

Jentsch, Carsten; Leucht, Anne; Meyer, Marco; Beering, Carina

Working Paper

Empirical characteristic functions-based estimation and distance correlation for locally stationary processes

Working Paper Series, No. 16-15

Provided in Cooperation with:

University of Mannheim, Department of Economics

Suggested Citation: Jentsch, Carsten; Leucht, Anne; Meyer, Marco; Beering, Carina (2016) : Empirical characteristic functions-based estimation and distance correlation for locally stationary processes, Working Paper Series, No. 16-15, University of Mannheim, Department of Economics, Mannheim, <https://nbn-resolving.de/urn:nbn:de:bsz:180-madoc-414388>

This Version is available at:

<https://hdl.handle.net/10419/149592>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

University of Mannheim / Department of Economics

Working Paper Series

***Empirical characteristic functions-based estimation and
distance correlation for locally stationary processes***

Carsten Jentsch Anne Leucht Marco Meyer Carina Beering

Working Paper 16-15

November 2016

EMPIRICAL CHARACTERISTIC FUNCTIONS-BASED ESTIMATION AND DISTANCE
CORRELATION FOR LOCALLY STATIONARY PROCESSES

Carsten Jentsch
Universität Mannheim
Abteilung Volkswirtschaftslehre
L 7, 3-5
D-68131 Mannheim
Germany
E-mail: cjentsch@mail.uni-mannheim.de

Anne Leucht, Marco Meyer, Carina Beering
Technische Universität Braunschweig
Institut für Mathematische Stochastik
Pockelsstraße 14
D-38106 Braunschweig
Germany
E-mail: a.leucht@tu-braunschweig.de, marco.meyer@tu-braunschweig.de,
c.beering@tu-braunschweig.de

Abstract

In this paper, we propose a kernel-type estimator for the local characteristic function of locally stationary processes. Under weak moment conditions, we prove joint asymptotic normality for local empirical characteristic functions. For time-varying linear processes, we establish a central limit theorem under the assumption of finite absolute first moments of the process. Additionally, we prove weak convergence of the local empirical characteristic process. We apply our asymptotic results to parameter estimation. Furthermore, by extending the notion of distance correlation of Székely, Rizzo and Bakirov (2007) to locally stationary processes, we are able to provide asymptotic theory for local empirical distance correlations. Finally, we provide a simulation study on minimum distance estimation for α -stable distributions and illustrate the pairwise dependence structure over time of log returns of German stock prices via local empirical distance correlations.

2010 Mathematics Subject Classification. Primary 62G05, 62G20; secondary 62M10, 62H20.

Keywords and Phrases. empirical characteristic function, local stationarity, time series, α -stable distributions, (local) distance correlation, minimum distance estimation, process convergence, asymptotic theory

Short title. Characteristic functions and distance correlation under local stationarity.

1. INTRODUCTION

The classical theory of characteristic functions (CFs) provides a powerful methodology to uniquely determine probability distributions of random variables and to characterize their properties. For an \mathbb{R}^d -valued random variable \underline{X} on a probability space (Ω, \mathcal{A}, P) , its CF $\varphi_{\underline{X}}$ is defined via the inverse Fourier transform of P , i.e.

$$\varphi_{\underline{X}}(\underline{s}) = E\left(e^{i\langle \underline{s}, \underline{X} \rangle}\right) = \int_{\Omega} e^{i\langle \underline{s}, \underline{X} \rangle} dP, \quad \underline{s} \in \mathbb{R}^d, \quad (1.1)$$

where $\langle \underline{a}, \underline{b} \rangle = \sum_{i=1}^k a_i b_i$ for $\underline{a}, \underline{b} \in \mathbb{R}^k$. Note that due to $|e^{i\langle \underline{s}, \underline{X} \rangle}| = 1$, the CF of a random variable \underline{X} does always exist. Hence, CFs are bounded and uniformly continuous. See e.g. Ushakov (1999) for general results on CFs. In particular, the definition of CFs does not require any moments of \underline{X} to be finite. Nevertheless, the existence of moments of \underline{X} translates nicely into smoothness of the CF in the origin. Several other features of the distribution $P^{\underline{X}}$ of \underline{X} indeed can be nicely described with the help of $\varphi_{\underline{X}}$. For example, stochastic independence of the components of a random vector $\underline{X} = (X_1, \dots, X_d)'$ is equivalent to the fact that the joint CF $\varphi_{\underline{X}}$ of \underline{X} equals the product of the marginal CFs φ_{X_i} of X_i , $i = 1, \dots, d$. That is, under independence, we have

$$\varphi_{\underline{X}}(\underline{s}) = E\left(e^{i\langle \underline{s}, \underline{X} \rangle}\right) = \prod_{i=1}^d E\left(e^{is_i X_i}\right) = \prod_{i=1}^d \varphi_{X_i}(s_i), \quad \underline{s} = (s_1, \dots, s_d)' \in \mathbb{R}^d. \quad (1.2)$$

In general, the CF of a random variable \underline{X} is a complex-valued function, but of particular interest is the case, where the CF is real-valued. A vanishing imaginary part of the CF translates to symmetry of the distribution of \underline{X} , i.e. $P^{\underline{X}} = P^{-\underline{X}}$. By exploiting such properties of CFs, several meaningful tools have been provided in statistical literature to investigate distributional features of data and to analyze independence or symmetry; see e.g. Bakirov, Rizzo, Székely (2006) or Leucht (2012) and references therein. In particular, parameter estimation methods based on empirical characteristic functions (ECFs) are very promising if the corresponding densities are complicated and higher order moments do not exist since classical maximum-likelihood estimation or method-of-moments can hardly be applied. CF based approaches have been considered, for example, by Press (1972), Höpfner and Rüschendorf (1999) or Xu and Knight (2010). Corresponding statistics are based on the ECF computed from the data. With data $\underline{X}_1, \dots, \underline{X}_T$ at hand, the ECF of \underline{X} is defined by

$$\widehat{\varphi}_{\underline{X}}(\underline{s}) = \frac{1}{T} \sum_{t=1}^T e^{i\langle \underline{s}, \underline{X}_t \rangle}, \quad \underline{s} \in \mathbb{R}^d. \quad (1.3)$$

The common approach to quantify dependence between random variables X and Y is to look at *linear* dependence measures such as covariances or correlations. Such measures might be misleading for certain statistical inference procedures as any dependence structure beyond the linear one will not be captured. That is, in general, i.e. without assuming a Gaussian distribution, covariances and correlations are suitable only to characterize uncorrelatedness, but not independence. Furthermore, these linear dependence measures require the second moments of random variables X and Y to exist, which may turn out to be too restrictive in practice. Moreover, at least fourth moments need to exist to derive central limit theorems (CLTs) for sample covariances or sample correlations.

In financial time series literature, the class of GARCH models is a prominent example

of time series processes, where uncorrelatedness and independence do not coincide *and* moment conditions are indeed restrictive; compare e.g. Francq and Zakoian (2010), Section 2.4.1 for the number of existing moments of a GARCH(1,1) process. In fact, when GARCH models are fitted to financial data, the resulting models usually tend to have only very few finite moments. Motivated by such observations, Szekely, Rizzo and Bakirov (2007) proposed an (unsigned) measure of dependence coined *distance correlation*; see also Szekely and Rizzo (2009). For two real-valued random vectors \underline{X} and \underline{Y} of dimensions p and q , respectively, their idea is to exploit the equivalence of stochastic independence and property (1.2) to consider a weighted L_2 -distance between the joint CF $\varphi_{(\underline{X}', \underline{Y}')'}((\underline{s}'_1, \underline{s}'_2)')$ and the product of the marginal CFs $\varphi_{\underline{X}}(\underline{s}_1)\varphi_{\underline{Y}}(\underline{s}_2)$. More precisely, they define the distance covariance $\mathcal{V}_{X,Y}$ by

$$\mathcal{V}_{X,Y}^2 = \int_{\mathbb{R}^{p+q}} \left| \varphi_{(\underline{X}', \underline{Y}')'}((\underline{s}'_1, \underline{s}'_2)') - \varphi_{\underline{X}}(\underline{s}_1)\varphi_{\underline{Y}}(\underline{s}_2) \right|^2 w(\underline{s}_1, \underline{s}_2) d\underline{s}_1 d\underline{s}_2, \quad (1.4)$$

where $w(\cdot, \cdot)$ is a (non-integrable) weight function defined by

$$w(\underline{s}_1, \underline{s}_2) = \left(c_p c_q |\underline{s}_1|_2^{1+p} |\underline{s}_2|_2^{1+q} \right)^{-1}, \quad \underline{s}_1 \in \mathbb{R}^p \setminus \{0\}, \quad \underline{s}_2 \in \mathbb{R}^q \setminus \{0\} \quad (1.5)$$

with $c_d = \pi^{(1+d)/2} / \Gamma((1+d)/2)$. The distance correlation $\mathcal{R}_{X,Y}$ is then easily deduced from (1.4) and defined as

$$\mathcal{R}_{X,Y}^2 = \begin{cases} \frac{\mathcal{V}_{X,Y}^2}{\sqrt{\mathcal{V}_X^2 \mathcal{V}_Y^2}}, & \mathcal{V}_X^2 \mathcal{V}_Y^2 > 0, \\ 0, & \mathcal{V}_X^2 \mathcal{V}_Y^2 = 0, \end{cases} \quad (1.6)$$

where $\mathcal{V}_X := \mathcal{V}_{X,X}$. This leads to an unsigned measure for dependence \mathcal{R} with $0 \leq \mathcal{R}_{X,Y} \leq 1$ and such that $\mathcal{R}_{X,Y} = 0$ is equivalent to \underline{X} and \underline{Y} being independent. Due to the specific choice of the weight function w in (1.5), it holds that \mathcal{R} is scale invariant; see Szekely, Rizzo and Bakirov (2007, p. 2771). In Szekely and Rizzo (2012), the authors show that the choice of w is essentially unique to maintain certain properties of \mathcal{V} that eventually lead e.g. to the desired scale-invariance of \mathcal{R} . Affinely invariant versions of distance correlations have been studied in Dueck, Edelmann, Gneiting and Richards (2014). With data $((\underline{X}'_1, \underline{Y}'_1)', \dots, (\underline{X}'_T, \underline{Y}'_T)')$ at hand, sample versions of \mathcal{V} and \mathcal{R} are defined by just replacing the CFs φ by ECFs $\hat{\varphi}$ in (1.4) and (1.6), respectively. Under mild assumptions of existing first absolute moments in an independent, identically distributed (i.i.d.) setup, Szekely, Rizzo and Bakirov (2007) show almost sure convergence of these empirical quantities to their theoretical counterparts \mathcal{V} and \mathcal{R} , respectively. Furthermore, they provide some limiting results that allow to construct tests for pairwise independence of X and Y . Zhou (2012) employed the concept of distance correlation to measure nonlinear dependence in time series by considering auto-distance correlation. In a recent paper, Davis, Matsui, Mikosch, and Wan (2016) provide general asymptotic theory for distance correlations of stationary time series.

Whereas most of the classical theory on CFs was originally developed for the case of i.i.d. random variables, many results have been established more recently also for stationary time series processes; see e.g. Leucht (2012), Knight and Yu (2002) or Yu (2004) for a review. However, in practice, it is rather unlikely to observe real time series data from a perfectly (strictly) stationary data generating processes (DGP), that is, whose entire dependence structure is invariant as time evolves. Also such an idealized stationary model may not be suitable at all to describe the data at least approximately. To approximate non-stationary processes locally (on segments) by stationary processes goes back to Priestley (1965), who proposed time series with evolutionary spectra representation. In a series of

papers, Dahlhaus (1996a,b, 1997) proposed the notion of local stationarity and provided an asymptotic (in-fill) framework that allows for meaningful and rigorous theory; see Dahlhaus (2012) for a recent review of the state-of-the-art. Locally stationary processes are indeed globally non-stationary, but locally they can be sufficiently well approximated by stationary processes. By rescaling the data to the unit interval and executing in-fill asymptotics, it is assured that locally more and more data is observed, which allows e.g. consistent estimation of model parameters. One common way in the literature is to define locally stationary processes $(\underline{X}_{t,T})_{t=1}^T$, $T \in \mathbb{N}$, via a linear MA(∞) representation with time-varying coefficients, that is,

$$\underline{X}_{t,T} = \underline{\mu} \left(\frac{t}{T} \right) + \sum_{j=-\infty}^{\infty} A_{t,T}(j) \varepsilon_{t-j}, \quad t = 1, \dots, T, \quad (1.7)$$

where $(\varepsilon_t)_{t \in \mathbb{Z}}$ is some d -dimensional noise process and $(A_{t,T}(j))_{t=1}^T$, $j \in \mathbb{Z}$, are coefficient matrices smoothly varying over time. We make use of the concept (1.7) and define d -variate linear locally stationary processes in Section 2, where we provide also detailed assumptions used throughout this paper.

So far, many results for locally stationary processes are based on spectral methods and therefore require the DGP to have at least finite second moments. In particular, this assumption might be violated for financial time series. In view of the far reaching statistical application possibilities of CFs, the main purpose of this paper is to extend the existing theory for strictly stationary processes to the case of locally stationary processes in the sense of (1.7). To extend the applicability of statistical procedures (as e.g. distance correlations or minimum distance estimation) established for the stationary case to this more general setup, we define the CF of locally stationary processes and propose a suitable kernel-type estimator generalizing (1.1) and (1.3), respectively, to the locally stationary case. In a second step we investigate asymptotic properties of the newly proposed estimator.

The paper is organized as follows. In Section 2, we present the basic notation of locally stationary processes, which allows us to define corresponding versions of the CF and the ECF. We provide regularity conditions imposed on the locally stationary process and deduce from these some important preliminary results that will be used frequently throughout this paper. Asymptotic theory for the kernel-based local ECF is provided in Section 3, where we first provide explicit formulas for the bias and for covariances of local ECFs. Based on these results, we are then able to prove a CLT as well as process convergence for local ECFs. In Section 4.1, we apply our results to derive theory for time-varying parameter estimation in locally stationary models, and in Section 4.2, we define local (auto-)distance correlations and prove consistency results. In Section 5, we provide a simulation study on two estimation procedures for α -stable distributions and illustrate the pairwise dependence structure over time of log returns of German stock prices via local empirical distance correlations. All proofs are deferred to Section 6.

Finally, let us fix some notation. Throughout the paper, we underline vector-valued quantities. Moreover, \bar{Z} denotes the (entry-wise) complex conjugate of a matrix Z and X' indicates the transpose of a vector or matrix X . Let $|\cdot|_1$ denote the "max column sum matrix norm", i.e. $|M|_1 = \max_{1 \leq j \leq r} \sum_{i=1}^n |m^{(i,j)}|$ for some $(n \times r)$ matrix $M = (m^{(i,j)})$. Further, let $\|M\|_1 := E(|M|_1)$ be the matrix-valued L^1 -norm. For n -dimensional vectors $v = (v_1, \dots, v_n)'$, its p -norm $|\cdot|_p$ is defined as usual as $|\cdot|_p := (\sum_{j=1}^n |v_j|^p)^{1/p}$. Note that

the matrix norm $|\cdot|_1$ is submultiplicative and it coincides with the usual 1-norm for n -dimensional vectors, i.e. for $r = 1$. For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ be the largest integer smaller as or equal to $x \in \mathbb{R}$ and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. Let $\Re x$ and $\Im x$ denote the real and imaginary parts of x . With \xrightarrow{d} , we denote convergence in distribution and with \xrightarrow{P} stochastic convergence.

2. EMPIRICAL CHARACTERISTIC FUNCTIONS FOR LOCALLY STATIONARY PROCESSES

2.1. Setup and preliminaries. Throughout this paper, let $(\underline{X}_{t,T})_{t=1}^T$ be a d -variate time series process that has a time-varying MA(∞) representation. That is, we have

$$\underline{X}_{t,T} = \underline{\mu}\left(\frac{t}{T}\right) + \sum_{j=-\infty}^{\infty} A_{t,T}(j) \underline{\varepsilon}_{t-j}, \quad (2.1)$$

where $\underline{\mu}(\cdot) = (\mu_1(\cdot), \dots, \mu_d(\cdot))'$ denotes a d -dimensional time-varying mean function, $(A_{t,T}(j))_{j \in \mathbb{Z}}$ are $(d \times d)$ coefficient matrices and $(\underline{\varepsilon}_t)_{t \in \mathbb{Z}}$ is a sequence of i.i.d. centered random vectors.

Of course the components of the coefficient matrices $A_{t,T}(j)$ have to decay sufficiently fast as $|j| \rightarrow \infty$ to assure existence of the process. In order to establish a meaningful statistical methodology, we require also the coefficients to vary slowly over time. More precisely, we make the following assumptions on the process $(\underline{X}_{t,T})_{t=1}^T$.

Assumption 2.1 (The process).

The process $(\underline{X}_{t,T})_{t=1}^T$, $T \in \mathbb{N}$, is of the form (2.1), where

- (i) $(\underline{\varepsilon}_t)_{t \in \mathbb{Z}}$ is i.i.d., mean zero and satisfies $E|\underline{\varepsilon}_0|_1 < \infty$,
- (ii) for some $k \in \mathbb{N}_0$ it holds:
 - (ii.1) $E|\underline{\varepsilon}_0|_k < \infty$.
 - (ii.2) There exists a real-valued, deterministic sequence $(l(j))_{j \in \mathbb{Z}}$ with $\sum_{j \in \mathbb{Z}} j^2 l^{-1}(j) < \infty$ and a constant $B < \infty$ such that

$$\sup_{t,T} |A_{t,T}(j)|_1 \leq \frac{B}{l(j)}. \quad (2.2)$$

Further, for each $j \in \mathbb{Z}$ there exists an entry-wise k -times continuously differentiable function $A(\cdot, j): [0, 1] \rightarrow \mathbb{R}^{d \times d}$ such that for all $p, q = 1, \dots, d$ and $s = 0, 1, \dots, k$, we have

$$\sup_u \left| \frac{\partial^s a^{(p,q)}(u, j)}{\partial u^s} \right| \leq \frac{B}{l(j)} \quad \text{and} \quad \sup_{t,T} T \left| A_{t,T}(j) - A\left(\frac{t}{T}, j\right) \right|_1 \leq \frac{B}{l(j)}, \quad (2.3)$$

where $A(u, j) = (a^{(p,q)}(u, j))_{p,q=1,\dots,d}$ and $A^{(s)}(u, j) = \frac{\partial^s}{\partial u^s} A(u, j) = \left(\frac{\partial^s}{\partial u^s} a^{(p,q)}(u, j) \right)_{p,q=1,\dots,d}$.

- (ii.3) Each component of $\underline{\mu}$ is k -times continuously differentiable.

The construction with $A_{t,T}(j)$ and $A(t/T, j)$ appears to be unnecessarily complicated. However, $A(t/T, j)$ is needed for rescaling and to be able to impose smoothness conditions, while $A_{t,T}(j)$ makes the class of processes rich enough; compare Dahlhaus (2012). For instance, a broad class of time-varying autoregressive (tvAR) processes of the form

$$X_{t,T} = r_{t,T}(1) X_{t-1,T} + \dots + r_{t,T}(p) X_{t-p,T} + \varepsilon_t, \quad t = 1, \dots, T,$$

is included. Cardinali and Nason (2010) coined the term *close pair* for $(A_{t,T}(j), A(t/T, j))$. Moreover, note that continuity and differentiability conditions on $[0, 1]$ in Assumption 2.1 shall be understood in the one-sided sense at the boundary points.

As an immediate consequence of Assumption 2.1 we get from (2.3) by straightforward calculations

$$\sup_{u \in [0,1]} |A(u, j)|_1 \leq \frac{dB}{l(j)}. \quad (2.4)$$

As discussed in Section 4 of Dahlhaus (2012), Assumption 2.1 forms a classical set of conditions for statistical inference based on local estimators for locally stationary processes. The basic idea is that in some neighborhood of a fixed time point $u = t/T$, the process $(\underline{X}_{t,T})_{t=1}^T$ can be locally approximated by (strictly) stationary processes $(\tilde{X}_t(u))_{t \in \mathbb{Z}}$ given by

$$\tilde{X}_t(u) = \underline{\mu}(u) + \sum_{j=-\infty}^{\infty} A(u, j) \varepsilon_{t-j}. \quad (2.5)$$

Note that this equation indeed has a strictly stationary solution under Assumption 2.1 for each fixed u . The following lemma provides results which will be used several times in the remainder of this paper. In particular, the uniform decay of the covariances of the CFs of $(\tilde{X}_t(u))_{t \in \mathbb{Z}}$ as well as the approximation of $\underline{X}_{t,T}$ by $\tilde{X}_t(t/T)$ are quantified.

Lemma 2.1 (Preliminary consequences of Assumption 2.1).

Suppose that Assumption 2.1 holds true for $k = 0$.

(i) Then, for all $(u_1, \underline{s}_1), (u_2, \underline{s}_2) \in [0, 1] \times \mathbb{R}^d$ and $h \in \mathbb{Z}$, we have

$$\begin{aligned} & \sup_{u_1, u_2 \in [0,1]} \left| \text{Cov}(\exp(i \langle \underline{s}_1, \tilde{X}_h(u_1) \rangle), \exp(i \langle \underline{s}_2, \tilde{X}_0(u_2) \rangle)) \right| \\ & \leq 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\varepsilon_0\|_1 \sum_{|j| > \lfloor |h/2| \rfloor} \frac{B}{l(j)}. \end{aligned}$$

(ii) It holds

$$\sup_{1 \leq t \leq T} \left\| \underline{X}_{t,T} - \tilde{X}_t(t/T) \right\|_1 \leq \frac{1}{T} \cdot \|\varepsilon_0\|_1 \sum_{j=-\infty}^{\infty} \frac{B}{l(j)} = \mathcal{O}(T^{-1}).$$

(iii) Suppose now that Assumption 2.1 holds true for $k = 1$. Then, for all $u, w \in [0, 1]$ and all $h \in \mathbb{Z}$, we have

$$\left\| \tilde{X}_h(u) - \tilde{X}_h(w) \right\|_1 \leq d \left(L + \|\varepsilon_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \frac{B}{l(j)} \right) \cdot |u - w|$$

for some $L < \infty$.

Remark 1.

There are more general definitions of locally stationary processes as e.g. proposed in Vogt (2012) that go beyond the linear representation in (2.1). However, we stick to the framework of linear locally stationary processes of (2.1) to avoid causality of the process as well as high-level assumptions. Nevertheless, we expect results similar to those displayed in Lemma 2.1 to hold e.g. under physical dependence-type conditions (as used e.g. in Zhou (2014)) or other weak dependence assumptions as these allow for similar proof techniques.

2.2. Local CF and ECF. We denote the CFs of $\tilde{X}_1(u)$ and $\underline{X}_{t,T}$ by $\varphi(u, \cdot)$ and $\varphi_{t,T}(\cdot)$, respectively, i.e. we define the *close pair of characteristic functions* by

$$\varphi(t/T, \underline{s}) := E \left(e^{i\langle \underline{s}, \tilde{X}_1(t/T) \rangle} \right) \quad \text{and} \quad \varphi_{t,T}(\underline{s}) := E \left(e^{i\langle \underline{s}, \underline{X}_{t,T} \rangle} \right) \quad (2.6)$$

for all $\underline{s} \in \mathbb{R}^d$. For fixed u and fixed pair (t, T) , $\varphi(u, \cdot)$ and $\varphi_{t,T}(\cdot)$ can be interpreted as *local characteristic functions* of the corresponding time series $(\tilde{X}_t(u))_t$ and $(\underline{X}_{t,T})_t$ at rescaled time u and t/T , respectively. The next lemma shows that $\varphi_{t,T}(\underline{s})$ can be approximated (uniformly) well by $\varphi(t/T, \underline{s})$.

Lemma 2.2 (Uniform approximation and derivatives of local CFs).

- (i) *Suppose Assumption 2.1 holds with $k = 0$. Then, for any $S \in (0, \infty)$ there exists a constant $C < \infty$ such that*

$$\sup_{\underline{s} \in [-S, S]^d, 1 \leq t \leq T} \left| \varphi_{t,T}(\underline{s}) - \varphi \left(\frac{t}{T}, \underline{s} \right) \right| \leq \frac{C}{T}.$$

- (ii) *Let Assumption 2.1 hold for some $k \geq 1$. Then, φ has continuous derivatives up to order k with respect to (w.r.t.) its first argument, that is, for all $u \in [0, 1]$,*

$$\varphi^{(j)}(u, \underline{s}) = \frac{\partial^j}{\partial u^j} \varphi(u, \underline{s}), \quad j = 0, \dots, k,$$

exist.

If we had observations $\tilde{X}_1(u), \dots, \tilde{X}_T(u)$ of the process $(\tilde{X}_t(u))_{t \in \mathbb{Z}}$ at hand, a natural estimator of $\varphi(u, \underline{s})$ would be the usual empirical CF from (1.3), that is,

$$\hat{\varphi}(u, \underline{s}) = \frac{1}{T} \sum_{t=1}^T e^{i\langle \underline{s}, \tilde{X}_t(u) \rangle},$$

but the process $(\tilde{X}_t(u))_{t \in \mathbb{Z}}$ cannot be observed. Of course, the CF of the variables that form the locally stationary process $(\underline{X}_{t,T})_{t=1}^T$, $T \in \mathbb{N}$, varies smoothly by Assumption 2.1 with the indices t and T . Therefore, the idea is to approximate the CF of $\underline{X}_{[uT], T}$ by its local sample analogue, that is,

$$\hat{\varphi}(u, \underline{s}) = \frac{1}{T} \sum_{t=1}^T K_b \left(\frac{t}{T} - u \right) e^{i\langle \underline{s}, \underline{X}_{t,T} \rangle}, \quad (2.7)$$

where $K_b(\cdot) = K(\cdot/b)/b$ for some kernel function K and b is a bandwidth parameter.

To derive asymptotic theory for the local ECF defined in (2.7), we make the following assumptions concerning the kernel K and the bandwidth b .

Assumption 2.2 (The kernel).

- (i) *The function $K: \mathbb{R} \rightarrow [0, \infty)$ is symmetric, non-negative, and Lipschitz continuous. Moreover, $\int K(u) du = 1$, and K has compact support $[-1, 1]$.*
(ii) *The sequence of bandwidths $b = b_T$ is non-negative and satisfies $b \rightarrow 0$ and $Tb^2 \rightarrow \infty$ as $T \rightarrow \infty$.*

This assumption is standard in nonparametric statistics. All asymptotic results below remain valid under weaker assumptions on K . For the sake of notational simplicity, we stick to Assumption 2.2. Note that Assumption 2.2 implies that it holds for all $m \geq 0$

$$\frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left|\frac{t}{T} - u\right|^m = \frac{1}{bT} \sum_{t=1}^T K\left(\frac{t/T - u}{b}\right) \left|\frac{t}{T} - u\right|^m = \mathcal{O}(b^m), \quad (2.8)$$

since all summands with $|t/T - u| > b$ vanish and since there are at most $\lfloor 2bT \rfloor + 1$ summands with $|t/T - u| \leq b$. Assertion (2.8) will be used multiple times in the remainder of this work.

3. ASYMPTOTIC THEORY FOR THE LOCAL ECF

In this section, we derive asymptotic results for the local ECF. First, in Section 3.1, we derive explicit expressions for bias, variance and covariances before we provide CLTs and process convergence for the local ECF in Sections 3.2 and 3.3, respectively.

3.1. Bias, Variance and Covariances. Under different sets of assumptions, we get the following representations for bias, variance and covariances of our kernel-type local ECF $\widehat{\varphi}(u, \underline{s})$ as defined in (2.7).

Lemma 3.1 (Bias, variance and covariances).

Suppose Assumption 2.1 holds with some $k \geq 1$ to be specified further below and Assumption 2.2 is satisfied. Let $(u, \underline{s}), (u_1, \underline{s}_1), (u_2, \underline{s}_2) \in (0, 1) \times \mathbb{R}^d$.

(i) *If $k \geq 1$, we have*

$$\sup_{\underline{s} \in [-S, S]^d} |E(\widehat{\varphi}(u, \underline{s})) - \varphi(u, \underline{s})| = \mathcal{O}(b) + \mathcal{O}((bT)^{-1}). \quad (3.1)$$

(ii) *If $k \geq 3$, we have*

$$E(\widehat{\varphi}(u, \underline{s})) - \varphi(u, \underline{s}) = b^2 \cdot \beta(u, \underline{s}) + \mathcal{O}((bT)^{-1}) + \mathcal{O}(b^3), \quad (3.2)$$

where $\beta(u, \underline{s}) = (1/2) \int_{-1}^1 K(x) x^2 dx \cdot \varphi^{(2)}(u, \underline{s})$.

(iii) *If $k \geq 1$, we have*

$$\text{Cov}(\widehat{\varphi}(u_1, \underline{s}_1), \widehat{\varphi}(u_2, \underline{s}_2)) = (bT)^{-1} V((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o((bT)^{-1}), \quad (3.3)$$

where

$$\begin{aligned} & V((u_1, \underline{s}_1), (u_2, \underline{s}_2)) \\ &= \mathbf{1}_{\{u_1=u_2\}} \cdot \int_{-1}^1 K^2(x) dx \cdot \sum_{h=-\infty}^{\infty} \text{Cov}(\exp(i \langle \underline{s}_1, \widetilde{\mathbf{X}}_h(u_1) \rangle), \exp(i \langle \underline{s}_2, \widetilde{\mathbf{X}}_0(u_2) \rangle)). \end{aligned} \quad (3.4)$$

If $\widehat{\varphi}(u_1, \underline{s}_1)$ and $\widehat{\varphi}(u_2, \underline{s}_2)$ are replaced by their real or imaginary parts, a result analogue to (3.3) remains true, as long as $\exp(i \langle \underline{s}_1, \widetilde{\mathbf{X}}_h(u_1) \rangle)$ and $\exp(i \langle \underline{s}_2, \widetilde{\mathbf{X}}_0(u_2) \rangle)$ are also replaced in (3.4) in the same way by their real or imaginary parts, respectively. See also the notation below in (3.5).

Note, that $V((u_1, \underline{s}_1), (u_2, \underline{s}_2))$ is well-defined and finite for all $(u_1, \underline{s}_1), (u_2, \underline{s}_2) \in [0, 1] \times \mathbb{R}^d$ by Lemma 2.1(i) since

$$\sum_{h=-\infty}^{\infty} \text{Cov}(\exp(i \langle \underline{s}_1, \widetilde{\mathbf{X}}_h(u_1) \rangle), \exp(i \langle \underline{s}_2, \widetilde{\mathbf{X}}_0(u_2) \rangle))$$

$$\begin{aligned}
&\leq \sum_{h=-\infty}^{\infty} 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \sum_{|j| > \lfloor |h|/2 \rfloor} \frac{B}{l(j)} \\
&\leq 8dB(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \sum_{|j| > 0} |j| \frac{1}{l(j)} < \infty.
\end{aligned}$$

In the following section, the results above will be invoked to establish a CLT for the real and imaginary parts of the estimator $\widehat{\varphi}(u, \underline{s})$. In particular, we will use the following notation:

$$\begin{aligned}
\text{Cov}(\Re \widehat{\varphi}(u_1, \underline{s}_1), \Re \widehat{\varphi}(u_2, \underline{s}_2)) &= (bT)^{-1} V_{\Re \Re}((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o((bT)^{-1}), \\
\text{Cov}(\Im \widehat{\varphi}(u_1, \underline{s}_1), \Im \widehat{\varphi}(u_2, \underline{s}_2)) &= (bT)^{-1} V_{\Im \Im}((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o((bT)^{-1}), \\
\text{Cov}(\Re \widehat{\varphi}(u_1, \underline{s}_1), \Im \widehat{\varphi}(u_2, \underline{s}_2)) &= (bT)^{-1} V_{\Re \Im}((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o((bT)^{-1}), \\
\text{Cov}(\Im \widehat{\varphi}(u_1, \underline{s}_1), \Re \widehat{\varphi}(u_2, \underline{s}_2)) &= (bT)^{-1} V_{\Im \Re}((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o((bT)^{-1}),
\end{aligned} \tag{3.5}$$

where $V_{\Re \Re}((u_1, \underline{s}_1), (u_2, \underline{s}_2))$ etc. are defined analogue to $V((u_1, \underline{s}_1), (u_2, \underline{s}_2))$, but with $\exp(i \cdot)$ replaced by $\cos(\cdot)$ and $\sin(\cdot)$ according to real and imaginary parts used, respectively.

3.2. Central limit theorem. The following theorem states results for the joint limiting distribution of local ECF of the marginal distribution of $(\underline{X}_{t,T})_{t=1}^T$, $T \in \mathbb{N}$.

Theorem 3.1 (CLT).

Suppose Assumption 2.1 holds with some $k \geq 1$ to be specified further below and Assumption 2.2 is satisfied. Let $J \in \mathbb{N}$ and $(u_j, \underline{s}_j) \in (0, 1) \times \mathbb{R}^d$, $j = 1, \dots, J$.

(i) If $k \geq 1$, we have

$$\sqrt{bT} \left(\left(\begin{array}{c} \Re \left(\widehat{\varphi}(u_j, \underline{s}_j) - E(\widehat{\varphi}(u_j, \underline{s}_j)) \right) \\ \Im \left(\widehat{\varphi}(u_j, \underline{s}_j) - E(\widehat{\varphi}(u_j, \underline{s}_j)) \right) \end{array} \right), j = 1, \dots, J \right) \xrightarrow{d} \mathcal{N}(\underline{0}, \mathbf{V}), \tag{3.6}$$

where $\mathbf{V} = (\mathbf{V}(m, n))_{m, n=1, \dots, J}$ is a $(2J \times 2J)$ block covariance matrix with

$$\mathbf{V}(m, n) = \begin{pmatrix} V_{\Re \Re}((u_m, \underline{s}_m), (u_n, \underline{s}_n)) & V_{\Re \Im}((u_m, \underline{s}_m), (u_n, \underline{s}_n)) \\ V_{\Im \Re}((u_m, \underline{s}_m), (u_n, \underline{s}_n)) & V_{\Im \Im}((u_m, \underline{s}_m), (u_n, \underline{s}_n)) \end{pmatrix}$$

(ii) If $k \geq 1$ and $b^3 T \rightarrow 0$ as $T \rightarrow \infty$ holds, we have

$$\sqrt{bT} \left(\left(\begin{array}{c} \Re \left(\widehat{\varphi}(u_j, \underline{s}_j) - \varphi(u_j, \underline{s}_j) \right) \\ \Im \left(\widehat{\varphi}(u_j, \underline{s}_j) - \varphi(u_j, \underline{s}_j) \right) \end{array} \right), j = 1, \dots, J \right) \xrightarrow{d} \mathcal{N}(\underline{0}, \mathbf{V}), \tag{3.7}$$

(iii) If $k \geq 3$ and $b^5 T \rightarrow C^2$ as $T \rightarrow \infty$ for some $C \in [0, +\infty)$, we have

$$\sqrt{bT} \left(\left(\begin{array}{c} \Re \left(\widehat{\varphi}(u_j, \underline{s}_j) - \varphi(u_j, \underline{s}_j) \right) \\ \Im \left(\widehat{\varphi}(u_j, \underline{s}_j) - \varphi(u_j, \underline{s}_j) \right) \end{array} \right), j = 1, \dots, J \right) \xrightarrow{d} \mathcal{N}(C\underline{\beta}, \mathbf{V}), \tag{3.8}$$

where the bias vector $\underline{\beta} = (\beta(u_1, \underline{s}_1), \dots, \beta(u_J, \underline{s}_J))'$ can be obtained from the results of Lemma 3.1(ii).

The CLT is not directly applicable if one aims to estimate the joint local CF of lagged variables $\underline{X}_{t,T}$ and $\underline{X}_{t+h,T}$, say. However, the latter result can be deduced as a corollary. We provide a counterpart to (ii) of Theorem 3.1 in the next section, where we additionally show weak convergence of the corresponding process on any compact hyperrectangle $[-S, S]^d$ (holding u fixed); see Corollary 3.1 below.

3.3. Weak convergence of the local ECF process. Starting from Csörgő (1981), there exists a wide range of articles considering the limit behavior of the (multivariate) empirical characteristic process in the case of a stationary DGP. As indicated there, one can in general not expect convergence on the whole real line but only on compact intervals that may increase with a certain rate as the sample size grows. We intend to generalize these results to locally stationary DGPs since they are very useful for statistical applications; see Section 4. While Csörgő (1981) provides a necessary and sufficient condition for weak convergence if the observations are i.i.d., we give a sufficient condition only. This is due to the fact that we allow for dependent observations.

We will verify distributional convergence of the real and imaginary part of the *local ECF process* defined as

$$\left\{ \sqrt{bT} (\hat{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) \right\}_{\underline{s} \in [-S, S]^d}$$

for any fixed $S \in (0, \infty)$ and any fixed $u \in (0, 1)$. Obviously, convergence of the finite dimensional distributions (fidis) is an immediate consequence of Theorem 3.1(ii) if $b^3 T \rightarrow 0$ as $T \rightarrow \infty$. It remains to prove tightness. As

$$\sup_{\underline{s} \in [-S, S]^d} \sqrt{bT} (E\hat{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) = \mathcal{O}(\sqrt{b^3 T}) + \mathcal{O}((bT)^{-1/2})$$

tends to zero if $b^3 T \rightarrow 0$ due to Lemma 3.1(i), for such choices of bandwidths, it remains to show tightness for the real and the imaginary part of

$$\left\{ \sqrt{bT} (\hat{\varphi}(u, \underline{s}) - E\hat{\varphi}(u, \underline{s})) \right\}_{\underline{s} \in [-S, S]^d}.$$

It turns out that the situation can be simplified even more, since this quantity can be approximated using the stationary accompanying process.

Lemma 3.2.

Suppose that Assumption 2.1 holds true for $k = 1$, Assumption 2.2 holds and that $b^3 T \rightarrow 0$. Then, as $T \rightarrow \infty$,

$$E \left(\sup_{\underline{s} \in [-S, S]^d} \sqrt{bT} |\hat{\varphi}(u, \underline{s}) - E\hat{\varphi}(u, \underline{s}) - [\tilde{\varphi}(u, \underline{s}) - E\tilde{\varphi}(u, \underline{s})]| \right) \rightarrow 0,$$

where $\tilde{\varphi}(u, \underline{s}) = \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \exp(i\langle \underline{s}, \tilde{X}_t(u) \rangle)$.

In comparison to the results in Sections 3.1 and 3.2, we obtain tightness of the real and imaginary part of the stationary approximating local ECF process under a slightly more restrictive moment condition of $\|\tilde{\varepsilon}_0\|_{1+\delta} < \infty$ for some $\delta > 0$ if the coefficients of the stationary accompanying process $\tilde{X}_t(u)$ decay sufficiently fast. More precisely, we get the following result.

Lemma 3.3 (Tightness).

Suppose that Assumption 2.1 holds true for $k = 1$, Assumption 2.2 holds, and that $b^3 T \rightarrow 0$. Further, assume that $\|\tilde{\varepsilon}_0\|_{1+\delta} < \infty$ for some $\delta \in (0, 1)$ and that $\sum_{j=1}^{\infty} j^m / l(j) < \infty$ for some

$m > 1 + 4/(3\delta)$. Then, for all $\lambda > 0$, we have for any $S \in (0, \infty)$

$$\begin{aligned} & \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left\{ \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \|\underline{s}_1 - \underline{s}_2\|_1 < r}} \left| \Re \left(\sqrt{bT} (\tilde{\varphi}(u, \underline{s}_1) - E\tilde{\varphi}(u, \underline{s}_1)) - \left[\sqrt{bT} (\tilde{\varphi}(u, \underline{s}_2) - E\tilde{\varphi}(u, \underline{s}_2)) \right] \right) \right| > \lambda \right\} = 0, \\ & \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left\{ \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \|\underline{s}_1 - \underline{s}_2\|_1 < r}} \left| \Im \left(\sqrt{bT} (\tilde{\varphi}(u, \underline{s}_1) - E\tilde{\varphi}(u, \underline{s}_1)) - \left[\sqrt{bT} (\tilde{\varphi}(u, \underline{s}_2) - E\tilde{\varphi}(u, \underline{s}_2)) \right] \right) \right| > \lambda \right\} = 0. \end{aligned} \quad (3.9)$$

From the previous lemmas we can conclude convergence of the local ECF process.

Theorem 3.2 (Process convergence).

Suppose that Assumption 2.1 holds true for $k = 1$, Assumption 2.2 holds, and that $b^3T \rightarrow 0$. Further, assume that $\|\underline{\varepsilon}_0\|_{1+\delta} < \infty$ for some $\delta > 0$ and that $\sum_{j=1}^{\infty} j^m/l(j) < \infty$ for some $m > 1 + 4/(3\delta)$. Then, it holds for all $u \in (0, 1)$ and all $S \in (0, \infty)$,

$$\left\{ \sqrt{bT} \begin{pmatrix} \Re(\tilde{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) \\ \Im(\tilde{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) \end{pmatrix} \right\}_{\underline{s} \in [-S, S]^d} \xrightarrow{d} \{Z(u, \underline{s})\}_{\underline{s} \in [-S, S]^d}. \quad (3.10)$$

where $\{Z(u, \underline{s})\}_{\underline{s} \in [-S, S]^d}$ is a centered Gaussian process with continuous sample paths, zero mean and covariance function \mathbf{V} , given by

$$\mathbf{V}(u, \underline{s}) = \begin{pmatrix} V_{\Re, \Re}(u, \underline{s}) & V_{\Re, \Im}(u, \underline{s}) \\ V_{\Re, \Im}(u, \underline{s}) & V_{\Im, \Im}(u, \underline{s}) \end{pmatrix}.$$

Here, convergence holds w.r.t. the uniform norm.

Remark 2.

One can obtain functional CLTs generalizing Theorem 3.1(i) and (iii) in a similar manner. Since most statistical applications require results of the type (3.10), we focus on this result only.

The local ECF $\hat{\varphi}(t/T, \cdot)$ as proposed in (2.7) is designed to analyze the *marginal* distribution of the process $(\underline{X}_{t,T})_{t=1}^T$ by estimating the corresponding local CF $\varphi(t/T, \cdot)$. In particular, it is not perfectly suitable to analyze the *joint* distribution of e.g. $((\underline{X}'_{t,T}, \underline{X}'_{t+h,T})')_{t=1}^T$. First, note that the bivariate process does not directly satisfy Assumption 2.1 even though $(\underline{X}_{t,T})_{t=1}^T$ does. Still, under some regularity conditions on $\underline{\mu}$ and $(A_{t,T}(j))_{j \in \mathbb{Z}}$, it holds

$$\underline{\mathbf{X}}_{t,h,T} := \begin{pmatrix} \underline{X}_{t,T} \\ \underline{X}_{t+h,T} \end{pmatrix} = \underline{\bar{\mu}} \left(\frac{t}{T} \right) + \sum_{j=-\infty}^{\infty} \bar{A}_{t,T}(j) \bar{\underline{\varepsilon}}_{t-j} + \mathcal{O}_P \left(\frac{1}{T} \right)$$

with

$$\underline{\bar{\mu}} \left(\frac{t}{T} \right) = \begin{pmatrix} \underline{\mu}(t/T) \\ \underline{\mu}(t/T) \end{pmatrix}, \quad \bar{A}_{t,T}(j) = \begin{pmatrix} A_{t,T}(j) & 0 \\ 0 & A_{t,T}(j+h) \end{pmatrix}, \quad \text{and} \quad \bar{\underline{\varepsilon}}_t = \begin{pmatrix} \underline{\varepsilon}_t \\ \underline{\varepsilon}_t \end{pmatrix}.$$

Second, if one would naively apply our estimator $\hat{\varphi}_h$ to $(\underline{\mathbf{X}}_{t,h,T})_{t=1}^{T-h}$, we would give more weight to the observations with index close to t than to observations with index close to $t+h$ due to the properties of the kernel K . To obtain better finite sample results, we propose to use

$$\hat{\varphi}_h(u, \underline{s}) = \frac{1}{T} \sum_{t=1}^T K_b \left(\frac{t+h/2}{T} - u \right) e^{i\langle \underline{s}, \underline{\mathbf{X}}_{t,h,T} \rangle}, \quad \underline{s} \in \mathbb{R}^{2d}, \quad (3.11)$$

instead, where we added the term $h/2$ in the kernel above to estimate

$$\varphi_h(u, \underline{s}) = Ee^{i\langle \underline{s}, \tilde{\mathbf{X}}_{0,h}(u) \rangle} = Ee^{i\langle \underline{s}, (\tilde{X}_0(u)', \tilde{X}_h(u)')' \rangle}$$

from those $\underline{X}_{t,T}$ and $\underline{X}_{t+h,T}$, where u is near to the middle of t and $t+h$, which is $t+h/2$. In general, $\varphi_h(u, \underline{s})$ can be interpreted as a local lagged CF. Although Theorem 3.2 can not be applied directly to $\hat{\varphi}_h$, we obtain a similar result for the modified estimator with the same arguments as before.

Corollary 3.1 (A functional CLT for lagged variables).

Suppose that the assumptions of Theorem 3.2 are satisfied. Then, for all fixed $h_0 \in \mathbb{Z}$,

$$\left\{ \sqrt{bT} \begin{pmatrix} \Re(\hat{\varphi}_{h_0}(u, \underline{s}) - \varphi_{h_0}(u, \underline{s})) \\ \Im(\hat{\varphi}_{h_0}(u, \underline{s}) - \varphi_{h_0}(u, \underline{s})) \end{pmatrix} \right\}_{\underline{s} \in [-S, S]^d} \xrightarrow{d} \{Z_{h_0}(u, \underline{s})\}_{\underline{s} \in [-S, S]^d}, \quad (3.12)$$

where $\{Z_{h_0}(u, \underline{s})\}_{\underline{s} \in [-S, S]^d}$ is a centered Gaussian process with continuous sample paths, zero mean and covariance function \mathbf{V}_{h_0} . Here, \mathbf{V}_{h_0} is defined similar to \mathbf{V} in Theorem 3.2 with $\tilde{X}_h(u)$ and $\tilde{X}_0(u)$ in Lemma 3.1(iii) substituted by $(\tilde{X}'_h(u), \tilde{X}'_{h+h_0}(u))'$ and $(\tilde{X}'_0(u), \tilde{X}'_{h_0}(u))'$, respectively.

4. APPLICATIONS TO STATISTICS

In this section, we exploit the theory derived in the previous section for the local ECF. We use it for the purpose of parameter estimation in Section 4.1 and apply it to the concept of distance correlation in Section 4.2.

4.1. Parameter estimation.

4.1.1. *Parameter estimation in α -stable distributions.* The univariate stable distribution with location parameter $\mu \in \mathbb{R}$, characteristic exponent $\alpha \in (0, 2]$, skewness parameter $\beta \in [-1, 1]$, and scale parameter $\gamma \geq 0$ is determined by the CF

$$\varphi_{\mu, \alpha, \beta, \gamma}(s) = \exp \{i\mu s - \gamma |s|^\alpha [1 + i\beta \operatorname{sgn}(s) f(s, \alpha)]\}, \quad s \in \mathbb{R}, \quad (4.1)$$

with

$$f(s, \alpha) = \begin{cases} \tan\left(\frac{\pi\alpha}{2}\right) & \text{if } \alpha \neq 1 \\ \frac{2}{\pi} \log |s| & \text{if } \alpha = 1 \end{cases}.$$

This class includes normal distributions ($\alpha = 2$) and Cauchy distributions ($\alpha = 1$ and $\beta = 0$) as special cases. In this sense, light and heavy tailed distributions are contained in this class of distributions. Moreover, also asymmetry can be modeled if $\alpha < 2$ and $\beta \neq 0$. Therefore, this parametric family is very appealing to model economic and financial data; see e.g. Borak, Misiolek, Weron (2005, Section 1.2.6) for a critical discussion. This class turns out to be especially suitable for linear locally stationary processes in the following sense. Suppose that the innovations have a (time-invariant) CF $\varphi^{(\varepsilon)}(s) = \varphi_{0, \alpha, \beta, \gamma}(s)$ with $\alpha > 1$, then $\varphi^{(a\varepsilon)}(s) = \varphi_{0, \alpha, \operatorname{sgn}(a)\beta, |a|^\alpha \gamma}(s)$ which in turn implies that the process $(\tilde{X}_t(u))_{t \in \mathbb{Z}}$ defined in (1.7) has a time-varying CF

$$\varphi(u, s) = \varphi_{\tilde{\mu}(u), \tilde{\alpha}(u), \tilde{\beta}(u), \tilde{\gamma}(u)}(s) \quad (4.2)$$

with

$$\begin{aligned}\tilde{\alpha}(u) &= \alpha, \\ \tilde{\mu}(u) &= \mu(u), \\ \tilde{\beta}(u) &= \beta \sum_{j=-\infty}^{\infty} \operatorname{sgn}(A(u, j)) |A(u, j)|^\alpha / \sum_{j=-\infty}^{\infty} |A(u, j)|^\alpha, \\ \tilde{\gamma}(u) &= \gamma \sum_{j=-\infty}^{\infty} |A(u, j)|^\alpha.\end{aligned}$$

Hence, $\tilde{X}_t(u)$ also admits a stable distribution with parameters $\tilde{\alpha}(u)$, $\tilde{\mu}(u)$, $\tilde{\beta}(u)$ and $\tilde{\gamma}(u)$. Fitting stable distributions to data requires suitable parameter estimators. Due to the lack of closed form representation of the density and the cdf for many parameter values, the application of classical maximum likelihood methods can only be carried out numerically. Moreover, higher order moments do not exist if $\alpha < 2$ such that the classical method of moments cannot be invoked. Therefore, procedures relying on the (empirical) CF seem to be more appealing. We introduce two methods in the present and the following section.

In this section, we adapt the plug-in estimators proposed by Press (1972) and references therein in the classical i.i.d. set-up. This approach has the great advantage that it provides closed-form estimators for all four parameters simultaneously, which is not the case for maximum likelihood estimators and the minimum distance estimators discussed in the following section. Based on our results on the ECF under local stationarity, we generalize their results to univariate processes satisfying Assumption 2.1 with $k = 1$. Note that the assumption of a finite first moment is equivalent to $\alpha > 1$ in the stable family. We adapt the plug-in approach of Press (1972) with $\alpha > 1$ and assume $\gamma > 0$ to assure identifiability of all parameters. Solving $|\varphi_{\tilde{\mu}(u), \tilde{\alpha}(u), \tilde{\beta}(u), \tilde{\gamma}(u)}(s)| = \exp(-\tilde{\gamma}(u)|s|^{\alpha(u)})$ for non-zero values s_1 and s_2 with $|s_1| \neq |s_2|$ gives in complete analogy to (2.1) and (2.2) in Press (1972) plug-in estimators for the characteristic exponent

$$\hat{\alpha}(u) = \frac{\log \left| \frac{\log |\hat{\varphi}(u, s_1)|}{\log |\hat{\varphi}(u, s_2)|} \right|}{\log |s_1| - \log |s_2|} \quad (4.3)$$

and for the scale parameter

$$\hat{\gamma}(u) = \exp \left(\frac{\log |s_1| \log(-\log |\hat{\varphi}(u, s_2)|) - \log |s_2| \log(-\log |\hat{\varphi}(u, s_1)|)}{\log |s_1| - \log |s_2|} \right). \quad (4.4)$$

We rely on the imaginary part of the logarithm of the CF, denoted by $\psi(u, s) = \arg(\varphi(u, s))$, in order to establish estimators for $\tilde{\beta}(u)$ and $\tilde{\mu}(u)$. Again in complete analogy to Press (1972), we obtain

$$\tilde{\mu}(u) = \frac{|s_4|^{\tilde{\alpha}(u)-1} \frac{\psi(u, s_3)}{s_3} - |s_3|^{\tilde{\alpha}(u)-1} \frac{\psi(u, s_4)}{s_4}}{|s_4|^{\tilde{\alpha}(u)-1} - |s_3|^{\tilde{\alpha}(u)-1}} \quad (4.5)$$

and

$$\tilde{\beta}(u) = \frac{\frac{\psi(u, s_3)}{s_3} - \frac{\psi(u, s_4)}{s_4}}{\left(|s_4|^{\tilde{\alpha}(u)-1} - |s_3|^{\tilde{\alpha}(u)-1} \right) \tilde{\gamma}(u) \tan(\pi \tilde{\alpha}(u)/2)} \quad (4.6)$$

for non-zero values s_3 and s_4 with $|s_3| \neq |s_4|$ which might coincide with s_1 and s_2 . In particular, note that β is not identifiable in the normal case, i.e. for $\alpha = 2$.

To establish the corresponding estimators $\hat{\mu}(u)$ and $\hat{\beta}(u)$, we substitute in (4.5) and (4.6)

the unknown parameters $\tilde{\alpha}$ and $\tilde{\gamma}$ by their estimates from (4.3) and (4.4), respectively, and it remains to establish the empirical counterpart of $\psi(u, s)$. Again applying the plug-in principle, we obtain

$$\hat{\psi}(u, s) = \arg(\hat{\varphi}(u, s)).$$

Consistency of the parameter estimators follows as a corollary of Theorem 3.1.

Corollary 4.1 (Consistency for plug-in estimators).

Suppose that Assumption 2.1 with $k = d = 1$ and Assumption 2.2 hold true and that $b^3 T \rightarrow 0$ as $T \rightarrow \infty$. Moreover, assume that $\tilde{X}_t(u)$ has a stable distribution with parameters $\mu(u) \in \mathbb{R}$, $\tilde{\alpha}(u) \in (1, 2)$, $\tilde{\beta}(u) \in [-1, 1]$ and $\tilde{\gamma}(u) > 0$, $u \in (0, 1)$. Then

$$(\hat{\mu}(u), \hat{\alpha}(u), \hat{\beta}(u), \hat{\gamma}(u)) \xrightarrow{P} (\mu(u), \tilde{\alpha}(u), \tilde{\beta}(u), \tilde{\gamma}(u)) \quad \forall u \in (0, 1).$$

Remark 3.

- (i) Asymptotic normality of the estimators can be deduced via the Δ method from Theorem 3.1. One can basically follow the lines of Section 3.4.1 in Ushakov (1999), where the i.i.d. case is considered. Since the calculations are straightforward but tedious, we skip details.
- (ii) The theory can be extended to symmetric multivariate stable distributions. Again one can follow the lines of Press (1972, Section 4).
- (iii) Of course, the choice of s_1 to s_4 has an impact on the performance of the estimators. A detailed discussion of this problem for the univariate i.i.d. problem can be found in Fielitz and Rozelle (1981). For s_1 and s_2 they propose values ranging from 0.2 to 1 and found out that good choices of s_3 and s_4 will additionally depend on the estimated values of the characteristic exponent and the scale parameter.

4.1.2. *Minimum distance estimation.* The estimation method described in the previous section depends on the choice of the second argument of the CF used there. This can be circumvented by not only evaluating the local ECF at fixed points but instead using an approach that smooths over all possible values. In the sequel, we consider the L_2 minimum distance estimator based on local ECFs for an unknown parameter, $\theta_0(u)$ say, defined as

$$\hat{\theta}(u) = \arg \min_{\theta \in \Theta} \int |\hat{\varphi}(u, s) - \varphi(u, s; \theta)|^2 w(s) ds, \quad (4.7)$$

where $\Theta \subseteq \mathbb{R}^p$ denotes a compact parameter space, w is a weight function assuring existence of the integral and $\varphi(u, \cdot; \theta)$ is the CF of $\tilde{X}(u) \sim P^{\theta(u)}$. Minimum distance estimators were considered e.g. by Höpfner and Rüschenendorf (1999) in a comparative overview of estimators for the characteristic exponent in stable distributions based on i.i.d. data and turned out to behave favorably.

In the sequel, we do not restrict ourselves to the stable family but allow for a broad parametric class of distribution. In order to establish asymptotic normality of the minimum distance estimator, we assume:

Assumption 4.1.

- (i) *The prerequisites of Theorem 3.2 hold true.*
- (ii) *The weight function w is integrable and strictly positive.*
- (iii) *The parameter $\theta(u) \in \Theta \subseteq \mathbb{R}^p$ is identifiable, i.e. for $\theta_1(u) \neq \theta_2(u) \in \Theta$ it holds $P^{\theta_1(u)} \neq P^{\theta_2(u)}$.*

(iv) $\varphi(u, s; \theta)$ is continuously partially differentiable w.r.t. θ and

$$\int_{\mathbb{R}} \sup_{\theta(u) \in \Theta} |\varphi^{(1)}(u, s; \theta(u))|_1^2 w(s) ds < \infty. \quad (4.8)$$

(v) Additionally, $\varphi(u, s; \theta)$ is twice continuously partially differentiable w.r.t. θ and

$$\int_{\mathbb{R}} \sup_{\theta(u) \in U(\theta_0)} |\varphi^{(2)}(u, s; \theta(u))|_1 w(s) ds < \infty \quad (4.9)$$

for some open neighbourhood $U(\theta_0)$ of θ_0 . It holds

$$\int_{\mathbb{R}} (1 + |s|^2) (1 + |\varphi^{(1)}(u, s; \theta_0(u))|_1^2 + |\varphi^{(2)}(u, s; \theta_0(u))|_1^2) w(s) ds < \infty. \quad (4.10)$$

Finally,

$$D_0 := \int_{\mathbb{R}} \left(\Re \varphi^{(1)}(u, s; \theta_0(u)) [\Re \varphi^{(1)}(u, s; \theta_0(u))]' + \Im \varphi^{(1)}(u, s; \theta_0(u)) [\Im \varphi^{(1)}(u, s; \theta_0(u))]' \right) w(s) ds$$

is invertible.

Under these assumptions we obtain consistency and asymptotic normality of the minimum distance estimator similar to Ushakov (1999, Section 3.4.2) who considered the i.i.d. case.

Theorem 4.1 (CLT for minimum distance estimators).

Under Assumption 4.1(i) - (iv) and for $u \in (0, 1)$, we have

- (i) $\hat{\theta}(u) \xrightarrow{P} \theta_0(u)$.
- (ii) If additionally Assumption 4.1(v) holds and $\theta_0(u)$ lies in the interior of Θ , then

$$\sqrt{bT}(\hat{\theta}(u) - \theta_0(u)) \xrightarrow{d} Z_{MDE}(u) \sim \mathcal{N}(0_p, \Sigma_{MDE}(u))$$

with

$$\begin{aligned} \Sigma_{MDE}(u) = \frac{1}{4} D_0^{-1} \iint_{\mathbb{R}^2} \left\{ & V_{\Re \Re}((u, s_1), (u, s_2)) \Re \varphi^{(1)}(u, s_1; \theta_0(u)) [\Re \varphi^{(1)}(u, s_2; \theta_0(u))]' \right. \\ & + V_{\Im \Im}((u, s_1), (u, s_2)) \Im \varphi^{(1)}(u, s_1; \theta_0(u)) [\Im \varphi^{(1)}(u, s_2; \theta_0(u))]' \\ & - V_{\Re \Im}((u, s_1), (u, s_2)) \left(\Re \varphi^{(1)}(u, s_1; \theta_0(u)) [\Im \varphi^{(1)}(u, s_2; \theta_0(u))]' \right. \\ & \left. \left. + \Im \varphi^{(1)}(u, s_1; \theta_0(u)) [\Re \varphi^{(1)}(u, s_2; \theta_0(u))]' \right) \right\} w(s_1) w(s_2) ds_1 ds_2 D_0^{-1}. \end{aligned}$$

Remark 4.

Under suitably modified assumptions, consistency can also be verified for L_p -minimum distance estimators with $p \in [1, \infty]$.

4.2. Applications to (auto-)distance correlation. Based on the concept of distance correlation as introduced by Szekely, Rizzo and Bakirov (2007) for i.i.d. data and the definition of auto-distance correlation for stationary time series considered in Zhou (2012) and more recently in Davis, Matsui, Mikosch, and Wan (2016), we propose a time-localized version of (auto-)distance correlation in this section. Let $(\underline{X}_{t,T})_{t=1}^T$ be a $d = p + q$ -dimensional time series process with $\underline{X}_{t,T} = (\underline{Y}'_{t,T}, \underline{Z}'_{t,T})'$. Then, we define the *local* (auto-)distance correlation $\mathcal{R}_{Y,Z}(u; h)$ at lag h between time series $(\underline{Y}_{t,T})_{t=1}^T$ and $(\underline{Z}_{t,T})_{t=1}^T$ of dimensions p and q , respectively, at rescaled time $u \in [0, 1]$ by

$$\mathcal{R}_{Y,Z}^2(u; h) = \frac{\mathcal{V}_{Y,Z}^2(u; h)}{\sqrt{\mathcal{V}_Y^2(u; 0)\mathcal{V}_Z^2(u; 0)}} \mathbf{1}(\mathcal{V}_Y^2(u; 0)\mathcal{V}_Z^2(u; 0) > 0). \quad (4.11)$$

Here, $\mathcal{V}_{Y,Z}(u; h)$ denotes the corresponding *local* (auto-)distance covariance defined by

$$\mathcal{V}_{Y,Z}^2(u; h) = \int_{\mathbb{R}^q} \int_{\mathbb{R}^p} |\varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y;0}(u; \underline{s}_1)\varphi_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d\underline{s}_1 d\underline{s}_2, \quad (4.12)$$

where w is the (non-integrable) weight function defined in (1.5). For $\underline{s}_1 \in \mathbb{R}^p$ and $\underline{s}_2 \in \mathbb{R}^q$, the CFs are defined as follows

$$\begin{aligned} \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) &= E \left(\exp \left(i \langle \underline{s}_1, \tilde{Y}_0(u) \rangle + i \langle \underline{s}_2, \tilde{Z}_h(u) \rangle \right) \right), \\ \varphi_{Y;0}(u; \underline{s}_1) &= E \left(\exp \left(i \langle \underline{s}_1, \tilde{Y}_0(u) \rangle \right) \right), \\ \varphi_{Z;h}(u; \underline{s}_2) &= E \left(\exp \left(i \langle \underline{s}_2, \tilde{Z}_h(u) \rangle \right) \right), \end{aligned}$$

where $(\tilde{Y}'_t(t/T), \tilde{Z}'_t(t/T))'_{t=1}^T$ is the stationary approximation to the process $((\underline{Y}'_{t,T}, \underline{Z}'_{t,T})'_{t=1}^T)$; compare (2.5). Also we set $\mathcal{V}_Y(u; h) := \mathcal{V}_{Y,Y}(u; h)$ and $\mathcal{V}_Z(u; h) := \mathcal{V}_{Z,Z}(u; h)$. With data $((\underline{Y}'_{t,T}, \underline{Z}'_{t,T})'_{t=1}^T)$ at hand, empirical versions $\hat{\mathcal{V}}_{Y,Z}(u; h)$ and $\hat{\mathcal{R}}_{Y,Z}(u; h)$ of $\mathcal{V}_{Y,Z}(u; h)$ and $\mathcal{R}_{Y,Z}(u; h)$, respectively, are obtained by simply replacing the CFs φ by suitable ECFs $\hat{\varphi}$ (and by adding a factor κ_T for technical reasons; see Lemma 4.2 and Remark 5 below). Hence, we define

$$\hat{\mathcal{R}}_{Y,Z}^2(u; h) = \frac{\hat{\mathcal{V}}_{Y,Z}^2(u; h)}{\sqrt{\hat{\mathcal{V}}_Y^2(u; 0)\hat{\mathcal{V}}_Z^2(u; 0)}} \mathbf{1}(\hat{\mathcal{V}}_Y^2(u; 0)\hat{\mathcal{V}}_Z^2(u; 0) > 0), \quad (4.13)$$

where

$$\hat{\mathcal{V}}_{Y,Z}^2(u; h) = \int_{\mathbb{R}^q} \int_{\mathbb{R}^p} |\kappa_T \hat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \hat{\varphi}_{Y;0}(u; \underline{s}_1)\hat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d\underline{s}_1 d\underline{s}_2 \quad (4.14)$$

and $\kappa_T = T^{-1} \sum_{t=1}^T K_b(\frac{t}{T} - u)$ and again we set $\hat{\mathcal{V}}_Y(u; h) := \hat{\mathcal{V}}_{Y,Y}(u; h)$, $\hat{\mathcal{V}}_Z(u; h) := \hat{\mathcal{V}}_{Z,Z}(u; h)$. With a slight abuse of notation (as $\hat{\varphi}_{Y;0}(u; \underline{s}_1)$ indeed also depends on h), the ECFs used in the definition above are defined for $h \geq 0$ as

$$\begin{aligned} \hat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) &= \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) \exp \left(i \langle \underline{s}_1, \underline{Y}_{t,T} \rangle + i \langle \underline{s}_2, \underline{Z}_{t+h,T} \rangle \right), \\ \hat{\varphi}_{Y;0}(u; \underline{s}_1) &= \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) \exp \left(i \langle \underline{s}_1, \underline{Y}_{t,T} \rangle \right), \\ \hat{\varphi}_{Z;h}(u; \underline{s}_2) &= \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) \exp \left(i \langle \underline{s}_2, \underline{Z}_{t+h,T} \rangle \right). \end{aligned} \quad (4.15)$$

The latter definitions follow the lines of Dahlhaus (2012), equation (9), where local kernel-type covariance estimators are defined in the same fashion. The special case of $p = q$

and $\underline{Y}_{t,T} = \underline{Z}_{t,T}$ for all t is of particular interest. In this case, $\mathcal{R}_{Y,Y}(u; h)$ measures the auto-distance correlation at lag h of the p -dimensional process $(\underline{Y}_{t,T})_{t=1}^T$. For example, if $p = 1$ we get the extension of the univariate case addressed in Zhou (2012) to the locally stationary framework. In Section 5.2, we apply the local (auto-)distance correlation to log returns of German stock prices to illustrate the pairwise dependence structure over time, i.e. for $p = q = 1$.

The definitions of (4.12) and (4.14) involve integrations over \mathbb{R}^{p+q} which becomes computationally very demanding already for small p and q . Nevertheless, a fundamental property of empirical distance covariances in the i.i.d. and stationary time series setups is that $\widehat{\mathcal{V}}_{Y,Z}(u; h)$ defined in (4.14) can actually be written in a much simpler form that gets rid of the computational burden of numerical integration; cf. Theorem 1 in Szekely, Rizzo and Bakirov (2007) and Proposition 1 in Zhou (2012). To achieve this, the crucial observation is the following lemma; see also Dueck, Edelman and Richards (2015) for a generalization.

Lemma 4.1 (Szekely, Rizzo and Bakirov (2007), Lemma 1).
If $0 < \alpha < 2$, then for all $\underline{x} \in \mathbb{R}^d$

$$\int_{\mathbb{R}^d} \frac{1 - \cos(\langle \underline{t}, \underline{x} \rangle)}{|\underline{t}|_2^{d+\alpha}} d\underline{t} = C(d, \alpha) |\underline{x}|_2^\alpha,$$

where

$$C(d, \alpha) = \frac{2\pi^{d/2}\Gamma(1 - \alpha/2)}{\alpha 2^\alpha \Gamma((d + \alpha)/2)}$$

and $\Gamma(\cdot)$ is the complete gamma function.

Setting $\alpha = 1$ in the result above and following essentially the steps in the proof of Theorem 1 in Szekely, Rizzo and Bakirov (2007), we get the following much simpler form of $\widehat{\mathcal{V}}_{Y,Z}(u; h)$.

Lemma 4.2 (Alternative representation of $\widehat{\mathcal{V}}_{X,Y}(u; h)$).

We have

$$\widehat{\mathcal{V}}_{Y,Z}^2(u; h) = \kappa_T^2 \widehat{S}_{1,Y,Z}(u; h) + \widehat{S}_{2,Y,Z}(u; h) - 2\kappa_T \widehat{S}_{3,Y,Z}(u; h), \quad (4.16)$$

where

$$\begin{aligned} \widehat{S}_{1,Y,Z}(u; h) &= \frac{1}{T^2} \sum_{t_1, t_2=1}^{T-h} K_b\left(\frac{t_1 + h/2}{T} - u\right) K_b\left(\frac{t_2 + h/2}{T} - u\right) |\underline{Y}_{t_1, T} - \underline{Y}_{t_2, T}|_2 |\underline{Z}_{t_1+h, T} - \underline{Z}_{t_2+h, T}|_2 \\ \widehat{S}_{2,Y,Z}(u; h) &= \left(\frac{1}{T^2} \sum_{t_1, t_2=1}^{T-h} K_b\left(\frac{t_1 + h/2}{T} - u\right) K_b\left(\frac{t_2 + h/2}{T} - u\right) |\underline{Y}_{t_1, T} - \underline{Y}_{t_2, T}|_2 \right) \\ &\quad \times \left(\frac{1}{T^2} \sum_{t_3, t_4=1}^{T-h} K_b\left(\frac{t_3 + h/2}{T} - u\right) K_b\left(\frac{t_4 + h/2}{T} - u\right) |\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}|_2 \right), \\ \widehat{S}_{3,Y,Z}(u; h) &= \frac{1}{T^3} \sum_{t_1, t_2, t_3=1}^{T-h} K_b\left(\frac{t_1 + h/2}{T} - u\right) K_b\left(\frac{t_2 + h/2}{T} - u\right) K_b\left(\frac{t_3 + h/2}{T} - u\right) \\ &\quad \times |\underline{Y}_{t_1, T} - \underline{Y}_{t_2, T}|_2 |\underline{Z}_{t_1+h, T} - \underline{Z}_{t_3+h, T}|_2. \end{aligned}$$

In the following theorem we prove consistency of empirical local (auto-)distance covariances and correlations defined in (4.14) and (4.13), respectively. These results relate to Theorem 2 and Corollary 1 in Szekely, Rizzo and Bakirov (2007), where almost sure convergence was established in an i.i.d. setup, to Theorem 1 and Corollary 1 in Zhou (2012), where consistency for stationary processes under a physical dependence condition was shown, and to Theorem 3.1 in Davis, Matsui, Mikosch, and Wan (2016), who prove consistency under strong mixing assumptions while allowing for a more flexible class of weight functions.

Theorem 4.2 (Consistency of $\widehat{\mathcal{V}}_{Y,Z}^2(u; h)$ and $\widehat{\mathcal{R}}_{Y,Z}^2(u; h)$).

Suppose that the assumptions of Theorem 3.2 hold true for $(\underline{X}_{t,T})_{t=1}^T$, $T \in \mathbb{N}$, with $\underline{X}_{t,T} = (\underline{Y}'_{t,T}, \underline{Z}'_{t,T})'$. Then, we have

$$\widehat{\mathcal{V}}_{Y,Z}^2(u; h) \xrightarrow{P} \mathcal{V}_{Y,Z}^2(u; h).$$

If $\text{Var}(\widetilde{X}(u)) > 0$ and $\text{Var}(\widetilde{Y}(u)) > 0$, then additionally

$$\widehat{\mathcal{R}}_{Y,Z}^2(u; h) \xrightarrow{P} \mathcal{R}_{Y,Z}^2(u; h).$$

Remark 5 (On the necessity of κ_T).

Taking a closer look at the proofs, it can be observed, why the introduction of κ_T in the definition of $\widehat{\mathcal{V}}_{Y,Z}(u; h)$ in (4.14) is indeed necessary. Without κ_T several terms in the proof of Lemma 4.2 do not cancel exactly such that the simple representation in (4.16) does not hold. More importantly, one can even show divergence of the integral defining $\widehat{\mathcal{V}}_{X,Y}$ if κ_T is not included.

5. NUMERICAL EXAMPLES

5.1. A simulation study: Parameter estimation of α -stable distributions. In this section we present results of a simulation study involving the two kinds of parameter estimation for locally stationary processes with α -stable distributions as described in Section 4.1: the moment-based plug-in method from Section 4.1.1, which in this section will simply be called 'moment method', as well as the minimum distance estimation from Section 4.1.2, abbreviated by 'MDE method' in the following. For the sake of simplicity, we will focus on centered and symmetric α -stable distributions, i.e. with mean $\mu = 0$ and skewness parameter $\beta = 0$.

The data were generated by the time-varying AR(1)-model

$$X_{t,T} = r\left(\frac{t}{T}\right) X_{t-1,T} + \varepsilon_t, \quad (5.1)$$

with $r(u) = 0.9 \cdot \sin(2\pi u)$, which can be locally approximated by the stationary process

$$\widetilde{X}_t(u) = \sum_{j=0}^{\infty} r(u)^j \varepsilon_{t-j},$$

corresponding to (2.5) with $\mu \equiv 0$ and

$$A(u, j) = r(u)^j \cdot \mathbf{1}_{\{j \geq 0\}},$$

cf. Dahlhaus (2012). Here, $(\varepsilon_t)_{t \in \mathbb{N}_0}$ was generated as an i.i.d. sequence with a centered and symmetric, α -stable marginal distribution, i.e. with CF (4.1) and

$$\mu = 0, \quad \alpha = 1.5, \quad \beta = 0, \quad \gamma = 0.5.$$

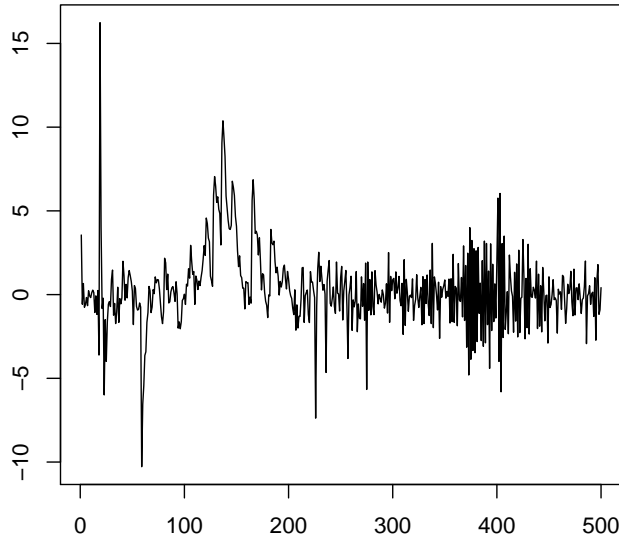


FIGURE 1. Sample path of length $T = 500$ from (5.1) with α -stable innovations and parameters $\mu = 0$, $\alpha = 1.5$, $\beta = 0$, $\gamma = 0.5$.

A sample path of length $T = 500$ from (5.1) with the aforementioned choice of parameter values is given in Figure 1. One can clearly observe the time-varying autoregressive structure as well as some peaks resulting from the fact that the distribution of the innovations has an infinite variance. The stationary processes $(\tilde{X}_t(u))_{t \in \mathbb{Z}}$ approximating (5.1) then possess the CFs (4.2) with

$$\tilde{\mu}(u) = 0, \quad \tilde{\alpha}(u) = 1.5, \quad \tilde{\beta}(u) = 0, \quad \tilde{\gamma}(u) = \frac{0.5}{1 - |0.9 \cdot \sin(2\pi u)|^{1.5}}.$$

First, we implemented the moment method from Section 4.1.1, based on a sample of length $T = 5000$. Here, $\tilde{\alpha}(u)$ and $\tilde{\gamma}(u)$ were estimated according to (4.3) and (4.4), respectively, with $s_1 = 0.6$ and $s_2 = 0.8$ which have shown a favorable behavior of the corresponding estimators in the classical i.i.d. case; see Fielitz and Rozelle (1981). Since only $\tilde{\gamma}(u)$ actually depends on its location $u \in [0, 1]$, we focus our analysis mainly on this parameter. The ECF was calculated with a bandwidth of $b_T = 0.25 \cdot T^{-1/5}$. Figure 2 shows the true function $\tilde{\gamma}(u)$ and boxplots of the estimators $\hat{\gamma}(u)$ at four different locations between $u = 0.2$ and $u = 0.8$, each boxplot based on $N = 100$ iterations. However, the performance of this moment-based method critically depends on the choice of parameters s_1 and s_2 . For a different choice of these values, namely $s_1 = 0.1$ and $s_2 = 0.11$, we get the results from Figure 3 which are obviously considerably worse.

Next, we turn to the results for the joint minimum distance estimation of the parameters $\tilde{\alpha}(u)$ and $\tilde{\gamma}(u)$ as introduced in Section 4.1.2, also based on a sample of length $T = 5000$. The minimum was determined simultaneously for both parameters by calculating the integrated distance (4.7) for each combination of $\tilde{\alpha}(u)$ - and $\tilde{\gamma}(u)$ -values on a grid of values within the interval $(1, 2]$ for $\tilde{\alpha}(u)$ and within $[0, 6]$ for $\tilde{\gamma}(u)$. For the sake of simplicity we used $w(s) = \mathbf{1}\{-0.5 \leq s \leq 0.5\}$ as a weight function, which suffices to ensure that all parameters are identifiable for the class of centered, symmetric α -stable distributions which are under consideration here. The results for $\tilde{\gamma}(u)$ are shown in Figure 4. One can see that, although using a simple weight function, the MDE method produces results comparable to the ones

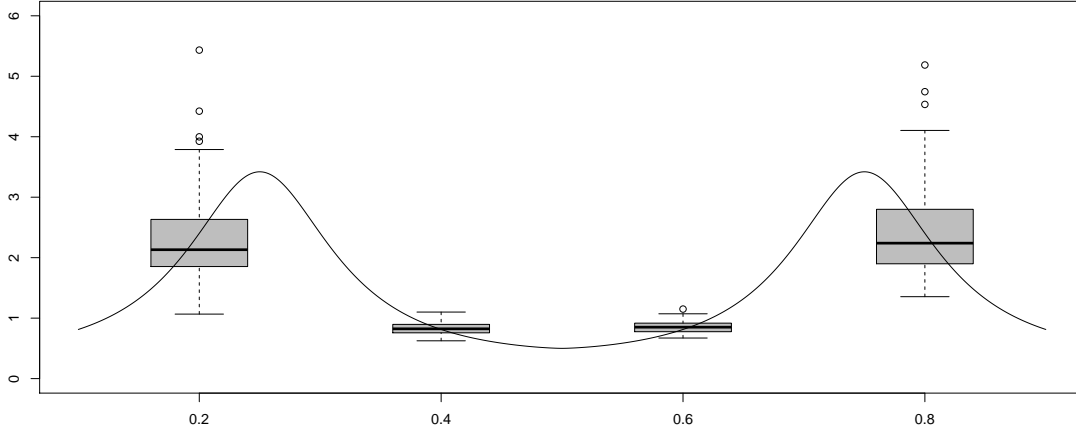


FIGURE 2. True function $\tilde{\gamma}(u)$ and boxplots of moment-based estimators at locations (from left to right) $u = 0.2$, $u = 0.4$, $u = 0.6$ and $u = 0.8$.

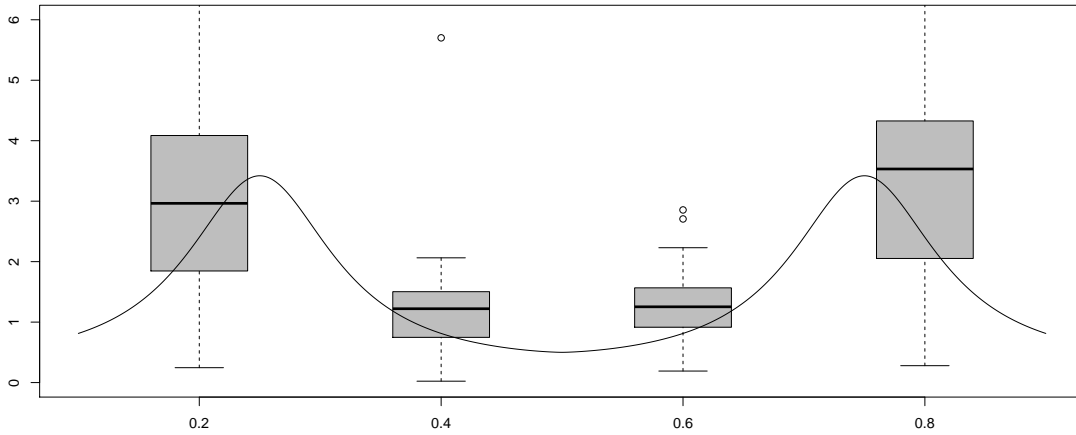


FIGURE 3. True function $\tilde{\gamma}(u)$ and boxplots of moment-based estimators (with 'bad' choice of parameters s_1 and s_2) at locations (from left to right) $u = 0.2$, $u = 0.4$, $u = 0.6$ and $u = 0.8$.

of the moment-based method, with a slightly higher variance but less bias. In addition, the MDE method does not require a choice of tuning parameters s_1 and s_2 – a choice that the moment-based method is rather sensitive to, as shown by Figure 3.

The results for estimation of the parameter $\tilde{\alpha}(u)$ were similar to the ones of $\tilde{\gamma}(u)$ shown in Figures 2 - 4. Since $\tilde{\alpha}(u)$ does not actually depend on u , we show results of the two estimation methods for $u = 0.4$, only, in Figure 5. As observed for the estimation of $\tilde{\gamma}(u)$ before, the MDE method for $\tilde{\alpha}(u)$ shows a slightly higher variance but less bias compared to the moment method with a 'good' choice of tuning parameters $s_1 = 0.6$ and $s_2 = 0.8$. However, taking the results for a 'bad' choice of $s_1 = 0.1$ and $s_2 = 0.11$ into account, the moment method is very sensitive w.r.t. these tuning parameters.

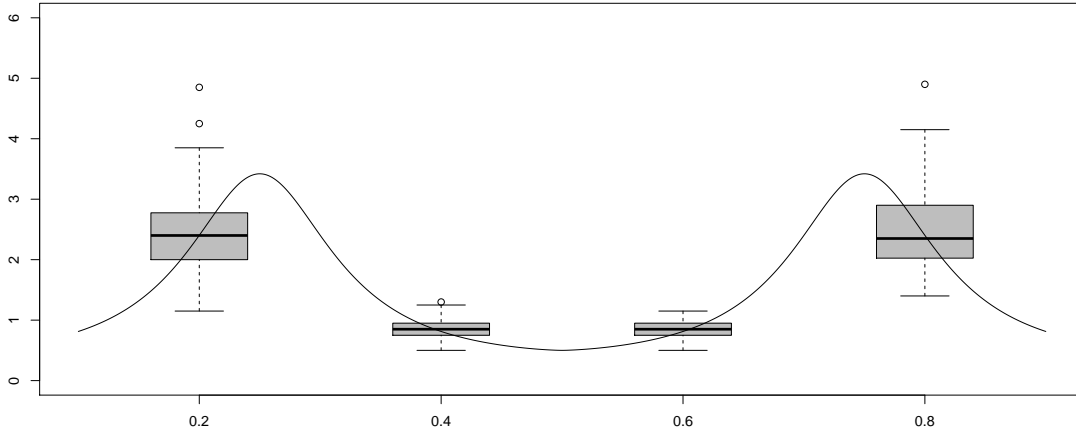


FIGURE 4. True function $\tilde{\gamma}(u)$ and boxplots of MDE estimators at locations (from left to right) $u = 0.2$, $u = 0.4$, $u = 0.6$ and $u = 0.8$.

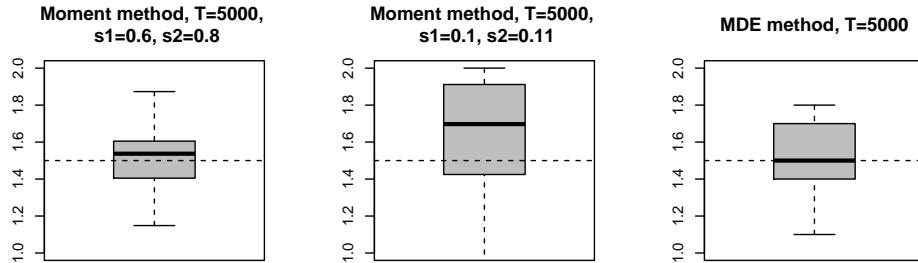


FIGURE 5. Results for $\hat{\alpha}(0.4)$ from the moment method with 'good' choice of parameters s_1 and s_2 (left), the moment method with 'bad' choice of parameters (center) and from the MDE method (right), each plot generated from a sample of size $T = 5000$. The true value $\alpha = 1.5$ is given by the dashed line.

5.2. A real data illustration: Measuring dependence of log returns from German stock prices by local (auto-)distance correlations. In this section, we illustrate the applicability of local distance correlations as introduced in Section 4.2 to measure dependence between financial time series. In particular, this measure allows to study the time-varying cross-dependence of two univariate locally stationary processes. Precisely, let $((Y_{1,T}, Z_{1,T})', \dots, (Y_{T,T}, Z_{T,T})')$ be bivariate time series data, where Y and Z are log-returns of German Stock Prices from July 1, 1991 till April 14, 2015 (6202 trading days). Based on this data, we compute local distance correlations

$$\hat{\mathcal{R}}_{Y,Z}(u; h), \quad h = 0, 1, 2,$$

over rescaled time as defined in (4.13). Note that the weight function w in (1.5) that is used in the definition of $\hat{\mathcal{V}}_{Y,Z}(u; h)$ simplifies to

$$w(t, s) = \frac{1}{\pi^2 t^2 s^2}, \quad t, s \in \mathbb{R} \setminus \{0\},$$

for this case of $p = q = 1$. We have used the bandwidth $b = 0.05$ and the kernel $K(x) = \frac{3}{4}(1 - x^2)\mathbb{1}(x \in [-1, 1])$. For rescaled time u , this choice of b allows us to compute

$\widehat{\mathcal{V}}_{Y,Z}(u; h)$ (and $\widehat{\mathcal{R}}_{Y,Z}(u; h)$) for $u \in [0.05, 0.95]$ without running into boundary issues.

In Figure 6, we show pairwise comparisons of log-returns of stock prices of the German automobile companies Volkswagen and BMW and the automotive supplier Continental together with the corresponding $\widehat{\mathcal{R}}_{Y,Z}(u; h)$ and local cross-correlation $\widehat{\rho}_{Y,Z}(u; h)$ plotted over time for different lags. The local cross-correlation $\widehat{\rho}_{Y,Z}(u; h)$ is defined by $\widehat{\rho}_{Y,Z}(u; h) = \widehat{\gamma}_{Y,Z}(u; h) / \widehat{\gamma}_{Y,Z}(u; 0)$, where

$$\widehat{\gamma}_{Y,Z}(u; h) = \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) (Y_{t,T} - \widehat{\mu}_{Y;0}(u))(Z_{t+h,T} - \widehat{\mu}_{Z;h}(u))$$

with $\widehat{\mu}_{Y;0}(u) = \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) Y_{t,T}$ and $\widehat{\mu}_{Z;h}(u) = \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h}{T} - u\right) Z_{t+h,T}$; see e.g. Dahlhaus (2012). The log-returns clearly reveal the turbulent times in the late 90s and early 2000s with several crises (Asian crisis, Russian crisis, Argentine crisis, Dot-com bubble) including the terror attacks on September 11, 2001 as well as the financial crisis beginning in 2007 which also caused the Euro crisis, which reached its climax around 2012. In terms of volatility, the crisis around 2008 appears to be more severe than those around 2000 and 2012. The third row of panels in Figure 6 shows contemporaneous local distance correlation and local correlation between log-return time series at lag $h = 0$. Cross-sectional dependence as well as its variation over time is clearly visible between all log-returns. Interestingly, the local correlation turns out to be close to the distance correlation most of the time. The only exceptions are between Volkswagen and Continental and between Volkswagen and BMW during the financial crisis in 2008, where local distance correlation differs considerably from local correlation and turns out to be much larger. This phenomenon is not visible between BMW and Continental. As can be seen in fourth and fifth row of panels in Figure 6, distance correlation as well as correlations drop considerably for lags $h > 0$. At the same time, the variation over time is less pronounced in comparison to $h = 0$. To explain some of the differences between distance correlation and correlation for $h > 0$, note that distance correlation is a non-negative measure whereas correlation can be negative.

The fact that distance correlation and correlation are indeed close for most of the time might be explained by Figure 1 in Szekely, Rizzo and Bakirov (2007), where the authors demonstrate that squared distance correlation and squared correlation are actually close under Gaussianity. Recall that this special case does not generate any non-linear dependence at all. To explain the phenomenon that distance correlation and correlation between Volkswagen and Continental as well as Volkswagen and BMW are not close during the financial crisis, observe that Volkswagen shows very dominant peaks at that time. This gives some strong evidence for the local distribution of log-returns of Volkswagen to be heavy tailed and, consequently, non-Gaussian. In this case, the correlations actually might not exist. Nevertheless, the sample correlation can always be computed and it is assured to be finite between -1 and 1 . However, it is not clear what the sample correlation is measuring in this case, whereas the sample distance correlation does exist under the less restrictive assumption of finite first absolute moments. Hence, it might be advisable in particular in times of heavy tailedness to look at the non-linear dependence measure of distance correlation instead of the linear dependence measure of correlations.

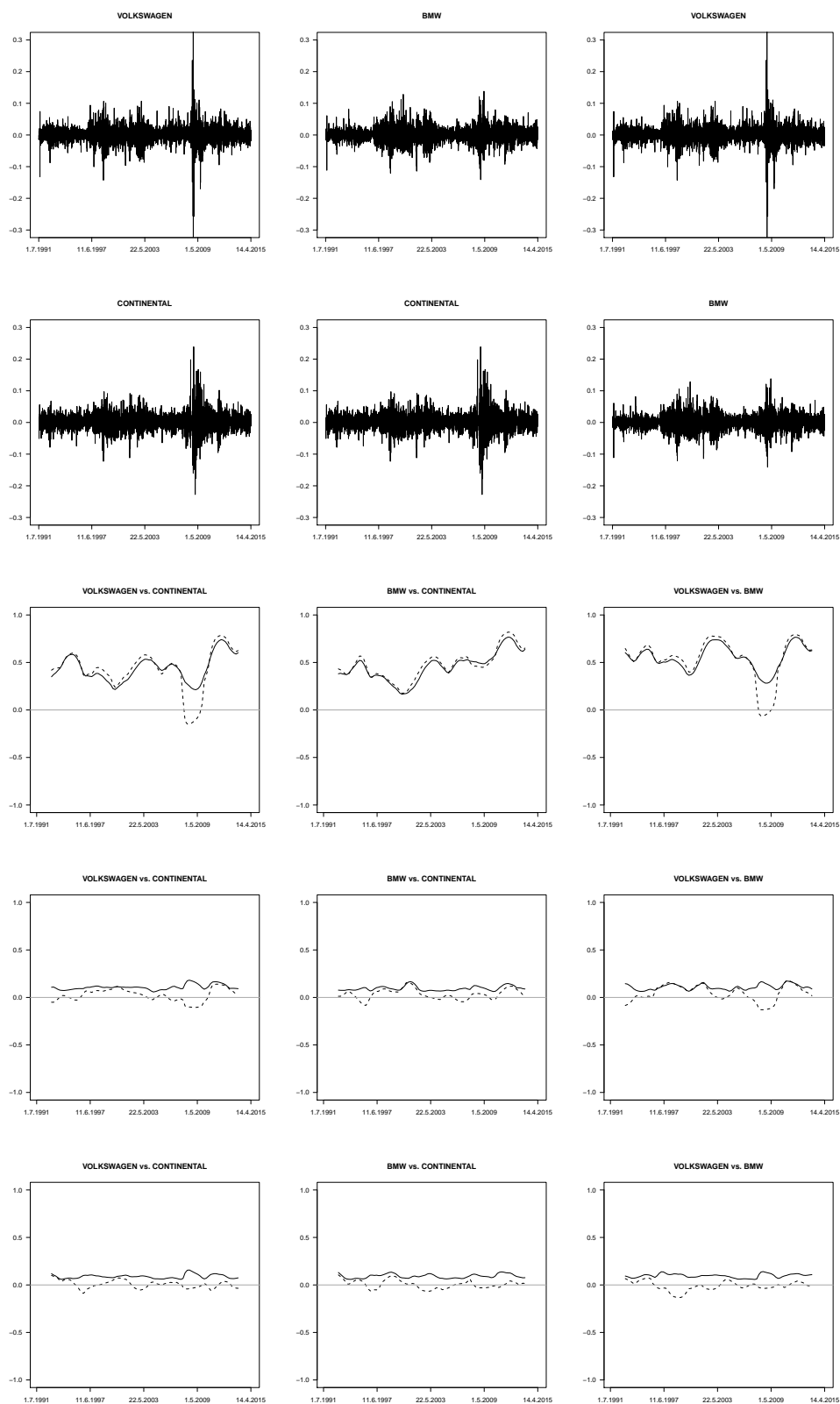


FIGURE 6. Log returns of Stock data of German companies, Volkswagen, BMW and Continental from July 1, 1991 till April 14, 2015 together with corresponding empirical local cross-distance correlation (solid) plots and empirical local cross-correlations (dashed) at lags $h = 0, 1, 2$ (from top to bottom).

6. PROOFS

Throughout this section C denotes a generic constant in $(0, \infty)$ that may change its value from line to line.

6.1. Proofs of Section 2.

Proof of Lemma 2.1. (i) We abbreviate $v = \lfloor |h/2| \rfloor$ and define $\tilde{X}_t^{(v)}(u) = \mu(u) + \sum_{|j| \leq v} A(u, j) \varepsilon_{t-j}$. In the following we use the notation $g_{\underline{s}}(\underline{x}) := \exp(i \langle \underline{s}, \underline{x} \rangle)$. Now we obtain from independence of the innovations, as well as $|\text{Cov}(X, Y)| \leq E|XY| + E|X|E|Y|$ for complex-valued X, Y and $|g_{\underline{s}}(\underline{x})| = 1$,

$$\begin{aligned} & \sup_{u_1, u_2 \in [0, 1]} |\text{Cov}(\exp(i \langle \underline{s}_1, \tilde{X}_h(u_1) \rangle), \exp(i \langle \underline{s}_2, \tilde{X}_0(u_2) \rangle))| \\ & \leq \sup_{u_1, u_2 \in [0, 1]} |\text{Cov}(g_{\underline{s}_1}(\tilde{X}_h(u_1)) - g_{\underline{s}_1}(\tilde{X}_h^{(v)}(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_2)))| \\ & \quad + \sup_{u_1, u_2 \in [0, 1]} |\text{Cov}(g_{\underline{s}_1}(\tilde{X}_h^{(v)}(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_2)) - g_{\underline{s}_2}(\tilde{X}_0^{(v)}(u_2)))| \\ & \leq \sup_{u_1 \in [0, 1]} 2E|g_{\underline{s}_1}(\tilde{X}_h(u_1)) - g_{\underline{s}_1}(\tilde{X}_h^{(v)}(u_1))|_1 + \sup_{u_2 \in [0, 1]} 2E|g_{\underline{s}_2}(\tilde{X}_0(u_2)) - g_{\underline{s}_2}(\tilde{X}_0^{(v)}(u_2))|_1. \end{aligned}$$

Now, we invoke a Taylor expansion of first order for $g_{\underline{s}}$. Note that the gradient $\nabla g_{\underline{s}}(\cdot)$ fulfills $|\nabla g_{\underline{s}}(\underline{x})|_1 \leq |\underline{s}|_1$. Using (2.4), the first summand on the right-hand side equals

$$\begin{aligned} 2 \sup_{u_1 \in [0, 1]} \|(\nabla g_{\underline{s}_1}(\xi_h))' \cdot (\tilde{X}_h(u_1) - \tilde{X}_h^{(v)}(u_1))\|_1 & \leq 2|\underline{s}_1|_1 \sup_{u_1 \in [0, 1]} \|\tilde{X}_h(u_1) - \tilde{X}_h^{(v)}(u_1)\|_1 \\ & \leq 2|\underline{s}_1|_1 \sum_{|j| > v} \sup_{u_1 \in [0, 1]} |A(u_1, j)|_1 \cdot \|\varepsilon_{h-j}\|_1 \\ & \leq 2d|\underline{s}_1|_1 \|\varepsilon_0\|_1 \sum_{|j| > v} \frac{B}{l(j)}, \end{aligned}$$

for suitable ξ_h between $\tilde{X}_h(u_1)$ and $\tilde{X}_h^{(v)}(u_1)$. An analogous bound for the second summand on the right-hand side above finally yields

$$\sup_{u_1, u_2 \in (0, 1]} \left| \text{Cov}(\exp(i \langle \underline{s}_1, \tilde{X}_h(u_1) \rangle), \exp(i \langle \underline{s}_2, \tilde{X}_0(u_2) \rangle)) \right| \leq 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\varepsilon_0\|_1 \sum_{|j| > v} \frac{B}{l(j)}.$$

(ii) Using representations (2.1) and (2.5), Assumption 2.1, submultiplicativity of the $|\cdot|_1$ -norm and the i.i.d. property of $(\varepsilon_t)_t$ we get

$$\begin{aligned} \left\| \underline{X}_{t, T} - \tilde{\underline{X}}_t(t/T) \right\|_1 & \leq \sum_{j=-\infty}^{\infty} |A_{t, T}(j) - A(t/T, j)|_1 \|\varepsilon_{t-j}\|_1 \\ & \leq \frac{1}{T} \|\varepsilon_0\|_1 \sum_{j=-\infty}^{\infty} \sup_{t, T} T |A_{t, T}(j) - A(t/T, j)|_1 \\ & \leq \frac{1}{T} \|\varepsilon_0\|_1 \sum_{j=-\infty}^{\infty} \frac{B}{l(j)} \\ & = \mathcal{O}(T^{-1}). \end{aligned}$$

(iii) All components of the mean function $\underline{\mu}$ are Lipschitz continuous due to the continuous differentiability condition in Assumption 2.1. Hence, we get

$$|\underline{\mu}(u) - \underline{\mu}(w)|_1 = \sum_{r=1}^d |\mu_r(u) - \mu_r(w)| \leq Ld \cdot |u - w|,$$

where L is the maximum of the d Lipschitz constants of the components of $\underline{\mu}(\cdot)$. Now, using representation (2.5) and a Taylor expansion for each entry $a^{(k,r)}(q, j)$ of matrix $A(q, j)$, it follows from Assumption 2.1(ii.2)

$$\begin{aligned} & \|\tilde{\underline{X}}_h(u) - \tilde{\underline{X}}_h(w)\|_1 \\ & \leq |\underline{\mu}(u) - \underline{\mu}(w)|_1 + \sum_{j=-\infty}^{\infty} |A(u, j) - A(w, j)|_1 \|\underline{\varepsilon}_{h-j}\|_1 \\ & \leq Ld|u - w| + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \max_{r=1, \dots, d} \sum_{k=1}^d |a^{(k,r)}(u, j) - a^{(k,r)}(w, j)| \\ & = Ld|u - w| + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \max_{r=1, \dots, d} \sum_{k=1}^d \left| \left(\frac{\partial a^{(k,r)}(q, j)}{\partial q} \Big|_{q=\xi_{k,r,j}} \right) (u - w) \right| \\ & \leq \left(Ld + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \cdot |u - w|, \end{aligned}$$

for suitable $\xi_{k,r,j}$ between $a^{(k,r)}(u, j)$ and $a^{(k,r)}(w, j)$. \square

Proof of Lemma 2.2. (i) By Lipschitz continuity of the function $\exp(i \cdot)$ with constant 1, we get from Lemma 2.1(ii)

$$\begin{aligned} \sup_{s \in [-S, S]^d, 1 \leq t \leq T} \left| \varphi_{t,T}(\underline{s}) - \varphi \left(\frac{t}{T}, \underline{s} \right) \right| & \leq \sup_{s \in [-S, S]^d, 1 \leq t \leq T} E \left| \left\langle \underline{s}, \underline{X}_{t,T} - \tilde{\underline{X}}_t \left(\frac{t}{T} \right) \right\rangle \right| \\ & \leq S \sup_{1 \leq t \leq T} \left\| \underline{X}_{t,T} - \tilde{\underline{X}}_t \left(\frac{t}{T} \right) \right\|_1 \quad (6.1) \\ & \leq \left(S \|\underline{\varepsilon}_0\|_1 \sum_{j=-\infty}^{\infty} \frac{B}{l(j)} \right) \frac{1}{T}. \end{aligned}$$

(ii) Let $k \geq 1$. Under Assumption 2.1(ii), we get existence of

$$\tilde{\underline{X}}_1^{(l)}(u) = \frac{\partial^l}{\partial u^l} \left(\tilde{\underline{X}}_1(u) \right) = \underline{\mu}^{(l)}(u) + \sum_{j=-\infty}^{\infty} A^{(l)}(u, j) \underline{\varepsilon}_{t-j}, \quad l = 0, \dots, k.$$

Similarly, by Lebesgue's dominated convergence theorem and as $\exp(i \langle \underline{s}, \tilde{\underline{X}}_1(\cdot) \rangle)$ is a composition of k -times continuously differentiable functions, we get also existence and continuity of

$$\varphi^{(l)}(u, \underline{s}) = \frac{\partial^l}{\partial u^l} (\varphi(u, \underline{s})) = E \left(\frac{\partial^l}{\partial u^l} \exp(i \langle \underline{s}, \tilde{\underline{X}}_0(u) \rangle) \right), \quad l = 0, \dots, k.$$

To see this, we consider $l = 1$ in detail first. Note that

$$\frac{\partial e^{i \langle \underline{s}, \tilde{\underline{X}}_0(u) \rangle}}{\partial u} = i e^{i \langle \underline{s}, \tilde{\underline{X}}_0(u) \rangle} \langle \underline{s}, \tilde{\underline{X}}_0^{(1)}(u) \rangle, \quad u \in [0, 1],$$

is absolutely integrable. The mean value theorem gives

$$\begin{aligned} & \lim_{h \rightarrow 0} \left| E \left(\frac{e^{i\langle \underline{s}, \tilde{X}_0(u+h) \rangle} - e^{i\langle \underline{s}, \tilde{X}_0(u) \rangle}}{h} \right) - E \left(i e^{i\langle \underline{s}, \tilde{X}_0(u) \rangle} \langle \underline{s}, \tilde{X}_0^{(1)}(u) \rangle \right) \right| \\ &= \lim_{h \rightarrow 0} \left| E \left(i e^{i\langle \underline{s}, \tilde{X}_0(u+\xi(h)) \rangle} \langle \underline{s}, \tilde{X}_0^{(1)}(u+\xi(h)) \rangle \right) - E \left(i e^{i\langle \underline{s}, \tilde{X}_0(u) \rangle} \langle \underline{s}, \tilde{X}_0^{(1)}(u) \rangle \right) \right| \end{aligned} \quad (6.2)$$

for some $\xi(h) = \xi(h, \omega) \rightarrow 0$ as $h \rightarrow 0$ for each fixed ω . The right-hand side of (6.2) equals zero by Lebesgue's dominated convergence theorem. Similarly, we obtain continuity of $\varphi^{(l)}(\cdot, \underline{s})$.

For $l = 2$ we obtain absolute integrability of

$$\frac{\partial^2 e^{i\langle \underline{s}, \tilde{X}_0(u) \rangle}}{\partial u^2} = i e^{i\langle \underline{s}, \tilde{X}_0(u) \rangle} \langle \underline{s}, \tilde{X}_0^{(2)}(u) \rangle - e^{i\langle \underline{s}, \tilde{X}_0(u) \rangle} \left(\langle \underline{s}, \tilde{X}_0^{(1)}(u) \rangle \right)^2, \quad u \in [0, 1],$$

from $E|\underline{\varepsilon}_0|_2^2 < \infty$. Hence, with the same arguments as before we can derive $\varphi^{(2)}(u, \underline{s})$. The proofs for higher order derivatives are analogous and therefore omitted. \square

6.2. Proofs of Section 3.

Proof of Lemma 3.1. (i) Bias – first part:

For the bias term we get

$$\begin{aligned} & E(\hat{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) \\ &= \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \varphi_{t,T}(\underline{s}) - \varphi(u, \underline{s}) \\ &= \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) (\varphi_{t,T}(\underline{s}) - \varphi(u, \underline{s})) + \left(\frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) - 1 \right) \cdot \varphi(u, \underline{s}). \end{aligned} \quad (6.3)$$

In order to get a bound for the second summand on the right-hand side, consider that we get for T large enough such that $-u/b \leq -1$ and $(1-u)/b \geq 1$

$$\int_0^1 \frac{1}{b} K\left(\frac{y-u}{b}\right) dy = \int_{-u/b}^{(1-u)/b} K(z) dz = \int_{-1}^1 K(z) dz = 1$$

by the standard substitution $z = (y-u)/b$ and from Assumption 2.2 (i). Therefore, for large T ,

$$\frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) - 1 = \frac{1}{b} \sum_{t=1}^T \int_{(t-1)/T}^{t/T} \left(K\left(\frac{t/T - u}{b}\right) - K\left(\frac{y-u}{b}\right) \right) dy. \quad (6.4)$$

Since K is Lipschitz, each of the integrals on the right-hand side can be bounded in absolute value by

$$\int_{(t-1)/T}^{t/T} C \cdot \left| \frac{t/T - u}{b} - y \right| dy \leq \int_{(t-1)/T}^{t/T} C \cdot \frac{1}{bT} dy = C \cdot \frac{1}{bT^2}.$$

Since $K(x)$ is zero for all $x \notin [-1, 1]$ it is easy to see that for T large enough at most $\lceil 2bT \rceil$ of the T summands in (6.4) are non-zero. Hence, (6.4) can be bounded in absolute value by

$$\frac{1}{b} \cdot \lceil 2bT \rceil \cdot \frac{C}{bT^2} = \mathcal{O}\left(\frac{1}{bT}\right).$$

Therefore, (6.3) is equal to

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) (\varphi_{t,T}(\underline{s}) - \varphi(u, \underline{s})) + \mathcal{O}\left(\frac{1}{bT}\right) \\ &= \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\varphi_{t,T}(\underline{s}) - \varphi\left(\frac{t}{T}, \underline{s}\right)\right) \end{aligned} \quad (6.5)$$

$$+ \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\varphi\left(\frac{t}{T}, \underline{s}\right) - \varphi(u, \underline{s})\right) + \mathcal{O}\left(\frac{1}{bT}\right). \quad (6.6)$$

In order to derive the bound for the bias from assertion (3.1), which is uniform in $\underline{s} \in [-S, S]^d$, notice that the $\mathcal{O}((bT)^{-1})$ term in (6.6) is uniform in \underline{s} since $|\varphi(u, \underline{s})| \leq 1$. Moreover, using Lemma 2.2(i), the term in (6.5) can be bounded in absolute value (uniformly in $\underline{s} \in [-S, S]^d$) by

$$\frac{C}{T} \sum_{t=1}^T \frac{1}{bT} K\left(\frac{t/T - u}{b}\right) = \frac{C}{T} \cdot \mathcal{O}(1) = \mathcal{O}(T^{-1}) \quad (6.7)$$

because of (2.8). Hence, assertion (3.1) holds if we can show that

$$\sup_{\underline{s} \in [-S, S]^d} \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left| \varphi\left(\frac{t}{T}, \underline{s}\right) - \varphi(u, \underline{s}) \right| = \mathcal{O}(b). \quad (6.8)$$

Lemma 2.2 (ii) ensures existence of $\varphi^{(1)}(u, \underline{s})$, which is continuous in both arguments and therefore bounded on compact sets. Hence, we have from the mean value theorem

$$\begin{aligned} & \sup_{\underline{s} \in [-S, S]^d} \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left| \varphi\left(\frac{t}{T}, \underline{s}\right) - \varphi(u, \underline{s}) \right| \\ & \leq \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \sup_{\underline{s} \in [-S, S]^d} \left| \varphi^{(1)}(\tilde{u}, \underline{s}) \left(\frac{t}{T} - u\right) \right| \\ & \leq \sup_{\underline{s} \in [-S, S]^d, u \in [0, 1]} \left| \varphi^{(1)}(u, \underline{s}) \right| \cdot \sum_{t=1}^T \frac{1}{bT} K\left(\frac{t/T - u}{b}\right) \left| \frac{t}{T} - u \right| = \mathcal{O}(b) \end{aligned}$$

for some \tilde{u} between u and t/T due to (2.8). It follows (6.8) which completes the proof of part (i).

(ii) *Bias – second part:*

Following exactly the lines of the proof of (i) up to (6.7) yields

$$E(\hat{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) = \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\varphi\left(\frac{t}{T}, \underline{s}\right) - \varphi(u, \underline{s})\right) + \mathcal{O}\left(\frac{1}{bT}\right).$$

By Lemma 2.2(ii) $\varphi(\cdot, \underline{s})$ is three times differentiable w.r.t. the first argument, hence, we can use a Taylor expansion of $\varphi(t/T, \underline{s})$ to derive for the leading term in (6.6)

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\varphi\left(\frac{t}{T}, \underline{s}\right) - \varphi(u, \underline{s})\right) \\ &= \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \varphi^{(1)}(u, \underline{s}) \left(\frac{t}{T} - u\right) + \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \frac{1}{2} \varphi^{(2)}(\tilde{u}, \underline{s}) \left(\frac{t}{T} - u\right)^2 \end{aligned} \quad (6.9)$$

where \tilde{u} is between u and t/T . Now we use that K is symmetric, cf. Assumption 2.2, i.e. $\int_{-1}^1 K(z) z dz = 0$, to derive for T large enough such that $-u/b \leq -1$ and $(1-u)/b \geq 1$,

$$\int_0^1 K_b(y-u)(y-u) dy = b \int_0^1 \frac{1}{b} K\left(\frac{y-u}{b}\right) \left(\frac{y-u}{b}\right) dy = b \int_{-1}^1 K(z) z dz = 0.$$

This can be used to derive for the first summand in (6.9) with the same arguments as used in (6.4)

$$\begin{aligned} & \varphi^{(1)}(u, \underline{s}) \cdot \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\frac{t}{T} - u\right) \\ &= \varphi^{(1)}(u, \underline{s}) \cdot \left(\frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\frac{t}{T} - u\right) - \int_0^1 K_b(y-u)(y-u) dy \right) \\ &= \varphi^{(1)}(u, \underline{s}) \cdot \sum_{t=1}^T \int_{(t-1)/T}^{t/T} \left(K\left(\frac{t/T-u}{b}\right) \left(\frac{t/T-u}{b}\right) - K\left(\frac{y-u}{b}\right) \left(\frac{y-u}{b}\right) \right) dy \\ &= \mathcal{O}(T^{-1}), \end{aligned}$$

since each of the integrals can be bounded by $C/(bT^2)$ as above (note that $K((y-u)/b) \cdot ((y-u)/b)$ is Lipschitz since K is zero outside $[-1, 1]$). Using this bound in (6.9) and applying (6.7) to (6.5) shows that the bias term (6.3) is equal to

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \frac{1}{2} \varphi^{(2)}(\tilde{u}, \underline{s}) \left(\frac{t}{T} - u\right)^2 + \mathcal{O}\left(\frac{1}{bT}\right) \\ &= \frac{1}{2} \varphi^{(2)}(u, \underline{s}) \cdot \int_0^1 K_b(y-u)(y-u)^2 dy + \mathcal{O}\left(\frac{1}{bT}\right) \\ & \quad + \frac{1}{2} \varphi^{(2)}(u, \underline{s}) \left(\frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\frac{t}{T} - u\right)^2 - \int_0^1 K_b(y-u)(y-u)^2 dy \right) \quad (6.10) \\ & \quad + \frac{1}{2T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\frac{t}{T} - u\right)^2 (\varphi^{(2)}(\tilde{u}, \underline{s}) - \varphi^{(2)}(u, \underline{s})). \quad (6.11) \end{aligned}$$

The expression in (6.10) can be bounded in absolute value by

$$\begin{aligned} & \frac{1}{2} |\varphi^{(2)}(u, \underline{s})| \cdot b \sum_{t=1}^T \int_{(t-1)/T}^{t/T} \left| K\left(\frac{t/T-u}{b}\right) \left(\frac{t/T-u}{b}\right)^2 - K\left(\frac{y-u}{b}\right) \left(\frac{y-u}{b}\right)^2 \right| dy \\ &= \mathcal{O}\left(\frac{b}{T}\right) \quad (6.12) \end{aligned}$$

with the same arguments as used before. Lemma 2.2 (ii) ensures that $\varphi^{(3)}(\cdot, \underline{s})$ is continuous and therefore bounded on $[0, 1]$. Hence, using the mean value theorem, expression (6.11) can be bounded in absolute value by

$$\begin{aligned} & \sup_{w \in [0, 1]} |\varphi^{(3)}(w, \underline{s})| \cdot \frac{1}{2T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left(\frac{t}{T} - u\right)^2 |\tilde{u} - u| \\ & \leq \sup_{w \in [0, 1]} |\varphi^{(3)}(w, \underline{s})| \cdot \frac{1}{2} \cdot \frac{1}{T} \sum_{t=1}^T K_b\left(\frac{t}{T} - u\right) \left| \frac{t}{T} - u \right|^3 = \mathcal{O}(b^3), \end{aligned}$$

due to (2.8). Inserting the derived bounds into (6.10) and (6.11) yields that the bias term equals

$$\frac{1}{2}\varphi^{(2)}(u, \underline{s}) \int_0^1 K_b(y-u)(y-u)^2 dy + \mathcal{O}\left(\frac{1}{bT}\right) + \mathcal{O}(b^3). \quad (6.13)$$

By substituting as before we get

$$\int_0^1 K_b(y-u)(y-u)^2 dy = b \int_0^1 K\left(\frac{y-u}{b}\right) \left(\frac{y-u}{b}\right)^2 dy = b^2 \int_{-1}^1 K(z) z^2 dz$$

and, with $\beta(u, \underline{s})$ as defined in Lemma 3.1, (6.13) yields

$$E(\widehat{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) = b^2 \cdot \beta(u, \underline{s}) + \mathcal{O}((bT)^{-1}) + \mathcal{O}(b^3),$$

which finishes the proof for (ii).

(iii) *Covariances:*

For the covariances we show

$$bT \text{Cov}(\widehat{\varphi}(u_1, \underline{s}_1), \widehat{\varphi}(u_2, \underline{s}_2)) = V((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o(1). \quad (6.14)$$

Using the notation $g_{\underline{s}}(\underline{x}) := \exp(i\langle \underline{s}, \underline{x} \rangle)$ we immediately get

$$\begin{aligned} & bT \text{Cov}(\widehat{\varphi}(u_1, \underline{s}_1), \widehat{\varphi}(u_2, \underline{s}_2)) \\ &= \frac{b}{T} \sum_{t_1, t_2=1}^T K_b\left(\frac{t_1}{T} - u_1\right) K_b\left(\frac{t_2}{T} - u_2\right) \text{Cov}(g_{\underline{s}_1}(\underline{X}_{t_1, T}), g_{\underline{s}_2}(\underline{X}_{t_2, T})). \end{aligned} \quad (6.15)$$

In order to replace $\underline{X}_{t_i, T}$ by $\widetilde{\underline{X}}_{t_i}(\frac{t_i}{T})$ on the right-hand side, we first show that

$$\text{Cov}(g_{\underline{s}_1}(\underline{X}_{t_1, T}), g_{\underline{s}_2}(\underline{X}_{t_2, T})) = \text{Cov}(g_{\underline{s}_1}(\widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})), g_{\underline{s}_2}(\underline{X}_{t_2, T})) + \mathcal{O}(T^{-1}) \quad (6.16)$$

uniformly for all t_1, t_2 . Using a Taylor expansion of first order for $g_{\underline{s}_1}$ we get

$$\begin{aligned} & \text{Cov}(g_{\underline{s}_1}(\underline{X}_{t_1, T}), g_{\underline{s}_2}(\underline{X}_{t_2, T})) \\ &= \text{Cov}(g_{\underline{s}_1}(\widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})), g_{\underline{s}_2}(\underline{X}_{t_2, T})) + \text{Cov}\left(\nabla g_{\underline{s}_1}(\underline{\xi})'(\underline{X}_{t_1, T} - \widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})), g_{\underline{s}_2}(\underline{X}_{t_2, T})\right), \end{aligned} \quad (6.17)$$

where $\underline{\xi}$ is between $\underline{X}_{t_1, T}$ and $\widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})$ and the gradient $\nabla g_{\underline{s}}(\cdot)$ fulfills

$$\nabla g_{\underline{s}}(\underline{x}) = (is_1 g_{\underline{s}}(\underline{x}), \dots, is_d g_{\underline{s}}(\underline{x}))' = i g_{\underline{s}}(\underline{x}) \cdot \underline{s}.$$

It holds $|g_{\underline{s}}(\underline{x})| = 1$ for all \underline{x} , which implies on the one hand $|\nabla g_{\underline{s}}(\underline{x})|_1 = |\underline{s}|_1$ and on the other hand that the second summand in (6.18) can be bounded in absolute value by

$$\begin{aligned} 2E \left| \nabla g_{\underline{s}_1}(\underline{\xi})'(\underline{X}_{t_1, T} - \widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})) \right| &\leq 2E \left(|\nabla g_{\underline{s}_1}(\underline{\xi})|_1 \cdot \left| \underline{X}_{t_1, T} - \widetilde{\underline{X}}_{t_1}(\frac{t_1}{T}) \right|_1 \right) \\ &\leq 2|\underline{s}_1|_1 \cdot \left\| \underline{X}_{t_1, T} - \widetilde{\underline{X}}_{t_1}(\frac{t_1}{T}) \right\|_1 = \mathcal{O}(T^{-1}) \end{aligned}$$

uniformly in t_1, t_2 , due to Lemma 2.1 (ii). This proves (6.16) and with exactly the same calculation the second argument can be replaced which yields

$$\text{Cov}(g_{\underline{s}_1}(\underline{X}_{t_1, T}), g_{\underline{s}_2}(\underline{X}_{t_2, T})) = \text{Cov}(g_{\underline{s}_1}(\widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})), g_{\underline{s}_2}(\widetilde{\underline{X}}_{t_2}(\frac{t_2}{T}))) + \mathcal{O}(T^{-1}) \quad (6.18)$$

uniformly in t_1, t_2 . Inserting this result into (6.15) gives

$$\begin{aligned} & bT \text{Cov}(\widehat{\varphi}(u_1, \underline{s}_1), \widehat{\varphi}(u_2, \underline{s}_2)) \\ &= \frac{b}{T} \sum_{t_1, t_2=1}^T K_b\left(\frac{t_1}{T} - u_1\right) K_b\left(\frac{t_2}{T} - u_2\right) \text{Cov}(g_{\underline{s}_1}(\widetilde{\underline{X}}_{t_1}(\frac{t_1}{T})), g_{\underline{s}_2}(\widetilde{\underline{X}}_{t_2}(\frac{t_2}{T}))) \end{aligned}$$

$$+ \mathcal{O}(T^{-1}) \cdot \frac{1}{bT} \sum_{t_1=1}^T K\left(\frac{t_1/T - u_1}{b}\right) \sum_{t_2=1}^T K\left(\frac{t_2/T - u_2}{b}\right).$$

The second summand on the right-hand side vanishes asymptotically with rate $\mathcal{O}(b)$ because of (2.8). Therefore, the desired assertion follows if we can show that

$$\frac{b}{T} \sum_{t_1, t_2=1}^T K_b\left(\frac{t_1}{T} - u_1\right) K_b\left(\frac{t_2}{T} - u_2\right) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_{t_1}\left(\frac{t_1}{T}\right)), g_{\underline{s}_2}(\tilde{X}_{t_2}\left(\frac{t_2}{T}\right))\right) = V((u_1, \underline{s}_1), (u_2, \underline{s}_2)) + o(1).$$

From stationarity of $(\tilde{X}_t(u))_t$, we get

$$\begin{aligned} & \frac{b}{T} \sum_{t_1, t_2=1}^T K_b\left(\frac{t_1}{T} - u_1\right) K_b\left(\frac{t_2}{T} - u_2\right) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_{t_1}\left(\frac{t_1}{T}\right)), g_{\underline{s}_2}(\tilde{X}_{t_2}\left(\frac{t_2}{T}\right))\right) \\ &= \frac{1}{bT} \sum_{h=-(T-1)}^{T-1} \sum_{t=\max\{1, 1-h\}}^{\min\{T, T-h\}} K\left(\frac{t+h}{b} - u_1\right) K\left(\frac{t}{b} - u_2\right) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h\left(\frac{t+h}{T}\right)), g_{\underline{s}_2}(\tilde{X}_0\left(\frac{t}{T}\right))\right). \end{aligned} \quad (6.19)$$

In the following, rather than the last right-hand side, we consider

$$\frac{1}{bT} \sum_{h=-(T-1)}^{T-1} \sum_{t=1}^T K\left(\frac{t+h}{b} - u_1\right) K\left(\frac{t}{b} - u_2\right) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h\left(\frac{t+h}{T}\right)), g_{\underline{s}_2}(\tilde{X}_0\left(\frac{t}{T}\right))\right) \quad (6.20)$$

with

$$\tilde{X}_v(z) = \begin{cases} \tilde{X}_v(1) & \text{if } z \geq 1 \\ \tilde{X}_v(0) & \text{if } z \leq 0 \end{cases}.$$

This can be justified by the fact that via Lemma 2.1 (i) the modulus of the difference of (6.20) and (6.20) can be bounded by

$$\begin{aligned} & \frac{1}{bT} \sum_{h=0}^{T-1} \sum_{t=T-h+1}^T K\left(\frac{t+h}{b} - u_1\right) K\left(\frac{t}{b} - u_2\right) \left| \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h\left(\frac{t+h}{T}\right)), g_{\underline{s}_2}(\tilde{X}_0\left(\frac{t}{T}\right))\right) \right| \\ & + \frac{1}{bT} \sum_{h=-(T-1)}^{-1} \sum_{t=1}^{-h} K\left(\frac{t+h}{b} - u_1\right) K\left(\frac{t}{b} - u_2\right) \left| \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h\left(\frac{t+h}{T}\right)), g_{\underline{s}_2}(\tilde{X}_0\left(\frac{t}{T}\right))\right) \right| \\ & \leq \frac{1}{bT} \sum_{h=-(T-1)}^{T-1} |h| C 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \sum_{|j| > \lfloor |h/2| \rfloor} \frac{B}{l(j)} \\ & \leq \frac{1}{bT} C 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \sum_{h=-(T-1)}^{T-1} \sum_{|j| > \lfloor |h/2| \rfloor} 2|j| \frac{B}{l(j)} \\ & \leq \frac{1}{bT} 16 C d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \sum_{j \in \mathbb{Z}} j^2 \frac{B}{l(j)}, \end{aligned}$$

since $K(\cdot)$ is bounded by a constant C . The right-hand side is asymptotically vanishing of order $\mathcal{O}((bT)^{-1})$ under the summability condition on $(l(j))_j$ stated in Assumption 2.1. In the next step, we replace the Riemann sum in (6.20) by its integral, i.e. we consider

$$\frac{1}{b} \sum_{h=-(T-1)}^{T-1} \int_0^1 K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h\left(y + \frac{h}{T}\right)), g_{\underline{s}_2}(\tilde{X}_0(y))\right) dy. \quad (6.21)$$

To be allowed to consider (6.21) in the following, we have to show that the difference of (6.20) and (6.21) vanishes asymptotically. By standard arguments, its modulus can be bounded by

$$\begin{aligned} & \frac{1}{b} \sum_{h=-(T-1)}^{T-1} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} \left| K\left(\frac{t+h}{T} - u_1\right) K\left(\frac{t}{T} - u_2\right) \text{Cov}\left(g_{s_1}(\tilde{X}_h(\frac{t+h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T}))\right) \right. \\ & \quad \left. - K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \text{Cov}\left(g_{s_1}(\tilde{X}_h(y + \frac{h}{T})), g_{s_2}(\tilde{X}_0(y))\right) \right| dy \end{aligned}$$

and by adding three zeros, it remains to show that

$$\begin{aligned} I &= \frac{1}{b} \sum_{h=-(T-1)}^{T-1} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} \left| K\left(\frac{t+h}{T} - u_1\right) - K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) \right| K\left(\frac{t}{T} - u_2\right) \\ & \quad \times \left| \text{Cov}\left(g_{s_1}(\tilde{X}_h(\frac{t+h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T}))\right) \right| dy, \\ II &= \frac{1}{b} \sum_{h=-(T-1)}^{T-1} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) \left| K\left(\frac{t}{T} - u_2\right) - K\left(\frac{y - u_2}{b}\right) \right| \\ & \quad \times \left| \text{Cov}\left(g_{s_1}(\tilde{X}_h(\frac{t+h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T}))\right) \right| dy, \\ III &= \frac{1}{b} \sum_{h=-(T-1)}^{T-1} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \\ & \quad \times \left| \text{Cov}\left(g_{s_1}(\tilde{X}_h(\frac{t+h}{T})) - g_{s_1}(\tilde{X}_h(y + \frac{h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T}))\right) \right| dy, \\ IV &= \frac{1}{b} \sum_{h=-(T-1)}^{T-1} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \\ & \quad \times \left| \text{Cov}\left(g_{s_1}(\tilde{X}_h(y + \frac{h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T})) - g_{s_2}(\tilde{X}_0(y))\right) \right| dy \end{aligned}$$

vanish asymptotically. By Lipschitz continuity of the kernel K , $b^2T \rightarrow \infty$ and the absolute summability of the covariances $\text{Cov}(g_{s_1}(\tilde{X}_h((t+h)/T)), g_{s_2}(\tilde{X}_0(t/T)))$, cf. Lemma 2.1 (i), the first two terms I and II are of order $\mathcal{O}((b^2T)^{-1})$ and vanish asymptotically. Hence, it remains to consider III and IV , where we focus on III only, as the arguments are completely analogous for IV . In order to consider $\lim_{T \rightarrow \infty} |III|$, note that

$$|III| \leq \sum_{h=-\infty}^{\infty} \mathbf{1}_{\{|h| \leq T-1\}} |f_T(h)|$$

with an obvious notation for $f_T(h)$. We want to apply Lebesgue's dominated convergence theorem and derive the following bound for $|f_T(h)|$ using Lemma 2.1 (i) and the fact that K is bounded by a constant C :

$$\begin{aligned} & |f_T(h)| \\ & \leq \frac{1}{b} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} C K\left(\frac{y - u_2}{b}\right) \\ & \quad \times \left(\left| \text{Cov}\left(g_{s_1}(\tilde{X}_h(\frac{t+h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T}))\right) \right| + \left| \text{Cov}\left(g_{s_1}(\tilde{X}_h(y + \frac{h}{T})), g_{s_2}(\tilde{X}_0(\frac{t}{T}))\right) \right| \right) dy \end{aligned}$$

$$\begin{aligned}
&\leq 2C 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \frac{1}{b} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} K\left(\frac{y-u_2}{b}\right) dy \sum_{|j| > \lfloor |h/2| \rfloor} \frac{B}{l(j)} \\
&\leq 4Cd(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \int_{-1}^1 K(z) dz \sum_{|j| > \lfloor |h/2| \rfloor} \frac{B}{l(j)}, \tag{6.22}
\end{aligned}$$

where the standard substitution $z = (y - u_2)/b$ was used. The bound on the right-hand side does not depend on T and is summable in h since

$$\sum_{h \in \mathbb{Z}} \sum_{|j| > \lfloor |h/2| \rfloor} \frac{B}{l(j)} \leq 4 \sum_{j \in \mathbb{Z}} |j| \frac{B}{l(j)} < \infty.$$

Hence, we can apply Lebesgue's dominated convergence theorem to III and obtain

$$\begin{aligned}
\lim_{T \rightarrow \infty} |III| &\leq \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{b} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \\
&\quad \times \left| \text{Cov}\left(g_{\underline{s}_1}\left(\tilde{X}_h\left(\frac{t+h}{T}\right)\right) - g_{\underline{s}_1}\left(\tilde{X}_h\left(y + \frac{h}{T}\right)\right), g_{\underline{s}_2}\left(\tilde{X}_0\left(\frac{t}{T}\right)\right)\right| dy \\
&\leq \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{b} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} C K\left(\frac{y - u_2}{b}\right) \\
&\quad \times 2E \left| g_{\underline{s}_1}\left(\tilde{X}_h\left(\frac{t+h}{T}\right)\right) - g_{\underline{s}_1}\left(\tilde{X}_h\left(y + \frac{h}{T}\right)\right) \right| dy, \tag{6.23}
\end{aligned}$$

because $|\text{Cov}(X, Y)| \leq 2E|X|$ if $|Y| = 1$. The expectation in (6.23) can be bounded using a Taylor expansion of $g_{\underline{s}_1}$ exactly as the one for the covariance term in (6.18). Thus, invoking Lemma 2.1 (iii), we get

$$\begin{aligned}
&E \left| g_{\underline{s}_1}\left(\tilde{X}_h\left(\frac{t+h}{T}\right)\right) - g_{\underline{s}_1}\left(\tilde{X}_h\left(y + \frac{h}{T}\right)\right) \right| \\
&= E \left| \nabla g_{\underline{s}_1}(\xi)' \left(\tilde{X}_h\left(\frac{t+h}{T}\right) - \tilde{X}_h\left(y + \frac{h}{T}\right) \right) \right| \\
&\leq |\underline{s}_1|_1 \cdot \left\| \tilde{X}_h\left(\frac{t+h}{T}\right) - \tilde{X}_h\left(y + \frac{h}{T}\right) \right\|_1 \\
&\leq |\underline{s}_1|_1 \left(Ld + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \cdot \left| \frac{t}{T} - y \right|, \tag{6.24}
\end{aligned}$$

where ξ is between $\tilde{X}_h((t+h)/T)$ and $\tilde{X}_h(y + (h/T))$. With the bound from (6.24) the expression in (6.23) can be bounded by

$$\begin{aligned}
&2C |\underline{s}_1|_1 \left(Ld + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \cdot \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{b} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} K\left(\frac{y - u_2}{b}\right) \left| \frac{t}{T} - y \right| dy \\
&\leq 2C |\underline{s}_1|_1 \left(Ld + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \cdot \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^1 \frac{1}{b} K\left(\frac{y - u_2}{b}\right) dy \\
&= 2C |\underline{s}_1|_1 \left(Ld + \|\underline{\varepsilon}_0\|_1 \cdot \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \cdot \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-1}^1 K(z) dz \\
&= 0.
\end{aligned}$$

Therefore, III vanishes asymptotically and, by the same arguments, it follows $IV = o(1)$. This allows us to consider (6.21) in the sequel.

We first consider the case $u_1 \neq u_2$. In order to obtain a preliminary result, consider the integral

$$\int_{-1}^1 K\left(z + \frac{h}{bT} + \frac{u_2 - u_1}{b}\right) K(z) dz$$

for a fixed $h \in \mathbb{Z}$. If $T \rightarrow \infty$, it follows $h/(bT) \rightarrow 0$ and $|(u_2 - u_1)/b| \rightarrow +\infty$. Therefore, there exists $T_0 \in \mathbb{N}$ such that for all $T \geq T_0$ it holds

$$\left|z + \frac{h}{bT} + \frac{u_2 - u_1}{b}\right| > 1 \quad \forall z \in [-1, 1].$$

Since K has support $[-1, 1]$, this immediately implies

$$\lim_{T \rightarrow \infty} \int_{-1}^1 K\left(z + \frac{h}{bT} + \frac{u_2 - u_1}{b}\right) K(z) dz = 0. \quad (6.25)$$

Now we turn back to (6.21). With the same calculations as in (6.22), we get for each $h \in \mathbb{Z}$ the bound

$$\begin{aligned} & \frac{1}{b} \int_0^1 K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \left| \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(y + \frac{h}{T})), g_{\underline{s}_2}(\tilde{X}_0(y))\right) \right| dy \\ & \leq C 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \int_{-1}^1 K(z) dz \sum_{j > \lfloor |h/2| \rfloor} \frac{B}{l(j)}, \end{aligned}$$

which is summable in h . This allows for the application of Lebesgue's dominated convergence theorem to (6.21) which, together with Lemma 2.1(i) and again $z = (y - u_2)/b$, yields

$$\begin{aligned} & \lim_{T \rightarrow \infty} \left| \frac{1}{b} \sum_{h=-(T-1)}^{T-1} \int_0^1 K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(y + \frac{h}{T})), g_{\underline{s}_2}(\tilde{X}_0(y))\right) dy \right| \\ & \leq \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{b} \int_0^1 K\left(\frac{y + \frac{h}{T} - u_1}{b}\right) K\left(\frac{y - u_2}{b}\right) \left| \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(y + \frac{h}{T})), g_{\underline{s}_2}(\tilde{X}_0(y))\right) \right| dy \\ & \leq 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\underline{\varepsilon}_0\|_1 \sum_{h=-\infty}^{\infty} \sum_{|j| > \lfloor |h/2| \rfloor} \frac{B}{l(j)} \lim_{T \rightarrow \infty} \int_{-1}^1 K\left(z + \frac{h}{bT} + \frac{u_2 - u_1}{b}\right) K(z) dz \\ & = 0, \end{aligned}$$

due to (6.25). This proves the first part of assertion (6.14), i.e.

$$bT \text{Cov}(\hat{\varphi}(u_1, \underline{s}_1), \hat{\varphi}(u_2, \underline{s}_2)) = o(1), \quad \forall u_1 \neq u_2.$$

In the second step we consider the case $u_1 = u_2$. Then, using the substitution $z = (y - u_1)/b$, (6.21) becomes

$$\begin{aligned} & \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 K\left(z + \frac{h}{bT}\right) K(z) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1 + bz + \frac{h}{T})), g_{\underline{s}_2}(\tilde{X}_0(u_1 + bz))\right) dz \\ & = \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 K^2(z) dz \cdot \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_1))\right) + R1 + R2 + R3, \quad (6.26) \end{aligned}$$

where

$$R1 = \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 \left(K\left(z + \frac{h}{bT}\right) - K(z)\right) K(z)$$

$$\begin{aligned}
& \times \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1 + bz + \frac{h}{T})), g_{\underline{s}_2}(\tilde{X}_0(u_1 + bz))\right) dz, \\
R2 &= \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 K^2(z) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1 + bz + \frac{h}{T})) - g_{\underline{s}_1}(\tilde{X}_h(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_1 + bz))\right) dz, \\
R3 &= \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 K^2(z) \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_1 + bz)) - g_{\underline{s}_2}(\tilde{X}_0(u_1))\right) dz.
\end{aligned}$$

From boundedness and Lipschitz continuity of K , as well as Lemma 2.1(i), $R1$ can be bounded in absolute value by

$$\begin{aligned}
& \mathcal{O}((bT)^{-1}) \sum_{h=-\infty}^{\infty} |h| 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\varepsilon_0\|_1 \sum_{|j| > \lfloor |h|/2 \rfloor} \frac{B}{l(j)} \\
& \leq \mathcal{O}((bT)^{-1}) 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\varepsilon_0\|_1 \sum_{h=-\infty}^{\infty} \sum_{|j| > \lfloor |h|/2 \rfloor} |j| \frac{B}{l(j)} \\
& = \mathcal{O}((bT)^{-1}) 2d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\varepsilon_0\|_1 \sum_{j \in \mathbb{Z}} j^2 \frac{B}{l(j)} \\
& = \mathcal{O}((bT)^{-1}).
\end{aligned}$$

Therefore, $R1$ vanishes asymptotically. For $R2$ we have

$$|R2| \leq \sum_{h=-\infty}^{\infty} \mathbb{1}_{\{|h| \leq T-1\}} |\tilde{f}_T(h)|$$

with an obvious notation for $\tilde{f}_T(h)$. Via Lemma 2.1(i) we have the bound

$$|\tilde{f}_T(h)| \leq \int_{-1}^1 K^2(z) dz \cdot 4d(|\underline{s}_1|_1 + |\underline{s}_2|_1) \|\varepsilon_0\|_1 \sum_{|j| > \lfloor |h|/2 \rfloor} \frac{B}{l(j)},$$

which is summable in h and not depending on T . Hence, we can apply Lebesgue's dominated convergence theorem which yields, again using the fact that $|\text{Cov}(X, Y)| \leq 2E|X|$ if $|Y| = 1$,

$$\lim_{T \rightarrow \infty} |R2| \leq \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \int_{-1}^1 K^2(z) 2E \left| g_{\underline{s}_1}(\tilde{X}_h(u_1 + bz + \frac{h}{T})) - g_{\underline{s}_1}(\tilde{X}_h(u_1)) \right| dz.$$

Following exactly the lines of (6.24), we can bound the right-hand side by

$$\begin{aligned}
& \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \int_{-1}^1 K^2(z) 2|\underline{s}_1|_1 \left(Ld + \|\varepsilon_0\|_1 \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \left| \frac{h}{T} + bz \right| dz \\
& \leq 2|\underline{s}_1|_1 \left(Ld + \|\varepsilon_0\|_1 \sum_{j=-\infty}^{\infty} \frac{Bd}{l(j)} \right) \int_{-1}^1 K^2(z) dz \sum_{h=-\infty}^{\infty} \lim_{T \rightarrow \infty} \left(\frac{|h|}{T} + |b| \right) \\
& = 0.
\end{aligned}$$

Therefore, $R2$ vanishes asymptotically and with exactly the same arguments, $R3$ also converges to zero. Applying these results to (6.26), it holds

$$\begin{aligned}
& bT \text{Cov}(\hat{\varphi}(u_1, \underline{s}_1), \hat{\varphi}(u_1, \underline{s}_2)) \\
& = \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 K^2(z) dz \cdot \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_1))\right) + o(1)
\end{aligned}$$

$$\begin{aligned}
&= \int_{-1}^1 K^2(z) dz \cdot \sum_{h=-\infty}^{\infty} \text{Cov}\left(g_{\underline{s}_1}(\tilde{X}_h(u_1)), g_{\underline{s}_2}(\tilde{X}_0(u_1))\right) + o(1) \\
&= V((u_1, \underline{s}_1), (u_1, \underline{s}_2)) + o(1).
\end{aligned} \tag{6.27}$$

The assertion for the real and imaginary parts follows immediately from $\Re x = (x + \bar{x})/2$ and $\Im x = (x - \bar{x})/(2i)$ and $\exp(ix) = \cos(x) + i \sin(x)$ which completes the proof of part (iii). \square

Proof of Theorem 3.1. For any $u \in (0, 1)$ and $\underline{s} \in [-S, S]^d$, we can write

$$\sqrt{bT} (\hat{\varphi}(u, \underline{s}) - \varphi(u, \underline{s})) = \sqrt{bT} (\hat{\varphi}(u, \underline{s}) - E[\hat{\varphi}(u, \underline{s})]) + \sqrt{bT} (E[\hat{\varphi}(u, \underline{s})] - \varphi(u, \underline{s})), \tag{6.28}$$

where the second summand on the right-hand side above converges to zero if $b^3T \rightarrow 0$ holds and to $C\beta(u, \underline{s})$ if $b^5T \rightarrow C^2$ holds, respectively, by Lemma 3.1(i,ii). Hence, it remains to show the CLT in (3.6). By Cramér-Wold device, this is equivalent to show for all $\underline{c} \in \mathbb{R}^{2J}$ the corresponding CLT

$$Z_T := \sqrt{bT} \underline{c}' \left(\left(\begin{array}{c} \Re(\hat{\varphi}(u_j, \underline{s}_j) - E(\hat{\varphi}(u_j, \underline{s}_j))) \\ \Im(\hat{\varphi}(u_j, \underline{s}_j) - E(\hat{\varphi}(u_j, \underline{s}_j))) \end{array} \right) \right)_{j=1, \dots, J} \xrightarrow{d} \mathcal{N}(0, \underline{c}' \mathbf{V} \underline{c}).$$

Now, we have to distinguish the two cases of a positive variance $\underline{c}' \mathbf{V} \underline{c}$ and a vanishing variance $\underline{c}' \mathbf{V} \underline{c} = 0$. In the latter case, we get from Lemma 3.1(iii) that $\text{Var}(Z) \rightarrow \underline{c}' \mathbf{V} \underline{c} = 0$ holds such that

$$Z_T \xrightarrow{d} \mathcal{N}(0, \underline{c}' \mathbf{V} \underline{c}) = \mathcal{N}(0, 0) = \delta_0$$

is degenerate. Now, suppose $\underline{c}' \mathbf{V} \underline{c} > 0$ holds. Further, let $(\underline{X}_{t,T}^{(M)})$ be the truncated version of $(\underline{X}_{t,T})$, i.e.

$$\underline{X}_{t,T}^{(M)} = \underline{\mu}\left(\frac{t}{T}\right) + \sum_{j=-M}^M A_{t,T}(j) \varepsilon_{t-j},$$

and define

$$Z_T^{(M)} := \sqrt{bT} \underline{c}' \left(\left(\begin{array}{c} \Re(\hat{\varphi}^{(M)}(u_j, \underline{s}_j) - E(\hat{\varphi}^{(M)}(u_j, \underline{s}_j))) \\ \Im(\hat{\varphi}^{(M)}(u_j, \underline{s}_j) - E(\hat{\varphi}^{(M)}(u_j, \underline{s}_j))) \end{array} \right) \right)_{j=1, \dots, J},$$

where $\hat{\varphi}^{(M)}(u, \underline{s})$ is defined analogue to $\hat{\varphi}(u, \underline{s})$, but with $\underline{X}_{t,T}$ replaced by $\underline{X}_{t,T}^{(M)}$. Further, we define the block covariance matrix \mathbf{V}_M analogous to \mathbf{V} , but based on

$$V_M((u_1, \underline{s}_1), (u_2, \underline{s}_2)) = \int_{-1}^1 K^2(x) dx \cdot \sum_{h=-M}^M \text{Cov}(\exp(i \langle \underline{s}_1, \tilde{X}_0(u_1) \rangle), \exp(i \langle \underline{s}_2, \tilde{X}_h(u_2) \rangle))$$

instead of $V((u_1, \underline{s}_1), (u_2, \underline{s}_2))$. Now, we can make use of Proposition 6.3.9 in Brockwell and Davis (1991) and we have to show

- (a) $\exists M_0 \in \mathbb{N} \forall M \geq M_0 : Z_T^{(M)} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \underline{c}' \mathbf{V}_M \underline{c})$ as $T \rightarrow \infty$,
- (b) $\underline{c}' \mathbf{V}_M \underline{c} \rightarrow \underline{c}' \mathbf{V} \underline{c}$ as $M \rightarrow \infty$,
- (c) $\forall \delta > 0 : \lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} P(|Z_T - Z_T^{(M)}| \geq \delta) = 0$.

First, we get immediately from Lemma 2.1 (i) that $\underline{c}'\mathbf{V}_{M\underline{c}} \rightarrow \underline{c}'\mathbf{V}\underline{c} < \infty$ as $M \rightarrow \infty$, which proves (b). Now turn to part (a). By using completely the same arguments as used to compute the variance in Lemma 3.1(iii), we can show that

$$\text{Var}(Z_T^{(M)}) = \underline{c}'\mathbf{V}_{M\underline{c}} + o(1).$$

From $\underline{c}'\mathbf{V}\underline{c} > 0$ and as $\mathbf{V}_M \rightarrow \mathbf{V}$ for $M \rightarrow \infty$, we get also $\underline{c}'\mathbf{V}_{M\underline{c}} > 0$ for all $M \geq M_0$ and M_0 sufficient large. Hence, it suffices to show that

$$\frac{Z_T^{(M)}}{\sqrt{\text{Var}(Z_T^{(M)})}} \xrightarrow{d} \mathcal{N}(0, 1) \text{ as } T \rightarrow \infty. \quad (6.29)$$

To prove this, we write

$$\begin{aligned} Z_T^{(M)} &= \sum_{t=1}^T \sqrt{\frac{b}{T}} \underline{c}' \left(K_b \left(\frac{t}{T} - u_j \right) \begin{pmatrix} \cos(\langle \underline{s}_j, \underline{X}_{t,T}^{(M)} \rangle) - E[\cos(\langle \underline{s}_j, \underline{X}_{t,T}^{(M)} \rangle)] \\ \sin(\langle \underline{s}_j, \underline{X}_{t,T}^{(M)} \rangle) - E[\sin(\langle \underline{s}_j, \underline{X}_{t,T}^{(M)} \rangle)] \end{pmatrix}, j = 1, \dots, J \right) \\ &= \sum_{t=1}^T Y_{t,T}^{(M)} \end{aligned}$$

with an obvious notation for $Y_{t,T}^{(M)}$, where we suppress the dependence on u_j and \underline{s}_j . Note that all summands in $Y_{t,T}^{(M)}$ with $|(t/T - u_j)/b| > 1$ are zero since the kernel K has compact support $[-1, 1]$. Consequently, as we consider u_1, \dots, u_J , at most $d_T = J(2\lfloor bT \rfloor + 1)$ (subsequent) of the summands $Y_{t,T}^{(M)}$ above fulfill $|(t/T - u_j)/b| \leq 1$ for at least one j . Let $Y_{t_1,T}^{(M)}, \dots, Y_{t_{d_T},T}^{(M)}$ denote these non-vanishing summands such that we can write

$$Z_T^{(M)} = \sum_{r=1}^{d_T} Y_{t_r,T}^{(M)}.$$

Note that $(Y_{t_r,T}^{(M)}, r = 1, \dots, d_T)$ forms a triangular array of centered $(2M)$ -dependent random variables such that we can use the CLT in Theorem 2.1 in Romano and Wolf (2000), which is tailor-made for m -dependent random variables. In the following, we adapt their notation and we have to check their Conditions (1) - (6). We refer to Romano and Wolf (2000) for details. Since

$$\begin{aligned} &E[|Y_{t_r,T}^{(M)}|^{2+\delta}] \\ &= E\left[\left| \sqrt{\frac{b}{T}} \underline{c}' \left(K_b \left(\frac{t_r}{T} - u_j \right) \begin{pmatrix} \cos(\langle \underline{s}_j, \underline{X}_{t_r,T}^{(M)} \rangle) - E[\cos(\langle \underline{s}_j, \underline{X}_{t_r,T}^{(M)} \rangle)] \\ \sin(\langle \underline{s}_j, \underline{X}_{t_r,T}^{(M)} \rangle) - E[\sin(\langle \underline{s}_j, \underline{X}_{t_r,T}^{(M)} \rangle)] \end{pmatrix}, j = 1, \dots, J \right) \right|^{2+\delta} \right] \\ &\leq C_\Delta (bT)^{-(1+\frac{\delta}{2})} \end{aligned}$$

for some finite constant C_Δ and any $\delta > 0$, we get that Condition (1) holds with $\Delta_T = C_\Delta (bT)^{-(1+\frac{\delta}{2})}$. As $|\text{Cov}(\cos(\langle \underline{s}_j, \underline{X}_{t_1,T}^{(M)} \rangle), \cos(\langle \underline{s}_j, \underline{X}_{t_2,T}^{(M)} \rangle))|, |\text{Cov}(\sin(\langle \underline{s}_j, \underline{X}_{t_1,T}^{(M)} \rangle), \sin(\langle \underline{s}_j, \underline{X}_{t_2,T}^{(M)} \rangle))|$ and $|\text{Cov}(\sin(\langle \underline{s}_j, \underline{X}_{t_1,T}^{(M)} \rangle), \cos(\langle \underline{s}_j, \underline{X}_{t_2,T}^{(M)} \rangle))|$ can always be bounded by 2, we get that $|\text{Cov}(Y_{t_{r_1}}^{(M)}, Y_{t_{r_2}}^{(M)})| = O((bT)^{-1})$. Hence, with $\gamma = 0$ in their notation, we have for all a and all $k \geq 1$ that

$$\frac{1}{k} \text{Var} \left(\sum_{r=a}^{a+k-1} Y_{t_r,T}^{(M)} \right) = \frac{1}{k} \sum_{r_1, r_2=1}^k \text{Cov}(Y_{t_{r_1+a-1},T}^{(M)}, Y_{t_{r_2+a-1},T}^{(M)})$$

$$\begin{aligned}
&= \sum_{h=\max(-(k-1), -2M)}^{\min(k-1, 2M)} \frac{1}{k} \sum_{s=\max(1, 1-h)}^{\min(k, k-h)} \text{Cov}(Y_{t_s+h+a-1, T}^{(M)}, Y_{t_s+a-1, T}^{(M)}) \\
&\leq C_K \cdot (bT)^{-1}
\end{aligned}$$

for some suitable positive and finite constant C_K . This means that Condition (2) holds with $K_T = C_K \cdot (bT)^{-1}$. Further, since $\underline{c}'\mathbf{V}_M\underline{c} > 0$ for all $M \geq M_0$ and M_0 sufficiently large, and as $\text{Var}\left(\sum_{r=1}^{d_T} Y_{t_r, T}^{(M)}\right) = \underline{c}'\mathbf{V}_M\underline{c} + o(1)$, we obtain for sufficiently large T , that there exists a strictly positive and finite constant C_L such that

$$\frac{1}{d_T} \text{Var}\left(\sum_{r=1}^{d_T} Y_{t_r, T}^{(M)}\right) \geq C_L \cdot (bT)^{-1}.$$

This means that Condition (3) holds with $L_T = C_L \cdot (bT)^{-1}$. Altogether, this leads to Conditions (4) and (5) being satisfied, i.e.

$$(4) \quad \frac{K_T}{L_T} = \frac{C_K(bT)^{-1}}{C_L(bT)^{-1}} = \frac{C_K}{C_L} = \mathcal{O}(1), \quad (5) \quad \frac{\Delta_T}{L_T^{1+\delta/2}} = \frac{C_\Delta (bT)^{-(1+\frac{\delta}{2})}}{(C_L(bT)^{-1})^{1+\delta/2}} = \mathcal{O}(1).$$

As their Condition (6) is trivially fulfilled as M is fixed here, this proves the CLT in (6.29) and completes part (a). Finally, to show (c), by Markov inequality, it suffices to consider $E\left(|Z_T - Z_T^{(M)}|^2\right)$ in more detail. To avoid lengthy notation, we treat only the case of $J = 1$ and consider only the real part. Similar to (6.27) in the proof of Lemma 3.1(iii), we get

$$\begin{aligned}
&bT \text{Var}\left(\Re\widehat{\varphi}(u, \underline{s}) - \Re\widehat{\varphi}^{(M)}(u, \underline{s})\right) \\
&= \sum_{h=-(T-1)}^{T-1} \int_{-1}^1 K^2(z) dz \text{Cov}\left(\Re g_{\underline{s}}(\widetilde{X}_h(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M)}(u)), \Re g_{\underline{s}}(\widetilde{X}_0(u)) - \Re g_{\underline{s}}(\widetilde{X}_0^{(M)}(u))\right) + o(1) \\
&=: I_{T, M} + o(1)
\end{aligned}$$

as $T \rightarrow \infty$, where $\widetilde{X}_t^{(M)}(u) = \underline{\mu}(u) + \sum_{j=-M}^M A(u, j) \varepsilon_{t-j}$ is the truncated version of $\widetilde{X}_t(u)$ and $g_{\underline{s}}(\underline{x}) := \exp(i\langle \underline{s}, \underline{x} \rangle)$. Hence, it suffices to prove

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} |I_{T, M}| = 0$$

by making use of Lebesgue's Theorem. By using the same approach as in the proof of Lemma 3.1(iii) to show the finiteness of the variance \mathbf{V} , the modulus of the covariances in $I_{T, M}$ above can be bounded by

$$\begin{aligned}
&\left| \text{Cov}\left(\Re g_{\underline{s}}(\widetilde{X}_h(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M)}(u)) - \left\{ \Re g_{\underline{s}}(\widetilde{X}_h^{(v)}(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M, v)}(u)) \right\}, \Re g_{\underline{s}}(\widetilde{X}_0(u)) - \Re g_{\underline{s}}(\widetilde{X}_0^{(M)}(u))\right) \right| \\
&+ \left| \text{Cov}\left(\Re g_{\underline{s}}(\widetilde{X}_h^{(v)}(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M, v)}(u)), g(\widetilde{X}_0(u)) - \Re g_{\underline{s}}(\widetilde{X}_0^{(M)}(u)) - \left\{ \Re g_{\underline{s}}(\widetilde{X}_0^{(v)}(u)) - \Re g_{\underline{s}}(\widetilde{X}_0^{(M, v)}(u)) \right\}\right) \right|,
\end{aligned}$$

where $\widetilde{X}_t^{(M, v)}(u) = \underline{\mu}(u) + \sum_{|j| < \min(M, v)} A(u, j) \varepsilon_{t-j}$ and $v = \lfloor |h/2| \rfloor$. The first summand above can be bounded by

$$\begin{aligned}
&4E\left(\left|\Re g_{\underline{s}}(\widetilde{X}_h(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M)}(u)) - \left\{ \Re g_{\underline{s}}(\widetilde{X}_h^{(v)}(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M, v)}(u)) \right\}\right|\right) \\
&\leq 4\left\{E\left(\left|\Re g_{\underline{s}}(\widetilde{X}_h(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(v)}(u))\right|\right) + E\left(\left|\Re g_{\underline{s}}(\widetilde{X}_h^{(M)}(u)) - \Re g_{\underline{s}}(\widetilde{X}_h^{(M, v)}(u))\right|\right)\right\}
\end{aligned}$$

$$\begin{aligned}
&\leq 4 \|\underline{s}\|_1 \|\underline{\varepsilon}_0\|_1 \left\{ \sum_{|j|>v} |A(u, j)|_1 + \sum_{M \geq |j| > \min(v, M)} |A(u, j)|_1 \right\} \\
&\leq 8 \|\underline{s}\|_1 \|\underline{\varepsilon}_0\|_1 \sum_{|j|>v} \sup_u |A(u, j)|_1 \\
&\leq 8d \|\underline{s}\|_1 \|\underline{\varepsilon}_0\|_1 \sum_{|j|>v} \frac{B}{l(j)}
\end{aligned}$$

and the second summand can be treated completely analogue. Altogether, the summands can be bounded by summable coefficients not depending on M , which allows to bound $\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} |I_{T, M}|$ applying Lebesgue's theorem as follows

$$\begin{aligned}
&\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} |I_{T, M}| \\
&\leq \lim_{M \rightarrow \infty} \lim_{T \rightarrow \infty} \sum_{h=-\infty}^{\infty} \int_{-1}^1 K^2(z) dz \left| \text{Cov} \left(\Re g_{\underline{s}}(\tilde{X}_h(u)) - \Re g_{\underline{s}}(\tilde{X}_h^{(M)}(u)), \Re g_{\underline{s}}(\tilde{X}_0(u)) - \Re g_{\underline{s}}(\tilde{X}_0^{(M)}(u)) \right) \right| \\
&\leq \sum_{h=-\infty}^{\infty} \int_{-1}^1 K^2(z) dz \left\{ \lim_{M \rightarrow \infty} \left| \text{Cov} \left(\Re g_{\underline{s}}(\tilde{X}_h(u)) - \Re g_{\underline{s}}(\tilde{X}_h^{(M)}(u)), \Re g_{\underline{s}}(\tilde{X}_0(u)) - \Re g_{\underline{s}}(\tilde{X}_0^{(M)}(u)) \right) \right| \right\} \\
&\leq 4 \sum_{h=-\infty}^{\infty} \int_{-1}^1 K^2(z) dz \left\{ \lim_{M \rightarrow \infty} E \left(\left| \Re g_{\underline{s}}(\tilde{X}_h(u)) - \Re g_{\underline{s}}(\tilde{X}_h^{(M)}(u)) \right| \right) \right\} \\
&\leq 4 \|\underline{s}\|_1 \|\underline{\varepsilon}_0\|_1 \sum_{h=-\infty}^{\infty} \int_{-1}^1 K^2(z) dz \left\{ \lim_{M \rightarrow \infty} \sum_{|j|>M} \sup_u |A(u, j)|_1 \right\} \\
&\leq 4d \|\underline{s}\|_1 \|\underline{\varepsilon}_0\|_1 \sum_{h=-\infty}^{\infty} \int_{-1}^1 K^2(z) dz \left\{ \lim_{M \rightarrow \infty} \sum_{|j|>M} \frac{B}{l(j)} \right\} \\
&= 0,
\end{aligned}$$

which proves part (c) and concludes this proof. \square

Proof of Lemma 3.2. Applying Lemma 2.1(ii) and (iii) we obtain

$$\begin{aligned}
&E \left(\sup_{s \in [-S, S]^d} \sqrt{bT} \left\{ \hat{\varphi}(u, \underline{s}) - E \hat{\varphi}(u, \underline{s}) - \frac{1}{T} \sum_{t=1}^T K_b \left(\frac{t}{T} - u \right) \left[\exp(i \langle \underline{s}, \tilde{X}_t(u) \rangle) - E \exp(i \langle \underline{s}, \tilde{X}_t(u) \rangle) \right] \right\} \right) \\
&\leq \frac{2S}{\sqrt{bT}} \sum_{t=1}^T K \left(\frac{t/T - u}{b} \right) \left\{ \left\| \underline{X}_{t, T} - \tilde{X}_t(t/T) \right\|_1 + \left\| \tilde{X}_t(t/T) - \tilde{X}_t(u) \right\|_1 \right\} \\
&\leq \mathcal{O} \left(\sqrt{\frac{b}{T}} \right) + \mathcal{O}(\sqrt{b^3 T}),
\end{aligned}$$

which vanishes asymptotically. \square

Proof of Lemma 3.3. We restrict ourselves to the first assertion since the second one can be derived in complete analogy. For notational simplicity, we define $\bar{g}_{\underline{s}}(\underline{X}) = \cos(\underline{s}' \underline{X}) - E \cos(\underline{s}' \underline{X})$ for any $\underline{s} \in \mathbb{R}^d$ and any \mathbb{R}^d -valued random variable \underline{X} . Recall that the number of nonvanishing summands is bounded by $[2bT] + 1$ (see end of Section 2) and that these summands are subsequent. Let t_1 denote the smallest of these indices. Inspired by Arcones, Yu (1994, Section 2), we divide this set of indices into blocks H_t, T_t (both of equal length

κ_T) for $t = 1, \dots, \mu_T$ and a remainder term R , where

$$\begin{aligned} H_t &:= \{i \mid 2(t-1)\kappa_T + t_1 \leq i < (2t-1)\kappa_T + t_1\}, \\ T_t &:= \{i \mid (2t-1)\kappa_T + t_1 \leq i < 2t\kappa_T + t_1\}, \\ R &:= \{i \mid 2\mu_T\kappa_T + t_1 \leq i < t_1 + \lceil 2bT \rceil\}. \end{aligned}$$

Here κ_T and μ_T have the following form for all $T \in \mathbb{N}$:

$$\kappa_T := \left\lfloor (bT)^{\frac{1}{2m}} \right\rfloor \quad \text{and} \quad \mu_T := \left\lfloor \frac{bT}{\kappa_T} \right\rfloor. \quad (6.30)$$

Further, let $\rho(\underline{s}_1, \underline{s}_2) = \|\underline{s}_1 - \underline{s}_2\|_1$. This gives

$$\begin{aligned} & \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| (bT)^{-\frac{1}{2}} \sum_{t=1}^{bT} K \left(\frac{t/T - u}{b} \right) [\bar{g}_{\underline{s}_1}(\tilde{X}_t(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_t(u))] \right| > \lambda \right) \\ & \leq \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in H_t} K \left(\frac{i/T - u}{b} \right) [\bar{g}_{\underline{s}_1}(\tilde{X}_i(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i(u))] \right| > \frac{\lambda}{3} \right) \\ & \quad + \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in T_t} K \left(\frac{i/T - u}{b} \right) [\bar{g}_{\underline{s}_1}(\tilde{X}_i(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i(u))] \right| > \frac{\lambda}{3} \right) \\ & \quad + \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| (bT)^{-\frac{1}{2}} \sum_{i \in R} K \left(\frac{i/T - u}{b} \right) [\bar{g}_{\underline{s}_1}(\tilde{X}_i(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i(u))] \right| > \frac{\lambda}{3} \right) \\ & =: \text{I} + \text{II} + \text{III}. \end{aligned} \quad (6.31)$$

As the second summand can be treated similarly to the first one, we focus on I and III in the sequel. For the third summand III of (6.31), we have

$$\begin{aligned} \text{III} & \leq \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \|K\|_\infty (bT)^{-\frac{1}{2}} \sum_{i \in R} |\bar{g}_{\underline{s}_1}(\tilde{X}_i(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i(u))| > \frac{\lambda}{3} \right) \\ & \leq \lim_{T \rightarrow \infty} \limsup P \left(C (bT)^{-1/2} \kappa_T > \frac{\lambda}{3} \right) \\ & = 0, \end{aligned}$$

as it follows from $m > 1$ that $\kappa_T = o((bT)^{1/2})$. In the next step we approximate the statistics under consideration by a statistic based on variables that are constructed such that the involved random variables with indices in different blocks H_1, \dots, H_{μ_T} are independent. To this end, we define

$$\tilde{X}_i^{(M)}(u) = \underline{\mu}(u) + \sum_{|k| \leq M} A(u, k) \varepsilon_{i-k}$$

for $M = \lfloor \kappa_T/2 \rfloor$. Using a first order Taylor expansion of \bar{g}_s , the first summand \mathbb{I} of (6.31) can be rewritten as follows

$$\begin{aligned}
\mathbb{I} &= \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in H_t} K \left(\frac{i/T - u}{b} \right) \left[\bar{g}_{\underline{s}_1}(\tilde{X}_i^{(M)}(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i^{(M)}(u)) \right. \right. \right. \\
&\quad \left. \left. \left. + \nabla \left(\bar{g}_{\underline{s}_1}(\underline{Y}_i^*) - \bar{g}_{\underline{s}_2}(\underline{Y}_i^*) \right) (\tilde{X}_i(u) - \tilde{X}_i^{(M)}(u)) \right] \right| > \frac{\lambda}{3} \right) \\
&\leq \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in H_t} K \left(\frac{i/T - u}{b} \right) \left[\bar{g}_{\underline{s}_1}(\tilde{X}_i^{(M)}(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i^{(M)}(u)) \right] \right| > \frac{\lambda}{6} \right) \\
&\quad + \limsup_{T \rightarrow \infty} P \left(2dS |K|_\infty (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in H_t} |\tilde{X}_i(u) - \tilde{X}_i^{(M)}(u)|_1 > \frac{\lambda}{6} \right) \\
&=: \lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} \mathbb{I}a + \limsup_{T \rightarrow \infty} \mathbb{I}b
\end{aligned} \tag{6.32}$$

for some \underline{Y}_i^* between $\tilde{X}_i(u)$ and $\tilde{X}_i^{(M)}(u)$. Now $\mathbb{I}b$ goes to 0 with $T \rightarrow \infty$ since Markov's inequality gives

$$\begin{aligned}
\limsup_{T \rightarrow \infty} \mathbb{I}b &\leq \limsup_{T \rightarrow \infty} \frac{12}{\lambda} dS |K|_\infty (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in H_t} \|\tilde{X}_i(u) - \tilde{X}_i^{(M)}(u)\|_1 \\
&= \limsup_{T \rightarrow \infty} C (bT)^{-\frac{1}{2}} \mu_T \sum_{i \in H_1} \sum_{|k| > M} \frac{1}{l(k)} \\
&\leq \limsup_{T \rightarrow \infty} C (bT)^{-\frac{1}{2}} \mu_T \kappa_T^{-m+1} \sum_{|k| > \frac{\kappa_T}{2}} \frac{k^m}{l(k)} \\
&= 0.
\end{aligned}$$

Hence, it remains to consider term $\mathbb{I}a$, which is more involved. Note that the random variables $(\tilde{X}_i^{(M)}(u))_{i \in H_{t_1}}$ and $(\tilde{X}_i^{(M)}(u))_{i \in H_{t_2}}$ are independent for $t_1 \neq t_2$, which will allow us to apply standard empirical process theory in the sequel. We adapt the arguments of Arcones, Yu (1994) and introduce some notation first. Let $\nu_T(\underline{s})$ and $\nu_T(\underline{s}_1, \underline{s}_2)$ with $\underline{s}, \underline{s}_1, \underline{s}_2 \in [-S, S]^d$ be defined as

$$\nu_T(\underline{s}) := (bT)^{-\frac{1}{2}} \sum_{t=1}^{\mu_T} \sum_{i \in H_t} K \left(\frac{i/T - u}{b} \right) \bar{g}_{\underline{s}}(\tilde{X}_i^{(M)}(u)) \quad \text{and} \quad \nu_T(\underline{s}_1, \underline{s}_2) := \nu_T(\underline{s}_1) - \nu_T(\underline{s}_2).$$

Now, we use a classical chaining argument inspired by Arcones, Yu (1994). To this end, we introduce a decreasing sequence

$$r_k := r2^{-k}, \quad k = 0, \dots, k_T,$$

for some r specified below and such that

$$(bT)^{3(1-m)/(4m)} \leq r_{k_T} \leq (bT)^{-1/(m\delta)}. \tag{6.33}$$

Note that for our choice of m such a k_T exists since for large T

$$(bT)^{3(1-m)/(4m)} \leq (bT)^{-1/(m\delta)} \Leftrightarrow m \geq 1 + \frac{4}{3\delta}.$$

Let $\mathcal{F}_k \subset [-S, S]^d$ for $k \in \mathbb{N}_0$ be a collection of indices with

$$\#\mathcal{F}_k = D_k = D(r_k, [-S, S]^d, \rho) \quad \text{and} \quad \sup_{\underline{s}_1 \in [-S, S]^d} \min_{\underline{s}_2 \in \mathcal{F}_k} \rho(\underline{s}_1, \underline{s}_2) < r_k,$$

where $D(u, [-S, S]^d, \rho) = \max\{\#\mathcal{T}_0 \mid \mathcal{T}_0 \subseteq [-S, S]^d, \rho(\underline{s}_1, \underline{s}_2) > u \ \forall \underline{s}_1 \neq \underline{s}_2 \in [-S, S]^d\}$ denotes the usual packing number defined e.g. in van der Vaart and Wellner (2000, Definition 2.2.3). Obviously, for any $u > 0$

$$D(u, [-S, S]^d, \rho) \leq \left(\frac{2Sd}{u} + 1\right)^d$$

and hence, $D_k = \mathcal{O}(r_k^{-d})$. Then there exists maps $\pi_k: [-S, S]^d \rightarrow \mathcal{F}_k$ such that

$$|\underline{s} - \pi_k \underline{s}|_1 \leq r_k \quad \forall \underline{s} \in [-S, S]^d.$$

Thereby, we have

$$\begin{aligned} & \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} |\nu_T(\underline{s}_1, \underline{s}_2)| \\ & \leq \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \{|\nu_T(\underline{s}_1) - \nu_T(\underline{s}_2) - \nu_T(\pi_{k_T} \underline{s}_1) + \nu_T(\pi_{k_T} \underline{s}_2)| + |\nu_T(\pi_0 \underline{s}_1) - \nu_T(\pi_0 \underline{s}_2)|\} \\ & \quad + \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} \left| \sum_{k=1}^{k_T} \nu_T(\pi_k \underline{s}_1) - \nu_T(\pi_{k-1} \underline{s}_1) - \nu_T(\pi_k \underline{s}_2) + \nu_T(\pi_{k-1} \underline{s}_2) \right| \\ & \leq 2 \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) \leq r_{k_T}}} |\nu_T(\underline{s}_1, \underline{s}_2)| + \sup_{\substack{\underline{s}_1, \underline{s}_2 \in \mathcal{F}_0 \\ \rho(\underline{s}_1, \underline{s}_2) \leq 2r}} |\nu_T(\underline{s}_1, \underline{s}_2)| + 2 \sum_{k=1}^{k_T} \sup_{\substack{\underline{s}_1 \in \mathcal{F}_k, \underline{s}_2 \in \mathcal{F}_{k-1} \\ \rho(\underline{s}_1, \underline{s}_2) \leq 3r_k}} |\nu_T(\underline{s}_1, \underline{s}_2)|. \end{aligned}$$

Moreover, we define for $k \in \mathbb{N}$ and for some C^* specified below

$$\lambda_k := r_k^{\frac{1}{6}} \vee \left(C^* r_k^{\frac{1}{3}} (\log D_k)^{\frac{1}{2}} \right) \quad (6.34)$$

and let r be small enough to ensure

$$2 \sum_{k=1}^{\infty} \lambda_k \leq \frac{\lambda}{18}. \quad (6.35)$$

Note that $D_k = \mathcal{O}(r_k^{-d})$ guaranties summability of $(\lambda_k)_k$. Hence, we have for Ia in (6.32)

$$\begin{aligned}
\underline{\text{Ia}} &= P \left\{ \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r}} |\nu_T(\underline{s}_1, \underline{s}_2)| > \frac{\lambda}{6} \right\} \\
&\leq P \left\{ 2 \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) \leq r_{k_T}}} |\nu_T(\underline{s}_1, \underline{s}_2)| > \frac{\lambda}{18} \right\} + P \left\{ 2 \sum_{k=1}^{k_T} \sup_{\substack{\underline{s}_1 \in \mathcal{F}_k, \underline{s}_2 \in \mathcal{F}_{k-1} \\ \rho(\underline{s}_1, \underline{s}_2) \leq 3r_k}} |\nu_T(\underline{s}_1, \underline{s}_2)| > 2 \sum_{k=1}^{k_T} \lambda_k \right\} \\
&\quad + P \left\{ \sup_{\substack{\underline{s}_1, \underline{s}_2 \in \mathcal{F}_0 \\ \rho(\underline{s}_1, \underline{s}_2) \leq 2r}} |\nu_T(\underline{s}_1, \underline{s}_2)| > \frac{\lambda}{18} \right\} \\
&=: \text{I} + \text{II} + \text{III}.
\end{aligned} \tag{6.36}$$

Hence, it remains to show asymptotic negligibility of I, II, and III in order to prove the lemma. We start with the second summand and apply Bernstein's inequality for sums of independent random variables to the outer sum in the definition of ν_T :

$$\begin{aligned}
\text{II} &\leq \sum_{k=1}^{k_T} P \left(\sup_{\substack{\underline{s}_1 \in \mathcal{F}_k, \underline{s}_2 \in \mathcal{F}_{k-1} \\ \rho(\underline{s}_1, \underline{s}_2) \leq 3r_k}} |\nu_T(\underline{s}_1, \underline{s}_2)| > \lambda_k \right) \\
&\leq \sum_{k=1}^{k_T} \sum_{\substack{\underline{s}_1 \in \mathcal{F}_k, \underline{s}_2 \in \mathcal{F}_{k-1} \\ \rho(\underline{s}_1, \underline{s}_2) \leq 3r_k}} P \left(\left| (bT)^{\frac{1}{2}} \nu_T(\underline{s}_1, \underline{s}_2) \right| > (bT)^{\frac{1}{2}} \lambda_k \right) \\
&\leq 2 \sum_{k=1}^{k_T} D_k D_{k-1} \exp \left\{ -\frac{1}{2} \frac{bT \lambda_k^2}{A_{II,k} + 4|K|_{\infty} \kappa_T \frac{(bT)^{\frac{1}{2}} \lambda_k}{3}} \right\} \\
&\leq 2 \sum_{k=1}^{k_T} \exp \left\{ 2 \log D_k - \frac{1}{2} \frac{bT \lambda_k^2}{A_{II,k} + 4|K|_{\infty} \kappa_T \frac{(bT)^{\frac{1}{2}} \lambda_k}{3}} \right\}.
\end{aligned} \tag{6.37}$$

Here,

$$A_{II,k} \geq \sup_{\substack{\underline{s}_1 \in \mathcal{F}_k, \underline{s}_2 \in \mathcal{F}_{k-1} \\ \rho(\underline{s}_1, \underline{s}_2) \leq 3r_k}} \text{Var} \left((bT)^{\frac{1}{2}} \nu_T(\underline{s}_1, \underline{s}_2) \right).$$

For the variance term, we obtain

$$\begin{aligned}
&bT \text{Var} \left(\nu_T(\underline{s}_1, \underline{s}_2) \right) \\
&\leq \text{Var} \left(\sum_{t=1}^{\mu_T} \sum_{i \in H_t} K \left(\frac{i/T - u}{b} \right) [\bar{g}_{\underline{s}_1}(\tilde{X}_i^{(M)}(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_i^{(M)}(u))] \right) \\
&\leq C \mu_T \sum_{i_1, i_2 \in H_1} \text{Cov} \left[\left(\bar{g}_{\underline{s}_1}(\tilde{X}_{i_1}^{(M)}(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_{i_1}^{(M)}(u)) \right), \left(\bar{g}_{\underline{s}_1}(\tilde{X}_{i_2}^{(M)}(u)) - \bar{g}_{\underline{s}_2}(\tilde{X}_{i_2}^{(M)}(u)) \right) \right].
\end{aligned} \tag{6.38}$$

In analogy to the proof of Lemma 2.1(i) we can bound the latter term by

$$bT \text{Var} \left(\nu_T(\underline{s}_1, \underline{s}_2) \right) \leq C \mu_T \sum_{i_1, i_2 \in H_1} \min \left\{ |\underline{s}_1 - \underline{s}_2|_1, \sum_{j > \lfloor |i_1 - i_2|/2 \rfloor} \frac{1}{l(j)} \right\}$$

$$\begin{aligned}
&\leq C \mu_T \sum_{t=0}^{\kappa_T-1} (\kappa_T - t) \min \left\{ |\underline{s}_1 - \underline{s}_2|_1, \sum_{j>[t/2]} \frac{1}{l(j)} \right\} \\
&\leq C bT \left[\sum_{t=0}^{R_0-1} |\underline{s}_1 - \underline{s}_2|_1 + \sum_{t=R_0}^{\kappa_T-1} t^{-3} \right] \\
&\leq C bT \left[R_0 |\underline{s}_1 - \underline{s}_2|_1 + \int_{R_0/2}^{\infty} u^{-3} du \right] \\
&\leq C bT \left[R_0 |\underline{s}_1 - \underline{s}_2|_1 + R_0^{-2} \right]
\end{aligned}$$

for any $R_0 \geq 2$. This holds for

$$R_0 := \left\lfloor |\underline{s}_1 - \underline{s}_2|_1^{-\frac{1}{3}} \right\rfloor,$$

if r is sufficiently small. Then we have

$$\sup_{\substack{\underline{s}_1 \in \mathcal{F}_k, \underline{s}_2 \in \mathcal{F}_{k-1} \\ \rho(\underline{s}_1, \underline{s}_2) \leq 3r_k}} \text{Var} \left((bT)^{\frac{1}{2}} \nu_T(\underline{s}_1, \underline{s}_2) \right) \leq C bT r_k^{2/3} =: A_{II,k}. \quad (6.39)$$

Hence, we obtain from (6.37) and (6.39)

$$\begin{aligned}
\Pi &\leq 2 \sum_{k=1}^{k_T} \exp \left\{ 2 \log D_k - \frac{1}{2} \frac{bT \lambda_k^2}{A_{II,k} + 4 \|K\|_{\infty} \kappa_T \frac{(bT)^{\frac{1}{2}} \lambda_k}{3}} \right\} \\
&\leq 2 \sum_{k=1}^{k_T} \exp \left\{ 2 \log D_k - C \frac{\lambda_k^2}{r_k^{2/3} + (bT)^{(1-m)/(2m)} \lambda_k} \right\} \\
&\leq 2 \sum_{k=1}^{k_T} \exp \left\{ 2 \log D_k - C \frac{\lambda_k^2}{r_k^{2/3} + r_{k_T}^{2/3} \lambda_k} \right\} \\
&\leq 2 \sum_{k=1}^{k_T} \exp \left\{ 2 \log D_k - C \frac{\lambda_k^2}{r_k^{2/3} + r_k^{2/3} \lambda} \right\} \\
&\leq 2 \sum_{k=1}^{k_T} \exp \left\{ 2 \log D_k - \bar{C} \lambda_k^2 r_k^{-2/3} \right\}
\end{aligned}$$

for some $\bar{C} \in (0, \infty)$. Setting $C^* = (4/\bar{C})^{1/2}$ in definition (6.34) of λ_k yields

$$\Pi \leq 2 \sum_{k=1}^{k_T} \exp \left\{ -\frac{\bar{C}}{2} \lambda_k^2 r_k^{-2/3} \right\} \leq 2 \sum_{k=1}^{\infty} \exp \left\{ -\frac{\bar{C}}{2} r_k^{-1/3} \right\} \xrightarrow{r \rightarrow 0} 0.$$

Analogously, we have for summand III in (6.36)

$$P \left\{ \sup_{\substack{\underline{s}_1, \underline{s}_2 \in \mathcal{F}_0 \\ \rho(\underline{s}_1, \underline{s}_2) \leq 2r}} |\nu_T(\underline{s}_1, \underline{s}_2)| > \lambda \right\} \leq 2 \exp \left\{ 2 \log D_0 - \frac{1}{2} \frac{bT \lambda^2}{A_{III,k} + 4 \|K\|_{\infty} \kappa_T \frac{(bT)^{\frac{1}{2}} \lambda}{3}} \right\},$$

where $A_{III,k} = C bT r^{2/3}$. Using the same arguments as before, we get asymptotic negligibility as $r \rightarrow 0$.

Finally, we look at the first summand I in (6.36). Applying Markov's inequality, it suffices to verify

$$\lim_{r \rightarrow 0} \limsup_{T \rightarrow \infty} E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{k_T}}} |\nu_T(\underline{s}_1, \underline{s}_2)| \right) = 0, \quad (6.40)$$

which then implies convergence of I and completes the proof. To this end, we introduce some further notation:

$$L_{t,T}(\underline{s}) := (bT)^{-\frac{1}{2}} \sum_{i \in H_t} K \left(\frac{i/T - u}{b} \right) \cos(\underline{s}' \tilde{X}_i^{(M)}(u)) \quad \text{and} \quad L_{t,T}^0(\underline{s}) := \zeta_t L_{t,T}(\underline{s})$$

for $\underline{s} \in [-S, S]^d$, where are $(\zeta_t)_{t \in \mathbb{N}}$ are i.i.d. Rademacher variables independent of $(\varepsilon_t)_{t \in \mathbb{Z}}$. Note that $(L_{t,T}(\underline{s}))_t$ is a sequence of independent random variables by construction. Hence, with a standard symmetrization lemma (see van der Vaart and Wellner (2000), Lemma 2.3.1) we get

$$\begin{aligned} & E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{k_T}}} |\nu_T(\underline{s}_1, \underline{s}_2)| \right) \\ &= E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{k_T}}} \left| \sum_{t=1}^{\mu_T} [L_{t,T}(\underline{s}_1) - EL_{t,T}(\underline{s}_1) - L_{t,T}(\underline{s}_2) + EL_{t,T}(\underline{s}_2)] \right| \right) \\ &\leq 2E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{k_T}}} \left| \sum_{t=1}^{\mu_T} [L_{t,T}^0(\underline{s}_1) - L_{t,T}^0(\underline{s}_2)] \right| \right). \end{aligned}$$

Note that $\sum_{t=1}^{\mu_T} L_{t,T}^0$ has sub-Gaussian increments conditionally on $L_{1,T}, \dots, L_{\mu_T, T}$, $T \in \mathbb{N}$. This is because for $\underline{s}_1, \underline{s}_2 \in [-S, S]^d$ and $\eta > 0$ we can apply Hoeffding's inequality to obtain

$$\begin{aligned} & P \left(\left| \sum_{t=1}^{\mu_T} L_{t,T}^0(\underline{s}_1) - L_{t,T}^0(\underline{s}_2) \right| > \hat{\rho}_{T,2}(\underline{s}_1, \underline{s}_2) \eta \mid L_{1,T}, \dots, L_{\mu_T, T} \right) \\ &\leq 2 \exp \left\{ - \frac{\hat{\rho}_{T,2}(\underline{s}_1, \underline{s}_2)^2 \eta^2}{2 \sum_{t=1}^{\mu_T} (L_{t,T}(\underline{s}_1) - L_{t,T}(\underline{s}_2))^2} \right\} \\ &= 2 \exp \left\{ - \frac{\eta^2}{2} \right\} \end{aligned}$$

with the random semimetric

$$\hat{\rho}_{T,2}(\underline{s}_1, \underline{s}_2) := \sqrt{\sum_{t=1}^{\mu_T} (L_{t,T}(\underline{s}_1) - L_{t,T}(\underline{s}_2))^2}, \quad \underline{s}_1, \underline{s}_2 \in [-S, S]^d.$$

Therefore, we aim at applying a maximal inequality for sub-Gaussian processes to prove (6.40). It turns out to be easier to use a slightly modified semimetric. To this end, first note that

$$\begin{aligned} (L_{t,T}(\underline{s}_1) - L_{t,T}(\underline{s}_2))^2 &\leq \left(|L_{t,T}(\underline{s}_1)| + |L_{t,T}(\underline{s}_2)| \right)^{1-\delta} |L_{t,T}(\underline{s}_1) - L_{t,T}(\underline{s}_2)|^{1+\delta} \\ &\leq 2^{1-\delta} |L_{t,T}|_{\infty}^{1-\delta} |L_{t,T}|_{\text{Lip}}^{1+\delta} \rho(\underline{s}_1, \underline{s}_2)^{1+\delta} \end{aligned} \quad (6.41)$$

for $\underline{s}_1, \underline{s}_2 \in [-S, S]^d$. Here, $|f|_{\text{Lip}}$ denotes the Lipschitz constant of a function f . We define

$$Q_T := 2^{(1-\delta)/2} \sqrt{\sum_{t=1}^{\mu_T} |L_{t,T}|_\infty^{1-\delta} |L_{t,T}|_{\text{Lip}}^{1+\delta}}$$

and in order to establish an upper bound for Q_T we derive

$$\|L_{t,T}\|_\infty = (bT)^{-\frac{1}{2}} \kappa_T \|K\|_\infty, \quad \text{and} \quad \|L_{t,T}\|_{\text{Lip}} \leq (bT)^{-\frac{1}{2}} \|K\|_\infty \sum_{i \in H_t} \left| \tilde{X}_i^{(M)}(u) \right|_1,$$

which give

$$Q_T \leq C \sqrt{\frac{\kappa_T^{1-\delta}}{bT} \sum_{t=1}^{\mu_T} \left(\sum_{i \in H_t} \left| \tilde{X}_i^{(M)}(u) \right|_1 \right)^{1+\delta}}.$$

Using the definition of Q_T we have

$$\hat{\rho}_{T,2}(\underline{s}_1, \underline{s}_2) \leq Q_T \rho(\underline{s}_1, \underline{s}_2)^{(1+\delta)/2} =: \check{\rho}_T(\underline{s}_1, \underline{s}_2).$$

Note that $\check{\rho}_T$ is again a random semimetric since we obtain from $(1+\delta)/2 \in (0, 1]$

$$\begin{aligned} \check{\rho}_T(\underline{s}_1, \underline{s}_2) &\leq Q_T (\rho(\underline{s}_1, \underline{s}_3) + \rho(\underline{s}_3, \underline{s}_2))^{(1+\delta)/2} \\ &= \check{\rho}_T(\underline{s}_1, \underline{s}_3) + \check{\rho}_T(\underline{s}_3, \underline{s}_2). \end{aligned}$$

Hence, we can apply Corollary 2.2.8 of van der Vaart and Wellner (2000) as follows

$$\begin{aligned} E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{k_T}}} \left| \sum_{t=1}^{\mu_T} [L_{t,T}^0(\underline{s}_1) - L_{t,T}^0(\underline{s}_2)] \right| \right) &= E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \check{\rho}_T(\underline{s}_1, \underline{s}_2) < Q_T r_{k_T}^{(1+\delta)/2}}} \left| \sum_{t=1}^{\mu_T} [L_{t,T}^0(\underline{s}_1) - L_{t,T}^0(\underline{s}_2)] \right| \right) \\ &\leq C E \left(\int_0^{Q_T r_{k_T}^{(1+\delta)/2}} \sqrt{\log D(u, [-S, S]^d, \check{\rho}_T)} \, du \right). \end{aligned}$$

The packing number can be calculated as

$$D(u, [-S, S]^d, \check{\rho}_T) = D\left(\left(\frac{u}{Q_T}\right)^{2/(1+\delta)}, [-S, S]^d, \rho\right) \leq \left(\frac{2Sd}{\left(\frac{u}{Q_T}\right)^{2/(1+\delta)}} + 1\right)^d,$$

which gives

$$\begin{aligned} E \left(\sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{k_T}}} \left| \sum_{t=1}^{\mu_T} [L_{t,T}^0(\underline{s}_1) - L_{t,T}^0(\underline{s}_2)] \right| \right) &\leq C E \left(\int_0^{Q_T r_{k_T}^{1+\delta}} \sqrt{\log \left(\frac{2Sd}{\left(\frac{u}{Q_T}\right)^{1+\delta}} + 1 \right)^d} \, du \right) \\ &= C E \left(Q_T \int_0^{r_{k_T}^{1+\delta}} \sqrt{\log \left(\frac{2Sd}{u^{1+\delta}} + 1 \right)^d} \, du \right). \end{aligned}$$

To sum up, we obtain by Jensen's inequality, the definitions of κ_T and r_{κ_T} , and $\log(x+1) \leq x$ for $x > 0$

$$\begin{aligned}
& E \left(\left| \sup_{\substack{\underline{s}_1, \underline{s}_2 \in [-S, S]^d \\ \rho(\underline{s}_1, \underline{s}_2) < r_{\kappa_T}}} \sum_{t=1}^{\mu_T} [L_{t,T}^0(\underline{s}_1) - L_{t,T}^0(\underline{s}_2)] \right| \right) \\
& \leq C E \left(\sqrt{\frac{\kappa_T^{1-\delta}}{bT} \sum_{t=1}^{\mu_T} \left(\sum_{i \in H_t} \left| \tilde{X}_i^{(M)}(u) \right|_1 \right)^{1+\delta}} \int_0^{r_{\kappa_T}^{\frac{1+\delta}{2}}} \sqrt{\log \left(\frac{2Sd}{u^{\frac{2}{1+\delta}}} + 1 \right)^d} du \right) \\
& \leq C \sqrt{\frac{\kappa_T^2}{bT} \mu_T} \left(\left\| \tilde{X}_i^{(M)}(u) \right\|_{1+\delta} \right)^{1+\delta} \int_0^{r_{\kappa_T}^{\frac{1+\delta}{2}}} \sqrt{\log \left(\frac{2Sd}{u^{\frac{2}{1+\delta}}} + 1 \right)} du \\
& \leq C \sqrt{\kappa_T} \int_0^{r_{\kappa_T}^{\frac{1+\delta}{2}}} u^{-\frac{1}{1+\delta}} du \\
& = C \sqrt{\kappa_T} r_{\kappa_T}^{\delta/2} \\
& \leq C (bT)^{-1/(4m)},
\end{aligned}$$

which obviously tends to zero and thus gives (6.40) which completes the proof. \square

Proof of Theorem 3.2. By Theorem 1.5.4 and Theorem 1.5.7 of van der Vaart and Wellner (2000), we have to show convergence of the fidis and tightness in the sense of Lemma 3.3. Convergence of the fidis follows from Theorem 3.1 and tightness from Lemma 3.3. By their Addendum 1.5.8 we can deduce continuity of the sample path of the limiting process. \square

Proof of Corollary 3.1. From Assumption 2.1(ii.2) and (ii.3) we obtain

$$\underline{\mathbf{X}}_{t,h,T} = \bar{\mu} \left(\frac{t}{T} \right) + \sum_{j=-\infty}^{\infty} \bar{A}_{t,T}(j) \bar{\varepsilon}_{t-j} + \mathcal{O}_P \left(\frac{1}{T} \right) =: \bar{X}_{t,T} + \mathcal{O}_P \left(\frac{1}{T} \right)$$

with $\bar{\mu}$ and $\bar{A}_{t,T}(j)$ satisfying Assumption 2.1 again. This leads to

$$\sqrt{bT} \sup_{\underline{s} \in \mathbb{R}^{2d}} \left| \hat{\varphi}_h(u, \underline{s}) - \frac{1}{T} \sum_{t=1}^T K_b \left(\frac{t+h/2}{T} - u \right) e^{i\langle \underline{s}, \bar{X}_{t,T} \rangle} \right| = o_P(1) \quad \forall u \in (0, 1).$$

Moreover, it follows from Lipschitz continuity of K that

$$\sqrt{bT} \sup_{\underline{s} \in \mathbb{R}^{2d}} \left| \frac{1}{T} \sum_{t=1}^T \left[K_b \left(\frac{t+h/2}{T} - u \right) - K_b \left(\frac{t}{T} - u \right) \right] e^{i\langle \underline{s}, \bar{X}_{t,T} \rangle} \right| = o_P(1) \quad \forall u \in (0, 1).$$

Since $\bar{X}_{t,T}$ satisfies Assumption 2.2 with $k = 1$, the assertion now follows from Theorem 3.2. \square

6.3. Proofs of Section 4.

Proof of Corollary 4.1. The parameter estimators are continuous functions of $\widehat{\varphi}$. Theorem 3.1(ii) yields (pointwise) consistency of $\widehat{\varphi}$ for φ . Hence the continuous mapping theorem implies consistency of the estimators for the parameters of the stable distribution. \square

Proof of Theorem 4.1. (i) First note that in analogy to the proof of Theorem 3.4.1 in Ushakov (1999),

$$\int \left| \varphi(u, s; \widehat{\theta}(u)) - \varphi(u, s; \theta_0(u)) \right|^2 w(s) ds \xrightarrow{P} 0 \quad \text{as } T \rightarrow \infty \quad (6.42)$$

follows from Theorem 3.2 and integrability of w . From this we can deduce that $\widehat{\theta}(u) \xrightarrow{P} \theta_0(u)$ as follows. First we get, that every sequence $(T_k)_k \subseteq \mathbb{N}$ contains a subsequence $(T_{k_l})_l$ such that (6.42) holds a.s. for this subsequence. Assume that for a subsequence $(T_{k_{l_m}})_m$ thereof, $(\widehat{\theta}_{T_{k_{l_m}}}(u))_m \rightarrow \bar{\theta}(u) \neq \theta_0(u)$ a.s. which implies that the corresponding CFs differ on an open set, say U_0 , in view of Assumption 4.1(iii). Thus, we have

$$\begin{aligned} & \int \left| \varphi(u, s; \widehat{\theta}_{T_{k_{l_m}}}(u)) - \varphi(u, s; \theta_0(u)) \right|^2 w(s) ds \\ & \geq \int_{U_0} \left| \varphi(u, s; \theta_0(u)) - \varphi(u, s; \bar{\theta}(u)) \right|^2 w(s) ds - \int_{\mathbb{R}} \left| \varphi(u, s; \widehat{\theta}_{T_{k_{l_m}}}(u)) - \varphi(u, s; \bar{\theta}(u)) \right|^2 w(s) ds \\ & > \eta \end{aligned}$$

under Assumption 4.1(iii) for all large k and some $\eta > 0$, which yields a contradiction.

(ii) We abbreviate the integral on the right-hand side of (4.7) by $D_T(u; \theta(u))$. The (uniform) consistency properties of our estimator $\widehat{\varphi}(u, s)$ implies that its first derivative w.r.t. θ satisfies $\widehat{D}_T^{(1)}(u; \widehat{\theta}(u)) = 0_p$ with probability tending to one. Hence, Taylor expansion of order two gives

$$\sqrt{bT}(\widehat{\theta}(u) - \theta_0(u)) = -\sqrt{bT} \left[D_T^{(2)}(u; \theta_0(u) + \xi(\widehat{\theta}(u) - \theta_0(u))) \right]^{-1} D_T^{(1)}(u; \theta_0(u))$$

for some random $\xi \in [-1, 1]$. By Assumption 4.1(iv), the order of differentiation and integration is commutable. Hence, in a neighborhood of $\theta_0(u)$ gradient and Hessian can be represented as

$$\begin{aligned} D_T^{(1)}(u; \theta(u)) &= -2 \int_{\mathbb{R}} \left[(\Re \widehat{\varphi}(u, s) - \Re \varphi(u, s; \theta(u))) \Re \varphi^{(1)}(u, s; \theta(u)) \right. \\ & \quad \left. + (\Im \widehat{\varphi}(u, s) - \Im \varphi(u, s; \theta(u))) \Im \varphi^{(1)}(u, s; \theta(u)) \right] w(s) ds \end{aligned}$$

and

$$\begin{aligned} D_T^{(2)}(u; \theta(u)) &= 2 \int_{\mathbb{R}} \left(\Re \varphi^{(1)}(u, s; \theta(u)) [\Re \varphi^{(1)}(u, s; \theta(u))]' + \Im \varphi^{(1)}(u, s; \theta(u)) [\Im \varphi^{(1)}(u, s; \theta(u))] \right)' \\ & \quad - (\Re \widehat{\varphi}(u, s) - \Re \varphi(u, s; \theta(u))) \Re \varphi^{(2)}(u, s; \theta(u)) \\ & \quad - (\Im \widehat{\varphi}(u, s) - \Im \varphi(u, s; \theta(u))) \Im \varphi^{(2)}(u, s; \theta(u)) \Big) w(s) ds, \end{aligned}$$

respectively. By part (i) and Theorem 3.2 $D_T^{(2)}(u; \theta_0(u) + \xi(\widehat{\theta}(u) - \theta_0(u))) \xrightarrow{P} 2D_0$, which is invertible by assumption. Hence, it remains to show that

$$\sqrt{bT} D_T^{(1)}(u; \theta_0(u)) \xrightarrow{d} 2D_0 Z_{MDE}.$$

Applying Theorem 3.2 and the continuous mapping theorem, we obtain

$$\begin{aligned} R_{S,T} &:= \sqrt{bT} \int_{[-S,S]^d} \left[(\Re \widehat{\varphi}(u, s) - \Re \varphi(u, s; \theta_0(u))) \Re \varphi^{(1)}(u, s; \theta_0(u)) \right. \\ &\quad \left. + (\Im \widehat{\varphi}(u, s) - \Im \varphi(u, s; \theta_0(u))) \Im \varphi^{(1)}(u, s; \theta_0(u)) \right] w(s) ds \\ \xrightarrow{d} R_S &:= \int_{[-S,S]^d} \left[Z_1(u, s) \Re \varphi^{(1)}(u, s; \theta_0(u)) + Z_2(u, s) \Im \varphi^{(1)}(u, s; \theta_0(u)) \right] w(s) ds \end{aligned}$$

for any $S \in (0, \infty)$, where $Z = (Z_1, Z_2)'$ is the limiting process in Theorem 3.2 applied to $\varphi(u, s) = \varphi(u, s; \theta_0)$. It follows from a usual Riemann approximation of the integral that the limit R_S is a centered normal random variable. By (4.10),

$$\text{Var}(R_S) \longrightarrow \text{Var}(D_0 Z_{MDE}) < \infty \quad \text{as } S \rightarrow \infty,$$

which implies that

$$R_S \xrightarrow{d} D_0 Z_{MDE}.$$

Now, the assertion follows from Proposition 6.3.9 in Brockwell and Davis (1991) if additionally

$$\lim_{S \rightarrow \infty} \limsup_{T \rightarrow \infty} P \left(|R_{S,T} - \sqrt{bT} D_T^{(1)}(u; \theta_0(u))| > \eta \right) = 0.$$

In view of (4.10), Lemma 3.1(i) implies

$$\lim_{S \rightarrow \infty} \limsup_{T \rightarrow \infty} E \left| R_{S,T} - \sqrt{bT} D_T^{(1)}(u; \theta_0(u)) \right|^2 = 0,$$

which then completes the proof. \square

Proof of Lemma 4.2. By expanding the squared term of the integrand in (4.14), we get

$$\begin{aligned} & \left| \kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2) \right|^2 \\ &= \left| \kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) \right|^2 - \kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) \overline{\widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)} \\ &\quad - \kappa_T \overline{\widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2)} \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2) + \left| \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2) \right|^2 \\ &= I + II + III + IV. \end{aligned}$$

From (4.15) and expanding $\exp(i \cdot) = \cos(\cdot) + i \sin(\cdot)$, we get for the first term

$$\begin{aligned} I &= \frac{1}{T^4} \sum_{t_1, t_2, t_3, t_4=1}^{T-h} \prod_{j=1}^4 K_b \left(\frac{t_j + h/2}{T} - u \right) \cos \left(\underline{s}'_1 (\underline{Y}_{t_3, T} - \underline{Y}_{t_4, T}) \right) \cos \left(\underline{s}'_2 (\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}) \right) \\ &\quad + R_I, \end{aligned}$$

where R_I is a remainder term that leads to vanishing terms for the integral in (4.14) (note that $\sin(\cdot)$ is an odd function). By using the identity

$$\cos(u) \cos(v) = 1 - (1 - \cos(u)) - (1 - \cos(v)) + (1 - \cos(u))(1 - \cos(v))$$

we get

$$\begin{aligned} I &= R_I + \frac{1}{T^4} \sum_{t_1, t_2, t_3, t_4=1}^{T-h} \prod_{j=1}^4 K_b \left(\frac{t_j + h/2}{T} - u \right) \left\{ 1 - \left(1 - \cos \left(\underline{s}'_1 (\underline{Y}_{t_3, T} - \underline{Y}_{t_4, T}) \right) \right) \right. \\ &\quad \left. - \left(1 - \cos \left(\underline{s}'_2 (\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}) \right) \right) + \left(1 - \cos \left(\underline{s}'_1 (\underline{Y}_{t_3, T} - \underline{Y}_{t_4, T}) \right) \right) \left(1 - \cos \left(\underline{s}'_2 (\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}) \right) \right) \right\} \\ &=: I_1 - I_2 - I_3 + I_4 + R_I. \end{aligned}$$

The same calculation for terms *II*, *III* and *IV* leads to similar expressions, where the first three terms cancel out, i.e. $I_j + II_j + III_j + IV_j = 0$ for $j = 1, 2, 3$ and $II_4 = III_4$. Note that the cancellation is due to the inclusion of the factor κ_T in the definition of (4.13). Altogether, this leads to

$$\begin{aligned}
& |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 \\
&= \frac{1}{T^4} \sum_{t_1, t_2, t_3, t_4=1}^{T-h} \prod_{j=1}^4 K_b\left(\frac{t_j + h/2}{T} - u\right) \left\{ \left(1 - \cos(\underline{s}'_1(\underline{Y}_{t_3, T} - \underline{Y}_{t_4, T}))\right) \left(1 - \cos(\underline{s}'_2(\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}))\right)\right. \\
&\quad \left. - 2 \left(1 - \cos(\underline{s}'_1(\underline{Y}_{t_2, T} - \underline{Y}_{t_4, T}))\right) \left(1 - \cos(\underline{s}'_2(\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}))\right)\right. \\
&\quad \left. + \left(1 - \cos(\underline{s}'_1(\underline{Y}_{t_1, T} - \underline{Y}_{t_2, T}))\right) \left(1 - \cos(\underline{s}'_2(\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T}))\right) \right\} + R,
\end{aligned} \tag{6.43}$$

where R is a remainder term that leads to vanishing terms for the integral in (4.14). Now, we plug the above into (4.14) and make use of Lemma 4.1 to get the claimed result. \square

Proof of Theorem 4.2. To prove consistency of $\widehat{\mathcal{V}}_{Y,Z}^2(u; h)$ we proceed in three steps and show:

1. For $\delta \in (0, 1)$

$$\begin{aligned}
\widehat{\mathcal{V}}_{Y,Z;\delta}^2(u; h) &:= \int_{D_\delta} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\
&\xrightarrow{P} \mathcal{V}_{Y,Z;\delta}^2(u; h) := \int_{D_\delta} |\varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y;0}(u; \underline{s}_1) \varphi_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2),
\end{aligned}$$

where $D_\delta = \{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid \delta \leq |\underline{s}_1|_2 \leq 1/\delta, \delta \leq |\underline{s}_2|_2 \leq 1/\delta\}$.

2. $\mathcal{V}_{Y,Z;\delta}^2(u; h) \rightarrow \mathcal{V}_{Y,Z}^2(u; h)$ with $\delta \rightarrow 0$.

3. $\lim_{\delta \rightarrow 0} \limsup_{T \rightarrow \infty} P\left(\left|\widehat{\mathcal{V}}_{Y,Z;\delta}^2(u; h) - \widehat{\mathcal{V}}_{Y,Z}^2(u; h)\right| > \varepsilon\right) = 0 \quad \forall \varepsilon > 0$.

Then, the assertion for distance covariance can be deduced from Proposition 6.3.9 in Brockwell and Davis (1991) and it remains to consider the distance correlation. Under the assumptions of the theorem $\mathcal{V}_Y^2(u; 0) \mathcal{V}_Z^2(u; 0) > \eta$ for some $\eta > 0$. Hence, with probability tending to one, we have $\mathbb{1}(\widehat{\mathcal{V}}_Y^2(u; 0) \widehat{\mathcal{V}}_Z^2(u; 0) > 0) = 1$. Thus, stochastic convergence of $\widehat{\mathcal{R}}_{Y,Z}^2(u; h)$ follows immediately from convergence of the distance covariance and it remains to carry out steps 1 to 3 mentioned above.

Step 1.

We rewrite

$$\begin{aligned}
\widehat{\mathcal{V}}_{Y,Z;\delta}^2(u; h) &= \int_{D_\delta} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) \\
&\quad + \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y;0}(u; \underline{s}_1) \varphi_{Z;h}(u; \underline{s}_2) \\
&\quad + \varphi_{Y;0}(u; \underline{s}_1) \varphi_{Z;h}(u; \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2).
\end{aligned}$$

By Cauchy-Schwarz inequality and recognizing that

$$\int_{D_\delta} |\varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y;0}(u; \underline{s}_1) \varphi_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) < \infty,$$

it remains to show that

$$\int_{D_\delta} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \xrightarrow{P} 0 \tag{6.44}$$

and

$$\int_{D_\delta} |\varphi_{Y;0}(u; \underline{s}_1) \varphi_{Z;h}(u; \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \xrightarrow{P} 0. \quad (6.45)$$

For (6.44) note that

$$\begin{aligned} & \int_{D_\delta} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\ & \leq 2 \int_{D_\delta} (\kappa_T - 1)^2 |\widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\ & \quad + 2 \int_{D_\delta} |\widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2). \end{aligned} \quad (6.46)$$

The first summand on the right-hand side tends to zero in probability since $\kappa_T \rightarrow 1$ and $\widehat{\varphi}_{Y,Z;h}$ is bounded (uniformly in T). Following the lines of the proof of Corollary 4.1, we obtain that $\sqrt{bT}(\widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \varphi_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2))_{(\underline{s}'_1, \underline{s}'_2)' \in D_\delta}$ converges in distribution to a complex centered Gaussian process with continuous sample paths. Together with the continuous mapping theorem this gives asymptotic negligibility of the second summand in (6.46). With the same arguments (6.45) can be obtained.

Step 2.

The assertion is an immediate consequence of the monotone convergence theorem.

Step 3.

We follow the lines of the proof of Theorem 2 in Szekely, Rizzo and Bakirov (2007). First, note that

$$\begin{aligned} & \widehat{\mathcal{V}}_{Y,Z}^2(u; h) - \widehat{\mathcal{V}}_{Y,Z;\delta}^2(u; h) \\ & = \int_{\mathbb{R}^{p+q} \setminus D_\delta} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2). \end{aligned}$$

and

$$\begin{aligned} \mathbb{R}^{p+q} \setminus D_\delta & = \{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 < \delta\} \cup \{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 > 1/\delta\} \\ & \quad \cup \{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_2|_2 < \delta\} \cup \{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_2|_2 > 1/\delta\}. \end{aligned}$$

Hence, for symmetry reasons it suffices to show that $\forall \varepsilon > 0$

$$\lim_{\delta \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 > 1/\delta\}} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) > \varepsilon \right) \quad (6.47)$$

and

$$\lim_{\delta \rightarrow 0} \limsup_{T \rightarrow \infty} P \left(\int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 < \delta\}} |\kappa_T \widehat{\varphi}_{Y,Z;h}(u; \underline{s}_1, \underline{s}_2) - \widehat{\varphi}_{Y;0}(u; \underline{s}_1) \widehat{\varphi}_{Z;h}(u; \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) > \varepsilon \right) \quad (6.48)$$

are equal to zero. For (6.47) we invoke expansion (6.43) of the integrand and Lemma 4.1

$$\begin{aligned} & \int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 > 1/\delta\}} |\kappa_T \widehat{\varphi}_{X,Y;h}(u, (\underline{s}_1, \underline{s}_2)) - \widehat{\varphi}_{X;0}(u, \underline{s}_1) \widehat{\varphi}_{Y;h}(u, \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\ & = \int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 > 1/\delta\}} \left| \frac{1}{T^4} \sum_{t_1, t_2, t_3, t_4=1}^{T-h} \prod_{j=1}^4 K_b \left(\frac{t_j + h/2}{T} - u \right) \right. \\ & \quad \times \left\{ \left(1 - \cos(\underline{s}'_1(\underline{Y}_{t_3, T} - \underline{Y}_{t_4, T})) \right) \left(1 - \cos(\underline{s}'_2(\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T})) \right) \right. \\ & \quad \left. \left. - 2 \left(1 - \cos(\underline{s}'_1(\underline{Y}_{t_2, T} - \underline{Y}_{t_4, T})) \right) \left(1 - \cos(\underline{s}'_2(\underline{Z}_{t_3+h, T} - \underline{Z}_{t_4+h, T})) \right) \right\} \right. \end{aligned}$$

$$\begin{aligned}
& + \left(1 - \cos(\underline{s}'_1(Y_{t_1,T} - Y_{t_2,T})) \right) \left(1 - \cos(\underline{s}'_2(Z_{t_3+h,T} - Z_{t_4+h,T})) \right) \Big| w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\
\leq & 8 \int_{|\underline{s}_1|_2 > 1/\delta} \frac{1}{c_p |\underline{s}_1|_2^{1+p}} d\underline{s}_1 \\
& \times \int_{\mathbb{R}^q} \frac{1}{T^4} \sum_{t_1, t_2, t_3, t_4=1}^{T-h} \prod_{j=1}^4 K_b\left(\frac{t_j + h/2}{T} - u\right) \left(1 - \cos\left(\underline{s}'_2(Z_{t_3+h,T} - Z_{t_4+h,T})\right) \right) \frac{1}{c_q |\underline{s}_2|_2^{1+q}} d\underline{s}_2 \\
\leq & C \int_{|\underline{s}_1|_2 > 1/\delta} \frac{1}{|\underline{s}_1|_2^{1+p}} d\underline{s}_1 \frac{1}{T^2} \sum_{t_3, t_4=1}^{T-h} \prod_{j=3}^4 K_b\left(\frac{t_j + h/2}{T} - u\right) |\underline{Z}_{t_3+h,T} - \underline{Z}_{t_4+h,T}| \\
\leq & C \int_{|\underline{s}_1|_2 > 1/\delta} \frac{1}{|\underline{s}_1|_2^{1+p}} d\underline{s}_1 \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) |\underline{Z}_{t+h,T}|.
\end{aligned}$$

Now, Markov's inequality gives

$$(6.47) \leq C \lim_{\delta \rightarrow 0} \int_{|\underline{s}_1|_2 > 1/\delta} \frac{1}{|\underline{s}_1|_2^{1+p}} d\underline{s}_1 \limsup_{T \rightarrow \infty} \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) E|\underline{Z}_{t+h,T}| = 0.$$

The verification of (6.48) is more involved. We denote by $\bar{\underline{X}}(u) = (\bar{\underline{Y}}'(u), \bar{\underline{Z}}'(u))'$ a copy of $\tilde{\underline{X}}_1(u)$ which is independent of the DGP $(\underline{X}_{t,T})_{t=1}^T$, $T \in \mathbb{N}$. Using the inequalities of Cauchy-Schwarz and Jensen, we obtain

$$\begin{aligned}
& \int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 < \delta\}} |\kappa_T \widehat{\varphi}_{X,Y;h}(u, (\underline{s}_1, \underline{s}_2)) - \widehat{\varphi}_{X;0}(u, \underline{s}_1) \widehat{\varphi}_{Y;h}(u, \underline{s}_2)|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\
& = \int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 < \delta\}} w(\underline{s}_1, \underline{s}_2) \left| \frac{\kappa_T}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) (e^{i\underline{s}'_1 Y_{t,T}} - E e^{i\underline{s}'_1 \bar{\underline{Y}}(u)}) (e^{i\underline{s}'_2 Z_{t+h,T}} - E e^{i\underline{s}'_2 \bar{\underline{Z}}(u)}) \right. \\
& \quad \left. - \frac{1}{T^2} \sum_{t_1, t_2=1}^{T-h} K_b\left(\frac{t_1 + h/2}{T} - u\right) K_b\left(\frac{t_2 + h/2}{T} - u\right) (e^{i\underline{s}'_1 Y_{t_1,T}} - E e^{i\underline{s}'_1 \bar{\underline{Y}}(u)}) (e^{i\underline{s}'_2 Z_{t_2,T}} - E e^{i\underline{s}'_2 \bar{\underline{Z}}(u)}) \right|^2 d(\underline{s}_1, \underline{s}_2) \\
& \leq 2 \int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 < \delta\}} \left(\frac{\kappa_T}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) |e^{i\underline{s}'_1 Y_{t,T}} - E e^{i\underline{s}'_1 \bar{\underline{Y}}(u)}|^2 \right) \\
& \quad \times \left(\frac{\kappa_T}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) |e^{i\underline{s}'_2 Z_{t+h,T}} - E e^{i\underline{s}'_2 \bar{\underline{Z}}(u)}|^2 \right) w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\
& \quad + 2 \int_{\{(\underline{s}'_1, \underline{s}'_2)' \in \mathbb{R}^{p+q} \mid |\underline{s}_1|_2 < \delta\}} \frac{1}{T^2} \left| \sum_{t_1=1}^{T-h} K_b\left(\frac{t_1 + h/2}{T} - u\right) (e^{i\underline{s}'_1 Y_{t_1,T}} - E e^{i\underline{s}'_1 \bar{\underline{Y}}(u)}) \right|^2 \\
& \quad \times \left| \sum_{t_2=1}^{T-h} K_b\left(\frac{t_2 + h/2}{T} - u\right) (e^{i\underline{s}'_2 Z_{t_2+h,T}} - E e^{i\underline{s}'_2 \bar{\underline{Z}}(u)}) \right|^2 w(\underline{s}_1, \underline{s}_2) d(\underline{s}_1, \underline{s}_2) \\
& \leq 4\kappa_T^2 \int_{|\underline{s}_1|_2 \leq \delta} \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) |e^{i\underline{s}'_1 Y_{t,T}} - E e^{i\underline{s}'_1 \bar{\underline{Y}}(u)}|^2 \frac{1}{c_p |\underline{s}_1|_2^{1+p}} d\underline{s}_1 \\
& \quad \times \int_{\underline{s}_2 \in \mathbb{R}^q} \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t + h/2}{T} - u\right) |e^{i\underline{s}'_2 Z_{t+h,T}} - E e^{i\underline{s}'_2 \bar{\underline{Z}}(u)}|^2 \frac{1}{c_q |\underline{s}_2|_2^{1+q}} d\underline{s}_2.
\end{aligned}$$

With the same arguments as in Székely, Rizzo and Bakirov (2007, page 2777f.), we obtain

$$\begin{aligned} & \int_{\mathbf{s}_2 \in \mathbb{R}^q} \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) |e^{is'_2 \mathbf{Z}_{t,T}} - Ee^{is'_2 \bar{\mathbf{Z}}(u)}|^2 \frac{1}{c_q |\mathbf{s}_2|_2^{1+q}} d\mathbf{s}_2 \\ & \leq \frac{2}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) (|\mathbf{Z}_{t+h,T}| + E|\bar{\mathbf{Z}}(u)|) \end{aligned}$$

and

$$\begin{aligned} & \int_{|\mathbf{s}_1|_2 \leq \delta} \frac{1}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) |e^{is'_1 \mathbf{Y}_{t,T}} - Ee^{is'_1 \bar{\mathbf{Y}}(u)}|^2 \frac{1}{c_p |\mathbf{s}_1|_2^{1+p}} d\mathbf{s}_1 \\ & \leq \frac{2}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) E(|\mathbf{Y}_{t,T} - \bar{\mathbf{Y}}(u)| G(|\mathbf{Y}_{t,T} - \bar{\mathbf{Y}}(u)|\delta) | \mathbf{Y}_{t,T}), \end{aligned}$$

where

$$G(y) = \int_{|\mathbf{z}|_2 \leq y} \frac{1 - \cos(z_1)}{|\mathbf{z}|_2^{1+p}} d\mathbf{z}, \quad \mathbf{z} = (z_1, \dots, z_p)'$$

Summing up, we have for (6.48) by Markov's inequality and Lebesgue's dominated convergence theorem

(6.48)

$$\begin{aligned} & \leq \lim_{\delta \rightarrow 0} \limsup_{T \rightarrow \infty} P\left(\frac{16}{T^2} \kappa_T^2 \sum_{t_1=1}^{T-h} K_b\left(\frac{t_1+h/2}{T} - u\right) (|\mathbf{Z}_{t_1+h,T}| + E|\bar{\mathbf{Z}}(u)|) \right. \\ & \quad \left. \times \sum_{t_2=1}^{T-h} K_b\left(\frac{t_2+h/2}{T} - u\right) E(|\mathbf{Y}_{t_2,T} - \bar{\mathbf{Y}}(u)| G(|\mathbf{Y}_{t_2,T} - \bar{\mathbf{Y}}(u)|\delta) | \mathbf{Y}_{t_2,T}) > \varepsilon\right) \\ & \leq \eta_M + \lim_{\delta \rightarrow 0} \limsup_{T \rightarrow \infty} P\left(\frac{M}{T} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) E(|\mathbf{Y}_{t,T} - \bar{\mathbf{Y}}(u)| G(|\mathbf{Y}_{t,T} - \bar{\mathbf{Y}}(u)|\delta) | \mathbf{Y}_{t,T}) > \varepsilon\right) \\ & \leq \eta_M + \lim_{\delta \rightarrow 0} \limsup_{T \rightarrow \infty} \frac{M}{T\varepsilon} \sum_{t=1}^{T-h} K_b\left(\frac{t+h/2}{T} - u\right) E(|\mathbf{Y}_{t,T} - \bar{\mathbf{Y}}(u)| G(|\mathbf{Y}_{t,T} - \bar{\mathbf{Y}}(u)|\delta)) \\ & \leq \eta_M \end{aligned}$$

with $\eta_M \rightarrow 0$ as $M \rightarrow \infty$. This leads to the desired result as G is bounded and $G(y) \rightarrow 0$ with $y \rightarrow 0$. \square

Acknowledgment . This research was partly funded by the German Research Foundation DFG via the Research Training Group RTG 1953, the Collaborative Research Center SFB 884 (Project B6), and the Volkswagen Foundation (Professorinnen für Niedersachsen des Niedersächsischen Vorab).

REFERENCES

- Arcones, M.A. and Yu, B. (1994). Central limit theorems for empirical and U -processes of stationary mixing sequences. *Journal of Theoretical Probability* **7**, 47–71.
- Bakirov, N.K., Rizzo, M.L., and Székely, G.J. (2006). A multivariate nonparametric test of independence. *Journal of Multivariate Analysis* **97**, 1742–1756.
- Borak, S., Misiorek, A. and Weron, R. (2005). Models for heavy tailed asset returns. In: *Statistical Tools in Finance and Insurance* Eds. Čížek, Härdle, Weron. Springer.

- Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer, New York.
- Cardinali, A. and Nason, G. (2010). Costationarity of locally stationary time series. *Journal of Time Series Econometrics* **2**, Article 1
- Csörgő, S. (1981) Multivariate empirical characteristic functions. *Z. Wahrsch. Verw. Gebiete* **55**, 203–229.
- Dahlhaus, R. (1996a). On the Kullback-Leibler information divergence for locally stationary processes. *Stoch. Proc. Appl.* **62**, 139–168.
- Dahlhaus, R. (1996b). Maximum likelihood estimation and model selection for locally stationary processes. *J. Nonpar. Statist.* **6**, 171–191.
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics* **25**, 1–37.
- Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28**, 1762–1794.
- Dahlhaus, R. (2012). Locally stationary processes. *Handbook of Statistics* **30**.
- Davis, R. A., Matsui, M. Mikosch, T., and Wan, P. (2016). Applications of Distance Correlation to Time Series. Submitted.
- Dueck, J., Edelman, D., Gneiting, T., and Richards, D. (2014). The affinely invariant distance correlation. *Bernoulli* **20**, 2305–2330.
- Dueck, J., Edelman, D., and Richards, D. (2015). A generalization of an integral arising in the theory of distance correlation. *Statist. Probab. Lett* **119**, 116–119.
- Franco, C. and Zakoian, J.-M. (2010). *GARCH Models*. Wiley, Chichester.
- Fielitz, B. D., and Rozelle, J. P. (1981). Method-of-moments estimators of stable distribution parameters. *Applied Mathematics and Computation* **8**, 303–320.
- Höpfner, R. and Rüschendorf, L. (1999). Comparison of estimators in stable models. *Mathematical and Computer Modelling* **29**, 145–160.
- Knight, J.L. and Yu, J. (2002). The empirical characteristic function in time series estimation. *Econometrics Theory* **18**, 691–721.
- Leucht, A. (2012). Characteristic function-based hypothesis tests under weak dependence. *Journal of Multivariate Analysis* **108**, 67–89.
- Press, J. S. (1972). Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association* **67**, 842–846.
- Priestley, M.B. (1965). Evolutionary spectra and non-stationary processes. *J. Roy. Statist. Soc. Ser. B* **27**, 204–237.
- Romano, J. P. and Wolf, M. (2000). A likelihood approximation for locally stationary processes. *Statist. Probab. Lett* **47**, 115–124.
- Szekely, G. J., Rizzo, M. L., and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics* **35**, 2769–2794.
- Szekely, G. J. and Rizzo, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics* **3**, 1236–1265.
- Szekely, G.J. and Rizzo, M.L. (2012). On the uniqueness of distance correlation. *Statist. Probab. Lett* **82**, 2278–2282.
- Ushakov, N. G. (1999). *Selected Topics in Characteristic Functions*. Utrecht: VSP.
- Van der Vaart, A. W. and Wellner, J. A. (2000). *Weak Convergence and Empirical Processes. With Applications to Statistics*. New York: Springer.
- Vogt, M. (2012). Nonparametric regression for locally stationary time series. *The Annals of Statistics* **40**, 2601–2633.
- Xu, D. and Knight, J. (2010). Continuous empirical characteristic function estimation of mixtures of normal parameters. *Econometric Reviews*, 25–50.

- Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometrics Reviews* **23**, 93–1223.
- Zhou, Z. (2012). Measuring nonlinear dependence in time series, a distance correlation approach. *J. Time Ser. Anal.* **33**, 438–457.
- Zhou, Z. (2014). Inference of weighted V -statistics for nonstationary time series and its applications. *The Annals of Statistics* **42**, 87–114.