Distance correlation for long-range dependent time series

Preprint · July 2021

Annika Betken
Ruhr-Universität Bochum

DISTANCE CORRELATION FOR LONG-RANGE DEPENDENT TIME SERIES

BY ANNIKA BETKEN[1] AND HEROLD DEHLING[2]

[1]*Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS), University of Twente, a.betken@utwente.nl*

[2]*Faculty of Mathematics, Ruhr-Universität Bochum, herold.dehling@rub.de*

We apply the concept of distance correlation for testing independence of long-range dependent time series. For this, we establish a non-central limit theorem for stochastic processes with values in an L_2-Hilbert space. This limit theorem is of a general theoretical interest that goes beyond the context of this article. For the purpose of this article, it provides the basis for deriving the asymptotic distribution of the distance covariance of subordinated Gaussian processes. Depending on the dependence in the data, the standardization and the limit of distance correlation vary. In any case, the limit is not feasible, such that test decisions are based on a subsampling procedure. We prove the validity of the subsampling procedure and assess the finite sample performance of a hypothesis test based on the distance covariance. In particular, we compare its finite sample performance to that of a test based on Pearson's sample correlation coefficient. For this purpose, we additionally establish convergence results for this dependence measure. Different dependencies between the vectors are considered. It turns out that only linear correlation is better detected by Pearson's sample correlation coefficient, while all other dependencies are better detected by distance correlation. An analysis with regard to cross-dependencies between the mean monthly discharges of three different rivers provides an application of the theoretical results established in this article.

CONTENTS

Primary 62G10; secondary 62H20.

Keywords and phrases: time series, distance correlation, long-range dependence, subsampling.

1. Introduction. Given observations $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ as initial segments of the stationary components of a bivariate, covariance stationary stochastic process (X_k, Y_k), $k \geq 1$, our goal is to decide on the testing problem

$$H_0\colon X_k,\ k \geq 1, \text{ and } Y_k,\ k \geq 1, \text{ are independent,} \tag{1}$$
$$H_1\colon X_k,\ k \geq 1, \text{ and } Y_k,\ k \geq 1, \text{ are dependent.}$$

It seems natural to base a test decision for this testing problem on some measure of the degree of dependence based on the observations $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$. The best-known such measure is Pearson's sample correlation coefficient

$$r_{X,Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}},$$

which is an estimator for the correlation $\rho_{X,Y}$ of the random variables X and Y, where (X, Y) is a generic random vector with the same joint distribution as any of the random vectors (X_k, Y_k), $k \geq 1$. Consequently, $\rho_{X,Y}$ measures the dependence between X and Y rather than the dependence between $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$. In practice, however, independence of X and Y typically goes along with independence of $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$, such that a decision on the considered testing problem can be based on the value of $r_{X,Y}$.

It is well-known that Pearson's correlation coefficient only measures the degree of linear dependence, while it is not sensitive to non-linear dependence. As a result, the random variables X and Y may be highly dependent although $\rho_{X,Y} = 0$. In a series of papers, Székely, Rizzo and Bakirov (2007) and Székely and Rizzo (2009, 2012, 2013, 2014) introduced distance covariance and distance correlation as alternative measures for the degree of dependence between two random vectors X and Y with values in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively. They define the distance covariance of X and Y as

$$\mathcal{V}^2(X, Y; w) := \int_{\mathbb{R}^p} \int_{\mathbb{R}^q} \left| \varphi_{X,Y}(s,t) - \varphi_X(s)\, \varphi_Y(t) \right|^2 w(s,t) ds\, dt,$$

where $\varphi_{X,Y}(s,t) := \mathrm{E}\left(\mathrm{e}^{i(\langle s, X\rangle + \langle t, Y\rangle)}\right)$ is the joint characteristic function of X, Y, $\varphi_X := \mathrm{E}(\mathrm{e}^{i\langle s, X\rangle})$ and $\varphi_Y := \mathrm{E}(\mathrm{e}^{\langle t, Y\rangle})$ are the characteristic functions of X and Y, and where $w(s,t)$ is some non-negative weight function on $\mathbb{R}^p \times \mathbb{R}^q$. Under the assumption that

$$w(s,t) > 0 \text{ for all } s \in \mathbb{R}^p,\ t \in \mathbb{R}^q,$$

the random variables X and Y are independent if and only if $\mathcal{V}^2(X, Y; w) = 0$. For a variety of reasons, usually a weight function

$$w(s,t) = \frac{c_{p,q}}{|s|^{\alpha+p} |t|^{\alpha+q}}$$

with parameter $\alpha \in (0, 2)$ and a constant $c_{p,q} > 0$ is considered. This choice of weight function guarantees scale invariance of the corresponding distance correlation

$$\mathcal{R}(X, Y; w) := \frac{\mathcal{V}^2(X, Y; w)}{\sqrt{\mathcal{V}^2(X, X; w) \mathcal{V}^2(Y, Y; w)}},$$

as well as invariance relative to orthogonal transformations of the random vectors, two further desirable properties of dependence measures. Given observations $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$, Székely, Rizzo and Bakirov (2007) propose to use an empirical version of $\mathcal{V}^2(X, Y, w)$ as test statistic for a test of independence. They define the empirical distance covariance as

$$\mathcal{V}_n^2(X, Y; w) := \int_{\mathbb{R}^p} \int_{\mathbb{R}^q} \left| \varphi^{(n)}_{X,Y}(s,t) - \varphi^{(n)}_X(s) \varphi^{(n)}_Y(t) \right|^2 w(s,t) ds\, dt,$$

where

$$\varphi_{X,Y}^{(n)}(s,t) := \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{i\langle s,X_j\rangle + i\langle t,Y_j\rangle}$$

denotes the joint empirical characteristic function of X, Y, and

$$\varphi_X^{(n)}(s) := \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{i\langle s,X_j\rangle},\ \varphi_Y^{(n)}(t) := \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{i\langle t,Y_j\rangle}$$

correspond to the empirical characteristic functions of X and Y. Székely, Rizzo and Bakirov (2007) derive the large sample distribution of $\mathcal{V}_n^2(X,Y;w)$ in the case of independent, identically distributed sequences X_k, $k \geq 1$, and Y_k, $k \geq 1$. Zhou (2012) extends the concept of distance correlation to auto-distance correlation in a time series as a tool to explore non-linear dependence within a time series. Davis et al. (2018) apply the auto-distance correlation function to stationary multivariate time series in order to measure lagged auto- and cross-dependencies in a time series. Under mixing assumptions, these authors establish asymptotic theory for the sample auto- and cross-dependencies. Matsui, Mikosch and Samorodnitsky (2017) extend the notion of distance covariance to continuous time stochastic processes defined on the same interval. Dehling et al. (2020) establish consistency of a bootstrap procedure for the sample distance covariance of two processes. Kroll (2021) derived the asymptotic distribution of the sample distance covariance under mixing conditions.

In the present paper, we investigate distance correlation for long-range dependent processes X_k, $k \geq 1$, and Y_k, $k \geq 1$. Such processes, also known as long memory processes, are commonly used as models for random phenomena that exhibit dependence at all scales that are not captured by common time series models such as ARMA processes. Pipiras and Taqqu (2017) present stochastic models and probabilistic theory for long-range dependent processes. Statistical methods for long-range dependent processes are presented in Beran et al. (2013), and in Surgailis, Koul and Giraitis (2012). Our mathematical analysis is based on novel theory for Hilbert space valued long-range dependent processes which we apply to the empirical characteristic functions.

In Section 3.1, we establish a non-central limit theorem for stochastic processes with values in an L_2-Hilbert space. This limit theorem is of independent interest and additionally provides the basis for deriving the asymptotic distribution of the distance covariance of subordinated Gaussian processes in Section 3.2. Depending on the dependence in the data, the standardization and the limit of distance correlation vary. In both cases, the limits are not feasible, such that test decisions are based on a subsampling procedure. We prove the validity of the corresponding subsampling procedure in Section 4.1 and assess the finite sample performance of a hypothesis test based on the distance covariance through simulations in Section 4.2. In particular, we compare its finite sample performance to that of a test based on the sample covariance. For this purpose, we also establish convergence results for this dependence measure. Different dependencies between the vectors are considered. It turns out that only linear correlation is better detected by Pearson's sample correlation, while all other dependencies are better detected by distance correlation. An analysis with regard to cross-dependencies between the mean monthly discharges of three different rivers in Section 5 provides an application of the theoretical results established in this article.

2. Preliminaries. The following section establishes the model assumptions that are taken as a basis of the theoretical results in this article.

2.1. *Gaussian subordination and Long-range dependence.* In the following sections, we will focus on the consideration of subordinated Gaussian time series, i.e., on random observations generated by transformations of Gaussian processes.

DEFINITION 2.1. Let ξ_t, $t \in T$, be a Gaussian process with index set T. A process Y_t, $t \in T$, satisfying $Y_t = G(\xi_t)$ for some measurable function $G : \mathbb{R} \longrightarrow \mathbb{R}$ is called subordinated Gaussian process.

REMARK 2.1. For any particular distribution function F, an appropriate choice of the transformation G in Definition 2.1 yields subordinated Gaussian processes with marginal distribution F. Moreover, there exist algorithms for generating Gaussian processes that, after suitable transformation, yield subordinated Gaussian processes with marginal distribution F and a predefined covariance structure; see Pipiras and Taqqu (2017).

2.1.1. *Univariate Hermite expansion.* The subordinated random variables $Y_t = G(\xi_t)$, $t \in T$, can be considered as elements of the Hilbert space $L^2(\mathbb{R}, \varphi(x)dx) = \mathcal{L}^2(\mathbb{R}, \varphi(x)dx)/\mathcal{N}$, where $\mathcal{L}^2(\mathbb{R}, \varphi(x)dx)$ denotes the space of all measurable, real-valued functions which are square-integrable with respect to the measure $\varphi(x)dx$ associated with the standard normal density function φ and $\mathcal{N} := \ker(\|\cdot\|_{L^2})$. For two functions $G_1, G_2 \in L^2(\mathbb{R}, \varphi(x)dx)$ the corresponding inner product is defined by

$$\langle G_1, G_2 \rangle_{L^2} := \int_{-\infty}^{\infty} G_1(x) G_2(x) \varphi(x) dx = \mathrm{E} G_1(X) G_2(X) \tag{2}$$

with X denoting a standard normally distributed random variable.
A collection of orthogonal elements in $L^2(\mathbb{R}, \varphi(x)dx)$ is given by the sequence of Hermite polynomials; see Pipiras and Taqqu (2017).

DEFINITION 2.2. For $n \geq 0$, the Hermite polynomial of order n is defined by

$$H_n(x) = (-1)^n \mathrm{e}^{\frac{1}{2}x^2} \frac{d^n}{dx^n} \mathrm{e}^{-\frac{1}{2}x^2}, \; x \in \mathbb{R}.$$

Orthogonality of the sequence H_n, $n \geq 0$, in $L^2(\mathbb{R}, \varphi(x)dx)$ follows from

$$\langle H_n, H_m \rangle_{L^2} = \begin{cases} n! & \text{if } n = m, \\ 0 & \text{if } n \neq m. \end{cases}$$

Moreover, it can be shown that the Hermite polynomials form an orthogonal basis of $L^2(\mathbb{R}, \varphi(x)dx)$. As a result, every $G \in L^2(\mathbb{R}, \varphi(x)dx)$ has an expansion in Hermite polynomials, i.e., for $G \in L^2(\mathbb{R}, \varphi(x)dx)$ and ξ standard normally distributed, we have

$$G(\xi) = \sum_{r=0}^{\infty} \frac{J_r(G)}{r!} H_r(\xi), \tag{3}$$

where the so-called *Hermite coefficient* $J_r(G)$ is given by

$$J_r(G) := \langle G, H_r \rangle_{L^2} = \mathrm{E} G(X) H_r(X).$$

Equation (3) holds in an L^2-sense, meaning

$$\lim_{n \to \infty} \left\| G(\xi) - \sum_{r=0}^{n} \frac{J_r(G)}{r!} H_r(\xi) \right\|_{L^2} = 0,$$

where $\|\cdot\|_{L^2}$ denotes the norm induced by the inner product $\langle\cdot,\cdot\rangle_{L^2}$.
Given the Hermite expansion (3), it is possible to characterize the dependence structure of subordinated Gaussian time series $G(\xi_n)$, $n \in \mathbb{N}$. In fact, it holds that

$$\mathrm{Cov}(G(\xi_1), G(\xi_{k+1})) = \sum_{r=1}^{\infty} \frac{J_r^2(G)}{r!} (\rho(k))^r, \tag{4}$$

where ρ denotes the auto-covariance function of ξ_n, $n \in \mathbb{N}$; see Pipiras and Taqqu (2017). Under the assumption that, as k tends to ∞, $\rho(k)$ converges to 0 with a certain rate, the asymptotically dominating term in the series (4) is the summand corresponding to the smallest integer r for which the Hermite coefficient $J_r(G)$ is non-zero. This index, which decisively depends on G, is called *Hermite rank*.

DEFINITION 2.3. Let $G \in L^2(\mathbb{R}, \varphi(x)dx)$ with $\mathrm{E}G(X) = 0$ for standard normally distributed X and let $J_r(G)$, $r \geq 0$, be the Hermite coefficients in the Hermite expansion of G. The smallest index $k \geq 1$ for which $J_k(G) \neq 0$ is called the Hermite rank of G, i.e.,

$$r := \min\{k \geq 1 : J_k(G) \neq 0\}.$$

2.1.2. *Bivariate Hermite expansion.* Let ξ_t, $t \in T$, be a bivariate Gaussian process with index set T. More precisely, assume that $\xi_t := (\xi_t^{(1)}, \xi_t^{(2)})$ are Gaussian random vectors with mean 0 and covariance matrix Σ. We write φ_Σ for the corresponding density. Given a measurable function $G : \mathbb{R}^2 \longrightarrow \mathbb{R}$, subordinated random variables $Y = G(\xi_t)$, $t \in T$, can be considered as elements of the Hilbert space $L^2(\mathbb{R}^2, \varphi_\Sigma) = \mathcal{L}^2(\mathbb{R}^2, \varphi_\Sigma)/\mathcal{N}$, where $\mathcal{L}^2(\mathbb{R}^2, \varphi_\Sigma)$ denotes the space of all measurable, real-valued functions which are square-integrable with respect to the measure associated with the density function φ_Σ and $\mathcal{N} := \ker(\|\cdot\|_{L^2})$. For two functions $G_1, G_2 \in L^2(\mathbb{R}^2, \varphi_{I_2})$ the corresponding inner product is defined by

$$\langle G_1, G_2\rangle_{L^2} := \int_{-\infty}^{\infty} G_1(x)G_2(x)\varphi_{I_2}(x)dx = \mathrm{E}G_1(X)G_2(X) \tag{5}$$

with X denoting a standard normally distributed, bivariate random vector.
Based on Hermite polynomials, as defined in the univariate case, we can define a collection of orthogonal elements in $L^2(\mathbb{R}^2, \varphi_{I_2})$.

DEFINITION 2.4. For $x, y \in \mathbb{R}$ and $p, q \in \mathbb{N}$ we call $H_{p,q}$, defined by

$$H_{p,q}(x, y) := H_p(x)H_q(y),$$

a bivariate Hermite polynomial of degree $k = p + q$.

The collection of all bivariate Hermite polynomials forms an othogonal basis of $L^2(\mathbb{R}^2, \varphi_{I_2})$; see Beran et al. (2013), p. 122. In particular, the L^2-expansion of $G \in L^2(\mathbb{R}^2, \varphi_{I_2})$ is given by

$$G(X, Y) = \sum_{k=0}^{\infty} \sum_{p+q=k} \frac{1}{p!q!} J_{p,q} H_{p,q}(X, Y)$$

with $J_{p,q} = \mathrm{E}\left(G(X, Y)H_{p,q}(X, Y)\right)$. Naturally, we define the Hermite rank of this bivariate expansion in analogy to its definition in the univariate case.

DEFINITION 2.5. The index $r := \min\{p + q : J_{p,q} \neq 0\}$ which yields the first non-zero Hermite coefficient is called the Hermite rank of G.

2.1.3. *Long-range dependence.* In the following sections, we study the asymptotic behavior of distance correlation for long-range dependent time series. The rate of decay of the auto-covariance function is crucial to the definition of long-range dependent time series. A relatively slow decay of the auto-covariances characterizes long-range dependent time series, while a relatively fast decay characterizes short-range dependent processes; see Pipiras and Taqqu (2017), p. 17.

DEFINITION 2.6. A (second-order) stationary, real-valued time series X_k, $k \in \mathbb{Z}$, is called long-range dependent if its auto-covariance function ρ satisfies

$$\rho(k) := \mathrm{Cov}(X_1, X_{k+1}) \sim k^{-D} L(k), \text{ as } k \to \infty,$$

with $D \in (0,1)$ for some slowly varying function L. We refer to D as long-range dependence (LRD) parameter.

It follows from (4) that subordination of long-range dependent Gaussian time series potentially generates time series whose auto-covariances decay faster than the auto-covariances of the underlying Gaussian process. In some cases, the subordinated time series is long-range dependent as well, in other cases subordination may even yield short-range dependence. Given that $\mathrm{Cov}(\xi_1, \xi_{k+1}) \sim k^{-D} L(k)$, as $k \to \infty$, for some slowly varying function L and $D \in (0,1)$ and given that $G \in L^2(\mathbb{R}, \varphi(x)dx)$ is a function with Hermite rank r, we have

$$\mathrm{Cov}(G(\xi_1), G(\xi_{k+1})) \sim J_r^2(G) r! k^{-Dr} L^r(k), \text{ as } k \to \infty.$$

It immediately follows that subordinated Gaussian time series $G(\xi_n)$, $n \in \mathbb{N}$, are long-range dependent with LRD parameter $D_G := Dr$ and slowly-varying function $L_G(k) = J_r^2(G) r! L^r(k)$ whenever $Dr < 1$.

3. Main results. Recall that, given observations $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ stemming from the stationary components of a bivariate, covariance stationary stochastic process (X_k, Y_k), $k \geq 1$, our goal to decide on the testing problem (1).
As a test statistic for the above testing problem we consider the empirical distance covariance defined by

$$\begin{aligned}\mathcal{V}_n^2(X, Y; \omega) &= \|\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)\|_2^2 \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \left|\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)\right|^2 \omega(s,t) ds dt,\end{aligned}$$

where

$$\omega(s,t) = \left(c^2 |s|^{\alpha+1} |t|^{\alpha+1}\right)^{-1}$$

for some constant c and $\alpha \in (0,2)$. Székely, Rizzo and Bakirov (2007) argue that $\alpha = 1$ corresponds to a natural choice for the weight function. In the following, we will therefore assume that $\omega(s,t) = \left(c^2 s^2 t^2\right)^{-1}$ and we will only consider $\mathcal{V}_n^2(X,Y) := \mathcal{V}_n^2(X,Y;\omega)$.
Since, in this case, $\mathcal{V}^2(X,Y) = 0$ if and only if X and Y are independent, a statistical test based on $\mathcal{V}_n^2(X,Y)$ rejects the hypothesis for values of $\mathcal{V}_n^2(X,Y)$ exceeding a predefined critical value. In order to set critical values, we approximate the distribution of the test statistic by its asymptotic distribution.
Preceding literature focuses on data generated by independent random variables or short-range dependent time series X_k, $k \geq 1$, and Y_k, $k \geq 1$. Among these, Matteson and Tsay (2017) propose a methodology for estimating and testing the existence of mutually

independent components in multivariate data. Jin and Matteson (2018) generalize distance covariance from measuring pairwise dependence to measuring mutual dependence. Most recently, Kroll (2021) established the asymptotic distribution of the empirical distance covariance on separable metric spaces under the assumption of absolutely regular data-generating processes. Some authors focus on applications to high-dimensional time series: Kong et al. (2017) establish a two-stage interaction identification method based on the concept of distance correlation. Yao, Zhang and Shao (2016) introduce a test constructed on the basis of the pairwise distance covariance for identifying dependence structures in high dimensional data. Asymptotic normality of the test statistic is shown under moment assumptions and restrictions on the growth rate of the dimension as a function of the sample size. Gao et al. (2019) develop central limit theorems and establish rates of convergence for a rescaled test statistic based on the bias-corrected distance correlation in high dimensions when both, sample size and dimensionality, diverge.

3.1. *A limit theorem for Hilbert space valued random elements.* Instead of considering the empirical distance correlation as a V-statistic, we aim at basing our theoretical results on limit theorems for Hilbert space valued random elements. To this end, recall that for a measure space $(S,\mathcal{S},\mu)$ the set $\mathcal{L}^2(S,\mu) := \{f : S \longrightarrow \mathbb{C} \text{ measurable} \mid \|f\|_2 < \infty\}$, equipped with $\|\cdot\|_2$, where $\|f\|_2 := \left(\int_S |f|^2 \, d\mu\right)^{\frac{1}{2}}$, is a semi-normed vector space. The quotient space $L^2(S,\mu) := \mathcal{L}^2(S,\mu)/\mathcal{N}$ with $\mathcal{N} := \ker(\|\cdot\|_2)$, equipped with $\|\cdot\|_2$, is then a normed vector space and, in particular, a Hilbert space.

For the analysis of the distance covariance of two random vectors X and Y, we note that $\mathcal{V}_n^2(X,Y) = \|\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)\|_2^2$, where

$$\|f(s,t)\|_2^2 = \int_{\mathbb{R}}\int_{\mathbb{R}} |f(s,t)|^2 \, \omega(s,t) ds dt$$

for $f \in \mathcal{L}^2(\mathbb{R}^2, \omega(s,t)dsdt)$.

In order to derive the asymptotic distribution of $\mathcal{V}_n^2(X,Y)$, we consider the following decomposition:

$$\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t) = -\left(\varphi_Y^{(n)}(t) - \varphi_Y(t)\right)\left(\varphi_X^{(n)}(s) - \varphi_X(s)\right) + \frac{1}{n}\sum_{j=1}^{n} \left(\exp(isX_j) - \varphi_X(s)\right)\left(\exp(itY_j) - \varphi_Y(t)\right).$$

The two summands on the right-hand side of this equation will be considered separately in the following sections. For an analysis of the first summand, we make use of limit theorems for Hilbert space valued random elements. For this, we consider $\varphi_X^{(n)}(s) - \varphi_X(s)$ and $\varphi_Y^{(n)}(t) - \varphi_Y(t)$ as elements of the Hilbert space $L^2(\mathbb{R}, \omega(s)ds)$, where $\omega(s) := (cs^2)^{-1}$, and $\left(\varphi_Y^{(n)}(t) - \varphi_Y(t)\right)\left(\varphi_X^{(n)}(s) - \varphi_X(s)\right)$ as an element of $L^2(\mathbb{R}^2, \omega(s,t)dsdt)$.

The following lemma yields theoretical justification for these considerations.

LEMMA 3.1. *Let X_k, $k \geq 1$, and Y_k, $k \geq 1$, with $\mathrm{E}\,|X_1| < \infty$ and $\mathrm{E}\,|Y_1| < \infty$, be stationary processes. Then, it holds that*

$$\int_{\mathbb{R}} \left|\varphi_X^{(n)}(s) - \varphi_X(s)\right|^2 \omega(s) ds < \infty, \quad \int_{\mathbb{R}} \left|\varphi_Y^{(n)}(t) - \varphi_Y(t)\right|^2 \omega(t) ds < \infty,$$

and

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)\right|^2\omega(s,t)dsdt<\infty.$$

The proof of Lemma 3.1 is based on arguments that have been established in Székely, Rizzo and Bakirov (2007). It can be found in the appendix.
In the following, we establish a non-central limit theorem for processes with values in an L_2-Hilbert space. Against the background of analyzing the distance covariance of time series, this limit theorem provides a basis for deriving the asymptotic distribution of the distance covariance of subordinated Gaussian processes. Yet, the limit theorem is not specially geared to this problem and can thus be considered of particular and independent interest.

THEOREM 3.1. *Let X_k, $k\geq 1$, be a real-valued, stationary, long-range dependent Gaussian process with $\operatorname{E}X_1=0$, $\operatorname{Var}(X_1)=1$, and*

$$\rho(k):=\operatorname{Cov}(X_1,X_{1+k})\sim k^{-D}L(k)$$

for the LRD parameter $D\in(0,1)$ and a slowly varying function L. Given a positive weight function $\omega:\ \mathbb{R}\longrightarrow\mathbb{R}_+$, consider the Hilbert space $S:=L^2(\mathbb{R},\omega(t)dt)$. Let $f:\mathbb{R}\longrightarrow S$ map $x\in\mathbb{R}$ to the function $t\mapsto f_t(x)$ with $(t,x)\mapsto f_t(x)$ measurable and $f_t\in L^2(\mathbb{R},\varphi(x)dx)$ for all $t\in\mathbb{R}$. Moreover, assume that the LRD parameter D meets the condition $0<D<\frac{1}{r}$, where r denotes the Hermite rank of the class of functions $\{f_t(X_1)-\operatorname{E}f_t(X_1),\ t\in\mathbb{R}\}$, i.e., $r:=\min\{q\geq 1: J_q(t)\neq 0 \text{ for some } t\in\mathbb{R}\}$ with $J_q(t)=\operatorname{E}(f_t(X_1)H_q(X_1))$. If $\operatorname{E}\left(\|f_X\|_2^2\right)<\infty$, where $f_X(t):=f_t(X_1)$, then

$$\left\|n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^n\left[\left(f_t(X_j)-\operatorname{E}f_t(X_j)\right)-\frac{1}{r!}J_r(t)H_r(X_j)\right]\right\|_2=o_P(1).$$

Moreover, it follows that

$$n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^n\left(f_t(X_j)-\operatorname{E}f_t(X_j)\right)\xrightarrow{\mathcal{D}}\frac{1}{r!}J_r(t)Z_{r,H}(1),\ t\in\mathbb{R},$$

where $Z_{r,H}$ denotes an r-th order Hermite process with $H=1-\frac{rD}{2}$, and $\xrightarrow{\mathcal{D}}$ denotes weak convergence in $L^2(\mathbb{R},\omega(t)dt)$.

Noting that

$$\operatorname{E}(\cos(X_1)X_1)=0 \text{ and } \operatorname{E}(\sin(X_1)X_1)=\exp\left(-\frac{s^2}{2}\right)s,$$

an application of Theorem 3.1 establishes the following corollary:

COROLLARY 3.1. *Under the assumptions of Theorem 3.1, we have*

$$n^{\frac{D_X}{2}}L_X^{-\frac{1}{2}}(n)\left\|\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)-J_1(s)\frac{1}{n}\sum_{i=1}^nX_i\right\|_2=o_P(1).$$

Moreover, it follows that

$$n^{\frac{D_X}{2}}L_X^{-\frac{1}{2}}(n)\operatorname{Re}\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)\xrightarrow{\mathcal{D}}0,\ s\in\mathbb{R},$$

while

$$n^{\frac{D_X}{2}} L_X^{-\frac{1}{2}}(n)\mathrm{Im}\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right) \xrightarrow{\mathcal{D}} \exp\left(-\frac{s^2}{2}\right) sZ,\ s\in\mathbb{R},$$

where Z is a standard normally distributed random variable.

REMARK 3.1. Note that, since $\omega(s)=(cs^2)^{-1}$,

$$\int \left|\exp\left(-\frac{s^2}{2}\right) s\right|^2 \omega(s)ds < \infty,$$

such that the limit in the above corollary takes values in $L^2(\mathbb{R},\omega(s)ds)$.

PROOF OF THEOREM 3.1. Since, due to Fubini's theorem,

$$\int_{\mathbb{R}} \mathrm{E}\left|f_t(X_1)\right|^2 \omega(t)dt = \mathrm{E}\left(\|f_X\|_2^2\right) < \infty$$

by assumption, we have $\mathrm{E}\left|f_t(X_1)\right|^2<\infty$ for almost every t, so that it is possible to expand the function f_t in Hermite polynomials, meaning that

$$f_t(X_j) - \mathrm{E}f_t(X_j) \overset{L^2}{=} \sum_{q=r}^{\infty} \frac{J_q(t)}{q!} H_q(X_j),$$

i.e.,

$$\lim_{n\to\infty}\left\| f_t(X_j) - \mathrm{E}f_t(X_j) - \sum_{q=r}^{n} \frac{J_q(t)}{q!} H_q(X_j)\right\|_{L^2} = 0,$$

where $\|\cdot\|_{L^2}$ denotes the norm induced by the inner product (2).
We will see that the first summand in the Hermite expansion of the function f_t determines the asymptotic behavior of the partial sum process.
To this end, we show L_2-convergence of

$$n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^{n}\left(f_t(X_j)-\mathrm{E}f_t(X_j)-\frac{1}{r!}J_r(t)H_r(X_j)\right).$$

Fubini's theorem yields

$$\mathrm{E}\left(\int_{\mathbb{R}}\left|n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^{n}\left(f_t(X_j)-\mathrm{E}f_t(X_j)-\frac{1}{r!}J_r(t)H_r(X_j)\right)\right|^2\omega(t)dt\right)$$

$$=\int_{\mathbb{R}}\mathrm{E}\left(\left|n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^{n}\left(f_t(X_j)-\mathrm{E}f_t(X_j)-\frac{1}{r!}J_r(t)H_r(X_j)\right)\right|^2\right)\omega(t)dt.$$

Since $\mathrm{E}\left(H_q(X_i)H_{q'}(X_j)\right)=0$ for $q\neq q'$ and $\mathrm{E}\left(H_q(X_i)H_q(X_j)\right)=q!\rho(i-j)^q$, we have

$$\mathrm{E}\left(\left|\sum_{j=1}^{n}\left(f_t(X_j)-\mathrm{E}f_t(X_j)\right)-\frac{1}{r!}J_r(t)\sum_{j=1}^{n}H_r(X_j)\right|^2\right)$$

$$
\begin{aligned}
&= \mathrm{E}\left(\left|\sum_{j=1}^{n}\sum_{q=r+1}^{\infty}\frac{1}{q!}J_q(t)H_q(X_j)\right|^2\right) \\
&= \sum_{q=r+1}^{\infty}\frac{1}{q!^2}\left|J_q(t)\right|^2\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}\left(H_q(X_i)H_q(X_j)\right) \\
&= \sum_{q=r+1}^{\infty}\frac{1}{q!}\left|J_q(t)\right|^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\rho(i-j)\right|^q \\
&\leq \sum_{q=r+1}^{\infty}\frac{1}{q!}\left|J_q(t)\right|^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\rho(i-j)\right|^{r+1}.
\end{aligned}
$$

In general, i.e., for an auto-covariance function $\rho(k) = k^{-D}L(k)$, as $k \to \infty$, where $0 < D < 1$ and where L is a slowly varying function, it holds that

$$
\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\rho(i-j)\right|^{r+1} = \mathcal{O}\left(n^{1\vee(2-(r+1)D)}L'(n)\right),
$$

where L' is some slowly varying function; see p. 1777 in Dehling and Taqqu (1989). As a result, the previous considerations establish

$$
\begin{aligned}
&\mathrm{E}\left(\int_{\mathbb{R}}\left|n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^{n}\left(f_t(X_j)-\mathrm{E}f_t(X_j)-\frac{1}{r!}J_r(t)H_r(X_j)\right)\right|^2\omega(t)dt\right) \\
&= \mathcal{O}\left(n^{rD-2}L^{-r}(n)n^{1\vee(2-(r+1)D)}L'(n)\int_{\mathbb{R}}\sum_{q=r+1}^{\infty}\frac{1}{q!}\left|J_q(t)\right|^2\omega(t)dt\right).
\end{aligned}
$$

Since $\sum_{q=r+1}^{\infty}\frac{1}{q!}\left|J_q(t)\right|^2 \leq \sum_{q=0}^{\infty}\frac{1}{q!}\left|J_q(t)\right|^2 = \mathrm{E}\left|f_t(X_1)\right|^2$, we conclude that the right-hand side of the above equality is

$$
\mathcal{O}\left(n^{-\min(1-rD,D)}\tilde{L}(n)\int_{\mathbb{R}}\mathrm{E}\left|f_t(X_1)\right|^2\omega(t)dt\right)
$$

for some slowly varying function $\tilde{L}$. This expression is $o(n^{-\delta})$ for some $\delta > 0$ as $D < \frac{1}{r}$ and $\int_{\mathbb{R}}\mathrm{E}\left|f_t(X_1)\right|^2\omega(t)dt = \mathrm{E}\left(\|f_X\|_2^2\right) < \infty$ by assumption. The assertion then follows from the fact that

$$
n^{\frac{rD}{2}-1}L^{-\frac{r}{2}}(n)\sum_{j=1}^{n}H_r(X_j) \xrightarrow{\mathcal{D}} Z_{r,H}(1);
$$

see Taqqu (1979) and Dobrushin and Major (1979). □

3.2. *Convergence of the distance covariance.* Since

$$
\mathcal{V}_n^2(X,Y) = \int_{\mathbb{R}}\int_{\mathbb{R}}\left|\varphi_{X,Y}^{(n)}(s,t)-\varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)\right|^2\omega(s,t)dsdt,
$$

a limit theorem for $\varphi_{X,Y}^{(n)}(s,t)-\varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)$ as an element of an L^2-Hilbert-space would yield convergence of the distance covariance by means of an application of the continuous mapping theorem. This section focuses on establishing corresponding results. The following theorem summarizes these.

THEOREM 3.2. *Let X_k, $k \geq 1$, and Y_k, $k \geq 1$, be two independent, real-valued, stationary, long-range dependent Gaussian processes with* $\mathrm{E}X_1 = \mathrm{E}Y_1 = 0$, $\mathrm{Var}\,(X_1) = \mathrm{Var}\,(Y_1) = 1$,

$$\rho_X(k) = \mathrm{Cov}(X_1, X_{1+k}) = k^{-D_X} L_X(k) \text{ and } \rho_Y(k) = \mathrm{Cov}(Y_1, Y_{1+k}) = k^{-D_Y} L_Y(k)$$

for $D_X, D_Y \in (0,1)$ *and slowly varying functions* L_X *and* L_Y.

1. *For* $D_X, D_Y \in (\frac{1}{2}, 1)$, *it holds that*

$$\sqrt{n}\left(\varphi^{(n)}_{X,Y}(s,t) - \varphi^{(n)}_X(s)\varphi^{(n)}_Y(t)\right) \xrightarrow{\mathcal{D}} Z(s,t),$$

where $Z(s,t) \sim \mathcal{CN}(0, \Gamma_{s,t}, C_{s,t})$ *with*

$$\Gamma_{s,t} := \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1)\overline{f_{s,t}(X_{k+1}, Y_{k+1})}\right), \tag{6}$$

$$C_{s,t} := \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1) f_{s,t}(X_{k+1}, Y_{k+1})\right),$$

where

$$f_{s,t}(X_j, Y_j) := (\exp(isX_j) - \varphi_X(s))\,(\exp(itY_j) - \varphi_Y(t)),$$

$$\mathrm{Cov}\left(Z(s,t), Z(s',t')\right) = \mathrm{E}\left(Z(s,t)\overline{Z(s',t')}\right) = \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1)\overline{f_{s',t'}(X_{k+1}, Y_{k+1})}\right),$$

$$\mathrm{Cov}\left(Z(s,t), \overline{Z(s',t')}\right) = \mathrm{E}\left(Z(s,t)Z(s',t')\right) = \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1) f_{s',t'}(X_{k+1}, Y_{k+1})\right).$$

2. *Assume that* $D := D_X = D_Y$ *and* $L := L_X = CL_Y$ *for some constant* C. *Then, for* $D \in \left(0, \frac{1}{2}\right)$, *it holds that*

$$n^D L^{-1}(n)\left(\varphi^{(n)}_{X,Y}(s,t) - \varphi^{(n)}_X(s)\varphi^{(n)}_Y(t)\right) \xrightarrow{\mathcal{D}}$$

$$st\exp\left(-\frac{s^2+t^2}{2}\right)\int_{[-\pi,\pi)^2}\left[\left(\frac{\mathrm{e}^{ix}-1}{ix}\right)\left(\frac{\mathrm{e}^{iy}-1}{iy}\right) - \frac{\mathrm{e}^{i(x+y)}-1}{i(x+y)}\right] Z_{G,X}(dx) Z_{G,Y}(dy),$$

where $Z_{G,X}$ *and* $Z_{G,Y}$ *are random spectral measures defined subsequently by* (9).

REMARK 3.2. 1. A complex Gaussian random variable Z is characterized by three parameters: the location parameter $\mu := \mathrm{E}Z$, the covariance matrix $\Gamma := \mathrm{E}[(Z - \mu)\overline{(Z-\mu)}]$, and the relation matrix $C := \mathrm{E}[(Z-\mu)(Z-\mu)^{\mathrm{T}}]$ with Z^{T} corresponding to the matrix transpose of Z. It is then denoted as $Z \sim \mathcal{CN}(\mu, \Gamma, C)$; see Goodman (1963).

2. The case $D_X, D_Y \in \left(0, \frac{1}{2}\right)$ is atypical in practice. Typically, we encounter $D_X, D_Y \in \left(\frac{1}{2}, 1\right)$. For this reason, we focus on the latter case.

In the following, we outline a step-by-step proof of Theorem 3.2. For this purpose, recall that

$$\varphi_{X,Y}^{(n)}(s,t)-\varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)=-\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)+\frac{1}{n}\sum_{j=1}^{n}\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(itY_j)-\varphi_Y(t)\right). \tag{7}$$

Under the assumption of independence or short-range dependence within the sequences X_k, $k\geq 1$, and Y_k, $k\geq 1$,

$$\sqrt{n}\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)=o_P(1),$$

i.e., with a corresponding normalization the first summand on the right-hand side of the above equation is asymptotically negligible, while the second summand determines the asymptotic distribution of the left-hand side. For long-range dependent time series X_k, $k\geq 1$, and Y_k, $k\geq 1$, the asymptotic behavior of the second summand depends on the values of D_X and D_Y, such that both summands may contribute to the limit distribution. The first summand, however, can be treated by an argument that is not sensitive to the values of these parameters. We consider the two summands separately. For these, we state intermediate results while detailed proofs are left to the appendix.
For the first summand, we prove the following result:

PROPOSITION 3.1. *Let X_k, $k\geq 1$, and Y_k, $k\geq 1$, be two independent, real-valued, stationary, long-range dependent Gaussian processes with* $\mathrm{E}X_1=\mathrm{E}Y_1=0$, $\mathrm{Var}\,(X_1)=\mathrm{Var}\,(Y_1)=1$,

$$\rho_X(k)=\mathrm{Cov}(X_1,X_{1+k})=k^{-D_X}L_X(k) \text{ and } \rho_Y(k)=\mathrm{Cov}(Y_1,Y_{1+k})=k^{-D_Y}L_Y(k)$$

for $D_X,D_Y\in(0,1)$ and slowly varying functions L_X and L_Y. Then, it holds that

$$\left\|\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)-J_1(s)\frac{1}{n}\sum_{j=1}^{n}X_jJ_1(t)\frac{1}{n}\sum_{j=1}^{n}Y_j\right\|_2=o_P\left(n^{-\frac{D_X+D_Y}{2}}L_X^{\frac{1}{2}}(n)L_Y^{\frac{1}{2}}(n)\right), \tag{8}$$

where $J_1(s)=i\exp\left(-\frac{s^2}{2}\right)s$. Moreover, it follows that

$$n^{\frac{D_X+D_Y}{2}}L_X^{-\frac{1}{2}}(n)L_Y^{-\frac{1}{2}}(n)\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)\xrightarrow{\mathcal{D}}st\exp\left(-\frac{s^2+t^2}{2}\right)Z_XZ_Y,$$

where Z_X, Z_Y are independent standard normally distributed random variables.

The asymptotic behavior of the second summand in the decomposition in formula (7) depends on the values of the long-range dependence parameters D_X and D_Y. For this reason, the following two propositions treat different values of these separately. Initially, we consider the case $D_X,D_Y\in(0,\frac{1}{2})$. Following this, we focus on the case $D_X,D_Y\in\left(\frac{1}{2},1\right)$.

PROPOSITION 3.2. *Let X_k, $k\geq 1$, and Y_k, $k\geq 1$, be two independent, real-valued, stationary, long-range dependent Gaussian processes with* $\mathrm{E}X_1=\mathrm{E}Y_1=0$,

$\text{Var}(X_1) = \text{Var}(Y_1) = 1$, $\rho_X(k) = k^{-D}L(k)$, *and* $\rho_Y(k) = Ck^{-D}L(k)$ *for a constant* C, *for* $D \in (0,1)$ *and a slowly varying function* L. *Then, for* $D \in (0, \frac{1}{2})$, *it holds that*

$$\left\| \frac{1}{n}\sum_{j=1}^{n} (\exp(isX_j) - \varphi_X(s))(\exp(itY_j) - \varphi_Y(t)) + \exp\left(-\frac{s^2}{2}\right) s \exp\left(-\frac{t^2}{2}\right) t \frac{1}{n}\sum_{j=1}^{n} X_jY_j \right\|_2 = o_P\left(n^{-D}L(n)\right).$$

Taking Proposition 3.1 into consideration and noting that

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \left| \exp\left(-\frac{s^2}{2}\right) s \exp\left(-\frac{t^2}{2}\right) t \right|^2 \omega(s,t)dsdt = \frac{\pi}{c^2},$$

it follows that the limit of

$$\begin{aligned} &n^{D_X+D_Y} L_X^{-1}(n) L_Y^{-1}(n) \mathcal{V}_n^2(X,Y) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \left| n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \left(\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t) \right) \right|^2 \omega(s,t)dsdt \end{aligned}$$

corresponds to the limit of

$$\frac{\pi}{c^2}\left(n^{\frac{D_X+D_Y}{2}-2} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \sum_{i=1}^{n}\sum_{j=1}^{n} X_iY_j - n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \sum_{j=1}^{n} X_jY_j \right)^2.$$

In order to derive the limit distribution of the above expression, we make use of the theory on spectral distributions established in Major (2020). For this, we consider the following representation of Gaussian random variables:

$$X_j = \int_{[-\pi,\pi)} \mathrm{e}^{ijx} Z_{G,X}(dx), \; Y_j = \int_{[-\pi,\pi)} \mathrm{e}^{ijy} Z_{G,Y}(dy), \tag{9}$$

where $Z_{G,X}$ and $Z_{G,Y}$ are corresponding random spectral measures determined by the positive semidefinite matrix-valued, even measure $(G_{j,j'})$, $1 \leq j, j' \leq 2$, on the torus $[-\pi, \pi)$ with coordinates $G_{j,j'}$ satisfying

$$\mathrm{E}(X_jX_{j+k}) = \int_{[-\pi,\pi)} \mathrm{e}^{ikx} G_{1,1}(dx), \qquad \mathrm{E}(Y_jY_{j+k}) = \int_{[-\pi,\pi)} \mathrm{e}^{ikx} G_{2,2}(dx),$$

$$\mathrm{E}(X_jY_{j+k}) = \int_{[-\pi,\pi)} \mathrm{e}^{ikx} G_{1,2}(dx) = 0, \qquad \mathrm{E}(Y_jX_{j+k}) = \int_{[-\pi,\pi)} \mathrm{e}^{ikx} G_{2,1}(dx) = 0.$$

PROPOSITION 3.3. *Let* X_k, $k \geq 1$, *and* Y_k, $k \geq 1$, *be two independent, real-valued, stationary, long-range dependent Gaussian processes with* $\mathrm{E}X_1 = \mathrm{E}Y_1 = 0$, $\text{Var}(X_1) = \text{Var}(Y_1) = 1$,

$$\rho_X(k) = \text{Cov}(X_1, X_{1+k}) = k^{-D_X}L_X(k) \textit{ and } \rho_Y(k) = \text{Cov}(Y_1, Y_{1+k}) = k^{-D_Y}L_Y(k)$$

for $D_X, D_Y \in (0,1)$ *and slowly varying functions* L_X *and* L_Y. *Then, it holds that*

$$\begin{aligned} &n^{\frac{D_X+D_Y}{2}-2} L_X^{-\frac{1}{2}} L_Y^{-\frac{1}{2}} \sum_{i=1}^{n}\sum_{j=1}^{n} X_iY_j - n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}} L_Y^{-\frac{1}{2}} \sum_{j=1}^{n} X_jY_j \xrightarrow{\mathcal{D}} \\ &\int_{[-\pi,\pi)^2} \left[\left(\frac{\mathrm{e}^{ix}-1}{ix} \right)\left(\frac{\mathrm{e}^{iy}-1}{iy} \right) - \frac{\mathrm{e}^{i(x+y)}-1}{i(x+y)} \right] Z_{G,X}(dx) Z_{G,Y}(dy). \end{aligned}$$

For $D_X, D_Y \in \left(\frac{1}{2}, 1\right)$, we derive the asymptotic distribution of the second summand in the decomposition (7) under the general assumption of subordinated Gaussian processes $X_k = G(\xi_k)$, $k \geq 1$, and $Y_k = G(\eta_k)$, $k \geq 1$. Under some restriction on the values of the parameters characterizing the dependence within X_k, $k \geq 1$, and Y_k, $k \geq 1$, the limit is a complex Gaussian random variable.

PROPOSITION 3.4. *Let $X_k = G(\xi_k)$, $k \geq 1$, and $Y_k = G(\eta_k)$, $k \geq 1$, where ξ_k, $k \geq 1$, and η_k, $k \geq 1$, are two independent, real-valued, stationary, long-range dependent Gaussian processes with $\mathrm{E}\xi_1 = \mathrm{E}\eta_1 = 0$, $\mathrm{Var}(\xi_1) = \mathrm{Var}(\eta_1) = 1$,*

$$\rho_\xi(k) = \mathrm{Cov}(\xi_1, \xi_{1+k}) = k^{-D_\xi} L_\xi(k) \text{ and } \rho_\eta(k) = \mathrm{Cov}(\eta_1, \eta_{1+k}) = k^{-D_\eta} L_\eta(k)$$

for $D_\xi, D_\eta \in (\frac{1}{2}, 1)$ and slowly varying functions L_ξ and L_η. Then, it holds that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\exp(isX_j) - \varphi_X(s)\right)\left(\exp(itY_j) - \varphi_Y(t)\right) \xrightarrow{\mathcal{D}} Z(s,t),$$

where $Z(s,t) \sim \mathcal{CN}(0, \Gamma_{s,t}, C_{s,t})$ with

$$\Gamma_{s,t} := \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1)\overline{f_{s,t}(X_{k+1}, Y_{k+1})}\right), \tag{10}$$

$$C_{s,t} := \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1) f_{s,t}(X_{k+1}, Y_{k+1})\right),$$

where

$$f_{s,t}(X_j, Y_j) := \left(\exp(isX_j) - \varphi_X(s)\right)\left(\exp(itY_j) - \varphi_Y(t)\right),$$

$$\mathrm{Cov}\left(Z(s,t), Z(s',t')\right) = \mathrm{E}\left(Z(s,t)\overline{Z(s',t')}\right) = \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1)\overline{f_{s',t'}(X_{k+1}, Y_{k+1})}\right),$$

$$\mathrm{Cov}\left(Z(s,t), \overline{Z(s',t')}\right) = \mathrm{E}\left(Z(s,t) Z(s',t')\right) = \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1, Y_1) f_{s',t'}(X_{k+1}, Y_{k+1})\right).$$

According to Proposition 3.1, we have

$$\sqrt{n}\left(\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)\right)$$
$$= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\exp(isX_j) - \varphi_X(s)\right)\left(\exp(itY_j) - \varphi_Y(t)\right) + o_P(1),$$

as $D_X + D_Y > 1$. A computation of the limit parameters $\Gamma_{s,t}$ and $C_{s,t}$ specified in formula (10) then establishes the following corollary:

COROLLARY 3.2. *Let X_k, $k \geq 1$, and Y_k, $k \geq 1$, be two independent, real-valued, stationary, long-range dependent Gaussian processes with $\mathrm{E}X_1 = \mathrm{E}Y_1 = 0$, $\mathrm{Var}(X_1) = \mathrm{Var}(Y_1) = 1$,*

$$\rho_X(k) = \mathrm{Cov}(X_1, X_{1+k}) = k^{-D_X} L_X(k) \text{ and } \rho_Y(k) = \mathrm{Cov}(Y_1, Y_{1+k}) = k^{-D_Y} L_Y(k)$$

for $D_X, D_Y \in (0,1)$ and slowly varying functions L_X and L_Y. Then, for $D_X, D_Y \in (\frac{1}{2}, 1)$ it holds that

$$\sqrt{n}\left(\varphi_{X,Y}^{(n)}(s,t) - \varphi_X^{(n)}(s)\varphi_Y^{(n)}(t)\right) \xrightarrow{\mathcal{D}} \mathcal{CN}(0, \Gamma_{s,t}, C_{s,t})$$

with

$$\Gamma_{s,t} = \sum_{k=-\infty}^{\infty} \left(\exp(-s^2(1-\rho_X(k))) - \exp(-s^2)\right)\left(\exp(-t^2(1-\rho_Y(k))) - \exp(-t^2)\right)$$

and

$$C_{s,t} = \sum_{k=-\infty}^{\infty} \left(\exp(-s^2(1+\rho_X(k))) - \exp(-s^2)\right)\left(\exp(-t^2(1+\rho_Y(k))) - \exp(-t^2)\right).$$

4. Finite sample performance. Based on our theoretical results, we conclude that for $D := D_X = D_Y \in \left(0, \frac{1}{2}\right)$, $n^D L^{-1}(n) V_n(X,Y)$ converges in distribution to a non-degenerate limit, while for $D_X, D_Y \in \left(\frac{1}{2}, 1\right)$, this holds for $\sqrt{n} V_n(X,Y)$. As in both cases the limits are not feasible, we base test decisions on a subsampling procedure.

4.1. *Subsampling.* For the theoretical results on subsampling, we consider a more general situation: given observations $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ stemming from real-valued time series X_k, $k \geq 1$, and Y_k, $k \geq 1$, our goal is to decide on the testing problem

H_0: X_k, $k \geq 1$, and Y_k, $k \geq 1$, are independent,

H_1: X_k, $k \geq 1$, and Y_k, $k \geq 1$, are dependent.

For this purpose, we consider a test statistic $T_n := T_n(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$, such that our goal is to approximate the distribution F_{T_n} of T_n under the assumption of independence. Therefore, the subsampling procedure has to be designed in such a way that it mimics the behavior of the test statistic for two independent time series, given both: data generated according to the model assumptions under the hypothesis as well as data generated according to the model assumptions under the alternative. To not destroy the dependence structure of the individual time series, it seems reasonable to consider blocks of observations. For this, we define blocks

$$B_{k,l_n} := (X_k, \ldots, X_{k+l_n-1}),\ C_{k,l_n} := (Y_k, \ldots, Y_{k+l_n-1}).$$

To mimic the behavior of two independent time series, it seems reasonable to compute the distance correlation of blocks that are far apart. For this reason, we compute the test statistic on blocks that are separated by a lag $d_n \in \{1, \ldots, n-l_n\}$, i.e., for $k = 1, \ldots, n-l_n-d_n$ we compute

$$T_{l_n,k} := T_{l_n}\left(X_k, \ldots, X_{k+l_n-1}, Y_{k+d_n}, \ldots, Y_{k+d_n+l_n-1}\right),\ k = 1, \ldots, m_n,$$

where $m_n := n - l_n - d_n$. As a result, we obtain multiple (though dependent) realizations of the test statistic T_{l_n}. Due to the fact that consecutive observations are chosen, the subsamples retain the dependence structure of the original sample, so that the empirical distribution function of $T_{l_n,1}, \ldots, T_{l_n,m_n}$, defined by

$$\widehat{F}_{m_n,l_n}(t) := \frac{1}{m_n}\sum_{k=1}^{m_n} 1_{\{T_{l_n,k} \leq t\}}, \tag{11}$$

can be considered as an appropriate estimator for F_{T_n}.
In order to establish the validity of the subsampling procedure, i.e., in order to show that the empirical distribution function of $T_{l_n,1}, \ldots, T_{l_n,m_n}$ can be considered as a suitable approximation of F_{T_n}, we aim at proving that the distance between $\widehat{F}_{m_n,l_n}$ and F_{T_n} vanishes as the number of observations tends to ∞.

DEFINITION 4.1. The subsampling procedure is said to be *consistent* if

$$\left|\widehat{F}_{m_n,l_n}(t) - F_{T_n}(t)\right| \xrightarrow{\mathcal{P}} 0, \text{ as } n \to \infty,$$

for all points of continuity t of F_T.

REMARK 4.1. If the subsampling procedure is consistent in the sense of Definition 4.1, and if F_T is continuous, the usual Glivenko-Cantelli argument for uniform convergence of empirical distribution functions implies that

$$\sup_{t\in\mathbb{R}}\left|\widehat{F}_{m_n,l_n}(t) - F_{T_n}(t)\right| \xrightarrow{\mathcal{P}} 0, \text{ as } n \to \infty.$$

In order to prove consistency, we have to make the following technical assumptions:

ASSUMPTION 1. Let ξ_k, $k \geq 1$, denote a stationary, long-range dependent Gaussian process with mean 0, variance 1, LRD parameter D and spectral density $f(\lambda) = |\lambda|^{D-1}L_f(\lambda)$ for a slowly varying function L_f which is bounded away from 0 on $[0,\pi]$. Moreover, assume that $\lim_{\lambda\to 0} L_f(\lambda) \in (0,\infty]$ exists.

ASSUMPTION 2. Let ξ_k, $k \geq 1$, denote a stationary, long-range dependent Gaussian process with mean 0, variance 1 and covariance function $\rho(k) := \operatorname{Cov}(\xi_1, \xi_{k+1}) = k^{-D}L_\rho(k)$ for some parameter $D \in (0,1)$ and some slowly varying function L_ρ. Assume that there exists a constant $K \in (0,\infty)$, such that for all $k \in \mathbb{N}$

$$\max_{k+1\leq j\leq k+2l-2} |L_\rho(k) - L_\rho(j)| \leq K\frac{l}{k}\min\{L_\rho(k), 1\}$$

for $l \in \{l_k, \ldots, k\}$.

Given Assumptions 1 and 2, consistency of the subsampling procedure is established by the following theorem:

THEOREM 4.1. *Given two independent, stationary, subordinated Gaussian LRD time series X_k, $k \geq 1$, and Yk, $k \geq 1$, each satisfying Assumptions 1 and 2 with LRD parameters D_X and D_Y and a series of (measurable) statistics $T_n = T_n(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$ that converge in distribution to a (non-degenerate) random variable T. Let F_T and F_{T_n} denote the distribution functions of T and T_n. Moreover, let l_n, $n \in \mathbb{N}$, be an increasing, divergent series of integers. If $l_n = \mathcal{O}\left(n^{(1+\min(D_X,D_Y))/2-\varepsilon}\right)$ for some $\varepsilon > 0$, then, for some $\beta \in (0,1)$,*

$$\sum_{k=\lfloor n^\beta\rfloor+1}^{m_n} \rho_{k,l_n,d_n} = o(m_n).$$

In particular, it follows that the sampling-window method is consistent under these assumptions.

The proof of Theorem 4.1 is based on arguments that have been established in Betken and Wendler (2018). It can be found in the appendix.

4.2. *Simulations.* So far, we focused on analyzing the asymptotic behavior of the distance covariance with respect to data (X_i, Y_i), $i = 1, \ldots, n$, stemming from long-range dependent time series X_k, $k \geq 1$, and Y_k, $k \geq 1$.
In the following, we will assess the finite sample performance of a hypothesis test based on the distance covariance. In particular, we will compare its finite sample performance to that of a hypothesis test based on the sample covariance

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}).$$

Prior to a comparison of the finite sample performance of the two dependence measures, we derive a limit theorem for the sample covariance complementing our main theoretical results stated in Theorem 3.2.

THEOREM 4.2. *Let X_k, $k \geq 1$, and Y_k, $k \geq 1$, be two independent, real-valued, stationary, long-range dependent Gaussian processes with* $\mathrm{E}X_1 = \mathrm{E}Y_1 = 0$, $\mathrm{Var}\,(X_1) = \mathrm{Var}\,(Y_1) = 1$, $\rho_X(k) = \mathrm{Cov}(X_1, X_{1+k}) = k^{-D_X} L_X(k)$, *and* $\rho_Y(k) = \mathrm{Cov}(Y_1, Y_{1+k}) = k^{-D_Y} L_Y(k)$. *For $D_X, D_Y \in (0, \frac{1}{2})$, it holds that*

$$n^{\frac{D_X + D_Y}{2}} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) \xrightarrow{\mathcal{D}} Z_X Z_Y,$$

where $\xrightarrow{\mathcal{D}}$ denotes weak convergence in $L^2\left(\mathbb{R}^2, \omega(s,t)dt\right)$ and where Z_X and Z_Y are two independent, standard normally distributed random variables. For $D_X, D_Y \in (\frac{1}{2}, 1)$, it holds that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \rho_X(k)\rho_Y(k).$$

In order to compare the performance of distance correlation to that of Pearson's correlation coefficient, we consider four different scenarios:

1. "linearly" correlated data, i.e., we simulate $k = 5000$ repetitions of $(X_1, \ldots, X_n, Y_1 \ldots, Y_n) := (2\Phi(Z_1) - 1, \ldots, 2\Phi(Z_{2n}) - 1)$, where $(Z_1, \ldots, Z_{2n})$ is multivariate normally distributed with mean 0 and covariance matrix

$$\sigma_{i,j} := \begin{cases} \rho(|i-j|) & \text{for } 1 \leq i, j \leq n \\ \sigma_{i-n, j-n} & \text{for } n+1 \leq i, j \leq 2n \\ r\sigma_{i-n, j} & \text{for } n+1 \leq i \leq 2n, 1 \leq j \leq n \\ r\sigma_{i, j-n} & \text{for } 1 \leq i \leq n, n+1 \leq j \leq 2n \end{cases},$$

   where

$$\rho(k) = \frac{1}{2}\left(|k+1|^{2H} - 2\,|k|^{2H} + |k-1|^{2H}\right);$$

   see Figure 1 for an illustration of different parameter combinations.

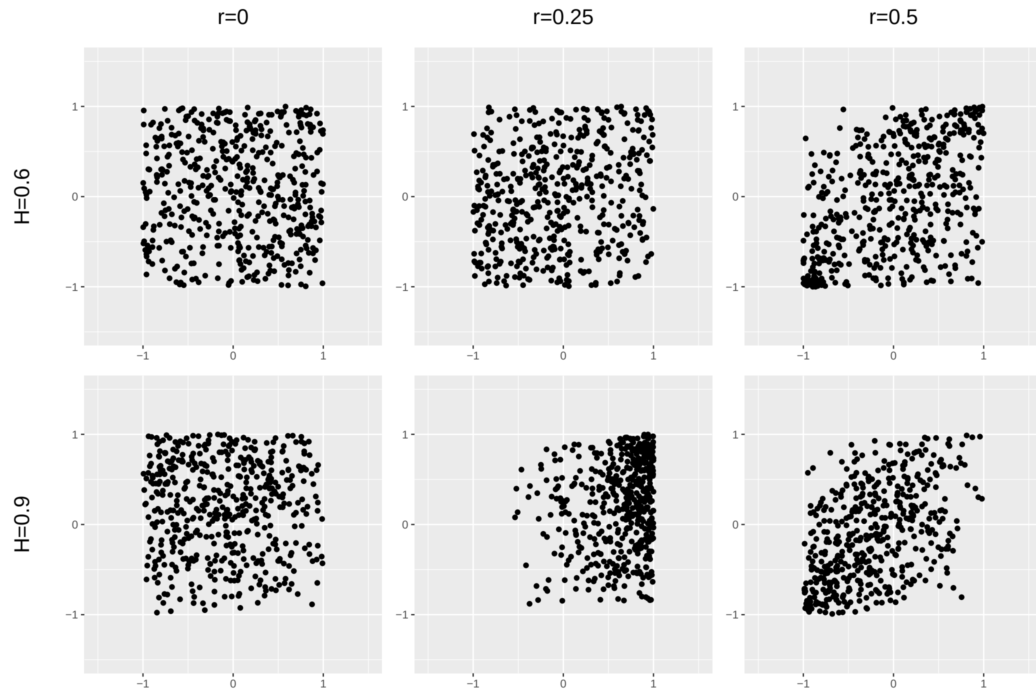

Fig 1: "Linearly" correlated data (X_i, Y_i), $i = 1, \ldots, 500$, with parameters H and r.

2. "parabolically" correlated data, i.e., we simulate $k = 5000$ repetitions of $(X_1, \ldots, X_n, Y_1 \ldots, Y_n)$, where $(X_1, \ldots, X_n) := (2\Phi(Z_1) - 1, \ldots, 2\Phi(Z_n) - 1)$, for fractional Gaussian noise Z_i, $i \geq 1$, with parameter H, and

$$Y_i := v\left(X_i^2 - \frac{1}{3}\right) + w\xi_i, \; w = \sqrt{1 - \frac{4}{15}v^2}, \tag{12}$$

where $\xi_1, \ldots, \xi_n$ are independent uniformly on $[-1, 1]$ distributed random variables; see Figure 2 for an illustration of different parameter combinations. The choice of the parameter w guarantees $\mathrm{E}Y_1 = \mathrm{E}X_1 = 0$ and $\mathrm{Var}\,(Y_1) = \mathrm{Var}\,(X_1) = \frac{1}{3}$.
3. "wavily" correlated data, i.e., we simulate $k = 5000$ repetitions of $(X_1, \ldots, X_n, Y_1 \ldots, Y_n)$, where $(X_1, \ldots, X_n) := (2\Phi(Z_1) - 1, \ldots, 2\Phi(Z_n) - 1)$, for fractional Gaussian noise Z_i, $i \geq 1$, with parameter H, and

$$Y_j := v\left(\left(X_j^2 - \frac{1}{3}\right)^2 - 3/45\right) + w\xi_j, \; w = \sqrt{1 - 242/4725v^2}, \tag{13}$$

where ξ_j, $1 \leq j \leq n$, are independent $\mathcal{U}[-1, 1]$ distributed random variables; see Figure 3 for an illustration of different parameter combinations. The choice of the parameter w guarantees $\mathrm{E}Y_1 = \mathrm{E}X_1 = 0$ and $\mathrm{Var}\,(Y_1) = \mathrm{Var}\,(X_1) = \frac{1}{3}$.
4. "rectangularly" correlated data, i.e., we simulate $k = 5000$ repetitions of $(X_1, \ldots, X_n, Y_1 \ldots, Y_n)$, where

$$\begin{pmatrix} X_1 & Y_1 \\ X_2 & Y_2 \\ \ldots & \ldots \\ X_n & Y_n \end{pmatrix} = \begin{pmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \\ \ldots & \ldots \\ \xi_n & \eta_n \end{pmatrix} \begin{pmatrix} \cos(\frac{\pi}{12}v) & -\sin(\frac{\pi}{12}v) \\ \sin(\frac{\pi}{12}v) & \cos(\frac{\pi}{12}v) \end{pmatrix} \tag{14}$$

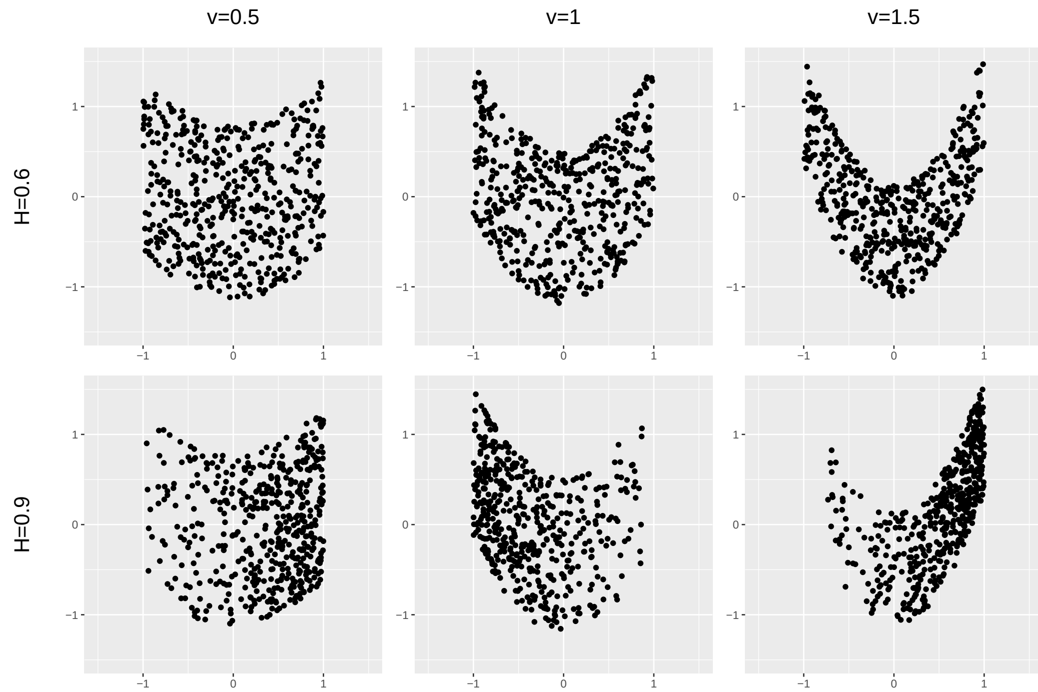

Fig 2: “Parabolically” correlated data (X_i, Y_i), $i = 1, \ldots, 500$, with parameters H and v.

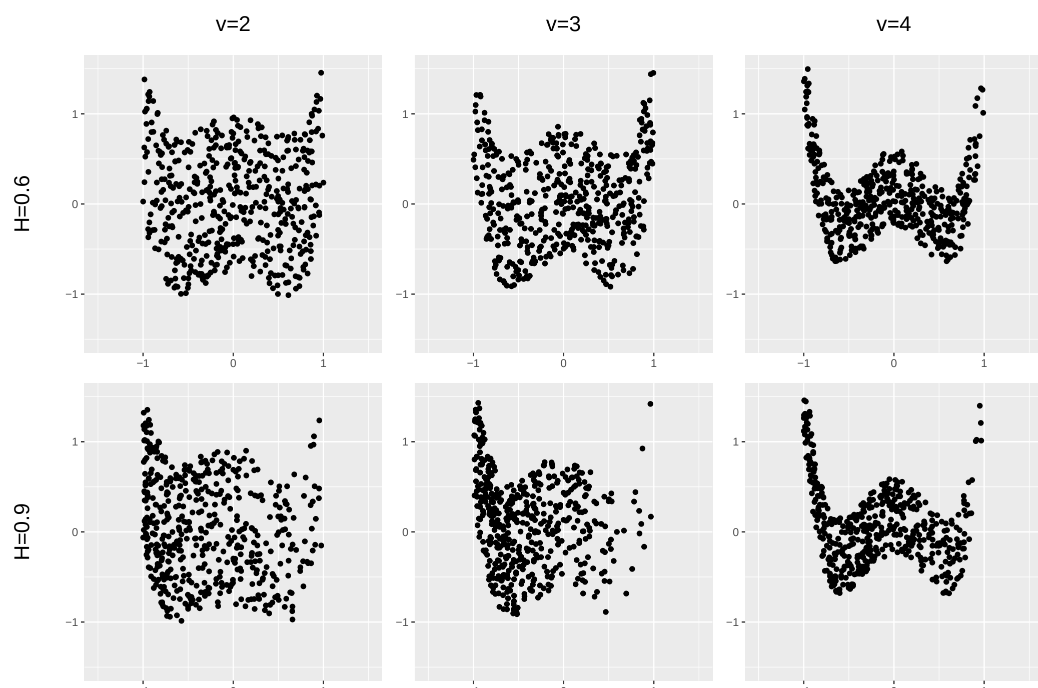

Fig 3: “Wavily” correlated data (X_i, Y_i), $i = 1, \ldots, 500$, with parameters H and v.

and for two independent fractional Gaussian noise sequences Z_i, $i \geq 1$, and $\tilde{Z}_i$, $i \geq 1$, each with parameter H, $(X_1, \ldots, X_n) := (2\Phi(Z_1) - 1, \ldots, 2\Phi(Z_n) - 1)$ and $(Y_1, \ldots, Y_n) := (2\Phi(\tilde{Z}_1) - 1, \ldots, 2\Phi(\tilde{Z}_n) - 1)$; see Figure 4 for an illustration of different parameter combinations.

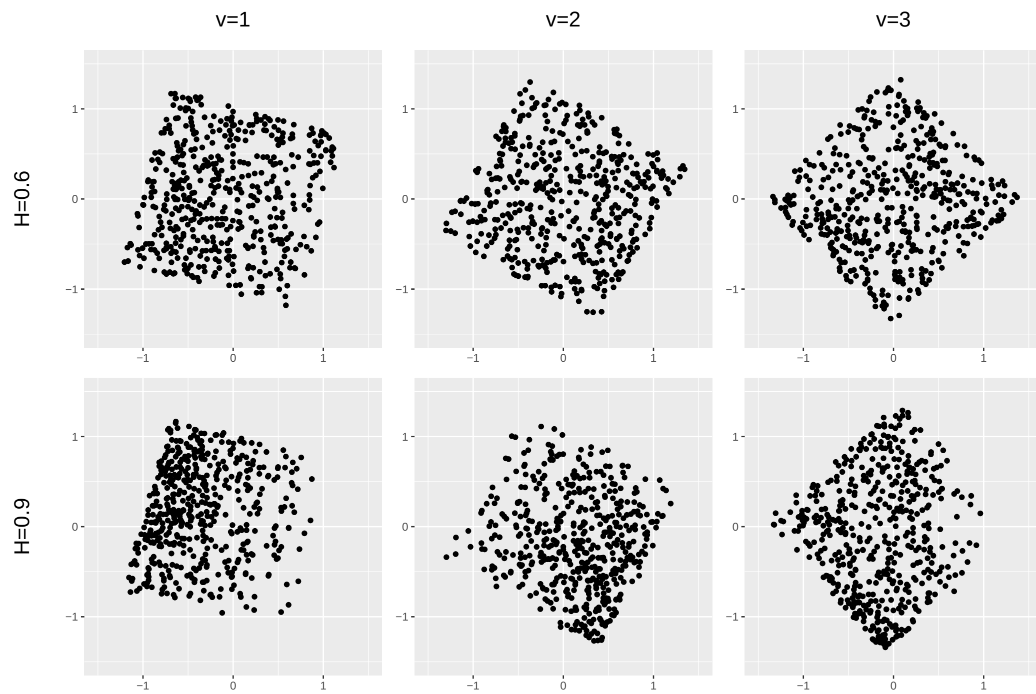

Fig 4: "Rectangularly" correlated data ($n = 500$, $H = 0.6$).

All calculations are based on $5,000$ realizations of simulated time series and test decisions are based on an application of the sampling-window method for a significance level of 5%, meaning that the values of the test statistics are compared to the 95%-quantile of the empirical distribution function $\widehat{F}_{m_n,l_n}$ defined by (11). The fractional Gaussian noise sequences are generated by the function `simFGN0` from the `longmemo` package in `R`. Detailed simulation results can be found in Tables 1 - 8 in the appendix. These display results for sample sizes $n = 100, 300, 500, 1000$, block lengths $l_n = \lfloor n^\gamma \rfloor$ with $\gamma \in \{0.4, 0.5, 0.6\}$, Hurst parameters $H = 0.6, 0.7, 0.8, 0.9$ and different values of the parameters r and v.

As a whole, the simulation results concur with the expected behaviour of hypothesis tests for independence of time series: For both testing procedures, an increasing sample size goes along with an improvement of the finite sample performance of the test, i.e., the empirical size (that can be found in the columns of Tables 1 and 2 superscribed by $r = 0$) approaches the level of significance and the empirical power increases; stronger deviations from the hypothesis, i.e., an increase of the parameters r and v leads to an increase of the rejection rates. Moreover, the testing procedures seem to be sensitive to a dependence within the single random vectors, as an increase of the Hurst parameter H results in significantly higher or lower rejection rates.

Both testing procedures tend to be oversized for small sample sizes. Tables 1 and 2 show that linear correlation as well as independence of two random vectors are slightly better

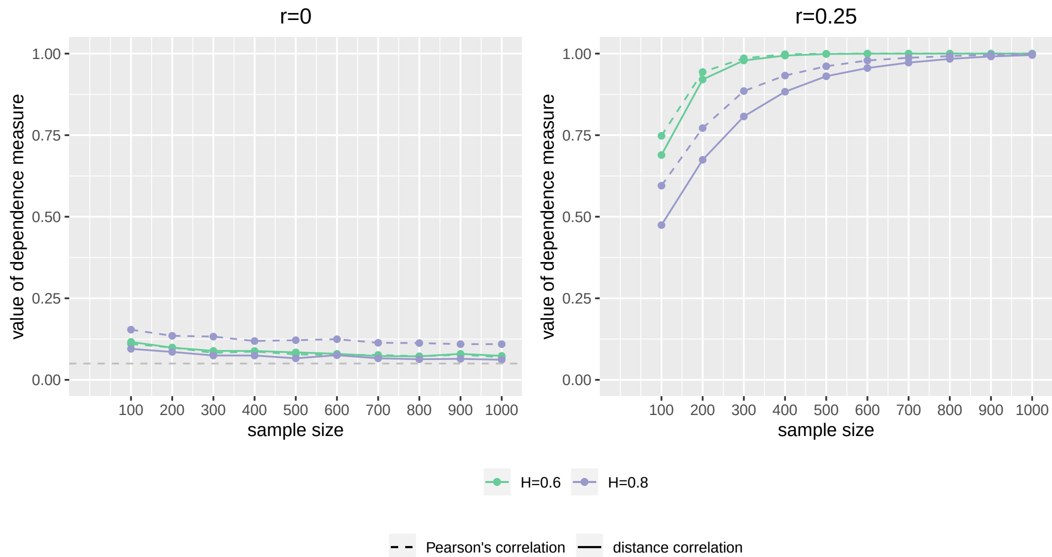

Fig 5: Rejection rates of the hypothesis tests resulting from Pearson's correlation and distance correlation obtained by subsampling based on "linearly" correlated fractional Gaussian noise time series X_j, $j = 1, \ldots, n$, Y_j, $j = 1, \ldots, n$ with block length $l_n = \lfloor \sqrt{n} \rfloor$, $d = 0.1n$, Hurst parameters $H = 0.6, 0.8$, and cross-correlation $\text{Cov}(X_i, Y_j) = r\text{Cov}(X_i, X_j)$, $1 \leq i, j \leq n$, with $r = 0$ and $r = 0.25$. The level of significance equals 5%.

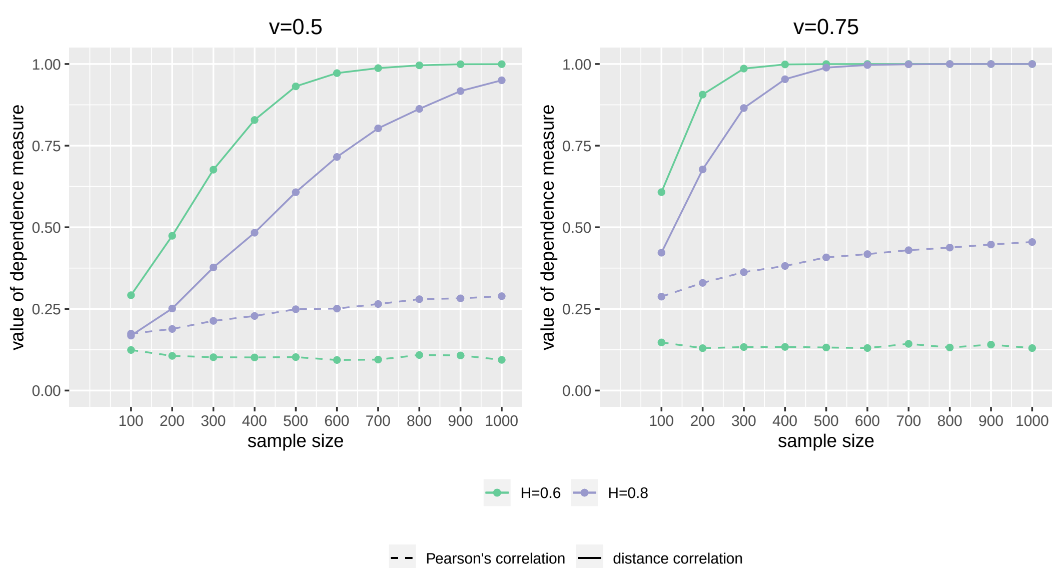

Fig 6: Rejection rates of the hypothesis tests resulting from Pearson's correlation and distance correlation obtained by subsampling based on "parabolically" correlated time series X_j, $j = 1, \ldots, n$, Y_j, $j = 1, \ldots, n$ according to (12) with block length $l_n = \lfloor \sqrt{n} \rfloor$, $d = 0.1n$, Hurst parameters $H = 0.6, 0.8$, and with $v = 0.5$ and $v = 0.75$.

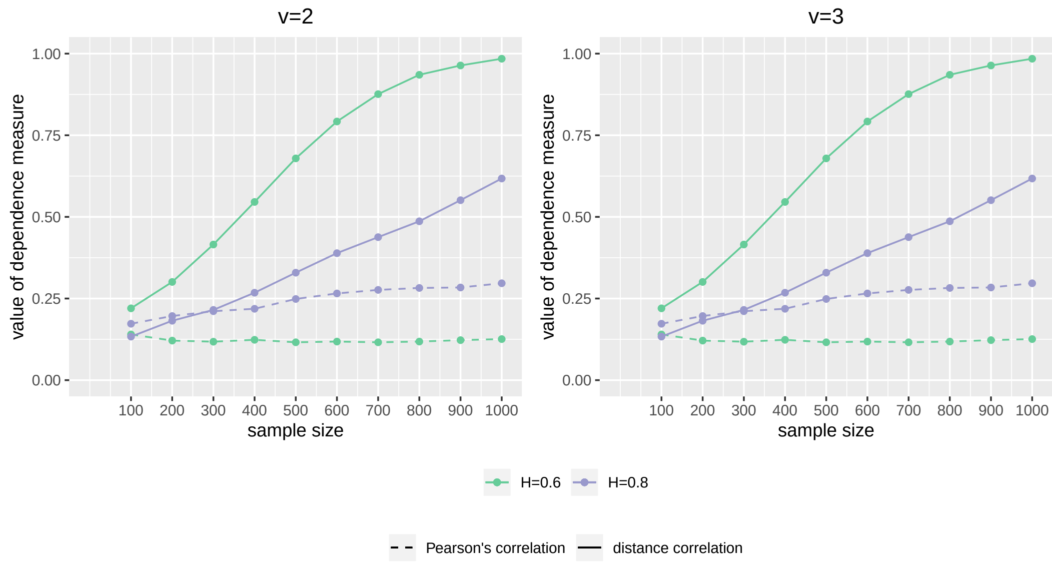

Fig 7: Rejection rates of the hypothesis tests resulting from Pearson's correlation and distance correlation obtained by subsampling based on "wavily" correlated time series X_j, $j = 1, \ldots, n$, Y_j, $j = 1, \ldots, n$ according to (13) with block length $l_n = \lfloor \sqrt{n} \rfloor$, $d = 0.1n$, Hurst parameters $H = 0.6, 0.8$, and with $v = 2$ and $v = 3$. The level of significance equals 5%.

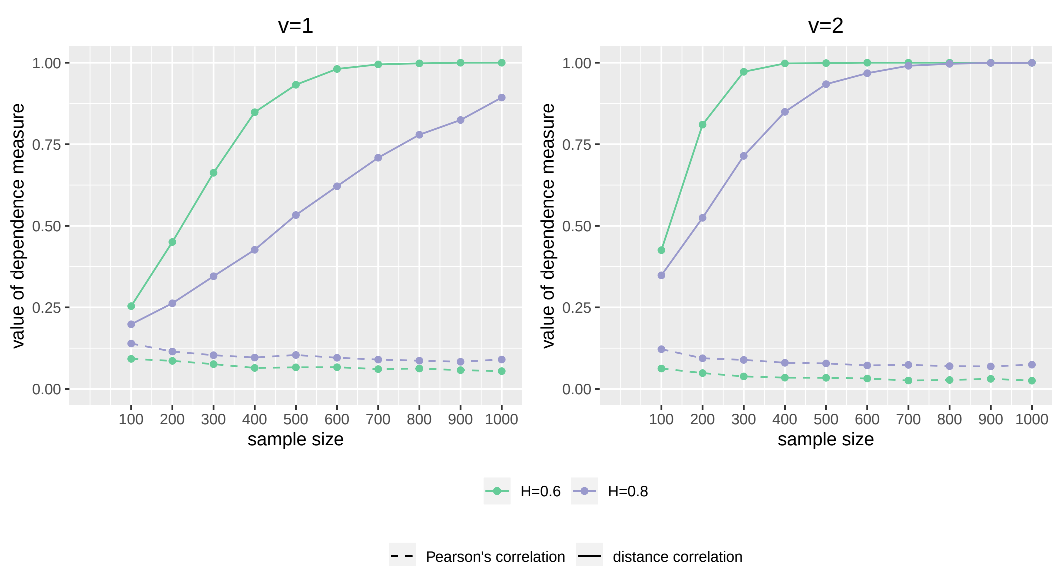

Fig 8: Rejection rates of the hypothesis tests resulting from Pearson's correlation and distance correlation obtained by subsampling based on "rectangularly" correlated time series X_j, $j = 1, \ldots, n$, Y_j, $j = 1, \ldots, n$ according to (14) with block length $l_n = \lfloor \sqrt{n} \rfloor$, $d = 0.1n$, Hurst parameters $H = 0.6, 0.8$, and with $v = 1$ and $v = 2$. The level of significance equals 5%.

detected by a test based on Pearson's correlation than by a test based on distance correlation. For linearly correlated data, an increase of dependence within the random vectors, i.e., an increase of the parameter H, results in a decrease of the empirical power of both testing procedures. For "parabolically" and "wavily" correlated data, this observation can only be made with respect to the finite sample performance of the test based on distance correlation. Most notably, in these two cases, the test based on distance correlation clearly outperforms the test based on Pearson's correlation in that it yields decisively higher empirical power. In addition to this, it seems remarkable that the test based on distance covariance interprets a rotation of data points generated by independent random vectors as dependence between the coordinates, while the test that is based on Pearson's correlation tends to classify these as being generated by independent random vectors.

5. Data example. In the following, the mean monthly discharges of three different rivers are analyzed with regard to cross-dependence between the corresponding data-generating processes by an application of the test statistics considered in the previous sections.
The data was provided by the Global Runoff Data Centre (GRDC) in Koblenz, Germany; see Global Runoff Data Centre (GRDC). The GRDC is an international archive currently comprising river discharge data of more than 9,900 stations from 159 countries.
The time series we are considering consist of $n = 96$ measurements of the mean monthly discharge from January 2000 to December 2007, i.e., a time period of 8 years, for the Amazon River, monitored at a station in São Paulo de Olivença, Brazil (corresponding to GRDC-No. 3623100), the Rhine, monitored at a station in Cologne, Germany (corresponding to GRDC-No. 6335060), and the Jutaí River, a tributary of the Amazon River, monitored at a station in Colocação Caxias (corresponding to GRDC-No. 3624201). (We chose the Jutaí River because its discharge volume compares to that of the Rhine.)
As the discharge volume of rivers is affected by seasonalities and trends, we eliminated these effects from the original data sets by the Small Trend Method, see Brockwell and Davis (1991), Chapter 1.4, p. 21, before our analysis. Figures 9, 10, and 11 depict the values of the detrended and deseasonalized time series.

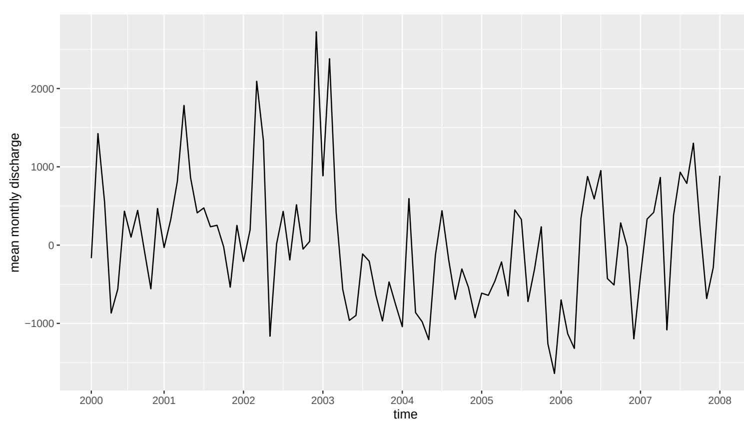

Fig 9: Detrended and deseasonalized mean monthly discharges of the river Rhine.

The mean monthly discharges of rivers typically display long-range dependence characterized by a Hurst parameter H that is close to 0.7, meaning the long-range dependence parameter D of the data-generating time series may be assumed to be close to 0.6. Under

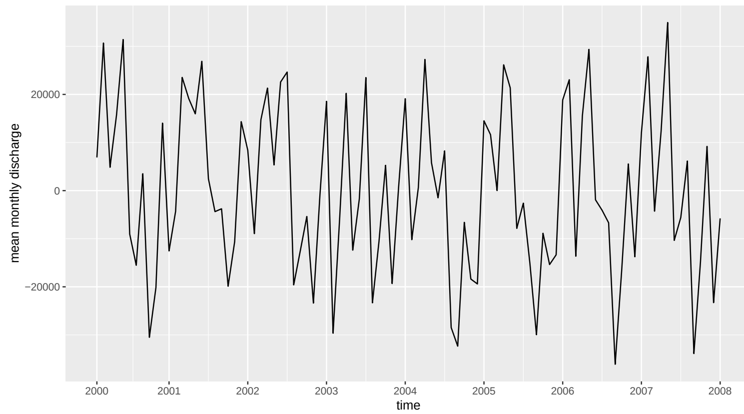

Fig 10: Detrended and deseasonalized mean monthly discharges of the river River Amazon.

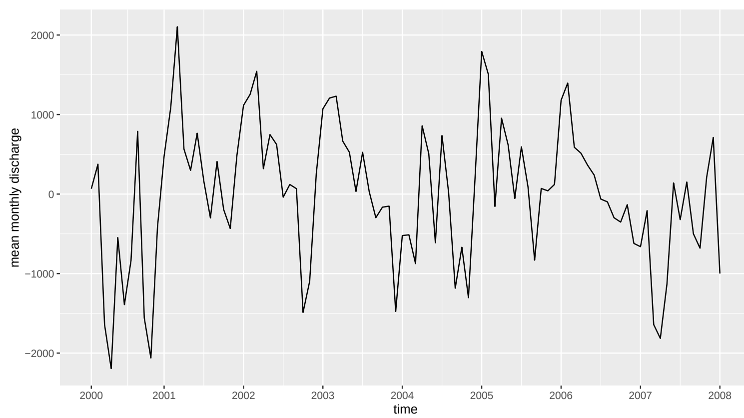

Fig 11: Detrended and deseasonalized mean monthly discharges of the Jutaí River.

a corresponding assumption, a test decision based on the distance correlation rejects the hypothesis for large values of

$$\sqrt{n}\int_{\mathbb{R}^p}\int_{\mathbb{R}^q}\left|\varphi^{(n)}_{X,Y}(s,t)-\varphi^{(n)}_X(s)\varphi^{(n)}_Y(t)\right|^2 s^{-2}t^{-2}dsdt,$$

while a test decision based on the sample covariance rejects the hypothesis for large values of

$$\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})\right|.$$

In our analysis, we apply both tests to the data. We base test decisions on an approximation of the distribution of the test statistics by the sampling-window method with block size $l = \lfloor\sqrt{n}\rfloor = 9$. As significance level we choose $\alpha = 5\%$.

As the Rhine is geographically separated from the other two rivers, we expect the tests to decide in favor of the hypothesis of independence, when applied to the discharges of the Rhine and one of the Brazilian rivers. Due to the fact that the Jutaí River is a tributary of the Amazon River, and due to the spatial proximity of the two measuring stations in Brazil, which are approximately 200 kilometers apart, we expect a test for independence of the discharge volumes of these two rivers to reject the hypothesis. In fact, both tests do not reject the hypothesis of two independent time series when applied to the Rhine's discharge and the Amazon River's or Jutaí River's discharge, respectively, and reject the hypothesis when applied to the discharges of the two Brazilian rivers.

Acknowledgments. The authors would like to thank the Global Runoff Data Centre for providing the considered data. Both authors were supported by Collaborative Research Center SFB 823 *Statistical modelling of non-linear dynamic processes*.

REFERENCES

ARCONES, M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *The Annals of Probability* **22** 2242 – 2274.

BERAN, J., FENG, Y., GHOSH, S. and KULIK, R. (2013). *Long-Memory Processes*. Springer-Verlag Berlin Heidelberg.

BETKEN, A. and WENDLER, M. (2018). Subsampling for General Statistics under Long Range Dependence. *Statistica Sinica* **28** 1199 – 1224.

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.

BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probability surveys* **2** 107 – 144.

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*. Springer-Verlag New York, Inc.

CREMERS, H. and KADELKA, D. (1986). On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L^p_E. *Stochastic Processes and their Applications* **21** 305 – 317.

DAVIS, R. A., MATSUI, M., MIKOSCH, T. and WAN, P. (2018). Applications of distance correlation to time series. *Bernoulli* **24** 3087 – 3116.

DEHLING, H. and TAQQU, M. S. (1989). The Empirical Process of some Long-Range Dependent Sequences with an Application to U -Statistics. *The Annals of Statistics* **17** 1767 – 1783.

DEHLING, H., MATSUI, M., MIKOSCH, T., SAMORODNITSKY, G. and TAFAKORI, L. (2020). Distance covariance for discretized stochastic processes. *Bernoulli* **26** 2758 – 2789.

DOBRUSHIN, R. L. and MAJOR, P. (1979). Non-central limit theorems for non-linear functionals of Gaussian fields. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50** 27 – 52.

GAO, L., FAN, Y., LV, J. and SHAO, Q.-M. (2019). Asymptotic Distributions of High-Dimensional Distance Correlation Inference.

GLOBAL RUNOFF DATA CENTRE (GRDC), G. 56068 KOBLENZ

GOODMAN, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics* **34** 152 – 177.

JIN, Z. and MATTESON, D. S. (2018). Generalizing distance covariance to measure and test multivariate mutual dependence via complete and incomplete V-statistics. *Journal of Multivariate Analysis* **168** 304 – 322.

KONG, Y., LI, D., FAN, Y. and LV, J. (2017). Interaction pursuit in high-dimensional multi-response regression via distance correlation. *The Annals of Statistics* **45** 897 – 922.

KROLL, M. (2021). Asymptotic Behaviour of the Empirical Distance Covariance for Dependent Data. *Journal of Theoretical Probability* 1 – 21.

MAJOR, P. (2020). The theory of Wiener–Ito integrals in vector valued Gaussian stationary random fields. Part I. *Moscow Mathematical Journal* **20** 749 – 812.

MATSUI, M., MIKOSCH, T. and SAMORODNITSKY, G. (2017). Distance covariance for stochastic processes. *Probability and Mathematical Statistics* **37** 355-372.

MATTESON, D. S. and TSAY, R. S. (2017). Independent component analysis via distance covariance. *Journal of the American Statistical Association* **112** 623 – 637.

PIPIRAS, V. and TAQQU, M. S. (2017). *Long-Range Dependence and Self-Similarity* **45**. Cambridge University Press.

SURGAILIS, D., KOUL, H. L. and GIRAITIS, L. (2012). *Large sample inference for long memory processes*. World Scientific Publishing Company.

SZÉKELY, 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.

SZÉKELY, G. J. and RIZZO, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics* **3** 1236 – 1265.

SZÉKELY, G. J. and RIZZO, M. L. (2012). On the uniqueness of distance covariance. *Statistics & Probability Letters* **82** 2278 – 2282.

SZÉKELY, G. J. and RIZZO, M. L. (2013). The distance correlation t-test of independence in high dimension. *Journal of Multivariate Analysis* **117** 193 – 213.

SZÉKELY, G. J. and RIZZO, M. L. (2014). Partial distance correlation with methods for dissimilarities. *The Annals of Statistics* **42** 2382 – 2412.

TAQQU, M. S. (1979). Convergence of Integrated Processes of Arbitrary Hermite Rank. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50** 53 – 83.

YAO, S., ZHANG, X. and SHAO, X. (2016). Testing mutual independence in high dimension via distance covariance. *Journal of the Royal Statistical Society: Series B* **80** 455 – 480.

ZHOU, Z. (2012). Measuring nonlinear dependence in time-series, a distance correlation approach. *Journal of Time Series Analysis* **33** 438 – 457.

APPENDIX: PROOFS

PROOF OF LEMMA 3.1. Define $u_j(s) := \exp(isX_j) - \varphi_X(s)$. Then, it follows that

$$\varphi_X^{(n)}(s) - \varphi_X(s) = \frac{1}{n}\sum_{j=1}^{n} u_j(s).$$

In order to show that $\varphi_X^{(n)}(s) - \varphi_X(s)$ is an element of the vector space $L^2(\mathbb{R}, \omega(s)ds)$, it thus suffices to show that $u_j \in L^2(\mathbb{R}, \omega(s)ds)$. For this, note that

$$\begin{aligned}|u_j(s)|^2 &= (\exp(isX_j) - \mathrm{E}(\exp(isX_j)))(\exp(-isX_j) - \mathrm{E}(-\exp(isX_j))) \\ &= 1 + \varphi_X(s)\bar{\varphi}_X(s) - \exp(isX_j)\bar{\varphi}_X(s) - \exp(-isX_j)\varphi_X(s).\end{aligned}$$

For $X' \stackrel{\mathcal{D}}{=} X$, independent of X, and E_X denoting the expected value taken with respect to X, we have

$$\begin{aligned}\varphi_X(s)\bar{\varphi}_X(s) &= \mathrm{E}\left[(\cos(sX) + i\sin(sX))\left(\cos(sX') - i\sin(sX')\right)\right] \\ &= \mathrm{E}\left[\cos(sX)\cos(sX') + \sin(sX)\sin(sX')\right] \\ &= \mathrm{E}\left[\cos(s(X - X'))\right],\end{aligned}$$

$$\begin{aligned}&\exp(isX_j)\bar{\varphi}_X(s) \\ &= (\cos(sX_j) + i\sin(sX_j))\,\mathrm{E}(\cos(sX) - i\sin(sX)) \\ &= \mathrm{E}_X(\cos(sX_j)\cos(sX) + \sin(sX)\sin(sX_j)) + i\sin(sX_j)\mathrm{E}(\cos(sX)) - i\cos(sX_j)\mathrm{E}(\sin(sX)) \\ &= \mathrm{E}_X(\cos(s(X_j - X))) + i\sin(sX_j)\mathrm{E}(\cos(sX)) - i\cos(sX_j)\mathrm{E}(\sin(sX)),\end{aligned}$$

and

$$\begin{aligned}&\exp(-isX_j)\varphi_X(s) \\ &= (\cos(sX_j) - i\sin(sX_j))\,\mathrm{E}(\cos(sX) + i\sin(sX)) \\ &= \mathrm{E}_X(\cos(sX_j)\cos(sX) + \sin(sX)\sin(sX_j)) - i\sin(sX_j)\mathrm{E}(\cos(sX)) + i\cos(sX_j)\mathrm{E}(\sin(sX)) \\ &= \mathrm{E}_X(\cos(s(X_j - X))) - i\sin(sX_j)\mathrm{E}(\cos(sX)) + i\cos(sX_j)\mathrm{E}(\sin(sX)).\end{aligned}$$

It follows that

$$|u_j(s)|^2 = 2\mathrm{E}_X[1 - \cos(s(X_j - X))] - \mathrm{E}\left[1 - \cos(s(X - X'))\right].$$

As a result, and according to Lemma 1 in Székely, Rizzo and Bakirov (2007), we arrive at

$$\begin{aligned}&\int_{\mathbb{R}} \frac{|u_j(s)|^2}{|s|^2} ds \\ &= \int_{\mathbb{R}} \frac{1}{|s|^2}\left(2\mathrm{E}_X[1 - \cos(s(X_j - X))] - \mathrm{E}\left[1 - \cos(s(X - X'))\right]\right) ds \\ &= \left(2\mathrm{E}_X\left[\int_{\mathbb{R}} \frac{1}{|s|^2}(1 - \cos(s(X_j - X)))\,ds\right] - \mathrm{E}\left[\int_{\mathbb{R}} \frac{1}{|s|^2}\left(1 - \cos(s(X - X'))\right) ds\right]\right) \\ &= 2\left(\mathrm{E}_X(|X_j - X|) - \mathrm{E}|X - X'|\right) \leq 2\left(|X_j| + \mathrm{E}|X|\right).\end{aligned}$$

Consequently, $\varphi_X^{(n)}(s) - \varphi_X(s)$ takes values in $L^2(\mathbb{R}, \omega(s)ds)$.

Furthermore, it holds that

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)\right|^2\omega(s,t)dsdt$$
$$=\int_{\mathbb{R}}\left|\varphi_X^{(n)}(s)-\varphi_X(s)\right|^2(cs^2)^{-1}ds\int_{\mathbb{R}}\left|\varphi_Y^{(n)}(t)-\varphi_Y(t)\right|^2(ct^2)^{-1}dt<\infty,$$

i.e., $\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)$ takes values in $L^2\left(\mathbb{R}^2,\omega(s,t)dsdt\right)$. □

PROOF OF PROPOSITION 3.1. Note that

$$\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)-J_1(s)\frac{1}{n}\sum_{i=1}^{n}X_iJ_1(t)\frac{1}{n}\sum_{j=1}^{n}Y_j$$
$$=\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)-J_1(t)\frac{1}{n}\sum_{j=1}^{n}Y_j\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)-J_1(s)\frac{1}{n}\sum_{i=1}^{n}X_i\right)$$
$$+\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)J_1(s)\frac{1}{n}\sum_{i=1}^{n}X_i+\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)J_1(t)\frac{1}{n}\sum_{j=1}^{n}Y_j.$$

For the first summand on the right-hand side of the above equation Corollary 3.1 implies

$$\left\|\varphi_Y^{(n)}(t)-\varphi_Y(t)-J_1(t)\frac{1}{n}\sum_{j=1}^{n}Y_j\right\|_2\left\|\varphi_X^{(n)}(s)-\varphi_X(s)-J_1(s)\frac{1}{n}\sum_{i=1}^{n}X_i\right\|_2$$
$$=o_P\left(n^{-\frac{D_X}{2}}L_X^{\frac{1}{2}}(n)n^{-\frac{D_Y}{2}}L_Y^{\frac{1}{2}}(n)\right).$$

For the second summand it follows from Corollary 3.1 and Theorem 5.6 in Taqqu (1979) that

$$\left\|\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)J_1(s)\frac{1}{n}\sum_{j=1}^{n}X_j\right\|_2=\left|\frac{1}{n}\sum_{j=1}^{n}X_j\right|\|J_1(s)\|_2\left\|\varphi_Y^{(n)}(t)-\varphi_Y(t)\right\|_2$$
$$=\mathcal{O}_P\left(n^{-\frac{D_X+D_Y}{2}}L_X^{\frac{1}{2}}(n)L_Y^{\frac{1}{2}}(n)\right).$$

An analogous result holds for the third summand. Thus, we established (8). Given (8), independence of X_k, $k\geq 1$, and Y_k, $k\geq 1$, implies

$$n^{\frac{D_X+D_Y}{2}}L_X^{-\frac{1}{2}}(n)L_Y^{-\frac{1}{2}}(n)\left(\varphi_Y^{(n)}(t)-\varphi_Y(t)\right)\left(\varphi_X^{(n)}(s)-\varphi_X(s)\right)$$
$$\xrightarrow{\mathcal{D}}-st\exp\left(-\frac{s^2+t^2}{2}\right)Z_XZ_Y,$$

where $\xrightarrow{\mathcal{D}}$ denotes weak convergence in $L^2\left(\mathbb{R}^2,\omega(s,t)dsdt\right)$ and and Z_X, Z_Y are independent standard normally distributed random variables. □

PROOF OF PROPOSITION 3.2. ´ Define

$$f_{s,t}(X_j,Y_j):=\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(itY_j)-\varphi_Y(t)\right).$$

Note that

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{E}\left[(\mathrm{Re} f_{s,t}(X_1,Y_1))^2\right]\omega(s,t)dsdt \leq \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{E}\left[|f_{s,t}(X_1,Y_1)|^2\right]\omega(s,t)dsdt < \infty,$$
$$\int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{E}\left[(\mathrm{Im} f_{s,t}(X_1,Y_1))^2\right]\omega(s,t)dsdt \leq \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{E}\left[|f_{s,t}(X_1,Y_1)|^2\right]\omega(s,t)dsdt < \infty$$

according to the proof of Lemma 3.1. In particular, this implies that $\mathrm{Re} f_{s,t}(X_1,Y_1)$ and $\mathrm{Im} f_{s,t}(X_1,Y_1)$ allow for an expansion in bivariate Hermite polynomials.
First, we will focus our considerations on $\mathrm{Re} f_{s,t}$. An expansion of $\mathrm{Re} f_{s,t}$ in Hermite polynomials is given by

$$\mathrm{Re} f_{s,t}\left(X_j,Y_j\right) - \mathrm{E}\mathrm{Re} f_{s,t}\left(X_1,Y_1\right) \overset{L^2}{=} \sum_{q=r}^{\infty}\sum_{k,l:k+l=q}\frac{J_{k,l}(s,t)}{k!l!}H_k(X_1)H_l(Y_1),$$

i.e.,

$$\lim_{n\to\infty}\left\|\mathrm{Re} f_{s,t}\left(X_j,Y_j\right) - \mathrm{E}\mathrm{Re} f_{s,t}\left(X_j,Y_j\right) - \sum_{q=r}^{\infty}\sum_{k,l:k+l=q}\frac{J_{k,l}(s,t)}{k!l!}H_k(X_j)H_l(Y_j)\right\|_{L^2} = 0,$$

where $\|\cdot\|_{L^2}$ denotes the norm induced by the inner product (5), $J_{k,l}(s,t) := \mathrm{E}\left(\mathrm{Re} f_{s,t}\left(X_1,Y_1\right)H_k(X_1)H_l(Y_1)\right)$ and $r := \min\{p+q : J_{p,q}(s,t) \neq 0\}$.
In order to compute the Hermite rank r and the corresponding Hermite coefficients $J_{k,l}(s,t)$, note that

$$\begin{aligned}
&\mathrm{Re} f_{s,t}(X_1,Y_1)\\
=&\mathrm{Re}\left(\exp(isX_1)-\varphi_X(s)\right)\left(\exp(itY_1)-\varphi_Y(t)\right)\\
=&\cos(sX_1+tY_1)-\cos(tY_1)\mathrm{E}\left(\cos(sX_1)\right)+\sin(tY_1)\mathrm{E}\left(\sin(sX_1)\right)-\cos(sX_1)\mathrm{E}(\cos(tY_1))\\
&+\sin(sX_1)\mathrm{E}(\sin(tY_1))+\mathrm{E}\cos(tY_1)\mathrm{E}(\cos(sX_1))-\mathrm{E}(\sin(tY_1))\mathrm{E}(\sin(sX_1)).
\end{aligned}$$

Since $\cos(x+y)=\cos(x)\cos(y)-\sin(x)\sin(y)$, it follows that

$$\mathrm{E}\left(\mathrm{Re} f_{s,t}(X_1,Y_1)H_l(X_1)H_0(Y_1)\right) = \mathrm{E}\left(\mathrm{Re} f_{s,t}(X_1,Y_1)H_0(X_1)H_l(Y_1)\right) = 0.$$

In particular, we have $J_{0,1}(s,t) = J_{1,0}(s,t) = J_{0,2}(s,t) = J_{2,0}(s,t) = 0$. Since $\sin(x+y) = \sin(x)\cos(y)+\cos(x)\sin(y)$,

$$\begin{aligned}
\mathrm{E}\left(\mathrm{Re} f_{s,t}(X_1,Y_1)H_1(X_1)H_1(Y_1)\right) &= -\mathrm{E}\left(\sin(sX_1)X_1\right)\mathrm{E}\left(\sin(tY_1)Y_1\right)\\
&= -\exp\left(-\frac{s^2}{2}\right)s\exp\left(-\frac{t^2}{2}\right)t \neq 0
\end{aligned}$$

almost everywhere. Therefore, the Hermite rank r of $\mathrm{Re} f_{s,t}$ corresponds to 2 almost everywhere. More precisely, it follows that the only summmand in the Hermite expansion of $\mathrm{Re} f_{s,t}(X_1,Y_1)$ that is of order $1 \leq q \leq 2$ corresponds to the coefficient $J_{1,1}(s,t)$ and the polynomial $H_1(x)H_1(y) = xy$. We will see that this summand determines the asymptotic behavior of the partial sum process.
To this end, we show L_2-convergence of

$$n^{D-1}L^{-1}(n)\sum_{j=1}^{n}\left(\mathrm{Re} f_{s,t}\left(X_j,Y_j\right) - J_{1,1}(s,t)X_jY_j\right).$$

Due to orthogonality of the Hermite polynomials in $L^2(\mathbb{R}^2, \varphi_{I_2})$, we have

$$\begin{aligned}\mathrm{E}\left(H_{k_1}(X_i)H_{l_1}(Y_i)H_{k_2}(X_j)H_{l_2}(Y_j)\right) =& \mathrm{E}\left(H_{k_1}(X_i)H_{k_2}(X_j)\right)\mathrm{E}\left(H_{l_1}(Y_i)H_{l_2}(Y_j)\right)\\ =& \begin{cases} k_1!l_1!\rho(i-j)^{k_1+l_1} & \text{if } k_1=k_2 \text{ and } l_1=l_2 \\ 0 & \text{else}\end{cases}.\end{aligned}$$

It follows that

$$\begin{aligned}&\mathrm{E}\left(\left|n^{D-1}L^{-1}(n)\sum_{j=1}^{n}\left(\mathrm{Re}f_{s,t}\left(X_j,Y_j\right)-J_{1,1}(s,t)X_jY_j\right)\right|^2\right)\\ =&\mathrm{E}\left(\left|n^{D-1}L^{-1}(n)\sum_{j=1}^{n}\sum_{q=3}^{\infty}\sum_{k,l:k+l=q}\frac{J_{k,l}(s,t)}{k!l!}H_k(X_j)H_l(Y_j)\right|^2\right)\\ =&n^{2D-2}L^{-2}(n)\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{q=3}^{\infty}\sum_{k,l:k+l=q}\frac{1}{k!l!}J_{k,l}^2(s,t)\left|\rho(i-j)\right|^q\\ \leq& n^{2D-2}L^{-2}(n)\sum_{q=3}^{\infty}\sum_{k,l:k+l=q}^{\infty}\frac{1}{k!l!}J_{k,l}^2(s,t)\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\rho(i-j)\right|^3.\end{aligned}$$

Recall that for an auto-covariance function $\rho(k) = k^{-D}L(k)$, it holds that

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\rho(i-j)\right|^3 = \mathcal{O}\left(n^{1\vee(2-3D)}L'(n)\right),$$

where L' is some slowly varying function.
As a result, the previous considerations establish

$$\begin{aligned}&\mathrm{E}\left(\int_{\mathbb{R}}\int_{\mathbb{R}}\left|n^{D-1}L^{-1}(n)\sum_{j=1}^{n}\left(\mathrm{Re}f_{s,t}\left(X_j,Y_j\right)-J_{1,1}(s,t)X_jY_j\right)\right|^2\omega(s,t)dsdt\right)\\ =&\mathcal{O}\left(n^{2D-2}L^{-2}(n)n^{1\vee(2-3D)}L'(n)\int_{\mathbb{R}}\int_{\mathbb{R}}\sum_{q=3}^{\infty}\sum_{k,l:k+l=q}^{\infty}\frac{1}{k!l!}J_{k,l}^2(s,t)\omega(s,t)dsdt\right).\end{aligned}$$

Since

$$\sum_{q=3}^{\infty}\sum_{k,l:k+l=q}^{\infty}\frac{1}{k!l!}J_{k,l}^2(s,t) \leq \sum_{q=0}^{\infty}\sum_{k,l:k+l=q}^{\infty}\frac{1}{k!l!}J_{k,l}^2(s,t) = \mathrm{E}\left[\left(\mathrm{Re}f_{s,t}(X_1,Y_1)\right)^2\right],$$

we conclude that the right-hand side of the above inequality corresponds to

$$\mathcal{O}\left(n^{-\min(1-2D,D)}\tilde{L}(n)\int_{\mathbb{R}}\int_{\mathbb{R}}\mathrm{E}\left(\mathrm{Re}f_{s,t}(X_1,Y_1)\right)^2\omega(s,t)dsdt\right) \tag{15}$$

for some slowly varying function $\tilde{L}$.
With respect to the corresponding Hermite coefficients, we note that

$$\begin{aligned}&\mathrm{Im}f_{s,t}(X_1,Y_1)\\ =&\mathrm{Im}\left(\exp(isX_1)-\varphi_X(s)\right)\left(\exp(itY_1)-\varphi_Y(t)\right)\\ =&\sin(sX_1+tY_1)-\cos(tY_1)\mathrm{E}(\sin(sX_1))-\mathrm{E}(\cos(sX_1))\sin(tY_1)-\cos(sX_1)\mathrm{E}(\sin(tY_1))\\ &-\mathrm{E}(\cos(tY_1))\sin(sX_1)+\mathrm{E}(\cos(sX_1))\mathrm{E}(\sin(tY_1))+\mathrm{E}(\cos(tY_1))\mathrm{E}(\sin(sX_1)).\end{aligned}$$

As a result, we have

$$\mathrm{E}\left(\mathrm{Im} f_{s,t}(X_1,Y_1)H_l(X_1)H_0(Y_1)\right)=\mathrm{E}\left(\mathrm{Im} f_{s,t}(X_1,Y_1)H_0(X_1)H_l(Y_1)\right)=0$$

for $l\in\mathbb{N}$ and

$$\mathrm{E}\left(\mathrm{Im} f_{s,t}(X_1,Y_1)H_1(X_1)H_1(Y_1)\right)=\mathrm{E}\left(\sin(sX_1+tY_1)H_1(X_1)H_1(Y_1)\right)=0.$$

Therefore, the Hermite rank of the imaginary part of $f_{s,t}$ is bigger than 2, such that an argument analogous to the considerations for the real part of $f_{s,t}$ implies that $\mathrm{Im} f_{s,t}$ is asymptotically negligible.
As a result, (15) is $o(n^{-\delta})$ for some $\delta>0$ if $D\in\left(0,\frac{1}{2}\right)$. □

PROOF OF PROPOSITION 3.3. It holds that

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}X_iY_j &=\sum_{j=1}^{n}\int_{[-\pi,\pi)}\left(\mathrm{e}^{ix}\right)^j Z_{G,X}(dx)\sum_{k=1}^{n}\int_{[-\pi,\pi)}\left(\mathrm{e}^{iy}\right)^k Z_{G,Y}(dy)\\ &=\int_{[-\pi,\pi)^2}\mathrm{e}^{i(x+y)}\left(\sum_{j=0}^{n-1}\left(\mathrm{e}^{ix}\right)^j\right)\left(\sum_{j=0}^{n-1}\left(\mathrm{e}^{iy}\right)^j\right)Z_{G,X}(dx)Z_{G,Y}(dy)\\ &=\int_{[-\pi,\pi)^2}\mathrm{e}^{i(x+y)}\left(\frac{\mathrm{e}^{ixn}-1}{\mathrm{e}^{ix}-1}\right)\left(\frac{\mathrm{e}^{iyn}-1}{\mathrm{e}^{iy}-1}\right)Z_{G,X}(dx)Z_{G,Y}(dy).\end{aligned}$$

By the change of variables formula, it follows that

$$\begin{aligned}&n^{\frac{D_X+D_Y}{2}-2}L_X^{-\frac{1}{2}}L_Y^{-\frac{1}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}X_iY_j\\ =&n^{D-2}L^{-1}(n)\int_{[-\pi,\pi)^2}\mathrm{e}^{i(x+y)}\left(\frac{\mathrm{e}^{ixn}-1}{\mathrm{e}^{ix}-1}\right)\left(\frac{\mathrm{e}^{iyn}-1}{\mathrm{e}^{iy}-1}\right)Z_{G,X}(dx)Z_{G,Y}(dy)\\ =&\int_{[-\pi,\pi)^2}\mathrm{e}^{i\frac{x+y}{n}}\frac{1}{n^2}\left(\frac{\mathrm{e}^{ix}-1}{\mathrm{e}^{i\frac{x}{n}}-1}\right)\left(\frac{\mathrm{e}^{iy}-1}{\mathrm{e}^{i\frac{y}{n}}-1}\right)Z_{G,X}^{(n)}(dx)Z_{G,Y}^{(n)}(dy),\end{aligned}$$

where

$$Z_{G,X}^{(n)}(A)=\sqrt{n^{D_X}L_X^{-1}(n)}Z_{G,X}\left(\frac{A}{n}\right)=n^{\frac{D_X}{2}}L_X^{-\frac{1}{2}}(n)Z_{G,X}\left(\frac{A}{n}\right)$$

and

$$Z_{G,Y}^{(n)}(A)=\sqrt{n^{D_Y}L_Y^{-1}(n)}Z_{G,Y}\left(\frac{A}{n}\right)=n^{\frac{D_Y}{2}}L_Y^{-\frac{1}{2}}(n)Z_{G,Y}\left(\frac{A}{n}\right).$$

Moreover, it holds that

$$\begin{aligned}\sum_{j=1}^{n}X_jY_j&=\sum_{j=1}^{n}\int_{[-\pi,\pi)}\mathrm{e}^{ixj}Z_{G,X}(dx)\int_{[-\pi,\pi)}\mathrm{e}^{iyj}Z_{G,Y}(dy)\\ &=\int_{[-\pi,\pi)^2}\sum_{j=1}^{n}\mathrm{e}^{i(x+y)j}Z_{G,X}(dx)Z_{G,Y}(dy)\end{aligned}$$

$$= \int_{[-\pi,\pi)^2} \mathrm{e}^{i(x+y)} \sum_{j=0}^{n-1} \mathrm{e}^{i(x+y)j} Z_{G,X}(dx) Z_{G,Y}(dy)$$

$$= \int_{[-\pi,\pi)^2} \mathrm{e}^{i(x+y)} \frac{\mathrm{e}^{i(x+y)n}-1}{\mathrm{e}^{i(x+y)}-1} Z_{G,X}(dx) Z_{G,Y}(dy).$$

Again, the change of variables formula yields

$$n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}} L_Y^{-\frac{1}{2}} \sum_{j=1}^{n} X_j Y_j = n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}} L_Y^{-\frac{1}{2}} \int\limits_{[-\pi,\pi)^2} \mathrm{e}^{i(x+y)} \left(\frac{\mathrm{e}^{i(x+y)n}-1}{\mathrm{e}^{i(x+y)}-1} \right) Z_{G,X}(dx) Z_{G,Y}(dy)$$

$$= \int\limits_{[-\pi,\pi)^2} \mathrm{e}^{i\frac{x+y}{n}} \left(\frac{\mathrm{e}^{i(x+y)}-1}{n\left(\mathrm{e}^{i\frac{x+y}{n}}-1\right)} \right) Z_{G,X}^{(n)}(dx) Z_{G,Y}^{(n)}(dy),$$

where

$$Z_{G,X}^{(n)}(A) = \sqrt{n^{D_X} L_X^{-1}(n)}\, Z_{G,X}\left(\frac{A}{n}\right) = n^{\frac{D_X}{2}} L_X^{-\frac{1}{2}}(n) Z_{G,X}\left(\frac{A}{n}\right)$$

and

$$Z_{G,Y}^{(n)}(A) = \sqrt{n^{D_Y} L_Y^{-1}(n)}\, Z_{G,Y}\left(\frac{A}{n}\right) = n^{\frac{D_Y}{2}} L_Y^{-\frac{1}{2}}(n) Z_{G,Y}\left(\frac{A}{n}\right).$$

With the results of Major (2020) it then follows that

$$n^{\frac{D_X+D_Y}{2}-2} L_X^{-\frac{1}{2}} L_Y^{-\frac{1}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} X_i Y_j - n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}} L_Y^{-\frac{1}{2}} \sum_{j=1}^{n} X_j Y_j$$

converges in distribution to

$$\int_{[-\pi,\pi)^2} \left[\left(\frac{\mathrm{e}^{ix}-1}{ix} \right) \left(\frac{\mathrm{e}^{iy}-1}{iy} \right) - \frac{\mathrm{e}^{i(x+y)}-1}{i(x+y)} \right] Z_{G,X}(dx) Z_{G,Y}(dy).$$

□

In order to prove Proposition 3.4, we apply the following theorem that corresponds to Theorem 2 in Cremers and Kadelka (1986) for stochastic processes with paths in $L^p(S,\mu)$, where $(S,\mathcal{S},\mu)$ is a σ-finite measure space, when choosing $p=2$:

THEOREM A.1. *Let $(S,\mathcal{S},\mu)$ be a σ-finite measure space and let ξ_k, $k \geq 1$, be a sequence of stochastic processes with paths in $L^2(S,\mu)$. Then*

$$\xi_n \xrightarrow{\mathcal{D}} \xi_0, \quad \text{where } \xrightarrow{\mathcal{D}} \text{ denotes convergence in } L^2(S,\mu),$$

provided the finite dimensional distributions of ξ_n converge weakly to those of ξ_0 almost everywhere and provided the following conditions hold: for some positive, μ-integrable function f, it holds that

$$\mathrm{E}\,|\xi_n(s)|^2 \leq f(s) \text{ for all } s \in S,\ n \in \mathbb{N}, \text{ and } \mathrm{E}\,|\xi_n(s)|^2 \longrightarrow \mathrm{E}\,|\xi_0(s)|^2 \text{ for all } s \in S.$$

PROOF OF PROPOSITION 3.4. Our goal is to show convergence of

$$Z_n(s,t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} f_{s,t}(X_j, Y_j).$$

In order to show convergence of the finite dimensional distributions we have to prove that for fixed k and $s_1, t_1, s_2, t_2, \ldots, s_k, t_k \in \mathbb{R}$

$$Z_n := (Z_{n,1}, Z_{n,2}, \ldots, Z_{n,k})^\top, \quad \text{where } Z_{n,i} := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} f_{s_i,t_i}(X_j, Y_j),$$

converges in distribution to the corresponding finite dimensional distribution of a complex Gaussian random variable $Z := (Z_1, Z_2, \ldots, Z_k)^\top$.
Due to the Cramér-Wold theorem, for this we have to show that for all $\lambda_1, \ldots, \lambda_k, \eta_1, \ldots \eta_k \in \mathbb{R}$

$$\sum_{i=1}^{k} \lambda_i \mathrm{Re}(Z_{n,i}) + \sum_{i=1}^{k} \eta_i \mathrm{Im}(Z_{n,i}) \xrightarrow{\mathcal{D}} \sum_{i=1}^{k} \lambda_i \, \mathrm{Re}(Z_i) + \sum_{i=1}^{k} \eta_i \mathrm{Im}(Z_i).$$

To this end, note that

$$\sum_{i=1}^{k} \lambda_i \mathrm{Re}(Z_{n,i}) + \sum_{i=1}^{k} \eta_i \mathrm{Im}(Z_{n,i}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} K(X_j, Y_j),$$

where

$$K(X_j, Y_j) := \sum_{i=1}^{k} \lambda_i \mathrm{Re}(f_{s_i,t_i}(X_j, Y_j)) + \sum_{i=1}^{k} \eta_i \mathrm{Im}(f_{s_i,t_i}(X_j, Y_j)).$$

In order to derive the asymptotic distribution of the above partial sum, we apply the following theorem that directly follows from Theorem 4 in Arcones (1994):

THEOREM A.2. *Let $\mathbf{X}_k = (X_k^{(1)}, X_k^{(2)})$, $k \geq 1$, be a stationary mean zero Gaussian sequence of $\mathbb{R}^2$-valued random vectors. Let f be a function on $\mathbb{R}^2$ with Hermite rank r. We define*

$$\rho^{(p,q)}(k) := \mathrm{E}\left(X_1^{(p)} X_{1+k}^{(q)}\right)$$

for $k \geq 1$. Suppose that

$$\sum_{k=-\infty}^{\infty} \left|\rho^{(p,q)}(k)\right|^r < \infty$$

for each $1 \leq p, q \leq 2$. Then, it holds that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} (f(\mathbf{X}_j) - \mathrm{E}f(\mathbf{X}_j)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 := \mathrm{E}\left[(f(\mathbf{X}_1) - \mathrm{E}f(\mathbf{X}_1))^2\right] + 2\sum_{k=1}^{\infty} \mathrm{E}\left[(f(\mathbf{X}_1) - \mathrm{E}f(\mathbf{X}_1))(f(\mathbf{X}_{1+k}) - \mathrm{E}f(\mathbf{X}_{1+k}))\right].$$

For an application of Theorem A.2, we have to compute the Hermite rank of K. Recall that the Hermite rank is defined by

$$r := \inf \{\tau \mid \exists\, l_1, l_2 \text{ with } l_1 + l_2 = \tau \text{ and } \mathrm{E}\left(K(X_1, Y_1) H_{l_1}(\xi_1) H_{l_2}(\eta_1)\right) \neq 0\}.$$

We have

$$\begin{aligned}
&\mathrm{E}\left(K(X_1, Y_1) H_{l_1}(\xi_1) H_{l_2}(\eta_1)\right)\\
&= \sum_{m=1}^{k} \lambda_m \mathrm{E}\left(\mathrm{Re} f_{s_m, t_m}(X_1, Y_1) H_{l_1}(\xi_1) H_{l_2}(\eta_1)\right)\\
&\quad + \sum_{m=1}^{k} \eta_m \mathrm{E}\left(\mathrm{Im} f_{s_m, t_m}(X_1, Y_1) H_{l_1}(\xi_1) H_{l_2}(\eta_1)\right).
\end{aligned}$$

Moreover, as shown in the proof of Proposition 3.2, it holds that

$$\begin{aligned}
&\mathrm{Re} f_{s,t}\,(X_1, Y_1)\\
&= \cos(sX_1 + tY_1) - \cos(tY_1)\mathrm{E}\left(\cos(sX_1)\right) + \sin(tY_1)\mathrm{E}\left(\sin(sX_1)\right) - \cos(sX_1)\mathrm{E}(\cos(tY_1))\\
&\quad + \sin(sX_1)\mathrm{E}(\sin(tY_1)) + \mathrm{E}\cos(tY_1)\mathrm{E}(\cos(sX_1)) - \mathrm{E}(\sin(tY_1))\mathrm{E}(\sin(sX_1))
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{Im} f_{s,t}\,(X_1, Y_1)\\
&= \sin(sX_1 + tY_1) - \cos(tY_1)\mathrm{E}(\sin(sX_1)) - \mathrm{E}(\cos(sX_1))\sin(tY_1) - \cos(sX_1)\mathrm{E}(\sin(tY_1))\\
&\quad - \mathrm{E}(\cos(tY_1))\sin(sX_1) + \mathrm{E}(\cos(sX_1))\mathrm{E}(\sin(tY_1)) + \mathrm{E}(\cos(tY_1))\mathrm{E}(\sin(sX_1)).
\end{aligned}$$

Since $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$, it follows that

$$\mathrm{E}\left(\mathrm{Re} f_{s_i,t_i}\,(X_1, Y_1)\, H_l(\xi_1) H_0(\eta_1)\right) = \mathrm{E}\left(\mathrm{Re} f_{s_i,t_i}\,(X_1, Y_1)\, H_0(\xi_1) H_l(\eta_1)\right) = 0,$$

while

$$\mathrm{E}\left(\mathrm{Im} f_{s_i,t_i}\,(X_1, Y_1)\, H_l(\xi_1) H_0(\eta_1)\right) = \mathrm{E}\left(\mathrm{Im} f_{s_i,t_i}\,(X_1, Y_1)\, H_0(\xi_1) H_l(\eta_1)\right) = 0$$

for all $l \in \mathbb{N}$. Therefore, the Hermite rank r of K is bigger than 1, such that for $D_\xi, D_\eta \in \left(\frac{1}{2}, 1\right)$

$$\sum_{k=-\infty}^{\infty} |\rho_\xi(k)|^r \leq \sum_{k=-\infty}^{\infty} |\rho(k)|^2 < \infty, \quad \sum_{k=-\infty}^{\infty} |\rho_\eta(k)|^r \leq \sum_{k=-\infty}^{\infty} |\rho(k)|^2 < \infty.$$

As a result, Theorem A.2 implies that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} K(X_j, Y_j) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 := \mathrm{E}\left[K(X_1, Y_1)^2\right] + 2 \sum_{k=1}^{\infty} \mathrm{E}\left[K(X_1, Y_1) K(X_{k+1}, Y_{k+1})\right].$$

According to Theorem A.1, for a proof of Proposition 3.4 it thus remains to show that for some positive, $\omega(s,t)dsdt$-integrable function f

$$\mathrm{E}\,|Z_n(s,t)|^2 \leq f(s,t) \text{ for all } (s,t) \in \mathbb{R}^2,\, n \in \mathbb{N},$$

and

$$\lim_{n\to\infty} \mathrm{E}\,|Z_n(s,t)|^2 = \mathrm{E}\,|Z(s,t)|^2 \text{ for all } (s,t)\in\mathbb{R}^2.$$

For this, note that due to independence of X_j, $j\in\mathbb{N}$, and Y_j, $j\in\mathbb{N}$, it holds that

$$\mathrm{E}\left(|Z_n(s,t)|^2\right) = \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{E}\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(-isX_k)-\varphi_X(-s)\right)$$
$$\mathrm{E}\left(\exp(itY_j)-\varphi_Y(t)\right)\left(\exp(-itY_k)-\varphi_Y(-t)\right).$$

An expansion in Hermite polynomials yields

$$\exp(isX_j)-\varphi_X(s) = \sum_{l=1}^{\infty}\frac{1}{l!}\mathrm{E}\left(\cos(sG(\xi_j))H_l(\xi_j)\right)H_l(\xi_j) + i\sum_{l=1}^{\infty}\frac{1}{l!}\mathrm{E}\left(\sin(sG(\xi_j))H_l(\xi_j)\right)H_l(\xi_j)$$

and

$$\exp(-isX_k)-\varphi_X(-s) = \sum_{l=1}^{\infty}\frac{1}{l!}\mathrm{E}\left(\cos(sG(\xi_j))H_l(\xi_j)\right)H_l(\xi_j) - i\sum_{l=1}^{\infty}\frac{1}{l!}\mathrm{E}\left(\sin(sG(\xi_j))H_l(\xi_j)\right)H_l(\xi_j).$$

Since

$$\mathrm{Cov}(H_l(\xi_j)H_m(\xi_k)) = \begin{cases}\rho_\xi^l(j-k)l! & \text{if } l=m\\ 0 & \text{if } l\neq m\end{cases},$$

we thus have

$$\begin{aligned}&\mathrm{E}\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(-isX_k)-\varphi_X(-s)\right)\\ =&\sum_{l=1}^{\infty}\frac{\left(\mathrm{E}\left(\cos(sG(\xi_1))H_l(\xi_1)\right)\right)^2+\left(\mathrm{E}\left(\sin(sG(\xi_1))H_l(\xi_1)\right)\right)^2}{l!}\rho_\xi^l(j-k).\end{aligned}$$

With $J_l(s) := \left(\mathrm{E}\left(\cos(sG(\xi_j))H_l(\xi_j)\right)\right)^2+\left(\mathrm{E}\left(\sin(sG(\xi_j))H_l(\xi_j)\right)\right)^2$, it holds that

$$\mathrm{E}\left(|Z_n(s,t)|^2\right) = \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\frac{J_l(s)J_m(t)}{l!m!}\rho_\xi^l(j-k)\rho_\eta^m(j-k).$$

For $l+m\geq 2$ and $D\in\left(\frac{1}{2},1\right)$, we have

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\rho_\xi^l(j-k)\rho_\eta^m(j-k) \leq \sum_{j=1}^{n}\sum_{k=1}^{n}\rho_\xi(j-k)\rho_\eta(j-k) = \mathcal{O}\left(n\right).$$

It follows that

$$\mathrm{E}\left(|Z_n(s,t)|^2\right) = \mathcal{O}\left(\sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\frac{J_l(s)J_m(t)}{l!m!}\right).$$

Since

$$\begin{aligned}\sum_{l=1}^{\infty}\frac{J_l(s)}{l!} =& \mathrm{E}\left(\exp(isX_1)-\varphi_X(s)\right)\left(\exp(-isX_1)-\varphi_X(-s)\right)\\ =&1-\varphi_X(s)\varphi_X(-s)\\ =&1-\left[\mathrm{E}\left(\cos(sG(\xi_1))\right)\right]^2-\left[\mathrm{E}\left(\sin(sG(\xi_1))\right)\right]^2,\end{aligned}$$

it follows by Lemma 1 in Székely, Rizzo and Bakirov (2007) that

$$\begin{aligned}\int_{\mathbb{R}} \sum_{l=1}^{\infty} \frac{J_l(s)}{l!}(cs^2)^{-1} ds &= \int_{\mathbb{R}} \left(1 - (\mathrm{E}\,(\cos(sG(\xi_j))))^2 - (\mathrm{E}\,(\sin(sG(\xi_j))))^2\right)(cs^2)^{-1} ds \\ &= \int_{\mathbb{R}} \mathrm{E}\,(1 - \cos(s(X - X')))\,(cs^2)^{-1} ds \\ &= \mathrm{E}\left(\int_{\mathbb{R}} (1 - \cos(s(X - X')))\,(cs^2)^{-1} ds\right) \\ &\leq C\mathrm{E}\,|X - X'| < \infty.\end{aligned}$$

As a result, we have

$$\mathrm{E}\,|Z_n(s,t)|^2 \leq f(s,t) \text{ for all } (s,t) \in \mathbb{R}^2,\ n \in \mathbb{N},$$

where, for some positive constant C, $f(s,t) := C \sum_{l=1}^{\infty} \frac{J_l(s)}{l!} \sum_{m=1}^{\infty} \frac{J_m(t)}{m!}$ is a positive, $\omega(s,t)dsdt$-integrable function .
Moreover, convergence of $\mathrm{E}\,|Z_n(s,t)|^2$ and $\mathrm{E}\,(Z_n(s,t))^2$ follows by the dominated convergence theorem. As limits we obtain

$$\begin{aligned}&\mathrm{E}\left(|Z_n(s,t)|^2\right) \\ &= \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \mathrm{E}\left(f_{s,t}(X_1,Y_1)\overline{f_{s,t}(X_{k+1},Y_{k+1})}\right) \\ &\longrightarrow \sum_{k=-\infty}^{\infty} \mathrm{E}\left(f_{s,t}(X_1,Y_1)\overline{f_{s,t}(X_{k+1},Y_{k+1})}\right),\end{aligned}$$

and

$$\begin{aligned}&\mathrm{E}\left((Z_n(s,t))^2\right) \\ &= \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \mathrm{E}\,(f_{s,t}(X_1,Y_1)f_{s,t}(X_{k+1},Y_{k+1})) \\ &\longrightarrow \sum_{k=-\infty}^{\infty} \mathrm{E}\,(f_{s,t}(X_1,Y_1)f_{s,t}(X_{k+1},Y_{k+1}))\,,.\end{aligned}$$

□

PROOF OF COROLLARY 3.2. Proposition 3.4 implies convergence to a complex Gaussian random variable. It therefore only remains to compute the limit parameters $\Gamma_{s,t}$ and $C_{s,t}$ specified in formula (10). For this, note that for a Gaussian random variable X, $\varphi_X(s) = \varphi_X(-s)$, and therefore

$$\begin{aligned}\mathrm{E}\,(\exp(isX_j) - \varphi_X(s))\,(\exp(-isX_k) - \varphi_X(-s)) &= \mathrm{E}\exp(is(X_j - X_k)) - \varphi_X^2(s) \\ &= \exp\left(-\frac{s^2\sigma_{j,k}^2}{2}\right) - \exp(-s^2),\end{aligned}$$

where $\sigma_{j,k}^2 = \mathrm{Var}(X_j - X_k) = 2(1 - \rho_X(j-k))$.

It follows that

$$\mathrm{E}\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(-isX_k)-\varphi_X(-s)\right)=\exp(-s^2(1-\rho_X(j-k)))-\exp(-s^2).$$

Therefore, we have

$$\Gamma_{s,t}=\sum_{k=-\infty}^{\infty}\left(\exp(-s^2(1-\rho_X(k)))-\exp(-s^2)\right)\left(\exp(-t^2(1-\rho_Y(k)))-\exp(-t^2)\right).$$

Moreover, it holds that

$$\begin{aligned}\left|\mathrm{E}\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(isX_k)-\varphi_X(s)\right)\right| &= \left|\mathrm{E}\exp(is(X_j+X_k))-\varphi_X^2(s)\right| \\ &= \left|\exp\left(-\frac{s^2\sigma_{j,k}^2}{2}\right)-\exp(-s^2)\right|,\end{aligned}$$

where $\sigma_{j,k}^2=\mathrm{Var}(X_j+X_k)=2(1+\rho_X(j-k))$.
It follows that

$$\mathrm{E}\left(\exp(isX_j)-\varphi_X(s)\right)\left(\exp(isX_k)-\varphi_X(s)\right)=\exp(-s^2(1+\rho_X(j-k)))-\exp(-s^2).$$

Therefore, we have

$$C_{s,t}=\sum_{k=-\infty}^{\infty}\left(\exp(-s^2(1+\rho_X(k)))-\exp(-s^2)\right)\left(\exp(-t^2(1+\rho_Y(k)))-\exp(-t^2)\right).$$

□

PROOF OF THEOREM 4.1. In order to establish the validity of the subsampling procedure, note that the triangular inequality yields

$$|\widehat{F}_{m_n,l_n}(t)-F_{T_n}(t)|\leq|\widehat{F}_{m_n,l_n}(t)-F_T(t)|+|F_T(t)-F_{T_n}(t)|. \tag{16}$$

The second term on the right-hand side of the inequality converges to 0 for all points of continuity t of F_T if the statistics T_n, $n\in\mathbb{N}$, are measurable and converge in distribution to a (non-degenerate) random variable T with distribution function F_T.
It remains to show that the first summand on the right-hand side of inequality (16) converges to 0 as well. As L^2-convergence implies convergence in probability, it suffices to show that $\lim_{n\to\infty}\mathrm{E}|\widehat{F}_{m_n,l_n}(t)-F_T(t)|^2=0$. For this purpose, we consider the following bias-variance decomposition:

$$\mathrm{E}\left(|\widehat{F}_{m_n,l_n}(t)-F_T(t)|^2\right)=\mathrm{Var}\left(\widehat{F}_{m_n,l_n}(t)\right)+\left|\mathrm{E}\widehat{F}_{m_n,l_n}(t)-F_T(t)\right|^2.$$

Stationarity and independence of the processes X_n, $n\in\mathbb{N}$, and Y_n, $n\in\mathbb{N}$, imply that $\mathrm{E}\widehat{F}_{m_n,l_n}(t)=F_{T_l}(t)$, so that, due to the convergence of T_n, the bias term of the above equation converges to 0 as l_n tends to ∞.
As a result, it remains to show that the variance term vanishes as n tends to ∞. Initially, note that

$$\begin{aligned}\mathrm{Var}\left(\widehat{F}_{m_n,l_n}(t)\right) &= \frac{1}{m_n}\mathrm{Var}\left(1_{\{T_{l_n,1}\leq t\}}\right)+\frac{2}{m_n^2}\sum_{k=2}^{m_n}(m_n-i+1)\mathrm{Cov}\left(1_{\{T_{l_n,1}\leq t\}},1_{\{T_{l_n,k}\leq t\}}\right) \\ &\leq \frac{2}{m_n}\sum_{k=1}^{m_n}\left|\mathrm{Cov}\left(1_{\{T_{l_n,1}\leq t\}},1_{\{T_{l_n,k}\leq t\}}\right)\right|\end{aligned}$$

due to stationarity.

Since $T_{l_n,k} = T_{l_n}\left(X_k, \ldots, X_{k+l_n-1}, Y_{k+d_n}, \ldots, Y_{k+d_n+l_n-1}\right)$,

$$\left|\mathrm{Cov}\left(1_{\{T_{l_n,1}\leq t\}}, 1_{\{T_{l_n,k}\leq t\}}\right)\right| \leq \rho\left(\sigma(X_i, Y_{d_n+i}, 1 \leq i \leq l_n), \sigma(X_j, Y_{d_n+j}, k \leq j \leq k+l_n-1)\right),$$

where $\sigma(X_i, Y_{d_n+i}, 1 \leq i \leq l_n)$, $\sigma(X_j, Y_{d_n+j}, k \leq j \leq k+l_n-1))$ resp., denotes the σ-field generated by the random variables $X_i, Y_{d_n+i}, 1 \leq i \leq l_n$, $X_j, Y_{d_n+j}, k \leq j \leq k+l_n-1)$, resp.

For $\beta \in (0,1)$, we split the sum of covariances into two parts:

$$\begin{aligned}
&\frac{1}{m_n}\sum_{k=1}^{m_n}\left|\mathrm{Cov}\left(1_{\{T_{l_n,1}\leq t\}}, 1_{\{T_{l_n,k}\leq t\}}\right)\right| \\
=&\frac{1}{m_n}\sum_{k=1}^{\lfloor n^\beta\rfloor}\left|\mathrm{Cov}\left(1_{\{T_{l_n,1}\leq t\}}, 1_{\{T_{l_n,k}\leq t\}}\right)\right| + \frac{1}{m_n}\sum_{k=\lfloor n^\beta\rfloor+1}^{m_n}\left|\mathrm{Cov}\left(1_{\{T_{l_n,1}\leq t\}}, 1_{\{T_{l_n,k}\leq t\}}\right)\right| \\
\leq&\frac{\lfloor n^\beta\rfloor}{m_n} + \frac{1}{m_n}\sum_{k=\lfloor n^\beta\rfloor+1}^{m_n}\rho\left(\sigma(X_i, Y_{d_n+i}, 1 \leq i \leq l_n), \sigma(X_j, Y_{d_n+j}, k \leq j \leq k+l_n-1)\right) \\
\leq&\frac{\lfloor n^\beta\rfloor}{m_n} + \frac{1}{m_n}\sum_{k=\lfloor n^\beta\rfloor+1}^{m_n}\rho_{k,l_n,d_n},
\end{aligned}$$

where

$$\rho_{k,l_n,d_n} := \rho\left(\sigma(X_i, Y_{d_n+i}, 1 \leq i \leq l_n), \sigma(X_j, Y_{d_n+j}, k \leq j \leq k+l_n-1)\right).$$

The first summand on the right-hand side of the inequality converges to 0 if $l_n = o(n)$ and $d_n = o(n)$. In order to show that the second summand converges to 0, a sufficiently good approximation to the sum of maximal correlations is needed. In particular, we have to show that

$$\sum_{k=\lfloor n^\beta\rfloor+1}^{m_n}\rho_{k,l_n,d_n} = o(m_n).$$

According to Bradley (2005), Theorem 5.1,

$$\rho(\mathcal{A}_1 \vee \mathcal{A}_2, \mathcal{B}_1 \vee \mathcal{B}_2) \leq \max\left\{\rho(\mathcal{A}_1, \mathcal{B}_1), \rho(\mathcal{A}_2, \mathcal{B}_2)\right\},$$

if $\mathcal{A}_1 \vee \mathcal{B}_1$ and $\mathcal{A}_2 \vee \mathcal{B}_2$ are independent.

As a result (and because of stationarity), it holds that

$$\rho_{k,l_n,d_n} \leq \max\{\rho_{k,l_n,X}, \rho_{k,l_n,Y}\},$$

where

$$\begin{aligned}
\rho_{k,l_n,X} :=& \rho\left(\sigma(X_i, 1 \leq i \leq l_n), \sigma(X_j, k \leq j \leq k+l_n-1)\right), \\
\rho_{k,l_n,Y} :=& \rho\left(\sigma(Y_i, 1 \leq i \leq l_n), \sigma(Y_j, k \leq j \leq k+l_n-1)\right).
\end{aligned}$$

For this reason, it suffices to show that

$$\sum_{k=\lfloor n^\beta\rfloor+1}^{m_n}\rho_{k,l_n,X} = \sum_{k=\lfloor n^\beta\rfloor+1}^{m_n}\rho_{k,l_n,Y} = o(m_n).$$

Betken and Wendler (2018) establish the following result:

LEMMA A.1 (Betken and Wendler (2018)). *Given a time series ξ_k, $k \geq 1$, satisfying Assumptions 1 and 2, there exist constants $C_1, C_2 \in (0,\infty)$, such that*

$$\rho\left(\sigma(\xi_i, 1 \leq i \leq l), \sigma(\xi_j, k+l \leq j \leq k+2l-1)\right) \leq C_1 l^D k^{-D} L_\rho(k) + C_2 l^2 k^{-D-1} \max\{L_\rho(k), 1\}$$

for all $k \in \mathbb{N}$ and all $l \in \{l_k, \ldots, k\}$.

We consider $\sum_{k=\lfloor n^\beta \rfloor+1}^{m_n} \rho_{k,l_n,X}$ only, since the same argument yields $\sum_{k=\lfloor n^\beta \rfloor+1}^{m_n} \rho_{k,l_n,Y} = o(m_n)$.
Let $\varepsilon > 0$. By assumption, $l_n \leq C_l n^\alpha$ for $\alpha := \frac{1}{2}(1+D_X) - \varepsilon$ and some constant $C_l \in (0,\infty)$. As a consequence of Potter's Theorem, for every $\delta > 0$, there exists a constant $C_\delta \in (0,\infty)$ such that $L_\rho(k) \leq C_\delta k^\delta$ for all $k \in \mathbb{N}$; see Theorem 1.5.6 in Bingham, Goldie and Teugels (1987).
Moreover, we choose $\beta > \alpha$ and n large enough such that $l_n < \frac{1}{2}\lfloor n^\beta \rfloor$. According to this, Lemma A.1 yields

$$\begin{aligned}
&\frac{1}{m_n} \sum_{k=\lfloor n^\beta \rfloor+1}^{m_n} \rho_{k,l_n,X} \\
&= \frac{1}{m_n} \sum_{k=\lfloor n^\beta \rfloor - l_n+1}^{m_n-l_n} \rho_{k+l_n,l_n,X} \\
&\leq C_1 l_n^{D_X} \frac{1}{m_n} \sum_{k=\lfloor n^\beta \rfloor - l_n+1}^{m_n-l_n} k^{-D_X} L_\rho(k) + C_2 \frac{l_n^2}{m_n} \sum_{k=\lfloor n^\beta \rfloor - l_n+1}^{m_n-l_n} k^{-D_X-1} \max\{L_\rho(k), 1\} \\
&\leq C_\delta C_1 \frac{l_n^{D_X}}{m_n} \sum_{k=\lfloor n^\beta \rfloor/2}^{m_n-l_n} k^{-D_X+\delta} + C_\delta C_2 \frac{l_n^2}{m_n} \sum_{k=\lfloor n^\beta \rfloor/2}^{m_n-l_n} k^{-D_X-1+\delta} \\
&\leq C\left(n^{D_X\alpha - \beta D_X + \beta\delta} + n^{2\alpha - \beta D_X - \beta + \delta\beta}\right)
\end{aligned}$$

for some constant $C \in (0,\infty)$. By definition of α and for a suitable choice of β, the right-hand side of the above inequality converges to 0. □

PROOF OF THEOREM 4.2. Note that

$$\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) = \sum_{i=1}^n X_i Y_i - \frac{1}{n} \sum_{i=1}^n X_i \sum_{i=1}^n Y_i.$$

Thus, for $D_X, D_Y \in \left(0, \frac{1}{2}\right)$, we have

$$\begin{aligned}
&n^{\frac{D_X+D_Y}{2}} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \\
=& n^{\frac{D_X+D_Y}{2}-1} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \sum_{i=1}^n X_i Y_i - n^{\frac{D_X+D_Y}{2}-2} L_X^{-\frac{1}{2}}(n) L_Y^{-\frac{1}{2}}(n) \sum_{i=1}^n X_i \sum_{j=1}^n Y_j.
\end{aligned}$$

According to the proof of Corollary 3.3 it holds that the first summand in the expression on the right-hand side of the above equation is $o_P(1)$. For this reason, the asymptotic distribution

is determined by the second summand, which converges in distribution to the product of two independent standard normally distributed random variables.
For $D_X, D_Y \in \left(\frac{1}{2}, 1\right)$ we have

$$\begin{aligned}
&\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_iY_i - n^{\frac{3}{2}-\frac{D_X+D_Y}{2}}L_Y^{\frac{1}{2}}(n)L_Y^{\frac{1}{2}}(n)n^{\frac{D_X}{2}-1}L_X^{-\frac{1}{2}}(n)\sum_{i=1}^{n}X_i n^{\frac{D_Y}{2}-1}L_Y^{-\frac{1}{2}}(n)\sum_{i=1}^{n}Y_i\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_iY_i + o_P(1).
\end{aligned}$$

According to Theorem A.2 $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_iY_i$ converges in distribution to a normally distributed random variables with expectation 0 and variance

$$\sigma^2 = \mathrm{E}\left(X_1^2\right)\mathrm{E}\left(Y_1^2\right) + 2\sum_{k=1}^{\infty}\mathrm{E}\left(X_1X_{k+1}\right)\mathrm{E}\left(Y_1Y_{k+1}\right) = \sum_{k=-\infty}^{\infty}\rho_X(k)\rho_Y(k).$$

□

APPENDIX: ADDITIONAL SIMULATION RESULTS

TABLE 1

Rejection rates of the hypothesis tests resulting from distance correlation obtained by subsampling based on "linearly" correlated fractional Gaussian noise time series $X_j,\ j=1,\dots,n$, $Y_j,\ j=1,\dots,n$ *with block length* l_n, $d=0.1n$, *Hurst parameters* H, *and cross-correlation* $\mathrm{Cov}(X_i,Y_j)=r\mathrm{Cov}(X_i,X_j)$, $1\le i,j\le n$. *The level of significance equals 5%.*

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$r=0$	$r=0.25$	$r=0.5$	$r=0$	$r=0.25$	$r=0.5$	$r=0$	$r=0.25$	$r=0.5$	$r=0$	$r=0.25$	$r=0.5$
$l_n=n^{0.4}$	100	0.131	0.713	0.999	0.175	0.745	0.997	0.099	0.492	0.971	0.024	0.117	0.558
	300	0.095	0.984	1.000	0.161	0.977	1.000	0.084	0.824	1.000	0.017	0.205	0.901
	500	0.089	1.000	1.000	0.148	0.998	1.000	0.072	0.939	1.000	0.015	0.292	0.974
	1000	0.081	1.000	1.000	0.131	1.000	1.000	0.065	0.996	1.000	0.012	0.460	0.998
$l_n=n^{0.5}$	100	0.116	0.689	0.998	0.155	0.706	0.996	0.095	0.474	0.965	0.034	0.145	0.620
	300	0.089	0.979	1.000	0.125	0.969	1.000	0.075	0.807	1.000	0.023	0.262	0.926
	500	0.084	0.999	1.000	0.126	0.998	1.000	0.066	0.931	1.000	0.021	0.337	0.981
	1000	0.073	1.000	1.000	0.112	1.000	1.000	0.062	0.996	1.000	0.018	0.543	0.999
$l_n=n^{0.6}$	100	0.123	0.683	0.997	0.150	0.679	0.994	0.103	0.478	0.960	0.045	0.180	0.670
	300	0.091	0.973	1.000	0.118	0.958	1.000	0.076	0.794	1.000	0.032	0.324	0.940
	500	0.086	0.999	1.000	0.116	0.995	1.000	0.071	0.927	1.000	0.029	0.420	0.984
	1000	0.074	1.000	1.000	0.097	1.000	1.000	0.062	0.996	1.000	0.030	0.605	0.999

TABLE 2

Rejection rates of the hypothesis tests resulting from Pearson's correlation obtained by subsampling based on "linearly" correlated fractional Gaussian noise time series $X_j,\ j=1,\dots,n$, $Y_j,\ j=1,\dots,n$ *with block length* l_n, $d=0.1n$, *Hurst parameters* H, *and cross-correlation* $\mathrm{Cov}(X_i,Y_j)=r\mathrm{Cov}(X_i,X_j)$, $1\le i,j\le n$. *The level of significance equals 5%.*

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$r=0$	$r=0.25$	$r=0.5$	$r=0$	$r=0.25$	$r=0.5$	$r=0$	$r=0.25$	$r=0.5$	$r=0$	$r=0.25$	$r=0.5$
$l_n=n^{0.4}$	100	0.125	0.770	1.000	0.172	0.772	0.999	0.174	0.646	0.992	0.113	0.354	0.852
	300	0.098	0.991	1.000	0.164	0.985	1.000	0.154	0.910	1.000	0.086	0.527	0.985
	500	0.083	1.000	1.000	0.160	0.999	1.000	0.138	0.973	1.000	0.073	0.613	0.998
	1000	0.085	1.000	1.000	0.139	1.000	1.000	0.139	0.999	1.000	0.064	0.792	1.000
$l_n=n^{0.5}$	100	0.110	0.748	0.999	0.165	0.740	0.999	0.154	0.595	0.985	0.107	0.338	0.833
	300	0.083	0.986	1.000	0.136	0.976	1.000	0.133	0.885	1.000	0.078	0.491	0.984
	500	0.078	1.000	1.000	0.125	0.999	1.000	0.122	0.961	1.000	0.074	0.594	0.996
	1000	0.069	1.000	1.000	0.124	1.000	1.000	0.109	0.999	1.000	0.068	0.758	1.000
$l_n=n^{0.6}$	100	0.120	0.730	0.999	0.148	0.720	0.996	0.144	0.574	0.978	0.110	0.339	0.818
	300	0.089	0.983	1.000	0.115	0.971	1.000	0.115	0.858	1.000	0.085	0.497	0.977
	500	0.084	1.000	1.000	0.118	0.996	1.000	0.102	0.954	1.000	0.074	0.587	0.994
	1000	0.072	1.000	1.000	0.101	1.000	1.000	0.095	0.996	1.000	0.065	0.739	1.000

TABLE 3

Rejection rates of the hypothesis tests resulting from distance correlation obtained by subsampling based on "parabolically" correlated time series $X_j,\ j = 1,\dots,n$, $Y_j,\ j = 1,\dots,n$ according to (12) with block length l_n, $d = 0.1n$, and Hurst parameter H. The level of significance equals 5%.

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$v=0.5$	$v=1$	$v=1.5$	$v=0.5$	$v=1$	$v=1.5$	$v=0.5$	$v=1$	$v=1.5$	$v=0.5$	$v=1$	$v=1.5$
$l_n=n^{0.4}$	100	0.312	0.916	1.000	0.373	0.944	1.000	0.140	0.698	0.997	0.005	0.174	0.631
	300	0.723	1.000	1.000	0.788	1.000	1.000	0.328	0.997	1.000	0.008	0.414	0.925
	500	0.950	1.000	1.000	0.962	1.000	1.000	0.545	1.000	1.000	0.017	0.572	0.998
	1000	1.000	1.000	1.000	1.000	1.000	1.000	0.927	1.000	1.000	0.056	0.805	1.000
$l_n=n^{0.5}$	100	0.292	0.897	1.000	0.348	0.917	1.000	0.168	0.724	0.994	0.015	0.285	0.741
	300	0.676	1.000	1.000	0.725	1.000	1.000	0.377	0.996	1.000	0.038	0.563	0.981
	500	0.932	1.000	1.000	0.942	1.000	1.000	0.607	1.000	1.000	0.076	0.749	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000	0.950	1.000	1.000	0.193	0.967	1.000
$l_n=n^{0.6}$	100	0.308	0.879	1.000	0.354	0.903	0.999	0.205	0.758	0.992	0.044	0.396	0.813
	300	0.665	1.000	1.000	0.706	1.000	1.000	0.449	0.996	1.000	0.117	0.722	0.990
	500	0.907	1.000	1.000	0.918	1.000	1.000	0.682	1.000	1.000	0.199	0.906	0.999
	1000	0.998	1.000	1.000	0.999	1.000	1.000	0.961	1.000	1.000	0.387	0.995	1.000

TABLE 4

Rejection rates of the hypothesis tests resulting from Pearson's correlation obtained by subsampling based on "parabolically" correlated time series $X_j,\ j=1,\dots,n$, $Y_j,\ j=1,\dots,n$ according to (12) with block length l_n, $d=0.1n$, and Hurst parameter H. The level of significance equals 5%.

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$v=0.5$	$v=0.75$	$v=1$	$v=0.5$	$v=0.75$	$v=1$	$v=0.5$	$v=0.75$	$v=1$	$v=0.5$	$v=0.75$	$v=1$
$l_n=n^{0.4}$	100	0.122	0.148	0.175	0.174	0.242	0.315	0.165	0.289	0.409	0.072	0.193	0.366
	300	0.118	0.136	0.185	0.197	0.281	0.384	0.196	0.358	0.468	0.093	0.308	0.467
	500	0.107	0.138	0.190	0.201	0.301	0.413	0.223	0.400	0.524	0.145	0.372	0.531
	1000	0.102	0.134	0.189	0.218	0.347	0.435	0.266	0.447	0.567	0.207	0.425	0.563
$l_n=n^{0.5}$	100	0.124	0.147	0.191	0.168	0.231	0.308	0.174	0.288	0.399	0.106	0.243	0.397
	300	0.102	0.133	0.181	0.180	0.265	0.353	0.213	0.363	0.475	0.152	0.370	0.506
	500	0.102	0.132	0.178	0.187	0.282	0.385	0.249	0.408	0.510	0.205	0.424	0.562
	1000	0.094	0.130	0.185	0.211	0.332	0.435	0.289	0.455	0.564	0.286	0.487	0.617
$l_n=n^{0.6}$	100	0.131	0.157	0.185	0.183	0.232	0.309	0.190	0.307	0.417	0.146	0.297	0.456
	300	0.109	0.141	0.187	0.175	0.267	0.362	0.235	0.383	0.486	0.233	0.435	0.566
	500	0.109	0.136	0.188	0.189	0.282	0.390	0.266	0.425	0.538	0.287	0.488	0.618
	1000	0.104	0.134	0.193	0.209	0.329	0.423	0.312	0.476	0.593	0.376	0.553	0.660

TABLE 5

Rejection rates of the hypothesis tests resulting from distance correlation obtained by subsampling based on "wavily" correlated time series $X_j,\ j=1,\dots,n$, $Y_j,\ j=1,\dots,n$ according to (13) with block length l_n, $d=0.1n$, and Hurst parameter H. The level of significance equals 5%.

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$v=2$	$v=3$	$v=4$	$v=2$	$v=3$	$v=4$	$v=2$	$v=3$	$v=4$	$v=2$	$v=3$	$v=4$
$l_n=n^{0.4}$	100	0.224	0.494	0.931	0.281	0.573	0.943	0.112	0.305	0.662	0.010	0.096	0.261
	300	0.442	0.964	1.000	0.538	0.970	1.000	0.184	0.598	0.993	0.030	0.158	0.387
	500	0.716	1.000	1.000	0.788	1.000	1.000	0.287	0.830	1.000	0.052	0.224	0.499
	1000	0.992	1.000	1.000	0.996	1.000	1.000	0.544	0.997	1.000	89.000	0.298	0.662
$l_n=n^{0.5}$	100	0.220	0.478	0.921	0.273	0.543	0.932	0.133	0.337	0.706	0.027	0.140	0.324
	300	0.415	0.941	1.000	0.500	0.953	1.000	0.215	0.645	0.993	0.064	0.225	0.489
	500	0.679	0.999	1.000	0.740	0.999	1.000	0.329	0.865	1.000	0.101	0.294	0.613
	1000	0.984	1.000	1.000	0.988	1.000	1.000	0.618	0.998	1.000	0.158	0.408	0.852
$l_n=n^{0.6}$	100	0.229	0.487	0.914	0.273	0.541	0.922	0.162	0.377	0.742	0.054	0.188	0.387
	300	0.431	0.919	1.000	0.491	0.931	1.000	0.270	0.702	0.994	0.116	0.305	0.626
	500	0.658	0.995	1.000	0.717	0.996	1.000	0.406	0.899	1.000	0.163	0.389	0.788
	1000	0.969	1.000	1.000	0.971	1.000	1.000	0.712	0.998	1.000	0.250	0.549	0.970

TABLE 6

Rejection rates of the hypothesis tests resulting from Pearson's correlation obtained by subsampling based on "wavily" correlated time series $X_j,\ j=1,\dots,n$, $Y_j,\ j=1,\dots,n$ according to (13) with block length l_n, $d=0.1n$, and Hurst parameter H. The level of significance equals 5%.

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$v=2$	$v=3$	$v=4$	$v=2$	$v=3$	$v=4$	$v=2$	$v=3$	$v=4$	$v=2$	$v=3$	$v=4$
$l_n=n^{0.4}$	100	0.144	0.186	0.282	0.199	0.290	0.430	0.168	0.315	0.476	0.066	0.183	0.351
	300	0.118	0.188	0.283	0.213	0.328	0.467	0.204	0.383	0.549	0.090	0.252	0.443
	500	0.120	0.176	0.294	0.223	0.362	0.497	0.238	0.424	0.588	0.127	0.306	0.499
	1000	0.128	0.193	0.300	0.244	0.391	0.531	0.287	0.483	0.631	0.167	0.370	0.568
$l_n=n^{0.5}$	100	0.140	0.184	0.283	0.185	0.273	0.417	0.173	0.319	0.464	0.087	0.213	0.370
	300	0.118	0.181	0.280	0.196	0.309	0.453	0.211	0.385	0.545	0.128	0.300	0.472
	500	0.116	0.175	0.292	0.211	0.347	0.487	0.249	0.429	0.590	0.168	0.350	0.530
	1000	0.126	0.187	0.296	0.229	0.374	0.516	0.297	0.492	0.633	0.228	0.429	0.597
$l_n=n^{0.6}$	100	0.148	0.188	0.298	0.188	0.276	0.413	0.188	0.330	0.473	0.120	0.251	0.398
	300	0.126	0.192	0.283	0.195	0.312	0.452	0.231	0.396	0.558	0.176	0.348	0.516
	500	0.122	0.179	0.300	0.214	0.343	0.485	0.271	0.447	0.602	0.224	0.408	0.566
	1000	0.128	0.190	0.299	0.224	0.372	0.510	0.325	0.506	0.653	0.302	0.495	0.645

TABLE 7

Rejection rates of the hypothesis tests resulting from distance correlation obtained by subsampling based on "rectangularly" correlated time series $X_j,\ j = 1,\dots,n$, $Y_j,\ j = 1,\dots,n$ according to (14) with block length l_n, $d = 0.1n$, and Hurst parameters H. The level of significance equals 5%.

		$H = 0.6$			$H = 0.7$			$H = 0.8$			$H = 0.9$		
	n	$v = 1$	$v = 2$	$v = 3$	$v = 1$	$v = 2$	$v = 3$	$v = 1$	$v = 2$	$v = 3$	$v = 1$	$v = 2$	$v = 3$
$l_n = n^{0.4}$	100	0.273	0.448	0.381	0.391	0.620	0.570	0.199	0.364	0.367	0.071	0.140	0.172
	300	0.733	0.990	0.985	0.819	0.995	0.994	0.362	0.744	0.734	0.073	0.199	0.246
	500	0.966	1.000	1.000	0.980	1.000	1.000	0.553	0.952	0.941	0.087	0.271	0.316
	1000	1.000	1.000	1.000	1.000	1.000	1.000	0.929	1.000	1.000	0.114	0.354	0.412
$l_n = n^{0.5}$	100	0.254	0.425	0.366	0.343	0.551	0.506	0.198	0.348	0.361	0.092	0.186	0.209
	300	0.663	0.972	0.962	0.720	0.977	0.978	0.345	0.714	0.702	0.104	0.250	0.303
	500	0.932	0.999	1.000	0.948	1.000	1.000	0.533	0.934	0.917	0.120	0.330	0.369
	1000	1.000	1.000	1.000	1.000	1.000	1.000	0.893	1.000	1.000	0.169	0.432	0.488
$l_n = n^{0.6}$	100	0.266	0.437	0.399	0.334	0.530	0.498	0.208	0.368	0.388	0.118	0.228	0.252
	300	0.646	0.946	0.937	0.663	0.942	0.950	0.368	0.709	0.697	0.142	0.317	0.372
	500	0.893	0.998	0.997	0.902	0.997	0.997	0.525	0.906	0.895	0.167	0.406	0.439
	1000	1.000	1.000	1.000	0.998	1.000	1.000	0.852	0.997	0.998	0.227	0.514	0.564

TABLE 8

Rejection rates of the hypothesis tests resulting from Pearson's correlation obtained by subsampling based on "rectangularly" correlated time series $X_j,\ j=1,\ldots,n$, $Y_j,\ j=1,\ldots,n$ according to (14) with block length l_n, $d=0.1n$, and Hurst parameters H. The level of significance equals 5%

		$H=0.6$			$H=0.7$			$H=0.8$			$H=0.9$		
	n	$v=1$	$v=2$	$v=3$	$v=1$	$v=2$	$v=3$	$v=1$	$v=2$	$v=3$	$v=1$	$v=2$	$v=3$
$l_n=n^{0.4}$	100	0.090	0.059	0.038	0.141	0.093	0.072	0.145	0.124	0.112	0.131	0.150	0.155
	300	0.078	0.037	0.023	0.114	0.070	0.050	0.116	0.097	0.082	0.104	0.126	0.135
	500	0.065	0.034	0.017	0.104	0.061	0.036	0.115	0.089	0.070	0.097	0.124	0.124
	1000	0.058	0.025	0.014	0.091	0.053	0.036	0.101	0.085	0.063	0.092	0.115	0.122
$l_n=n^{0.5}$	100	0.092	0.063	0.042	0.131	0.091	0.065	0.139	0.122	0.108	0.131	0.153	0.155
	300	0.076	0.038	0.025	0.104	0.059	0.045	0.103	0.089	0.076	0.102	0.123	0.132
	500	0.066	0.034	0.019	0.093	0.053	0.030	0.104	0.078	0.063	0.094	0.122	0.122
	1000	0.055	0.025	0.015	0.077	0.042	0.028	0.090	0.074	0.056	0.090	0.107	0.119
$l_n=n^{0.6}$	100	0.113	0.079	0.058	0.138	0.102	0.078	0.148	0.131	0.122	0.146	0.166	0.170
	300	0.090	0.052	0.038	0.106	0.071	0.052	0.109	0.092	0.077	0.111	0.131	0.142
	500	0.075	0.042	0.028	0.091	0.055	0.034	0.106	0.084	0.064	0.106	0.128	0.127
	1000	0.060	0.034	0.022	0.075	0.046	0.034	0.092	0.072	0.057	0.098	0.115	0.125