}$ can be written as the sum of martingale differences. First consider $\mathcal{S}_{T,1}$, this can be written as $\mathcal{S}_{T,1} = \frac{1}{T^2}\sum_{k=1}^{T} U_{k,1}$, where with a change of variables we have

$$\begin{aligned}U_{k,1} &= \int\sum_{j_1=0}^{T-k}M_{j_1}(k)\sum_{j_2,j_3,j_4}\sum_{r,t_1\in\tilde{\mathcal{A}}_{k,1}}\lambda_M(r)^2\\ &\quad M^{(y)}_{j_2}(k+j_1+r-j_2)M^{(x)}_{j_3}(t_2-j_3)M^{(y)}_{j_4}(t_2+r-j_4)dF_0(x)dF_0(y)\end{aligned}$$

and $\tilde{\mathcal{A}}_{k,1} = \{(r,t_2) \text{ such that } (k > k+j_1+r-j_2 > t_2-j_3 > t_2+r-j_4)\}$. Using a similar argument we can decompose $\mathcal{S}_{T,i}$ as $\mathcal{S}_{T,i} = \frac{1}{T^2}\sum_{k=1}^{T} U_{k,i}$ (and $U_{k,i}$ is defined similar to above). Therefore, altogether $\mathcal{S}_T$ is the sum of martingale differences, where $\mathcal{S}_T = \frac{1}{T^2}\sum_{k=1}^{T}\sum_{i=1}^{24} U_{k,i}$, and $\sum_{i=1}^{24} U_{k,i} \in \sigma(X_k, X_{k-1}, \ldots)$ are the martingale differences. Therefore under the conditions in Theorem II.4 we have

$$\mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T) = \mathcal{S}_T + O_p(\frac{1}{T} + \frac{M^{1/2}}{T^2}) = \frac{1}{T^2}\sum_{k=1}^{T}\sum_{i=1}^{24} U_{k,i} + O_p(\frac{1}{T} + \frac{M^{1/2}}{T^2}).$$

These approximations will allow us to use the martingale central limit theorem to prove asymptotic normality, which requires the following lemma.

Lemma II.4 *Suppose that Assumption IV.1 holds. Then for all* $1 \leq i \leq 24$ *and* $1 \leq k \leq T$ *we have*

$$\|\sum_{i=1}^{24} U_{k,i}\|_q = O(T^{1/2}M^{1/2}), \tag{2.17}$$

$$\begin{aligned} &\frac{1}{T^2 M}\sum_{k=1}^{T}\mathbb{E}(\sum_{i=1}^{24} U_{k,i}^2) \\ &\to \frac{4}{M}\int\int \Delta_M(\theta_1-\theta_2)^2\prod_{i=1}^{2} G(x_1,x_2;\theta_i)G(y_1,y_2;\theta_i)d\theta_i\prod_{j=1}^{2} dF_0(x_j)dF_0(y_j) \end{aligned} \tag{2.18}$$

and

$$\frac{1}{T^2 M}\sum_{k=1}^{T}\left[\mathbb{E}\left((\sum_{i=1}^{24} U_{k,i})^2|\mathcal{F}_{k-1}\right) - \mathbb{E}(\sum_{i=1}^{24} U_{k,i})^2\right] \xrightarrow{\mathcal{P}} 0. \tag{2.19}$$

PROOF. To prove the result we concentrate on $U_{k,1}$, a similar proof applies to the other terms. By using the Hölder inequality, for any $q \geq 2$, we obtain

$$\begin{aligned} \|U_{k,1}\|_q \quad \leq \quad & \int\sum_{j_1=0}^{T-k}\|M_{j_1}(k)\|_{4q}\|\sum_{j_2,j_3,j_4}\sum_{r,t_1\in\tilde{\mathcal{A}}_{k,1}}\lambda_M(r)^2 M_{j_2}^{(y)}(k+j_1+r-j_2)\times \\ & M_{j_3}^{(x)}(t_2-j_3)M_{j_4}^{(y)}(t_2+r-j_4)\|_{4q/3}dF_0(x)dF_0(y). \end{aligned}$$

Now by repeated use of Burkhölder's inequality we have $\|U_{k,1}\|_q = O(M^{1/2}T^{1/2})$, using a similar method we obtain a similar bound for $\|U_{k,i}\|_q$, this gives (2.17). The proof of (2.18) follows from the proof of Theorem II.4 (noting that the asymptotic variance of $\mathcal{Q}_{\mathcal{T}}$ is determined by the variance of $\mathcal{S}_T$).

To prove (2.19), we consider only the $U_{k,1}$ (the proof involving the other terms in similar). For brevity we write $U_{k,1}$ as

$$U_{k,1} = \int \sum_{j_1=0}^{T-k} M_{j_1}^{(x)}(k) N_{j_1,k-1,1}^{(x,y)} dF_0(x) dF_0(y),$$

where

$$N_{j_1,k-1,1}^{(x,y)} \quad = \quad \sum_{j_2,j_3,j_4} \sum_{r,t_1 \in \tilde{\mathcal{A}}_{k,1}} \lambda_M(r)^2 M_{j_2}^{(y)}(k+j_1+r-j_2) M_{j_3}^{(x)}(t_2-j_3) M_{j_4}^{(y)}(t_2+r-j_4).$$

Noting that $N_{j_1,k-1,1}^{(x,y)} \in \mathcal{F}_{k-1}$ we have

$$\begin{aligned}
& \frac{1}{T^2 M} \sum_{k=1}^{T} \big(\mathbb{E}(U_{k,1}^2 | \mathcal{F}_{k-1}) - \mathbb{E}(U_{k,1})^2\big) = \\
& \frac{1}{T^2 M} \sum_{k=1}^{T} \int \sum_{j_1,j_2=0}^{T-k} \big(\mathbb{E}(M_{j_1}^{(x_1)}(k) M_{j_2}^{(x_2)}(k) | \mathcal{F}_{k-1}) - \mathbb{E}(M_{j_1}^{(x_1)}(k) M_{j_2}^{(x_2)}(k))\big) \\
& N_{j_1,k-1,1}^{(x_1,y_1)} N_{j_2,k-1,1}^{(x_2,y_2)} \prod_{i=1}^{2} dF_0(x_i) dF_0(y_i) \\
+ \quad & \frac{1}{T^2 M} \sum_{k=1}^{T} \int \sum_{j_1,j_2=0}^{T-k} \mathbb{E}(M_{j_1}^{(x_1)}(k) M_{j_2}^{(x_2)}(k))\big) \\
& \big(N_{j_1,k-1,1}^{(x_1,y_1)} N_{j_2,k-1,1}^{(x_2,y_2)} - \mathbb{E}(N_{j_1,k-1,1}^{(x_1,y_1)} N_{j_2,k-1,1}^{(x_2,y_2)})\big) \prod_{i=1}^{2} dF_0(x_i) dF_0(y_i).
\end{aligned}$$

Now by using similar methods to the iterative martingale methods detailed in the proof of Theorem II.4, we can show that the $\|\cdot\|_q$-norm ($q \geq 2$) of the above converges to zero, thus we have (2.19). $\square$

PROOF of Theorem II.2 Using the above we have

$$\mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T) = \frac{1}{T^2} \sum_{k=1}^{T} \sum_{i=1}^{24} U_{k,i} + O_p(\frac{1}{T} + \frac{M^{1/2}}{T^2}),$$

thus $\mathcal{Q}_T - \mathbb{E}(\mathcal{Q}_T)$ can be written as the sum of martingales plus a smaller order term. Therefore to prove asymptotic normality of $\mathcal{Q}_T$ we can use the martingale central limit, for this we need to verify (a) the conditional Lindeberg condition $\frac{1}{T^2M}\sum_{k=1}^{T}\mathbb{E}(|\sum_{i=1}^{24}U_{k,i}|^2 I(\frac{1}{TM^{1/2}}|\sum_{i=1}^{24}U_{k,i}| > \varepsilon)|\mathcal{F}_{k-1}) \xrightarrow{\mathcal{P}} 0$ for all $\varepsilon > 0$, (b) that $\frac{1}{T^2M}\sum_{k=1}^{T}\mathbb{E}(|\sum_{i=1}^{24}U_{k,i}|^2|\mathcal{F}_{k-1}) - \frac{T^2}{M}\mathrm{var}(Q_T) \xrightarrow{\mathcal{P}} 0$.

To verify the conditional Lindeberg condition, we observe that the Cauchy-Schwartz and Markov's inequalities give

$$\begin{aligned} & \frac{1}{T^2M}\sum_{k=1}^{T}\mathbb{E}\big(|\sum_{i=1}^{24}U_{k,i}|^2 I(\frac{1}{TM^{1/2}}|\sum_{i=1}^{24}U_{k,i}| > \varepsilon)|\mathcal{F}_{k-1}\big) \\ \leq\ & \frac{1}{\varepsilon T^4M^2}\sum_{k=1}^{T}\mathbb{E}(|\sum_{i=1}^{24}U_{k,i}|^4|\mathcal{F}_{k-1}) := B_T. \end{aligned}$$

By using (2.17), the expectation of the above is $\mathbb{E}(B_T) = O(\frac{1}{T})$. As B_T is a non-negative random variable, this implies $B \xrightarrow{\mathcal{P}} 0$ as $T \to \infty$. Thus we have shown that the Lindeberg condition is satisfied. To prove (b) we note that

$$\begin{aligned} & \frac{1}{T^2M}\sum_{k=1}^{T}\mathbb{E}\Big(|\sum_{i=1}^{24}U_{k,i}|^2|\mathcal{F}_{k-1}\Big) - \frac{T^2}{M}\mathrm{var}(Q_T) \\ =\ & \frac{1}{T^2M}\sum_{k=1}^{T}\Big[\mathbb{E}\big(|\sum_{i=1}^{24}U_{k,i}|^2|\mathcal{F}_{k-1}\big) - \mathbb{E}\big(|\sum_{i=1}^{24}U_{k,i}|^2\big)\Big] \\ +\ & \frac{1}{T^2M}\sum_{k=1}^{T}\mathbb{E}(|\sum_{i=1}^{24}U_{k,i}|^2) - \frac{T^2}{M}\mathrm{var}(Q_T). \end{aligned}$$

By using (2.18) and (2.19) the above converges to zero in probability. Thus we have verified the conditions of the martingale central limit theorem and we have the desired result. $\square$

7.3. *Proof of Theorem II.3*

Since the limiting distribution of $\mathcal{Q}_T$ is determined by $\mathcal{Q}_{T,2}$, we rewrite $\mathcal{Q}_{T,2}$ as in the proof of Theorem IV.2. We observe that

$$
\begin{aligned}
&\mathcal{Q}_{T,2}\\
&= \frac{2}{T}\Re\int\sum_k \Lambda_T(x,y;\omega_k)\big\{J_T(x;\omega_k)\overline{J}_T(y;\omega_k) - \mathbb{E}(J_T(x;\omega_k)\overline{J}_T(y;\omega_k))\big\}dF_0(x)dF_0(y)\\
&= \int\frac{2}{T}\sum_{t,\tau}\lambda_M(t-\tau)^2 D_{t-\tau,T}(x,y)(Z_t(x)Z_\tau(y) - \mathbb{E}(Z_t(x)Z_\tau(y)))dF_0(x)dF_0(y)\\
&= \int\frac{2}{T}\sum_{t,\tau}\lambda_M(t-\tau)^2 D_{t-\tau,T}(x,y)(\tilde{Z}_T(x)\tilde{Z}_\tau(y) - C_r(x,y))dF_0(x)dF_0(y) + O_p(\frac{1}{T}),
\end{aligned}
$$

where $\Lambda_T(x,y;\omega_s) = \sum_r \lambda_M(r)^2(\frac{T-|r|}{T})\big(C_{r,1}(x,y) - C_{r,0}(x,y)\big)\exp(ir\omega_k)$, $D_{r,T}(x,y) = (\frac{T-|r|}{T})\big(C_{r,1}(x,y) - C_{r,0}(x,y)\big)$ and $\tilde{Z}_t(x) = I(X_t \le x) - F_1(x)$.

PROOF of Theorem II.3 Now we observe that under the stated assumptions of the theorem we have that the quantile covariances under the null decay at the rate $\sup_{x,y}|C_{r,0}(x,y)| \le K|r|^{-(2+\delta)}$ (for some $\delta > 0$) and $\sup_{x,y}|C_{r,1}(x,y)| \le K|r|^{-s}$ (for some $s > 2$). Thus by definition of $D_{r,T}(\cdot)$, we have $\sup_{x,y}|\lambda_M(r)D_{r,T}(x,y)| \le K|r|^{-\min(2+\delta,s)}$. Thus we can write $\mathcal{Q}_{T,2}$ as

$$
\begin{aligned}
&\mathcal{Q}_{T,2}\\
= &\int\frac{2}{T}\sum_{t,\tau}\lambda_M(t-\tau)^2 D_{t-\tau,T}(x,y)(\tilde{Z}_T(x)\tilde{Z}_\tau(y) - \mathbb{E}(\tilde{Z}_T(x)\tilde{Z}_\tau(y)))dF_0(x)dF_0(y)\\
+ &O_p(\frac{1}{T}),
\end{aligned}
$$

where we observe that terms where $|t-\tau| > 2M$, are zero. Thus with the Bernstein blocking arguments for quadratic forms used to prove Theorem IV.2, we can show asymptotic normality of the above. This proves (2.5). Finally to prove (2.6), we note that $\mathcal{Q}_T = \mathcal{Q}_{T,2} + E_{T,2} + O_p(\frac{M^{1/2}}{T} + \frac{M}{T} + \frac{1}{M^{s-1}})$, by using (2.5), this immediately leads to (2.6). $\square$

CHAPTER III

PERIODICITIES AND OTHER FEATURES ON THE DOMAIN OF A TIME SERIES

1. Introduction

Often spectral methods are used to analysis a (stationary) time series, because it may exhibit periodicities or patterns which can easily be interpreted using the Fourier transform of the autocovariances. Brillinger (1981) and Shumway and Stoffer (2006), Chapter 4, eloquently describe spectral analysis of a time series and its applications, for example it can be used to identify the dominant frequencies in a system and identify the linear time series model. However, despite its advantage, there are disadvantages in using the autocovariance function as the basis in spectral analysis. The autocovariance function only measures the average interaction between elements of a time series, but cannot identify differences which may lie on the domain of the time series. Consider the following toy example, suppose $\{X_t\}$ is a stationary time series where

$$X_t = \varepsilon_t + I(r_1 \leq \varepsilon_t \leq r_2)Z_t + I(-r_2 \leq \varepsilon_t \leq -r_1)Y_t, \tag{3.1}$$

where $\{\varepsilon_t\}$ are independent, identically distributed (iid) random variables with variance σ^2 and $\{Z_t\}$ and $\{Y_t\}$ are two independent stationary, linear time series, with spectral densities $f_Z(\omega)$ and $f_Y(\omega)$ respectively. Both are independent of $\{\varepsilon_t\}$. The spectral density of $\{X_t\}$ is $f_X(\omega) = \sigma^2 + P(r_1 \leq \varepsilon_0 \leq r_2)^2 f_Z(\omega) + P(-r_2 \leq \varepsilon_0 \leq -r_1)^2 f_Y(\omega)$, which is not particularly informative about the underlying model. Furthermore, it is not clear whether the dominant frequencies in f_Y and f_Z arise in f_X.

This is an example of a time series where it may be more valuable to understand the interactions between different regions of the domain of the time series that the autocovariance.

Our objective is to define a spectral density which measures associations between different parts of the domain of the time series. More precisely, we assume that the time series is stationary and define the association spectral density as

$$g_{\mathcal{S}}(x,y;\omega) = \frac{1}{2\pi}\Big(\sum_{r\neq 0}\big(f_r(x,y) - f(x)f(y)\big)\exp(ir\omega) - f(x)f(y)\Big),$$

$f_r(\cdot)$ and $f(\cdot)$ is the marginal and joint density of X_0 and (X_0, X_r) respectively. To understand how $g_{\mathcal{S}}$ may help, we return to example (3.1). It can be shown that the association spectral density in this example is

$$\begin{aligned} g_{\mathcal{S},X}(x,y;\omega) &= \int_{-r_2}^{-r_1} g_{\mathcal{S},Y}(x-e_1, y-e_2) f_\varepsilon(e_1) f_\varepsilon(e_2) de_1 de_2 \\ &+ \int_{r_1}^{r_2} g_{\mathcal{S},Z}(x-e_1, y-e_2) f_\varepsilon(e_1) f_\varepsilon(e_2) de_1 de_2 + k(x,y) \end{aligned}$$

where $g_{\mathcal{S},Y}$ and $g_{\mathcal{S},Z}$ are the association spectral densities of $\{Y_t\}$ and $\{Z_t\}$ respectively and f_ε the marginal density of $\{\varepsilon_t\}$ and $k(x,y) = \frac{1}{2\pi}\Big((\int_{r_1}^{r_2} f_Z(x-e) f_\varepsilon(e)de)(\int_{-r_2}^{-r_1} f_Y(y-\varepsilon) f_\varepsilon(e)de) + (\int_{r_1}^{r_2} f_Z(y-\varepsilon) f_\varepsilon(e)de)(\int_{-r_2}^{-r_1} f_Y(x-\varepsilon) f_\varepsilon(e)de)\Big)$. To understand how $g_{\mathcal{S}}$ behaves for different x and y, we consider the association spectral density of a stationary Gaussian time series $X_t = \sum_{j\geq 0} a_j \varepsilon_{t-j}$, which has an explicit form such as

$$\begin{aligned} & g_{\mathcal{S},Gaussian} \\ = \; & \frac{1}{(2\pi\sigma_0)^2}\exp\big(-\frac{x^2+y^2}{2\sigma_0^2}\big) \times \\ & \Big(\sum_{r\neq 0}\Big\{\frac{1}{\sqrt{1-\rho_r^2}}\exp\big(-\frac{\rho_r^2(x^2+y^2) - 2\rho_r xy}{2\sigma_0^2(1-\rho_r^2)}\big) - 1\Big\}\exp(ir\omega) - 1\Big) \end{aligned}$$

where $\sigma_0^2 = \sigma^2 \sum_{j=0}^{\infty} a_j^2$ and $\rho_r = \sum_{r=0}^{\infty} a_j a_{j+|r|}$.

We observe that it tends to be largest when x and y are close to zero (since X_t has the largest of lying there). In general this behavior is true for non-Gaussian linear time series, that is the association spectral density of a linear time series tends to be largest about zero. This has an interesting repercussion on $g_{\mathcal{S},X}$. $g_{\mathcal{S},X}$ will have different behaviors depending on the x and y. If x and y lie close to the interval $[-r_2, -r_1]$, $g_{\mathcal{S},X}(\cdot)$ is dominated by $g_{\mathcal{S},Y}$, on the other hand for x and y close to the interval $[r_1, r_2]$, $g_{\mathcal{S},X}$ will be dominated by $g_{\mathcal{S},Z}$. Thus the association spectral density suggests that X_t is a mixture of two time series. This example motivates the use of the association spectral density as a tool in explanatory data analysis.

In general, the association spectral density can be considered as a means of depicting the general dependence between pairs of random variables in a time series, which is usually called serial dependence. Several methods have been proposed to generalize the spectral density to serial dependence. For example, Hong (1998) defines generalized spectral density, which is the Fourier transform of the characteristic function of pair-wise dependent data, which he uses to test for goodness of fit. More recently, Li (2008), Hagemann (2011) and Dette et al. (2011) define a generalize spectral density based on the cumulative distribution functions, which does not easily represent the associations and periodicities between different parts of the domain of the time series. In this paper, we address this issue. The purpose of this paper is to introduce the association spectral density as a means of studying the pairwise dependence structure of a time series at different scales and locations.

In Section 2 we consider in detail the properties of the association spectral density and propose a method of estimation. In Section 3 we derive the asymptotic sampling properties of the estimator. Some simulation results can be found in Section 4, and all the proofs are in Section 5.

2. The association covariance and association spectral density

2.1. Motivation

In this section we motivate the association covariance and the corresponding association spectral density. The autocorrelation function gives information about the average interaction between any two random variables in the time series. However, it is too 'global' to give information about how different domains of the time series may interact and influence each other. In order to illicit this type of information we can transform the time series $\{X_t\}$ by windowing it around regions of the domain that are of interest. Suppose that we are interested in the interaction of the time series around x and y, then by transforming X_t as $\{(\frac{1}{b}W(\frac{x-X_t}{b}), \frac{1}{b}W(\frac{y-X_t}{b})\}_t$, where $W : [-1, 1] \to \mathbb{R}$ is a positive symmetric kernel and b a window length, and can consider the cross correlation of this transformed time series. Inspecting the covariance, we observe that depending on how close the kernel $W(\cdot)$ is to the rectangular kernel we have

$$\begin{aligned}
& \mathrm{cov}\big(W_b(x - X_t), W_b(y - X_\tau)\big) \\
= \;& \mathbb{E}\big(W_b(x - X_t)W_b(y - X_\tau)\big) - \mathbb{E}\big(W_b(x - X_t))\mathbb{E}(W_b(y - X_\tau))\big) \\
\approx \;& P(x - b < X_t < x + b, y - b < X_\tau < y + b) \\
- \;& P(x - b < X_t < x + b)P(y - b < X_\tau < y + b).
\end{aligned}$$

Therefore, we observe that a positive $\mathrm{cov}(W_b(x - X_t), W_b(y - X_\tau))$ implies the probability that X_t and X_τ lie in the neighborhood of x and y respectively, is greater than independent events (thus positive association), whereas a negative covariance suggests the opposite. However, the magnitude of association depends on various factors that we now explore. Under the condition that $\{X_t\}$ is ψ-mixing (see Lemma

III.2, below) it can be shown that

$$|\text{cov}(W_b(x - X_t), W_b(y - X_\tau))| \leq C|t - \tau|^{-s}\mathbb{E}(W_b(x - X_0))\mathbb{E}(W_b(y - X_0)),$$

where $W_b(x) = \frac{1}{b}W(\frac{x}{b})$, $\mathbb{E}(W_b(\frac{x-X_0}{b})) = \int \frac{1}{b}W(\frac{x-y}{b})f(y)dy$, f is the marginal density of X_t and s the ψ-mixing size. We can see from the above that the covariance is determined by two factors (a) the region that is windowed and (b) the ψ-mixing size s. For example, the covariances $\text{cov}(W_b(x_1 - X_0), W_b(y - X_r))$ and $\text{cov}(W_b(x_2 - X_0), W_b(y - X_r))$ both decay at the same rate ($O(|r|^{-s})$), however if x_1 lies in the tails and x_2 lies close to the mode, then $\mathbb{E}(W_b(x_1 - X_t))$ will be smaller than $\mathbb{E}(W_b(x_2 - X_t))$, subsequently it is likely that $\text{cov}(W_b(x_1 - X_0), W_b(y - X_r))$ will be smaller than $\text{cov}(W_b(x_2 - X_0), W_b(y - X_r))$. Therefore, to make a fairer comparison between these two covariances, we can standardize, by defining the following pseudo association correlation

$$\text{cor}_p(W_b(x - X_0), W_b(y - X_r) = \frac{\text{cov}(W_b(x - X_0), W_b(y - X_r))}{\mathbb{E}(W_b(x - X_0))\mathbb{E}(W_b(y - X_0)}. \tag{3.2}$$

Comparing (3.2) with the usual correlation $\text{cor}(W_b(x - X_0), W_b(y - X_r))$, we observe that since $\text{var}(W_b(x - X_t)) = O(b^{-1})$, then $\text{cor}(W_b(x - X_0), W_b(y - X_r)) = O(b|r|^{-s})$, whereas $\mathbb{E}(W_b(x - X_0)) = O(1)$, therefore cor_p, is in some sense invariant to the scale b.

Having defined the association covariance and pseudo association correlation, we now consider the corresponding spectral density. We recall that the for autoregressive processes, the spectral density is often used to check for the order of the process (Parzen (1974)) and to look for periodicities and patterns in the autocovariance structure. However, it could well be that periodicities that may arise in certain regions of the domain of the random variables are averaged out in the regular spectral density. This suggests that we should search for patterns that may arise in the cross covariance of $\{W_b(x - X_t), W_b(y - X_t)\}_t$, thus we define its Fourier transform

$$g_b(x, y; \omega) = \frac{1}{2\pi}\sum_r \text{cov}\big(W_b(x - X_0), W_b(y - X_r)\big)\exp(ir\omega).$$

However, in the following section we show that the limit of $\text{cov}\big(W_b(x - X_0), W_b(y - X_0)\big)$ as $b \to 0$, is not well defined, this means that the limit of $g_b(x, y; \omega)$ as a function, won't be well defined. Therefore we a define a shifted version of g_b

$$\begin{aligned} g_{\mathcal{S},b}(x, y; \omega) &= g_b(x, y; \omega) - \frac{1}{2\pi}\mathbb{E}\big(W_b(x - X_0)W_b(y - X_0)\big) \\ &= \frac{1}{2\pi}\sum_{r\neq 0} \text{cov}\big(W_b(x - X_0), W_b(y - X_r)\big)\exp(ir\omega) \\ &\quad -\frac{1}{2\pi}\mathbb{E}\big(W_b(x - X_0)\big)\mathbb{E}\big(W_b(y - X_0)\big), \end{aligned}$$

whose limit is well defined. Since g_b and $g_{\mathcal{S},b}$ are simply shifts of each other, their shapes are same. Moreover, when the intervals $[x - b, x + b]$ and $[y - b, y + b]$ do not intersect, then $g_b(x, y; \omega) = g_{\mathcal{S},b}(x, y; \omega)$. The standardized association spectral density is defined similarly

$$\begin{aligned} h_{\mathcal{S},b}(x, y; \omega) &= \frac{1}{2\pi}\left(\sum_{r\neq 0} \text{cor}_p(W_b(x - X_0), W_b(y - X_r))\exp(ir\omega) - 1\right) \\ &= \frac{g_{\mathcal{S},b}(x, y; \omega)}{\mathbb{E}\big(W_b(x - X_0)\big)\mathbb{E}\big(W_b(y - X_0)\big)}. \end{aligned}$$

Using the above argument $h_{\mathcal{S},b}$ may be a useful tool for comparing association spectral for different values of x, y and b.

Remark III.1 *The correlation association spectral density is defined as*

$$h_{\mathcal{S}}(x, y; \omega) = \frac{1}{2\pi}\left(\sum_{r\neq 0} \big(\frac{f_r(x, y)}{f(x)f(y)} - 1\big)\exp(ir\omega) - 1\right).$$

However it is possible to generalise the definition. For example, one can also consider *monotonic tranforms of the psuedo-correlation such as the log transform*

$$\ell(x,y;\omega) = \frac{1}{2\pi}\sum_{r}\log\big(\frac{f_r(x,y)}{f(x)f(y)}\big)\exp(ir\omega).$$

For the Gaussian time series, the above has the form

$$\ell(x,y;\omega) = -\frac{1}{2\pi}\sum_{r}\left(\frac{1}{2}\log(1-\rho_r^2) + \frac{1}{2\sigma_0^2(1-\rho_r^2)}\big(\rho_r^2(x^2+y^2) - 2\rho_r xy\big)\right)\exp(ir\omega),$$

thus $\ell(x,y;\omega)$ is finite for any short-memory Gaussian time series. In general, using Assumption III.1 (ii) we have that $|\log(\frac{f_r(x,y)}{f(x)f(y)})| \leq \log(C|r|^{-s}+1)$, thus if $s > 1$ then $\ell(x,y;\omega)$ is well defined. However, in general it is difficult to estimate $\ell(\cdot)$, therefore in the following sections we will focus on $g_{\mathcal{S}}$ and $h_{\mathcal{S}}$.

In Figures 9 and 10 we plot the association spectral densities for two AR(1) ($X_t = 0.6X_{t-1} + Z_t$, $X_t = 0.9X_{t-1} + Z_t$) with independent, identically distributed (iid) Gaussian innovations Z_t. The diagonals are of $g_{\mathcal{S}}(x,x;\omega)$, the lower triangle contains the real part of $g_{\mathcal{S}}(x,y;\omega)$ and the upper triangle the imaginary part of $g_{\mathcal{S}}(x,y;\omega)$. For better understanding the behaviors of $g_{\mathcal{S}}(\cdot)$ and $h_{\mathcal{S}}(\cdot)$, we plot their inverse Fourier transforms $\{f_r(x,y)1_{r\neq 0} - f(x)f(y)\}_r$ and $\{\frac{f_r(x,y)}{f(x)f(y)}1_{r\neq 0} - 1\}_r$ in Figures 11 and 12. The shapes and magnitudes of $g_{\mathcal{S}}(x,x;\omega)$ are similar for all x, but the magnitude of $h_{\mathcal{S}}(x,x;\omega)$ for $x = 50\%$ percentile is much smaller than the others. This observation hints us that the magnitude of $g_{\mathcal{S}}(x,x;\omega)$ for $x = 50\%$ percentile mainly comes from the large value of $f(x)$ and it can be confirmed in Figure 11 and 12. Except $r = 0$, we can interpret $\{\frac{f_r(x,y)}{f(x)f(y)}1_{r\neq 0} - 1\}_r$ as if it is the limit pseudo association correlation defined in (3.2) as b goes to 0. Similar to the usual auto-correlation function of AR(1) process, its rate of decay is determined by its AR coefficient and this rate of decay specifies the steepnesses of the curves in $h_{\mathcal{S}}(x,x;\omega)$ for $x = 10, 90\%$ percentile. For the AR(1) model with $\phi > 0$, when X_t lies in the either direction of the tail area, X_{t+1} will likely fall onto the tail area preserving the sign and there's only small probability that X_{t+1} is in the tail area of the opposite direction. These behaviors could be captured in positive values of $\{\frac{f_r(x,x)}{f(x)f(x)}1_{r\neq 0} - 1\}_r$ and negative values of $\{\frac{f_r(x,y)}{f(x)f(y)}1_{r\neq 0} - 1\}_r$ especially between 10% and 90% percentile, and we can see that the standardized association spectral densities have distinct looks reflecting these features fairly well.

Figures 13 and 14 illustrate the association spectral densities for ARCH(1) ($X_t = \sigma_t Z_t$ with $\sigma_t^2 = 1/1.9 + 0.9X_{t-1}^2$) and the squared ARCH, with independent, identically distributed (iid) Gaussian innovations Z_t. We first observe that the AR and ARCH association spectral densities are very different and the association spectral densities for ARCH are not flat suggesting that though ARCH process is uncorrelated overall, there definitely exist the correlations between certain regions of the time domains. Furthermore, recalling that the AR and ARCH squared have the same spectral density (if the moments of the ARCH squared exist), there is a large difference between the association spectral density of the AR and the ARCH squared in Figure 10 and 14.

2.2. ***Properties of*** $g_b(x, y; \omega)$

We now consider how g_b behaves for different b, focusing on the case that $b \to 0$. First we define two close related quantities

$$
\begin{aligned}
g(x, y; \omega) &:= \frac{1}{2\pi}\left(\sum_{r\neq 0}\big(f_r(x,y) - f(x)f(y)\big)\exp(ir\omega) + \big(f(x)\delta_x(y) - f(x)f(y)\big)\right) \\
g_{\mathcal{S}}(x, y; \omega) &= \frac{1}{2\pi}\left(\sum_{r\neq 0}\big(f_r(x,y) - f(x)f(y)\big)\exp(ir\omega) - f(x)f(y)\right),
\end{aligned}
$$

where $f(\cdot)$ and f_r denotes the densities of $\{X_t\}$ and (X_0, X_r), respectively and $\delta_x(y)$ the Dirac delta function. Observe that for $x \neq y$, $g_{\mathcal{S}}(x,y;\omega) = g(x,y;\omega)$. We call $g_{\mathcal{S}}(x,y;\omega)$ the association spectral density. Since the definition of g involves a Dirac

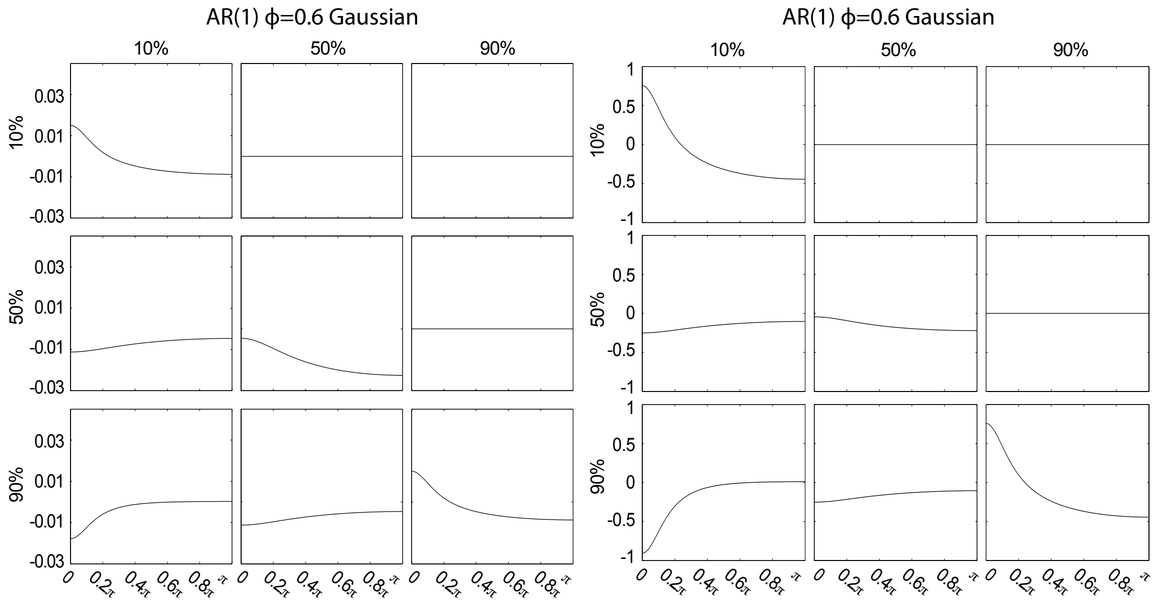

Fig. 9. The association spectral density of $X_t = 0.6X_{t-1} + Z_t$; (Left) Nonstandardized, (Right) Standardized

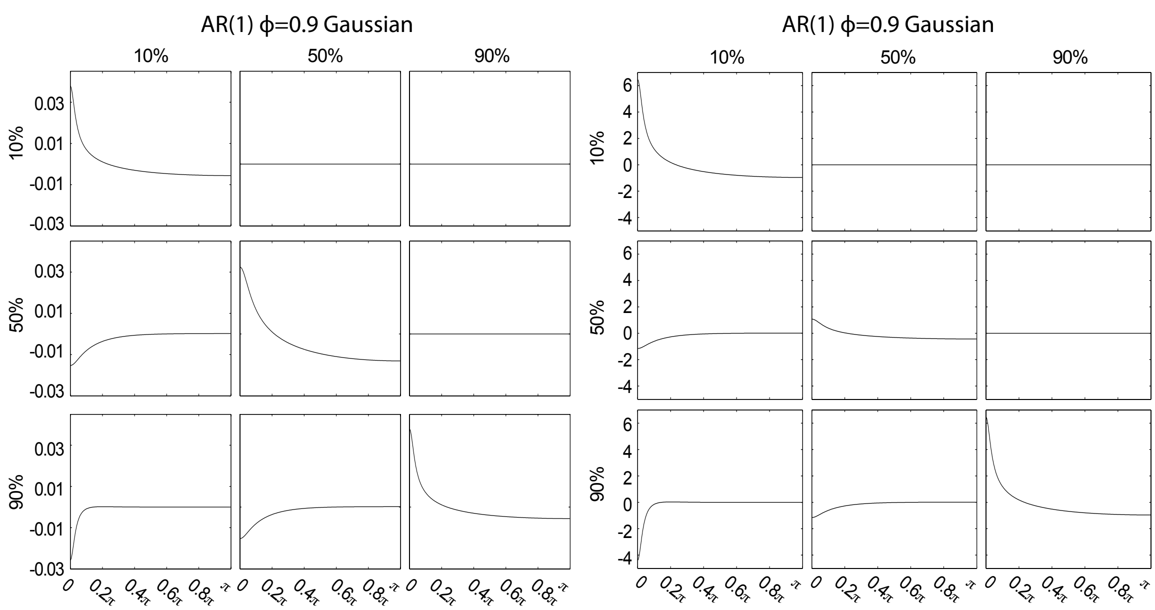

Fig. 10. The association spectral density of $X_t = 0.9X_{t-1} + Z_t$; (Left) Nonstandardized, (Right) Standardized

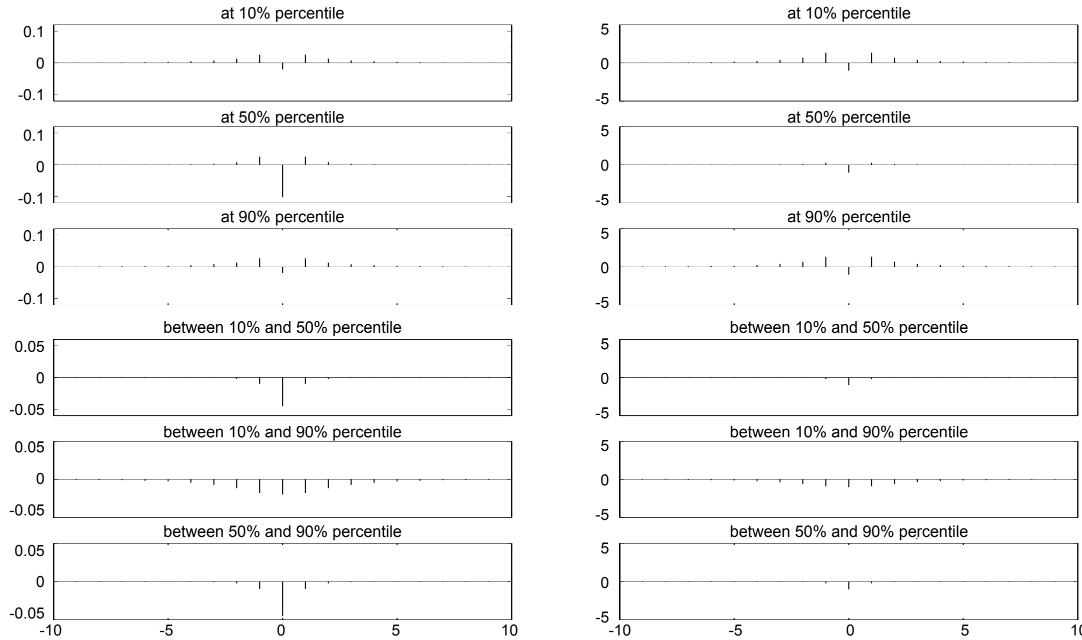

Fig. 11. $X_t = 0.6X_{t-1} + Z_t$; (Left) $\{f_r(x, y)1_{r\neq 0} - f(x)f(y)\}_r$, (Right) $\{\frac{f_r(x,y)}{f(x)f(y)}1_{r\neq 0} - 1\}_r$

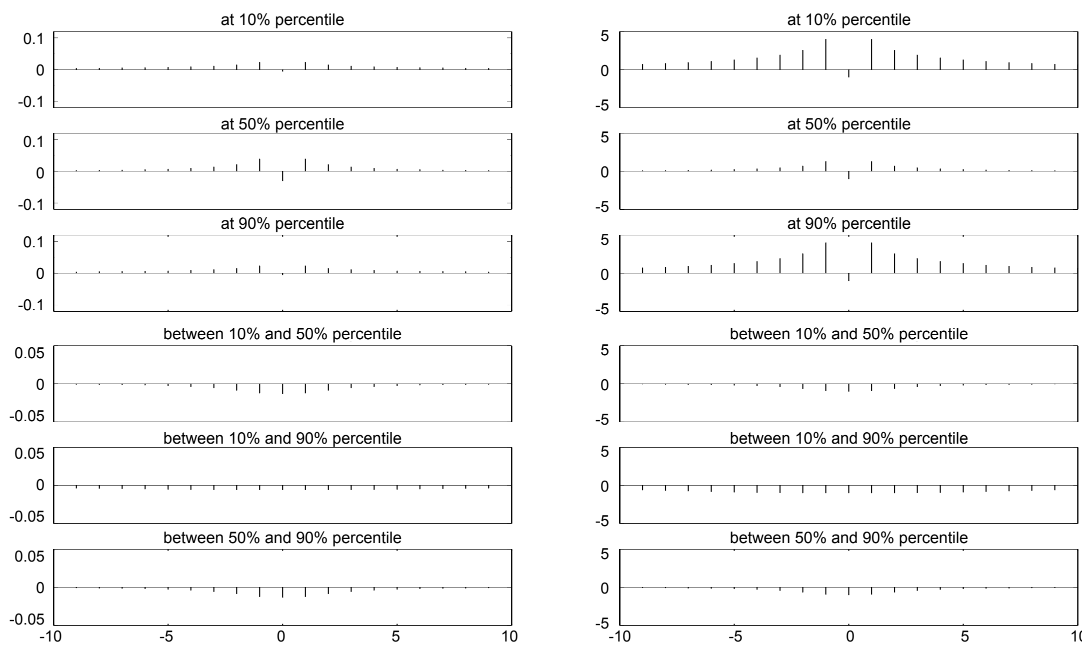

Fig. 12. $X_t = 0.9X_{t-1} + Z_t$; (Left) $\{f_r(x, y)1_{r\neq 0} - f(x)f(y)\}_r$, (Right) $\{\frac{f_r(x,y)}{f(x)f(y)}1_{r\neq 0} - 1\}_r$

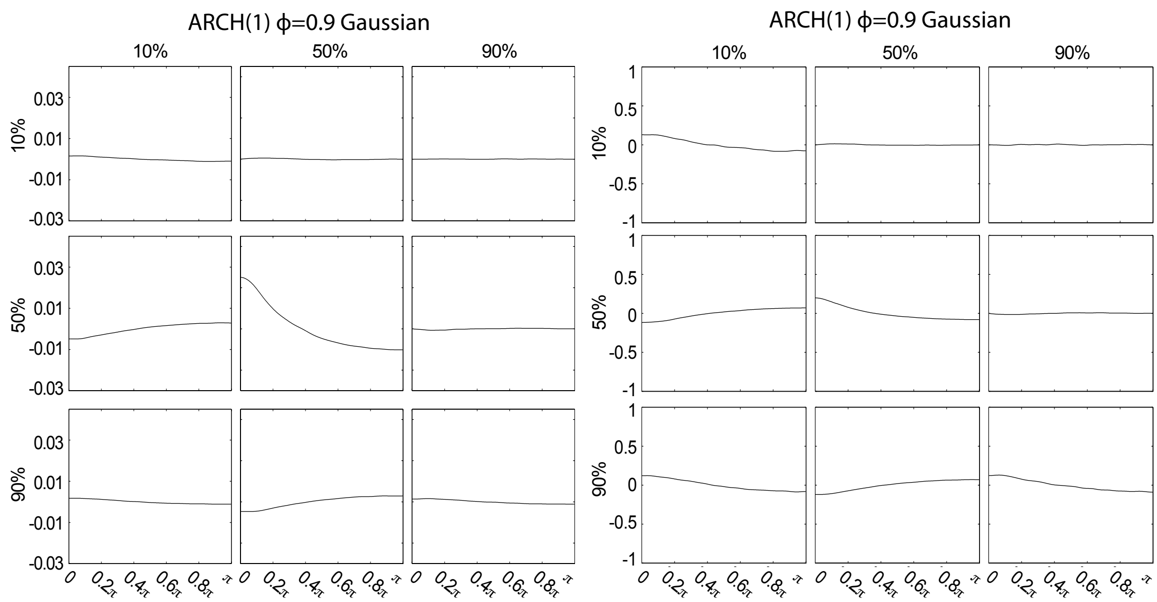

Fig. 13. The association spectral density of $X_t = \sigma_t Z_t$, where $\sigma_t^2 = 1/1.9 + 0.9X_{t-1}^2$; (Left) Nonstandardized, (Right) Standardized

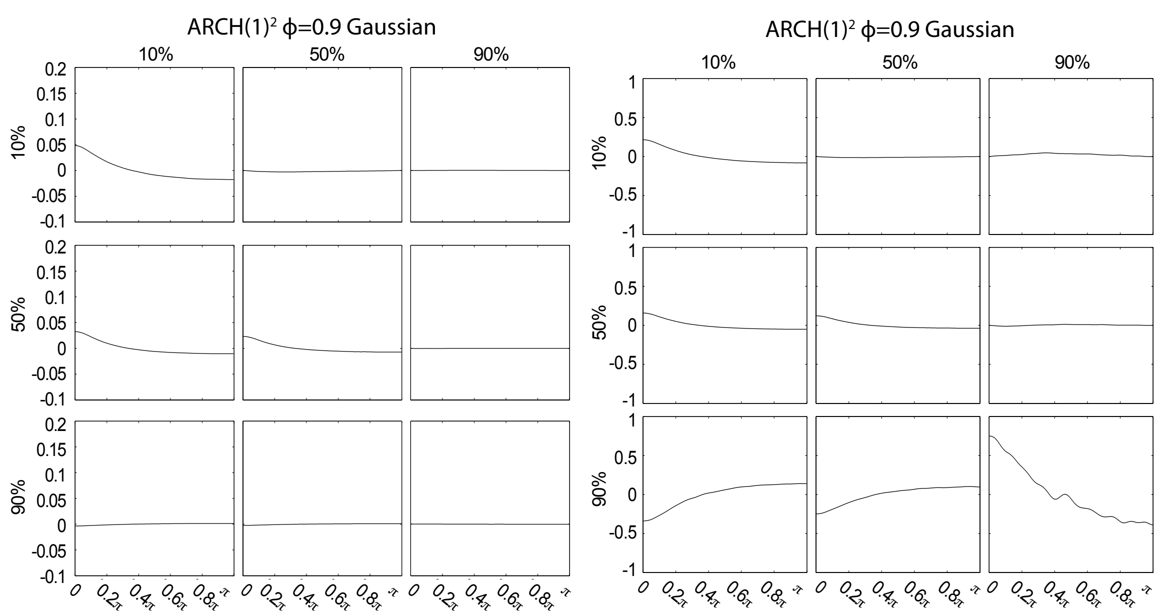

Fig. 14. The association spectral density of $X_t^2 = \sigma_t^2 Z_t^2$, where $\sigma_t^2 = 1/1.9 + 0.9X_{t-1}^2$; (Left) Nonstandardized, (Right) Standardized

delta function, it is not well defined, whereas we show in Lemma III.2 that $g_{\mathcal{S}}$ is finite for all x and y. However, g is a generalized function and is closely related to the usual spectral density function. In particular, for any transformation of the time series $\{h_1(X_t), h_2(X_t)\}_t$, its cross spectral density can be written as

$$f_{h_1h_2}(\omega) = \int\int h_1(x)h_2(y)g(x,y;\omega)dxdy.$$

Returning to the spectral density of $\{W_b(x-X_t), W_b(y-X_t)\}_t$, $g_b(x,y;\omega)$, we see that by using the above argument we have

$$g_b(x,y;\omega) = \int\int \frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})g(x,y;\omega)dxdy.$$

If $x \neq y$, then for a small enough b, $g(x,y;\omega)$ is continuous in the neighborhood of $[x-,b,x+b] \times [y-b,y+b]$, therefore by using Bochner's Theorem (see Bochner (1955)) we have the following result.

Lemma III.1 *Suppose Assumption III.1(ii) and (iii) are satisfied. Then for all* $x, y \in \mathbb{R}$

$$\int\int W_b(x-u)W_b(y-v)g_{\mathcal{S}}(u,v;\omega)dudv \to g_{\mathcal{S}}(x,y;\omega)$$

as $b \to 0$.

2.3. Estimation

In order to estimate the association spectral density $g_{\mathcal{S}}$, we define the following discrete Fourier transform (DFT)

$$J_{b,T}(x;\omega) \quad = \quad \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T}\left(\frac{1}{b}W(\frac{x-X_t}{b}) - \bar{W}_b(x)\right)\exp(it\omega),$$

where $\bar{W}_b(x) = \frac{1}{T}\sum_t W_b(x - X_t)$ is the estimate of the marginal density $f(x)$. Motivated by the classical estimator of the cross spectral density we use $\widehat{g}_{b,T}(x, y; \omega)$ as an estimator of $g_b(x, y; \omega)$, where

$$\begin{aligned} \widehat{g}_{b,T}(x, y; \omega_s) &= \frac{1}{2\pi}\sum_{r=-M}^{M} \lambda(\frac{r}{M})\widehat{c}_r(x, y)\exp(ir\omega_s) \\ &= \sum_k K_M(\omega_s - \omega_k)J_{b,T}(x; \omega_k)\overline{J_{b,T}(y; \omega_k)} \end{aligned} \tag{3.3}$$

with $\omega_s = \frac{2\pi s}{T}$,

$$\widehat{c}_r(x, y) = \frac{1}{T}\sum_{t=1}^{T-|r|} \left(\frac{1}{b}W(\frac{x - X_t}{b}) - \bar{W}_b(x)\right)\left(\frac{1}{b}W(\frac{y - X_{t+r}}{b}) - \bar{W}_b(y)\right), \tag{3.4}$$

$\lambda[-1, 1] \to \mathbb{R}$ is the lag window and $K_M(\omega)$ is the corresponding spectral window with

$$K_M(\omega) = \frac{1}{T}\sum_{r=-M}^{M} \lambda(\frac{r}{M})e^{ir\omega}.$$

If we keep b fixed and $\{X_t\}$ is an α-mixing time series, then asymptotic arguments for spectral density estimators (see, for example, Rosenblatt (1984) and Chapter IV), can be used to show $\widehat{g}_{b,T}(x, y; \omega) \xrightarrow{\mathcal{P}} g_b(x, y; \omega)$ (as well as asymptotic normality). However, if we are interested in determining the limit for small windows about x and y, then we first need to establish what we are interested in estimating. More precisely, from Lemma III.1 we know that for $x \neq y$, as $b \to 0$, then g_b goes to g, but for $x = y$ case the limit of $g_b(x, x; \omega)$ is not a function in the strict sense. However, the limit of $g_{b,\mathcal{S}}$, which is the association spectral density $g_{\mathcal{S}}$, is well defined for all x and y. Noting that $g_{\mathcal{S}}$ and g are the same up to a shift, we now consider and estimator of $g_{\mathcal{S}}$. Motivated by the estimation scheme for g_b, we propose $\widehat{g}_{\mathcal{S},T}$ as an estimator of

$g_{\mathcal{S}}$, where

$$\begin{aligned}
\widehat{g}_{\mathcal{S},T}(x,y;\omega_s) &= \frac{1}{2\pi}\sum_{r=-M,r\neq 0}^{M}\lambda(\frac{r}{M})\exp(ir\omega_s)\widehat{c}_r(x,y) - \lambda(0)\frac{1}{2\pi}\overline{W}_b(x)\overline{W}_b(y) \\
&= \sum_k K_M(\omega_s-\omega_k)J_{b,T}(x;\omega_k)\overline{J_{b,T}(y;\omega_k)} \\
&- \lambda(0)\frac{1}{2\pi T}\sum_{t=1}^{T}\frac{1}{b^2}W(\frac{x-X_t}{b})W(\frac{y-X_t}{b}) \qquad (3.5)
\end{aligned}$$

In order to make comparisons for different values of x and y that takes into account that the marginal distributions of x and y we can estimate standardised association spectral density, $h_{\mathcal{S}}$, with

$$\widehat{h}_{\mathcal{S},T}(x,y;\omega_s) = \frac{\widehat{g}_{\mathcal{S},T}(x,y;\omega_s)}{\bar{W}_b(x)\bar{W}_b(y)}.$$

In order to construct confidence intervals for $g_{\mathcal{S}}$ and $h_{\mathcal{S}}$ we derive the asymptotic properties of their estimators, noting that unlike the usual spectral density we need to consider the limit of $\widehat{g}_{\mathcal{S},T}(x,y;\omega_s)$ and $\widehat{h}_{\mathcal{S},T}(x,y;\omega_s)$ as $b\to 0$ and $T\to\infty$. This means that the usual methods used to prove consistency and asymptotic normality of thes spectral density estimator do not directly apply in this case and the rates of convergence will change.

3. Sampling properties of the estimator

In order to prove the results we require the following assumptions.

Assumption III.1 *(i) Let us suppose that $\{X_t\}$ is a strictly stationary α-mixing time series, ie.*

$$\sup_{\substack{A\in\sigma(X_0,X_{-1},\ldots)\\ B\in\sigma(X_t,X_{t+1},\ldots)}}\left|P(A\cap B)-P(A)P(B)\right| \leq C|t|^{-\alpha},$$

where $\alpha > 0$.

(ii) $\{X_t\}$ *is 2-ψ-mixing, with*

$$\sup_{\substack{A\in\sigma(X_0)\\ B\in\sigma(X_t)}} \left|\frac{P(A\cap B)}{P(A)P(B)} - 1\right| \leq Ct^{-s},$$

for some $s > 2$.

(iii) *The densities* f, f_r *and their derivatives exist.*

(iv) $\sup_{x,y}\left|\frac{\partial\{f_r(x,y)-f(x)f(y)\}}{\partial x}\right| \leq Cr^{-(1+\varepsilon)}$ *and* $\sup_{x,y}\left|\frac{\partial\{f_r(x,y)-f(x)f(y)\}}{\partial y}\right| \leq Cr^{-(1+\varepsilon)}$, *for some* $\varepsilon > 0$.

(v) *We assume that the lag window is symmetric and has the following form*

$$\lambda(x) = (1-|x|^u)\cdot 1_{(-1,1)}(x) \qquad \textit{for } u > 0.$$

We will show that under Assumption III.1 (ii) and (iii) the association spectral density $g_{\mathcal{S}}(x,y;\omega)$ is well-defined and its estimator $\widehat{g}_{\mathcal{S},T}(x,y;\omega)$ is asymptotically unbiased. Furthermore, if Assumption III.1 (iv) is satisfied, we can obtain the convergence rate for the bias of $\widehat{g}_{\mathcal{S},T}(x,y;\omega)$. Assumption III.1(i) (under some assumptions on the size α) is used to show asymptotic normality of $\widehat{g}_{\mathcal{S},T}(x,y;\omega)$.

Lemma III.2 *Suppose Assumption III.1(ii) holds. Then we have*

(i) *For all* x *and* y, *we have* $|f_r(x,y) - f(x)f(y)| \leq Cf(x)f(y)|r|^{-s}$.

(ii) *For* $r \neq 0$ *we have*

$\sup_{x,y}\left|cov\big(W_b(x-X_0), W_b(y-X_r)\big)\right| \leq C\mathbb{E}\big(W_b(x-X_0)\big)\mathbb{E}\big(W_b(y-X_0)\big)|r|^{-s}$.

(iii) $\sup_{x,y}\sum_{r\neq 0}|f_r(x,y) - f(x)f(y)| < \infty$.

To obtain the sampling properties of $\hat{g}_{\mathcal{S},T}(\cdot)$, we first replace $\bar{W}_b(x)$ and $\bar{W}_b(y)$ in $\widehat{g}_{\mathcal{S},T}(x,y;\omega)$ with its expectation and define the quantities

$$\begin{aligned}
\tilde{J}_{b,T}(x;\omega) &= \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T}\Big(\frac{1}{b}W(\frac{x-X_t}{b}) - \mathbb{E}\big(\frac{1}{b}W(\frac{x-X_t}{b})\big)\Big)\exp(it\omega) \\
\tilde{c}_r(x,y) &= \frac{1}{T}\sum_{t=1}^{T-|r|}\Big(\frac{1}{b}W(\frac{x-X_t}{b}) - \mathbb{E}\big(\frac{1}{b}W(\frac{x-X_t}{b})\big)\Big)\times \\
&\quad \Big(\frac{1}{b}W(\frac{y-X_{t+r}}{b}) - \mathbb{E}\big(\frac{1}{b}W(\frac{y-X_{t+r}}{b})\big)\Big) \\
\tilde{g}_{\mathcal{S},T}(x,y;\omega) &= \sum_k K_M(\omega-\omega_k)\tilde{J}_{b,T}(x;\omega_k)\overline{\tilde{J}_{b,T}(y;\omega_k)} \\
&- \frac{1}{2\pi T}\sum_{t=1}^{T}\frac{1}{b^2}W(\frac{x-X_t}{b})W(\frac{y-X_t}{b}).
\end{aligned}$$

The following lemma gives the bound on the difference between $\widehat{g}_{\mathcal{S},T}(x,y;\omega_k)$ and $\tilde{g}_{\mathcal{S},T}(x,y;\omega_k)$.

Lemma III.3 *Suppose Assumption III.1(i-iii),(v) holds (with $\alpha > 6$) and $M \leq T^{1/2}$. Then we have*

$$\|\widehat{g}_{\mathcal{S},T}(x,y;\omega_k) - \tilde{g}_{\mathcal{S},T}(x,y;\omega_k)\|_2 = O(\frac{M}{bT}),$$

where $\omega_k = \frac{2\pi k}{T}$.

We use the following result to obtain an expression for the mean and variance of the estimators, where it can be considered as a variant of Brillinger (1981) which covers the cumulants of DFTs of stationary time series on triangular arrays.

Lemma III.4 *Suppose Assumption III.1 (i-ii) holds with $\alpha > 6$. Then for $1 \leq k \leq T$ we have*

(i)

$$cov\big(\tilde{J}_{b,T}(x;\omega_{k_1}), \tilde{J}_{b,T}(y;\omega_{k_2})\big) = \begin{cases} g_b(x,y;\omega_{k_1}) + O(\frac{1}{T}), & \omega_{k_1} = \omega_{k_2} \\ O(\frac{1}{T}) & \omega_{k_1} \neq \omega_{k_2} \end{cases}$$

(ii)

$$cum\big(\tilde{J}_{b,T}(x;w_{k_1}), \tilde{J}_{b,T}(y;-w_{k_1}), \tilde{J}_{b,T}(x;w_{k_2}), \tilde{J}_{b,T}(y;-w_{k_2})\big) = O(\frac{1}{b^3 T})$$

Lemma III.2 implies that for $r \neq 0$,

$$\mathrm{cov}\big(\frac{1}{b}W(\frac{x-X_0}{b}), \frac{1}{b}W(\frac{y-X_r}{b})\big) \leq \mathbb{E}\big(W_b(x-X_0)\big)\mathbb{E}\big(W_b(y-X_0)\big)|r|^{-s}.$$

It is straightforward to see that $\mathrm{cov}\big(\frac{1}{b}W(\frac{x-X_0}{b}), \frac{1}{b}W(\frac{x-X_0}{b})\big) = O(\frac{1}{b})$, and this leads to $g_b(x,x;\omega) = O(\frac{1}{b})$. For $x \neq y$, $g_b(x,y;\omega) = O\big(\mathbb{E}\big(W_b(x-X_0)\big)\mathbb{E}\big(W_b(y-X_0)\big)\big)$. These results give us bounds for covariances of the DFTs in Lemma III.4.

In order to obtain expressions for the asymptotic variance we define

$$c_b(x,y;\omega) = \Re(g_b(x,y;\omega)) \quad \text{and} \quad q_b(x,y;\omega) = \Im(g_b(x,y;\omega)).$$

Lemma III.5 *Suppose Assumption III.1(i-iii) is satisfied with $\alpha > 6$, and $b^{-1} << M$. Then for $0 < \omega < \pi$ we have*

$$\begin{aligned} & var\big(\Re\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) \\ = \; & \frac{1}{2}\sum_{k=1}^{T} K_M(\omega-\omega_k)^2\big(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + c_b(x,y;\omega_k)^2 - q_b(x,y;\omega_k)^2\big) \\ + \; & O\big(\frac{1}{T} + \frac{1}{b^3 T} + \frac{M^{1/2}}{b^{5/2} T}\big), \end{aligned}$$

$$
\begin{aligned}
& var\big(\Im\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) \\
= \;& \frac{1}{2}\sum_{k=1}^{T} K_M(\omega-\omega_k)^2\big(g_b(x,x;\omega_k)g_b(y,y;\omega_k)+q_b(x,y;\omega_k)^2-c_b(x,y;\omega_k)^2\big) \\
+ \;& O\big(\frac{1}{T}+\frac{1}{b^3T}+\frac{M^{1/2}}{b^{5/2}T}\big),
\end{aligned}
$$

and

$$
\begin{aligned}
cov\big(\Re\tilde{g}_{\mathcal{S},T}(x,y;\omega),\Im\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) &= \sum_{k=1}^{T} K_M(\omega-\omega_k)^2 c_b(x,y;\omega_k)q_b(x,y;\omega_k) \\
&+ O\big(\frac{1}{T}+\frac{1}{b^3T}+\frac{M^{1/2}}{b^{5/2}T}\big).
\end{aligned}
$$

In order to estimate the above variances, we replace c_b and q_b with $\widehat{c}_{b,T}(x,y;\omega) = \Re\widehat{g}_{b,T}(x,y;\omega)$ and $\widehat{q}_{b,T}(x,y;\omega) = \Im\widehat{g}_{b,T}(x,y;\omega)$ respectively.

In the following lemma we obtain the bias of the estimator.

Lemma III.6 *Suppose Assumption III.1(ii,iii,v) is satisfied. Let us suppose that* $M \to \infty$ *and* $b \to 0$ *as* $T \to \infty$.

(i) Then we have $\mathbb{E}(\tilde{g}_{\mathcal{S},T}(x,y;\omega)) \to g_{\mathcal{S}}(x,y;\omega)$.

(ii) If in addition Assumption III.1(iv) is satisfied, then we have

$$
\mathbb{E}(\tilde{g}_{\mathcal{S},T}(x,y;\omega)) = g_{\mathcal{S}}(x,y;\omega) + O\big(\frac{1}{M^{s-1}}+\frac{1}{T}+b\big).
$$

Using Lemmas III.5 and III.6 we can obtain the mean squared error of the estimator.

Lemma III.7 *Suppose Assumption III.1 holds with* $\alpha > 6$. *Then for* $0 < \omega < \pi$ *we* *have*

$$
\begin{aligned}
& \mathbb{E}\Big(\Re\widehat{g}_{\mathcal{S},T}(x,y;\omega) - \Re g_{\mathcal{S}}(x,y;\omega)\Big)^2 \\
= \; & \frac{1}{2}\sum_{k=1}^{T} K_M(\omega-\omega_k)^2\big(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + c_b(x,y;\omega_k)^2 - q_b(x,y;\omega_k)^2\big) + \\
& O\big(\frac{1}{T} + \frac{1}{b^3T} + \frac{M^{1/2}}{b^{5/2}T} + (\frac{1}{M^{s-1}} + \frac{1}{T} + b)^2\big),
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{E}\Big(\Im\widehat{g}_{\mathcal{S},T}(x,y;\omega) - \Im g_{\mathcal{S}}(x,y;\omega)\Big)^2 \\
= \; & \frac{1}{2}\sum_{k=1}^{T} K_M(\omega-\omega_k)^2\big(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + q_b(x,y;\omega_k)^2 - c_b(x,y;\omega_k)^2\big) + \\
& O\big(\frac{1}{T} + \frac{1}{b^3T} + \frac{M^{1/2}}{b^{5/2}T} + (\frac{1}{M^{s-1}} + \frac{1}{T} + b)^2\big),
\end{aligned}
$$

In the following result we show asymptotic normality, this allows us to obtain pointwise confidence intervals for $g_{\mathcal{S}}$.

Theorem III.1 *Let us suppose that Assumption III.1 holds (with $\alpha > 14$), $b^{-1} << M$ and $\sqrt{\frac{b^2T}{M}}\big(\frac{M^{1/2}}{b^{5/2}T} + b^2\big) \to 0$. Then for $0 < \omega < \pi$ we have*

$$
V_{T,b}(x,y;\omega)^{-1/2}\begin{pmatrix} \Re\widehat{g}_{\mathcal{S},T}(x,y;\omega) - \Re g_{\mathcal{S}}(x,y;\omega) \\ \Im\widehat{g}_{\mathcal{S},T}(x,y;\omega) - \Im g_{\mathcal{S}}(x,y;\omega) \end{pmatrix} \xrightarrow{D} \mathcal{N}\big(0, I_2\big)
$$

where

$$
V_{T,b}(x,y;\omega) = \sum_{k=1}^{T} K_M(\omega-\omega_k)^2 \begin{pmatrix} A_b(x,y;\omega_k) & C_b(x,y;\omega_k) \\ C_b(x,y;\omega_k) & B_b(x,y;\omega_k) \end{pmatrix},
$$

$$
\begin{aligned}
A_b(x,y;\omega_k) &= \frac{1}{2}(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + c_b(x,y;\omega_k)^2 - q_b(x,y;\omega_k)^2), \\
B_b(x,y;\omega_k) &= \frac{1}{2}(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + q_b(x,y;\omega_k)^2 - c_b(x,y;\omega_k)^2),
\end{aligned}
$$

and

$$C_b(x,y;\omega_k) = c_b(x,y;\omega_k) q_b(x,y;\omega_k).$$

Using the theorem above and that $\bar{W}_b(x)\bar{W}_b(y) \xrightarrow{\mathcal{P}} f(x)f(y)$ the asymptotic normality of the standardized association density immediately follows.

Corollary III.1 *Let us suppose that Assumption III.1 holds (with $\alpha > 14$), $f(x) > 0$, $f(y) > 0$, $b^{-1} << M$ and $\sqrt{\frac{b^2 T}{M}}\big(\frac{M^{1/2}}{b^{5/2}T} + b^2\big) \to 0$. Then for $0 < \omega < \pi$ we have*

$$f(x)f(y)V_{T,b}(x,y;\omega)^{-1/2} \begin{pmatrix} \Re\widehat{h}_{\mathcal{S},T}(x,y;\omega) - \Re h_{\mathcal{S}}(x,y;\omega) \\ \Im\widehat{h}_{\mathcal{S},T}(x,y;\omega) - \Im h_{\mathcal{S}}(x,y;\omega) \end{pmatrix} \xrightarrow{D} \mathcal{N}\big(0, I_2\big),$$

where $V_{T,b}(x,y;\omega)$ is defined in Theorem III.1.

The above results allow us to make piecewise confidence intervals for $g_{\mathcal{S}}(\cdot)$ and $h_{\mathcal{S}}(\cdot)$.

4. Simulations

In this section we conduct a simulation study. In order to see the convergence of asymptotic normality of the estimators, we construct the pointwise confidence intervals using Theorem III.1 Corolloary III.1. We consider two AR(1) model $X_t = \phi X_{t-1} + \varepsilon_t$, $\phi = 0.9, 0.6$, where $\{\varepsilon_t\}$ iid standard Gaussian random variables. In our simulations we use the Bartlett window for $\lambda_M(\cdot)$ and Epanechnikov kernel for $W(\cdot)$ and b was set as $T^{-1/13}$.

In Figure 15, 16 17 and 18 we plot the the confidence intervals and their true values. The confidence intervals for $\Re g_{\mathcal{S}}$ are on the diagonal and lower triangle and for $\Im g_{\mathcal{S}}$ on the upper triangle. In these all plots, we could see that the confidence intervals captures the true values fairly well with moderate sample sizes confirming the asymptotic normality.

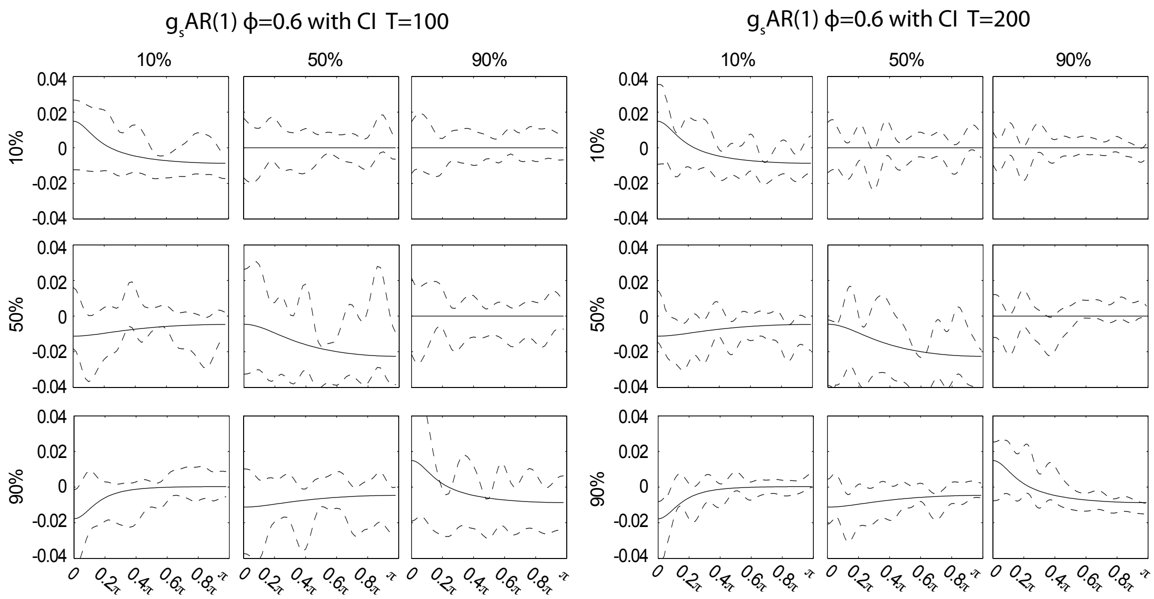

Fig. 15. The confidence intervals of $g_{\mathcal{S}}$ in $X_t = 0.6X_{t-1} + \varepsilon_t$; (Left) $T = 100$, (Right) $T = 200$

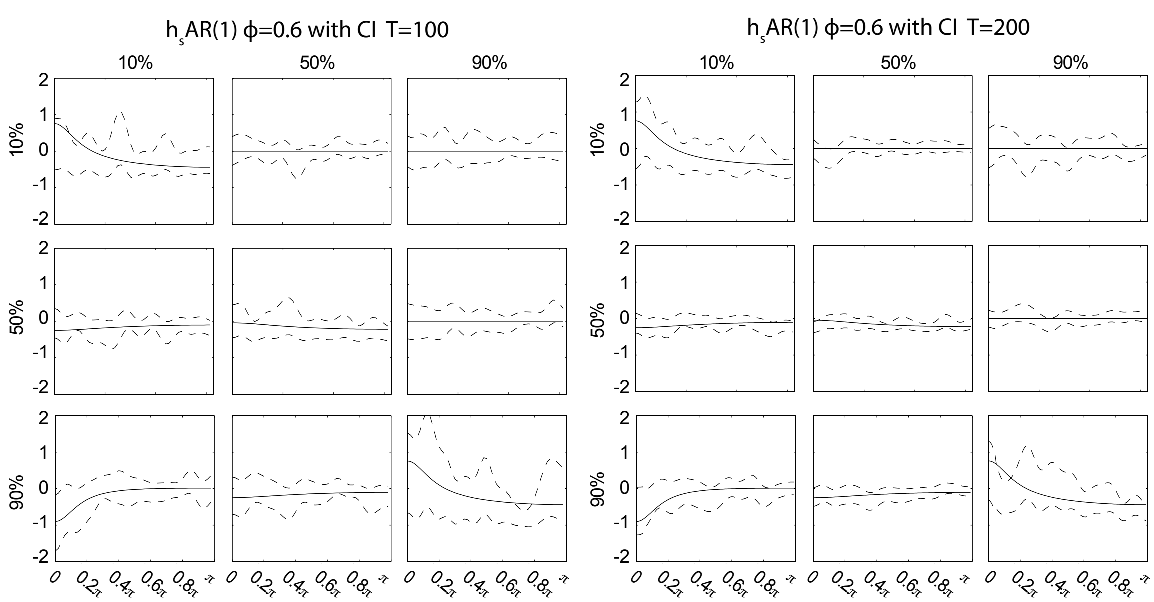

Fig. 16. The confidence intervals of $h_{\mathcal{S}}$ in $X_t = 0.6X_{t-1} + \varepsilon_t$; (Left) $T = 100$, (Right) $T = 200$

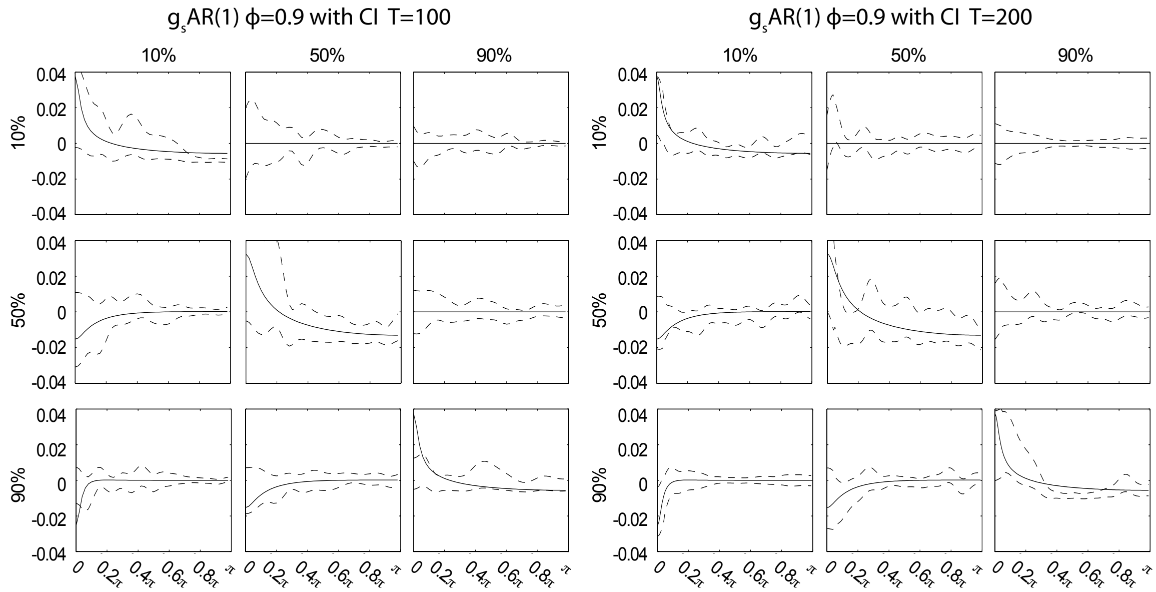

Fig. 17. The confidence intervals of $g_{\mathcal{S}}$ in $X_t = 0.9X_{t-1} + \varepsilon_t$; (Left) $T = 100$, (Right) $T = 200$

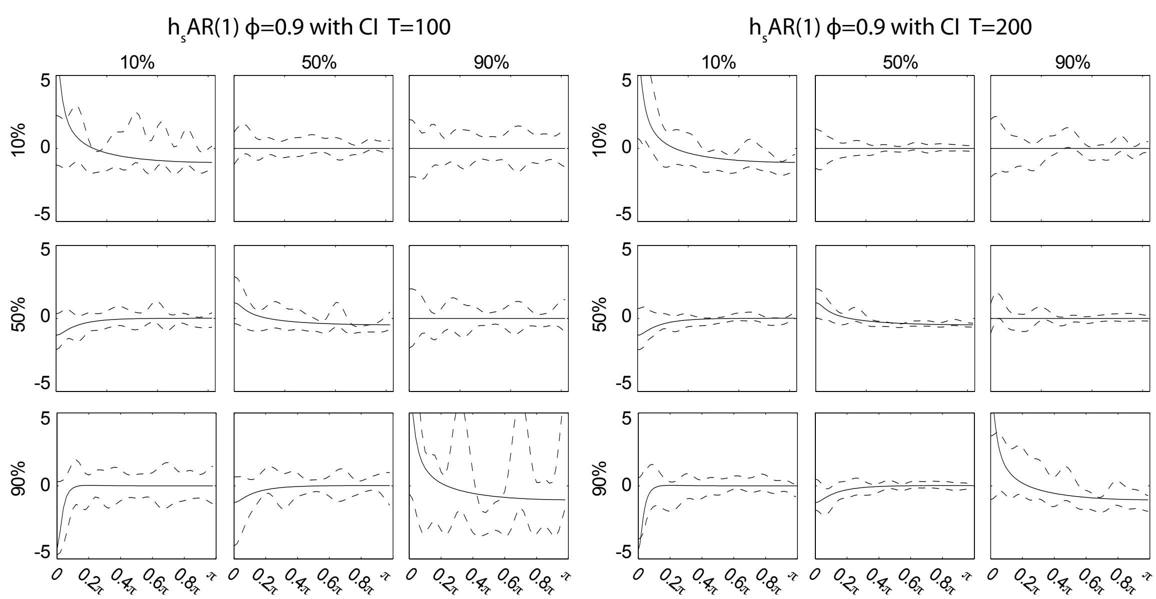

Fig. 18. The confidence intervals of $h_{\mathcal{S}}$ in $X_t = 0.9X_{t-1} + \varepsilon_t$; (Left) $T = 100$, (Right) $T = 200$

5. Proofs

5.1. Proof of Lemma III.2

The proof of Lemma III.2 (i) hinges on the relationship between the distribution function and the density function. We define the function H such that $H(x,y) = G(X \leq x, Y \leq y)$. Using this, we recall that the partial derivatives of H are defined as

$$
\begin{aligned}
\frac{\partial H(x,y)}{\partial x} &= \lim_{h\to 0}\frac{1}{h}\Big(H(x+h,y) - H(x,y)\Big) \\
&= \lim_{h\to 0}\frac{1}{h}G(x \leq X \leq x+h, Y \leq y) \\
\frac{\partial^2 H(x,y)}{\partial x \partial y} &= \lim_{\substack{h_1\to 0\\ h_2\to 0}}\frac{1}{h_1h_2}\Big(\big\{H(x+h_1,y+h_2) - H(x,y+h_2)\big\} \\
&- \big\{H(x+h_1,y) - H(x,y)\big\}\Big) \\
&= \lim_{\substack{h_1\to 0\\ h_2\to 0}}\frac{1}{h_1h_2}G(x \leq X \leq x+h_1, y \leq Y \leq y+h_2). \qquad (3.6)
\end{aligned}
$$

We now use the above to prove (i). We recall the definition of $f_r(x,y) - f(x)f(y)$ is

$$
f_r(x,y) - f(x)f(y) = \frac{\partial^2 F_r(x,y)}{\partial x \partial y} - \frac{\partial^2 F(x)F(y)}{\partial x \partial y}.
$$

Let $H(x,y) = G(X_0 \leq x, X_r \leq y) = P(X_0 \leq x, X_r \leq y) - P(X_0 \leq x)P(X_r \leq y)$. Under Assumption III.1(ii), for all h_1 and h_2, we have the following bound

$$
\begin{aligned}
& |G(x \leq X_0 \leq x+h_1, y \leq X_r \leq y+h_2)| \\
\leq \;& CP(x \leq X \leq x+h_1)P(y \leq X_r \leq y+h_2)|r|^{-s}.
\end{aligned}
$$

Therefore, by substituting the above into (3.6), and letting $h_1, h_2 \to 0$, we have the required result.

To prove Lemma III.2 (ii) we use the bound of Lemma III.2 (i) in the definition of $\mathrm{cov}\big(W_b(x-X_0), W_b(y-X_r)\big)$ to give

$$\begin{aligned}
&\big|\mathrm{cov}\big(W_b(x-X_0), W_b(y-X_r)\big)\big| \\
= \;& \Big|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{b^2}W(\frac{x-z_1}{b})W(\frac{y-z_2}{b})\big(f_r(z_1,z_2)-f(z_1)f(z_2)\big)dz_1dz_2\Big| \\
\leq \;& C|r|^{-s}\int\int\frac{1}{b^2}W(\frac{x-z_1}{b})W(\frac{y-z_2}{b})f(z_1)f(z_2)dz_1dz_2 \\
\leq \;& C\mathbb{E}\big(\frac{1}{b}W(\frac{x-X_0}{b})\big)\mathbb{E}\big(\frac{1}{b}W(\frac{y-X_0}{b})\big)|r|^{-s},
\end{aligned}$$

as desired.

The proof of Lemma III.2 (iii) follows immediately from Lemma III.2 (i).

5.2. Proof of Lemma III.3

We first observe that

$$\begin{aligned}
& J_{b,T}(x;\omega_k)\overline{J_{b,T}(y;\omega_k)} - \tilde{J}_{b,T}(x;\omega_k)\overline{\tilde{J}_{b,T}(y;\omega_k)} \\
= & \begin{cases} 0 & \omega_k \neq 0,\pi \\ T\big(\bar{W}_b(x)-\mathbb{E}(W_b(x-X_0))\big)\big(\bar{W}_b(y)-\mathbb{E}(W_b(y-X_0))\big) & \text{otherwise} \end{cases}.
\end{aligned}$$

Substituting the above into $\tilde{g}_{\mathcal{S},T}(x,y;\omega_s)-\widehat{g}_{\mathcal{S},T}(x,y;\omega_s)$ gives

$$\begin{aligned}
&\tilde{g}_{\mathcal{S},T}(x,y;\omega_s)-\widehat{g}_{\mathcal{S},T}(x,y;\omega_s) \\
&= TK_M(\omega)\big(\bar{W}_b(x)-\mathbb{E}(W_b(x-X_0))\big)\big(\bar{W}_b(y)-\mathbb{E}(W_b(y-X_0))\big).
\end{aligned}$$

Using $\sup_x \mathbb{E}\big(\bar{W}_b(x)-\mathbb{E}(W_b(x-X_0))\big)^2 = O(\frac{1}{bT})$ and $K_M(\cdot)=O(\frac{M}{T})$, we obtain the required result. $\square$

5.3. Proof of Lemma III.4

To prove Lemma III.4, we first prove the following lemma which gives the bound on the $k-$th order cumulant of $\{\frac{1}{b}W(\frac{x-X_t}{b})\}$ and its summation.

Lemma III.8 *Let us suppose that Assumption III.1(i) is satisfied with the mixing size $\alpha > 2k$.*

(i) Then for $0 = t_0 \leq t_1 \leq \ldots \leq t_k$, the $(k+1)$-th order cumulant is

$$\begin{aligned}&\left|cum(\frac{1}{b}W(\frac{z_0 - X_0}{b}), \frac{1}{b}W(\frac{z_1 - X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k - X_{t_k}}{b}))\right| \\ &\leq C \min\left(\frac{1}{b^{k-s}}, \frac{1}{b^{k+1}} \prod_{i=1}^{k} |t_i - t_{i-1}|^{-\alpha/k}\right)\end{aligned}$$

where s are the number of different $\{t_i\}_{i=1}^k$ (for example, if t_i are all different, then $s = k$, on the other hand if all are the same then $s = 0$).

(ii) Moreover

$$\sum_{t_1,\ldots,t_k=-\infty}^{\infty} \left|cum(\frac{1}{b}W(\frac{z_0 - X_0}{b}), \frac{1}{b}W(\frac{z_1 - X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k - X_{t_k}}{b}))\right| \leq \frac{1}{b^k}.$$

PROOF. We first prove (i). To prove the first part of the inequality on the RHS we treat $\text{cum}\big(\frac{1}{b}W(\frac{z_0-X_0}{b}), \frac{1}{b}W(\frac{z_1-X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k-X_{t_k}}{b})\big)$ as an integral. Let us suppose that the distinct $(X_0, X_{t_1}, \ldots, X_{t_k})$ are $(X_0, \ldots, X_{r_s})$ and the joint densities of $(X_0, X_{t_1}, \ldots, X_{r_s})$ are bounded, then by a change of variables it is straightforward to show that

$$\left|\text{cum}(\frac{1}{b}W(\frac{z_0 - X_0}{b}), \frac{1}{b}W(\frac{z_1 - X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k - X_{t_k}}{b}))\right| \leq \frac{C}{b^{k-s}}.$$

To obtain the second bound on the RHS we use Statulevicius and Jakimavicius (1988), Theorem 3, part (1a), where it is shown that for every t_i we have

$$\begin{aligned}&\left|\text{cum}(\frac{1}{b}W(\frac{z_0 - X_0}{b}), \frac{1}{b}W(\frac{z_1 - X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k - X_{t_k}}{b}))\right| \\ &\leq C\frac{k!}{b^{k+1}(k+1)} \max_x |W(x)|^{k+1} |t_{i+1} - t_i|^{-\alpha}.\end{aligned}$$

Taking the k-th root of the above and applying it to every t_i we obtain the second bound on the RHS of (i).

We start to prove (ii) by partitioning the summand

$$
\begin{aligned}
&\sum_{t_1,\ldots,t_k=-\infty}^{\infty} \big|\mathrm{cum}(\frac{1}{b}W(\frac{z_0-X_0}{b}), \frac{1}{b}W(\frac{z_1-X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k-X_{t_k}}{b}))\big| \\
= & \sum_{s=0}^{k} C_s \sum_{t_1<\ldots<t_s} \big|\mathrm{cum}(\frac{1}{b}W(\frac{z_0-X_0}{b}), \frac{1}{b}W(\frac{z_1-X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_s-X_{t_s}}{b}))\big| \\
\leq & \sum_{s=0}^{k} C_s \sum_{0<t_1<\ldots<t_s} \min\left(\frac{1}{b^{k-s}}, \frac{1}{b^{k+1}} \prod_{i=1}^{k} |t_i - t_{i-1}|^{-\alpha/k}\right),
\end{aligned}
$$

where $\{C_s\}$ are finite constants. Considering the inner summand of the above term, for all $h > 0$ we obtain the bound

$$
\begin{aligned}
&\sum_{0<t_1<\ldots<t_s} \min\left(\frac{1}{b^{k-s}}, \frac{1}{b^{k+1}} \prod_{i=1}^{k} |t_i - t_{i-1}|^{-\alpha/k}\right) \\
\leq & \sum_{r_1,\ldots,r_s} \min\left(\frac{1}{b^{k-s}}, \frac{1}{b^{k+1}} \prod_{i=1}^{k} |r_i|^{-\alpha/k}\right) \\
\leq & \sum_{r_1,\ldots,r_s=1}^{h} \frac{1}{b^{k-s}} + \sum_{r_1, \text{ or } ,r_2, \text{ or } ,\ldots,r_s>h} \frac{1}{b^{k+1}} \prod_{i=1}^{k} |r_i|^{-\alpha/k} \\
\leq & \left(\frac{h^s}{b^{k-s}} + \frac{h^{-\alpha/k+1}}{b^{k+1}}\right).
\end{aligned}
$$

As the above bound holds for all h, let $h = b^{-1}$, this gives the bound

$$
\sum_{0<t_1<\ldots<t_s} \min\left(\frac{1}{b^{k-s}}, \frac{1}{b^{k+1}} \prod_{i=1}^{k} |t_{i+1} - t_i|^{-\alpha/k}\right) \leq C(\frac{1}{b^k} + \frac{b^{\alpha/k-1}}{b^{k+1}}).
$$

Thus, by assumption we have $\alpha > 2k$ which gives

$$
\sum_{t_1,\ldots,t_k=\infty}^{\infty} \big|\mathrm{cum}(\frac{1}{b}W(\frac{z_0-X_0}{b}), \frac{1}{b}W(\frac{z_0-X_{t_1}}{b}), \ldots, \frac{1}{b}W(\frac{z_k-X_{t_k}}{b}))\big| \leq \frac{C}{b^k},
$$

which gives the desired result. $\square$

PROOF of Lemma III.4 To prove (i), we expand $\mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(y;\omega_{k_2})\big)$

$$\begin{aligned}
& \mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(y;\omega_{k_2})\big) \\
= & \frac{1}{T}\sum_{t,\tau=1}^{T}\mathrm{cov}\big(W_b(x-X_t),W_b(y-X_\tau)\big)\exp(it\omega_{k_1}-i\tau\omega_{k_2}) \\
= & \sum_{r=-(T-1)}^{T-1}\Big[\big\{\mathrm{cov}\big(W_b(x-X_0),W_b(y-X_r)\big)\exp(-ir\omega_{k_2})\big\}\times \\
& \big\{\frac{1}{T}\sum_{t=\max(1,1-r)}^{\min(T,T-r)}\exp(it(\omega_{k_1}-\omega_{k_2}))\big\}\Big] \\
= & \sum_{r=-(T-1)}^{T-1}\Big[\big\{\mathrm{cov}\big(W_b(x-X_0),W_b(y-X_r)\big)\exp(-ir\omega_{k_2})\big\}\times \\
& \big\{\frac{1}{T}\sum_{t=1}^{T}\exp(it(\omega_{k_1}-\omega_{k_2}))\big\}\Big]+O(\frac{1}{T}) \\
= & \begin{cases} \sum_{r=-(T-1)}^{T-1}\mathrm{cov}\big(W_b(x-X_0),W_b(y-X_r)\big)\exp(-ir\omega_{k_2})+O(\frac{1}{T}) & \omega_{k_1}=\omega_{k_2} \\ O(\frac{1}{T}) & \omega_{k_1}\neq\omega_{k_2} \end{cases}
\end{aligned}$$

For $k_1=k_2$, we obtain

$$\begin{aligned}
& |g_b(x,y;\omega_{k_1})-\mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(y;\omega_{k_1})\big)| \\
\leq & \sum_{|r|\geq T}|\mathrm{cov}\big(W_b(x-X_0),W_b(y-X_r)\big)|+O(\frac{1}{T}) \\
\leq & C\cdot\sum_{|r|\geq T}|r|^{-s}+O(\frac{1}{T}) \\
= & O(\frac{1}{T}).
\end{aligned}$$

To prove (ii), we use Lemma IV.1 (ii) with $k = 3$, and this gives

$$\begin{aligned}
& \left|\text{cum}(\tilde{J}_{b,T}(x; w_{k_1}), \tilde{J}_{b,T}(y; -w_{k_1}), \tilde{J}_{b,T}(x; w_{k_2}), \tilde{J}_{b,T}(y; -w_{k_2}))\right| \\
\leq \; & \frac{1}{(2\pi T)^2} \sum_{t_1,t_2,t_3,t_4=1}^{T} \left|\text{cum}(W_b(x - X_{t_1}), W_b(x - X_{t_2}), W_b(y - X_{t_3}), W_b(x - X_{t_4}))\right| \\
= \; & O(\frac{1}{b^3 T}).
\end{aligned}$$

□

5.4. Proof of Lemma III.5

We introduce a quantity comparable to $\tilde{g}_{\mathcal{S},T}(\cdot)$, since its asymptotic variance could be obtained easier than $\tilde{g}_{\mathcal{S},T}(\cdot)$. Let

$$\tilde{g}_{b,T}(x, y; \omega) = \sum_k K_M(\omega - \omega_k) \tilde{J}_{b,T}(x; \omega_k) \overline{\tilde{J}_{b,T}(y; \omega_k)}.$$

In the following lemma we show that the variance of $\tilde{g}_{\mathcal{S},T}(\cdot)$ and $\tilde{g}_{b,T}(\cdot)$, are asymptotically equivalent, and in Lemma we obtain the variance and the covariance of $\tilde{g}_{\mathcal{S},T}(\cdot)$.

Lemma III.9 *Suppose Assumption III.1(i,ii,v) is satisfied with $\alpha > 6$ and $b^{-1} << M$. Then we have*

$$var\big(\tilde{g}_{\mathcal{S},T}(x, y; \omega)\big) = var\big(\tilde{g}_{b,T}\big(x, y; \omega)\big) + O(\frac{M^{1/2}}{b^{5/2} T})$$

PROOF. We recall that

$$\begin{aligned}
& \tilde{g}_{\mathcal{S},T}(x, y; \omega) \\
= \; & \sum_{k=1}^{T} K_M(\omega - \omega_k) \tilde{J}_{b,T}(x; \omega_k) \overline{\tilde{J}_{b,T}(y; \omega_k)} - \frac{1}{2\pi}\big(\tilde{c}_0(x, y) + \bar{W}_b(x)\bar{W}_b(y)\big) \\
= \; & \tilde{g}_{b,T}(x, y; \omega) - \frac{1}{2\pi}\big(\tilde{c}_0(x, y) + \bar{W}_b(x)\bar{W}_b(y)\big) \qquad (3.7)
\end{aligned}$$

We now obtain the variance of $\tilde{c}_0(x,y)$ and $\bar{W}_b(x)\bar{W}_b(y)$. We first note that

$$\tilde{c}_0(x,y) = \frac{2\pi}{T}\sum_{k=1}^{T}\tilde{J}_{b,T}(x;\omega_k)\overline{\tilde{J}_{b,T}(y;\omega_k)} \quad (3.8)$$

$$\bar{W}_b(x) = \sqrt{\frac{2\pi}{T}}\tilde{J}_{b,T}(x;0) + \mathbb{E}\big(W_b(x-X_0)\big). \quad (3.9)$$

Combining the results of Lemma III.4 with (3.8), we obtain the bound for $\mathrm{var}(\tilde{c}_0(x,y))$.

$$\begin{aligned}
&\mathrm{var}\big(\tilde{c}_0(x,y)\big)\\
=\ & \frac{(2\pi)^2}{T^2}\Big(\sum_{k_1,k_2}\mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(x;\omega_{k_2})\big)\mathrm{cov}\big(\tilde{J}_{b,T}(y;-\omega_{k_1}),\tilde{J}_{b,T}(y;-\omega_{k_2})\big)\\
+\ & \sum_{k_1,k_2}\mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(x;-\omega_{k_2})\big)\mathrm{cov}\big(\tilde{J}_{b,T}(y;-\omega_{k_1}),\tilde{J}_{b,T}(y;\omega_{k_2})\big)\\
+\ & \sum_{k_1,k_2}\mathrm{cum}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(x;-\omega_{k_1}),\tilde{J}_{b,T}(y;\omega_{k_2}),\tilde{J}_{b,T}(y;-\omega_{k_2})\big)\Big)\\
=\ & O(\frac{1}{b^2T}+\frac{1}{b^3T}) = O(\frac{1}{b^3T}). \quad (3.10)
\end{aligned}$$

To obtain the order of $\mathrm{var}(\bar{W}_b(x)\bar{W}_b(y))$, we use (3.9).

$$\begin{aligned}
&\mathrm{var}\big(\bar{W}_b(x)\bar{W}_b(y)\big) = \quad (3.11)\\
&\mathrm{var}\big(\frac{2\pi}{T}\tilde{J}_T(x;0)\tilde{J}_T(y;0)+\sqrt{\frac{2\pi}{T}}\{\tilde{J}_T(x;0)\mathbb{E}\big(W_b(y-X_0)\big)+\tilde{J}_T(y;0)\mathbb{E}\big(W_b(x-X_0)\big)\}\big)
\end{aligned}$$

With straightforward application of Lemma III.4, we find that

$$\begin{aligned}
\mathrm{var}\big(\tilde{J}_T(x;0)\tilde{J}_T(y;0)\big) &= O(\frac{1}{b^2})\\
\mathrm{var}\big(\tilde{J}_T(x;0)\big) &= O(\frac{1}{b}).
\end{aligned}$$

Plugging the above bounds into (3.11) leads to

$$\mathrm{var}\big(\bar{W}_b(x)\bar{W}_b(y)\big) = O(\frac{1}{bT}). \quad (3.12)$$

Therefore, expanding the expression (3.7) with the bounds in (3.10), (3.12) and $\mathrm{var}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) = O(\frac{M}{b^2T})$ gives us

$$\begin{aligned}
& \mathrm{var}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) \\
= \;& \mathrm{var}\big(\tilde{g}_{b,T}(x,y;\omega)\big) + \frac{1}{(2\pi)^2}\big\{\mathrm{var}(\tilde{c}_0(x,y)) + \mathrm{var}\big(\bar{W}_b(x)\bar{W}_b(y)\big)\big\} \\
+ \;& \frac{1}{2\pi}\big\{\mathrm{cov}\big(\tilde{g}_{b,T}(x,y;\omega),\tilde{c}_0(x,y)\big) + \mathrm{cov}\big(\tilde{g}_{b,T}(x,y;\omega),\bar{W}_b(x)\bar{W}_b(y)\big) \\
+ \;& \mathrm{cov}\big(\tilde{c}_0(x,y),\bar{W}_b(x)\bar{W}_b(y)\big)\big\} \\
= \;& \mathrm{var}(\tilde{g}_{b,T}(x,y;\omega)) + O(\frac{1}{b^3T}) + O(\frac{1}{bT}) \\
+ \;& O(\frac{M^{1/2}}{bT^{1/2}}\cdot\frac{1}{b^{3/2}T^{1/2}}) + O(\frac{M^{1/2}}{bT^{1/2}}\cdot\frac{1}{b^{1/2}T^{1/2}}) + O(\frac{1}{b^{3/2}T^{1/2}}\cdot\frac{1}{b^{1/2}T^{1/2}}) \\
= \;& \mathrm{var}(\tilde{g}_{b,T}(x,y;\omega)) + O(\frac{M^{1/2}}{b^{5/2}T})
\end{aligned}$$

which is the desired result. $\square$

Lemma III.10 *Suppose Assumption III.1(i-iii) is satisfied with $\alpha > 6$, and $b^{-1} << M$. Then we have*

(i)

$$var\big(\tilde{g}_{b,T}(x,y;\omega)\big) = \begin{cases} \sum_{k=1}^{T} K_M(\omega-\omega_k)^2 g_b(x,x;\omega_k)g_b(y,y;\omega_k) + O\big(\frac{1}{b^3T}\big) & 0<\omega<\pi \\ \sum_{k=1}^{T} K_M(-\omega_k)^2\big(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + |g_b(x,y;\omega_k)|^2\big) + O\big(\frac{1}{b^3T}\big) & \omega = 0 \end{cases}$$

(ii)

$$cov\big(\tilde{g}_{b,T}(x,y;\omega), \overline{\tilde{g}_{b,T}(x,y;\omega)}\big) = \begin{cases} \sum_{k=1}^{T} K_M(\omega-\omega_k)^2 g_b(x,y;\omega_k)^2 + O\big(\frac{1}{b^3T}\big) & 0<\omega<\pi \\ \sum_{k=1}^{T} K_M(-\omega_k)^2\big(g_b(x,x;\omega_k)g_b(y,y;\omega_k) + g_b(x,y;\omega_k)^2\big) + O\big(\frac{1}{b^3T}\big) & \omega = 0 \end{cases}$$

PROOF. We first prove $\mathrm{var}\big(\tilde{g}_{b,T}(x,y;\omega)\big)$. By expanding $\mathrm{var}(\tilde{g}_{b,T}(x,y;\omega))$ we have

$$\mathrm{var}\big(\tilde{g}_{b,T}(x,y;\omega)\big) = I + II + III,$$

where

$$\begin{aligned}
I &= \sum_{k_1,k_2=1}^{T} K_M(\omega-\omega_{k_1})K_M(\omega-\omega_{k_2}) \\
&\quad \mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(x;\omega_{k_2})\big)\mathrm{cov}\big(\overline{\tilde{J}_{b,T}(y;\omega_{k_1})},\overline{\tilde{J}_{b,T}(y;\omega_{k_2})}\big) \\
II &= \sum_{k_1,k_2=1}^{T} K_M(\omega-\omega_{k_1})K_M(\omega-\omega_{k_2}) \\
&\quad \mathrm{cov}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\overline{\tilde{J}_{b,T}(y;\omega_{k_2})}\big)\mathrm{cov}\big(\overline{\tilde{J}_{b,T}(y;\omega_{k_1})},\tilde{J}_{b,T}(x;\omega_{k_2})\big) \\
III &= \sum_{k_1,k_2=1}^{T} K_M(\omega-\omega_{k_1})K_M(\omega-\omega_{k_2}) \\
&\quad \mathrm{cum}\big(\tilde{J}_{b,T}(x;\omega_{k_1}),\tilde{J}_{b,T}(y;-\omega_{k_1}),\tilde{J}_{b,T}(x;\omega_{k_2}),\tilde{J}_{b,T}(y;-\omega_{k_2})\big).
\end{aligned}$$

With Lemma III.4 (i), we obtain limiting expressions for I and II

$$\begin{aligned}
I &= \sum_{k} K_M(\omega-\omega_k)^2 g_b(x,x;\omega_k)g_b(y,y;\omega_k) + O(\frac{1}{T}) \\
II &= \sum_{k} K_M(\omega-\omega_k)K_M(\omega-\omega_{T-k})|g_b(x,y;\omega_k)|^2 + O(\frac{1}{T}) \\
&= \begin{cases} O(\frac{1}{T}) & 0<\omega<\pi \\ \sum_k K_M(\omega_k)K_M(\omega_{T-k})|g_b(x,y;\omega_k)|^2 + O(\frac{1}{T}) & \omega = 0 \end{cases}
\end{aligned}$$

, and Lemma III.4 (ii) immedately gives the bound $O(\frac{1}{b^3T})$ to III. This proves (i), and the proof of (ii) is similar and we omit the details. $\square$

We use Lemma III.9 and Lemma III.10 to prove Lemma III.5.

PROOF of Lemma III.5 By using that

$$\begin{aligned}\tilde{c}_{\mathcal{S},T}(x,y;\omega) &= \Re\tilde{g}_{\mathcal{S},T}(x,y;\omega) = \frac{1}{2}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega) + \overline{\tilde{g}_{\mathcal{S},T}(x,y;\omega)}\big) \\ \tilde{q}_{\mathcal{S},T}(x,y;\omega) &= \Im\tilde{g}_{\mathcal{S},T}(x,y;\omega) = \frac{i}{2}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega) - \overline{\tilde{g}_{\mathcal{S},T}(x,y;\omega)}\big),\end{aligned}$$

Lemma III.5 is an immediate corrolary from Lemma III.9 and III.10. □

5.5. Proof of Lemma III.6

Lemma III.11 *Suppose Assumption III.1 (ii,iii,v) holds, then*

$$\mathbb{E}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) = \int\int \frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})g_{\mathcal{S}}(u,v;\omega)dudv + O\big(\frac{1}{M^{s-1}} + \frac{1}{T}\big).$$

Furthermore if Assumption III.1 (iv) is satisfied, then we have

$$\Big|\int\int W_b(x-u)W_b(y-v)g_{\mathcal{S}}(u,v;\omega)dudv - g_{\mathcal{S}}(x,y;\omega)\Big| = O(b). \tag{3.13}$$

PROOF. To prove the result we observe that

$$\begin{aligned}&\mathbb{E}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) = \frac{1}{2\pi}\int\int\frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b}) \\ \times\ &\Big(\sum_{r\neq 0}\lambda(\frac{r}{M})(\frac{T-|r|}{T})\{f_r(u,v) - f(u)f(v)\}\exp(ir\omega) - f(u)f(v)\Big)dudv.\end{aligned}$$

Using this expansion we have

$$\Big|\mathbb{E}\big(\tilde{g}_{\mathcal{S},T}(x,y;\omega)\big) - \int\int\frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})g_{\mathcal{S}}(u,v;\omega)dudv\Big| = I + II,$$

where

$$
\begin{aligned}
I &= \frac{1}{2\pi}\int\int \frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b}) \\
&\times \Bigg(\sum_{0<|r|\leq M}\big(\lambda(\frac{r}{M})(\frac{T-|r|}{T})-1\big)\{f_r(u,v)-f(u)f(v)\}\exp(ir\omega)\Bigg)dudv. \\
II &= \frac{1}{2\pi}\int\int \frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})\sum_{|r|>M}\exp(ir\omega)\{f_r(u,v)-f(u)f(v)\}dudv.
\end{aligned}
$$

Thus we have the bounds

$$
\begin{aligned}
I &\leq C\sup_{u,v}\sum_{0<|r|\leq M}\big|\lambda(\frac{r}{M})(\frac{T-|r|}{T})-1\big|\big|f_r(u,v)-f(u)f(v)\big| \\
II &\leq C\sup_{u,v}\sum_{|r|>M}\big|f_r(u,v)-f(u)f(v)\big|.
\end{aligned}
$$

Finally, we use Lemma III.2(i) and that $\lambda(\frac{r}{M})=1-|\frac{r}{M}|^u$ to obtain

$$
\begin{aligned}
I &\leq C\sum_{r=-M}^{M}|\lambda(\frac{r}{M})-1||r|^{-(2+\varepsilon)}+\frac{C}{T}\sum_{r=-M}^{M}|r\lambda(\frac{r}{M})|\cdot|r|^{-s} \\
&\leq C\sum_{r=-M}^{M}|\frac{r}{M}|^u|r|^{-s}+\frac{C}{T}\sum_{r=-M}^{M}|r|^{-(s-1)}=O(\frac{1}{M^s}+\frac{1}{T}).
\end{aligned}
$$

To bound II we use Lemma III.2(i) to obtain

$$
|II| \leq C\cdot\sum_{|r|>M}|r|^{-s}=O(\frac{1}{M^{s-1}}).
$$

The above bounds for I and II give the desired result.

To prove (3.13) we make a Taylor expansion of $g_{\mathcal{S}}(u,v;\omega)$ about (x,y) to give

$$
\begin{aligned}
&g_{\mathcal{S}}(u,v;\omega)-g_{\mathcal{S}}(x,y;\omega) \\
=\ &(u-x)\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial x}\rfloor_{(x,y)=(\bar{x},\bar{y})}+(v-y)\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial y}\rfloor_{(x,y)=(\tilde{x},\tilde{y})}. \qquad (3.14)
\end{aligned}
$$

Now under Assumption III.1(iv), by exchanging sum and derivative, we have

$$\begin{aligned} \left|\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial x}\right| &= \left|\sum_{r\neq 0}\left(\frac{\partial f_r(x,y)}{\partial x} - f(y)\frac{\partial f(x)}{\partial x}\right)\exp(ir\omega) + \frac{\partial f(x)f(y)}{\partial x}\right| \\ &\leq K\cdot\sum_{r\neq 0}|r|^{-(1+\varepsilon)} < \infty. \end{aligned}$$

and a similar bound holds for $\left|\frac{\partial g(x,y;\omega)}{\partial y}\right|$. Plugging the above bounds into (3.14) leads to

$$\begin{aligned} &\left|\int\int\frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})\big(g_{\mathcal{S}}(u,v;\omega) - g_{\mathcal{S}}(x,y;\omega)\big)\right|dudv \\ \leq\ & \sup_{x,y}\left|\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial x}\right|\int\int\frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})|x-u|dudv \\ &+\sup_{x,y}\left|\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial y}\right|\int\int\frac{1}{b^2}W(\frac{x-u}{b})W(\frac{y-v}{b})|y-v|dudv \\ \leq\ & C\big(\sup_{x,y}\left|\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial y}\right| + \sup_{x,y}\left|\frac{\partial g_{\mathcal{S}}(x,y;\omega)}{\partial x}\right|\big)b, \end{aligned}$$

where C is a finite constant. Thus we have (3.13). □

PROOF of Lemma III.6 The proof of Lemma III.6 follows immediately from Lemma III.11. □

5.6. Proof of Theorem III.1

We now show asymptotic normality of $\tilde{g}_{\mathcal{S},T}(x,y,\omega)$. To do so, we define the partial sum

$$B^{(u)}_{T,S_T} = \frac{1}{T}\sum_{t=u+1}^{S_T+u}\sum_{\tau\neq t}\lambda(\frac{t-\tau}{M})\Big(Z_{b,t}(x)Z_{b,\tau}(y) - \mathbb{E}(Z_{b,t}(x)Z_{b,\tau}(y))\Big)\exp(i(t-\tau)\omega) \tag{3.15}$$

where $Z_{b,t}(x) = \frac{1}{b}W(\frac{x-X_t}{b}) - \mathbb{E}(\frac{1}{b}W(\frac{x-X_t}{b}))$, and note that $\tilde{g}_{\mathcal{S},T}$ can be written as as the quadratic form

$$\begin{aligned}
\tilde{g}_{\mathcal{S},T}(x,y;\omega) &= \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{\tau\neq t}\lambda(\frac{t-\tau}{M})Z_{b,t}(x)Z_{b,\tau}(y)\exp(i(t-\tau)\omega) \\
&- \frac{1}{2\pi}\mathbb{E}\big(W_b(x-X_0)\big)\mathbb{E}\big(W_b(y-X_0)\big).
\end{aligned} \tag{3.16}$$

Lemma III.12 *(i) Suppose Assumption III.1(i-iii,v) holds with $\alpha \geq 6$ and $b^{-1} << M$. Then we have*

$$\mathbb{E}\bigg(\sqrt{\frac{Tb^2}{M}}\Re B^{(u)}_{T,S_T}\bigg)^2 = O(\frac{S_T}{T}) \quad \mathbb{E}\bigg(\sqrt{\frac{Tb^2}{M}}\Im B^{(u)}_{T,S_T}\bigg)^2 = O(\frac{S_T}{T})$$

(ii) Suppose Assumption III.1 holds with $\alpha \geq 14$ and $b^{-1} << M$. Then we have

$$\mathbb{E}\bigg(\sqrt{\frac{Tb^2}{M}}\Re B^{(u)}_{T,S_T}\bigg)^4 = O(\frac{S_T}{T})^2 \qquad \mathbb{E}\bigg(\sqrt{\frac{Tb^2}{M}}\Im B^{(u)}_{T,S_T}\bigg)^4 = O(\frac{S_T}{T})^2.$$

PROOF. The proof of (i) is a straightforward application of Lemma IV.1 for the case $k=1$ and $k=3$, we omit the details.

We now prove (ii). Let $\kappa_4(X)$ denote the fourth order cumulant of the random variable X, then we recall that

$$\mathbb{E}\bigg(\sqrt{\frac{Tb^2}{M}}B^{(u)}_{T,S_T}\bigg)^4 = 3\mathrm{var}\bigg(\sqrt{\frac{Tb^2}{M}}B^{(u)}_{T,S_T}\bigg)^2 + \kappa_4\bigg(\sqrt{\frac{Tb^2}{M}}B^{(u)}_{T,S_T}\bigg).$$

It follows from (i) that $\mathrm{var}\big(\sqrt{\frac{Tb^2}{M}}B^{(u)}_{T,S_T}\big)^2 = O\big(\frac{S_T}{T}\big)^2$. We now obtain a bound for $\kappa_4\bigg(\sqrt{\frac{Tb^2}{M}}B^{(u)}_{T,S_T}\bigg)$. Expanding the cumulant we have

$$\begin{aligned}
\big|\kappa_4\big(B^{(u)}_{T,S_T}\big)\big| &= \frac{1}{T^4}\sum_{t_1,t_2,t_2,t_4=u+1}^{S_T+u}\sum_{\tau_1,\tau_2,\tau_3,\tau_4}\big|\lambda(\frac{t_1-\tau_1}{M})\lambda(\frac{t_2-\tau_2}{M})\lambda(\frac{t_3-\tau_3}{M})\lambda(\frac{t_4-\tau_4}{M})\big| \\
&\times \big|\mathrm{cum}(Y_{b,t_1}Z_{b,\tau_1},Y_{b,t_2}Z_{b,\tau_2},Y_{b,t_3}Z_{b,\tau_3},Y_{b,t_4}Z_{b,\tau_4})\big|.
\end{aligned}$$

We decompose $\left|\text{cum}(Y_{b,t_1}Z_{b,\tau_1}, Y_{b,t_2}Z_{b,\tau_2}, Y_{b,t_3}Z_{b,\tau_3}, Y_{b,t_4}Z_{b,\tau_4})\right|$ into the sum of indecomposable partitions (see, for example, Brillinger (1981)). Therefore, by using Lemma IV.1 and $b^{-1} << M$ we have that $\left|\kappa_4\big(B^{(u)}_{T,S_T}\big)\right| = o\big(\frac{S_T}{T}\big)^2$. This together with (i) gives (ii). $\square$

The above is moments bound on the partial sums which will be useful for applying the central limit theorem for quadratic forms of mixing random variables.

PROOF of Theorem III.1. To prove the result we note that $\tilde{g}_{\mathcal{S},T}(x,y,\omega)$ can be written as a quadratic form (see (3.16)). Using identical arguments to those in Theorem IV.2, with Lemma III.12 on moments of partial sums replacing Lemma IV.4, we can prove asymptotic normality of $\tilde{g}_{\mathcal{S},T}(x,y,\omega)$. $\square$

CHAPTER IV

A NOTE ON GENERAL QUADRATIC FORMS OF NONSTATIONARY TIME SERIES

1. Introduction

The study of the asymptotic theory of statistics often involves quadratic forms which have the general form

$$W_T = \frac{1}{T} \sum_{t,\tau=1}^{T} G_{t,\tau} h(X_t, X_\tau), \tag{4.1}$$

where $\{X_t\}$ is a stochastic process, $h(\cdot)$ is a function and $\{G_{t,\tau}\}$ are weights, which vary according to the application. Various statistical methods depend on the asymptotic sampling distribution of above statistic.

In view of its importance, several authors have studied W_T for the particular case $h(X_t, X_\tau) = X_t X_\tau$ under various assumptions on the stochastic process $\{X_t\}$. For example, Mikosch (1990), Götze and Tikhomirov (1999) and the references therein, analysis W_T under the assumption that $\{X_t\}$ are iid random variables. Kokoszka and Taqqu (1997) and Bhansali, Giraitis, and Kokoszka (2007) relax the independence assumption and establish asymptotic normality of W_T under the assumption that $\{X_t\}$ is a realisation from stationary, linear time series. Rosenblatt (1984) allows for nonlinear time series, by assuming that $\{X_t\}$ are α-mixing. In particular, he shows asymptotic normality of W_T under the assumption that $\{X_t\}$ is a strictly stationary α-mixing time series and has absolutely summability eight order cumulants. The generalising to mixing random variables, allow $\{X_t\}$ to be a non-linear time series, but the cumulant assumptions are quite strong. Recently, Gao and Anh (2000) relax the moment assumptions by considering geometric mixing $\{X_t\}$ and Lin (2009) considers the case $\{X_t\}$ is the sum of stationary α-mixing random variables. It should be mentioned, that there are other methods for measuring dependence. For example, Wu and Shao (2007) show asymptotic normality when $\{X_t\}$ can be written as a function of the innovations and satisfies the assumption of physical dependence. The study of the general quadratic form given in (4.1) can also arise in several applications, including nonparametric estimators, but has received less attention. One reason for this is that techniques used in the articles mentioned above cannot be directly applied to (4.1). Moreover, the underlying assumption in all the above references is that the process $\{X_t\}$ is strictly stationary.

In the analysis of nonstationary time series (which is possibly nonlinear), quadratic forms of the above type do occur, for example estimators of the time-varying spectral density involve quadratic forms (see, for example, Dahlhaus (2000) and Dwivedi and Subba Rao (2011)). In this paper, our objective is to study the asymptotic theory of general quadratic forms for nonstationary processes.

In Section 2 we show asymptotic normality of the general quadratic form under some moment assumptions and α-mixing of the stochastic process (which includes both nonstationary and nonlinear processes). By using Ibragimov-type inequalities (see Statulevicius and Jakimavicius (1988)) which link cumulants to the mixing rate, we avoid direct assumptions on the summability of the cumulants. The assumptions allow the weights $G_{t,\tau}$ to also depend on T, thus including the case of spectral density estimators. In Section 3 we derive some results on cumulants and moments of the quadratic form. We use mixingale and near-epoch dependent methods to prove the results in this section, these techniques may also be of independent interest. All the proofs can be found in Section 4.

2. The quadratic form

Let us suppose that $\{X_{t,T}; 1 \leq t \leq T\}$ is a time series which we do not assume to be stationary. By allowing $X_{t,T}$ to depend on T, the results below cover the case of triangular arrays, and in particular allows for locally stationary time series. We will assume that for all t, $\mathbb{E}(X_{t,T}) = 0$ and for all t, T $0 < \text{var}(X_{t,T}) < \infty$. This condition excludes degenerate cases by ensuring that $\{X_{t,T}\}$ does not converge to a non-random sequence but always has a bounded variance. In this paper we consider general quadratic forms of the type

$$Q_T = \frac{1}{T}\sum_{t,\tau=1}^{T} G_{t,\tau} h(X_{t,T}, X_{\tau,T}), \tag{4.2}$$

where we do not impose any conditions on the function $h : \mathbb{R}^2 \to \mathbb{R}$. By allowing this amount of generality on $h(\cdot)$, we need to assume that the weights $G_{t,\tau}$ decay to zero, in the sense that $\sup_{t,T}\sum_{\tau}|G_{t,\tau}| < \infty$. For example, if $h(X_{t,T}, X_{\tau,T}) = (X_{t,T} + X_{\tau,T})$, then for the variance of Q_T to decay to zero as $T \to \infty$, we require such a condition on the weights. In order to relax this condition on the weights $\{G_{t,\tau}\}$, a stronger condition on $h(\cdot)$ is required. Therefore, in addition to the above, we will also consider quadratic forms which have the multiplicative form $h(X_{t,T}, X_{\tau,T}) = X_{t,T}X_{\tau,T}$:

$$Q_{T,M} = \frac{1}{T}\sum_{t,\tau=1}^{T} G_{t,\tau,M} X_{t,T} X_{\tau,T} \tag{4.3}$$

where for some $0 < \alpha < 1$, $M := M(T) = T^{\alpha}$, and for $|t - \tau| > M$, then $G_{t,\tau,M} = 0$.

We now state some conditions, which we use to prove asymptotic normality of Q_T and $Q_{T,M}$.

Assumption IV.1 *(i) Let us suppose that $\{X_{t,T}\}$ is an α-mixing time series such* *that*

$$\sup_k \sup_{\substack{A\in\sigma(X_{t+k,T},X_{t+1+k,T},\ldots)\\ B\in\sigma(X_{k,T},X_{k-1,T},\ldots)}} |P(A\cap B)-P(A)P(B)| \leq \alpha(t),$$

where $\alpha(t)$ *are the mixing coefficients which satisfy* $\alpha(t) \leq K|t|^{-s}$ *for some* $s > 0$.

(ii) *(a)* *For* Q_T *defined in (4.2), we suppose* $|G_{t,\tau}| \leq C|t-\tau|^{-\delta}$ $(\delta > 2)$ *and* $c_1\frac{G}{T} \leq var(Q_T) \leq c_2\frac{G}{T}$ *(for some* $0 < c_1 \leq c_2 < \infty$*), were* $G = \sup_t \sum_\tau |G_{t,\tau}| < \infty$.

(b) *For* $Q_{T,M}$ *defined in (4.3), we suppose that* $G_{t,\tau,M} = 0$ *for* $|t-\tau| > M$ *and for all* T, $c_1\frac{G_M}{T} \leq var(Q_{T,M}) \leq c_2\frac{G_M}{T}$ *(for some* $0 < c_1 \leq c_2 < \infty$*), where* $G_M = \sup_t \sum_\tau |G_{t,\tau,M}|^2$ *and* $\inf_M G_M > 0$.

(iii) *(a)* *For some* $r > 2s/(s-2) > 0$*, we have* $\sup_{t,\tau,T} \mathbb{E}|h(X_{t,T}, X_{\tau,T})|^r < \infty$.

(b) *For some* $r > 4s/(s-6) > 0$*, we have* $\sup_{t,T} \mathbb{E}|X_{t,T}|^r < \infty$.

Before stating the asymptotic sampling properties of the the quadratic forms, some comments on the assumptions are in order. To prove asymptotic normality of Q_T and $Q_{T,M}$ we have to treat the cases differently and use a slightly different set of conditions. This is primarily because we need to obtain moment bounds for each of these terms (see Lemmas IV.2 and IV.3 in Section 3). The details can be found in Section 4 but to give a flavour of the methods, to bound Q_T we treat $\{\sum_\tau G_{t,\tau}h(X_{t,T}, X_{\tau,T})\}_t$ as a stochastic process with decaying dependence structure and use the notion of L_2-NED together with martingale methods to obtain the moment bounds. However, in the case of $Q_{T,M}$, despite Assumption IV.1 (i, ii (b)) (in particular the mixing and $G_{t,\tau,M} = 0$ for $|t-\tau| > M$) implying that the dependence in the sequence $\{\sum_\tau G_{t,\tau,M}X_{t,T}X_{\tau,T}\}$ decays the further apart the ts, the same methods used to bound Q_T, when applied to $Q_{T,M}$ gives sub-optimal bounds. Instead we use iterative martingale methods to obtain the optimal moment bounds for $Q_{T,M}$. We observe that in the case that $|G_{t,\tau}| \leq C|t-\tau|^{-\delta}$ ($\delta > 2$) and $g(X_{t,T}, X_{\tau,T}) = X_{t,T}X_{\tau,T}$, then Assumption IV.1 (iii (a)) is slightly weaker than Assumption IV.1 (iii (b)). As the assumptions on $Q_{T,M}$ allow $\sum_{\tau}|G_{t,\tau,M}| \to \infty$ as $T \to \infty$, we require that the fourth order cumulants are absolutely summable, see Remark IV.1 below. Finally, several time series, both stationary and nonstationary, satisfy the α-mixing conditions given in Assumption IV.1(i), see, for example, Tjøstheim (1990), Doukhan (1994), Cline and Pu (1999), Bradley (2007) and Fryzlewicz and Subba Rao (2011).

Remark IV.1 *(i) The variance of $Q_{T,M}$ is*

$$var(Q_{T,M}) = \frac{1}{T^2}\sum_{t_1,\tau_1=1}^{T}\sum_{t_2,\tau_2=1}^{T} G_{t_1,\tau_1,M}G_{t_2,\tau_2,M}\Big[cov(X_{t_1,T}, X_{t_2,T})cov(X_{\tau_1,T}, X_{\tau_2,T}) + cov(X_{t_1,T}, X_{\tau_2,T})cov(X_{\tau_1,T}, X_{t_2,T}) + cum(X_{t_1,T}, X_{\tau_1,T}, X_{t_2,T}, X_{\tau_2,T})\Big]. \quad (4.4)$$

(ii) If $G_M = O(T^{\alpha})$ (where $0 < \alpha < 1$), it can be shown that under Assumption IV.1(iiib) the fourth order cumulant term in (4.4) is asymptotically negligible with respect to the covariances terms.

We now derive the limiting distribution of Q_T and $Q_{T,M}$.

Theorem IV.1 *Suppose Assumption IV.1(i, ii (a), iii (a)) is satisfied. Let $var(Q_T) = V_T$, then we have $V_T^{-1/2}\big(Q_T - \mathbb{E}(Q_T)\big) \xrightarrow{D} \mathcal{N}(0,1)$ as $T \to \infty$.*

Theorem IV.2 *Suppose Assumption IV.1(i, ii (b), iii (b)) is satisfied. Let $var(Q_{T,M}) = V_T$, then we have $V_T^{-1/2}\big(Q_{T,M} - \mathbb{E}(Q_{T,M})\big) \xrightarrow{D} \mathcal{N}(0,1)$ as $T \to \infty$.*

The above results are for quadratic forms of univariate time series. As multivariate time series arise in several applications we now give an analogous result for multivariate time series, noting that the proof is almost identical to the univariate case.

Corollary IV.1 *Let us suppose that* $\{\underline{X}_{t,T}\}$ *is a d-dimensional vector time series, which is mixing*

$$\sup_k \sup_{\substack{A\in\sigma(\underline{X}_{t+k,T},\underline{X}_{t+1+k,T},\ldots)\\ B\in\sigma(\underline{X}_{k,T},\underline{X}_{k-1,T},\ldots)}} |P(A\cap B)-P(A)P(B)| \leq \alpha(t),$$

where $\alpha(t)$ *are the mixing coefficients and are such that* $\alpha(t) \leq K|t|^{-s}$ *where* $s>0$*, and suppose there exists some* $r > \frac{4s}{s-6}$*, such that* $\sup_{t,T}\mathbb{E}(\sum_{j=1}^{d}|X_{t,T,j}|)^r < \infty$ *(where* $|\cdot|$ *denotes the Euclidean norm of a vector or matrix). Define the quadratic form*

$$Q_T = \frac{1}{T}\sum_{t,\tau=1}^{T} \underline{X}'_{t,T}\boldsymbol{G}_{t,\tau,M}\underline{X}_{\tau,T},$$

where $\{\boldsymbol{G}_{t,\tau,M}\}$ *is a* $d\times d$ *matrix which satisfies* $\boldsymbol{G}_{t,\tau,M}=0$ *(for* $|t-\tau|>M$*). We assume there exists* $0<c_1\leq c_2<\infty$ *such that* $c_1G_M/T \leq var(Q_T) \leq c_2G_M/T$ *(*$G_M=\sup_t\sum_\tau|G_{t,\tau,M}|$*). Then we have* $V_T^{-1/2}(Q_T-\mathbb{E}(Q_T)) \xrightarrow{D} \mathcal{N}(0,1)$*, where* $V_T = var(Q_T)$*.*

3. Some bounds on cumulants and moments

In this section we state some bounds on the sums of moments and cumulants. These results will be used to prove Theorems IV.1 and IV.2. We mention that the techniques used in the proof of the results may also be of independent interest.

The following two lemmas concern summability of the higher order cumulants of a stochastic process. We first state a bound for the sum of cumulants based on the mixing rate. This result is motivated by Neumann (1996), Remark 3.1. Let $\|X\|_p = (\mathbb{E}(|X|^p)^{1/p}$ and K denote a finite generic constant.

Lemma IV.1 *Let us suppose that $\{X_{t,T}\}$ is a α-mixing time series with rate $\{\alpha(t)\}$. If $t_1 \leq t_2 \leq \ldots \leq t_k$, then we have*

(i)

$$|cum(X_{t_1,T}, \ldots, X_{t_k,T})| \leq C_k \sup_{t,T} \||X_{t,T}\|_r^k \prod_{i=2}^{k} \alpha(t_i - t_{i-1})^{\frac{1-k/r}{k-1}}, \tag{4.5}$$

(ii)

$$\begin{aligned} & \sup_{t_1} \sum_{t_2,\ldots,t_k=1}^{\infty} |cum(X_{t_1,T}, \ldots, X_{t_k,T})| \\ \leq \;& C_k \sup_{t,T} \|X_{t,T}\|_r^k \Big(\sum_t \alpha(t)^{\frac{1-k/r}{k-1}}\Big)^{k-1} < \infty, \end{aligned} \tag{4.6}$$

(iii) For all $2 \leq j \leq k$, we have

$$\begin{aligned} & \sup_{t_1} \sum_{t_2,\ldots,t_k=1}^{\infty} (1+|t_j|)|cum(X_{t_1,T}, \ldots, X_{t_k,T})| \\ \leq \;& C_k \sup_{t,T} \|X_{t,T}\|_r^k \Big(\sum_t \alpha(t)^{\frac{1-k/r}{k-1}}\Big)^{k-1} < \infty, \end{aligned} \tag{4.7}$$

where C_k is a finite constant which depends only on k.

Using the lemma above, the following corollary on the absolute summability of the fourth order cumulants immediately follows.

Corollary IV.2 *Suppose that $\{X_{t,T}\}$ is a α-mixing time series which satisfies Assumption IV.1(i), where $\alpha(t) \leq K \cdot |t|^{-s}$.*

(i) Let us suppose that $r > 4s/(s-3)$ and $\sup_{t,T} \mathbb{E}|X_{t,T}|^r < \infty$, then we have $|cov(X_{t,T}, X_{\tau,T})| \leq C|t-\tau|^{-\frac{(s+3)}{2}}$ and
$\sup_{t_1} \sum_{t_2,t_3,t_4=-\infty}^{\infty} |cum(X_{t_1,T}, X_{t_2,T}, X_{t_3,T}, X_{t_4,T})| < \infty$.

(ii) Let us suppose that $r > 4s/(s-6)$ and $\sup_{t,T} \mathbb{E}|X_{t,T}|^r < \infty$, then we have $|cov(X_{t,T}, X_{\tau,T})| \leq C|t-\tau|^{-\frac{(s+6)}{2}}$ and for all $2 \leq j \leq 4$,

$\sup_{t_1} \sum_{t_2,t_3,t_4=-\infty}^{\infty}(1+|t_j|)|cum(X_{t_1,T}, X_{t_2,T}, X_{t_3,T}, X_{t_4,T})| < \infty.$

It is worth mentioning that Assumption IV.1 (ii (a)) is weaker than those in Corollary IV.2, this is because we do not require absolute summability of the fourth order cumulants in order for $\mathrm{var}(Q_T) = O(T^{-1})$.

In order to use a blocking argument to prove Theorems IV.1 and IV.2, we need to partition the data such that Q_T can be written as a sum of random variables which are non-intersecting. This is immediately possible with $Q_{T,M}$ but not Q_T, Thus we now define a close approximation of Q_T which satisfies this condition. Let

$$\tilde{Q}_{T,M} = \frac{1}{T}\sum_{t,\tau=1}^{T} I(\frac{t-\tau}{M})G_{t,\tau}h(X_{t,T}, X_{\tau,T}) \tag{4.8}$$

where $M = T^{1/2+\gamma}$ for some $0 < \gamma < 1/2$ and $I(x) = 1$ for $x \in [-1, 1]$ and zero elsewhere. Since $|G_{t,\tau}| \leq K|t-\tau|^{-\delta}$ $(\delta > 2)$ we have

$$Q_T = \tilde{Q}_{T,M} + O_p(T^{-1/2-\gamma}), \tag{4.9}$$

and $\mathrm{var}(\sqrt{T}Q_T) = \mathrm{var}(\sqrt{T}\tilde{Q}_{T,M}) + O(T^{-\gamma})$. We will show that $\mathrm{var}(\sqrt{T}Q_T) = O(1)$, thus Q_T and the truncated $\tilde{Q}_{T,M}$ are asymptotically equivalent. The results concerning $\tilde{Q}_{T,M}$ and $Q_{T,M}$ are largely the same, the only difference are the proofs, thus to unify notation, we let $Q_{T,M} := \tilde{Q}_{T,M}$ and $G_{t,\tau,M} = I(\frac{t-\tau}{M})G_{t,\tau}$, and state under what conditions we obtain the each result.

We now define sub-blocks of $Q_{T,M}$, which will be used to prove Theorem IV.2. Let

$$Y_{t,T} = \sum_{\tau<t} G_{t,\tau,M}h(X_{t,T}, X_{\tau,T}) + \sum_{\tau\leq t} G_{\tau,t,M}h(X_{\tau,T}, X_{t,T}). \tag{4.10}$$

In the case that $h(X_{t,T}, X_{\tau,T}) = X_{t,T}X_{\tau,T}$ the above is

$$Y_{t,M} = \sum_{\tau=1}^{t} F_{t,\tau,M} X_{t,T} X_{\tau,T} \quad \text{where} \quad F_{t,\tau,M} = \begin{cases} G_{t,t,M} & t = \tau \\ (G_{t,\tau,M} + G_{\tau,t,M}) & t \neq \tau \end{cases}.$$

To use the Bernstein blocking argument we define a sub-block of S_T. Let

$$B_{T,S_T}^{(u)} = \frac{1}{T} \sum_{t=u+1}^{S_T+u} Y_{t,T}, \tag{4.11}$$

noting that $B_{T,T}^{(0)} = Q_{T,M}$. Lemma IV.1 can be used to obtain bounds for $\mathrm{var}(B_{T,S_T}^{(u)})$ and other integer moments of $B_{T,S_T}^{(u)}$. However, in order to prove asymptotic normality under relatively weak assumptions we will require bounds on non-integer moments of $B_{T,S_T}^{(u)}$, which use more subtle arguments. The actual proof used to obtain the bounds differs, depending on whether we use Assumption IV.1 (ii (a), iii (a)) or IV.1 (ii (b), iii (b)). Thus we state the results separately.

Lemma IV.2 *Suppose Assumption IV.1 (i, ii(a), iii(a)) holds and let* $\mathcal{F}_t = \sigma(X_{t,T}, X_{t-1,T}, \ldots)$. *If* $\sup_{t,\tau,T} \|h(X_{t,T}, X_{\tau,T})\|_r < \infty$ *for some* $r > q$ *, then*

$$\begin{aligned} &\left\| Y_{t,T} - \mathbb{E}(Y_{t,T}|\mathcal{F}_{t-j}) \right\|_q \\ &\leq K\left(j^{-(\delta-1)} \sup_{\tau} \|h(X_{t,T}, X_{\tau,T})\|_q + \sup_{\tau} \|h(X_{t,T}, X_{\tau,T})\|_r j^{-s(\frac{1}{q}-\frac{1}{r})} \right), \end{aligned} \tag{4.12}$$

and almost surely $Y_{t,T} = \sum_j N_{j,T}(t-j)$ *where*

$$N_{j,T}(t-j) = \mathbb{E}(Y_{t,T}|\mathcal{F}_{t-j}) - \mathbb{E}(Y_{t,T}|\mathcal{F}_{t-j-1}).$$

Let $q \geq 2$ *and* $B_{T,S_T}^{(u)}$ *be defined as in (4.11). Suppose the above conditions are satisfied, then we have*

$$\left\| B_{T,S_T}^{(u)} \right\|_q \leq KT^{-1} S_T^{1/2} \sum_{j=1}^{\infty} \left(\frac{1}{j^{\delta-1}} + \frac{1}{j^{s(\frac{1}{q}-\frac{1}{r})}} \right). \tag{4.13}$$

Lemma IV.3 *Suppose Assumption IV.1 (i, ii(b), iii (b)) hold and let* $\mathcal{F}_{t,T} = \sigma(X_{t,T}, X_{t-1,T}, \ldots)$ *and denote* $\mathbb{E}(Z|\mathcal{F}_{j,T}) = \mathbb{E}_j(Z)$*. If for some* $r > q$ *we have* $\sup_{t,T} \|X_{t,T}\|_r < \infty$*, then we obtain the bound*

$$\|\mathbb{E}_{t-j}(X_{t,T}) - \mathbb{E}_{t-j-1}(X_{t,T})\|_q \leq 4(2^{1/q}+1)\alpha(j)^{\frac{1}{q}-\frac{1}{r}}\|X_{t,T}\|_r, \tag{4.14}$$

and $X_{t,T}$ *almost surely admits the representation*

$$X_{t,T} = \sum_{j=0}^{\infty} \big(\mathbb{E}_{t-j}(X_{t,T}) - \mathbb{E}_{t-j-1}(X_{t,T})\big).$$

Let $M_j(t-j) = \mathbb{E}_{t-j}(X_{t,T}) - \mathbb{E}_{t-j-1}(X_{t,T})$*. If for some* $\tilde{r}/2 > r > q$ *we have* $\sup_{t,T} \|X_{t,T}\|_{\tilde{r}} < \infty$*, then*

$$\begin{aligned} & \big\|\mathbb{E}_{t-j_1-i}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)) - \mathbb{E}_{t-j_1-i-1}(M_{j_1}(t-j_1)M_{j_2}(t-j_1))\big\|_q \\ \leq \;& K\|X_{t,T}\|_{\tilde{r}}^2 \alpha(j_1)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\alpha(j_2)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\alpha(i)^{\frac{1}{q}-\frac{1}{r}}. \end{aligned} \tag{4.15}$$

Let $q \geq 2$ *and* $B^{(u)}_{T,S_T}$ *be defined as in (4.11). If there exists, an* $\tilde{r}$*, such that* $\sup_{t,T} \|X_{t,T}\|_{\tilde{r}} < \infty$*, where* $\tilde{r}/2 > r > q$*, then we have*

$$\big\|B^{(u)}_{T,S_T}\big\|_q \leq KT^{-1}S_T^{1/2}\Big[G_M^{1/2}\big(\sum_{j=1}^{\infty}\frac{1}{j^{s(\frac{1}{2q}-\frac{1}{r})}}\big)^2 + \big(\sum_{j=1}^{\infty}\frac{1}{j^{s(\frac{1}{q}-\frac{1}{r})}}\big)\big(\sum_{j=1}^{\infty}\frac{1}{j^{s(\frac{1}{2r}-\frac{1}{\tilde{r}})}}\big)^2\Big] \tag{4.16}$$

A simple application of the lemmas above is to derive bounds for the moments of the quadratic form $Q_{T,M}$ (since $Q_{T,M}$ is a special case of $B^{(u)}_{T,S_T}$, with $u = 0$ and $S_T = T$). By using the arguments in Lemma IV.4, below, it can be shown that for some $\epsilon > 0$, we have $\|Q_{T,M}\|_{2+\epsilon} \leq K/T^{1/2}$ (under Assumption IV.1 (i, ii (a), iii (a)) and $\|Q_{T,M}\|_{2+\epsilon} \leq KG_M^{1/2}/T^{1/2}$ (under Assumption IV.1 (i, ii (b), iii (b)).

4. Proofs

4.1. Proofs of results in Section 2

To do the analysis, we start by rewriting $Q_{T,M} - \mathbb{E}(Q_{T,M})$ as

$$Q_{T,M} - \mathbb{E}(Q_{T,M}) \;=\; \frac{1}{T}\sum_{t,\tau=1}^{T} G_{t,\tau,M}\big(h(X_{t,T}, X_{\tau,T}) - \mathbb{E}(h(X_{t,T}, X_{\tau,T})\big) = \sum_{t=1}^{T} Y_{t,T},$$

where $Y_{t,T}$ is defined in (4.10). To prove asymptotic normality we use a classical Bernstein blocking argument. Here we partition $\{Y_{t,T}; t = 1, \ldots, T\}$ into the sum of small and large blocks. Let $U_{i,T}$ and $V_{i,T}$ denote the big blocks and small blocks respectively, where

$$U_{i,T} = \sum_{t=ir_T+1}^{ir_T+p_T} Y_{t,T}, \qquad V_{i,T} = \sum_{t=ir_T+p_T+1}^{(i+1)r_T} Y_{t,T},$$

$p_T >> q_T >> M$ and $r_T = (p_T + q_T)$. Let $k_T = T/(p_T + q_T)$ and $q_T/(p_T + q_T) \to 0$ as $T \to \infty$. For the purpose of proving the results below we wil assume that $k_T = O((\log T)^{1/2})$. Using the above notation we let $Q_{T,M} - \mathbb{E}(Q_{T,M}) = \mathcal{S}_{k_T} + \mathcal{R}_{k_T}$, where

$$\mathcal{S}_{k_T} = \sum_{i=1}^{k_T} U_{i,T} \quad \text{and} \quad \mathcal{R}_{k_T} = \sum_{i=1}^{k_T} V_{i,T}.$$

Since $p_T >> q_T$, we will show that $\mathrm{var}\big(\sqrt{\frac{T}{G_M}}\mathcal{R}_{k_T}\big) \to 0$. We first obtain moment bounds for $\{U_{i,T}\}$ and $\{V_{i,T}\}$. We note that under Assumption IV.1 (i, ii (a), iii (a)), that $G_M := G \leq K\sum_{j=1}^{\infty} j^{-\delta} < \infty$.

Lemma IV.4 *Let us suppose Assumptions IV.1 holds. Then for some $\delta > 0$ we have*

$$\big\|U_{i,T}\big\|_{2+\delta} = O\big(\frac{p_T^{1/2}G_M^{1/2}}{T}\big) \qquad \big\|V_{i,T}\big\|_{2+\delta} = O\big(\frac{q_T^{1/2}G_M^{1/2}}{T}\big). \tag{4.17}$$

PROOF. We use Lemmas IV.2 and Lemma IV.3 to prove the result, with $p_T = S_T$ and $u = ir_T$. We first prove the result under Assumption IV.1 (i,ii (a),iii (a)). By applying Lemma IV.2 for $q = 2+\delta$ and $r > 2+\delta$ we have

$$\left\|B_{T,S_T}^{(u)}\right\|_{2+\delta} \leq K \sup_{t,\tau,T} \|g(X_{t,T}, X_{\tau,T})\|_r T^{-1} S_T^{1/2} \sum_{j=1}^{\infty} \left(\frac{1}{j^{\delta-1}} + \frac{1}{j^{s(\frac{1}{2+\delta} - \frac{1}{r})}} \right).$$

Thus the above bound is finite for $r > s(2+\delta)/(s-2-\delta)$. In other words, if $r > 2s/(s-2)$, there exists a δ, such that $\left\|B_{T,S_T}^{(u)}\right\|_{2+\delta} = O(\frac{q_T^{1/2} G_M^{1/2}}{T})$. To apply Lemma IV.3 for $q = 2+\delta$, then for some $\tilde{r}/2 > r > 2+\delta$ we have

$$\begin{aligned} &\left\|U_{i,T}\right\|_{2+\delta} \\ \leq\ & KT^{-1} p_T^{1/2} \|X_{t,T}\|_{\tilde{r}} \left(G_M^{1/2} \Big(\sum_{j=1}^{\infty} \frac{1}{j^{s(\frac{1}{2(2+\delta)} - \frac{1}{\tilde{r}})}} \Big)^2 + \Big(\sum_{j=1}^{\infty} \frac{1}{j^{s(\frac{1}{(2+\delta)} - \frac{1}{r})}} \Big) \Big(\sum_{j=1}^{\infty} \frac{1}{j^{s(\frac{1}{2r} - \frac{1}{\tilde{r}})}} \Big)^2 \right). \end{aligned}$$

In order to ensure that the right hand side of the above is finite, $\tilde{r}$ should satisfy the conditions

$$\frac{1}{2(2+\delta)} - \frac{1}{\tilde{r}} > \frac{1}{s}, \quad \frac{1}{2+\delta} - \frac{1}{r} > \frac{1}{s} \quad \text{and} \quad \frac{1}{2r} - \frac{1}{\tilde{r}} > \frac{1}{s},$$

which implies

$$\tilde{r} > \frac{2(2+\delta)s}{(s-2(2+\delta)} \quad \text{and} \quad \tilde{r} > \frac{2s(2+\delta))}{(s-3(2+\delta)}. \tag{4.18}$$

Thus by Assumption IV.1 (iii (b)) (we recall there exists an r such that $r > 4s/(s-6)$ and $\sup_{t,T} \|X_{t,T}\|_r < \infty$), there exists a $\tilde{r}$ and $\delta > 0$, such that (4.18) is satisfied. Thus for both cases, (4.17) holds for some $\delta > 0$. The proof of $\left\|V_{i,T}\right\|_{2+\delta} = O\big(\frac{q_T^{1/2} G_M^{1/2}}{T}\big)$ is the same, hence we omit the details. □

We now show that the contribution of the sum of small blocks, $\mathcal{R}_{k_T}$, is negligible with respect to the entire sum $Q_{T,M} - \mathbb{E}(Q_{T,M})$.

Lemma IV.5 *Suppose Assumption IV.1 holds and $q_T/(p_T+q_T) \to 0$ as $T \to \infty$.*

Then we have

$$\left|cov\big(V_{i_1,T},V_{i_2,T}\big)\right| \leq C\alpha\bigg(|i_1-i_2|p_T-M\bigg)^{1-\frac{2}{2+\delta}}\big(\frac{G_M q_T}{T^2}\big) \tag{4.19}$$

and

$$var\bigg(\sqrt{\frac{T}{G_M}}\mathcal{R}_{k_T}\bigg) \leq C\frac{q_T}{(p_T+q_T)} \to 0, \tag{4.20}$$

as $T\to\infty$, where C is a finite constant.

PROOF. Define the sigma-algebras $\mathcal{G}_{i_2}^{\infty}=\sigma(Y_{i_2r_T+p_T+1,T},Y_{i_2r_T+p_T+2,T},\ldots)$ and $\mathcal{G}_{-\infty}^{i_1}=\sigma(Y_{(i_1+1)r_T,T},Y_{(i_1+1)r_T-1,T},\ldots)$. To prove (4.19) for $i_2>i_1$ we use Ibragimov's inequality to obtain

$$\begin{aligned}
\left|\mathrm{cov}\big(V_{i_1,T},V_{i_2,T}\big)\right| &\leq C\Big\{\sup_{A\in\mathcal{G}_{i_2}^{\infty},B\in\mathcal{G}_{-\infty}^{i_1}}\big|P(A\cap B)-P(A)P(B)\big|\Big\}^{1-\frac{2}{2+\delta}}\big\|V_{i_1,T}\big\|_{2+\delta}^2\\
&\leq C\Big\{\alpha\big((i_2-i_1-1)r_T+p_T+1-M\big)\Big\}^{1-\frac{2}{2+\delta}}\big\|V_{i_1,T}\big\|_{2+\delta}^2\\
&\leq C\alpha\big((i_2-i_1)p_T-M\big)^{1-2/(2+\delta)}\big\|V_{i_1,T}\big\|_{2+\delta}^2. \qquad (4.21)
\end{aligned}$$

This gives (4.19).

To prove (4.20) we substitute (4.19) into $\mathrm{var}(\mathcal{R}_{k_T})=\sum_{i_1,i_2=1}^{k_T}\mathrm{cov}\big(V_{i_1,T},V_{i_2,T}\big)$ and use that $\|V_{i,T}\|_{2+\delta}=O(q_T^{1/2}G_M^{1/2}/T)$ to get

$$\mathrm{var}\big(\sqrt{\frac{T}{G_M}}\mathcal{R}_{k_T}\big) \leq C\frac{q_T}{T}\bigg(\sum_{i=1}^{k_T}1+2\sum_{i_1<i_2}^{k_T}\alpha\big(|i_1-i_2|p_T-M\big)^{1-2/(2+\delta)}\bigg).$$

Now by using that the mixing rate $\alpha(t)\leq Kt^{-s}$ and $k_T=T/(p_T+q_T)$ we have

$$\begin{aligned}
\mathrm{var}\big(\sqrt{\frac{T}{G_M}}\mathcal{R}_{k_T}\big) &\leq K\frac{q_T}{p_T+q_T}\bigg(1+\sum_{r=1}^{k_T}(rp_T-M)^{-s(1-\frac{2}{(2+\delta)})}\bigg)\\
&\leq K\frac{q_T}{p_T+q_T}\bigg(1+(p_T-M)^{-s(1-\frac{2}{2+\delta})}k_T\bigg).
\end{aligned}$$

Since $k_T=(\log T)^{1/2}$, we have $((p_T-M))^{-s(1-\frac{2}{2+\delta})}k_T<\infty$, which gives (4.20). □

Using that $Q_{T,M} = \mathcal{S}_T + \mathcal{R}_T$ and $\mathrm{var}(Q_{T,M}) := V_T = O(\frac{G_M}{T})$, the above result implies that $\mathrm{var}Q_{T,M}{}^{-1/2}\mathcal{R}_T = o(1)$ and

$$V_T^{-1/2}(Q_{T,M} - \mathbb{E}(Q_{T,M})) = V_T^{-1/2}\mathcal{S}_{k_T} + o_p(1). \tag{4.22}$$

We now show normality of $\mathcal{S}_{k_T}$. We do this by replacing $\mathcal{S}_{k_T}$ with $\tilde{\mathcal{S}}_{k_T} = \sum_i \tilde{U}_{i,T}$, where $\tilde{U}_{i,T}$ and $U_{i,T}$ have identical distributions, but $\{\tilde{U}_{i,T}\}$ are independent random variables. Below we show that the distributions of $\mathcal{S}_{k_T}$ and $\tilde{\mathcal{S}}_{k_T}$ are asymptotically equivalent.

We require the following general theorem, which gives a bound on the differences of characteristic functions of sums mixing and independent random variables. A potentially useful aspect of this result, is that we allow for the mixing rate to change with T.

Theorem IV.3 *Suppose $\{Z_{t,T}\}$ is an α-mixing sequence which for $t < \tau + s_T$ satisfies*

$$\sup_{\substack{A\in\sigma(Z_{t,T},Z_{t-1,T},\ldots)\\ B\in\sigma(Z_{\tau,T},Z_{\tau+1,T},\ldots)}} \big|P(A\cap B) - P(A)P(B)\big| \leq a(|t-\tau| - s_T). \tag{4.23}$$

Let $W_{i,T} = \sum_{t=ir_T+1}^{ir_T+p_T} Z_{t,T}$, where $r_T = p_T + q_T$ and $\{\tilde{W}_{i,T}\}$ be independent random variables where the marginal distributions of $\tilde{W}_{i,T}$ and $W_{i,T}$ are the same. Then, for any $x \in \mathbb{R}$, we have

$$\left|\mathbb{E}\bigg(\exp(ix\sum_{j=1}^{k_T} W_{j,T})\bigg) - \prod_{j=1}^{k_T}\mathbb{E}\bigg(\exp(ix\tilde{W}_{j,T})\bigg)\right| \leq Ck_T a(q_T - s_T),$$

where C is a finite constant.

PROOF. By expanding $\mathbb{E}\Big(\exp(ix\sum_{j=1}^{k_T} W_{j,T})\Big) - \prod_{j=1}^{k_T}\mathbb{E}\Big(\exp(ix\tilde{W}_{j,T})\Big)$, we have

$$\begin{aligned} D_T &= \Big|\mathbb{E}\Big(\exp(ix\sum_{j=1}^{k_T} W_{j,T})\Big) - \prod_{j=1}^{k_T}\mathbb{E}\Big(\exp(ix\tilde{W}_{j,T})\Big)\Big| \\ &\leq \sum_{s=1}^{k_T-1} \Big|\prod_{r=1}^{s-1}\mathbb{E}\big(\exp(ixW_{r,T})\big)\Big|\Big|\mathrm{cov}\Big(\exp(ixW_s), \exp(ix\sum_{j=s+1}^{k_T} W_j)\Big)\Big|, \end{aligned}$$

(to simplify notation we denote $\prod_{r=1}^{0} A_r = 1$). From the definition of $W_{i,T}$ and by using Ibragimov's inequality (for bounded random variables) it is straightforward to show that

$$D_T \leq \sum_{s=1}^{k_T-1} \sup_{\substack{A\in\sigma(Z_{(s+1)r_T+1,T}, Z_{(s+1)r_T+2,T},\ldots) \\ B\in\sigma(Z_{sr_T+p_T,T}, Z_{sr_T+p_T-1})}} |P(A\cap B) - P(A)P(B)|| \leq Ck_T a(q_T - s_T).$$

The above gives the required result. $\square$

Lemma IV.6 *Suppose that Assumption IV.1 holds, and we choose p_T and q_T such that $p_T >> q_T >> M$ and $k_T = (\log T)^{1/2}$, where $k_T = T/(p_T + q_T)$, then the asymptotic distributions of $V_T^{-1/2}(Q_{T,M} - \mathbb{E}(Q_{T,M}))$ and $V_T^{-1/2}\tilde{\mathcal{S}}_{k_T}$ are equivalent.*

PROOF. From (4.22) we have $V_T^{-1/2}(Q_{T,M} - \mathbb{E}(Q_{T,M})) = V_T^{-1/2}\mathcal{S}_{k_T} + o_p(1)$. By using Theorem IV.3 with $Z_{t,T} := Y_{t,T} = T^{-1}\sum_{\tau=\max(t-M,1)}^{t} F_{t,\tau,M}(X_{t,T}X_{\tau,T} - \mathbb{E}(X_{t,T}X_{\tau,T}))$ and $W_{i,T} := U_{i,T}$ we have

$$|\Phi_{k_T}(x) - \tilde{\Phi}_{k_T}(x)| \leq k_T\alpha(q_T - M),$$

where $\Phi_{k_T}(\cdot)$ and $\tilde{\Phi}_{k_T}(\cdot)$ are the characteristic functions of $\mathcal{S}_{k_T}$ and $\tilde{\mathcal{S}}_{k_T}$. Since $p_T >> q_T >> M$ and $k_T = (\log T)^{1/2}$, and under Assumption IV.1 (i) we have that $|\Phi_{k_T}(x) - \tilde{\Phi}_{k_T}(x)| \to 0$. Since the characteristic functions converge, we obtain the required result. $\square$

We now show asymptotic normality of $V_T^{-1/2}\tilde{\mathcal{S}}_{k_T}$, this result together with the above lemma will give Theorems IV.1 and IV.2.

Lemma IV.7 *Suppose Assumption IV.1 is satisfied. Then we have*

$$V_T^{-1/2}\tilde{\mathcal{S}}_{k_T} \xrightarrow{D} \mathcal{N}(0,1).$$

PROOF. We will use the central limit theorem for independent random variables. Due to the independence of $\tilde{U}_{i,T}$ it is straightforward to show $\frac{1}{T}\sum_{i=1}^{k_T}\mathbb{E}(\tilde{U}_{i,T}^2) \to V_T$, hence it remains to verify Lindeberg's condition. By using (4.17) we have $\sum_{i=1}^{k_T}\mathbb{E}\big[(V_T^{-1/2}|\tilde{U}_{i,T}|)^{2+\delta}\big] \leq K\big(\frac{p_T}{T}\big)^{\delta/2} \to 0$, as $T \to \infty$. Thus Lindeberg's condition is fulfilled and we have asymptotic normality of $\tilde{\mathcal{S}}_{k_T}$. □

PROOF of Theorem IV.1 To prove the result we show that Q_T we use that $Q_T = Q_{T,T^{1/2+\gamma}} + O_p(T^{-1/2-\gamma})$, where

$$Q_{T,T^{1/2+\gamma}} = \frac{1}{T}\sum_{t,\tau=1}^{T} I(\frac{t-\tau}{T^{1/2+\gamma}})G_{t,\tau,T}X_{t,T}X_{\tau,T}.$$

Thus by (4.22) we have

$$\mathrm{var}(Q_T)^{-1/2}\big(Q_T - \mathbb{E}(Q_T)\big) = \mathrm{var}(Q_T)^{-1/2}\mathcal{S}_{T,T^{1/2+\gamma}} + o_p(1), \tag{4.24}$$

and $\mathrm{var}(\sqrt{T}Q_T) = \mathrm{var}(\sqrt{T}Q_{T,T^{1/2+\gamma}}) + O(T^{-\gamma})$. We observe that $Q_{T,T^{1/2+\gamma}}$ satisfies representation (4.3) and Assumption IV.1, thus by applying Lemma IV.7, then $V_T^{-1/2}(Q_{T,T^{1/2+\gamma}} - \mathbb{E}(Q_T)) \xrightarrow{D} \mathcal{N}(0,1)$. Therefore from (4.24) we have $V_T^{-1/2}(Q_T - \mathbb{E}(Q_T)) \xrightarrow{D} \mathcal{N}(0,1)$, which gives the desired result. □

PROOF of Theorem IV.2 By using Lemma IV.6, it is straightforward to show that $V_T^{-1/2}(Q_{T,M} - \mathbb{E}(Q_{T,M}))$ and $V_T^{-1/2}\tilde{\mathcal{S}}_{k_T}$ have asymptotically the same distribution. Now by using the same arguments as in the proof of Theorem IV.1 we obtain the result. $\square$

4.2. Proofs of results in Section 3

PROOF of Lemma IV.1 To prove the lemma we apply a result from Statulevicius and Jakimavicius (1988), Theorem 3, part (2), which states that if $t_1 \leq t_2 \leq \ldots \leq t_k$, then for all $2 \leq i \leq k$ we have $\left|\mathrm{cum}(X_{t_1,T}, X_{t_2,T}, \ldots, X_{t_k,T})\right| \leq 3(k-1)!2^{k-1}\alpha(t_i - t_{i-1})^{1-\frac{k}{r}} \sup_{t,T} \|X_{t,T}\|_r^k$.

To prove (i), we use a method similar to the proof of Neumann (1996), Remark 3.1. By taking the $(k-1)$th root of the above for all $2 \leq i \leq k$ we have

$$\left|\mathrm{cum}(X_{t_1,T}, X_{t_2,T}, \ldots, X_{t_k,T})\right|^{\frac{1}{k-1}} \leq C_k^{1/(k-1)}\alpha(t_i - t_{i-1})^{\frac{1-k/r}{k-1}} \sup_{t,T} \|X_{t,T}\|_r^{\frac{k}{k-1}},$$

where $C_k = 3(k-1)!2^{k-1}$. Since the above bound holds for all i, multiplying the above over i gives

$$\left|\mathrm{cum}(X_{t_1,T}, X_{t_2,T}, \ldots, X_{t_k,T})\right| \leq C_k \sup_{t,T} \|X_{t,T}\|_r^k \prod_{i=2}^{k} \alpha(t_i - t_{i-1})^{\frac{1-k/r}{k-1}}, \qquad (4.25)$$

thus proving (i) of the lemma.

To prove (ii), we rewrite $\sum_{t_2,\ldots,t_k=1}^{\infty}$ as the sum of orderings, that is $\sum_{t_2,\ldots,t_k=1}^{\infty} = k!\sum_{1=t_2\leq\ldots\leq t_k}^{\infty}$. Now since the number of orderings is finite, we can use (i) to obtain

$$\sum_{t_2,\ldots,t_k=1}^{\infty} \left|\mathrm{cum}(X_{t_1,T}, X_{t_2,T}, \ldots, X_{t_k,T})\right| \leq C_k \sup_{t,T} \|X_{t,T}\|_r^k \Big\{\sum_r \alpha(r)^{\frac{(1-k/r)}{(k-1)}}\Big\}^{k-1} < \infty,$$

which gives (4.6).

To prove (iii) we use a similar argument to obtain

$$\begin{aligned}
&\sum_{t_2,\ldots,t_k=1}^{\infty}(1+|t_j|)\big|\mathrm{cum}(X_{t_1,T},X_{t_2,T},\ldots,X_{t_k,T})\big| \\
\leq &\sum_{1\leq t_2<\ldots<t_k<\infty}(1+|t_j|)\big|\mathrm{cum}(X_{t_1,T},X_{t_2,T},\ldots,X_{t_k,T})\big| \\
= & \; k!\sum_{r_2,\ldots,r_k=1}^{\infty}(1+\sum_{i=2}^{j}|r_i|)\big|\mathrm{cum}(X_{t_1,T},X_{t_2,T},\ldots,X_{t_k,T})\big|,
\end{aligned}$$

substituting (4.25) into the above gives the result. □

PROOF of Lemma IV.2 To prove the result we use the notion of Near Epoch Dependence. This requires bounding

$$\big\|Y_{t,T}-\mathbb{E}(Y_{t,T}|\mathcal{F}_{t-j})\|_q = A_1 + A_2$$

where

$$\begin{aligned}
A_1 &= \Big\|\sum_{\tau<t}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})-\mathbb{E}(\sum_{\tau<t}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})|\mathcal{F}_{t-j})\Big\|_q \\
A_2 &= \Big\|\sum_{\tau\leq t}G_{\tau,t,M}h(X_{\tau,T},X_{t,T})-\mathbb{E}(\sum_{\tau\leq t}G_{\tau,t,M}h(X_{\tau,T},X_{t,T})|\mathcal{F}_{t-j})\Big\|_q.
\end{aligned}$$

As the derivation of bounds on A_1 and A_2 are identical, we shall focus on A_1. We first observe that by using the Minkowski inequality we have

$$A_1 = \Big\|\sum_{\tau=1}^{t-1}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})-\mathbb{E}(\sum_{\tau<t}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})|\mathcal{F}_{t-j})\Big\|_q \leq I + II,$$

where

$$\begin{aligned}
I &= \big\|\sum_{\tau=1}^{t-1}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})-\mathbb{E}(\sum_{\tau<t}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})|\mathcal{F}_{t-j/2}^{t})\big\|_q \\
II &= \big\|\mathbb{E}\big(\mathbb{E}(\sum_{\tau=1}^{t-1}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})|\mathcal{F}_{t-j/2}^{t})|\mathcal{F}_{t-j}\big)-\mathbb{E}\big(\sum_{\tau=1}^{t-1}G_{t,\tau,M}h(X_{t,T},X_{\tau,T})\big)\big\|_q.
\end{aligned}$$

and $\mathcal{F}^t_{t-j/2} = \sigma(X_{t,T}, X_{t-1,T}, \ldots, X_{t-j/2,T})$. To bound I we note that for $t > \tau$ and all j we have

$$\begin{aligned}
& \Big\| \sum_{\tau=1}^{t-1} G_{t,\tau,M} h(X_{t,T}, X_{\tau,T}) - \mathbb{E}\big(\sum_{\tau=1}^{t-1} G_{t,\tau,M} h(X_{t,T}, X_{\tau,T}) | \mathcal{F}^t_{t-j}\big) \Big\|_q \\
= & \Big\| \sum_{k=j} G_{t,t-k} \big\{ h(X_{t,T}, X_{t-k,T}) - \mathbb{E}(h(X_{t,T}, X_{t-k,T}) | \mathcal{F}^t_{t-j}) \big\} \Big\|_q \\
\leq & \sum_{k=j} |G_{t,t-k}| \Big\| \big\{ h(X_{t,T}, X_{t-k,T}) - \mathbb{E}(h(X_{t,T}, X_{t-k,T}) | \mathcal{F}^t_{t-j}) \big\} \Big\|_q \\
\leq & K \sup_{\tau,T} \| h(X_{t,T}, X_{\tau,T}) \|_q \sum_{k=j}^{\infty} k^{-\delta} = K \sup_{\tau,T} \| h(X_{t,T}, X_{\tau,T}) \|_q (j/2)^{-(\delta-1)}.
\end{aligned}$$

where we use that $|G_{t,t-k,M}| \leq K|t-\tau|^{\delta}$. Furthermore, to bound II we use that $\mathbb{E}(\sum_{\tau=1}^{t-1} G_{t,\tau,M} h(X_{t,T}, X_{\tau,T}) | \mathcal{F}^t_{t-j/2}) \in \mathcal{F}^t_{t-j/2}$ together with Ibragimov's inequality to obtain

$$II \leq K(j/2)^{-s(\frac{1}{q}-\frac{1}{r})} \sum_{\tau<t} |G_{t,\tau,M}| \| h(X_{t,T}, X_{\tau,T}) \|_r$$

Thus altogether we have

$$A_1 \leq K \Big(j^{-(\delta-1)} \sup_{t,\tau,T} \| h(X_{t,T}, X_{\tau,T}) \|_q + \sup_{t,\tau,T} \| h(X_{t,T}, X_{\tau,T}) \|_r j^{-s(\frac{1}{q}-\frac{1}{r})} \Big).$$

A similar bound also applies to A_2, thus altogether this gives

$$\big\| Y_{t,T} - \mathbb{E}(Y_{t,T} | \mathcal{F}_{t-j}) \big\|_q \leq K \Big(j^{-(\delta-1)} \| h(X_{t,T}, X_{t-j,T}) \|_q + \| h(X_{t,T}, X_{t-j,T}) \|_r j^{-s(\frac{1}{q}-\frac{1}{r})} \Big),$$

and we have shown the first part of the required result.

To show the second part we note that since $\big\| Y_{t,T} - \mathbb{E}(Y_{t,T} | \mathcal{F}_{t-j}) \big\|_q \to 0$ as $T \to \infty$, thus we almost surely have the representation $Y_{t,T} - \mathbb{E}(Y_{t,T}) = \sum_j N_{j,T}(t-j)$, where

$$N_{j,T}(t-j) = \mathbb{E}(Y_{t,T} | \mathcal{F}_{t-j}) - \mathbb{E}(Y_{t,T} | \mathcal{F}_{t-j-1}).$$

Thus substituting the above into $\| B^{(u)}_{T,S_T} \|_q$ and using the Burkholder inequality we have

$$\begin{aligned}
B^{(u)}_{T,S_T} &= \Big\| \sum_{t=u}^{S_T+u} \sum_{j=0}^{\infty} N_{j,T}(t-j) \Big\|_q \\
&\leq \sum_{j=0}^{\infty} \Big\| \sum_{t=u}^{S_T+u} N_{j,T}(t-j) \Big\|_q \\
&\leq \sum_{j=0}^{\infty} \Big(\sum_{t=u}^{S_T+u} \|N_{j,T}(t-j)\|_q^2 \Big)^{1/2} \\
&\leq S_T^{1/2} \Big(\sum_{j=1}^{\infty} (j^{-\delta+2} + j^{-s(\frac{1}{q}-\frac{1}{r})+1}) \Big),
\end{aligned}$$

as required. □

PROOF of Lemma IV.3 The proof of (4.14) follows immediately from Ibragimov's inequality (Ibragimov (1962)) (see also Davidson (1994), Theorem 14.2). Using this we note that since $X_{t,T} = \mathbb{E}(X_{t,T}|\mathcal{F}_{t,T})$ and $\mathbb{E}(X_{t,T}|\mathcal{F}_{t-j}) \to 0$ as $j \to \infty$, almost surely we have

$$X_{t,T} = \sum_{j=0}^{\infty} \big(\mathbb{E}_{t-j}(X_{t,T}) - \mathbb{E}_{t-j-1}(X_{t,T}) \big). \tag{4.26}$$

To prove (4.15), we use Ibragimov's and Chebyshev's inequalities and (4.14) to obtain

$$\begin{aligned}
& \Big\| \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)|\mathcal{F}_{t-j_1-i}) - \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)|\mathcal{F}_{t-j_1-i-1}) \Big\|_q \\
\leq\ & 2\Big\| \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)|\mathcal{F}_{t-j_1-i}) - \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)) \Big\|_q \\
\leq\ & 4(2^{1/q}+1)\|M_{j_1}(t-j_1)M_{j_2}(t-j_1)\|_r \alpha(i)^{\frac{1}{q}-\frac{1}{r}} \\
\leq\ & 12\|M_{j_1}(t-j_1)\|_{2r}\|M_{j_2}(t-j_1)\|_{2r}\alpha(i)^{\frac{1}{q}-\frac{1}{r}} \\
\leq\ & 12^3 \sup_{t,T} \|X_{t,T}\|_{\tilde{r}}^2 \alpha(j_1)^{\frac{1}{2r}-\frac{1}{\tilde{r}}} \alpha(j_2)^{\frac{1}{2r}-\frac{1}{\tilde{r}}} \alpha(i)^{\frac{1}{q}-\frac{1}{r}},
\end{aligned}$$

where $\tilde{r}/2 > r > q$. Now we prove (4.16). By substituting (4.26) into $B^{(u)}_{T,S_T}$ and using the above notation for conditional expectations we have

$$
\begin{aligned}
B_{T,S_T}^{(u)} &= T^{-1}\sum_{t=u+1}^{S_T+u}\sum_{\tau} F_{t,\tau,M}\big(X_{t,T}X_{\tau,T} - \mathbb{E}(X_{t,T}X_{\tau,T})\big) \\
&= T^{-1}\sum_{j_1,j_2=0}^{\infty}\sum_{t=u+1}^{S_T+u}\sum_{\tau=\max(t-M,1)}^{t} F_{t,\tau,M}\Big(M_{j_1}(t-j_1)M_{j_2}(\tau-j_2) \\
&- \mathbb{E}\big(M_{j_1}(t-j_1)M_{j_2}(\tau-j_2)\big)\Big).
\end{aligned}
$$

Partitioning the above sum into various cases and using Minkowski's inequality gives

$$
\begin{aligned}
\big\|B_{T,S_T}^{(u)}\big\|_q &= T^{-1}\sum_{j_1,j_2=0}^{\infty}\Big\|\sum_{t=u+1}^{S_T+u}\sum_{\tau} F_{t,\tau,M}\Big(M_{j_1}(t-j_1)M_{j_2}(\tau-j_2) \\
&- \mathbb{E}\big(M_{j_1}(t-j_1)M_{j_2}(\tau-j_2)\big)\Big)\Big\|_q \\
&\leq I + II + III,
\end{aligned}
$$

where

$$
\begin{aligned}
I &= T^{-1}\sum_{j_1,j_2=0}^{\infty}\Big\|\sum_{t=u+1}^{S_T+u}\sum_{\tau<t-j_1+j_2} F_{t,\tau,M}M_{j_1}(t-j_1)M_{j_2}(\tau-j_2)\Big\|_q \\
II &= T^{-1}\sum_{j_1,j_2=0}^{\infty}\Big\|\sum_{\tau}\sum_{t<\tau-j_2+j_1} F_{t,\tau,M}M_{j_1}(t-j_1)M_{j_2}(\tau-j_2)\Big\|_q \\
III &= T^{-1}\Big\|\sum_{j_1,j_2=0}^{\infty}\sum_{t=u+1}^{S_T+u} F_{t,t-j_1+j_2,M}\Big(M_{j_1}(t-j_1)M_{j_2}(t-j_1) \\
&- \mathbb{E}\big(M_{j_1}(t-j_1)M_{t-j_1+j_2}(t-j_1)\big)\Big)\Big\|_q.
\end{aligned}
$$

We observe that $\big\{\sum_{\tau<t-j_1+j_2} F_{t,\tau,M}M_{j_1}(t-j_1)M_{j_2}(\tau-j_2)\big\}_t$ and $\big\{\sum_{t<\tau-j_2+j_1}^{T} F_{t,\tau,M}M_{j_1}(t-j_1)M_{j_2}(\tau-j_2)\big\}_\tau$ are martingale differences. Therefore by using the Burkholder-Rosenthal inequality twice together with Cauchy-Schwarz, for

$q \geq 2$ we have

$$\begin{aligned} I &\leq T^{-1} \sum_{j_1,j_2=0}^{\infty} \Big(\sum_{t=u+1}^{S_T+u} \Big\| \sum_{\tau < t-j_1+j_2} F_{t,\tau,M} M_{j_1}(t-j_1) M_{j_2}(\tau - j_2) \Big\|_q^2 \Big)^{1/2} \\ &\leq T^{-1} \sum_{j_1,j_2=0}^{\infty} \Big(\sum_{t=u+1}^{S_T+u} \big\| M_{j_1}(t-j_1) \big\|_{2q}^2 \Big\| \sum_{\tau < t-j_1+j_2} F_{t,\tau,M} M_{j_2}(\tau - j_2) \Big\|_{2q}^2 \Big)^{1/2} \\ &\leq T^{-1} \sum_{j_1,j_2=0}^{\infty} \Big(\sum_{t=u+1}^{S_T+u} \big\| M_{j_1}(t-j_1) \big\|_{2q}^2 \sum_{\tau < t-j_1+j_2} |F_{t,\tau,M}|^2 \big\| M_{j_2}(\tau - j_2) \big\|_{2q}^2 \Big)^{1/2}. \end{aligned}$$

Using (4.14) we have $\|M_j(t-j)\|_{2q} \leq C\alpha(j)^{\frac{1}{2q}-\frac{1}{r}}$. Substituting these bounds into I and under Assumption IV.1 (i) we have

$$\begin{aligned} I &\leq T^{-1} C \Big(\sum_{j=0}^{\infty} \alpha(j)^{\frac{1}{2q}-\frac{1}{r}} \Big)^2 \Big(\sum_{t=u+1}^{S_T+u} \sum_{\tau} |F_{t,\tau,M}|^2 \Big)^{1/2} \\ &\leq T^{-1} S_T^{1/2} K \Big(\sum_{j=0}^{\infty} \alpha(j)^{\frac{1}{2q}-\frac{1}{r}} \Big)^2 \sup_t \Big(\sum_{\tau} |F_{t,\tau,M}|^2 \Big)^{1/2} \\ &\leq K T^{-1} S_T^{1/2} G_M^{1/2} \Big(\sum_{j=0}^{\infty} \alpha(j)^{\frac{1}{2q}-\frac{1}{r}} \Big)^2. \end{aligned} \tag{4.27}$$

Using the same methods we have

$$II \leq K T^{-1} S_T^{1/2} G_M^{1/2} \Big(\sum_{j=0}^{\infty} \alpha(j)^{\frac{1}{2q}-\frac{1}{r}} \Big)^2. \tag{4.28}$$

Finally we obtain a bound for III. This requires a more delicate analysis since $\{M_j(t-j)M_{t-(\tau-j)}(t-j) - \mathbb{E}(M_j(t-j)M_{t-(\tau-j)}(t-j))\}$ are not necessarily martingale differences over t. We first represent $M_j(t-j_1)M_{j_2}(t-j_1) - \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1))$ as the sum of martingale differences. Since $\mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)|\mathcal{F}_{t-j_1-i}) \xrightarrow{a.s.} \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1))$, as $i \to \infty$, we have

$$M_j(t-j_1)M_{\tau-(t-j)}(t-j) - \mathbb{E}(M_j(t-j)M_{\tau-(t-j)}(t-j)) = \sum_{i=0}^{\infty} A_{j_1,j_2;i}(t-j_1-i),$$

almost surely, where

$$\begin{aligned}
& A_{j_1,j_2;i}(t-j_1-i) \\
= \ & \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)|\mathcal{F}_{t-j_1-i}) - \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1)|\mathcal{F}_{t-j_1-i-1}).
\end{aligned}$$

Substituting this into III and using Minkowski's inequality gives

$$\begin{aligned}
& III = \\
& \Big\|T^{-1}\sum_{j_1,j_2=0}^{\infty}\sum_{t=u+1}^{u+S_T} F_{t,t-j_2+j_2,M}\big(M_{j_1}(t-j_1)M_{j_2}(t-j_1) - \mathbb{E}(M_{j_1}(t-j_1)M_{j_2}(t-j_1))\big)\Big\|_q \\
& \leq T^{-1}\sum_{j_1,j_2=0}^{\infty}\sum_{i=0}^{\infty}\Big\|\sum_{t=u+1}^{u+S_T} F_{t,t-j_1+j_2,M}A_{j_1,j_2;i}(t-j_1-i)\Big\|_q.
\end{aligned}$$

We observe that since $A_{j_1,j_2;i}(t-j_1-i) \in \sigma(X_{t-j_1-i}, X_{t-j_1-i-1}, \ldots)$ and $\mathbb{E}(A_{j_1,j_2;i}(t-j_1-i)|\sigma(X_{t-j-i-1}, X_{t-j-i-2}, \ldots)) = 0$, then $\{A_{j_1,j_2;i}(t-j_1-i)\}_t$ are martingale differences. Therefore by using the Burkholder-Rosenthal and Hölder on the above yields

$$III \ \leq \ T^{-1}\sum_{j_1,j_2=0}^{\infty}\sum_{i=0}^{\infty}\Big(\sum_{t=u+1}^{u+S_T}|F_{t,t-j_1+j_2,M}|^2\big\|A_{j_1,j_2;i}(t-j_1-i)\big\|_q^2\Big)^{1/2}$$

Substituting (4.15) into III gives

$$\begin{aligned}
III \leq CT^{-1} & \sum_{j_1,j_2=0}^{\infty}\sum_{i=0}^{\infty}\Big\{\sum_{t=u+1}^{u+S_T}|F_{t,t-j_1+j_2,M}|^2\big(\|X_t\|_{\tilde{r}}^2\alpha(j_1)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\alpha(j_2)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\alpha(i)^{\frac{1}{q}-\frac{1}{r}}\big)^2\Big\}^{1/2} \\
\leq \ & T^{-1}\sum_{j_1,j_2=0}^{\infty}\sum_{i=0}^{\infty}\alpha(i)^{\frac{1}{q}-\frac{1}{r}}\alpha(j_1)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\alpha(j_2)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\Big\{\sum_{t=u+1}^{u+S_T}\sup_{\tau}F_{t,\tau,M}^2\Big\}^{1/2} \\
\leq \ & CT^{-1}S_T^{1/2}\Big(\sum_{i=0}^{\infty}\alpha(i)^{\frac{1}{q}-\frac{1}{r}}\Big)\Big(\sum_{i=0}^{\infty}\alpha(i)^{\frac{1}{2r}-\frac{1}{\tilde{r}}}\Big)^2 \qquad (4.29)
\end{aligned}$$

Finally, we substitute (4.27), (4.28) and (4.29) into $\|B_{T,S_T}^{(u)}\|_q$ to obtain (4.16). $\square$

CHAPTER V

SUMMARY

In this dissertation, we explore new analytic tools for nonlinear time series mainly focusing on frequency domain approach. We propose two new spectral densities which can describe the dependence structure and periodicities of nonlinear time series.

In Chapter II, we introduce the quantile spectral density which captures serial dependence in time series data without requiring linearity and certain moment assumptions. We estimate the quantile spectral density using L_2 methods and derive the sampling properties of the estimator. We develop a goodness-of-fit test using the quantile spectral density and propose a bootstrap method for estimating the finite sampling distribution of the test statistic under the null hypothesis. Through some simulations and real data example, we illustrate how this new method can be used for linear and nonlinear time series analysis.

In Chapter III, we propose the association spectral density which can detect periodicities on different parts of the domain of the time series. We consider the properties of the association spectral density and propose a method of estimation. The asymptotic properties of the estimator is derived and some simulation result is given.

In Chapter IV, we consider general quadratic forms of nonstationary, α-mixing time series and derive asymptotic normality of these forms under some moment assumptions. In order to show asymptotic normality of the generalized quadratic form, we obtain some bounds on moments and cumulants using mixingale and near-epoch dependent methods.

REFERENCES

Anderson, T. W. (1993). Goodness of fit tests for spectral distributions. *Annals of Statistics* **21**, 830–847.

Anderson, T. W. (1997). Goodness-of-fit tests for autoregressive processes. *Journal of Time Series Analysis* **18**, 321–339.

Andrews, D. F. and Stafford, J. E. (1998). Iterated full partitions. *Statistics and Computing* **8**, 189-192.

Berkowitz, J. and Diebold, F. X. (1998). Bootstrapping multivariate spectra. *Review of Economics and Statistics* **80**, 664–666.

Bhansali, R., Giraitis, L., and Kokoszka, P. (2007). Approximations and limit theory for quadratic forms of linear variables stochastic processes and their applications. *Stochastic Processes and their Applications* **117**, 71-95.

Blum, J. R., Kiefer, J., and Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *Annals of Mathematical Statistics* **32**, 485–498.

Bochner, S. (1955). *Harmonic analysis and the theory of probability.* Berkeley: University of California Press.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.

Box, G. E. P. and Jenkins, G. M. (1970). *Times series analysis. Forecasting and control.* San Francisco: Holden-Day.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association* **65**, 1509–1526.

Bradley, R. C. (2007). *Introduction to strong mixing conditions.* Heber City: Kendrick Press.

Brillinger, D. (1981). *Time series, data analysis and theory.* San Francisco: SIAM.

Chen, W. W. and Deo, R. S. (2004). A generalized Portmanteau goodness-of-fit test for time series models. *Econometric Theory* **20**, 382–416.

Cline, D. and Pu, H. (1999). Geometric ergodicity of nonlinear time series. *Statistica Sinica* **9**, 1103-118.

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Annals of Statistics* **28**, 1762-1794.

Davidson, J. (1994). *Stochastic Limit Theory.* Oxford: Oxford University Press.

Dette, H. and Paparoditis, E. (2009). Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities. *Journal of the Royal Statistical Society, Series B* **71**, 831–857.

Dette, H., Hallin, M., Kley, T., and Volgushav, S. (2011). Of copulas, quantile, ranks and spectra: and l_1 approach to spectral analysis. *arxiv:1111.7205.*

Doukhan, P. (1994). *Mixing, properties and examples.* New York: Springer.

Dwivedi, Y. and Subba Rao, S. (2011). A test for second order stationarity based on the discrete fourier transform. *Journal of Time Series Analysis* **32**, 68-91.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987–1007.

Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates. *Annals of Statistics* **20**, 121–145.

Fryzlewicz, P. and Subba Rao, S. (2011). On mixing properties of ARCH and time-varying ARCH processes. *Bernoulli* **13**, 320-346.

Gao, J. and Anh, V. (2000). A central limit theorem for random quadratic form of strictly stationary processes. *Statistics and Probability Letters* **49**, 69-79.

Götze, F. and Tikhomirov, A. N. (1999). Asymptotic distribution of quadratic forms. *Annals of Probability* **27**, 1072-1098.

Hagemann, A. (2011). Robust spectral analysis. *arxiv:111.1965v1, Arxiv, e-prints.*

Hallin, M. and Puri, M. L. (1992). Rank tests for time series analysis: a survey. In *New directions in time series analysis, Part I* (Vol. 45, pp. 111–153). New York: Springer.

Hoeffding, W. (1948). A non-parametric test of independence. *Annals of Mathematical Statistics* **19**, 546–557.

Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* **64**, 837–864.

Hong, Y. (1998). Testing for pairwise serial independence via the empirical distribution function. *Journal of the Royal Statistical Society, Series B* **60**, 429–453.

Hong, Y. (1999). Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *Journal of the American Statistical Association* **94**, 1201–1220.

Hong, Y. (2000). Generalized spectral tests for serial dependence. *Journal of the Royal Statistical Society, Series B* **62**, 557–574.

Hong, Y. and Lee, T.-H. (2003). Diagnostic checking for the adequacy of nonlinear time series models. *Econometric Theory* **19**, 1065–1121.

Hurvich, C. and Zeger, S. (1987). Frequency domain bootstrap methods for time series. *New York University, Graduate School of Business Adminstration.*

Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory of Probability and its Applications* **7**, 349-82.

Kokoszka, P. and Taqqu, M. S. (1997). The asympototic behavior of quadratic forms in heavy-tailed strongly dependent randon variables. *Stochastic Processes and their Applications* **66**, 21-40.

Li, T.-H. (2008). Laplace periodograms for time series analysis. *Journal of the American Statistical association* **103**, 757-768.

Lin, F. (2009). The asymptotic behaviour of quadratic forms in ϕ-mixing random variables. *Journal of Computational and Applied Mathematics* **233**, 437-448.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 297–303.

Mikosch, T. (1990). Functional limit theorems for random quadratic forms. *Stochastic Processes and their Application* **37**, 81-98.

Milhøj, A. (1981). A test of fit in time series models. *Biometrika* **68**, 177–187.

Neumann, M. (1996). Spectral density estimation via nonlinear wavelet methods for stationary nongaussian time series. *Journal of Time Series Analysis* **17**, 601-633.

Neumann, M. H. and Paparoditis, E. (2008). Goodness-of-fit tests for Markovian time series models: central limit theory and bootstrap approximations. *Bernoulli* **14**, 14–46.

Parzen, E. (1974). Some recent advances in time series modeling. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control* **AC-19**, 723–730.

Pinkse, J. (1998). A consistent nonparametric test for serial independence. *Journal of Econometrics* **84**, 205–231.

Priestley, M. B. (1981). *Spectral analysis and time series.* London: Academic Press.

Rosenblatt, M. (1984). Asymptotic normality, strong mixing and spectral density estimators. *Annals of Probability* **12**, 1167-1180.

Shumway, R. H. and Stoffer, D. S. (2006). *Time series analysis and its applications: With R examples* (Second ed.). New York: Springer.

Skaug, H. J. and Tjøstheim, D. (1993). A nonparametric test of serial independence based on the empirical distribution function. *Biometrika* **80**, 591–602.

Smith, B. and Field, C. (2001). Symbolic cumulant calculations for frequency domain time series. *Statistics and Computing* **11**, 75–82.

Stafford, J. E. (1994). Automating the partition of indexes. *Journal of Computational and Graphical Statistics* **3**, 249–259.

Statulevicius, V. and Jakimavicius. (1988). Estimates of semiivariant and centered moments of stochastic processes with mixing: I. *Lithuanian Mathematical Journal* **28**, 226-238.

Tjøstheim, D. (1990). Nonlinear time series and Markov chains. *Advances in Applied Probability* **22**, 587–611.

Tsay, R. S. (2005). *Analysis of financial time series* (2nd ed.). Hoboken: Wiley.

Velilla, S. (1994). A goodness-of-fit test for autoregressive moving-average models based on the standardized sample spectral distribution of the residuals. *Journal of Time Series Analysis* **15**, 637–647.

Wu, W. B. and Shao, X. (2007). A limit theorem for quadratic forms and its applications. *Econometric Theory* **23**, 930-951.

Zivot, E. (2009). Practical issues in the analysis of univariate garch models. In T. G. Andersen, R. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of financial time series.* New York: Springer.

APPENDIX A

SUPPLEMENT TO THE PROOF OF LEMMA II.1 IN CHAPTER II

We want to show that the third part of the variance of $\mathcal{Q}_T$ in Lemma II. 1 in Chpater II is in smaller order than the first and second term which are $O(\frac{M}{T^2})$. We recall that

$$\begin{aligned} II_3 &= \frac{1}{T^2}\sum_{s_1,s_2=1}^{T}\int\sum_{k_1,k_2,k_3,k_4}\mathrm{cum}(J_{k_1,x_1}\bar{J}_{k_1,y_1}, J_{k_2,x_1}\bar{J}_{k_2,y_1}, \bar{J}_{k_3,x_2}J_{k_3,y_2}, \overline{J}_{k_4,x_2}J_{k_4,y_2}) \\ & \prod_{i=1}^{2}K_M(\omega_{s_1}-\omega_{k_i})\prod_{i=3}^{4}K_M(\omega_{s_2}-\omega_{k_i})dF_0(x_1)dF_0(y_1)dF_0(x_2)dF_0(y_2). \end{aligned}$$

Let $\{X_{ij}\}$ be the ith row and jth column element of the matrix $\begin{pmatrix} J_{k_1,x_1} & \overline{J}_{k_1,y_1} \\ J_{k_2,x_1} & \overline{J}_{k_2,y_1} \\ \overline{J}_{k_3,x_2} & J_{k_3,y_2} \\ \overline{J}_{k_4,x_2} & J_{k_4,y_2} \end{pmatrix}$ and $Y_i = \prod_{j=1}^{2} X_{ij}$ for $i = 1,\dots,4$. With these notations, the above cumulant is represented as $\mathrm{cum}(Y_1,Y_2,Y_3,Y_4)$ and Theorem 2.3.2 in Brillinger (1981) gives

$$\mathrm{cum}(Y_1,Y_2,Y_3,Y_4) = \sum_{\nu_k}\mathrm{cum}(X_{ij}; ij\in\nu_{k,1})\cdots\mathrm{cum}(X_{ij}; ij\in\nu_{k,p})$$

where the sum is taken over all indecomposable partitions of the two way table of indicies $\{i,j\}$, $i=1,\dots,4, j=1,2$. There are 3915 all indecomposable partitions of the above matrix, thus it is infeasible to find them by hand. We used the mathematica routine by Andrews and Stafford (1998) for this purpose.

For one indecomposable partition ν_k, we let $n(\nu_k)$ be $\{n(\nu_{k,1}),\dots,n(\nu_{k,p})\}$ where $n(A)$ is the number of elements in the set A and $n(\nu_{k,i}) \geq n(\nu_{k,i+1})$. There are 22 ways of $n(\nu_k)$, but using $\mathbb{E}(J_{k_i,x_i}) = 0$, we only need to consider the partitions such that $\min(n(\nu_k)) \geq 2$, *i.e.*. $n(\nu_k)$ is either one of {8}, {6, 2}, {5, 3}, {4, 4}, {4, 2, 2}, {3, 3, 2}, {2, 2, 2, 2}. With these notations, we can seperate II_3 in the following way.

$$\begin{aligned} II_3 &= \sum_{j=1}^{7} \sum_{\{\nu_k : n(\nu_k)=\mathcal{A}_j\}} \frac{1}{T^2} \sum_{s_1,s_2=1}^{T} \int \sum_{k_1,k_2,k_3,k_4} \prod_{i=1}^{2} K_M(\omega_{s_1} - \omega_{k_i}) \prod_{i=3}^{4} K_M(\omega_{s_2} - \omega_{k_i}) \times \\ & \quad \big(\mathrm{cum}(X_{ij}; ij \in \nu_{k,1}) \cdots \mathrm{cum}(X_{ij}; ij \in \nu_{k,p})\big) dF_0(x_1) dF_0(y_1) dF_0(x_2) dF_0(y_2) \\ &:= \sum_{j=1}^{7} II_{3,j} \end{aligned}$$

where $\mathcal{A}_j$ is the j-th element of

$$\mathcal{A} = \big\{\{8\}, \{6,2\}, \{5,3\}, \{4,4\}, \{4,2,2\}, \{3,3,2\}, \{2,2,2,2\}\big\}.$$

For these partitions, we apply Theorem 3.4.3 in Brillinger (1981) which gives us the bounds of the cumulants.

We start by $II_{3,1}$ with the case $n(\nu_k) = \{8\}$ where there's only one partition. Theorem 3.4.3 in Brillinger (1981) gives us

$$\mathrm{cum}(J_{k_1,x_1}, \overline{J}_{k_1,y_1}, J_{k_2,x_1}, \overline{J}_{k_2,y_1}, \overline{J}_{k_3,x_2}, J_{k_3,y_2}, \overline{J}_{k_4,x_2}, J_{k_4,y_2}) = O(\frac{1}{T^3}),$$

which leads to

$$II_{3,1} = O(\frac{1}{T^3}).$$

The one example of the cumulant terms in $II_{3,2}$ is

$$\mathrm{cum}(J_{k_1,x_1}, \overline{J}_{k_2,y_1})\mathrm{cum}(\overline{J}_{k_1,y_1} J_{k_2,x_1}, \overline{J}_{k_3,x_2}, J_{k_3,y_2}, \overline{J}_{k_4,x_2}, J_{k_4,y_2}) = \begin{cases} O(\frac{1}{T^2}) & k_1 = k_2 \\ O(\frac{1}{T^4}) & otherwise. \end{cases}$$

This leads to

$$\frac{1}{T^2}\sum_{s_1,s_2}\int\sum_{k_1,k_3,k_4} K_M(\omega_{s_1}-\omega_{k_1})K_M(\omega_{s_1}-\omega_{k_1})K_M(\omega_{s_2}-\omega_{k_3})K_M(\omega_{s_2}-\omega_{k_4})$$
$$\mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_2,y_1})\mathrm{cum}(\overline{J}_{k_1,y_1}J_{k_2,x_1},\overline{J}_{k_3,x_2},J_{k_3,y_2},\overline{J}_{k_4,x_2},J_{k_4,y_2})dF_0(x_1)\cdots dF_0(y_2)$$
$$=O(\frac{M}{T^3}).$$

With the same method, we can obtain $II_{3,2}=O(\frac{1}{T^3})$.

We consider the one partition in $II_{3,3}$

$$\mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1},J_{k_2,x_1})\mathrm{cum}(\overline{J}_{k_2,y_1},\overline{J}_{k_3,x_2},J_{k_3,y_2},\overline{J}_{k_4,x_2},J_{k_4,y_2})$$
$$=\begin{cases}\frac{1}{T^2}f_3(\omega_{k_1},-\omega_{k_1})f_5(-\omega_{k_2},\omega_{k_3},-\omega_{k_3},-\omega_{k_4}) & k_2=T\\ O(\frac{1}{T^4}) & otherwise.\end{cases}$$

, and this immediately leads to $II_{3,3}=O(\frac{1}{T^3})$.

In $II_{3,4}$, one example of sepeartions having the largest order is

$$\mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1},J_{k_2,x_1},\overline{J}_{k_3,x_2})\mathrm{cum}(\overline{J}_{k_2,y_1},J_{k_3,y_2},\overline{J}_{k_4,x_2},J_{k_4,y_2})$$
$$=\begin{cases}\frac{1}{T^2}f_4(\omega_{k_1},-\omega_{k_1},\omega_{k_2})f_4(-\omega_{k_2},\omega_{k_2},-\omega_{k_4}) & k_2=k_3\\ O(\frac{1}{T^4}) & otherwise.\end{cases}$$

, and this leads to

$$\frac{1}{T^2}\sum_{s_1,s_2}\sum_{k_1,k_2=k_3,k_4} K_M(\omega_{s_1}-\omega_{k_1})K_M(\omega_{s_1}-\omega_{k_2})K_M(\omega_{s_2}-\omega_{k_3})K_M(\omega_{s_2}-\omega_{k_4})$$
$$\mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1},J_{k_2,x_1},\overline{J}_{k_3,x_2})\mathrm{cum}(\overline{J}_{k_2,y_1},J_{k_3,y_2},\overline{J}_{k_4,x_2},J_{k_4,y_2})$$
$$=\frac{1}{T}\int\cdots\int W_M(\omega_{s_1}-\theta_1)W_M(\omega_{s_1}-\theta_2)W_M(\omega_{s_2}-\theta_2)W_M(\omega_{s_2}-\theta_4)$$
$$\frac{1}{T^2}f_4(\theta_1,-\theta_1,\theta_2)f_4(-\theta_2,\theta_2,-\theta_4)d\theta_1 d\theta_2 d\theta_4 d\omega_{s_1}d\omega_{s_2}$$
$$=\frac{1}{T^3}\int\cdots\int W_M(\omega_{s_1}-\theta_1)W_M(\omega_{s_1}-\theta_2)W_M(\omega_{s_2}-\theta_2)W_M(\omega_{s_2}-\theta_4)$$
$$f_4(\theta_1,-\theta_1,\theta_2)f_4(-\theta_2,\theta_2,-\theta_4)d\theta_1 d\theta_2 d\theta_4 d\omega_{s_1}d\omega_{s_2}$$
$$=O(\frac{M}{T^3})$$

From the above, we have $II_{3,4}=O(\frac{M}{T^3})$.

The one example of the cumulant terms in $II_{3,5}$ is

$$\mathrm{cum}(J_{k_1,x_1},J_{k_2,x_1})\mathrm{cum}(\overline{J}_{k_1,y_1},\overline{J}_{k_3,x_2})\mathrm{cum}(\overline{J}_{k_2,y_1},J_{k_3,y_2},\overline{J}_{k_4,x_2},J_{k_4,y_2})$$
$$=\begin{cases}\frac{1}{T}G(x_1,x_1,\omega_{k_1})G(y_1,x_2,-\omega_{k_1})f_4(\omega_{k_1},-\omega_{k_1},-\omega_{k_4}) & k_1+k_2=T,k_2=k_3\\ O(\frac{1}{T^3}) & k_1+k_2=T\text{ exclusively or }k_2=k_3\\ O(\frac{1}{T^4}) & otherwise.\end{cases}$$

With the similar argument above, we obtain $II_{3,5}<O(\frac{1}{T^3})$.

$$\frac{1}{T^2}\sum_{s_1,s_2}\sum_{\substack{k_1+k_2=T\\k_2=k_3,k_4}} K_M(\omega_{s_1}-\omega_{k_1})K_M(\omega_{s_1}-\omega_{k_2})K_M(\omega_{s_2}-\omega_{k_3})K_M(\omega_{s_2}-\omega_{k_4})$$
$$\mathrm{cum}(J_{k_1,x_1},\overline{J}_{k_1,y_1})\mathrm{cum}(\overline{J}_{k_3,x_2},J_{k_3,y_2})\mathrm{cum}(J_{k_2,x_1},\overline{J}_{k_2,y_1},\overline{J}_{k_4,x_2},J_{k_4,y_2})$$
$$=\frac{1}{T^3}\int\int\int\int W_M(\omega_{s_1}-\theta_1)W_M(\omega_{s_1}+\theta_1)W_M(\omega_{s_2}+\theta_1)W_M(\omega_{s_2}-\theta_4)$$
$$G(x_1,x_1,\theta_1)G(y_1,x_2,-\theta_1)f_4(\theta_1,-\theta_1,-\theta_4)d\theta_1 d\theta_4 d\omega_{s_1}d\omega_{s_2}<O(\frac{1}{T^3})$$

One term in $II_{3,6}$ is

$$\begin{aligned} & \text{cum}(J_{k_1,x_1}, \overline{J}_{k_2,y_1})\text{cum}(\overline{J}_{k_1,y_1}, \overline{J}_{k_3,x_2}, J_{k_3,x_2})\text{cum}(J_{k_2,x_1}, \overline{J}_{k_4,x_2}, J_{k_4,y_2}) \\ = & \begin{cases} \frac{1}{T}G(x_1, y_1, \omega_{k_1}) f_3(\omega_{k_1}, -\omega_{k_3}) f_3(\omega_{k_2}, -\omega_{k_4}) & k_1 = k_2 = T \\ O(\frac{1}{T^3}) & \textit{otherwise.} \end{cases} \end{aligned}$$

, and this leads to $II_{3,6} = O(\frac{1}{T^3})$.

There are 48 partitions in $II_{3,7}$. All these 48 partitions have 3 constraints on (k_1, k_2, k_3, k_4) to maintain their largest order. For example,

$$\begin{aligned} & \text{cum}(J_{k_1,x_1}, J_{k_2,x_1})\text{cum}(\overline{J}_{k_1,y_1}, \overline{J}_{k_3,x_2})\text{cum}(\overline{J}_{k_2,y_1}, \overline{J}_{k_4,x_2})\text{cum}(J_{k_3,y_2}, J_{k_4,y_2}) \\ = & \begin{cases} O(1) & k_1 + k_2 = T, k_2 = k_3, k_3 + k_4 = T \\ O(\frac{1}{T^2}) & \textit{otherwise} \end{cases} . \end{aligned}$$

Applying Lemma II.3 in Chpater II, we obtain the following.

$$\begin{aligned} & \frac{1}{T^2} \sum_{s_1,s_2} \sum_{k_1,k_2,k_3,k_4} K_M(\omega_{s_1} - \omega_{k_1}) K_M(\omega_{s_1} - \omega_{k_2}) K_M(\omega_{s_2} - \omega_{k_3}) K_M(\omega_{s_2} - \omega_{k_4}) \\ & \times \big\{ \text{cum}(J_{k_1,x_1}, J_{k_2,x_1})\text{cum}(\overline{J}_{k_1,y_1}, \overline{J}_{k_3,x_2})\text{cum}(\overline{J}_{k_2,y_1}, \overline{J}_{k_4,x_2})\text{cum}(J_{k_3,y_2}, J_{k_4,y_2}) \big\} \\ = & \frac{1}{T^3} \int\int\int W_M(\omega_{s_1} - \theta_1) W_M(\omega_{s_1} + \theta_1) W_M(\omega_{s_2} - \theta_1) W_M(\omega_{s_2} + \theta_1) \\ & \times G(x_1, x_1, \theta_1) G(y_1, x_2, -\theta_1) G(y_1, x_2, \theta_1) G(y_2, y_2, -\theta_1) d\theta_1 d\omega_{s_1} d\omega_{s_2} \\ = & O(\frac{M}{T^3}) \end{aligned}$$

We observe that the largest term in II_3 is $O(\frac{M}{T^3})$ which is smaller than the order of II_1 and II_2, $O(\frac{M}{T^2})$. This completes the proof. □

VITA

Jun Bum Lee was born in Seoul, South Korea. He received his Bachelor of Science degree in statistics from Seoul National University in 2004. He was admitted to the graduate program in the Department of Statistics at Texas A&M University in year 2006. He received his Master's degree and Ph.D. in May 2009 and May 2012, respectively. His address is Department of Statistics, Texas A&M University, TAMU 3143, College Station, TX 77843-3143.