Ordinal Pattern Analysis: Limit Theorems for Multivariate Long-Range Dependent Gaussian Time Series and a Comparison to Multivariate Dependence Measures

DISSERTATION

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

vorgelegt von
Ines Nüßgen geb. Münker, M.Sc.

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen
Siegen 2021

gedruckt auf alterungsbeständigem holz- und säurefreiem Papier

1. Gutachter: Prof. Dr. Alexander Schnurr, Universität Siegen
2. Gutachter: Prof. Dr. Herold Dehling, Ruhr-Universität Bochum

Tag der mündlichen Prüfung: 09. Juli 2021

“Education is neither eastern nor western. Education is education and it’s the right of every human being.”

Malala Yousafzai

Acknowledgments

First of all, I would like thank all those who have contributed in various ways to the success of this work over the past years.

I thank my supervisor Alexander Schnurr for introducing me to the interesting field of ordinal pattern analysis, for providing me the opportunity to do a PhD, the alongside trust and for encouraging me to do independent mathematical research. I would like to thank him, Herold Dehling and Annika Betken for the successful collaboration, for many interesting discussions and hours working together during our stays in Bochum and Siegen. Furthermore, I am appreciative for Marie-Christine Düker's helpful advice on multivariate long-range dependence.

The financial support from the "Deutsche Forschungsgemeinschaft (DFG)" is gratefully acknowledged.

I would like to say thank you to Matthias Reuber, Kevin Berk, Anke Kramer, Stephie Wied, Annabel Roth, Tim Fabian Korzeniowski, Franziska Wollny and Sebastian Rickelhoff for countless discussions on all kinds of topics as well as for their contributions to the careful proofreading of this thesis, great company at conferences and constant distraction from work.

In general, there was a very familiar atmosphere at the stochastics group at University of Siegen where I truly enjoyed working. Special thanks go to Marco Oesting and Dustin Kremer for helpful advice at any time and for sharing their dedication to mathematics.

Sincere thanks go to my family: to my parents-in-law, Gerda and Toni, for their generosity and great help, especially in the last year. To my two brothers, Sven and Tobias, for many fruitful and not so fruitful discussions and for sometimes necessary point-blank honesty. To my parents Antje and Wolfram for their lifelong and unconditional support and their unbounded faith in me.

Last, but certainly not least, I would like to thank my husband Max for encouraging me in hard times by always listening and enduring my ups and downs with remarkable patience. For his positive attitude towards life. You and Finn make my life so much brighter every single day!

Abstract

Ordinal pattern analysis provides a possibility to study dependence structures in multivariate time series with few assumptions on the underlying stochastic model. Focussing on a univariate time series, we discuss the concept of ordinal pattern probabilities that deals with the occurance of one fixed ordinal pattern within this time series. Based on this method, the dependence within the time series is investigated. Turning to the multivariate case, ordinal pattern dependence allows us to compare data sets by studying the probability of coincident ordinal patterns at the same points in time. Applying these two approaches we are able to detect linear as well as non-linear dependence. We extend the theoretical framework for estimators in the context of these two concepts to multivariate long-range dependent Gaussian time series, allowing for pure long-range dependence as well as for mixed cases of short- and long-range dependent univariate components. We provide limit theorems for functionals with Hermite rank 1 and 2, as it turns out that the estimators in the context of ordinal pattern analysis are represented by these two classes. For functionals with Hermite rank 2, the asymptotic distribution is non-Gaussian and follows a Rosenblatt distribution. Further, we investigate the differences in the asymptotics considering multivariate stationary Gaussian time series and multivariate Gaussian time series with stationary increments, which is less restrictive. A generalization to more flexible models in the context of ordinal pattern dependence is also provided. The first part of this thesis closes with a simulation study that illustrates the theoretical results. The second part of this work puts ordinal pattern dependence in the perspective of multivariate dependence measures. We compare ordinal pattern dependence to classical dependence measures like Pearson's ρ and Kendall's τ. By precisely distinguishing between measures that arise in a time series context and models that study dependence between or within multivariate random vectors, we identify differences and provide relations between ordinal pattern dependence and the classical approaches. Finally, a simulation study and a real-world data analysis in the field of hydrology that emphasizes the practical value of ordinal pattern dependence complete this work.

Zusammenfassung

Die ordinale Muster-Analyse bietet eine Möglichkeit Abhängigkeitsstrukturen in multivariaten Zeitreihen mit wenigen Annahmen an das zugrundeliegende stochastische Modell zu untersuchen. Wir stellen das Konzept ordinaler Muster-Wahrscheinlichkeiten vor, welches sich mit dem Auftreten eines festen ordinalen Musters innerhalb dieser Zeitreihe beschäftigt und, auf dessen Basis, die Abhängigkeiten innerhalb der Zeitreihe beschreibt. Wenden wir uns dem multivariaten Fall zu, erlaubt uns die ordinale Muster-Abhängigkeit Datensätze zu vergleichen, indem wir die Wahrscheinlichkeit übereinstimmender ordinaler Muster zu den gleichen Zeitpunkten untersuchen. Durch die Anwendung dieses Modells sind wir in der Lage, sowohl lineare als auch nicht-lineare Abhängigkeiten zu erkennen. Wir erweitern den theoretischen Hintergrund der Schätzer im Kontext dieser beiden Konzepte auf multivariate langzeitabhängige Gaußsche Zeitreihen, wobei wir sowohl reine Langzeitabhängigkeit als auch gemischte Fälle von kurz- und langzeitabhängigen univariaten Komponenten berücksichtigen. Zunächst werden Grenzwertsätze für Funktionen mit Hermite-Rang 1 und 2 bewiesen, da sich herausstellt, dass Schätzer im Kontext der ordinalen Muster-Analyse durch diese Klassen repräsentiert werden. Die asymptotische Verteilung ist im Fall von Hermite-Rang 2 nicht Gaußsch, sondern folgt einer Rosenblatt-Verteilung. Weiterhin untersuchen wir die Unterschiede in der Asymptotik der Schätzer, einerseits für zugrundeliegende multivariate stationäre Gaußsche Zeitreihen und andererseits für multivariate Gaußsche Zeitreihen mit stationären Zuwächsen, was ein weniger restriktives Modell darstellt. Eine Verallgemeinerung der Resultate für flexiblere und abgewandelte Modelle im Kontext der ordinalen Muster-Abhängigkeit wird ebenfalls betrachtet. Der erste Teil der Arbeit schließt mit einer Simulationsstudie, welche die theoretischen Resultate veranschaulicht. Der zweite Teil der Dissertation befasst sich mit der Einordnung der ordinalen Muster-Abhängigkeit in die Klasse der multivariaten Abhängigkeitsmaße. Wir vergleichen die ordinale Muster-Abhängigkeit mit klassischen Abhängigkeitsmaßen wie Pearsons ρ und Kendalls τ. Durch die genaue Unterscheidung zwischen Maßen, welche sich auf einen Zeitreihenkontext beziehen, und Modellen, welche Abhängigkeiten zwischen oder innerhalb von multivariaten Zufallsvektoren untersuchen, identifizieren wir Unterschiede und stellen Beziehungen zwischen der ordinalen Muster-Abhängigkeit und den klassischen Ansätzen her. Schließlich runden eine Simulationsstudie sowie eine Analyse realer Daten aus dem Bereich der Hydrologie diese Arbeit ab, welche den praktischen Nutzwert der ordinalen Muster-Abhängigkeit unterstreicht.

Contents

1 Introduction

1.1 Historical background and motivation

In times of big data, a model that analyzes dependence structures between data sets depending on few technical assumptions is in great demand. Reducing high dimensional and complex data to minimal information that is still substantial enough to yield a significant and convincing answer to the questions of interest, is of strong need for applications with regard to the computational effort. This work focuses on dependence structures in multidimensional data sets that naturally contain a large amout of data points and that do not exhibit the well studied case of independence between observations. The method that we want to investigate in order to detect and classify this dependence is *ordinal pattern analysis*. It is known to be robust concerning disturbances in the data and invariant under monotone transformations of the time series under consideration, which are properties of high practical interest. Therefore, ordinal pattern analysis and in particular ordinal pattern dependence are of vital importance, especially in contrast to classical dependence measures, since the latter ones tend to overestimate dependence of data with measurement disturbances and to underestimate data that do not exhibit linear dependence.

In the literature, ordinal analysis was introduced in [5] in 2002, followed by [4] and [6]. They investigated order patterns that are closely linked to ordinal patterns, which in turn were discussed in [39]. In the following years, ordinal analysis became increasingly popular in applications. In neurological science this method was used to investigate brain activity in [38], [53] and [70]. Cardiological research in the field of heart rate variability based on ordinal patterns was discussed in [28].

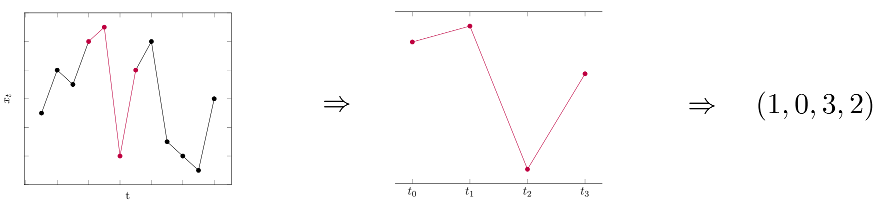

Figure 1.1: Example of the extraction of an ordinal pattern in a given data set.

In general, the method works as follows: we consider a data set $(x_t)_{t=0,1,\dots}$ and we need to fix $h \in \mathbb{N}$ to describe the length of the ordinal pattern we are interested in. Note that the number of data points taken into account is given by $h+1$. In Figure 1.1 we extract one pattern from the data using $h = 3$. The permutation is obtained by sorting the absolute values of the data points from largest to smallest and writing down the indices of the time points from left to right. This yields a large reduction of complexity because the permutation only describes the relative positions of the data points, but does not give any information concerning the absolute values or the L^2-distance between the original data points. Therefore, this results in a strong decrease of information necessary to do further analysis using this method.

Next, we turn to the stochastic framework and focus on the determination of the probability of one fixed ordinal pattern arising in a univariate time series. In [58], a heuristic estimator was proposed for a class of processes including those exhibiting short-range dependence, that is, the autocorrelations of the stochastic process under consideration have a power-law decay. This implicates absolutely summable autocorrelations. The estimator counts the occurence of the fixed ordinal pattern in fixed windows of time that are stepwisely shifted along the timeline of the data set. Then, the sum is divided by the maximal number of possibly occurenced pattern. In [60], the distribution of ordinal patterns was investigated for Gaussian processes with stationary increments, including the long-range dependent setting, that is, slowly hyperbolically decaying autocorrelations. In both papers limit theorems for the investigated estimators were given. However, in the latter article, the case of pure long-range dependence, which means that the autocorrelations of the estimators under consideration are still slowly hyperbolically decaying, was omitted.

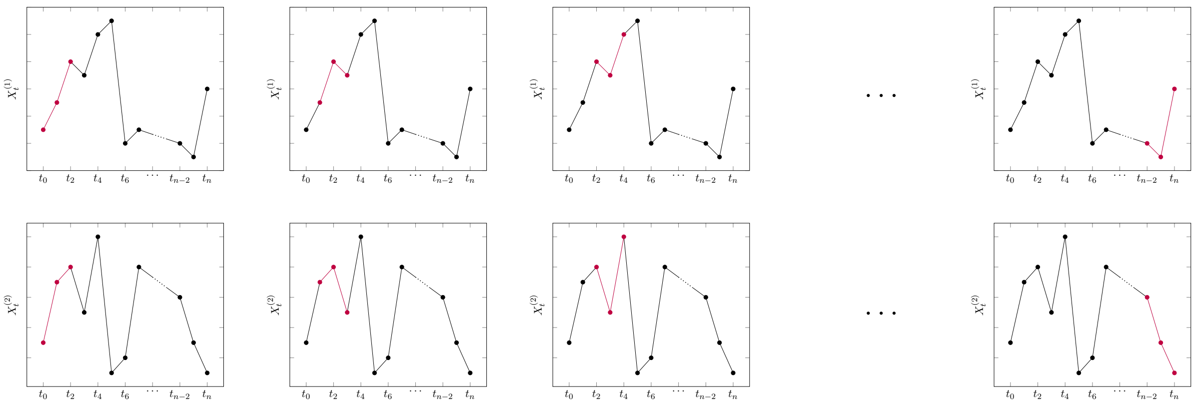

Figure 1.2: Illustration of estimation of ordinal pattern dependence.

The next approach that we want to present is ordinal pattern dependence. Originally, it was introduced in [57], to detect and model dependencies between two time series based on the co-movement of the data. In the mentioned article the method was applied to financial time series. We are no longer considering a fixed pattern, but we focus on the occurence of coincident ordinal patterns in the same moving window of two time series and, therefore, turn to the multidimensional case. The estimation works analogously to the previous case by comparing whether the pattern coincide in consecutive moving windows. An illustration of this idea is given in Figure 1.2. This method has already proved to be useful in applications. It was used to analyze hydrological data sets as discussed in [24] and [52], where the latter one dealt with ordinal pattern dependence in the context of extreme value theory.
From today's point of view, [58] is the only article that dealt with the investigation of the theoretical framework of estimators of ordinal pattern dependence. They introduced a suitable normalization to obtain a measure that is comparable to classical dependence measures and

provided asymptotic results for estimators of ordinal pattern dependence and further parameters in this context in a theoretical setting including short-range dependence. Hence, long-range dependent processes were not tackled in the context of ordinal pattern dependence at all.

In order to do so, we need to take a closer look at limit theorems for long-range dependent time series. In the univariate case this is a well investigated field, beginning almost fifty years ago with [64] and [15], who considered limit theorems for non-linear functionals of univariate Gaussian time series. This theory was extended in [3], where limit theorems for vectors of different functionals were derived, however, the underlying process remains to be univariate Gaussian. The multivariate approach was first considered in [31] and [2] and later in [19] who focused on limit theorems with underlying vector-valued Gaussian processes including long-range dependence. In the case of pure long-range dependence resulting in asymptotic distribution that is no longer Gaussian, they restricted theirselves to the case of independent entries within each Gaussian random vector. Recently, with [18] and [23] two articles were published that dealt with multivariate linear long-range dependent time series and hence softened the Gaussian assumption.

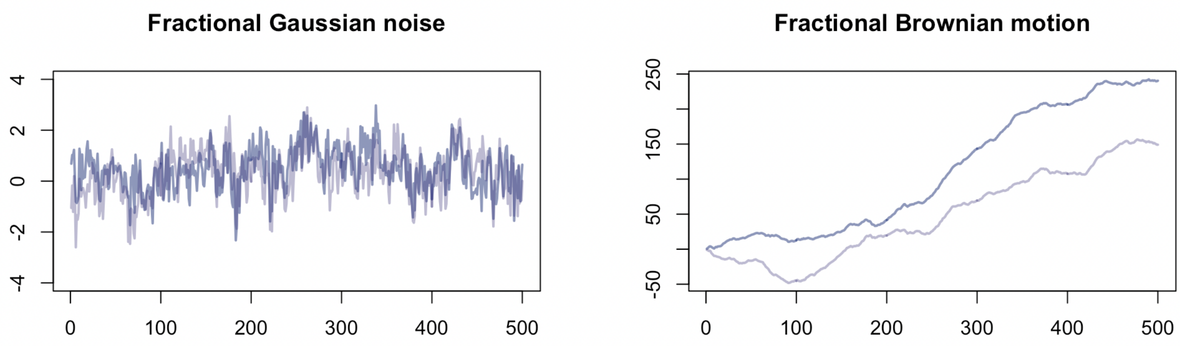

Figure 1.3: Sample paths of two correlated fractional Gaussian noise processes and of the corresponding fractional Brownian motions.

In general, the class of long-range dependent processes appears in many applications, e.g. hydrologic data sets are often known to exhibit this property. Recently long-range dependent processes gained popularity since they arose in the context of artificial intelligence, see [8] and neural networks, see [69]. Both topics are seminal for developing technology in the next decades.
Two explicit examples of commonly used long-range dependent processes are briefly highlighted: The fractional extensions of the well-known ARMA time series, so called FARIMA processes, that were used in various applications as network traffic prediction in [59] or in the field of biophysics in [17], are included in this setting. Fractional Gaussian noise as increment process of fractional Brownian motion, enjoyed increasing popularity in the last years in applications as neurological science, for example to model EEG data, see [36] or in the context of Alzheimer's disease in [46]. Therefore, the connection of ordinal analysis to long-range dependence does not only yield an interesting theoretical challenge but also happens for practical reasons. In Figure 1.3 sample paths of two correlated fractional Gaussian noise processes and of the corresponding fractional

Brownian motions are presented for $n = 500$ data points and Hurst paramter $H = 0.9$. The theoretical details of these processes are explained in Section 2.3.

The aim of this work is to investigate ordinal analysis in the context of pure long-range dependent time series. We can restrict ourselves to Gaussian time series, since ordinal patterns are not affected by monotone transformations and, therefore, the data sets can be transformed to exhibiting a Gaussian distribution. Further preprocessing of real-world data happens in the case of ties, as a small white noise disturbance does not destroy the ordinal structure either. Moreover, we are interested in mixed-cases, in which short-range as well as long-range dependent processes might occur. Summing up, the intended contribution of this thesis is to complement the theoretical framework of ordinal analysis in the case of long- and short-range dependent processes. In applications the motivation of this completion is given in the possibility to provide a robust measure to e.g. hydrological and neurological data sets that often happen to contain small noises or outliers, which biase the classical dependence measures.

A question that naturally arises is the one of comparing ordinal pattern dependence to classical dependence measures. The first sticking point here is to make ordinal pattern dependence comparable to univariate and multivariate dependence measures. In particular, the latter ones are not defined in a time series context in general. Hence, the goal of this thesis is to light up the theoretical background of those classical measures and of ordinal pattern dependence to work out the differences. A further focus is on providing examples to emphasize and illustrate these differences in a practical context, as well as on obtaining theoretical relations between ordinal pattern dependence and classical dependence measures.

1.2 Outline of this thesis

In the first part of this work, our aim is to close the gap between multivariate time series exhibiting different dependence structures as long-range dependence and ordinal analysis. In Chapter 2, we guide the reader through the theoretical background. We mathematically introduce the topic of long-range dependence and the linked processes described above and build the framework needed to tackle our theoretical challenges by providing the mathematical tools. Fully prepared, theoretical results in form of limit theorems for multivariate Gaussian time series are derived in Chapter 3. In the first two sections, we focus on limit theorems with Hermite rank 1 and 2, respectively, as they arise in the context of ordinal pattern analysis later on. The last section deals with a special type of functionals, namely those who are uniquely determined by the increments of the underlying process. In this case, a twist happens to the limit distribution proven in the first two sections.
We turn to the main chapter given by Chapter 4 that seeks to combine ordinal pattern analysis to the asymptotic theory for multivariate time series exhibiting the dependence structures investigated in Chapter 3. An extensive and illustrative introduction to ordinal patterns in general is provided in Section 4.1. Then, we turn to the estimation and simulation of ordinal pattern probabilities in Section 4.2 which is represented by an underlying univariate Gaussian

process. The multivariate extension that allows for componentwise different dependencies arises in Section 4.3. Here, we introduce standard ordinal pattern dependence and investigate different estimators. To increase practical applicability, we modify and generalize the model of ordinal pattern dependence in the final section of this chapter.

The second part of this thesis deals with the comparison of ordinal pattern dependence to different, possibly multivariate, dependence measures. First, a pilot study awakens the interest on this topic by clearly indicating differences between ordinal pattern dependence and classical measures of dependence. Then, we turn to the theoretical challenges and provide results to make ordinal pattern dependence comparable to other dependence measures. This chapter is completed by simulations to get a more illustrative understanding.

A real data study is given in Chapter 6. We emphasize the practical impact of this work by analyzing a data set that fits into the theoretical background such that the theoretical results derived in the combination of long-range dependence to ordinal pattern analysis hold. In comparison to the other dependence measures discussed in the previous chapter, we confirm ordinal pattern dependence as possibility to compare dependence structures between two data sets and highlight the advantages of robustness and computational efficiency.

2 Mathematical preliminaries

In this chapter we introduce the mathematical framework needed to step into the theory we want to develop in this thesis. In general, we assume that in the background we have a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ as basis for mathematical operations. Furthermore, the convergence results in this work hold for $n \to \infty$ if not declared differently.
All simulation studies presented in this thesis are derived by computing on a MacBook Pro 2,3 GHz Dual-Core Intel Core i5.

We turn to the framework of stochastic processes. First, basic information concerning univariate as well as multivariate Gaussian processes in the context of short- and long-range dependence is given. We continue with turning to the role of Hermite polynomials and later on introduce the class of Hermite-Rosenblatt processes after giving a brief excursus on Wiener-Itô integration.

2.1 Univariate stochastic processes

We begin with the basic definition of a stochastic process.

Definition 2.1 *(Stochastic process, [16], Definition 1.2.1. and Remark 1)*
A stochastic process is a family of random variables $(X_j)_{j \in J}$, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Remark, that in time series analysis, the index set J is a set of time points. In most cases in this work we use $J = \mathbb{Z}$, unless it is declared differently. If the index set J is countably infinite, we call the stochastic process *discrete* and we use the notation in the definition above. In order to mark the case when T is uncountably infinite and the process is called *continuous* over time, we use the notation $(X(t))_{t \in \mathbb{R}}$ to distinguish this class of processes from the discrete ones.

We continue with stationarity as an important property of a stochastic process and, therefore, turn to the definition of the autocovariance-function first:

Definition 2.2 *(Autocovariance-function, [16], Definition 1.3.1)*
If $(X_j)_{j \in \mathbb{Z}}$ is a stochastic process such that $\operatorname{Var}(X_j) < \infty$ for each $j \in \mathbb{Z}$, then the autocovariance function $\gamma_X(\cdot, \cdot)$ of $(X_j)_{j \in \mathbb{Z}}$ is defined by

$$\gamma_X(r, s) := Cov(X_r, X_s) = \mathbb{E}\left[(X_r - \mathbb{E}(X_r))(X_s - \mathbb{E}(X_s))\right], \ r, s \in \mathbb{Z}.$$

We are now able to define stationarity.

Definition 2.3 *(Weak stationarity, [16], Definition 1.3.2.)*
The time series $(X_j)_{j\in\mathbb{Z}}$ *is said to be (weak) stationary if*

(i) $\mathbb{E}|X_j|^2 < \infty$ *for all* $j \in \mathbb{Z}$,

(ii) there exists $m \in \mathbb{R}: \quad \mathbb{E}(X_j) = m$ *for all* $j \in \mathbb{Z}$,

(iii) $\gamma_X(r,s) = \gamma_X(r+t, s+t),$ *for all* $r, s, t \in \mathbb{Z}$.

This yields a possibility to redefine the covariance function, since for a stationary $(X_j)_{j\in\mathbb{Z}}$ it holds that $\gamma_X(r,s) = \gamma_X(r-s,0)$.

Definition 2.4 *(Covariance- and correlation-function for stationary processes, [16], Remark 2)*
For a stationary process $(X_j)_{t\in\mathbb{Z}}$, *we redefine*

$$\gamma_X(l) := \gamma_X(l,0) = Cov(X_j, X_{j+l}),$$

for all $j, l \in \mathbb{Z}$.
The function $\gamma_X(\cdot)$ *is referred to as the autocovariance function of* $(X_j)_{j\in\mathbb{Z}}$ *and* $\gamma_X(l)$ *as its value at lag* l. *The autocorrelation function of* (X_j) *is defined analogously as the function whose value at lag* l *is*

$$r_X(l) := \frac{\gamma_X(l)}{\gamma_X(0)} = \operatorname{Cor}(X_j, X_{j+l}).$$

For completeness, we also give the definition of strict stationarity.

Definition 2.5 *(Strict stationarity, [16], Definition 1.3.3.)*
The stochastic process $(X_j)_{j\in\mathbb{Z}}$ *is said to be strictly stationary if the joint distributions of* $(X_{j_1}, \dots, X_{j_k})^t$ *and* $(X_{j_1+l}, \dots, X_{j_k+l})^t$ *are the same for all* $k \in \mathbb{N}$ *and* $j_1, \dots, j_k, l \in \mathbb{Z}$.

We now turn to dependence properties of stationary stochastic processes, so called short- and long-range dependence, that play a cruicial role in the course of this thesis, as the title already indicates. Note that there are several non-equivalent definitions of long-range dependence in the literature, for details see [54], Sec. 2.1. Therefore, we focus on one definition that is commonly used.

Definition 2.6 *(Long-range dependence, [54], Sec. 2.1 Condition II)*
A weak stationary time series $(X_j)_{j\in\mathbb{Z}}$ *is called long-range dependent (LRD) if its autocorrelation function satisfies*

$$r_X(k) = L_2(k)k^{2d-1}, \quad k = 0, 1, \dots,$$

with $d \in \left(0, \frac{1}{2}\right)$ *and* L_2 *is a slowly varying function at infinity.*

We contuinue with short-range dependence.

Definition 2.7 *(Short-range dependence, [54], Definition 2.3.1)*
A weak stationary time series $(X_j)_{j\in\mathbb{Z}}$ *is called short-range dependent (SRD) if its autocorrelations are absolutely summable; that is, if*

$$\sum_{k=-\infty}^{\infty} |r_X(k)| < \infty.$$

Note that the definition of long-range dependence above implies that the autocorrelations are not absolutely summable, see Proposition 2.2.4 in [54].
We introduce a special kind of short-range dependence time series in the next definition.

Definition 2.8 *(Antipersistance, [54], Definition 2.3.2)*
A short-range dependent time series $(X_j)_{j\in\mathbb{Z}}$ is called antipersistent if

$$\sum_{k=-\infty}^{\infty} r_X(k) = 0.$$

A property of a stochastic process that is closely linked to long-range dependence is introduced in the next definition.

Definition 2.9 *(Self-similarity, [54], Definition 2.5.1)*
A stochastic process $(X(t))_{t\in\mathbb{R}}$ is called self-similar or H-self-similar if there is $H > 0$ such that, for all $c > 0$ and $t \in \mathbb{R}$

$$X(ct) \stackrel{\mathcal{D}}{=} c^H X(t).$$

To close this section of general information on univariate stochastic processes, we consider the class of linear processes and give a result that incorporates stationary purely non-deterministic processes into this particular class.

Definition 2.10 *(Linear process, [9], Sec. 4.2.4)*
We call the stochastic process $(X_j)_{j\in\mathbb{Z}}$ a causal linear process if it has the representation

$$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \tag{2.1}$$

where $\sum_{j=0}^{\infty} a_j^2 < \infty$ and $(\varepsilon_t)_{t\in\mathbb{Z}}$ are i.i.d. zero mean random variables with $\operatorname{Var}(\varepsilon_1) = \sigma_\varepsilon^2 < \infty$.

The Wold-decomposition as it is discussed in [40], Sec. 1.4. states that it is possible to decompose a stationary purely non-deterministic stochastic process X_j into a linear representation as given in (2.1).

2.2 Multivariate stochastic processes

It is possible to extend the theory of univariate stochastic processes to a multidimensional model, that yields better possibilities to describing and imaging the complexity of real-world phenomenons. Especially when not only the dependence structure within one process but the interdependence between several processes is of interest, the multivariate generalization yields a huge amendment.

Definition 2.11 *(d-dimensional stochastic process, [16], Sec. 11.1)*
A vector-valued stochastic process $(X_j)_{j\in J} = \left(X_j^{(1)}, \ldots, X_j^{(d)}\right)^t$ is called d-dimensional stochastic process.

Similar to the univariate case, in general, we use $J = \mathbb{Z}$ in the following. Analogously the concept of stationarity can be extended to the multivariate case. The following passage is taken from [16], Sec. 11.1.

Definition 2.12 *(Stationarity of a multivariate time series, [16], Definition 11.1.1)*
The d-dimensional stochastic process $(X_j)_{j\in\mathbb{Z}}$ *with mean vector*

$$\mu_j = \mathbb{E}\left(X_j\right) = \left(\mu_j^{(1)}, \ldots, \mu_j^{(d)}\right)^t$$

and covariance matrix

$$\Gamma\left(j+k, j\right) = \mathbb{E}\left[\left(X_{j+k} - \mu_{j+k}\right)\left(X_j - \mu_j\right)^t\right]$$

is said to be stationary if μ_j *and* $\Gamma\left(j+k, j\right)$, $k = 0, 1, \ldots$ *are independent of* j.

Therefore, we use the notation

$$\begin{aligned} \mu :&= \mathbb{E}\left(X_0\right) = \left(\mu^{(1)}, \ldots, \mu^{(d)}\right)^t, \\ \Gamma_X(k) :&= \mathbb{E}\left[\left(X_k - \mu\right)\left(X_0 - \mu\right)^t\right] \end{aligned}$$

for stationary time series. The entries of the covariance matrix are determined by the values of the cross-covariance functions $\gamma^{(p,q)}(\cdot)$, that are for a stationary time series defined by

$$\gamma^{(p,q)}(k) := \mathbb{E}\left[\left(X_0^{(p)} - \mu\right)\left(X_k^{(q)} - \mu\right)^t\right],$$

for $p, q = 1, \ldots, d$ and $k = 0, 1, \ldots$.
Note that for $p = q$, we obtain $\gamma^{(p,p)}(\cdot)$ as auto-covariance function of the marginal univariate process $\left(X_j^{(p)}\right)_{j\in\mathbb{Z}}$ whose stationarity is implied in the one of the vector-valued process $(X_j)_{j\in\mathbb{Z}}$. Analogously to the univariate case it is possible to norm the cross-covariance function and get the cross-correlation function

$$r^{(p,q)}(k) := \frac{\gamma^{(p,q)}(k)}{\left(\gamma^{(p,p)}(0)\gamma^{(p,q)}(0)\right)^{\frac{1}{2}}}, \quad k \in \mathbb{Z}.$$

These introduced functions exhibit the following properties:

(i) $r^{(p,q)}(k) = r^{(q,p)}(-k)$,

(ii) $r^{(p,p)}(0) = 1$,

(iii) $\left|r^{(p,q)}(k)\right| \leq 1$, $p, q = 1, \ldots, d$.

The correlation $r^{(p,q)}(0)$ is the correlation between $X_j^{(p)}$ and $X_j^{(q)}$ which is generally not equal to 1 if $p \neq q$. It is also possible that $\left|r^{(p,q)}(k)\right| > \left|r^{(p,q)}(0)\right|$ if $p \neq q$.

We follow the definition in [37] and define multivariate long-range dependence analogously to the univariate case.

Definition 2.13 *(Multivariate long-range dependence, [37], Definition 2.1)*
A multivariate stationary time series $(X_j)_{j\in\mathbb{Z}}$ is long-range dependent if its autovariance matrix function $\Gamma_X(k)$ satisfies

$$\Gamma_X(k) = k^{D-\frac{1}{2}I_d} L(k) k^{D-\frac{1}{2}I_d} = \left(L_{p,q}(k) k^{d_p+d_q-1}\right)_{p,q=1,\dots,d},$$

where $k^{D-\frac{1}{2}I_d} := \operatorname{diag}\left(k^{d_1-1/2}, \dots, k^{d_d-1/2}\right)$ for a diagonal matrix $D = \operatorname{diag}(d_1, \dots, d_d)$ with $d_p \in \left(0, \frac{1}{2}\right)$, $p = 1, \dots, d$. Furthermore, $L(k) = (L_{p,q}(k))_{p,q=1,\dots,d}$ is an $\mathbb{R}^{d\times d}$-valued function satisfying

$$L(k) \sim L = (L_{p,q})_{p,q=1,\dots,d}, \quad (k \to \infty),$$

where $L_{p,q} \in \mathbb{R}$ and $L_{p,p} \neq 0$, $p = 1, \dots, d$.

As well, a multivariate extension for short-range dependence is given in their paper.

Definition 2.14 *(Multivariate short-range dependence, [37], Remark 2.3)*
A multivariate stationary time series is short-range dependent if its autovariance matrix function satisfies

$$\sum_{n=-\infty}^{\infty} \|\Gamma_X(n)\|_F < \infty,$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

We have introduced the mathematical description of particular dependence structures of multivariate stationary stochastic processes without any further assumptions on the finite dimensional marginal distributions of the stochastic processes. In the next section, we specify the class of stationary multivariate processes under consideration and turn to the well-known class of Gaussian processes.

2.3 Gaussian processes

In this section Gaussian processes are introduced: the univariate case as well as its multivariate extension. We shed light on essential properties and several useful relations, as well as on popular examples of Gaussian processes. One focus here is the connection of the previous definitions of long-range dependence and self-similarity to Gaussian processes. As this class of processes relies on the multivariate Gaussian distribution that happens to be uniquely determined by the first two moments, long-range dependence as a condition on the covariance function has a strong impact here.

We start with the essential distribution needed to be able to classify the class of Gaussian processes, namely the multivariate normal distribution.

Definition 2.15 *(Multivariate normal distribution, [16], Definition 1.6.1, Rem. 4, Prop. 1.6.4)*
The random vector $Y = \left(Y^{(1)}, \dots, Y^{(d)}\right)^t$ is said to be multivariate normal, or to follow a

multivariate normal distribution, if and only if there exist a column vector $\mu \in \mathbb{R}^d$*, a matrix* $B \in \mathbb{R}^{d\times d}$ *and a random vector* $U = \left(U^{(1)}, \dots, U^{(d)}\right)^t$ *with independent standard normal components, such that*

$$Y = \mu + BU.$$

We write $Y \sim \mathcal{N}(\mu, \Sigma_Y)$*, with* $\Sigma_Y = BB^t$*. If* $\mu = (0, \dots, 0)^t$*, the random vector* Y *is called centered.*
If $\det(\Sigma_Y) > 0$ *then* Y *has the density*

$$f_Y(y) = (2\pi)^{-\frac{d}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(y-\mu)^t \Sigma_Y^{-1}(y-\mu)\right].$$

Two density functions of a bivariate Gaussian distribution are shown in Figure 2.1.

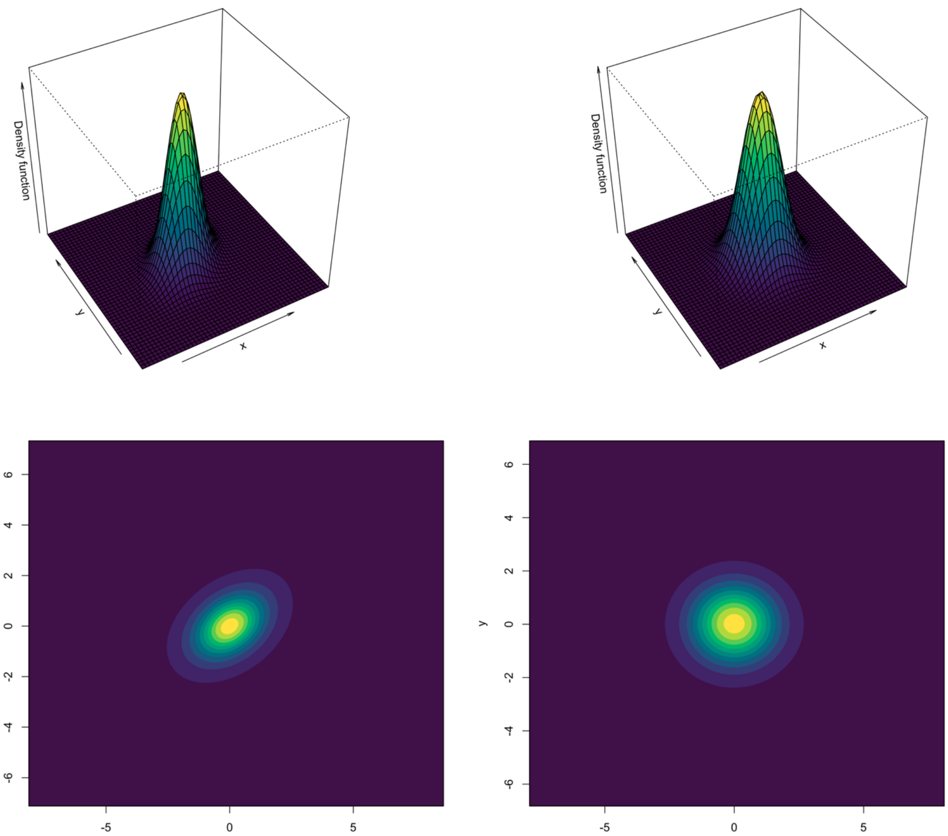

Figure 2.1: Plots of the pdf and the corresponding contour lines of a bivariate Gaussian distribution for Cor $\left(Y^{(1)}, Y^{(2)}\right) = 0.8$ (left) and Cor $\left(Y^{(1)}, Y^{(2)}\right) = 0$ (right), created by using the R-packages "ks" and "viridis" and [26], Sec. 3.1.

If $Y \sim \mathcal{N}(\mu, \Sigma_Y)$ it holds that

$$(i) \quad \left(Y^{(1)}, \dots, Y^{(d)}\right)^t \overset{\mathcal{D}}{=} \left(-Y^{(1)}, \dots, -Y^{(d)}\right)^t \tag{2.2}$$

$$(ii) \quad \left(Y^{(1)}, \dots, Y^{(d)}\right)^t \overset{\mathcal{D}}{=} \left(Y^{(d)}, \dots, Y^{(1)}\right)^t. \tag{2.3}$$

We denote the symmetry property in (2.2) as *space symmetry* and in (2.3) as *time symmetry*.
We are now fully equipped to introduce a class of stochastic processes that play an outstanding role in the course of this thesis.

Definition 2.16 *(Gaussian process, [16], Definition 1.3.4)*
The process $(X_j)_{j\in\mathbb{Z}}$ *is called a Gaussian process if and only if the finite-dimensional distributions of* $(X_j)_{j\in\mathbb{Z}}$ *are all multivariate normal. If* $\mathbb{E}(X_j) = 0$ *for all* $j \in \mathbb{Z}$ *the Gaussian process is called centered.*

The next remark highlights a property of Gaussian processes in the context of stationarity.

Remark 2.17 *([16], p.13)*
If $(X_j)_{j\in\mathbb{Z}}$ *is a stationary Gaussian process then* $(X_j)_{j\in\mathbb{Z}}$ *is strictly stationary, since for all* $k \in \mathbb{N}$ *and all* $h, j_1, j_2, \ldots \in \mathbb{Z}$*, the random vectors* $(X_{j_1}, \ldots, X_{j_k})^t$ *and* $(X_{j_1+h}, \ldots, X_{j_k+h})^t$ *have the same mean and covariance matrix, and hence the same distribution.*

A very well-known and common example of a Gaussian process is given by Brownian motion:

Definition 2.18 *(Standard Brownian motion, [16], Definition 1.7.1)*
Standard Brownian motion is a stochastic process $(B(t))_{t\in\mathbb{R}}$ *satisfying the conditions*

$$
\begin{aligned}
&(i) \quad B(0) = 0 \\
&(ii) \quad B(t_2) - B(t_1), B(t_3) - B(t_2), \ldots, B(t_n) - B(t_{n-1}) \ \textit{are independent} \\
&\qquad \textit{for every } n \in \{4, 5, \ldots\} \textit{ and every } t = (t_1, \ldots, t_n) \textit{ such that } 0 < t_1 < t_2 < \ldots < t_n, \\
&(iii) \quad B(t) - B(s) \sim \mathcal{N}(0, t-s) \textit{ for } t \geq s,
\end{aligned}
$$

with $t, t_1, \ldots, t_n, s \in \mathbb{R}$.

The corresponding increment process is the archetype of independent Gaussian time series. It is called Gaussian white noise due to the Gaussian innovations.

In the context of short- as well as long-range dependence it is possible to give an extension to *fractional* Brownian motion, that yields the possibility to represent these particular dependence structures.

Definition 2.19 *(Fractional Brownian motion, [9], Definition 3.23)*
A Gaussian stochastic process $B_H(u)$*,* $(u \in \mathbb{R})$ *with mean zero is called a fractional Brownian motion with Hurst parameter* $H \in (0,1)$ *if its covariance function is given by*

$$\gamma_H(u,v) = Cov(B_H(u), B_H(v)) = \frac{\sigma^2}{2}\left[|u|^{2H} + v^{2H} - |u-v|^{2H}\right], \quad (u, v \in \mathbb{R}).$$

Note that fractional Brownian motion is the only Gaussian process that is H-self-similar and has stationary increments, see [54], Definition 2.6.2. We take a closer look on the increment process.

Definition 2.20 *(Fractional Gaussian noise, [9], Example 2.2)*
Let $B_H(u)$ $(u \in \mathbb{R})$ *be a fractional Brownian motion. The corresponding increment process* $W_{H,t} = B_H(t) - B_H(t-1)$*,* $t \in \mathbb{Z}$ *is called fractional Gaussian noise with the explicit autocovariance structure*

$$\gamma_H(k) = \frac{\sigma^2}{2}\left(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}\right).$$

Note that for the autocovariance function of fractional Gaussian noise it holds that

$$\gamma_H(k) \simeq \sigma^2 H(2H-1)|k|^{2H-2} \quad (k \to \infty),$$

see [9], p. 34, where $a_n \simeq b_n \Leftrightarrow \lim_{n\to\infty} \frac{a_n}{b_n} = 1$.
We observe, that fractional Gaussian noise is long-range dependent if $H \in \left(\frac{1}{2}, 1\right)$. We have introduced univariate Gaussian processes, given popular examples and established a connection to long-range dependence.

As the title of this thesis already indicates, the main focus of this work is on multivariate or vector-valued Gaussian processes. Therefore, we give an extension of the univariate results to the vector-valued ones.

Definition 2.21 *(Vector-valued Gaussian process, [43], p.424)*
A vector-valued stochastic process $(X_j)_{j\in\mathbb{Z}}$ with $X_j = \left(X_j^{(1)}, \ldots, X_j^{(d)}\right)^t$ is called a (vector-valued) Gaussian process if all finite-dimensional marginal distributions are multivariate Gaussian. If $\mathbb{E}(X_j) = (0, \ldots, 0)^t$ for all $j \in \mathbb{Z}$ the (vector-valued) Gaussian process is called centered.

Note that this distribution does explicitly not imply that d one-dimensional Gaussian processes can be plugged together to a d-dimensional Gaussian process, since the conditions on d-dimensional vector-valued Gaussian processes are stronger.
Well-known examples of multivariate Gaussian processes can be found analogously to the examples in the univariate case introduced above. While the d-dimensional extension of the univariate Brownian motion is straightforward, see [43], Sec. C.8.2, the generalizations of fractional Brownian motions and fractional Gaussian noises to the multivariate case claim an extension of self-similarity to so-called operator self-similarity for the d-dimensional case first. We present the definition, for details we refer to [54], Sec. 9.3.

Definition 2.22 *(Vector operator fractional Brownian motion, [54], Definition 9.3.1)*
A vector-valued ($\mathbb{R}^d$-valued) process $(B_H(t))_{t\in\mathbb{R}}$ is called vector operator fractional Brownian motion if it is Gaussian, has stationary increments and is operator self-similar.
Stationarity of increments means that, for any $h \in \mathbb{R}$

$$(B_H(t+h) - B(h))_{t\in\mathbb{R}} \overset{\mathcal{D}}{=} (B_H(t) - B_H(0))_{t\in\mathbb{R}}.$$

For the definition of operator self-similarity, we shall take: for any $c > 0$,

$$(B_H(ct))_{t\in\mathbb{R}} \overset{\mathcal{D}}{=} \left(c^H B_H(t)\right)_{t\in\mathbb{R}},$$

with $H \in \mathbb{R}^{d\times d}$ and $c^H := \mathrm{e}^{(\log c)H} := \sum_{k=0}^{\infty} \frac{(\log c)^k H^k}{k!}$.

In the next result, we restrict ourselves to the case that self-similarity matrix $H = \operatorname{diag}(H_1, \ldots, H_d)$ is a diagonal-matrix, which yields $c^H = \operatorname{diag}\left(c^{H_1}, \ldots, c^{H_d}\right)$, such that the entries H_k of H satisfy $0 < \Re(H_k) < 1$, $k = 1, \ldots, d$, see [54], p.472 and p.476. Note that $\Re(z)$ denotes the real part of $z \in \mathbb{C}$. Following [54], Sec. 9.4.3, we extend the univariate definition of fractional Gaussian noise as given in Definition 2.20 to the vector-valued case.

Definition 2.23 *(Vector fractional Gaussian noise, [54], p.503)*
We define a vector fractional Gaussian noise series as

$$W_{H,n} = B_H(n) - B_H(n-1), \quad n \in \mathbb{Z}.$$

The series $(W_{H,n})_{n\in\mathbb{Z}}$ is Gaussian, zero mean and stationary. For the autocovariance matrix function of $(W_{H,n})_{n\in\mathbb{Z}}$, one can show

$$\gamma_W^{(p,q)}(k) = \frac{\sigma_p\sigma_q}{2}\left(\rho_{pq} - \text{sign}(k)\eta_{pq}\right)\left[(k+1)|k+1|^{H_p+H_q} + (k-1)|k-1|^{H_p+H_q} - 2|k|^{H_p+H_q}\right],$$

with $\sigma_p^2 := \mathbb{E}\left(B_H^{(p)}(1)^2\right)$, $\rho_{pq} = \text{Cor}\left(B_H^{(p)}(1)B_H^{(q)}(1)\right)$ and

$$\sigma_p\sigma_q\eta_{pq} = 4C_2\left(\frac{H_p+H_q}{2}\right)\Im\left(c_{pq}\right),$$

where $C_2(H) = \frac{\Gamma(2-2H)\sin(H\pi)}{2H(1-2H)}$. Furthermore, $\Im(z)$ denotes the imaginary part of $z \in \mathbb{C}$. Moreover, $C = AA^\star$ and A denotes the matrix given in the spectral representation of $B_H(t)$ in Section 2.7.2 and A^* the corresponding Hermitian transpose, for details, see [54], Theorem 9.3.2. It is possible, to obtain an asymptotic result for the behaviour of the covariance function analogously to the univariate case:

$$\gamma_W^{(p,q)}(k) \simeq \frac{\sigma_p\sigma_q}{2}\left(\rho_{pq} - \text{sign}(k)\eta_{pq}\right)\left(H_p+H_q\right)\left(H_p+H_q+1\right)|k|^{H_p+H_q-2}, \quad (k\to\infty), \tag{2.4}$$

if $H_p + H_q \neq 1$ and $p, q = 1, \ldots, d$.

Note that the class of stationary Gaussian processes is included in the class of linear processes, more specifically a stationary Gaussian process is a linear process that has standard normally distributed innovations, see [9], p. 216. However, the class of linear processes is much more general, see [9], Sec. 4.2.4. The corresponding vector-valued result is given in [43], p.424, where the innovations correspond to a d-dimensional vector-valued white noise process with each vector consisting of i.i.d standard normally distributed entries.
We follow [37] and introduce the commonly used multivariate long-range dependence condition of linear processes. However, we restrict ourselves to the case of Gaussian innovations.
A multivariate linear process $(X_n)_{n\in\mathbb{Z}}$ with finite variance has the representation

$$X_n = \sum_{k=0}^{\infty} \Psi_k \epsilon_{n-k}, \tag{2.5}$$

where $\left(\Psi_k = (\psi_{pq,k})_{p,q=1,\ldots,d}\right)_{k\in\mathbb{N}_0}$ are $d\times d$-matrices such that $\|\Psi_k\|_F < \infty$ and $(\epsilon_n)_{n\in\mathbb{Z}}$ is a d-variate white noise, such that $\epsilon_n \sim \mathcal{N}(0, I_d)$. The representation in (2.5) is called causal (one-sided) representation. For details on multivariate linear processes, see [54], Sec. 3 and Sec. 4.

We specify two cases concerning the dependence structure of this class of processes.
The first one is given by the sequence Ψ_k of real-valued $d\times d$ matrices fulfills for all $p, q = 1, \ldots d$ that

$$\psi_{pq,k} = G_{pq}(k)|k|^{d_p-1}, \quad k \in \mathbb{N}, \tag{2.6}$$

where $d_p \in \left(0, \frac{1}{2}\right)$, $p = 1, \ldots, d$ and $G(k) = \left(G_{pq}(k)\right)_{p,q=1,\ldots,d}$ is an $\mathbb{R}^{d \times d}$-valued function satisfying

$$G(k) \simeq A^+, \quad \text{as } k \to \infty,$$

for a $d \times d$ real-valued matrix A^+ that is assumed to have full rank.
The second one is given by

$$d_p < 0, \quad p = 1, \ldots, d, \qquad \sup_{k \in \mathbb{N}} |G_{pq}(k)| \leq \beta, \tag{2.7}$$

for some constant $\beta > 0$. Additionally the matrix $\sum_{k \in \mathbb{N}_0} \left(\psi_{pq,k}\right)_{p,q=1,\ldots,d} \in \mathbb{R}^{d \times d}$ is assumed to have full rank.
The next two results incorporate these two assumptions into the multivariate setting of long- and short-range dependence.

Proposition 2.24 *([54], Proposition 3.1)*
Let $(X_n)_{n \in \mathbb{Z}}$ be a d-dimensional Gaussian process given in (2.5) and let (2.6) hold. Then, $(X_n)_{n \in \mathbb{Z}}$ is multivariate long-range dependent in the sense of Definition 2.13.

Proposition 2.25 *([23], Proposition 2.2)*
Let $(X_n)_{n \in \mathbb{Z}}$ be a d-dimensional Gaussian process given in (2.5) and let (2.7) hold. Then, $(X_n)_{n \in \mathbb{Z}}$ is multivariate short-range dependent in the sense of Definition 2.14.

There is a direct connection of $L_{p,q}$, $p, q = 1, \ldots, d$ in Definition 2.13 to the matrix A^+, see [54], p.8 (37), (38), that fully specifies the asymptotics of the autocovariance matrix and, therefore, the long-range dependence condition as given in Definition 2.13. The existence of these matrices A^+ and A^- for fixed d_j, $j = 1, \ldots, d$ is further described in [54], p. 9. An example is provided in [54], Section 5, where they deal with multivariate FARIMA $(0, d, 0)$ processes. However, it is important to note, that the class of multivariate stationary Gaussian processes that fulfill (2.6) is only a subset of multivariate long-range dependent Gaussian processes in the sense of Definition 2.13.

2.4 Framework of Hermite polynomials

Hermite polynomials play an essential part in determining the limit distribution of functionals of Gaussian processes. This is due to the fact that suitably normalized Hermite polynomials form an orthonormal basis of $L_2(\mathbb{R}, \varphi)$, with φ being the density of the standard normal distribution. We recall the definition of this Hilbert space:

$$L_2(\mathbb{R}, \varphi) := \left\{ f : \mathbb{R} \to \mathbb{R}, \int_{\mathbb{R}} f^2(x)\varphi(x)\mathrm{d}x < \infty \right\}.$$

Therefore, it is possible to expand functions of Gaussian random variables belonging to the Hilbert space above as Fourier series using Hermite polynomials as orthonormal basis. For details on basics of Fourier transformations and Fourier series, see [61]. In this section we take a closer look at these relations and describe it precisely.

Definition 2.26 *(Hermite polynomial, [9], Definition 3.1)*
The j-th Hermite polynomial $H_j(x)$, $j = 0, 1, \ldots$, is defined as

$$H_j(x) := (-1)^j \exp\left(\frac{x^2}{2}\right) \frac{\mathrm{d}^j}{\mathrm{d}x^j} \exp\left(-\frac{x^2}{2}\right).$$

To get an impression, we give the first four Hermite polynomials:
$H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - x$.

The orthogonality of Hermite polynomials is given alongside of Sec. 3.1.2 in [9]:

$$\langle H_j, H_k \rangle = \int_{-\infty}^{\infty} H_j(x) H_k(x) \varphi(x) \mathrm{d}x = \delta_{jk} j!.$$

Therefore, what remains to show for the result that the family of Hermite polynomials is an orthogonal basis of $L^2(\mathbb{R}, \varphi)$ is completeness, which is done in [9] in the proof of Lemma 3.3. Therefore, we can state:

Lemma 2.27 *(Orthogonal basis, [9], Lemma 3.3)*
The family of Hermite polynomials $\{H_j,\ j = 0, 1, \ldots\}$ forms an orthogonal basis in $L^2(\mathbb{R}, \varphi)$.

Equipped with this property, we can use that for a function $f \in L^2(\mathbb{R}, \varphi)$ and a standard normal random variable U, we obtain a (in the L^2-sense) unique decomposition in terms of Hermite polynomials:

Lemma 2.28 *(Hermite decomposition, [9], Lemma 3.4)*
Let $U \sim \mathcal{N}(0, 1)$, and let f be such that $\mathbb{E}(f(U) = 0)$ and $\mathbb{E}(f^2(U)) < \infty$. Then, $f(U)$ has the (L^2-) unique representation

$$f(U) = \sum_{k=1}^{\infty} g_k H_k(U) = \sum_{k=1}^{\infty} \frac{c_k}{k!} H_k(U),$$

with Hermite coefficients $g_k = \frac{c_k}{k!}$ and $c_k = \langle f, H_k \rangle = \mathbb{E}(f(U) H_k(U))$. In the course of this thesis, we refer to c_k as Hermite coefficients.

A crucial parameter in the context of limit theorems for long-range dependent time series is the lowest value of k with a nonzero Hermite coefficient, which is called *Hermite rank*:

Definition 2.29 *(Hermite rank, [9], Definition 3.2)*
Let U be a standard normal random variable, and f be a function such that $\mathbb{E}(f(U)) = 0$ and $\mathbb{E}(f^2(U)) < \infty$. Then, the Hermite rank m of f is the smallest integer $k \geq 1$, such that

$$c_k = \mathbb{E}(f(U) H_k(U)) \neq 0.$$

The corresponding result that emphasizes the importance of the Hermite rank in the asymptotics of functionals of stationary Gaussian time series, is given in the well-known *Reduction Theorem*, originally proven in [64], Theorem 4.1.

Theorem 2.30 *(Reduction Theorem, [54], Theorem 5.3.3.)*
Let $(X_n)_{n\in\mathbb{Z}}$ be a Gaussian stationary time series which is long-range dependent in the sense of Definition 2.6. Suppose that $\mathbb{E}(X_n) = 0$ and $\mathbb{E}(X_n^2) = 1$. Let f be a function with Hermite rank $m \geq 1$ and let c_m be the first nonzero coefficient in the Hermite expansion. H_m is the Hermite polynomial of order m. Then, the asymptotic distribution of suitably normalized $f(x) = c_m H_m(x)$ is the same as for the general f with Hermite rank m.

In the course of this thesis, we are dealing with multivariate Gaussian processes and, therefore, the univariate consideration of Hermite polynomials forming an orthonormal basis is not sufficient for us. Fortunately, it is possible to extend the ideas presented above to the multivariate case, i.e., considering Gaussian random vectors $X = \left(X^{(1)}, \ldots, X^{(d)}\right)^t \sim \mathcal{N}(0, \Sigma_d)$, where Σ_d denotes the corresponding covariance matrix. It turn out that for our aim we can restrict ourselves to the less complicated case of Gaussian random vectors with independent entries, such that $\Sigma_d = I_d$. The results of the following paragraph rely on Sec. 3.2 in [9]. Note that on pages 119-121 they introduce multivariate Hermite polynomials in terms of covariance matrices Σ_d of any order, but close these considerations with the statement that general multivariate Hermite polynomials are too complicated to be suitable in the context of limit theorems. Fortunately, to our aim we can focus on the case $\Sigma_d = I_d$, which is given in details in the following.
We are dealing with a random vector $X \sim \mathcal{N}(0, I_d)$, which has independent $\mathcal{N}(0,1)$ entries.

Definition 2.31 *(Multivariate Hermite polynomial, [9], p. 122)*
Let $d \in \mathbb{N}$. We define as d-dimensional Hermite polynomial

$$H_k(x) := H_{k_1,\ldots,k_d}(x) := H_{k_1,\ldots,k_d}(x_1, \ldots, x_d) = \prod_{j=1}^{d} H_{k_j}(x_j),$$

with $k = (k_1, \ldots, k_d) \in \mathbb{N}_0^d \setminus \{(0, \ldots, 0)\}$.

Remarkt that the case $k = (0, \ldots, 0)$ is excluded here due to the assumption $\mathbb{E}(f(X)) = 0$. Analogously to the univariate case, the family of multivariate Hermite polynomials $\{H_{k_1,\ldots,k_d},\ k_1, \ldots, k_d \in \mathbb{N}\}$ forms an orthogonal basis of $L^2\left(\mathbb{R}^d, \varphi_{I_d}\right)$, which is defined as

$$L^2\left(\mathbb{R}^d, \varphi_{I_d}\right) := \left\{f : \mathbb{R}^d \to \mathbb{R},\ \int_{\mathbb{R}^d} f^2(x_1, \ldots, x_d)\, \varphi(x_1) \ldots \varphi(x_d)\, \mathrm{d}x_d \ldots \mathrm{d}x_1 < \infty\right\}.$$

The parameter φ_{I_d} denotes the density of the d-dimensional standard normal distribution, which is already divided into the product of the univariate densities φ in the formula above.
We denote the Hermite coefficients by

$$C(f, X, k) := C(f, I_d, k) := \langle f, H_k \rangle = \mathbb{E}(f(X) H_k(X)).$$

The Hermite rank $m(f, I_d)$ of f with respect to the distribution $\mathcal{N}(0, I_d)$ is defined as the largest integer m, such that

$$\mathbb{E}\left(f(X) \prod_{j=1}^{d} H_{k_j}\left(X^{(j)}\right)\right) = 0 \text{ for all } 0 < k_1 + \ldots k_d < m.$$

Having these preparatory results in mind, we derive a the multivariate Hermite expansion given by

$$f(X) - \mathbb{E}f(X) = \sum_{k_1+\ldots+k_d \geq m(f,I_d)} \frac{C(f,X,k)}{k_1!\ldots k_d!} \prod_{j=1}^{d} H_{k_j}\left(X^{(j)}\right). \tag{2.8}$$

A multivariate extension of the Reduction Theorem 2.30 is proved in [2], Theorem 6, for the case that $\Sigma_d = I_d$.

Finally following [9], Sec. 3.2 it is necessary to determine the Hermite rank of a function f with respect to the original random vector $X \sim \mathcal{N}(0, \Sigma_d)$ in order to get the right asymptotic result. We denote this Hermite rank by $m(f, \Sigma_d)$ here. We give a first result dealing with the relation between $m(f, I_d)$ and $m(f, \Sigma_d)$.

Lemma 2.32 *([9], Lemma 3.7)*
Let $X \sim \mathcal{N}(0, \Sigma_d)$ and $\Sigma_d = AA^t$, such that $\tilde{X} = A^{-1}X \sim \mathcal{N}(0, I_d)$. Then, it holds that $m(f, \Sigma_d) = m(f \circ A, I_d)$.

In [10], Chapter 5, two results are proved that deal with a simplification of the determination of the Hermite rank above. We cite the latter one here, which prove itself to be very useful and is used at various occasions in the course of this thesis.

Lemma 2.33 *[10], Lemma 5.4)*
Let $f : \mathbb{R}^d \to \mathbb{R}$ be square-integrable with respect to $\mathcal{N}(0, I_d)$ and let $\Sigma_d = AA^t$ be a $d \times d$ positive definite covariance matrix such that $\Sigma_d^{-1} - I_d$ is positive semidefinite. Then

$$m(f \circ A, I_d) \leq m(f, I_d).$$

An important note here is that the additional assumption that $\Sigma_d^{-1} - I_d$ is positive semidefinite is a technical restriction to the covariance matrix. However, since we deal with ordinal patterns in this work, we anticipate one very advantageous property of ordinal patterns in general here, namely that they are not affected by monotone transformations. Therefore, we are always in the comfortable situation to scale the covariance matrix of the considered random vectors, such that the technical condition above is fulfilled. For further details, see [10], Remark 5.5.

One intriguing question that explains the crucial role of the Hermite rank in the context of asymptotics of *long-range dependent* Gaussian time series is whether a long-range dependent process is still long-range dependent after being transformed by f. The following proposition yields the answer.

Proposition 2.34 *([54], Proposition 5.2.4)*
Let $(X_j)_{j \in \mathbb{Z}}$ be a stationary Gaussian time series with $\mathbb{E}(X_j) = 0$ and $\mathbb{E}\left(X_j^2\right) = 1$, which is long-range dependent in the sense of Definition 2.6. Let $f \in L^2(\mathbb{R}, \varphi)$ be a function with Hermite rank m. Then, for $k \to \infty$,

$$Cov(f(X_k), f(X_0)) \sim g_m^2 m! (L_2(k))^m k^{(2d_1-1)m}.$$

Table 2.1 provides an overview which relates the long-range dependence parameter d_1 to the Hermite rank $m(f, \Sigma_d)$ and as well provides the connection between d_1 and the long-range dependence parameter d_f of the transformed process. Finally, we introduce an essential lemma

$m(f, \Sigma_d)$	d_1	d_f	Range of d_f
1	$\left(0, \frac{1}{2}\right)$	d_1	$\left(0, \frac{1}{2}\right)$
2	$\left(\frac{1}{4}, \frac{1}{2}\right)$	$2d_1 - \frac{1}{2}$	$\left(0, \frac{1}{2}\right)$
3	$\left(\frac{1}{3}, \frac{1}{2}\right)$	$3d_1 - 1$	$\left(0, \frac{1}{2}\right)$

Table 2.1: Range of d_1 which ensures that $(f(X_j))$ is still long-range dependent and the corresponding value of d_f, [54], Table 5.1.

for investigations of dependence structures of limiting processes and random variables in the course of this work.

Lemma 2.35 *([9], Lemma 3.5)*
For a pair of jointly standard normal random variables $\left(U^{(1)}, U^{(2)}\right)^t$ *with covariance* $\rho = Cov\left(U^{(1)}, U^{(2)}\right)$*, we have*

$$Cov\left(H_m\left(U^{(1)}\right) H_m\left(U^{(2)}\right)\right) = m!\rho^m, \quad m \in \mathbb{N},$$

whereas for $j \neq k$

$$Cov\left(H_j\left(U^{(1)}\right) H_k\left(U^{(2)}\right)\right) = 0.$$

2.5 Integrals with respect to random measures

In this section, we do an excursus in the field of integrals with respect to random measures. We discuss the fundamentals as construction and essential properties. Furthermore, we connect this approach to our framework of limit theorems as it appears that the limiting processes arising in the context of Gaussian long-range dependence are expressed as multiple Wiener-Itô integrals. For an overview of random measures, we refer to [35].

2.5.1 Single integrals with respect to random measures

In this paragraph, the results presented are derived from [54], Section B.1. We begin with a general description of the setting: Let $(E, \mathcal{E})$ be a measurable space, that is, E is a set and $\mathcal{E}$ is a σ-field of subsets of E. Single integrals with respect to random measures are written as

$$I(f) := \int_E f(x) M(\mathrm{d}x).$$

Here, f is a deterministic, possibly complex-valued measurable function on E. Furthermore, M is a suitable, possibly complex-valued random measure on $(E, \mathcal{E})$ that is defined with respect to a deterministic measure m on $(E, \mathcal{E})$, called a *control measure* of M. The role of the control measure will be specified in the following considerations. We denote $\mathcal{E}_0 := \{A \in \mathcal{E} : m(A) < \infty\}$.

Single integrals with respect to random measures are defined in the following way, similarly to the general construction of integrals with respect to measures in the well-known measure and integration theory, see [7].
A function f is called *simple*, if it can be written as

$$f(x) = \sum_{k=1}^{n} a_k \mathbf{1}_{A_k}(x),$$

where $a_k \in \mathbb{C}$ and $A_k \in \mathcal{E}_0$, $k = 1, \ldots, n$ are pairwise disjoint. For simple functions f the integral $I(f)$ is defined as

$$I(f) = \sum_{k=1}^{n} a_k M\left(A_k\right). \tag{2.9}$$

We continue with the construction of single integrals with respect to random measures by restricting ourselves to integrals with respect to random measures with orthogonal increments, as they are the random measures of interest in the course of this work.

Definition 2.36 *(Random measure with orthogonal increments, [54], Definition B.1.1)*
A set function $Z : \mathcal{E}_0 \to L^2(\Omega)$ with control measure m is called a random measure with orthogonal increments on $(E, \mathcal{E})$, if Z is σ-additive and

$$\mathbb{E}Z(A) = 0,$$
$$\mathbb{E}\left(Z\left(A_1\right)\overline{Z\left(A_2\right)}\right) = m\left(A_1 \cap A_2\right).$$

Considering a simple function f, set $I(f)$ as in (2.9). For f, g both being simple functions, by using Definition 2.36, we obtain

$$\mathbb{E}\left(I(f)\overline{I(g)}\right) = \int_E f(x)\overline{g(x)}m(\mathrm{d}x) = \langle f, g\rangle_{L^2(E,m)}.$$

We extend the definition of $I(f)$ to functions $f \in L^2(E, m)$ now. For $f \in L^2(E, m)$, there is a sequence of simple functions f_n, $n \geq 1$, such that $\|f - f_n\|_{L^2(E,m)} \to 0$, as $n \to \infty$. Then, f_n, $n \geq 1$ is also a Cauchy sequence in $L^2(E, m)$ and it is possible to show that $I\left(f_n\right)$, $n \geq 1$ is a Cauchy sequence in $L^2\left(\Omega\right)$, which implies that is has a limit in $L^2(\Omega)$ that is defined as the integral $I(f)$:

$$I(f) = \lim_{n\to\infty} \left(L^2(\Omega)\right) I\left(f_n\right).$$

Note that this limit is independent of the approximating sequence f_n. For details to the construction of single integrals with respect to random measures with orthogonal increments, see [54], Sec. B.1 and [45].
We introduce a particular case of random measures with orthogonal increments now.

Definition 2.37 *(Gaussian random measure, [54], Definition B.1.2)*
A random measure B on $(E, \mathcal{E})$ with orthogonal increments is called Gaussian if any vector $\left(B\left(A_1\right), \ldots, B\left(A_n\right)\right)$, $A_k \in \mathcal{E}$, $k = 1, \ldots, n$, is multivariate Gaussian.

Analogously to general random measures with orthogonal increments, the integral $I(f)$ with respect to Gaussian measure B is defined for all $f \in L^2(E,m)$. An additional property of these integrals is that any vector $(I(f_1), \ldots, I(f_n))$, $f_k \in L^2(E,m)$ $k = 1, \ldots, n$ is also multivariate Gaussian.

Concerning complex-valued measures, we introduce Hermitian Gaussian random measures. Note that m is a symmetric measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ in the sense that

$$m(A) = m(-A), \text{ for } A \in \mathcal{B}(\mathbb{R}),$$

where

$$-A = \{x \in \mathbb{R} : (-x) \in A\}.$$

Definition 2.38 *(Hermitian Gaussian random measure, [54], Definition B.1.3)*
An Hermitian Gaussian random measure $\tilde{B}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a complex-valued Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with a symmetric control measure m, such that

$$\overline{\tilde{B}(A)} = \tilde{B}(-A), \quad A \in \mathcal{B}(\mathbb{R}).$$

The symmetry of the control measure is often written as $m(\mathrm{d}x) = m(-\mathrm{d}x)$ and the condition in Definition 2.38 as $\overline{\tilde{B}(\mathrm{d}x)} = \tilde{B}(-\mathrm{d}x)$. For a detailed discussion on the properties of Hermitian Gaussian random measure, we refer to [54], Sec. B.1.2.

There is an important connection between the real-valued Gaussian random measure B on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with Lebesgue control measure $\mathrm{d}u$ and the Hermitian Gaussian random measure $\tilde{B}$ with Lebesgue control measure $\mathrm{d}x$, see [54], p.591: Let $f \in L^2(\mathbb{R}, \mathrm{d}u)$ and let

$$\hat{f}(x) = \int_{\mathbb{R}} \mathrm{e}^{ixu} f(u) \mathrm{d}u$$

be the Fourier transform of f. For details on the construction of Fourier transforms for $f \in L^2(\mathbb{R}, \mathrm{d}u)$ we refer to [54], Sec. A.1.2. It holds that

$$\int_{\mathbb{R}} f(u) B(\mathrm{d}u) \overset{\mathcal{D}}{=} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(x) \tilde{B}(\mathrm{d}x). \tag{2.10}$$

The integral on the left-hand side is said to be in the *time domain*, the one on the right-hand side in the *spectral domain*.

2.5.2 Multiple Wiener-Itô integrals

In this paragraph, we introduce the definition of multiple integrals with respect to Gaussian measures. They are also known as multiple Wiener-Itô integrals. The results used here are given in [54], Sec. B.2.

Definition 2.39 *(Multiple Wiener-Itô integral, [54], (B.2.1))*
Let B be a real-valued Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with Lebesgue control measure $\mathrm{d}u$, then the multiple Wiener-Itô integral is written as

$$I_k(f) := \int_{\mathbb{R}^k}' f(u_1, \ldots, u_k)\, B(\mathrm{d}u_1) \ldots B(\mathrm{d}u_k),$$

where $f \in L^2\left(\mathbb{R}^k, \mathrm{d}u_1, \ldots, \mathrm{d}u_k\right)$. The prime in $\int_{\mathbb{R}^k}'$ denotes that the integration excludes the diagonals.

The construction of the multiple Wiener-Itô integral comes alongside with the construction of the single integral in the previous section, for details see [54], p. 597f. or [9], Sec. 3.7.1.2. Note however, that the simple function here is defined in such a way that it vanishes on the diagonal to assure that the integration excludes the diagonals. This is due to the fact, that otherwise, if $x_1 = x_2$, then, informally, $\mathbb{E}\left(B(\mathrm{d}x_1)\, B(\mathrm{d}x_1)\right) = \mathrm{d}x_1$ so that the diagonal would contribute $\int_{\mathbb{R}} g(x_1, x_1)\, \mathrm{d}x_1$ yielding a non-zero mean in general, see [9], p.193.

It is possible to represent the multiple Wiener-Itô integral in the spectral domain, for details on the construction see [54], Sec. B.2. Let $\tilde{B}$ be a Hermitian Gaussian random measure and $g \in L^2\left(\mathbb{R}^k, \mathrm{d}x_1, \ldots, \mathrm{d}x_k\right)$. Then, the multiple Wiener-Itô integral in the spectral domain is given by

$$\hat{I}_k(g) = \int_{\mathbb{R}^k}'' g(x_1, \ldots, x_k)\, \tilde{B}(\mathrm{d}x_1) \ldots \tilde{B}(\mathrm{d}x_k).$$

The double prime in $\int_{\mathbb{R}^k}''$ excludes the diagonals $|x_i| = |x_j|$, $i \neq j$ in the integration.

It is possible to derive an analogous result to (2.10) for multiple Wiener-Itô integrals, following [54], p. 600, (B.2.16): we denote by

$$\hat{f}(x_1, \ldots, x_k) = \int_{\mathbb{R}^k} \mathrm{e}^{i(x_1 u_1 + \ldots + x_k u_k)} f(u_1, \ldots, u_k)\, \mathrm{d}u_1 \ldots \mathrm{d}u_k$$

the Fourier transform of f. Then

$$I_k(f) \stackrel{\mathcal{D}}{=} \frac{1}{(2\pi)^{k/2}} \hat{I}_k\left(\hat{f}\right).$$

2.6 Spectral representations of stochastic processes

There exists a way to categorize stationary stochastic processes is the spectral integral representation. We begin with the spectral representation of univariate processes, then briefly turning to the multivariate case so we can apply this representation to Hermite-Rosenblatt processes in the next section. The results are given in [9], Section 4.1.3, following [16], Chapter 4.
Herglotz's Theorem ([16], Theorem 4.3.1) yield the cruicial relation between the autocovariance function $\gamma_X(k)$ of a zero-mean stationary stochastic process $(X_t)_{t\in\mathbb{Z}}$ and a so called spectral distribution function F:

$$\gamma_X(k) = \int_{-\pi}^{\pi} \mathrm{e}^{ik\lambda} F(\mathrm{d}\lambda), \quad k \in \mathbb{Z}.$$

The stationary time series $(X_t)_{t\in\mathbb{Z}}$ itself, has a spectral representation of the form

$$X_t(\omega) = \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} M_X(\mathrm{d}\lambda, \omega), \tag{2.11}$$

where $M_X(\cdot, \omega)$ is a spectral measure. For simplicity we write $M_X(\mathrm{d}\lambda)$ instead of $M_X(\mathrm{d}\lambda, \omega)$ in the following, but, as for all random measures, the dependence on $\omega \in \Omega$ should be beared in mind. The spectral measure is a complex-valued zero mean stochastic process on $(-\pi, \pi]$ with (a.s.) right-continuous sample paths and *uncorrelated* but not necessarily independent increments with a variance that is directly related to F. More specifically

$$\mathrm{Cov}\left(M_X(\mathrm{d}\lambda), M_X(\mathrm{d}v)\right) = 0, \quad (\lambda \neq v),$$
$$\mathrm{Var}(M_X(\mathrm{d}\lambda)) = F(\mathrm{d}\lambda).$$

In particular, if the spectral density f exists, (for details, see [16], Sec. 4.3), we can write $\mathrm{Var}(M_X(\mathrm{d}\lambda)) = f_X(\lambda)\mathrm{d}\lambda$. In this case, in the $L^2(\Omega)$-sense, $f_X(\lambda) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma_X(\lambda)\mathrm{e}^{-i\lambda k}$ and, therefore, the spectral density f_X and the covariance function γ_X are Fourier pairs as it is stated in the Wiener-Khinchin-Theorem, for details see [68].

The difference between the spectral distribution F and the spectral measure M_X in general, is that the first one only describes the linear dependence within the process, while the latter one provides full information in the sense of of the probability distribution of sample paths. It is possible to express a white noise process $(\varepsilon_t)_{t\in\mathbb{Z}}$ by using the spectral measure M_ε, such that $\mathrm{Var}\left((M_\varepsilon(\mathrm{d}\lambda)\right) = \frac{\sigma^2}{2\pi}\mathrm{d}\lambda$. White noise implies uncorrelated observations $(\varepsilon_t)_t$, however, in general, this does not imply the independence of $(\varepsilon_t)_t$.

A zero-mean, purely non-deterministic stationary process always has a Wold-decomposition, see [16], Theorem 5.7.1,

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} = A(B)\varepsilon_t, \quad (t \in \mathbb{Z}),$$

with white noise innovations ε_t and $A(z) = \sum a_j z^j$ such that $\sum_{j=0}^{\infty} a_j^2 < \infty$ and B denotes the backshift-operator such that $B^k X_t = X_{t-k}$. Combining the Wold-decomposition with the spectral representation in (2.11) we obtain

$$X_t = \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} A\left(\mathrm{e}^{-i\lambda}\right) M_\varepsilon(\mathrm{d}\lambda), \quad (t \in \mathbb{Z}),$$

or analogously

$$M_X(\mathrm{d}\lambda) = A\left(\mathrm{e}^{-i\lambda}\right) M_\varepsilon(\mathrm{d}\lambda).$$

First, note that in general when considering linear processes, the innovations ε_t are uncorrelated but not necessarily independent. We now want to apply the representations above to the special case, that M_ε and, therefore, M_X are Gaussian spectral measures. Gaussian processes are included in the class of linear processes, they exhibit independent innovations fulfilling $\varepsilon_t \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$. For simplicity we assume $\sigma_\varepsilon^2 = 1$. Then, it holds for a Gaussian process $(X_t)_t$,

$$X_t = \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} M_X(\mathrm{d}\lambda),$$

where the Gaussian random measure $M_X(\lambda)$ is defined by

$$M_X(\mathrm{d}\lambda) = A\left(\mathrm{e}^{-i\lambda}\right) M_\varepsilon(\mathrm{d}\lambda) =: \sqrt{2\pi}a(\lambda)M_\varepsilon(\mathrm{d}\lambda),$$

such that the spectral density can be written as $f_X(\lambda) = |a(\lambda)|^2$.
These univariate results can be extended analogously to the multivarate case and follow [16], Sec. 11.8. In the next paragraph we consider $(X_t)_{t\in\mathbb{Z}}$ as a d-dimensional stationary process.

Theorem 2.40 *([16], Theorem 11.8.1)*
$\Gamma_X(\cdot)$ is the covariance matrix function of a d-dimensional stationary process $(X_t)_{t\in\mathbb{Z}}$, if and only if

$$\Gamma_X(h) = \int_{(-\pi,\pi]} \mathrm{e}^{ih\lambda} F(\mathrm{d}\lambda), \quad h = 0, \pm 1, \pm 2, \dots,$$

where $F(\cdot)$ is a $d \times d$-matrix distribution function on $(-\pi, \pi]$. F is called the spectral distribution matrix of $(X_t)_{t\in\mathbb{Z}}$.

To turn to the spectral representation of $(X_t)_{t\in\mathbb{Z}}$, we have to consider the multivariate extension of the spectral measure to a vector-valued spectral measure M_X with orthogonal increments. For details, see [16], p.439. Equipped with these multivariate generalizations, we obtain the next result.

Theorem 2.41 *(Spectral Representation Theorem, [16], Theorem 11.8.2)*
If $(X_t)_{t\in\mathbb{Z}}$ is a stationary d-dimensional sequence with mean zero and spectral distribution matrix $F(\cdot)$, then there exists a right-continuous orthogonal increment random measure $M_X(\lambda)$, $\lambda \in (-\pi, \pi]$, such that

(i) $\mathbb{E}\left[(M_X(\lambda) - M_X(-\pi))(M_X(\lambda) - M_X(-\pi))^*\right] = F(\lambda), \quad -\pi \le \lambda \le \pi$

(ii) $X_t = \int_{(-\pi,\pi]} \mathrm{e}^{it\lambda} M_X(\mathrm{d}\lambda)$ *with probability* 1.

Multivariate Gaussian time series can be represented analogously to the univariate case, by using the linear representation as given in (2.5). Finally let us remark that there is a difference between the spectral representation of a stationary discrete stochastic process and integral representations of a time continuous process, as it appears in the next section. In order to point out the differences, we present Table 2.2, given in [54], Table 1.1. The time domain representation of the time continuous process is related to the spectral representation in (2.10).

Components in representation	Discrete process	Time continuous process
Dependent stochastic process	X_t	$X(t)$
Underlying space	$(-\pi, \pi]$	E
Deterministic functions	$\mathrm{e}^{it\lambda}$	$f_t(u)$
Uncorrelated (or independent) random measure	$M_X(\mathrm{d}\lambda)$	$Z(\mathrm{d}u)$

Table 2.2: Components in representations of discrete and time continuous stochastic processes.

2.7 Integral representations of Hermite-Rosenblatt processes

We discussed the basic definitions and properties of integrals with respect to random measures in Section 2.5. Furthermore, we introduced the spectral representation of stochastic processes. Therefore, we are now able to introduce the class of processes that arises in the limits of functionals of long-range dependent sequences considered in this work.
In the following, we provide the integral representation of Hermite-Rosenblatt processes in the time domain as well as in the spectral domain.

2.7.1 Univariate Hermite-Rosenblatt processes

To get a first intuitive impression, we begin with the relation between the Wiener-Itô integral as single integral with respect to a Gaussian random measure and standard Brownian motion, as it is given in [54], Example 4.1.5.
Take $E = \mathbb{R}$, $m(\mathrm{d}u) = \mathrm{d}u$ and $f_t(u) = \frac{\mathbf{1}_{[0,t)}(u)}{\sqrt{t}}$, then we obtain

$$I_1(f_t) = \int_{\mathbb{R}} f_t(u) B(\mathrm{d}u) = \frac{1}{\sqrt{t}} \int_0^t B(\mathrm{d}u) = \frac{B(t)}{\sqrt{t}},$$

where $B(x)$ is a Brownian motion.

We define the Hermite-Rosenblatt process of order m in the time domain.

Definition 2.42 *(Hermite-Rosenblatt process in the time domain, [9], Definition 3.24)*
Let $m \geq 1$ be an integer and $H \in \left(\frac{1}{2}, 1\right)$. Set $H_0 = 1 - \frac{1-H}{m} \in \left(1 - \frac{1}{2m}, 1\right)$, so that $H = 1 - m(1 - H_0)$. The Hermite-Rosenblatt process $(Z_{m,H_0}(t))_{t\in\mathbb{R}}$ of order m is defined as

$$Z_{m,H_0}(t) = \frac{K(m, H_0)}{m!} \int_{\mathbb{R}^m}' \left(\int_0^t \prod_{j=1}^m (s - u_j)_+^{H_0 - \frac{3}{2}} \mathrm{d}s \right) B(\mathrm{d}u_1) \dots B(\mathrm{d}u_m),$$

with $u_+ = \max\{0, u\}$. Furthermore, $\frac{K(m,H_0)}{m!}$ is a normalizing constant.

Note that H is the self-similarity parameter and the process is called standard, if
$\mathbb{E}\left(Z_{m,H_0}(1)^2\right) = 1$, which is fulfilled, if

$$K^2(m, H_0) = \frac{m! \,(2m(H_0 - 1) + 1)\,(m(H_0 - 1) + 1)}{\left(\int_0^\infty (x + x^2)^{H_0 - 3/2} \mathrm{d}x\right)^m}.$$

The Hermite process of order $m = 1$ is fractional Brownian motion as given in Definition 2.19, the Hermite process of order $m = 2$ is called Rosenblatt-process.
The connection to long-range dependent processes is given by the observation, that the range of H_0 coincides with the range of the long-range dependence parameter d_1 such that the transformed process by a function f with Hermite rank m is still long-range dependent, see Table 2.1, with $H_0 = d_1 + \frac{1}{2}$.
An interesting approach to verify the standard variance of the Hermite process above which is based on Lemma 2.35 and on Mc Kean's representation of a Hermite-Rosenblatt process in [47] is given in [9], p. 196f.

It is possible to express the Hermite-Rosenblatt process via an integral representation in the spectral domain.

Proposition 2.43 *(Hermite-Rosenblatt process in the spectral domain, [54], Proposition 4.2.4) The Hermite process $(Z_{m,H_0}(t))_{t\in\mathbb{R}}$ of order $m \geq 1$ can be represented as*

$$Z_{m,H_0}(t) = b_{m,H_0} \int''_{\mathbb{R}^m} \frac{\mathrm{e}^{it(x_1+\ldots+x_m)}-1}{i\,(x_1+\ldots+x_m)} \prod_{j=1}^{m} |x_j|^{\frac{1}{2}-H_0}\, \tilde{B}\,(\mathrm{d}x_1)\ldots\tilde{B}\,(\mathrm{d}x_m)\,,$$

where $\tilde{B}\,(\mathrm{d}x)$ is an Hermitian Gaussian random measure as defined in Definition 2.38 with control measure $\mathrm{d}x$. b_{m,H_0} is a normalizing constant, for details see [54], p. 234. Note that the double prime in the integral means that the integration disregards the hyperplanes $|x_i| = |x_j|$ for $i \neq j$. The process $(Z_{m,H_0}(t))_{t\in\mathbb{Z}}$ is called standard, if

$$b_{m,H_0} = \left(\frac{(m\,(H_0-1)+1)\,(2m\,(H_0-1)+1)}{m!\,\left[2\Gamma\,(2-2H_0)\sin\left(\left(H_0-\frac{1}{2}\right)\pi\right)\right]^m} \right)^{\frac{1}{2}}.$$

Hermite-Rosenblatt processes arise as limit processes of functionals of univariate long-range dependent Gaussian processes, the order m is depending on the Hermite rank of the function f used. Another parameter that highly determines the limiting process is whether the long-range dependence parameter d_1 of the underlying process that corresponds to $d_1 = H_0 - \frac{1}{2}$ is in the suitable range, such that the transformed process is still long-range dependent, see Table 2.1. However, in the course of this thesis we are not only interested in the limits of univariate long-range dependent Gaussian processes, but also in the limits of their multivariate extensions.

2.7.2 Matrix-valued Hermite-Rosenblatt processes

In this section, we turn to the integral representations of multivariate generalizations of Hermite-Rosenblatt processes, however, we restrict ourselves to the cases $m = 1$ and $m = 2$ here. The first extension we already met in this chapter is given in the operator fractional Brownian motion defined in Definition 2.22, which corresponds to $m = 1$. In the context of multivariate linear processes, operator fractional Brownian motions arise as multivariate extension of the univariate fractional Brownian motion as limit of the sample mean process. For details, see [23], Theorem 3.3.

We assume in the following, that $H \in \mathbb{R}^{d\times d}$ is a symmetric matrix and the eigenvalues H_k, $k = 1, \ldots d$, of H fulfill $0 < H_k < 1$, for all k. Then, the operator fractional Brownian motion in Definition 2.22 exhibits the spectral integral representation

$$B_H(t) = \int_{\mathbb{R}} \frac{\mathrm{e}^{itx}-1}{ix} \left(x_+^{-(H-(1/2)I_d)} A + x_-^{-(H-(1/2)I_d)} A^\star \right) \tilde{B}(\mathrm{d}x),$$

where A is a matrix with complex-valued entries, $A^\star$ denotes its complex conjugate, $x_+ = \max\{x, 0\}$, $x_- = \max\{-x, 0\}$ and $\tilde{B}(\mathrm{d}x)$ is a suitable multivariate complex-valued Gaussian measure, for details, see [21], Section 3.1.

We follow [23], Section 2 and introduce the time domain representation of operator fractional Brownian motion as defined in Definition 2.22 and the multivariate extension of the Rosenblatt process, that corresponds to the matrix-valued Hermite-Rosenblatt process for $m = 2$.
We continue with the integral representation of operator fractional Brownian motion in the time domain and we define a multivariate real-valued Gaussian random measure

$$B(\mathrm{d}x) = \left(B^{(1)}(\mathrm{d}x), \dots, B^{(d)}(\mathrm{d}x)\right)^t,$$

that satisfies

$$\mathbb{E}B(\mathrm{d}x) = 0,$$
$$\mathbb{E}B(\mathrm{d}x)(B(\mathrm{d}x))^t = I_d\ \mathrm{d}x$$

and

$$\mathbb{E}B^{(k)}(\mathrm{d}x)B^{(l)}(\mathrm{d}y) = 0,\ x \neq y, \quad k, l = 1, \dots, d. \tag{2.12}$$

Theorem 2.44 *(OFBM, time domain representation, [21], Theorem 3.2)*
If for the eigenvalues of H it holds that $H_k \neq \frac{1}{2}$, $k = 1, \dots, d$, the operator fractional Brownian motion in Definition 2.22 admits the integral representation

$$B_H(t) = \int_{\mathbb{R}} \Bigg(\left((t-u)_+^{H-(1/2)I_d} - (-u)_+^{H-(1/2)I_d}\right) M^+ \\ + \left((t-u)_-^{H-(1/2)I_d} - (-u)_-^{H-(1/2)I_d}\right) M^- \Bigg) B\left(d_u\right).$$

For the formulas of the matrices $M^+ \in \mathbb{R}^{d\times d}$ and $M^- \in \mathbb{R}^{d\times d}$, we refer to [21], Theorem 3.2.

The cross-covariance structure of this process is given in Theorem 2.2 in [41].

We continue with the matrix-valued representation of the Rosenblatt process, i.e., the Hermite-Rosenblatt process in the case $m = 2$. We follow the approach in [23], p. 5399, where this generalization is given in terms of double Wiener-Itô integrals. The R^{d^2}-valued process $(Z_{2,H}(t))_{t\in[0,1]}$ has the following representation in the time domain if the eigenvalues of the symmetric matrix H satisfy $H_p + H_q \in \left(0, \frac{1}{2}\right)$, $p, q = 1, \dots, d$:

$$Z_{2,H}(t) = \int_{\mathbb{R}^2}' f_{H,t}\left(x_1, x_2\right) \operatorname{vec}\left(B\left(\mathrm{d}x_1\right) B\left(\mathrm{d}x_2\right)^t\right), \tag{2.13}$$

where the vec operator transforms a matrix into a vector by stacking the columns of the matrix underneath each other. $f_{H,t}$ is given by $f_{H,t}\left(x_1, x_2\right) := f_{H,t,M^+,M^-}\left(x_1, x_2\right)$, with

$$f_{H,t,M^+,M^-}\left(x_1, x_2\right) = \sum_{s_1,s_2\in\{+,-\}} \int_0^t \left(\left(v-x_2\right)_{s_2}^{H-I_d} \otimes \left(v-x_1\right)_{s_1}^{H-I_d}\right) \left(M^{s_2} \otimes M^{s_1}\right) \mathrm{d}v, \tag{2.14}$$

where $\otimes$ denotes the Kronecker product, $M^+, M^- \in \mathbb{R}^{d\times d}$.
Further details on this process, as operator self-similarity and stationary increments, are given in Lemma 5.15 in [23]. On p. 5414 the covariance structure of this process is also provided.

We turn to the multivariate extension of a spectral random measure which is needed to construct the vector-valued Rosenblatt process. One approach based on [2], Theorem 6 is found in [9], Sec. 4.4.1.1. They proposed the case of independent univariate Hermitian Gaussian random measures. We define a multivariate Hermitian-Gaussian random measure $\tilde{B}(\mathrm{d}\lambda)$ with independent entries by

$$\tilde{B}(\mathrm{d}\lambda) = \left(\tilde{B}^{(1)}(\mathrm{d}\lambda), \dots, \tilde{B}^{(d)}(\mathrm{d}\lambda)\right)^t, \tag{2.15}$$

where $\tilde{B}^{(p)}(\mathrm{d}\lambda)$ is a univariate Hermitian-Gaussian random measure as defined in Definition 2.38. The multivariate Hermitian-Gaussian random measure $\tilde{B}(\mathrm{d}\lambda)$ satisfies

$$\begin{aligned} \mathbb{E}\left(\tilde{B}(\mathrm{d}\lambda)\right) &= 0, \\ \mathbb{E}\left(\tilde{B}(\mathrm{d}\lambda)\tilde{B}(\mathrm{d}\lambda)^*\right) &= I_d \ \mathrm{d}\lambda \end{aligned}$$

and

$$\mathbb{E}\left(\tilde{B}^{(p)}(\mathrm{d}\lambda_1)\overline{\tilde{B}^{(q)}(\mathrm{d}\lambda_2)}\right) = 0, \quad |\lambda_1| \neq |\lambda_2|, \quad p,q = 1,\dots,d,$$

where $\tilde{B}(\mathrm{d}\lambda)^* = \left(\overline{B^{(1)}(\mathrm{d}\lambda)}, \dots, \overline{B^{(d)}(\mathrm{d}\lambda)}\right)$ denotes the Hermitian transpose of $\tilde{B}(\mathrm{d}\lambda)$. Thus, following [2], Theorem 6, we can state the spectral representation of the matrix-valued Rosenblatt process $Z_{2,H}(t)$, $t \in [0,1]$ as

$$Z_{2,H}(t) = \left(Z_{2,H}^{(p,q)}(t)\right)_{p,q=1,\dots,d}$$

where each entry of the matrix is given by

$$Z_{2,H}^{(p,q)}(t) = \int_{\mathbb{R}^2}'' \frac{\exp\left(it\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)} \tilde{B}^{(p)}\left(\mathrm{d}\lambda_1\right) \tilde{B}^{(q)}\left(\mathrm{d}\lambda_2\right).$$

2.7.3 Rosenblatt distribution

As we see in the course of this chapter, the Rosenblatt distribution plays a crucial role in the asymptotics of functionals of stationary long-range dependent time series. Note that using Definition 2.42 for $m = 2$ and without loss of generality $t = 1$, we obtain a random variable that follows the univariate Rosenblatt distribution. For a nice historical overview of the Rosenblatt process as well as the Rosenblatt distribution, see [63]. In [66], properties of the univariate Rosenblatt distribution were discussed. In Figure 2.2, we cite their results for the probability density function and the cumulated density function for different long-range dependence parameters, in order to compare their results to our convergence results later on. Note that D in their approach corresponds to $d_1 = \frac{1-D}{2}$.

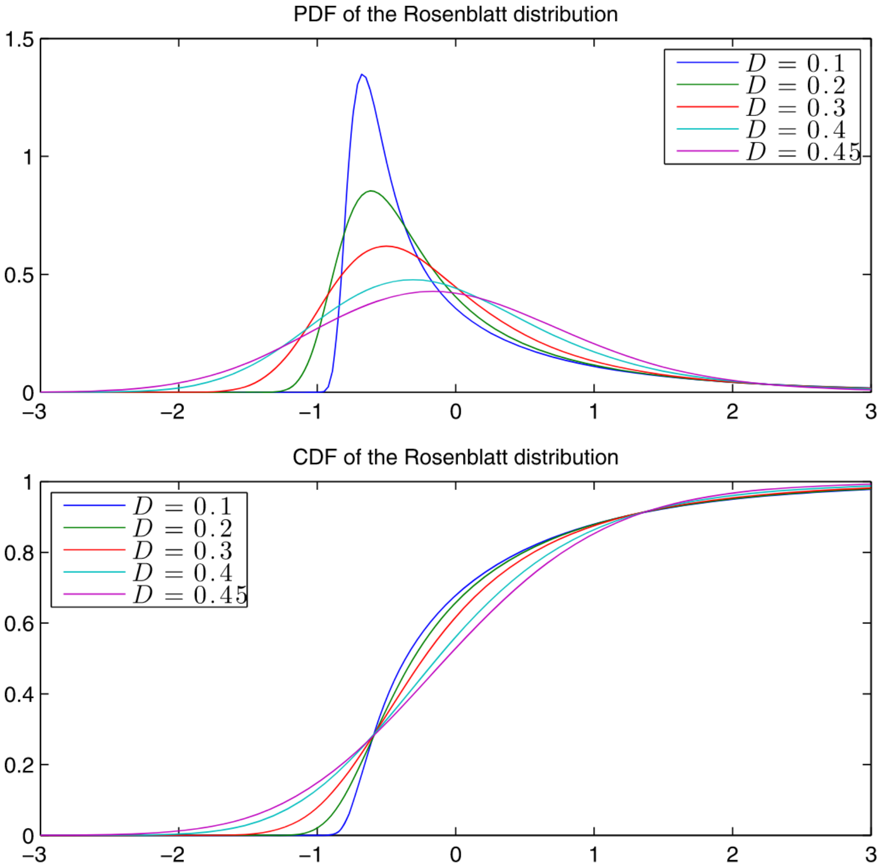

Figure 2.2: Plots of the pdf and cdf of a Rosenblatt distributed random variable, taken from [66], Figure 1, p.23.

The characteristic function of a standard Rosenblatt distributed random variable Z, i.e., with mean 0 and unit variance is given in the following power series which is only convergent near the origin, see [66], p.983:

$$\phi(\vartheta) = \exp\left(\frac{1}{2}\sum_{k=2}^{\infty}\left(2i\vartheta\,(2C_2)^{-1/2}\right)^k \frac{c_k}{k}\right),$$

where

$$c_k = \int_0^1 \ldots \int_0^1 |x_1 - x_2|^{2d_1-1}\,|x_2 - x_3|^{2d_1-1} \ldots |x_k - x_1|^{2d_1-1}\,\mathrm{d}x_1 \ldots \mathrm{d}x_k \tag{2.16}$$

and

$$C_2 := \frac{1}{2d_1\,(4d_1 - 1)}.$$

As we have seen in Definition 2.6 for the long-range dependence parameter d_1 it holds that $d_1 \in \left(0, \frac{1}{2}\right)$. It is very interesting to consider the extremes for $d \to \frac{1}{2}^-$ (limit case of strong long-range dependence) and $d \to 0^+$ (limit case of vanishing long-range dependence), where the notation $x \to y^+$ and $x \to y^-$ describes the limit from the right and from the left, respectively. For $d \to \frac{1}{2}^-$ the Rosenblatt distribution is a chi-squared distribution with mean 0 and variance 1. In the other extreme, if $d \to 0^+$ we observe a standard-normal distribution. The technical details to these two observations are given in [66], p. 983.

A special property of the Rosenblatt distribution is that it is uniquely determined by its cumulants κ_k, $k \in \mathbb{N}$, see in this context [62], Theorem 3c. The exact formula is given in [66], (10): $\kappa_1 = 0$ and for $k \geq 2$:

$$\kappa_k = 2^{k-1}(k-1)!\,(2C_2)^{-k/2}\,c_k. \tag{2.17}$$

We use the cumulants later on to compare limit distributions arising in the context of functionals with Hermite rank 2 of multivariate long-range dependent Gaussian processes to the Rosenblatt distribution.

In this section, the most relevant background information was given that is important and helpful in the course of this thesis. After recalling univariate and multivariate Gaussian processes, we took a closer look on short-range as well as long-range dependence. Subsequently, Hermite polynomials, the Hermite expansion and the Hermite rank were defined, as important tool in the context of long-range dependent Gaussian processes. The derivation of Wiener-Itô integrals was given and the time domain as well as the spectral domain representation were presented. The Hermite-Rosenblatt process, that turn out to be cruicial in the context of limit theorems for functionals with a certain Hermite rank, was introduced, as well as the closely linked Rosenblatt distribution. Having this mathematical background in mind, we are now able to turn to limit theorems for vector-valued Gaussian time series exhibiting long-range dependence and/or short-range dependence.

3 Limit theorems for functionals of long-range dependent multivariate Gaussian time series

In this chapter, we want to develop the theoretical framework we need to connect ordinal pattern analysis to long-range dependent time series and to mixed cases of short- and long-range dependence. Here, we keep the notation as general as possible and adapt it to ordinal analysis in the following two chapters. Having the historical background of limit theorems in the context of long-range dependent time series discussed in the introduction in mind, we recall the specific mathematical framework we consider in this work and give the incorporation in the literature. Note that in the course of the proofs of the limit theorems, which we derive in this chapter, several purple segments appear. The calculations or ideas discussed within these segments are not necessary for a stringent logical proof of the result, however, they are used to give further helpful comments to illustrate the argumentation.

In general, for $d \in \mathbb{N}$ we consider a stationary d-dimensional Gaussian time series $(Y_j)_{j\in\mathbb{Z}}$, with

$$Y_j := \left(Y_j^{(1)}, \ldots, Y_j^{(d)}\right)^t \tag{3.1}$$

such that $\mathbb{E}\left(Y_j^{(p)}\right) = 0$ and $\mathbb{E}\left(\left(Y_j^{(p)}\right)^2\right) = 1$ for all $j \in \mathbb{Z}$ and $p = 1, \ldots, d$. Furthermore, we require the cross-correlation function to fulfill $\left|r^{(p,q)}(k)\right| < 1$ for $p, q = 1, \ldots, d$ and $k \geq 1$, where the componentwise cross-correlation functions $r^{(p,q)}(k)$ are given by $r^{(p,q)}(k) = \mathbb{E}\left(Y_j^{(p)} Y_{j+k}^{(q)}\right)$ for each $p, q = 1, \ldots, d$ and $k \in \mathbb{Z}$. For each random vector Y_j we denote the covariance matrix by Σ_d, since it is independent of j due to stationarity. Therefore, we have $\Sigma_d = \left(r^{(p,q)}(0)\right)_{p,q=1,\ldots,d}$. This general setting holds for $(Y_j)_{j\in\mathbb{Z}}$ in Chapter 3 and Chapter 4.

We specify the dependence structure of $(Y_j)_{j\in\mathbb{Z}}$ and turn to long-range dependence: we assume that for the cross-correlation function $r^{(p,q)}(k)$ for each $p, q = 1, \ldots, d$, it holds that

$$r^{(p,q)}(k) = L_{p,q}(k) k^{d_p + d_q - 1}, \tag{3.2}$$

with $L_{p,q}(k) \to L_{p,q}$ $(k \to \infty)$ for finite constants $L_{p,q} \in [0, \infty)$ with $L_{p,p} \neq 0$, where the matrix $L = (L_{p,q})_{p,q=1,\ldots,d}$ has full rank, is symmetric and positive definite. Furthermore, the parameters $d_p, d_q \in \left(0, \frac{1}{2}\right)$ are called long-range dependence parameters. Therefore, $(Y_j)_{j\in\mathbb{Z}}$ fulfills the multivariate long-range dependence assumption given in Definition 2.13.

Hence, we observe a contrast to [9], p. 291, where it is assumed that for each $p, q = 1, \ldots, d$ one has the property

$$r^{(p,q)}(k) = b_{p,q} k^{2d-1} L(k),$$

for a parameter $d \in \left(0, \frac{1}{2}\right)$, a slowly varying function $L(k)$ and finite constants $b_{p,q}$ of which at least one is not equal to zero. This approach is based on the consideration of the one-dimensional definition of long-range dependence, see Definition 2.6, and hence $b_{p,p} = 1$ for $p = 1, \ldots, d$.

Note that the contrast to the property in (3.2) is that on the one hand, we no longer have the existence of a slowly varying function $L(k)$ such that $\lim_{k\to\infty} L(k) = \infty$, which seems a legitimate restrictions since the assumptions are still very appropriate for stastical inference, see [54], Section 9.4.1. On the other hand, the long-range dependence parameters depend on p and q and therefore, we are able to regard a much larger class of long-range dependent Gaussian processes. Note that the autocorrelation functions are given by $r^{(p,p)}(k) = L_{p,p} k^{2d_p-1}$, $p = 1, \ldots, d$ and, therefore, the parameters d_p and d_q of the autocorrelation functions of the corresponding marginal one-dimensional Gaussian processes $\left(Y_j^{(p)}\right)_{j\in\mathbb{Z}}$ and $\left(Y_j^{(q)}\right)_{j\in\mathbb{Z}}$ respectively, uniquely determine the long-range dependence parameter of the corresponding cross-correlation function $r^{(p,q)}(\cdot)$.

The processes we want to consider have a particular structure, namely for $h \in \mathbb{N}$, we obtain for fixed $j \in \mathbb{Z}$:

$$Y_{j,h} := \left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}, Y_j^{(2)}, \ldots, Y_{j+h-1}^{(2)}, \ldots, Y_j^{(d)}, \ldots, Y_{j+h-1}^{(d)}\right)^t \in \mathbb{R}^{dh}. \tag{3.3}$$

The following relation holds between the extendend process $(Y_{j,h})_{j\in\mathbb{Z}}$ and the primarily regarded process $(Y_j)_{j\in\mathbb{Z}}$. For all $k = 1, \ldots, dh$, $j \in \mathbb{Z}$ we have

$$Y_{j,h}^{(k)} = Y_{j+(k \mod h)-1}^{\lfloor \frac{k-1}{h} \rfloor + 1}, \tag{3.4}$$

where $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \le x\}$.

Example 3.1 *Let us consider $h = 2$ and $d = 3$. We obtain the vector*

$$Y_{j,2} = \left(Y_j^{(1)}, Y_{j+1}^{(1)}, Y_j^{(2)}, Y_{j+1}^{(2)}, Y_j^{(3)}, Y_{j+1}^{(3)}\right)^t.$$

If we want to determine $Y_{j,2}^{(3)}$ now, we use the formula in (3.4) and since $(3 \mod 2) = 1$ and $\left\lfloor \frac{2}{2} \right\rfloor = 1$, hence, we obtain $Y_{j,2}^{(3)} = Y_j^{(2)}$.

Note that the process $(Y_{j,h})_{j\in\mathbb{Z}}$ is still a centered Gaussian process since all finite-dimensional marginals of $(Y_j)_{j\in\mathbb{Z}}$ follow a normal distribution. Stationarity is preserved, too, since for all $p, q = 1, \ldots, dh$, $p \le q$ and $k \in \mathbb{Z}$ the cross-correlation function $r^{(p,q,h)}(k)$ of the process $(Y_{j,h})_{j\in\mathbb{Z}}$

is given by

$$
\begin{aligned}
r^{(p,q,h)}(k) &= \mathbb{E}\left(Y_{j,h}^{(p)} Y_{j+k,h}^{(q)}\right) \\
&= \mathbb{E}\left(Y_{j+(p \mod h)-1}^{\left\lfloor \frac{p-1}{h} \right\rfloor+1} Y_{j+k+(q \mod h)-1}^{\left\lfloor \frac{q-1}{h} \right\rfloor+1}\right) \\
&= r^{\left(\left\lfloor \frac{p-1}{h} \right\rfloor+1, \left\lfloor \frac{q-1}{h} \right\rfloor+1\right)}(k + ((q-p) \mod h)). \qquad (3.5)
\end{aligned}
$$

The covariance matrix $\Sigma_{d,h}$ of $Y_{j,h}$ has the following structure:

$$
\begin{aligned}
(\Sigma_{d,h})_{\substack{p,q=1,\ldots,d,\\ p\leq q}} &= \left(r^{(p,q,h)}(0)\right)_{\substack{p,q=1,\ldots,dh,\\ p\leq q,}}, \\
(\Sigma_{d,h})_{\substack{p,q=1,\ldots,d,\\ p>q}} &= \left(r^{(q,p,h)}(0)\right)_{\substack{p,q=1,\ldots,dh,\\ q<p}}.
\end{aligned}
$$

Hence, we arrive at

$$
\Sigma_{d,h} = \left(\Sigma_h^{(p,q)}\right)_{1\leq p,q\leq d}, \qquad (3.6)
$$

where $\Sigma_h^{(p,q)} = \mathbb{E}\left(\left(Y_1^{(p)}, \ldots, Y_h^{(p)}\right)^t \left(Y_1^{(q)}, \ldots, Y_h^{(q)}\right)\right) = \left(r^{(p,q)}(i-k)\right)_{1\leq i,k\leq h}$, $p, q = 1, \ldots, d$. Note that $\Sigma_h^{(p,q)} \in \mathbb{R}^{h\times h}$ and $r^{(p,q)}(k) = r^{(q,p)}(-k)$, $k \in \mathbb{Z}$ since we are studying cross-correlation functions, see Definition 2.12.

So finally we have to show that based on the assumptions on $(Y_j)_{j\in\mathbb{Z}}$ we are still dealing with a long-range dependent process.

Hence, we have to consider the cross-correlations again

$$
\begin{aligned}
r^{(p,q,h)}(k) &= r^{\left(\left\lfloor \frac{p-1}{h} \right\rfloor+1, \left\lfloor \frac{q-1}{h} \right\rfloor+1\right)}(k + ((q-p) \mod h)) \\
&= r^{(p^*,q^*)}(k + m^*) \\
&\simeq r^{(p^*,q^*)}(k) \quad (k\to\infty), \qquad (3.7)
\end{aligned}
$$

since $p^*, q^* \in \{1,\ldots,d\}$ and $m^* \in \{0,\ldots,h-1\}$, with $p^* := \left\lfloor \frac{p-1}{h} \right\rfloor + 1$, $q^* := \left\lfloor \frac{q-1}{h} \right\rfloor + 1$ and $m^* = (q-p) \mod h$.

Therefore, we are still dealing with a multivariate long-range dependent Gaussian process as defined in Definition 2.13. We see in the proofs of the following limit theorems that the crucial parameters that determine the asymptotic distribution are the long-range dependence parameters d_p, $p = 1, \ldots, d$ of the original process $(Y_j)_{j\in\mathbb{Z}}$ and therefore, we omit the detailed description of the parameters d_{p^*} here.

It is important to remark that the extended process $(Y_{j,h})_{j\in\mathbb{Z}}$ is also long-range dependent in the sense of [2], p. 2259, since

$$
\begin{aligned}
\lim_{k\to\infty} \frac{k^D r^{(p,q,h)}(k)}{L(k)} &= \lim_{k\to\infty} \frac{k^D r^{(p^*,q^*)}(k)}{L(k)} \\
&= \lim_{k\to\infty} \frac{k^D L_{p^*,q^*} k^{d_{p^*}+d_{q^*}-1}}{L(k)} \\
&= b_{p^*,q^*}, \qquad (3.8)
\end{aligned}
$$

with

$$D := \min_{p^* \in \{1,\dots,d\}} \{1 - 2d_{p^*}\} \in (0,1) \tag{3.9}$$

and $L(k)$ can be chosen as any constant $L_{p,q}$ that is not equal to zero, so for simplicity we assume without loss of generality $L_{1,1} \neq 0$ and, therefore, $L(k) = L_{1,1}$, since the condition in [2] only requires convergence to a finite constant b_{p^*,q^*}. Therefore, we may apply the results in [2] in the following.
We define the following set, which is needed in the proofs of the theorems of this chapter.

$$P^* := \{p \in \{1,\dots,d\} : \ d_p \geq d_q, \text{ for all } q \in \{1,\dots,d\}\} \tag{3.10}$$

and denote the corresponding long-range dependence parameter to each $p \in P^*$ by

$$d^* := d_p, \quad p \in P^*.$$

Before turning to limit theorems, we introduce a possibility to decompose the d-dimensional Gaussian process $(Y_j)_{j\in\mathbb{Z}}$ using the Cholesky decomposition, see [27]. Based on the definition of the multivariate normal distribution, see Definition 2.15, we find an upper triangular matrix $\tilde{A}$, such that $\tilde{A}\tilde{A}^t = \Sigma_d$. Then it holds, that

$$Y_j \overset{\mathcal{D}}{=} \tilde{A}U_j^*, \tag{3.11}$$

where U_j^* is a d-dimensional Gaussian process where each U_j^* has independent and identically $\mathcal{N}(0,1)$ distributed entries. We want to assure that $\left(U_j^*\right)_{j\in\mathbb{Z}}$ preserves the long-range dependent structure of $(Y_j)_{j\in\mathbb{Z}}$. Since we know from Definition 2.13, that

$$\mathbb{E}\left(Y_j Y_{j+k}\right) = \Gamma_Y(k) \simeq k^{D-\frac{1}{2}I_d} L k^{D-\frac{1}{2}I_d} \quad (k \to \infty),$$

the process $\left(U_j^*\right)$ has to fulfill

$$\mathbb{E}\left(U_j^* U_{j+k}^*\right) = \Gamma_{U^*}(k) \simeq k^{D-\frac{1}{2}I_d} L_U k^{D-\frac{1}{2}I_d} \quad (k \to \infty), \tag{3.12}$$

with $L = \tilde{A} L_{U^*} \tilde{A}^t$.
Then it holds for all $n \in \mathbb{N}$ that

$$(Y_j, \ j = 1,\dots,n) \overset{\mathcal{D}}{=} \left(\tilde{A}U_j^*, \ j = 1,\dots,n\right). \tag{3.13}$$

Note that the assumption in (3.12) is only well-defined because we assumed $\left|r^{(p,q)}(k)\right| < 1$ for $k \geq 1$ and $p,q = 1,\dots,d$ in (3.1). This becomes clear in the following considerations. In the proofs of the theorems in this chapter, we do not only need a decomposition of Y_j, but also of $Y_{j,h}$. As $Y_{j,h}$ is still a multivariate Gaussian process, the covariance matrix of $Y_{j,h}$ given by $\Sigma_{d,h}$ is positive definite. Hence, it is possible to find a upper triangular matrix A, such that $AA^t = \Sigma_{d,h}$. It holds that

$$Y_{j,h} \overset{\mathcal{D}}{=} AU_{j,h}$$

for

$$U_{j,h} = \left(U^{(1)}_{(j-1)h+1}, \ldots, U^{(1)}_{jh}, \ldots, U^{(d)}_{(j-1)h+1}, \ldots, U^{(d)}_{jh}\right)^t.$$

The random vector $U_{j,h}$ consists of $(d \cdot h)$ independent and standard normally distributed random variables. We notice the different structure of $U_{j,h}$ compared to $Y_{j,h}$. We assure that for consecutive j the entries in $U_{j,h}$ are all different while there are identical entries, for example in $Y_{1,h} = \left(Y^{(1)}_1, Y^{(1)}_2, \ldots, Y^{(d)}_h\right)^t$ and $Y_{2,h} = \left(Y^{(1)}_2, \ldots, Y^{(d)}_h, Y^{(d)}_{h+1}\right)^t$. This complicates our aim that

$$(Y_{j,h}, j = 1, \ldots, n)^t \overset{\mathcal{D}}{=} (AU_{j,h},\ j = 1, \ldots, n)^t \tag{3.14}$$

holds.

The special structure of $(Y_{j,h})_{j\in\mathbb{Z}}$, namely, that it is consisting of h consecutive entries of each marginal process $\left(Y^{(p)}_j\right)$, $p = 1, \ldots, d$, alongside with the dependence between two random vectors in the process $(Y_{j,h})$, has to be reflected in the covariance matrix of $(U_{j,h},\ j = 1, \ldots, n)$. Hence, we need to check whether such a vector $(U_{j,h},\ j = 1, \ldots, n)$ exists, i.e., if there is a positiv semi-definite matrix that fulfills these conditions. We define $\mathbf{A}$ as a block diagonal matrix with A as main-diagonal blocks and all off-diagonal blocks as $dh \times dh$-zero matrix. We denote the covariance matrix of $(Y_{j,h}, j = 1, \ldots, n)^t$ by $\Sigma_{Y,n}$ and define the following matrix:

$$\Sigma_{U,n} := \text{inv}\left(\mathbf{A}\right)\Sigma_{Y,n}\text{inv}\left(\mathbf{A}^t\right). \tag{3.15}$$

We know that $\Sigma_{Y,n}$ is positive semi-definite for all $n \in \mathbb{N}$ because (Y_j) is a Gaussian process. Mathematically described that means

$$x^t\Sigma_{Y,n}x \geq 0, \tag{3.16}$$

for all $x = (x_1, \ldots, x_{nhd})^t \in \mathbb{R}^{nhd}$. We conclude

$$\begin{aligned} x^t\Sigma_{U,n}x &= x^t\text{inv}\left(\mathbf{A}\right)\Sigma_{Y,n}\text{inv}\left(\mathbf{A}^t\right)x \\ &= \left(\text{inv}\left(\mathbf{A}^t\right)x\right)^t\Sigma_{Y,n}\left(x^t\text{inv}\left(\mathbf{A}\right)\right) \\ &\overset{(3.16)}{\geq} 0. \end{aligned}$$

Therefore, $\Sigma_{U,n}$ is a positive semi-definite matrix for all $n \in \mathbb{N}$ and the random vector

$$(U_{j,h},\ j = 1, \ldots, n)^t\mathcal{N} \sim (0, \Sigma_{U,n})$$

exists and (3.14) holds. Note that we do not have any further information on the dependence structure within the process (U_j), in general, this process does neither exhibit long-range dependence nor independence nor stationarity.

Example 3.2 *We consider the case $h = 2$ and $d = 1$, such that*

$$Y_{j,2} = \left(Y^{(1)}_j, Y^{(1)}_{j+1}\right)^t.$$

We denote $r_1 := r^{(1,1)}(1)$ *and* $r_2 := r^{(1,1)}(2)$ *for simplicity. Hence, for each* j *the covariance matrix* $\Sigma_{1,2}$ *of* $(Y_{j,h})$ *is given by*

$$\Sigma_{1,2} = \begin{pmatrix} 1 & r_1 \\ r_1 & 1 \end{pmatrix}.$$

The upper triangular matrix A *such that* $AA^t = \Sigma_{1,2}$ *is obtained by using the Cholesky decomposition:*

$$A = \begin{pmatrix} 1 & 0 \\ r_1 & \sqrt{1-r_1^2} \end{pmatrix}.$$

We arrive at

$$Y_{j,h} \overset{\mathcal{D}}{=} AU_{j,h},$$

for a random vector $U_{j,h} \sim \mathcal{N}(0, I_{1,2})$*. We turn to the joint distribution of* $(Y_{1,h}, Y_{2,h})^t$*. Let* $\boldsymbol{0}$ *be the* 2×2*-dimensional zero matrix. Our goal is*

$$(Y_{1,h}, Y_{2,h})^t = (Y_1, Y_2, Y_2, Y_3) \overset{\mathcal{D}}{=} \begin{pmatrix} A & \boldsymbol{0} \\ \boldsymbol{0} & A \end{pmatrix} (U_1, U_2, U_3, U_4)^t, \tag{3.17}$$

with a multivariate normal vector $(U_1, U_2, U_3, U_4)^t$*.*
A first idea is to choose $U_2 = U_3$*. However, this is not reasonable, since we obtain*

$$(X_2, X_2)^t \overset{\mathcal{D}}{=} \left(r_1 U_1 + \sqrt{1-r_1^2}U_2, U_2\right)^t.$$

That yields

$$\mathbb{E}\left(X_2^2\right) = 1 - r_1^2 \neq 1 \text{ if } r \neq 0,$$

since U_1 *and* U_2 *are independent by assumption.*
The next approach is to assume (U_1, U_2, U_3, U_4) *as a random vector with independent entries. This is not reasonable either, as* (3.17) *yields*

$$r_2 = \mathbb{E}(Y_1 Y_3) = \mathbb{E}(U_1 U_3) = 0,$$

and r_2 *does not equal zero in general.*
A third idea might be to use $U_3 = r_1 U_1 + \sqrt{1-r_1^2}U_2$ *and to choose* U_4 *to be independent. Besides the calculations for the dependence of* U_4 *on* U_1 *and* U_2*, this approach does not work for* $h > 2$*.*

Therefore, we observe that the existence of the vector (U_1, U_2, U_3, U_4) *and the appropriate choice of the covariance matrix are non-trivial problems. For the solution we derived in* (3.15)*, we provide an example for* $r_1 = 0.5$ *and* $r_2 = 0.2$*. Then*

$$\Sigma_{U,4} = \begin{pmatrix} 1 & 0 & 0.5 & -0.058 \\ 0 & 1 & 0.866 & 0.03 \\ 0.5 & 0.866 & 1 & 0 \\ -0.058 & 0.03 & 0 & 1 \end{pmatrix}.$$

We observe that $(U_1, \ldots, U_4)$ *is not stationary.*

According to these calculations, we may use the property in (3.14) in the proofs in this chapter, however, we need to retransform to $Y_{j,h}$ and later on to the entries $Y_j = \left(Y_j^{(1)}, \ldots, Y_j^{(d)}\right)^t$ again, as we cannot benefit from the property of $U_{j,h}$ that for fixed j we have independent entries, as we do no longer have any information on the dependence structure or even stationarity of the process $(U_j)_{j\in\mathbb{Z}}$.

We face this challenge in the following and prove limit theorems for functionals of a multivariate stationary Gaussian process as described by $(Y_{j,h})_{j\in\mathbb{Z}}$. Due to the considerations in Section 2.4 we assume that the studied functionals f fulfill $\mathbb{E}\left(f^2\left(Y_{j,h}\right)\right) < \infty$ and have a certain Hermite rank, which is crucial to determine the asymptotic distribution. For simplicity, we also assume $\mathbb{E}\left(f\left(Y_{j,h}\right)\right) = 0$ in the proofs.

3.1 Limit theorems for functionals with Hermite rank 1

We begin with the first result for functions that exhibit Hermite rank 1. As we have seen in Table 2.1, we do not need to distinguish between different ranges of d^*, because the transformed process is still long-range dependent. In the following result, we prove a central limit theorem for functionals f with Hermite rank 1 applied to the long-range dependent Gaussian process $(Y_{j,h})_{j\in\mathbb{Z}}$. In doing so, we shed light on the impact of the special structure of $(Y_{j,h})_{j\in\mathbb{Z}}$ as described in (3.3).

Theorem 3.3 *Let $(Y_j)_{j\in\mathbb{Z}}$ be a d-dimensional stationary Gaussian process as defined in* (3.1) *that exhibits long-range dependence in the sense of Definition 2.13. For $h \in \mathbb{N}$ we fix*

$$Y_{j,h} := \left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}, \ldots, Y_j^{(d)}, \ldots, Y_{j+h-1}^{(d)}\right)^t \in \mathbb{R}^{dh}$$

with $Y_{j,h} \sim \mathcal{N}\left(0, \Sigma_{d,h}\right)$ and $\Sigma_{d,h}$ as described in (3.6)*. Let $f : \mathbb{R}^{dh} \to \mathbb{R}$ be a function with Hermite rank* 1 *such that the set of discontinuity points D_f is a Null set with respect to the dh-dimensional Lebesgue measure. Furthermore we assume f fulfills $\mathbb{E}\left(f^2\left(Y_{j,h}\right)\right) < \infty$. Then, it holds that*

$$n^{-(d^*+1/2)} C_1^{-1/2} \sum_{j=1}^{n} \left(f\left(Y_{j,h}\right) - \mathbb{E}\left(f\left(Y_{j,h}\right)\right)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \left(\sum_{p,q\in P^*} \sum_{i,l=1}^{h} \alpha_i^{(p)} \alpha_l^{(q)} L_{p,q}\right)\right)$$

with $C_1 := \frac{1}{d^*(2d^*+1)}$ and $\alpha_i^{(p)} = \alpha_{i+(p-1)h}$ for $i = 1, \ldots, h$, $p = 1, \ldots, d$. Further we have $\alpha = (\alpha_1, \ldots, \alpha_{dh})^t = \Sigma_{d,h}^{-1} c$, $c = \mathbb{E}\left(f\left(Y_{1,h}\right) Y_{1,h}\right)$ and P^*, d^* defined as in (3.10)*.*

Proof. Without loss of generality, we assume $\mathbb{E}\left(f\left(Y_{j,h}\right)\right) = 0$. We incorporate the idea in the proof of Theorem 5.6 in [10]. The property in (3.14) yields for a continuous function f applying the continuous mapping theorem:

$$\sum_{j=1}^{n} f\left(Y_{j,h}\right) \stackrel{\mathcal{D}}{=} \sum_{j=1}^{n} f\left(AU_{j,h}\right) \tag{3.18}$$

We now want to study the asymptotic behavior of the partial sum $\sum_{j=1}^{n} f^*\left(U_{j,h}\right)$ where $f^*\left(U_{j,h}\right) := f\left(AU_{j,h}\right) \stackrel{\mathcal{D}}{=} f\left(Y_{j,h}\right)$. The equality in distribution has no impact on the asymptotic distribution of $\sum_{j=1}^{n} f(Y_{j,h})$: we use (3.14) and apply the continuous mapping theorem. We follow the Portmanteau Theorem. All bounded and continuous functions are included in the set of continuous functions we observe that the appropriately normalized left-hand side in (3.18) converges to the same distribution as the right-hand side.

The Hermite rank of $f^*(U_{j,h})$ can be calculated analogously, since $m(f, \Sigma_{d,h}) = m(f^*, I_{dh})$ for $f^*(U_{j,h}) = f \circ A(U_{j,h}) = f(Y_{j,h})$, see Lemma 2.32. According to the multivariate generalization of the Reduction Theorem in [2], p.2261, we know that the first order term in the Hermite expansion of f determines the asymptotic behavior, due to f having Hermite rank 1. We can state the following:

$$\sum_{j=1}^{n} f(Y_{j,h}) \overset{\mathcal{D}}{=} \sum_{j=1}^{n} f^*(U_{j,h}) = \sum_{j=1}^{n}\sum_{i=1}^{dh} \mathbb{E}\left(f^*(U_{j,h})\, U_{j,h}^{(i)}\right) U_{j,h}^{(i)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right), \tag{3.19}$$

where for the remainder R_n with $R_n = o_{\mathbb{P}}\left(n^{d^*+1/2}\right)$ it holds that

$$\lim_{n\to\infty} \frac{R_n}{n^{d^*+1/2}} \overset{\mathbb{P}}{\to} 0.$$

Define the vector of Hermite coefficients in the Hermite expansion with respect to $U_{j,h}$ by $b := \mathbb{E}(f^*(U_{j,h})\, U_{j,h}) \in \mathbb{R}^{dh}$ which we can express by

$$b = \mathbb{E}(f^*(U_{j,h})\, U_{j,h}) = \mathbb{E}(f(Y_{j,h})\, U_{j,h}) = \mathbb{E}\left((f(Y_{j,h})\, A^{-1} Y_{j,h}\right) = A^{-1} c,$$

with $c \in \mathbb{R}^{dh}$ being the vector of the *original* first order Hermite coefficients with respect to $Y_{j,h}$ given by $c := \mathbb{E}(f(Y_{j,h})\, Y_{j,h})$. Note that b and c are in general not coinciding.
Hence, we can continue to study (3.19):

$$\begin{aligned}
\sum_{j=1}^{n}\sum_{i=1}^{dh} \mathbb{E}\left(f^*(U_{j,h})\, U_{j,h}^{(i)}\right) U_{j,h}^{(i)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right) &= \sum_{j=1}^{n}\sum_{i=1}^{dh} b_i U_{j,h}^{(i)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right)\\
&= \sum_{j=1}^{n} b^t U_{j,h} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right)\\
&\overset{\mathcal{D}}{=} \sum_{j=1}^{n} c^t \left(A^{-1}\right)^t A^{-1} Y_{j,h} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right)\\
&= \sum_{j=1}^{n} \left(\Sigma_{d,h}^{-1} c\right)^t Y_{j,h} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right)\\
&= \sum_{j=1}^{n}\sum_{i=1}^{dh} \alpha_i Y_{j,h}^{(i)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right),
\end{aligned}$$

with b_i, c_i and α_i, $i = 1, \ldots, dh$ being the entries of b, c and α respectively, where $\alpha := \Sigma_{d,h}^{-1} c$.
Now we can use the special structure of $Y_{j,h}$ namely that it is compound of the d-dimensional Gaussian process $(Y_j)_{j\in\mathbb{Z}}$ by $Y_{j,h} = \left(Y_j^{(1)}, Y_{j+1}^{(1)}, \ldots, Y_{j+h-1}^{(1)}, \ldots, Y_j^{(d)}, \ldots, Y_{j+h-1}^{(d)}\right)^t$. Then, we obtain

$$\sum_{j=1}^{n}\sum_{i=1}^{dh} \alpha_i Y_{j,h}^{(i)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right) = \sum_{j=1}^{n}\sum_{p=1}^{d}\sum_{i=1}^{h} \alpha_i^{(p)} Y_{j+i-1}^{(p)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right)$$

$$
\begin{aligned}
&= \sum_{p=1}^{d} \sum_{i=1}^{h} \alpha_i^{(p)} \sum_{j=1}^{n} Y_{j+i-1}^{(p)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right) \\
&= \sum_{p=1}^{d} \sum_{i=1}^{h} \alpha_i^{(p)} \sum_{j=i}^{n+i} Y_{j}^{(p)} + o_{\mathbb{P}}\left(n^{d^*+1/2}\right) \\
&= \sum_{p=1}^{d} \Bigg(\left(\sum_{i=1}^{h} \alpha_i^{(p)} \right) \left(\sum_{j=1}^{n} Y_j^{(p)} \right) - \sum_{i=2}^{h} \alpha_i^{(p)} \left(\sum_{j=1}^{i-1} Y_j^{(p)} \right) \\
&\qquad + \sum_{i=2}^{h} \alpha_i^{(p)} \left(\sum_{j=n+1}^{n+i-1} Y_j^{(p)} \right) \Bigg) + o_{\mathbb{P}}\left(n^{d^*+1/2}\right) \\
&= \sum_{p=1}^{d} \left(\sum_{i=1}^{h} \alpha_i^{(p)} \right) \left(\sum_{j=1}^{n} Y_j^{(p)} \right) + o_{\mathbb{P}}\left(n^{d^*+1/2}\right) \qquad (3.20)
\end{aligned}
$$

where $\alpha^{(p)} = \left(\alpha_1^{(p)}, \ldots, \alpha_h^{(p)}\right)^t$, $p = 1, \ldots, d$ with $\alpha_i^{(p)} = \alpha_{i+(p-1)h}$ for $i = 1, \ldots, h$ such that $\alpha = \left(\alpha^{(1)}, \ldots, \alpha^{(d)}\right)$.

For simplicity consider $d = 2$ and $h = 2$. Then, we have $\alpha = (\alpha_1, \ldots, \alpha_4)^t$ and now divide this into $\alpha^{(1)} := (\alpha_1, \alpha_2)^t$ and $\alpha^{(2)} := (\alpha_3, \alpha_4)^t$, such that $\alpha = \left(\alpha^{(1)}, \alpha^{(2)}\right)^t$. We can also observe that for example for $i = 1$ and $p = 2$ the formula above is verified, since $\alpha_1^{(2)} = \alpha_3$.
An important notice here is that unfortunately it is not possible to (easily) further divide $\alpha = \Sigma_{d,h}^{-1} c$ into a combination of $\Sigma_h^{(p,q)}$ and a division of c depending on p, since we have $\Sigma_{d,h} = \left(\Sigma_h^{(p,q)}\right)_{1 \le p,q \le d}$, which, however, does not yield $\Sigma_{d,h}^{-1} = \left(\left(\Sigma_h^{(p,q)}\right)^{-1}\right)_{1 \le p,q \le d}$ in general, but a more technical structure, so we need to continue with the notation of α.

We can now define

$$
\tilde{Y}_j := \sum_{p=1}^{d} \left(\sum_{i=1}^{h} \alpha_i^{(p)} \right) Y_j^{(p)}, \qquad (3.21)
$$

so in terms of (3.20) we have

$$
\sum_{p=1}^{d} \left(\sum_{i=1}^{h} \alpha_i^{(p)} \right) \left(\sum_{j=1}^{n} Y_j^{(p)} \right) = \sum_{j=1}^{n} \tilde{Y}_j.
$$

Note that $\left(\tilde{Y}_1, \ldots, \tilde{Y}_n\right)^t$ is multivariate Gaussian for each $n \in \mathbb{N}$ since

$$\begin{pmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \\ \vdots \\ \tilde{Y}_n \end{pmatrix} = M \begin{pmatrix} Y_1^{(1)} \\ Y_1^{(2)} \\ \vdots \\ Y_1^{(d)} \\ \vdots \\ Y_n^{(1)} \\ Y_n^{(2)} \\ \vdots \\ Y_n^{(d)} \end{pmatrix}$$

with

$$M := \begin{pmatrix} \sum_{i=1}^h \alpha_i^{(1)} & \sum_{i=1}^h \alpha_i^{(2)} & \ldots & \sum_{i=1}^h \alpha_i^{(d)} & 0 & 0 & \ldots & 0 & \cdots & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & \sum_{i=1}^h \alpha_i^{(1)} & \sum_{i=1}^h \alpha_i^{(2)} & \ldots & \sum_{i=1}^h \alpha_i^{(d)} & \cdots & 0 & 0 & \ldots & 0 \\ & & & & & & & & \ddots & & & & \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \cdots & \sum_{i=1}^h \alpha_i^{(1)} & \sum_{i=1}^h \alpha_i^{(2)} & \ldots & \sum_{i=1}^h \alpha_i^{(d)} \end{pmatrix}$$

and, therefore, this is just a linear transformation of a multivariate Gaussian random vector.
We now need to modify the one-dimensional Gaussian process $\left(\tilde{Y}_j\right)_{j \in \mathbb{Z}}$ in order to get a limit distribution based on Theorem 5.3.1 in [54]. First we have to ensure that for a fixed j, $\tilde{Y}_j$ is a standard Gaussian random variable. Therefore, we compute the variance ($\mathbb{E}\left(\tilde{Y}_j\right) = 0$ follows directly by definition):

$$\mathbb{E}\left(\tilde{Y}_j^2\right) = \mathbb{E}\left(\sum_{p=1}^d \sum_{q=1}^d \left(\sum_{i=1}^h \alpha_i^{(p)}\right)\left(\sum_{k=1}^h \alpha_k^{(q)}\right) Y_j^{(p)} Y_j^{(q)}\right) = \sum_{p,q=1}^d \left(\sum_{i=1}^h \alpha_i^{(p)}\right)\left(\sum_{k=1}^h \alpha_k^{(q)}\right) r^{(p,q)}(0),$$

and therefore, we obtain

$$\tilde{Y}_{j,normed} := \frac{\tilde{Y}_j}{\sqrt{\sum_{p,q=1}^d \left(\sum_{i=1}^h \alpha_i^{(p)}\right)\left(\sum_{k=1}^h \alpha_k^{(q)}\right) r^{(p,q)}(0)}},$$

where we explicitly exclude the case that $\mathbb{E}\left(\tilde{Y}_j^2\right) = \sum_{p,q=1}^d \left(\sum_{i=1}^h \alpha_i^{(p)}\right)\left(\sum_{k=1}^h \alpha_k^{(q)}\right) r^{(p,q)}(0) = 0$, since in this case $\tilde{Y}_j = 0$ $\mathbb{P}$-a.s. for all $j \in \mathbb{Z}$ and the convergence problem gets trivial.
We have to determine the cross-correlation function of the process $(Y_{j,normed})_{j \in \mathbb{Z}}$ in order to find the right normalization constants in terms of the limit theorem.

$$\begin{aligned} \mathbb{E}\left(\tilde{Y}_{j,normed}\tilde{Y}_{j+k,normed}\right) &= \frac{\mathbb{E}\left(\tilde{Y}_j \tilde{Y}_{j+k}\right)}{\sum_{p,q=1}^d \left(\sum_{i=1}^h \alpha_i^{(p)}\right)\left(\sum_{l=1}^h \alpha_l^{(q)}\right) r^{(p,q)}(0)} \\ &= \frac{\mathbb{E}\left(\sum_{p,q=1}^d \sum_{i,l=1}^h \alpha_i^{(p)} \alpha_l^{(q)} Y_j^{(p)} Y_{j+k}^{(q)}\right)}{\sum_{p,q=1}^d \left(\sum_{i=1}^h \alpha_i^{(p)}\right)\left(\sum_{l=1}^h \alpha_l^{(q)}\right) r^{(p,q)}(0)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\sum_{p,q=1}^{d}\sum_{i,l=1}^{h}\alpha_i^{(p)}\alpha_l^{(q)}r^{(p,q)}(k)}{\sum_{p,q=1}^{d}\left(\sum_{i=1}^{h}\alpha_i^{(p)}\right)\left(\sum_{l=1}^{h}\alpha_l^{(q)}\right)r^{(p,q)}(0)} \\
&\simeq \frac{\sum_{p,q\in P^*}\sum_{i,l=1}^{h}\alpha_i^{(p)}\alpha_l^{(q)}L_{p,q}k^{2d^*-1}}{\sum_{p,q=1}^{d}\left(\sum_{i=1}^{h}\alpha_i^{(p)}\right)\left(\sum_{l=1}^{h}\alpha_l^{(q)}\right)r^{(p,q)}(0)},
\end{aligned}
\tag{3.22}
$$

with P^* and d^* defined as in (3.10). So note that the univariate Gaussian process $(Y_{j,normed})_{j\in\mathbb{Z}}$ is long-range dependent in the sense of Definition 2.6 with long-range dependence parameter d^* and slowly varying function $L_*(k)$ such that $\lim_{k\to\infty} L_*(k) = L_* := \frac{\sum_{p,q\in P^*}\sum_{i,l=1}^{h}\alpha_i^{(p)}\alpha_l^{(q)}L_{p,q}}{\sum_{p,q=1}^{d}\left(\sum_{i=1}^{h}\alpha_i^{(p)}\right)\left(\sum_{l=1}^{h}\alpha_l^{(q)}\right)r^{(p,q)}(0)}$. We can now conclude using Theorem 5.3.1 in [54], that

$$
n^{-(d^*+1/2)}C_1^{-1/2}\sum_{j=1}^{n}\tilde{Y}_{j,normed} \xrightarrow{\mathcal{D}} \mathcal{N}(0, L_*),
$$

with $C_1 := \frac{1}{d^*(2d^*+1)}$. Using this result we can determine the limit variance of (3.20) with a suitable normalization:

$$
\begin{aligned}
& n^{-(d^*+1/2)}C_1^{-1/2}\sum_{j=1}^{n}\tilde{Y}_{j,normed} \xrightarrow{\mathcal{D}} \mathcal{N}(0, L_*), \\
\Leftrightarrow\quad & n^{-(d^*+1/2)}C_1^{-1/2}\left(\sum_{p,q=1}^{d}\left(\sum_{i=1}^{h}\alpha_i^{(p)}\right)\left(\sum_{k=1}^{h}\alpha_k^{(q)}\right)r^{(p,q)}(0)\right)^{-1/2}\sum_{j=1}^{n}\tilde{Y}_j \xrightarrow{\mathcal{D}} \mathcal{N}(0, L_*) \\
\Leftrightarrow\quad & n^{-(d^*+1/2)}C_1^{-1/2}\sum_{j=1}^{n}\tilde{Y}_j \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, L_*\left(\sum_{i=1}^{h}\alpha_i^{(p)}\right)\left(\sum_{k=1}^{h}\alpha_k^{(q)}\right)r^{(p,q)}(0)\right)
\end{aligned}
$$

and finally we obtain

$$
n^{-(d^*+1/2)}C_1^{-1/2}\sum_{j=1}^{n}\tilde{Y}_j \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \left(\sum_{p,q\in P^*}\sum_{i,l=1}^{h}\alpha_i^{(p)}\alpha_l^{(q)}L_{p,q}\right)\right).
$$

We do not need a further distinction of cases with the previously excluded case that $\tilde{Y}_j = 0$ $\mathbb{P}$-a.s. for all $j \in \mathbb{Z}$, since the right-hand side of the above limit yields a degenerated limit distribution in this case, because of $\mathbb{E}\left(\tilde{Y}_j\tilde{Y}_{j+k}\right) = 0$ for all k and, therefore, in this case $\sum_{p,q\in P^*}\sum_{i,l=1}^{h}\alpha_i^{(p)}\alpha_l^{(q)}L_{p,q} = 0$, see [9], p. 300. □

It is possible to soften the assumption in Theorem 3.3 regarding the required long-range dependence of the underlying multivariate Gaussian process.

Corollary 3.4 *Instead of demanding in the assumptions of Theorem 3.3 that Definition 2.13 holds for all d_p, $p = 1, \dots, d$, it is indeed sufficient to assume that we have at least one $p \in \{1, \dots, d\}$, such that $d_p \in \left(0, \frac{1}{2}\right)$. For the other parameters we can also allow for $d_p \in (-\infty, 0)$. Then, the statement of Theorem 3.3 remains valid.*

Proof. The crucial point in the proof of Theorem 3.3 that determines the asymptotic distribution is given in (3.22). The asymptotic equality there is dominated by the largest paramter

d^* as defined in (3.10). Since we assumed that we have at least one parameter d_p such that the corresponding one-dimensional process $Y_j^{(p)}$ exhibits long-range dependence in the sense of Definition 2.6, we ensured $d^* \in \left(0, \frac{1}{2}\right)$ and hence need no adjustments in the proof. □

Let us remark, that if one considers the extended case of a functional limit theorem for functions with Hermite rank 1 as it is done in [54], Theorem 5.3.1 for the univariate case or [23], Theorem 3.3 for the multivariate extension, the limiting process is given by the univariate fractional Brownian motion as defined in Definition 2.19 and analogously the operator fractional Brownian motion as given in Theorem 2.44, since this is the Hermite-Rosenblatt process of order $m = 1$.

By now we have proved a limit theorem for a functional with Hermite rank 1 considering the special structure of the dh-dimensional process needed with h consecutive entries for each of the d univariate marginal process, as it appears in the context of ordinal analysis later on. As it is desirable in applications, we may allow for a pure long-range dependent structure of the underlying process, as well as for mixed cases of short- and long-range dependent components. In the next section, we complement the theoretical framework by considering functionals with Hermite rank 2, as they appear in the context of estimation of ordinal pattern probabilities as well as ordinal pattern dependence.

3.2 Limit theorems for functionals with Hermite rank 2

In this section, we focus on a limit theorem for functionals with Hermite rank 2. In the univariate results, as given in [54], Theorem 5.3.1, the limit is no longer Gaussian, but follows the Rosenblatt distribution. Some of the properties that characterize the univariate Rosenblatt distribution are given in Section 2.7.3. Therefore, it is an interesting research challenge to gain more information on the distribution of the random vector that arises as limit in the multivariate case considered here.

We begin with a preceding lemma that deals with the asymptotic distribution of the sample cross-covariance matrix.

Lemma 3.5 *Let $(Y_j)_{j\in\mathbb{Z}}$ be a d-dimensional Gaussian process as defined in* (3.1) *that fulfills* (3.2) *with $d_1 = \ldots = d_d = d^*$, such that*

$$\Gamma_Y(k) = \mathbb{E}\left(Y_j Y_{j+k}^t\right) \simeq L k^{2d^*-1}, \quad (k \to \infty).$$

Let C_2 be a normalization constant,

$$C_2 = \frac{1}{2d^*\left(4d^* - 1\right)}$$

and let B_Y be an upper triangular matrix, such that

$$B_Y B_Y^t = L.$$

Further, for $l \in \mathbb{N}$ we have

$$\hat{\Gamma}_{Y,n}(l) = \frac{1}{n-l}\sum_{j=1}^{n-l} Y_j Y_{j+l}^t.$$

Then, for $h \in \mathbb{N}$ it holds that

$$\left(n^{1-2d^*}(C_2)^{-1/2}(B_Y \otimes B_Y)^{-1} \operatorname{vec}\left(\hat{\Gamma}_n(l) - \Gamma(l)\right),\ l = 0, \ldots, h-1\right)$$
$$\xrightarrow{D} \left(\operatorname{vec}\left(Z^{(p,q)}_{2,d^*+1/2}(1)\right)_{p,q=1,\ldots,d},\ l = 0, \ldots, h-1\right),$$

where $Z^{(p,q)}_{2,d^*+1/2}(1)$ has the spectral domain representation

$$Z^{(p,q)}_{2,d^*+1/2}(1) = K_{p,q}(d^*) \int''_{\mathbb{R}^2} \frac{\exp\left(i\left(\lambda_1 + \lambda_2\right)\right) - 1}{i\left(\lambda_1 + \lambda_2\right)} |\lambda_1 \lambda_2|^{-d^*} \tilde{B}^{(p)}(\mathrm{d}\lambda_1)\, \tilde{B}^{(q)}(\mathrm{d}\lambda_2)$$

where

$$K^2_{p,q}(d^*) = \begin{cases} \frac{1}{2C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p = q \\ \frac{1}{C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p \neq q. \end{cases}$$

and $\tilde{B}(\mathrm{d}\lambda) = \left(\tilde{B}^{(1)}(\mathrm{d}\lambda), \ldots, \tilde{B}^{(d)}(\mathrm{d}\lambda)\right)$ is a multivariate Hermitian-Gaussian random measure as defined in (2.15).

Proof. First, we can use (3.11):

$$Y_j \overset{\mathcal{D}}{=} \tilde{A} U^*_j,$$

such that $\left(U^*_j\right)$ is a multivariate Gaussian process with $U^*_j \sim \mathcal{N}(0, I_d)$ and $\left(U^*_j\right)$ is still long-range dependent, see (3.12). It is possible to decompose the sample cross-covariance matrix $\hat{\Gamma}_{Y,n}(l) - \Gamma_Y(l)$ with respect to (Y_j) at lag l given by

$$\hat{\Gamma}_{Y,n}(l) - \Gamma_Y(l) = \frac{1}{n-l} \sum_{j=1}^{n-l} Y_j Y^t_{j+l} - \mathbb{E}\left(Y_j Y^t_{j+l}\right)$$

to

$$\hat{\Gamma}_{Y,n}(l) - \Gamma_Y(l) \overset{\mathcal{D}}{=} \tilde{A}\left(\hat{\Gamma}_{U^*,n}(l) - \Gamma_{U^*}(l)\right) \tilde{A}^t,$$

where we define the sample cross-covariance matrix $\hat{\Gamma}_{U^*,n}(l) - \Gamma_{U^*}(l)$ with respect to $\left(U^*_j\right)$ at lag l by

$$\hat{\Gamma}_{U^*,n}(l) - \Gamma_{U^*}(l) = \frac{1}{n-l} \sum_{j=1}^{n-l} U^*_j U^*_{j+l} - \mathbb{E}\left(U^*_j U^*_{j+l}\right).$$

Each entry of

$$\hat{\Gamma}_{U^*,n}(l) - \Gamma_{U^*}(l) = \left(\hat{r}^{(p,q)}_{n,U^*}(l) - r^{(p,q)}_{U^*}(l)\right)_{p,q=1,\ldots,d}$$

is given by

$$\hat{r}^{(p,q)}_{n,U^*}(l) - r^{(p,q)}_{U^*}(l) := \sum_{j=1}^{n} U^{*\ (p)}_j U^{*\ (q)}_{j+l} - \mathbb{E}\left(U^{*\ (p)}_j U^{*\ (q)}_{j+l}\right).$$

Following [22], proof of Lemma 7.4, the limit distribution of

$$\left(\hat{\Gamma}_{U^*,n}(l) - \Gamma_{U^*}(l),\ l = 0, \dots, h-1\right)$$

is equal to the limit distribution of

$$\left(\hat{\Gamma}_{U^*,n}(0) - \Gamma_{U^*}(0),\ l = 0, \dots, h-1\right).$$

We recall the assumption that $d^* = d_p$ for all $p = 1, \dots, d$. We follow [2], Theorem 6 and use the Cramer-Wold device: Let $a_{1,1}, a_{1,2}, \dots, a_{d,d} \in \mathbb{R}$. We are interested in the asymptotic behaviour of

$$\begin{aligned} & n^{1-2d^*} \sum_{p,q=1}^{d} a_{p,q} \left(\hat{r}_{n,U}^{(p,q)}(0) - r_U^{(p,q)}(0)\right) \\ =& n^{-2d^*} \sum_{j=1}^{n} \sum_{p,q=1}^{d} a_{p,q} \left(U_j^{*\ (p)} U_j^{*\ (q)} - \mathbb{E}\left(U_j^{*\ (p)} U_j^{*\ (q)}\right)\right). \end{aligned}$$

We consider the function

$$f\left(U_j^*\right) = \sum_{p,q=1}^{d} a_{p,q} \left(U_j^{*\ (p)} U_j^{*\ (q)} - \mathbb{E}\left(U_j^{*\ (p)} U_j^{*\ (q)}\right)\right) \tag{3.23}$$

and may apply Theorem 6 in [2]. Using the Hermite decomposition of f as given in (2.8), we observe that f and, therefore, $a_{p,q}$, $p, q = 1, \dots, d$, only affects the Hermite coefficients. Indeed, using Lemma 2.35 the Hermite coefficients reduce to $a_{p,q}$ for each summand on the right-hand side in (3.23). Hence, we can state

$$n^{-2d^*} \sum_{j=1}^{n} \sum_{p,q=1}^{d} a_{p,q} \left(U_j^{*\ (p)} U_j^{*\ (q)} - \mathbb{E}\left(U_j^{*\ (p)} U_j^{*\ (q)}\right)\right) \tag{3.24}$$

$$\xrightarrow{\mathcal{D}} \sum_{p,q=1}^{d} a_{p,q} Z_{2,d^*+1/2}^{(p,q)}(1), \tag{3.25}$$

where $Z_{2,d^*+1/2}^{(p,q)}(1)$ has the spectral domain representation

$$Z_{2,d^*+1/2}^{(p,q)}(1) = K_{p,q}(d^*) \int_{\mathbb{R}^2}'' \frac{\exp\left(i\left(\lambda_1 + \lambda_2\right)\right) - 1}{i\left(\lambda_1 + \lambda_2\right)} \left|\lambda_1 \lambda_2\right|^{-d^*} \tilde{B}^{(p)}\left(\mathrm{d}\lambda_1\right) \tilde{B}^{(q)}\left(\mathrm{d}\lambda_2\right) \tag{3.26}$$

where

$$K_{p,q}^2(d^*) = \begin{cases} \frac{1}{2C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p = q \\ \frac{1}{C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p \neq q. \end{cases}$$

and $\tilde{B}(\mathrm{d}\lambda) = \left(\tilde{B}^{(1)}(\mathrm{d}\lambda), \dots, \tilde{B}^{(d)}(\mathrm{d}\lambda)\right)$ is an appropriate multivariate Hermitian-Gaussian random measure. Thus, we proved convergence in distribution of the sample-cross correlation matrix:

$$n^{1-2d^*}\left(\hat{\Gamma}_{U^*,n}(0) - \Gamma_{U^*}(0)\right) \xrightarrow{D} \left(Z_{2,d^*+1/2}^{(p,q)}(1)\right)_{p,q=1,\dots,d}.$$

We take a closer look at the covariance matrix of vec $\left(\hat{\Gamma}_{U^*,n}(0) - \Gamma_{U^*}(0)\right)$. Following [23], Lemma 5.7, we observe

$$\begin{aligned} &n^{1-2d^*} \left(4d^* \left(4d^* - 1\right)\right)^{1/2} \operatorname{Cov}\left(\operatorname{vec}\left(\hat{\Gamma}_{U^*,n}(0) - \Gamma_{U^*}(0)\right), \operatorname{vec}\left(\hat{\Gamma}_{U^*,n}(0) - \Gamma_{U^*}(0)\right)\right) \\ = &\left(I_{d^2} + K_{d^2}\right)\left(L_{U^*} \otimes L_{U^*}\right), \end{aligned}$$

with L_{U^*} as defined in (3.13) and $\otimes$ denotes the Kronecker product. Furthermore K_d denotes the commutation matrix that that transforms vec(A) into vec $\left(A^t\right)$ for $A \in \mathbb{R}^{d\times d}$. For details see [44].

Hence, the covariance matrix of the vector of the sample cross-covariances is fully specified by the knowledge of L_{U^*} as it arises in the context of long-range dependence in (3.13).

We obtain a relation between L and L_{U^*}, since

$$\Gamma_Y(\cdot) = \tilde{A}\Gamma_U(\cdot)\tilde{A}^t.$$

Both

$$\Gamma_Y(k) \simeq L k^{2d^*-1} \ (k \to \infty)$$

and

$$\Gamma_{U^*}(k) \simeq L_{U^*} k^{2d^*-1} \ (k \to \infty)$$

hold and we obtain

$$L = \tilde{A} L_{U^*} \tilde{A}^t.$$

We study the covariance matrix of vec $\left(\hat{\Gamma}_{Y,n}(0) - \Gamma_Y(0)\right)$:

$$\begin{aligned} &n^{1-2d^*} \left(4d^* \left(4d^* - 1\right)\right)^{1/2} \operatorname{Cov}\left(\operatorname{vec}\left(\hat{\Gamma}_{Y,n}(0) - \Gamma_Y(0)\right), \operatorname{vec}\left(\hat{\Gamma}_{Y,n}(0) - \Gamma_Y(0)\right)^t\right) \\ \to \quad &\left(I_{d^2} + K_{d^2}\right)\left(L \otimes L\right) \qquad (3.27) \\ = \quad &\left(I_{d^2} + K_{d^2}\right)\left(\tilde{A} L_{U^*} \tilde{A}^t\right) \otimes \left(\tilde{A} L_{U^*} \tilde{A}^t\right) \\ = \quad &\left(I_{d^2} + K_{d^2}\right)\left(\tilde{A} \otimes \tilde{A}\right) \cdot \left(L_{U^*} \otimes L_{U^*}\right) \cdot \left(\tilde{A}^t \otimes \tilde{A}^t\right). \end{aligned}$$

Let B_{U^*} be an upper triangular matrix, such that

$$B_{U^*} B_{U^*}^t := L_{U^*}.$$

We know that such a matrix exists because L_{U^*} is positive definite. Analogously, we define B_Y:

$$B_Y := \tilde{A} B_{U^*}.$$

Then, it holds that

$$B_Y B_Y^t = L.$$

We arrive at

$$
\begin{aligned}
& n^{1-2d^*}\left(C_2\right)^{-1/2}\left(B_Y \otimes B_Y\right)^{-1} \operatorname{vec}\left(\hat{\Gamma}_{Y,n}(0)-\Gamma_Y(0)\right) \\
& \stackrel{\mathcal{D}}{=} n^{1-2d^*}\left(C_2\right)^{-1/2}\left(B_{U^*} \otimes B_{U^*}\right)^{-1}(A \otimes A)^{-1} \operatorname{vec}\left(\tilde{A}\left(\hat{\Gamma}_{U^*,n}(0)-\Gamma_{U^*}(0)\right) \tilde{A}^t\right) \\
& = n^{1-2d^*}\left(C_2\right)^{-1/2}\left(B_{U^*} \otimes B_{U^*}\right)^{-1} \operatorname{vec}\left(\hat{\Gamma}_{U^*,n}(0)-\Gamma_{U^*}(0)\right) \\
& \xrightarrow{D} \operatorname{vec}\left(Z_{2,d^*+1/2}^{(p,q)}(1)\right)_{p,q=1,\ldots,d},
\end{aligned}
$$

where $Z_{2,d^*+1/2}^{(p,q)}(1)$ has the spectral domain representation

$$
Z_{2,d^*+1/2}^{(p,q)}(1)=K_{p,q}(d^*) \int_{\mathbb{R}^2}^{\prime\prime} \frac{\exp\left(i\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)}\left|\lambda_1 \lambda_2\right|^{-d^*} \tilde{B}^{(p)}\left(\mathrm{d}\lambda_1\right) \tilde{B}^{(q)}\left(\mathrm{d}\lambda_2\right)
$$

where

$$
K_{p,q}^2(d^*)=\begin{cases} \frac{1}{2C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p=q \\ \frac{1}{C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p\neq q. \end{cases}
$$

and $\tilde{B}(\mathrm{d}\lambda)=\left(\tilde{B}^{(1)}(\mathrm{d}\lambda),\ldots,\tilde{B}^{(d)}(\mathrm{d}\lambda)\right)$ is a multivariate Hermitian-Gaussian random measure as defined in (2.15). Note that the standardization on the left-hand side is appropriate since the covariance matrix of $\operatorname{vec}\left(Z_{2,d^*+1/2}(1)\right)$ is given by

$$
\begin{aligned}
\mathbb{E}\Bigg(K^2(d^*) \int_{\mathbb{R}^2}^{\prime\prime} \int_{\mathbb{R}^2}^{\prime\prime} E_{\lambda_1,\lambda_2} \overline{E_{\lambda_3,\lambda_4}} \operatorname{vec}\left(\tilde{B}\left(\mathrm{d}\lambda_1\right)\left(\tilde{B}\left(\mathrm{d}\lambda_2\right)\right)^t\right) & \\
\left(\operatorname{vec}\left(\overline{\tilde{B}\left(\mathrm{d}\lambda_3\right)\left(\tilde{B}\left(\mathrm{d}\lambda_4\right)\right)^t}\right)\right)^t\Bigg). & \qquad (3.28)
\end{aligned}
$$

by denoting

$$
E_{\lambda_1,\lambda_2}:=\frac{\exp\left(i\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)}\left|\lambda_1 \lambda_2\right|^{-d^*}.
$$

We observe

$$
\begin{aligned}
& \mathbb{E}\left(\operatorname{vec}\left(\tilde{B}\left(\mathrm{d}\lambda_1\right) \tilde{B}\left(\mathrm{d}\lambda_2\right)^t\right)\left(\operatorname{vec}\left(\overline{\tilde{B}\left(\mathrm{d}\lambda_3\right)\left(\tilde{B}\left(\mathrm{d}\lambda_4\right)\right)^t}\right)\right)^t\right) \\
& =\begin{cases} I_{d^2} \mathrm{d}\lambda_1 \mathrm{d}\lambda_2, & \left|\lambda_1\right|=\left|\lambda_3\right| \neq\left|\lambda_2\right|=\left|\lambda_4\right|, \\ K_{d^2} \mathrm{d}\lambda_1 \mathrm{d}\lambda_2, & \left|\lambda_1\right|=\left|\lambda_4\right| \neq\left|\lambda_2\right|=\left|\lambda_3\right|, \end{cases} \qquad (3.29)
\end{aligned}
$$

following [23], (27). Neither the case $|\lambda_1|=|\lambda_2|$ nor $|\lambda_3|=|\lambda_4|$ has to be incorporated as the diagonals are excluded in the integration in (3.28). □

Corollary 3.6 *Under the assumptions of Lemma 3.5, there is a different representation of the limit random vector. For $h\in\mathbb{N}$ we obtain*

$$
\left(n^{1-2d^*}\left(C_2\right)^{-1/2} \operatorname{vec}\left(\hat{\Gamma}_n(l)-\Gamma(l)\right),\ l=0,\ldots,h-1\right) \xrightarrow{D}\left(\operatorname{vec}\left(Z_{2,d^*+1/2}(1)\right)\ l=0,\ldots,h-1\right),
$$

where $\text{vec}\left(Z_{2,d^*+1/2}(1)\right)$ *has the spectral domain representation*

$$\text{vec}\left(Z_{2,d^*+1/2}(1)\right) = D_{K(d^*)} \int_{\mathbb{R}^2}^{\prime\prime} \frac{\exp\left(i\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)} \left|\lambda_1\lambda_2\right|^{-d^*} \text{vec}\left(\tilde{B}_L\left(\mathrm{d}\lambda_1\right)\tilde{B}_L\left(\mathrm{d}\lambda_2\right)^t\right).$$

The matrix $D_{K(d^*)}$ *is a diagonal matrix,*

$$D_{K(d^*)} = \text{diag}\left(\text{vec}\left(K\left(d^*\right)\right)\right),$$

and $K(d^*) = \left(K_{p,q}(d^*)\right)_{p,q=1,\ldots,d}$ *is such that*

$$K^2(d^*)_{p,q} = \begin{cases} \frac{1}{2C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p=q \\ \frac{1}{C_2(2\Gamma(1-2d^*)\sin(\pi d^*))^2}, & p\neq q. \end{cases}$$

Furthermore, $\tilde{B}_L(\mathrm{d}\lambda)$ *is a multivariate Hermitian-Gaussian random measure that fulfills*

$$\mathbb{E}\left(\tilde{B}_L(\mathrm{d}\lambda)\tilde{B}_L(\mathrm{d}\lambda)^*\right) = L\ \mathrm{d}\lambda.$$

Proof. The proof is an immediate consequence of Lemma 3.5 using $\tilde{B}_L(\mathrm{d}\lambda) = B_Y\tilde{B}(\mathrm{d}\lambda)$ with $B_YB_Y^t = L$ and $\tilde{B}(\mathrm{d}\lambda)$ as defined in (2.15). □

Remark 3.7 *The application of Theorem 6 of [2] in* (3.25) *coincides with the result in [45], Theorem 1.2A, who discussed the results of [2] in details. For the function in* (3.23) *the result is uncontroversial, as it holds that*

$$f\left(U_j^*\right) = \sum_{p=1}^{d} a_{p,p}H_2\left(U_j^{*,\ (p)}\right) + \sum_{p,q=1,p>q}^{d} a_{p,q}H_1\left(U_j^{*\ (p)}\right)H_1\left(U_j^{*\ (q)}\right)$$

and hence f *has the representation as required in [45], (1.4).*

Remark 3.8 *An analogous result of the convergence of the sample covariance matrix for different lags* $l = 0,\ldots,h-1$ *has been proved in [23], Theorem 4.1 for a multivariate linear process that may have both long-range and short-range dependent components that fulfill stronger long-range dependence conditions, which are given in* (2.6)*. In this context, the aforementioned matrix* B_Y *is the limit of the coefficients in the linear representation in* (2.1)*. The Gaussian assumption is not used in this article. There is an intersection with our framework. Each multivariate Gaussian process has a linear representation as described in* (2.1) *with Gaussian innovations. However, the stronger assumption in* (2.6) *is not fulfilled for each multivariate Gaussian process that is long-range dependent in the sense of Definition 2.13. Nevertheless, FARIMA*$(0,d,0)$ *processes as popular class of Gaussian processes in applications fulfill the stronger long-range dependence condition in* (2.6)*, see [37], Sec. 5.*

Remark 3.9 *The joint convergence in distribution of the sample covariance matrix to a matrix-valued Hermite-Rosenblatt process as defined in* (2.13) *for a bivariate Gaussian process was proved in [25], Theorem 6.2. Their result is consistent with Lemma 3.5 and on the one hand it confirms that the asymptotic distribution is not depending on the marginal covariance matrix*

Σ_2. *On the other hand, the strong impact of the matrix L in Definition 2.13 on the covariance structure is echoed in their theorem. In [25], Lemma 6.4, they provide the cumulants of* $Z^{(1,2)}(1)$ *given by the spectral domain representation*

$$Z^{(1,2)}_{2,d^*+1/2}(1) = K(d) \int''_{\mathbb{R}^2} \frac{\exp\left(i\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)} \left|\lambda_1\lambda_2\right|^{-d^*} \tilde{B}^{(1)}\left(\mathrm{d}\lambda_1\right) \tilde{B}^{(2)}\left(\mathrm{d}\lambda_2\right)$$

for a bivariate Hermitian-Gaussian random measure that fulfills

$$\mathbb{E}\left(\tilde{B}(\mathrm{d}\lambda)\tilde{B}(\mathrm{d}\lambda)^*\right) = \begin{pmatrix} \mathrm{d}\lambda & \rho\,\mathrm{d}\lambda \\ \rho\,\mathrm{d}\lambda & \mathrm{d}\lambda \end{pmatrix}$$

with $\rho \in [-1,1]$*:* $\kappa_{1,2}(1) = 0$ *and for* $k \geq 2$*:*

$$\kappa_{1,2}(k) = (2C_2)^{-k/2}\frac{(k-1)!}{2} \times \sum_{m_1=1}^{2} \cdots \sum_{m_k=1}^{2} \int_0^1 \cdots \int_0^1 h_{m_1,m_2}\left(s_1,s_2\right) h_{m_2,m_3}\left(s_2,s_3\right) \ldots h_{m_k,m_1}\left(s_k,s_1\right) \mathrm{d}s_1 \ldots \mathrm{d}s_k.$$

with

$$h_{12}\left(s_1,s_2\right) = h_{21}\left(s_1,s_2\right) = \left|s_1-s_2\right|^{2d^*-1} \text{ and } h_{11}\left(s_1,s_2\right) = h_{22}\left(s_1,s_2\right) = \rho\left|s_1-s_2\right|^{2d^*-1}.$$

The cumulants uniquely determine the distribution, since it has a representation in terms of a multiple Wiener-Itô integral of order 2*, see [25], p.106f. Only in the case* $\rho = 1$*, we observe the standard Rosenblatt distribution as described in* (2.17) *and in the case* $\rho = 0$*, we obtain*

$$\kappa_{1,2}(m) = 0, \quad m \in 2\mathbb{N}-1$$
$$\kappa_{1,2}(m) = (2C_2)^{-k/2}(k-1)!c_k, \quad m \in 2\mathbb{N}$$

with c_k *as defined in* (2.16)*.*

We are now fully equipped to turn to a limit theorem for a functional f with Hermite rank 2 applied to the long-range dependent Gaussian process $\left(Y_{j,h}\right)_{j\in\mathbb{Z}}$.

Theorem 3.10 *Let* $\left(Y_j\right)_{j\in\mathbb{Z}}$ *be a stationary Gaussian process as defined in* (3.1) *that fulfills* (3.2) *for* $d_p \in \left(\frac{1}{4},\frac{1}{2}\right)$*,* $p = 1\ldots,d$*. For* $h \in \mathbb{N}$ *we fix*

$$Y_{j,h} := \left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)},\ldots,Y_j^{(d)},\ldots,Y_{j+h-1}^{(d)}\right)^t \in \mathbb{R}^{dh}$$

with $Y_{j,h} \sim \mathcal{N}\left(0,\Sigma_{d,h}\right)$ *and* $\Sigma_{d,h}$ *as described in* (3.6)*. Let* $f : \mathbb{R}^{dh} \to \mathbb{R}$ *be a function with Hermite rank* 2 *such that the set of discontinuity points* D_f *is a Null set with respect to the dh-dimensional Lebesgue measure. Furthermore, we assume* f *fulfills* $\mathbb{E}\left(f^2\left(Y_{j,h}\right)\right) < \infty$*. Then,*

$$n^{-2d^*}(C_2)^{-\frac{1}{2}} \sum_{j=1}^{n} \left(f\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(d)}\right) - \mathbb{E}\left(f\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(d)}\right)\right)\right) \xrightarrow{\mathcal{D}} \sum_{p,q\in P^*} \tilde{\alpha}^{(p,q)} Z^{(p,q)}_{2,d^*+1/2}(1), \tag{3.30}$$

where

$$Z^{(p,q)}_{2,d^*+1/2}(1) = K_{p,q}\left(d^*\right) \int''_{\mathbb{R}^2} \frac{\exp\left(i\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)} \left|\lambda_1\lambda_2\right|^{-d^*} \tilde{B}^{(p)}_L\left(\mathrm{d}\lambda_1\right) \tilde{B}^{(q)}_L\left(\mathrm{d}\lambda_2\right).$$

The matrix $K\left(d^*\right)$ is given in Corollary 3.6. Moreover, $\tilde{B}_L(\mathrm{d}\lambda)$ is a multivariate Hermitian-Gaussian random measure with $\mathbb{E}\left(\tilde{B}_L(\mathrm{d}\lambda)\tilde{B}_L(\mathrm{d}\lambda)^*\right) = L\ \mathrm{d}\lambda$ and L as defined in (3.2)*. Furthermore, $C_2 := \frac{1}{2d^*(4d^*-1)}$ is a normalizing constant and*

$$\tilde{\alpha}^{(p,q)} := \sum_{i,k=1}^{h} \alpha^{(p,q)}_{i,k}$$

where $\alpha^{(p,q)}_{i,k} = \alpha_{i+(p-1)h,k+(q-1)h}$ for each $p,q \in P^*$ and $i,k = 1,\ldots,h$ and

$$\left(\alpha_{i,k}\right)_{1\le i,k\le dh} = \Sigma^{-1}_{d,h} C \Sigma^{-1}_{d,h}$$

where C denotes the matrix of second order Hermite coefficients, given by

$$C = \left(c_{i,k}\right)_{1\le i,k\le dh} = \mathbb{E}\left(Y_{1,h}\left(f\left(Y_{1,h}\right)-\mathbb{E}\left(f\left(Y_{1,h}\right)\right)\right)Y^t_{1,h}\right).$$

Proof. Without loss of generality, we assume $\mathbb{E}\left(f\left(Y_{j,h}\right)\right)=0$. Following the argumentation in [10], Theorem 5.9, we first remark that $Y_{j,h} \overset{\mathcal{D}}{=} AU_{j,h}$ with $U_{j,h}$ and A as described in (3.14) and (3.15). We now want to study the asymptotic behavior of the partial sum $\sum_{j=1}^{n} f^*\left(U_j\right)$ where $f^*\left(U_{j,h}\right) := f\left(AU_{j,h}\right) \overset{\mathcal{D}}{=} f\left(Y_{j,h}\right)$. Since $m\left(f^*, I_{dh}\right) = m\left(f\circ A, I_{dh}\right) = m\left(f,\Sigma_{d,h}\right) = 2$, see Lemma 2.32, hence, we know by [2], Theorem 6, that these partial sums are dominated by the second order terms in the Hermite expansion of f^*:

$$\sum_{j=1}^{n} f^*\left(U_{j,h}\right) \sum_{j=1}^{n} \sum_{l_1+\ldots+l_{dh}=2} \mathbb{E}\left(f^*\left(U_{j,h}\right) H_{l_1,\ldots,l_{dh}}\left(U_{j,h}\right)\right) H_{l_1,\ldots,l_{dh}}\left(U_{j,h}\right) + o_{\mathbb{P}}\left(n^{2d^*}\right).$$

This follows from the multivariate extension of the Reduction Theorem as proved in [2]. We obtain

$$\begin{aligned}
&\sum_{l_1+\ldots+l_{dh}=2} \mathbb{E}\left(f^*\left(U_{j,h}\right) H_{l_1,\ldots,l_{dh}}\left(U_{j,h}\right)\right) H_{l_1,\ldots,l_{dh}}\left(U_{j,h}\right) \\
=&\sum_{i=1}^{dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)\left(\left(U^{(i)}_{j,h}\right)^2-1\right)\right)\left(\left(U^{(i)}_{j,h}\right)^2-1\right) + \sum_{1\le i,k\le dh, i\ne k} \mathbb{E}\left(f^*\left(U_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h} \\
=&\sum_{i=1}^{dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)\left(U^{(i)}_{j,h}\right)^2\right)\left(\left(U^{(i)}_{j,h}\right)^2-1\right) + \sum_{1\le i,k\le dh, i\ne k} \mathbb{E}\left(f^*\left(U_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h},
\end{aligned}$$

since $\mathbb{E}\left(f^*\left(U_{j,h}\right)\right) = \mathbb{E}\left(f\left(Y_{j,h}\right)\right) = 0$. This results in:

$$\begin{aligned}
&\sum_{i=1}^{dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)\left(U^{(i)}_{j,h}\right)^2\right)\left(\left(U^{(i)}_{j,h}\right)^2-1\right) + \sum_{1\le i,k\le dh, i\ne k} \mathbb{E}\left(f^*\left(U_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h} \\
&= \sum_{1\le i,k\le dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h}\right)U^{(i)}_{j,h}U^{(k)}_{j,h} - \sum_{i=1}^{dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)\left(U^{(i)}_{j,h}\right)^2\right). \qquad (3.31)
\end{aligned}$$

Note that

$$\sum_{1\le i,k\le dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)U_{j,h}^{(i)}U_{j,h}^{(k)}\right)\mathbb{E}\left(U_{j,h}^{(i)}U_{j,h}^{(k)}\right) = \sum_{i=1}^{dh}\mathbb{E}\left(f^*\left(U_{j,h}\right)\left(U_{j,h}^{(i)}\right)^2\right) \tag{3.32}$$

since the entries of $U_{j,h}$ are independent for fixed j and identically $\mathcal{N}(0,1)$ distributed. So the subtrahend in (3.31) equals the expected value of the minuend.
Define $B := (b_{i,k})_{1\le i,k\le dh} \in \mathbb{R}^{(dh)\times(dh)}$ with $b_{i,k} := \mathbb{E}\left(f^*\left(U_{j,h}\right)U_{j,h}^{(i)}U_{j,h}^{(k)}\right) = \mathbb{E}\left(f^*\left(U_1\right)U_1^{(i)}U_1^{(k)}\right)$ since we are considering a stationary process. We obtain

$$B = \mathbb{E}\left(U_{j,h}f^*\left(U_{j,h}\right)U_{j,h}^t\right) = \mathbb{E}\left(A^{-1}Y_{j,h}f\left(Y_{j,h}\right)Y_{j,h}^t\left(A^{-1}\right)^t\right).$$

Hence, we can state the following:

$$\begin{aligned}
\sum_{1\le i,k\le dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)U_{j,h}^{(i)}U_{j,h}^{(k)}\right)U_{j,h}^{(i)}U_{j,h}^{(k)} &= U_{j,h}^t B U_{j,h} \\
&\overset{\mathcal{D}}{=} Y_{j,h}^t\left(A^{-1}\right)^t B A^{-1} Y_{j,h} \\
&= Y_{j,h}^t\left(A^{-1}\right)^t A^{-1}\mathbb{E}\left(Y_{j,h}f\left(Y_{j,h}\right)Y_{j,h}^t\right)\left(A^{-1}\right)^t A^{-1}Y_{j,h} \\
&= Y_{j,h}^t\Sigma_{d,h}^{-1}\mathbb{E}\left(Y_{j,h}f\left(Y_{j,h}\right)Y_{j,h}^t\right)\Sigma_{d,h}^{-1}Y_{j,h} \\
&= Y_{j,h}^t\mathbb{A}Y_{j,h} \\
&= \sum_{1\le i,k\le dh} Y_j^{(i)}Y_j^{(k)}\alpha_{ik},
\end{aligned} \tag{3.33}$$

where we define $\mathbb{A} := (\alpha_{ik})_{1\le ik\le dh} := \Sigma_{d,h}^{-1}C\Sigma_{d,h}^{-1}$, with $C := \mathbb{E}\left(Y_{j,h}f\left(Y_{j,h}\right)Y_{j,h}^t\right)$ as the matrix of second order Hermite coefficients, in contrast to B now with respect to the original considered process $\left(Y_{j,h}\right)_{j\in\mathbb{Z}}$.

Remembering the special structure of $Y_{j,h} = \left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)},\ldots,Y_j^{(d)},Y_{j+h-1}^{(d)}\right)^t$, namely that $Y_{j,h}^{(k)} = Y_{j+(k \mod h)-1}^{\left(\lfloor\frac{k-1}{h}\rfloor+1\right)}$, $k=1,\ldots,dh$ we can see that

$$\begin{aligned}
\sum_{j=1}^n\sum_{1\le ik\le dh} Y_{j,h}^{(i)}Y_{j,h}^{(k)}\alpha_{ik} &= \sum_{j=1}^n\sum_{1\le ik\le dh} Y_{j+(i \mod h)-1}^{\left(\lfloor\frac{i-1}{h}\rfloor+1\right)}Y_{j+(k \mod h)-1}^{\left(\lfloor\frac{k-1}{h}\rfloor+1\right)}\alpha_{ik} \\
&= \sum_{j=1}^n\sum_{p,q=1}^d\sum_{i,k=1}^h Y_{j+i-1}^{(p)}Y_{j+k-1}^{(q)}\alpha_{ik}^{(p,q)},
\end{aligned} \tag{3.34}$$

where we divide

$$\mathbb{A} = \begin{pmatrix} \mathbb{A}^{(1,1)} & \mathbb{A}^{(1,2)} & \ldots & \mathbb{A}^{(1,d)} \\ \mathbb{A}^{(2,1)} & \mathbb{A}^{(2,2)} & \ldots & \mathbb{A}^{(2,d)} \\ \vdots & \vdots & \vdots & \\ \mathbb{A}^{(d,1)} & \mathbb{A}^{(d,2)} & \ldots & \mathbb{A}^{(d,d)} \end{pmatrix},$$

with $\mathbb{A}^{(p,q)} = \left(\alpha_{i,k}^{(p,q)}\right)_{1\le i,k\le h} \in \mathbb{R}^{h\times h}$ such that $\alpha_{i,k}^{(p,q)} = \alpha_{i+(p-1)h,k+(q-1)h}$ for each $p,q=1,\ldots,d$ and $i,k=1,\ldots,h$.

We can now split the considered sum in (3.34) in a way such that we are able to express it in terms of sample cross-covariances afterwards. In order to do so, we define the sample cross-covariance at lag l by

$$\hat{r}_n^{(p,q)}(l) := \frac{1}{n}\sum_{j=1}^{n-l} X_j^{(p)} X_{j+l}^{(q)}$$

for $p, q = 1, \ldots, d$.
Note that in the case $h = 1$, it follows directly that

$$\sum_{j=1}^{n}\sum_{p,q=1}^{d}\sum_{i,k=1}^{h} Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)} \alpha_{ik}^{(p,q)} = \sum_{p,q=1}^{d} \alpha_{1,1}^{(p,q)} \sum_{j=1}^{n} Y_j^{(p)} Y_j^{(q)} = n \sum_{p,q=1}^{d} \hat{r}_n^{(p,q)}(0).$$

The case $h = 2$ has to be regarded separately, too, and we obtain

$$\begin{aligned}
&\sum_{j=1}^{n}\sum_{p,q=1}^{d}\sum_{i,k=1}^{2} Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)} \alpha_{ik}^{(p,q)} \\
&= \sum_{p,q=1}^{d} \left(\alpha_{1,1}^{(p,q)} \sum_{j=1}^{n} Y_j^{(p)} Y_j^{(q)} + \alpha_{1,2}^{(p,q)} \sum_{j=1}^{n} Y_j^{(p)} Y_{j+1}^{(q)} + \alpha_{2,1}^{(p,q)} \sum_{j=1}^{n} Y_{j+1}^{(p)} Y_j^{(q)} + \alpha_{2,2}^{(p,q)} \sum_{j=1}^{n} Y_{j+1}^{(p)} Y_{j+1}^{(q)} \right) \\
&= \sum_{p,q=1}^{d} \left(\alpha_{1,1}^{(p,q)} n \hat{r}_n^{(p,q)}(0) + \alpha_{1,2}^{(p,q)} \left(n \hat{r}_n^{(p,q)}(1) + \underbrace{Y_n^{(p)} Y_{n+1}^{(q)}}_{\bigstar} \right) + \alpha_{2,1}^{(p,q)} \left(n \hat{r}_n^{(q,p)}(1) + \underbrace{Y_{n+1}^{(p)} Y_n^{(q)}}_{\bigstar} \right) \right. \\
&\qquad \left. + \alpha_{2,2}^{(p,q)} \left(n \hat{r}_n^{(p,q)}(0) + \underbrace{Y_{n+1}^{(p)} Y_{n+1}^{(q)}}_{\bigstar} - \underbrace{Y_1^{(p)} Y_1^{(q)}}_{\bigstar} \right) \right),
\end{aligned}$$

Note that for each of the terms labeled by $\bigstar$ the following holds for $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$:

$$n^{-2d^*} \bigstar \xrightarrow{\mathbb{P}} 0, \ (n \to \infty).$$

We use this property later on when dealing with the asymptotics of the term in (3.34).
Finally we consider the term in (3.34) for $h \geq 3$ and arrive at

$$\begin{aligned}
&\sum_{j=1}^{n}\sum_{p,q=1}^{d}\sum_{i,k=1}^{h} Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)} \alpha_{ik}^{(p,q)} \\
&= \sum_{p,q=1}^{d}\sum_{i,k=1}^{h} \alpha_{ik}^{(p,q)} \sum_{j=i}^{n+i-1} Y_j^{(p)} Y_{j+k-i}^{(q)} \\
&= \sum_{p,q=1}^{d}\sum_{l=0}^{h-1}\sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \sum_{j=i}^{n+i-1} Y_j^{(p)} Y_{j+l}^{(q)} \\
&\qquad + \sum_{p,q=1}^{d}\sum_{l=-(h-1)}^{-1}\sum_{i=1-l}^{h} \alpha_{i,i+l}^{(p,q)} \sum_{j=i}^{n+i-1} Y_j^{(p)} Y_{j+l}^{(q)} \\
&= \sum_{p,q=1}^{d}\sum_{l=0}^{h-1}\sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \sum_{j=i}^{n+i-1} Y_j^{(p)} Y_{j+l}^{(q)}
\end{aligned}$$

$$
+\sum_{p,q=1}^{d}\sum_{l=1}^{h-1}\sum_{i=1}^{h-l}\alpha_{i+l,i}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j+l}^{(p)}Y_{j}^{(q)} \tag{3.35}
$$

$$
=\sum_{p,q=1}^{d}\sum_{i=1}^{h}\alpha_{i,i}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j}^{(p)}Y_{j}^{(q)}
$$

$$
+\sum_{p,q=1}^{d}\sum_{l=1}^{h-1}\sum_{i=1}^{h-l}\left(\alpha_{i,i+l}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j}^{(p)}Y_{j+l}^{(q)}+\alpha_{i+l,i}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j+l}^{(p)}Y_{j}^{(q)}\right) \tag{3.36}
$$

$$
=\sum_{p,q=1}^{d}\left(\alpha_{1,1}^{(p,q)}\sum_{j=1}^{n}Y_{j}^{(p)}Y_{j}^{(q)}+\sum_{i=2}^{h}\alpha_{i,i}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j}^{(p)}Y_{j}^{(q)}\right)
$$

$$
+\sum_{p,q=1}^{d}\sum_{l=1}^{h-2}\Bigg(\left(\alpha_{1,1+l}^{(p,q)}\sum_{j=1}^{n}Y_{j}^{(p)}Y_{j+l}^{(q)}+\alpha_{1+l,1}^{(p,q)}\sum_{j=1}^{n}Y_{j+l}^{(p)}Y_{j}^{(q)}\right)
$$

$$
+\sum_{i=2}^{h-l}\left(\alpha_{i,i+l}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j}^{(p)}Y_{j+l}^{(q)}+\alpha_{i+l,i}^{(p,q)}\sum_{j=i}^{n+i-1}Y_{j+l}^{(p)}Y_{j}^{(q)}\right)\Bigg)
$$

$$
+\sum_{p,q=1}^{d}\left(\alpha_{1,h}^{(p,q)}\sum_{j=1}^{n}Y_{j}^{(p)}Y_{j+h-1}^{(q)}+\alpha_{h,1}^{(p,q)}Y_{j+h-1}^{(p)}Y_{j}^{(q)}\right) \tag{3.37}
$$

$$
=\sum_{p,q=1}^{d}\left(\alpha_{1,1}^{(p,q)}n\hat{r}_{n}^{(p,q)}(0)+\sum_{i=2}^{h}\alpha_{i,i}^{(p,q)}\left(\underbrace{\sum_{j=n+1}^{n+i-1}Y_{j}^{(p)}Y_{j}^{(q)}}_{\bigstar}+n\hat{r}_{n}^{(p,q)}(0)-\underbrace{\sum_{j=1}^{i-1}Y_{j}^{(p)}Y_{j}^{(q)}}_{\bigstar}\right)\right)
$$

$$
+\sum_{p,q=1}^{d}\sum_{l=1}^{h-2}\Bigg(\left(\alpha_{1,1+l}^{(p,q)}n\hat{r}_{n}^{(p,q)}(l)+\alpha_{1+l,1}^{(p,q)}n\hat{r}_{n}^{(q,p)}(l)\right)
$$

$$
+\sum_{i=2}^{h-l}\Bigg(\alpha_{i,i+l}^{(p,q)}\left(\underbrace{\sum_{j=n-l+1}^{n+i-1}Y_{j}^{(p)}Y_{j+l}^{(q)}}_{\bigstar}+n\hat{r}_{n}^{(p,q)}(l)-\underbrace{\sum_{j=1}^{i-1}Y_{j}^{(p)}Y_{j+l}^{(q)}}_{\bigstar}\right)
$$

$$
+\alpha_{i+l,i}^{(p,q)}\left(\underbrace{\sum_{j=n-l+1}^{n+i-1}Y_{j+l}^{(p)}Y_{j}^{(q)}}_{\bigstar}+n\hat{r}_{n}^{(q,p)}(l)-\underbrace{\sum_{j=1}^{i-1}Y_{j+l}^{(p)}Y_{j}^{(q)}}_{\bigstar}\right)\Bigg)\Bigg)
$$

$$
+\sum_{p,q=1}^{d}\Bigg(\alpha_{1,h}^{(p,q)}\left(\underbrace{\sum_{j=n-h+2}^{n}Y_{j}^{(p)}Y_{j+h-1}^{(q)}}_{\bigstar}+n\hat{r}_{n}^{(p,q)}(h-1)\right) \tag{3.38}
$$

$$
+\alpha_{h,1}^{(p,q)}\left(\underbrace{\sum_{j=n-h+2}^{n}Y_{j+h-1}^{(p)}Y_{j}^{(q)}}_{\bigstar}+n\hat{r}_{n}^{(q,p)}(h-1)\right)\Bigg). \tag{3.39}
$$

In order to illustrate the calculations up to (3.35), we briefly consider the following example with $h = 3$, $n = 5$ and fixed $p, q \in \{1, \dots, d\}$.
Starting with

$$
\begin{aligned}
&\sum_{i,k=1}^{3} \alpha_{ik}^{(p,q)} \sum_{j=i}^{5+i-1} Y_j^{(p)} Y_{j+k-i}^{(q)} \\
=&\alpha_{1,1}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_j^{(q)} + \alpha_{1,2}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_{j+1}^{(q)} + \alpha_{1,3}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_{j+2}^{(q)} \\
&+ \alpha_{2,1}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_{j-1}^{(q)} + \alpha_{2,2}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_j^{(q)} + \alpha_{2,3}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_{j+1}^{(q)} \\
&+ \alpha_{3,1}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_{j-2}^{(q)} + \alpha_{3,2}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_{j-1}^{(q)} + \alpha_{3,3}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_j^{(q)} \\
=&\alpha_{1,1}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_j^{(q)} + \alpha_{2,2}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_j^{(q)} + \alpha_{3,3}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_j^{(q)} \\
&+ \alpha_{1,2}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_{j+1}^{(q)} + \alpha_{2,3}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_{j+1}^{(q)} \\
&+ \alpha_{1,3}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_{j+2}^{(q)} \\
&+ \alpha_{2,1}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_{j-1}^{(q)} + \alpha_{3,2}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_{j-1}^{(q)} \\
&+ \alpha_{3,1}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_{j-2}^{(q)} \\
=&\alpha_{1,1}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_j^{(q)} + \alpha_{2,2}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_j^{(q)} + \alpha_{3,3}^{(p,q)} \sum_{j=3}^{7} Y_j^{(p)} Y_j^{(q)} \\
&+ \alpha_{1,2}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_{j+1}^{(q)} + \alpha_{2,3}^{(p,q)} \sum_{j=2}^{6} Y_j^{(p)} Y_{j+1}^{(q)} \\
&+ \alpha_{1,3}^{(p,q)} \sum_{j=1}^{5} Y_j^{(p)} Y_{j+2}^{(q)} \\
&+ \alpha_{2,1}^{(p,q)} \sum_{j=1}^{5} Y_{j+1}^{(p)} Y_j^{(q)} + \alpha_{3,2}^{(p,q)} \sum_{j=2}^{6} Y_{j+1}^{(p)} Y_j^{(q)} \\
&+ \alpha_{3,1}^{(p,q)} \sum_{j=1}^{5} Y_{j+2}^{(p)} Y_j^{(q)} \\
=&\sum_{l=0}^{2} \sum_{i=1}^{3-l} \alpha_{i,i+l}^{(p,q)} \sum_{j=i}^{5+i-1} Y_j^{(p)} Y_{j+l}^{(q)} + \sum_{l=1}^{2} \sum_{i=1}^{3-l} \alpha_{i+l,i}^{(p,q)} \sum_{j=i}^{5+i-1} Y_{j+l}^{(p)} Y_j^{(q)}.
\end{aligned}
$$

For the calculations in (3.36) it is necessary to separate the case $l = 0$ and for those in (3.37) to separate the cases $i = 1$ and $l = h - 1$ from the rest of the corresponding sums in order to get a precise differentation between the sample cross-correlations and the terms denoted by ★ in (3.39).

Again for each of the terms labeled by ★ it holds for $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$:

$$n^{-2d^*} \bigstar \xrightarrow{\mathbb{P}} 0, \ (n \to \infty),$$

since each ★ describes a sum with a finite number (independent of n) of summands. Therefore, we continue to express the terms denoted by ★ by $o_{\mathbb{P}}\left(n^{2d^*}\right)$.
With these calculations we are able to re-express the partial sum, whose asymptotics we are interested in, in terms of the sample cross-correlations of the original long-range dependent process $(Y_j)_{j\in\mathbb{Z}}$.
Finally, the previous calculations lead to

$$\begin{aligned}
&\sum_{j=1}^{n} f\left(Y_{j,h}\right) \\
\overset{\mathcal{D}}{=}\ & \sum_{j=1}^{n} f^*\left(U_{j,h}\right) \\
\overset{(3.31)}{=} & \sum_{j=1}^{n} \left(\sum_{1\le i,k\le dh} \mathbb{E}\left(f^*\left(U_{j,h}\right) U_{j,h}^{(i)} U_{j,h}^{(k)}\right) U_{j,h}^{(i)} U_{j,h}^{(k)} - \sum_{i=1}^{dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)\left(U_{j,h}^{(i)}\right)^2\right)\right) + o_{\mathbb{P}}(n^{2d^*}) \\
\underset{(3.34)}{\overset{\mathcal{D}}{=}} & \sum_{j=1}^{n} \sum_{p,q=1}^{d} \sum_{i,k=1}^{h} \alpha_{ik}^{(p,q)} \left(Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)} - \mathbb{E}\left(Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)}\right)\right) + o_{\mathbb{P}}(n^{2d^*}),
\end{aligned} \tag{3.40}$$

where (3.40) follows, since (3.32) yields

$$\begin{aligned}
\sum_{i=1}^{dh} \mathbb{E}\left(f^*\left(U_{j,h}\right)\left(U_{j,h}^{(i)}\right)^2\right) = & \sum_{1\le i,k\le dh} \mathbb{E}\left(f^*\left(U_{j,h}\right) U_{j,h}^{(i)} U_{j,h}^{(k)}\right) \mathbb{E}\left(U_{j,h}^{(i)} U_{j,h}^{(k)}\right) \\
\overset{(3.34)}{=} & \sum_{p,q=1}^{d} \sum_{i,k=1}^{h} \alpha_{ik}^{(p,q)} \mathbb{E}\left(Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)}\right).
\end{aligned}$$

Taking the parts containing the sample cross-correlations into account, we derive

$$\begin{aligned}
&\sum_{j=1}^{n} \sum_{p,q=1}^{d} \sum_{i,k=1}^{h} \alpha_{ik}^{(p,q)} \left(Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)} - \mathbb{E}\left(Y_{j+i-1}^{(p)} Y_{j+k-1}^{(q)}\right)\right) + o_{\mathbb{P}}(n^{2d^*}) \\
\overset{(3.39)}{=} & \sum_{p,q=1}^{d} \left(\alpha_{1,1}^{(p,q)} n \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) + \sum_{i=2}^{h} \alpha_{i,i}^{(p,q)} n \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right)\right) \\
&+ \sum_{p,q=1}^{d} \sum_{l=1}^{h-2} \Bigg(\left(\alpha_{1,1+l}^{(p,q)} n \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l)\right) + \alpha_{1+l,1}^{(p,q)} n \left(\hat{r}_n^{(q,p)}(l) - r^{(q,p)}(l)\right)\right) \\
&\qquad + \sum_{i=2}^{h-l} \left(\alpha_{i,i+l}^{(p,q)} n \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l)\right) + \alpha_{i+l,i}^{(p,q)} n \left(\hat{r}_n^{(q,p)}(l) - r^{(q,p)}(l)\right)\right)\Bigg) \\
&+ \sum_{p,q=1}^{d} \left(\alpha_{1,h}^{(p,q)} n \left(\hat{r}_n^{(p,q)}(h-1) - r^{(p,q)}(h-1)\right) + \alpha_{h,1}^{(p,q)} n \left(\hat{r}_n^{(q,p)}(h-1) - r^{(q,p)}(h-1)\right)\right)
\end{aligned} \tag{3.41}$$

$$+ o_{\mathbb{P}}(n^{2d^*})$$

$$= n \sum_{p,q=1}^{d} \left(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right) + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q)} \left(\hat{r}_n^{(q,p)}(l) - r^{(q,p)}(l) \right) \right) \\ + o_{\mathbb{P}}(n^{2d^*}). \tag{3.42}$$

We take a closer look at the impact of each long-range dependence parameter d_p, $p = 1, \dots, d$ to the convergence of this sum. The setting we are considering does not allow for a normalization depending on p and q for each cross-correlation $\left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right)$, $l = 0, \dots, h-1$ but we need to find a normalization for all $p, q = 1, \dots, d$. Hence, we need to remember the set $P^* := \{p \in \{1, \dots, d\} : \ d_p \geq d_q \forall q \in \{1, \dots, d\}\}$ and the parameter $d^* = \max_{p=1,\dots,d} d_p$, such that for each $p \in P^*$ we have $d_p = d^*$. For each $p, q \in \{1, \dots, d\}$ with $(p, q) \notin P^* \times P^*$ and $l = 0, \dots, h-1$, we conclude that

$$\begin{aligned} \mathbb{E}\left(\left(n^{1-2d^*} \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right) \right)^2 \right) &= n^{2(d_p+d_q-2d^*)} \mathbb{E}\left(\left(n^{1-d_p-d_q} \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right) \right)^2 \right) \\ &= n^{2d_p+2d_q-4d^*} C_2 \left(L_{p,p} L_{q,q} + L_{p,q} L_{q,p} \right) \\ &\xrightarrow{(n \to \infty)} 0, \end{aligned} \tag{3.43}$$

since $d_p + d_q - 2d^* < 0$.
This implies that

$$n^{1-2d^*} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0) \right) \xrightarrow{\mathbb{P}} 0$$

Hence, using Slutsky's theorem, the crucial parameters that determine the normalization and, therefore, the limit distribution of (3.44) are given in P^*. We have an equal long-range dependence parameter d^* to regard for all $p \in P^*$. Applying Lemma 3.5, we obtain the following, by using the symmetry in $l = 0$ of the cross correlation function $r^{(p,q)}(0) = r^{(q,p)}(0)$ for $p, q \in P^*$:

$$\begin{aligned} &\sum_{p,q=1}^{d} \left(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right) + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q)} \left(\hat{r}_n^{(q,p)}(l) - r^{(q,p)}(l) \right) \right) \\ &= \sum_{p,q \in P^*} \left(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0) \right) + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q)} \left(\hat{r}_n^{(q,p)}(0) - r^{(q,p)}(0) \right) \right) + o_{\mathbb{P}}(n^{2d^*-1}) \\ &= \sum_{p,q \in P^*} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0) \right) \left(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q)} \right) + o_{\mathbb{P}}(n^{2d^*}) \\ &= \sum_{p,q \in P^*} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0) \right) \left(\sum_{i,k=1}^{h} \alpha_{i,k}^{(p,q)} \right) + o_{\mathbb{P}}(n^{2d^*}) \\ &= \sum_{p,q \in P^*} \tilde{\alpha}^{(p,q)} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0) \right) + o_{\mathbb{P}}(n^{2d^*}), \end{aligned} \tag{3.44}$$

by defining $\tilde{\alpha}^{(p,q)} := \sum_{i,k=1}^{h} \alpha_{i,k}^{(p,q)}$.

At first glance it might be interesting to transform the term in (3.44) using

$$\sum_{p,q=1}^{d} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) = \frac{1}{n}\sum_{j=1}^{n} \left(\sum_{p=1}^{d} Y_j^{(p)}\right)^2 - \mathbb{E}\left(\sum_{p=1}^{d} Y_j^{(p)}\right)^2.$$

This would be a huge simplification, since we would be able to boil down the convergence problem to the second Hermite polynomial of a (suitably normalized) one-dimensional Gaussian process, as in the case of Hermite rank 1, see (3.21).
However, this approach is not reasonable, as we can see in Example 4.33, since for $d = 2$ and $h = 1$, we obtain $\tilde{\alpha}^{(1,1)}\tilde{\alpha}^{(2,2)} \neq \left(\tilde{\alpha}^{(1,2)}\right)^2$ and therefore, we are not able to simplify the term in (3.44) to the square of the sum weighted by $\tilde{\alpha}^{(p,q)}$ as it is indicated in the idea above. Therefore, we have to continue with the convergence results of cross-covariances in the d-dimensional case.

Applying the continuous mapping theorem given in [65], Theorem 2.3 to the result in Corollary 3.6 we arrive at

$$\begin{aligned} n^{-2d^*}\left(C_2\right)^{-1/2}\sum_{j=1}^{n} f\left(Y_{j,h}\right) &= n^{-2d^*}\left(n\sum_{p,q=1}^{d}\tilde{\alpha}^{(p,q)}\left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) + o_{\mathbb{P}}(n^{2d^*})\right) \\ &= n^{1-2d^*}\left(C_2\right)^{-1/2}\sum_{p,q=1}^{d}\tilde{\alpha}^{(p,q)}\left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) + o_{\mathbb{P}}(1) \\ &\xrightarrow{\mathcal{D}} \sum_{p,q\in P^*}\tilde{\alpha}^{(p,q)} Z_{2,d^*+1/2}^{(p,q)}(1), \end{aligned}$$

where

$$Z_{2,d^*+1/2}^{(p,q)}(1) = K_{p,q}\left(d^*\right)\int_{\mathbb{R}^2}^{\prime\prime} \frac{\exp\left(i\left(\lambda_1+\lambda_2\right)\right)-1}{i\left(\lambda_1+\lambda_2\right)} \left|\lambda_1\lambda_2\right|^{-d^*} \tilde{B}_L^{(p)}\left(\mathrm{d}\lambda_1\right)\tilde{B}_L^{(q)}\left(\mathrm{d}\lambda_2\right).$$

The matrix $K\left(d^*\right)$ is given in Corollary 3.6. Moreover, $\tilde{B}_L(\mathrm{d}\lambda)$ is a multivariate Hermitian-Gaussian random measure with $\mathbb{E}\left(B_L(\mathrm{d}\lambda)B_L(\mathrm{d}\lambda)^*\right) = L\ \mathrm{d}\lambda$ and L as defined in (3.2). □

It is even possible to extend this result to multivariate Gaussian processes $\left(Y_j\right)_{j\in\mathbb{Z}}$, that have long-range dependent, as well as short-range dependent components. Note that the assumption in Theorem 3.10, that for all $p = 1\ldots, d$ we demand $d_p \in \left(\frac{1}{4}, \frac{1}{2}\right)$ is necessary to ensure

$$g^{(p,q)}\left(Y_j^{(1)},\ldots,Y_j^{(d)}\right) := \frac{1}{n}\sum_{j=1}^{n} Y_j^{(p)}Y_j^{(q)} - r^{(p,q)}(0)$$

to be still long-range dependent.
However, we may allow for long-range dependence parameters $d_p + d_q \in \left(0, \frac{1}{2}\right)$, such that the transformed process $g^{(p,q)}\left(Y_j^{(1)},\ldots,Y_j^{(d)}\right)$ is short-range dependent, or even for parameters $d_p + d_q \in (-\infty, 0)$ such that at least one of the one-dimensional processes $\left(Y_j^{(p)}\right)_{j\in\mathbb{Z}}$ or $\left(Y_j^{(q)}\right)_{j\in\mathbb{Z}}$ is short-range dependent from the very first. Therefore, we may soften the conditions of Theorem 3.10 in the following corollary:

Corollary 3.11 *Instead of demanding in the assumptions of Theorem 3.10 that* (3.2) *holds for* $(Y_j)_{j\in\mathbb{Z}}$ *with the addition that for all* $p = 1, \ldots, d$ *we have* $d_p \in \left(\frac{1}{4}, \frac{1}{2}\right)$*, we may use the following condition:*
We assume that

$$r^{(p,q)}(k) = k^{d_p+d_q-1} L_{p,q}(k) \quad (k \to \infty)$$

with $L_{p,q}(k)$ *as given in* (3.2)*, but we do no longer assume* $d_p \in \left(\frac{1}{4}, \frac{1}{2}\right)$ *for all* $p = 1, \ldots, d$ *but soften the assumption to* $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$ *and for* $d_p \neq d^*$*,* $p = 1, \ldots, d$ *we allow for* $d_p \in (-\infty, 0) \cup \left(0, \frac{1}{4}\right]$*. Then, the statement of Theorem 3.10 remains valid.*

Proof. We assumed $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$, because otherwise we leave the long-range dependent setting, since we are studying functionals with Hermite rank 2 and the transformed process would no longer be long-range dependent and limit theorems for functionals of short-range dependent processes would hold, see Theorem 4 in [2]. This choice of d^* assures that the multivariate generalization of the Reduction theorem as it is used in the proof of Theorem 3.10 still holds for these softened assumptions, as explained in (3.8) and (3.9).
We turn to the asymptotics of $g^{(p,q)}(Y_j)$. We obtain for all $p, q \in \{1, \ldots, d\} \setminus P^*$,i.e., excluding $d_p = d_q = d^*$ and for all $l = 0, \ldots, h-1$ as in (3.43), that

$$\begin{aligned}\mathbb{E}\left(\left(n^{1-2d^*}\left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l)\right)\right)^2\right) &= n^{2(d_p+d_q-2d^*)}\mathbb{E}\left(\left(n^{1-d_p-d_q}\left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l)\right)\right)^2\right)\\ &= n^{2d_p+2d_q-4d^*}C_2\left(L_{p,p}L_{q,q} + L_{p,q}L_{q,p}\right)\\ &\xrightarrow{(n\to\infty)} 0, \end{aligned} \tag{3.45}$$

since $d_p + d_q - 2d^* < 0$.
This implies that

$$n^{1-2d^*}\left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) \xrightarrow{\mathbb{P}} 0.$$

Applying Slutsky's theorem, we observe that only $p, q \in P^*$ have an impact on the convergence behaviour as it is given in (3.44) and hence, the result in Theorem 3.10 holds. □

Remark 3.12 *The joint asymptotic distribution of the sample-covariance matrix of a multivariate linear process* $(Y_j)_{j\in\mathbb{Z}}$ *in the mixed case of short- and long-range dependent components in the sense of* (2.6) *and* (2.7) *is in proved in Theorem 4.1 in [23]. There, the set* $\{1, \ldots, d\}^2$ *is split into two sets*

$$P_1 = \left\{p, q) \in \{1, \ldots, d\} : \quad d_p + d_q \in \left(\frac{1}{2}, 1\right)\right\},$$
$$P_2 = \left\{p, q) \in \{1, \ldots, d\} : \quad d_p + d_q \in \left(-\infty, \frac{1}{2}\right)\right\},$$

such that P_1 *expresses long-range dependence and results in convergence of the sample-covariance matrix to the matrix-valued Rosenblatt process, whereas for parameters* $(p, q) \in P_2$ *it holds that*

$\left(r^{(p,q)}\right)^2$ *is absolutely summable and the limiting process is a Brownian motion.*
Note that in (3.45) *we explicitly include that there exists* $p \in \{1, \dots, d\}$, *such that* $d_p = \frac{1}{4}$, *but* $d^* \neq \frac{1}{4}$. *For limit theorems on this case it deserves special attention, since the convergence rate of the variance changes, for details, see [18], p.58 and [32], Theorem 4.*

Applying the result of Theorem 3.10 we can see that in the case $\#P^* = 1$, we only obtain one standard Rosenblatt random variable. In the case $\#P^* = 2$, which turns out to be most reasonable one in our application, we derive the linear combination of $Z^{(1,1)}_{2,d^*+1/2}(1)$, $Z^{(1,2)}_{2,d^*+1/2}(1)$ and $Z^{(2,2)}_{2,d^*+1/2}(1)$ each weighted by the corresponding $\tilde{\alpha}^{(p,q)}$, $p, q = 1, 2$. $Z^{(1,1)}_{2,d^*+1/2}(1)$ and $Z^{(2,2)}_{2,d^*+1/2}(1)$ are Rosenblatt random variables again and we obtain a mixing term $Z^{(1,2)}_{2,d^*+1/2}(1)$ that is further investigated in the simulation study in Section 4.3.5.
However, with a mild technical assumption on the covariances of the one-dimensional marginal Gaussian processes that is often fulfilled in applications, there is another way of normalizing the partial sum on the right-hand side in Theorem 3.10, this time explicity for the case $\#P^* = 2$ and $h \in \mathbb{N}$, such that the limit can be expressed in terms of two standard Rosenblatt random variables. This yields the possibility to further study the dependence structure between these two random variables. In the following theorem we assume $\#P^* = d = 2$ for the reader's convenience.

Theorem 3.13 *Under the same assumptions as in Theorem 3.10 with* $\#P^* = d = 2$ *and* $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$ *and the additional condition that* $r^{(1,1)}(l) = r^{(2,2)}(l)$, *for* $l = 0, \dots, h-1$, *and* $L_{1,1} + L_{2,2} \neq L_{1,2} + L_{2,1}$, *it holds that*

$$
\begin{aligned}
n^{-2d^*} (C_2)^{-\frac{1}{2}} &\sum_{j=1}^{n} \left(f\left(Y_j^{(1)}, \dots, Y_{j+h-1}^{(d)}\right) - \mathbb{E} f\left(Y_j^{(1)}, \dots, Y_{j+h-1}^{(d)}\right)\right) \\
&\xrightarrow{\mathcal{D}} \left(\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}\right) \frac{L_{2,2} - L_{2,1} - L_{1,2} + L_{1,1}}{2} Z^{*}_{2,d^*+1/2}(1) \\
&\quad + \left(\tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)}\right) \frac{L_{2,2} + L_{2,1} + L_{1,2} + L_{1,1}}{2} Z^{**}_{2,d^*+1/2}(1)
\end{aligned}
$$

with $C_2 := \frac{1}{2d^*(4d^*-1)}$ *being the same normalizing factor as in Theorem 3.10,* $(\alpha_{i,k})_{1\leq i,k\leq dh} = \Sigma^{-1}_{d,h} C \Sigma^{-1}_{d,h}$ *and* $C = (c_{i,k})_{1\leq i,k \leq dh} = \mathbb{E}\left(Y_{1,h}\left(f\left(Y_{1,h}\right) - \mathbb{E} f\left(Y_{1,h}\right)\right) Y^t_{1,h}\right)$. *Note that* $Z^{*}_{2,d^*+1/2}(1)$ *and* $Z^{**}_{2,d^*+1/2}(1)$ *are both standard Rosenblatt random variables whose covariance is given by*

$$
Cov\left(Z^{*}_{2,d^*+1/2}(1), Z^{**}_{2,d^*+1/2}(1)\right) = \frac{\left(L_{2,2} - L_{1,1}\right)^2}{\left(L_{1,1} + L_{2,2}\right)^2 - \left(L_{1,2} + L_{2,1}\right)^2}. \tag{3.46}
$$

Proof. We follow the proof of Theorem 3.10 until (3.44), in order to obtain a limit distribution that can be expressed by the sum of two standard Rosenblatt random variables:

$$
\begin{aligned}
&\sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) \\
=&\frac{1}{n} \sum_{j=1}^{n} \sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} \left(Y_j^{(p)} Y_j^{(q)} - r^{(p,q)}(0)\right)
\end{aligned}
$$

$$=\frac{1}{n}\sum_{j=1}^{n}\left(Y_j^{(1)},Y_j^{(2)}\right)\begin{pmatrix}\tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)}\\ \tilde{\alpha}^{(2,1)} & \tilde{\alpha}^{(2,2)}\end{pmatrix}\left(Y_j^{(1)},Y_j^{(2)}\right)^t$$

$$-\mathbb{E}\left(\left(Y_j^{(1)},Y_j^{(2)}\right)\begin{pmatrix}\tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)}\\ \tilde{\alpha}^{(2,1)} & \tilde{\alpha}^{(2,2)}\end{pmatrix}\left(Y_j^{(1)},Y_j^{(2)}\right)^t\right). \tag{3.47}$$

We remember that $\tilde{\alpha}^{(p,q)}=\sum_{i,k=1}^{h}\alpha_{i,k}^{(p,q)}=\sum_{i,k=1}^{h}\alpha_{i+(p-1)h,k+(q-1)h}$ for $p,q=1,2$ and $\mathbb{A}=(\alpha_{i,k})_{1\leq i,k\leq 2h}=\Sigma_{2,h}^{-1}C\Sigma_{2,h}^{-1}$. Since $\Sigma_{2,h}^{-1}$ is the inverse of the covariance matrix $\Sigma_{2,h}$ of $Y_{1,h}$ it is a symmetric matrix. The matrix of second order Hermite coefficients C has the representation $C=\mathbb{E}\left(Y_{j,h}f\left(Y_{j,h}\right)Y_{j,h}^t\right)$ and, therefore, $c_{i,k}=\mathbb{E}\left(Y_{j,h}^{(i)}Y_{j,h}^{(k)}f\left(Y_{j,h}\right)\right)=c_{k,i}$ for each $i,k=1,\dots,2h$. Then, $\mathbb{A}$ is a symmetric matrix, too, since $\mathbb{A}^t=\left(\Sigma_{2,h}^{-1}C\Sigma_{2,h}^{-1}\right)^t=\left(\Sigma_{2,h}^{-1}\right)^tC^t\left(\Sigma_{2,h}^{-1}\right)^t=\mathbb{A}$. We can now show that $\begin{pmatrix}\tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)}\\ \tilde{\alpha}^{(2,1)} & \tilde{\alpha}^{(2,2)}\end{pmatrix}$ is a symmetric matrix, i.e., $\tilde{\alpha}^{(1,2)}=\tilde{\alpha}^{(2,1)}$. To this end, we define $\mathbb{I}_p=(0,0,\dots,0,1,\dots,1,0,\dots,0)^t\in\mathbb{R}^{2h}$ such that $\mathbb{I}_p^{(i)}=1$ only if $i=(p-1)h+1,\dots,ph$, $p=1,2$. Then, we obtain

$$\tilde{\alpha}^{(1,2)}=\sum_{i,k=1}^{h}\alpha_{i,k}^{(1,2)}=\left(\tilde{\alpha}^{(1,2)}\right)^t=\left(\mathbb{I}_1^t\mathbb{A}\mathbb{I}_2\right)^t=\mathbb{I}_2^t\mathbb{A}\mathbb{I}_1=\tilde{\alpha}^{(2,1)}.$$

We now apply the new assumption that $r^{(1,1)}(l)=r^{(2,2)}(l)$, for $l=0,\dots,h-1$ and show $\tilde{\alpha}^{(1,1)}=\tilde{\alpha}^{(2,2)}$ with the symmetry features of the multivariate normal distribution discussed in (2.2) and in (2.3), since $c_{i,j}=c_{2h-i+1,2h-j+1}$, $i,j=1,\dots,2h$.
We have to study

$$\tilde{\alpha}^{(2,2)}=\left(\mathbb{I}_2^t\mathbb{A}\mathbb{I}_2\right)^t=\mathbb{I}_2^t\Sigma_{2,h}^{-1}C\Sigma_{2,h}^{-1}\mathbb{I}_2.$$

Since $\Sigma_{2,h}^{-1}=(g_{i,k})_{1\leq i,k\leq 2h}$ is a symmetric and persymmetric matrix, we have $g_{i,k}=g_{k,i}$ and $g_{i,k}=g_{2h-i+1,2h-k+1}$ for $i,k=1,\dots,2h$. Then, we obtain

$$\begin{aligned}\mathbb{I}_2^t\Sigma_{2,h}^{-1}&=\left(\sum_{i=h+1}^{2h}g_{i,1},\dots,\sum_{i=h+1}^{2h}g_{i,2h}\right)\\&=\left(\sum_{i=1}^{h}g_{i+h,1},\dots,\sum_{i=1}^{h}g_{i+h,2h}\right)\\&=\left(\sum_{i=1}^{h}g_{h-i+1,2h},\dots,\sum_{i=1}^{h}g_{h-i+1,1}\right)\\&=\left(\sum_{i=1}^{h}g_{i,2h},\dots,\sum_{i=1}^{h}g_{i,1}\right)\\&=\left(\sum_{i=1}^{h}g_{2h,i},\dots,\sum_{i=1}^{h}g_{1,i}\right)\\&=:\left(\tilde{g}_{2h},\dots,\tilde{g}_1\right).\end{aligned}$$

Note that

$$\Sigma_{2,h}^{-1}\mathbb{I}_1=\left(\sum_{i=1}^{h}g_{1,i},\dots,\sum_{i=1}^{h}g_{2h,i}\right)^t=\left(\tilde{g}_1,\dots,\tilde{g}_{2h}\right)^t.$$

Then, we arrive at

$$
\begin{aligned}
\tilde{\alpha}^{(2,2)} = \left(\mathbb{I}_2^t \mathbb{A} \mathbb{I}_2\right)^t &= \mathbb{I}_2^t \Sigma_{2,h}^{-1} C \Sigma_{2,h}^{-1} \mathbb{I}_2 \\
&= \sum_{i,k=1}^{2h} \tilde{g}_{2h-i+1} \tilde{g}_{2h-k+1} c_{i,k} \\
&= \sum_{i,k=1}^{2h} \tilde{g}_{2h-i+1} \tilde{g}_{2h-k+1} c_{2h-i+1,2h-k+1} \\
&= \sum_{i,k=1}^{2h} \tilde{g}_i \tilde{g}_k c_{i,k} \\
&= \mathbb{I}_1^t \Sigma_{2,h}^{-1} C \Sigma_{2,h}^{-1} \mathbb{I}_1 \\
&= \tilde{\alpha}^{(1,1)}.
\end{aligned}
$$

So we have to deal with a special type of 2×2-matrix, since the original matrix in the formula (3.44), namely $\begin{pmatrix} \tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)} \\ \tilde{\alpha}^{(2,1)} & \tilde{\alpha}^{(2,2)} \end{pmatrix}$ has now reduced to $\begin{pmatrix} \tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)} \\ \tilde{\alpha}^{(1,2)} & \tilde{\alpha}^{(1,1)} \end{pmatrix}$.

Finally, we know that any real-valued symmetric matrix A can be decomposed via diagonalization into an orthogonal matrix V and a diagonal matrix D, where the entries of the latter one are determined via the eigenvalues of A, for details, see [12], p. 327.

We can explicity give formulas for the entries of these matrices here:

$$
V = \begin{pmatrix} -2^{-1/2} & 2^{-1/2} \\ 2^{-1/2} & 2^{-1/2} \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 = \tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)} & 0 \\ 0 & \lambda_2 = \tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)} \end{pmatrix},
$$

such that

$$
VDV = \begin{pmatrix} \frac{\lambda_1+\lambda_2}{2} & \frac{\lambda_2-\lambda_1}{2} \\ \frac{\lambda_2-\lambda_1}{2} & \frac{\lambda_1+\lambda_2}{2} \end{pmatrix} = \begin{pmatrix} \tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)} \\ \tilde{\alpha}^{(1,2)} & \tilde{\alpha}^{(1,1)} \end{pmatrix}.
$$

So continuing with (3.47), we now have the representation

$$
\begin{aligned}
&\frac{1}{n} \sum_{j=1}^{n} \left(Y_j^{(1)}, Y_j^{(2)}\right) \begin{pmatrix} \tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)} \\ \tilde{\alpha}^{(2,1)} & \tilde{\alpha}^{(2,2)} \end{pmatrix} \left(Y_j^{(1)}, Y_j^{(2)}\right)^t \\
&\qquad\qquad - \mathbb{E}\left(\left(Y_j^{(1)}, Y_j^{(2)}\right) \begin{pmatrix} \tilde{\alpha}^{(1,1)} & \tilde{\alpha}^{(1,2)} \\ \tilde{\alpha}^{(2,1)} & \tilde{\alpha}^{(2,2)} \end{pmatrix} \left(Y_j^{(1)}, Y_j^{(2)}\right)^t\right) \\
=&\frac{1}{n} \sum_{j=1}^{n} \left(Y_j^{(1)}, Y_j^{(2)}\right) VDV \left(Y_j^{(1)}, Y_j^{(2)}\right)^t - \mathbb{E}\left(\left(Y_j^{(1)}, Y_j^{(2)}\right) VDV \left(Y_j^{(1)}, Y_j^{(2)}\right)^t\right) \\
=&\frac{1}{n} \sum_{j=1}^{n} \frac{\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}}{2} \left(\left(Y_j^{(2)} - Y_j^{(1)}\right)^2 - \mathbb{E}\left(Y_j^{(2)} - Y_j^{(1)}\right)^2\right) \\
&\quad + \frac{1}{n} \sum_{j=1}^{n} \frac{\tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)}}{2} \left(\left(Y_j^{(1)} + Y_j^{(2)}\right)^2 - \mathbb{E}\left(Y_j^{(1)} + Y_j^{(2)}\right)^2\right) \\
=&\frac{1}{n} \sum_{j=1}^{n} \left(\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}\right) \left(1 - r^{(1,2)}(0)\right) \left(\left(\frac{Y_j^{(2)} - Y_j^{(1)}}{\sqrt{2 - 2r^{(1,2)}(0)}}\right)^2 - 1\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{n}\sum_{j=1}^{n}\left(\tilde{\alpha}^{(1,1)}+\tilde{\alpha}^{(1,2)}\right)\left(1+r^{(1,2)}(0)\right)\left(\left(\frac{Y_j^{(1)}+Y_j^{(2)}}{\sqrt{2+2r^{(1,2)}(0)}}\right)^2-1\right)\\
=&\frac{1}{n}\left(\tilde{\alpha}^{(1,1)}-\tilde{\alpha}^{(1,2)}\right)\left(1-r^{(1,2)}(0)\right)\sum_{j=1}^{n}H_2\left(Y_j^{*}\right)\\
&+\frac{1}{n}\left(\tilde{\alpha}^{(1,1)}+\tilde{\alpha}^{(1,2)}\right)\left(1+r^{(1,2)}(0)\right)\sum_{j=1}^{n}H_2\left(Y_j^{**}\right),
\end{aligned}
\tag{3.48}
$$

with $Y_j^* := \frac{Y_j^{(2)}-Y_j^{(1)}}{\sqrt{2-2r^{(1,2)}(0)}}$ and $Y_j^{**} := \frac{Y_j^{(1)}+Y_j^{(2)}}{\sqrt{2+2r^{(1,2)}(0)}}$.
Now note that

$$
\mathbb{E}\left(Y_j^*Y_j^{**}\right)=\mathbb{E}\left(\frac{Y_j^{(2)}-Y_j^{(1)}}{\sqrt{2-2r^{(1,2)}(0)}}\frac{Y_j^{(1)}+Y_j^{(2)}}{\sqrt{2+2r^{(1,2)}(0)}}\right)=0.
$$

Therefore, we created a bivariate long-range dependent Gaussian process, since

$$
\begin{pmatrix}-1 & 1\\ 1 & 1\end{pmatrix}\left(Y_j^{(1)},Y_j^{(2)}\right)^t=\left(Y_j^*,Y_j^{**}\right)^t\sim\mathcal{N}\left(0,I_2\right)
$$

with cross-covariance function

$$
\begin{aligned}
r_*^{(1,2)}(k):=\mathbb{E}\left(Y_j^*Y_{j+k}^{**}\right)&=\mathbb{E}\left(\frac{Y_j^{(2)}-Y_j^{(1)}}{\sqrt{2-2r^{(1,2)}(0)}}\frac{Y_{j+k}^{(1)}+Y_{j+k}^{(2)}}{\sqrt{2+2r^{(1,2)}(0)}}\right)\\
&=\frac{r^{(2,1)}(k)+r^{(2,2)}(k)-r^{(1,1)}(k)-r^{(1,2)}(k)}{2\sqrt{\left(1-r^{(1,2)}(0)\right)\left(1+r^{(1,2)}(0)\right)}}\\
&\simeq\frac{L_{2,2}+L_{2,1}-L_{1,2}-L_{1,1}}{2\sqrt{\left(1-r^{(1,2)}(0)\right)\left(1+r^{(1,2)}(0)\right)}}k^{2d^*-1}.
\end{aligned}
\tag{3.49}
$$

Note that the covariance functions have the following asymptotic behaviour:

$$
\begin{aligned}
r_*^{(1,1)}(k):=\mathbb{E}\left(Y_j^*Y_{j+k}^{*}\right)&=\mathbb{E}\left(\frac{Y_j^{(2)}-Y_j^{(1)}}{\sqrt{2-2r^{(1,2)}(0)}}\frac{Y_{j+k}^{(2)}-Y_{j+k}^{(1)}}{\sqrt{2-2r^{(1,2)}(0)}}\right)\\
&=\frac{r^{(2,2)}(k)-r^{(2,1)}(k)-r^{(1,2)}(k)+r^{(1,1)}(k)}{2-2r^{(1,2)}(0)}\\
&\simeq\underbrace{\frac{L_{2,2}-L_{2,1}-L_{1,2}+L_{1,1}}{2-2r^{(1,2)}(0)}}_{=:L_{1,1}^*}k^{2d^*-1}
\end{aligned}
$$

and analogously

$$
r_*^{(2,2)}(k):=\mathbb{E}\left(Y_j^{**}Y_{j+k}^{**}\right)\simeq\underbrace{\frac{L_{2,2}+L_{2,1}+L_{1,2}+L_{1,1}}{2+2r^{(1,2)}(0)}}_{=:L_{2,2}^*}k^{2d^*-1}.
$$

We can now apply the result of [2], Theorem 6, since we created a bivariate Gaussian process with independent entries for fixed j. Note that for the function we apply here, namely

$\tilde{f}\left(Y_j^*, Y_j^{**}\right) = H_2\left(Y_j^*\right) + H_2\left(Y_j^{**}\right)$ the weighting factors in [2], Theorem 6 reduce to $e_{1,1} = e_{2,2} = 1$ and $e_{1,2} = e_{2,1} = 0$. These weighting factors fit into the result in [2], (3.6) and (3.7), that even yields joint convergence of the vector of both univariate summands, $\left(H_2\left(Y_j^*\right), H_2\left(Y_j^{**}\right)\right)$, suitably normalized, to a vector of two (dependent) Rosenblatt random variables. Since the long-range dependence property in Definition 2.13 is more specific than in [2], p. 2259, (3.1) (see considerations in (3.8)), we are able to scale the variances of each Rosenblatt random variable to 1 and give the covariance between them, by using the normalization given in [9], Theorem 4.3. We obtain

$$\begin{aligned}
&n^{-2d*}\left(2C_2\right)^{-1/2}\left(\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}\right)\left(1 - r^{(1,2)}(0)\right)\sum_{j=1}^{n} H_2\left(Y_j^*\right) \\
&\qquad + n^{-2d*}\left(2C_2\right)^{-1/2}\left(\tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)}\right)\left(1 + r^{(1,2)}(0)\right)\sum_{j=1}^{n} H_2\left(Y_j^{**}\right) \\
&\xrightarrow{\mathcal{D}} \left(\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}\right)\left(1 - r^{(1,2)}(0)\right) L_{1,1}^* Z_{2,d^*+1/2}^*(1) + \left(\tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)}\right)\left(1 + r^{(1,2)}(0)\right) L_{2,2}^* Z_{2,d^*+1/2}^{**} \\
&= \left(\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}\right)\frac{L_{2,2} - L_{2,1} - L_{1,2} + L_{1,1}}{2} Z_{2,d^*+1/2}^*(1) \\
&\qquad + \left(\tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)}\right)\frac{L_{2,2} + L_{2,1} + L_{1,2} + L_{1,1}}{2} Z_{2,d^*+1/2}^{**}(1)
\end{aligned}$$

with $C_2 := \frac{1}{2d^*(4d^*-1)}$ being the same normalizing factor as in Theorem 3.10.
We observe that $Z_{2,d^*+1/2}^*(1)$ and $Z_{2,d^*+1/2}^{**}(1)$ are both standard Rosenblatt random variables. Following Corollary 3.6, their covariance is given by

$$\begin{aligned}
\operatorname{Cov}\left(Z_{2,d^*+1/2}^*(1), Z_{2,d^*+1/2}^{**}(1)\right) &= \frac{\left(L_{1,2}^* + L_{2,1}^*\right)^2}{L_{1,1}^* L_{2,2}^*} \\
&= \frac{2\left(\left(L_{2,2} - L_{1,1}\right)^2\right)}{4\left(1 - r^{(1,2)}(0)\right)\left(1 + r^{(1,2)}(0)\right)}\left(L_{1,1}^* L_{2,2}^*\right)^{-1} \\
&= \frac{\left(L_{2,2} - L_{1,1}\right)^2}{\left(L_{1,1} + L_{2,2}\right)^2 - \left(L_{1,2} + L_{2,1}\right)^2}.
\end{aligned}$$

Note that $\left(L_{1,1} + L_{2,2}\right)^2 - \left(L_{1,2} + L_{2,1}\right)^2 \neq 0$ is fulfilled since $L_{1,1} + L_{2,2} \neq L_{1,2} + L_{2,1}$. □

Corollary 3.14 *Adding the condition that $r^{(1,1)}(k) = r^{(2,2)}(k)$ and $r^{(1,2)}(k) = r^{(2,1)}(k)$ for all $k \in \mathbb{Z}$ to the assumptions of Theorem 3.13, we obtain that the two Rosenblatt random variables $Z_{2,d^*+1/2}^*(1)$ and $Z_{2,d^*+1/2}^{**}(1)$ in the limit are independent.*

Proof. Using (3.49) in the proof of the preceding theorem, we know that

$$\begin{aligned}
r_*^{(1,2)}(k) := \mathbb{E}\left(Y_j^* Y_{j+k}^{**}\right) &= \mathbb{E}\left(\frac{Y_j^{(2)} - Y_j^{(1)}}{\sqrt{2 - 2r^{(1,2)}(0)}}\frac{Y_{j+k}^{(1)} + Y_{j+k}^{(2)}}{\sqrt{2 + 2r^{(1,2)}(0)}}\right) \\
&= \frac{r^{(2,1)}(k) + r^{(2,2)}(k) - r^{(1,1)}(k) - r^{(1,2)}(k)}{2\sqrt{\left(1 - r^{(1,2)}(0)\right)\left(1 + r^{(1,2)}(0)\right)}} \\
&= 0,
\end{aligned}$$

due to the added assumption. Therefore, the processes $\left(Y_j^*\right)_{j\in\mathbb{Z}}$ and $\left(Y_j^{**}\right)_{j\in\mathbb{Z}}$ are independent. Hence, the two corresponding limit random variables are independent, too, applying Levy's continuity theorem. □

Remark 3.15 *In the results concerning functionals f of multivariate Gaussian processes with Hermite rank 2, we always assume that $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$, such that the transformed processes $f\left(Y_{j,h}\right)$ is still long-range dependent, see Table 2.1. In the case that $d^* \in (-\infty, 0) \cup \left(0, \frac{1}{4}\right)$ and, therefore, the transformed process is no longer long-range dependent, although the original process is long-range dependent if $d^* \in \left(0, \frac{1}{4}\right)$, a central limit theorem holds by applying Theorem 4 in [2]:*

$$n^{-\frac{1}{2}} \sum_{j=1}^{n} \left(f\left(Y_{j,h}\right) - \mathbb{E}\left(f\left(Y_{j,h}\right)\right)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2\right),$$

with

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}\left[\left(f\left(Y_{1,h}\right) - \mathbb{E}\left(f\left(Y_{1,h}\right)\right)\right)\left(f\left(Y_{1+k,h}\right) - \mathbb{E}\left(f\left(Y_{1+k,h}\right)\right)\right)\right],$$

for further details on this formula, see [54], Theorem 5.4.1.
We may apply [2], Theorem 4, since for each $p, q = 1, \ldots, dh$ the squared cross-correlations $\left(r^{(p,q,h)}(k)\right)^2$ of the process $\left(Y_{j,h}\right)_{j\in\mathbb{Z}}$ as given in (3.7) with $p^*, q^* \in \{1, \ldots, d\}$ and $m^* = 0, \ldots, h-1$, are summable in this setting:

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=-n}^{n} \left(r^{(p,q,h)}(k)\right)^2 &= \lim_{n\to\infty} \sum_{k=-n}^{n} \left(r^{(p^*,q^*)}(k+m^*)\right)^2 \\
&= \lim_{n\to\infty} \sum_{k=-n}^{n} \left(r^{(p^*,q^*)}(k)\right)^2 + \underbrace{\sum_{l=1}^{m^*} \left(r^{(p^*,q^*)}(n+l)\right)^2}_{\xrightarrow{(n\to\infty)} 0} - \underbrace{\sum_{l=0}^{m^*-1} \left(r^{(p^*,q^*)}(-n+l)\right)^2}_{\xrightarrow{(n\to\infty)} 0} \\
&= \lim_{n\to\infty} \left(r^{(p^*,q^*)}(0)\right)^2 + \left(L_{p^*,q^*}^2 + L_{q^*,p^*}^2\right) \sum_{k=1}^{n} k^{2(d_{p^*}+d_{q^*}-1)} \\
&\leq \lim_{n\to\infty} \left(r^{(p^*,q^*)}(0)\right)^2 + \left(L_{p^*,q^*}^2 + L_{q^*,p^*}^2\right) \sum_{k=1}^{n} k^{4d^*-2} \\
&< \infty,
\end{aligned}$$

since $4d^* - 2 \in (-\infty, -1)$.

Hence, we proved a limit theorem for functionals with Hermite rank 1 and two limit theorems for Hermite rank 2. In both cases a generalization to mixed cases of short- and long-range dependent components within the multivariate Gaussian process is given. In the following section, we investigate the special case that the considered functionals f fulfill the property that they can be uniquely determined by the increments of the marginal processes. It turns out that this has a surprising effect on the asymptotic distribution.

3.3 Limit theorems for functionals of long-range dependent multivariate Gaussian time series that can be expressed in terms of the increment processes

In this section we show that for a small class of functionals, we obtain a degenerated asymptotic distribution in the limit theorems derived in the two preceding sections. Therefore, we slightly change the notation for the reader's convenience, with the following chapters dealing with applications to ordinal pattern analysis in mind. We now assume similar to the framework introduced at the beginning of Chapter 3 that we have a d-dimensional Gaussian process $(X_j)_{j\in\mathbb{Z}}$ as defined in (3.1) that is long-range dependent in the sense of Definition 2.13. Again we define $(X_{j,h+1})_{j\in\mathbb{Z}}$ with $X_{j,h+1} := \left(X_j^{(1)}, \ldots, X_{j+h}^{(1)} \ldots, X_j^{(d)}, X_{j+h}^{(d)}\right)$. So all consideration from (3.4) to (3.10) remain valid for this process. We define the increment process $(Y_{j,h})_{j\in\mathbb{Z}}$ of $(X_{j,h+1})_{j\in\mathbb{Z}}$ by

$$\begin{aligned} Y_{j+1,h} :&= \left(Y_{j+1}^{(1)}, \ldots, Y_{j+h}^{(1)} \ldots, Y_{j+1}^{(d)} \ldots, Y_{j+h}^{(d)}\right)^t \\ &= \left(X_{j+1}^{(1)} - X_j^{(1)}, \ldots, X_{j+h}^{(1)} - X_{j+h-1}^{(1)}, \ldots, X_{j+1}^{(d)} - X_j^{(d)}, \ldots, X_{j+h}^{(d)} - X_{j+h-1}^{(d)}\right)^t. \end{aligned}$$

Note that in the next two lemmas, the assumption of long-range dependence of $(X_{j,h+1})_{j\in\mathbb{Z}}$ is not required, however, in the following corollaries it is necessary, since the asymptotic distributions in the limit theorems we derived in Section 3.1 and Section 3.2, where long-range dependence is the crucial condition, turn out to be trivial.

Lemma 3.16 *Let $(X_{j,h+1})_{j\in\mathbb{Z}}$ be a stationary Gaussian process with $X_{j,h+1} \sim \mathcal{N}(0, \Sigma_{d,h+1})$ and $f : \mathbb{R}^{d(h+1)} \to \mathbb{R}$ be a square-integrable function such that $\mathbb{E}f(X_{j,h+1}) = 0$. Further assume that f has Hermite rank 1. If f only depends on the increment vector of $X_{j,h+1}$ defined by*

$$\begin{aligned} Y_{j+1,h} &= \left(Y_{j+1}^{(1)}, \ldots, Y_{j+h}^{(1)} \ldots, Y_{j+1}^{(d)} \ldots, Y_{j+h}^{(d)}\right)^t \\ &= \left(X_{j+1}^{(1)} - X_j^{(1)}, \ldots, X_{j+h}^{(1)} - X_{j+h-1}^{(1)}, \ldots, X_{j+1}^{(d)} - X_j^{(d)}, \ldots, X_{j+h}^{(d)} - X_{j+h-1}^{(d)}\right)^t \end{aligned}$$

so there exists a function $\tilde{f}$ such that $f(X_{j,h+1}) = \tilde{f}(Y_{j+1,h})$, then it holds that

$$\sum_{i=1}^{h+1} \alpha_{i+(p-1)(h+1)} = 0,$$

for each $p = 1, \ldots, d$ and $\alpha = \left(\alpha_1, \ldots, \alpha_{d(h+1)}\right)^t = \Sigma_{d,h+1}^{-1} c$ with $c = \mathbb{E}(f(X_{0,h+1}) X_{0,h+1})$.

Proof. We follow the argumentation in [10], Lemma 5.7.
We define the $(d(h+1)-1) \times (d(h+1))$-matrix $\mathcal{U}$ by

$$\mathcal{U} := \begin{pmatrix} \multicolumn{3}{c}{\mathbf{Inc}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\dots} & \multicolumn{3}{c}{\mathbf{0}} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{Inc}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\dots} & \multicolumn{3}{c}{\mathbf{0}} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{Inc}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\dots} & \multicolumn{3}{c}{\mathbf{0}} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\ddots} & \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\vdots} \\ \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\vdots} & \multicolumn{3}{c}{\ddots} & \multicolumn{3}{c}{\vdots} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\mathbf{0}} & \multicolumn{3}{c}{\dots} & \multicolumn{3}{c}{\mathbf{Inc}} \end{pmatrix}$$

where

$$\mathbf{Inc} := \begin{pmatrix} -1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(h+1)\times h}$$

and $\mathbf{0} \in \mathbb{R}^{(h+1)\times h}$ being the $((h+1)\times h)$-zero matrix.
So note that (among other possibilities) $\mathcal{U}_{i,k} = 0$ if $i = p(h+1)$, $p = 1, \dots, d-1$ since we are never considering "increments" between two processes, meaning for example $X_j^{(1)} - X_{j+1}^{(2)}$ what one might first assume looking at the special structure of $X_{j,h+1}$.
We observe that $\mathcal{U}X_{j,h+1} = Y_{j+1,h}$ for all $j \in \mathbb{Z}$.
Note that there exists a positive definite symmetric matrix $\Sigma_{d,h+1}^{\frac{1}{2}}$ that fulfills $\Sigma_{d,h+1} = \Sigma_{d,h+1}^{\frac{1}{2}}\Sigma_{d,h+1}^{\frac{1}{2}}$. With $\Sigma_{d,h+1}^{-\frac{1}{2}}$ we denote the inverse of $\Sigma_{d,h+1}^{\frac{1}{2}}$. Then, the random vector $U_{j,h+1} = \Sigma_{d,h+1}^{-\frac{1}{2}}X_{j,h+1}$ has independent $\mathcal{N}(0,1)$ entries. Then, we obtain

$$\begin{aligned} \alpha &= \Sigma_{d,h+1}^{-1}\mathbb{E}\left(\tilde{f}\left(Y_{j+1,h}\right)X_{j,h+1}\right) \\ &= \Sigma_{d,h+1}^{-1}\mathbb{E}\left(\tilde{f}\left(\mathcal{U}X_{j,h+1}\right)\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right) \\ &= \mathbb{E}\left(\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right). \end{aligned}$$

We now define $\mathbb{I}^{(p)} := (0,0,\dots,0,1,1,\dots,1,0,\dots,0,0)^t \in \mathbb{R}^{d(h+1)}$ for each $p = 1, \dots, d$, such that $\mathbb{I}_i^{(p)} = 1$ only if $i = (h+1)(p-1)+1, \dots, (h+1)p$. We then arrive at

$$\sum_{i=1}^{h+1} \alpha_{i+(p-1)(h+1)} = \left(\mathbb{I}^{(p)}\right)^t \alpha = \mathbb{E}\left(\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right). \tag{3.50}$$

We have

$$\begin{aligned}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}\right)\left(\left(\mathbb{I}^{(p)}\right)^t\Sigma_{d,h+1}^{-\frac{1}{2}}\right)^t &= \mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}\Sigma_{d,h+1}^{-\frac{1}{2}}\mathbb{I}^{(p)}\\ &= \mathcal{U}\mathbb{I}^{(p)}\\ &= 0,\end{aligned}$$

and using the proof of Lemma 5.7 in [10], we proved that $\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}$ and $\left(\mathbb{I}^{(p)}\right)^t\Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}$ are independent random variables and, therefore, using (3.50) we arrive at

$$\begin{aligned}\mathbb{E}\left(\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\left(\mathbb{I}^{(p)}\right)^t\Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right) &= \mathbb{E}\left(\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\right)\mathbb{E}\left(\left(\mathbb{I}^{(p)}\right)^t\Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right)\\ &= 0,\end{aligned}$$

since $\mathbb{E}\left(U_{j,h+1}\right) = 0$. □

Lemma 3.17 *Let $\left(X_{j,h+1}\right)_{j\in\mathbb{Z}}$ be a stationary Gaussian process with $X_{j,h+1} \sim \mathcal{N}\left(0, \Sigma_{d,h+1}\right)$ and $f : \mathbb{R}^{d(h+1)} \to \mathbb{R}$ be a square-integrable function such that $\mathbb{E}\left(f\left(X_{j,h+1}\right)\right) = 0$. Further assume that f has Hermite rank 2. If f only depends on the increments of the underlying random vector, so there exists a function $\tilde{f}$ such that $f\left(X_{j,h+1}\right) = \tilde{f}\left(Y_{j+1,h}\right)$, it then holds that*

$$\tilde{\alpha}^{(p,q)} = \sum_{i,k=1}^{h+1}\alpha_{i,k}^{(p,q)} = 0$$

for each $p = 1, \ldots, d$.

Proof. We extend the idea of the proof of Lemma 3.16 to functionals with Hermite rank 2. We use the same matrix $\mathcal{U}$ as in the proof of Lemma 3.16 so we can use the relation $\mathcal{U}X_{j,h+1} = Y_{j+1,h}$ for all $j \in \mathbb{Z}$ in this setting, too.
We also use the existence of the positive definite symmetric matrix $\Sigma_{d,h+1}^{\frac{1}{2}}$ again as described in the proof above. Then, we obtain

$$\begin{aligned}\alpha &= \Sigma_{d,h+1}^{-1}\mathbb{E}\left(X_{j,h+1}\tilde{f}\left(Y_{j+1,h}\right)X_{j,h+1}^t\right)\Sigma_{d,h+1}^{-1}\\ &= \mathbb{E}\left(\Sigma_{d,h+1}^{-1}X_{j,h+1}\tilde{f}\left(Y_{j+1,h}\right)X_{j,h+1}^t\Sigma_{d,h+1}^{-1}\right)\\ &= \mathbb{E}\left(\Sigma_{d,h+1}^{-1}X_{j,h+1}\tilde{f}\left(\mathcal{U}X_{j,h+1}\right)X_{j,h+1}^t\Sigma_{d,h+1}^{-1}\right)\\ &= \mathbb{E}\left(\Sigma_{d,h+1}^{-1}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\left(\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)^t\Sigma_{d,h+1}^{-1}\right)\\ &= \mathbb{E}\left(\Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)U_{j,h+1}^t\Sigma_{d,h+1}^{-\frac{1}{2}}\right).\end{aligned}$$

Using again

$$\mathbb{I}^{(p)} := (0, 0, \ldots, 0, 1, 1, \ldots, 1, 0, \ldots, 0, 0)^t \in \mathbb{R}^{d(h+1)}$$

We justify that if $\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}$ is independent of $\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}$ for each $p = 1, \ldots, d$, then we can conclude that $\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}$ is independent of $\left(\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}, \left(\mathbb{I}^{(q)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right)$ by a calculation using the characteristical function for multivariate random vectors. First note, that $\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}, \left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}, \left(\mathbb{I}^{(q)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right)$ is multivariate Gaussian and hence we simplify the notation here to the general setting of a multivariate Gaussian vector $(X_1, X_2, X_3) \sim \mathcal{N}\left(0, \Sigma_3 = \begin{pmatrix} \Sigma_1 & 0 & 0 \\ 0 & \Sigma_2 & \Sigma_{2,3} \\ 0 & \Sigma_{2,3} & \Sigma_3 \end{pmatrix}\right)$, where each component X_i, $i = 1, 2, 3$ is a $d-dimensional$ Gaussian vector itself.

For all $t_1, t_2, t_3 \in \mathbb{R}^d$, the characteristical function is then given by

$$\begin{aligned}
\varphi_{(X_1,X_2,X_3)}(t_1, t_2, t_3) &= \exp\left(-\frac{1}{2}(t_1, t_2, t_3)\,\Sigma_3\,(t_1, t_2, t_3)^t\right) \\
&= \exp\left(-\frac{1}{2}\left(t_1\Sigma_1 (t_1)^t + (t_2, t_3)\begin{pmatrix} \Sigma_2 & \Sigma_{2,3} \\ \Sigma_{2,3} & \Sigma_3 \end{pmatrix}(t_2, t_3)^t\right)\right) \\
&= \varphi_{(X_1)}(t_1)\,\varphi_{(X_2,X_3)}(t_2, t_3).
\end{aligned}$$

for each $p = 1, \ldots, d$ and the already proved independence of $\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}$ and $\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}$ for each $p = 1, \ldots, d$, we arrive at

$$\begin{aligned}
\tilde{\alpha}^{(p,q)} &= \sum_{i,k=1}^{h+1} \alpha_{i,k}^{(p,q)} \\
&= \left(\mathbb{I}^{(p)}\right)^t \alpha \mathbb{I}^{(q)} \\
&= \mathbb{E}\left(\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\left(\left(\mathbb{I}^{(q)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right)^t\right) \\
&= \mathbb{E}\left(\tilde{f}\left(\mathcal{U}\Sigma_{d,h+1}^{\frac{1}{2}}U_{j,h+1}\right)\right)\mathbb{E}\left(\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\left(\left(\mathbb{I}^{(q)}\right)^t \Sigma_{d,h+1}^{-\frac{1}{2}}U_{j,h+1}\right)^t\right) \\
&= \mathbb{E}\left(f(X_{j,h+1})\right)\left(\mathbb{I}^{(p)}\right)^t \Sigma_{d,h+1}^{-1}\mathbb{I}^{(q)} \\
&= 0,
\end{aligned}$$

for $p, q = 1, \ldots, d$, since we assumed $\mathbb{E}\left(f(X_{j,h+1})\right) = 0$. □

We can now apply these two lemmas to the limit theorems we derived in Section 3.1 and Section 3.2. As we see later on, the special case that the function f is uniquely determined by the increment process plays an important role in the context of ordinal pattern analysis.

Corollary 3.18 *Let $(X_{j,h+1})_{j\in\mathbb{Z}}$ be a stationary long-range dependent Gaussian process and $f : \mathbb{R}^{d(h+1)} \to \mathbb{R}$ that both fulfill the assumptions in Theorem 3.3. If f only depends on the increment process $(Y_{j+1,h})_{j\in\mathbb{Z}}$, so there exists a function $\tilde{f}$ such that $f(X_{j,h+1}) = \tilde{f}(Y_{j+1,h})$, for all $j \in \mathbb{Z}$, then*

$$n^{-(d^*+1/2)} C_1^{-1/2} \sum_{j=0}^{n-1} \left(f(X_{j,h+1}) - \mathbb{E} f(X_{j,h+1}) \right) \xrightarrow{\mathcal{D}} \delta_0,$$

where δ_0 denotes the Dirac measure in 0.

We derive an analogous result in the case of Hermite rank 2.

Corollary 3.19 *Let $(X_{j,h+1})_{j\in\mathbb{Z}}$ be a stationary long-range dependent Gaussian process and $f : \mathbb{R}^{d(h+1)} \to \mathbb{R}$ that both fulfill the assumptions in Theorem 3.10. If f only depends on the increment process $(Y_{j+1,h})_{j\in\mathbb{Z}}$, so there exists a function $\tilde{f}$ such that $f(X_{j,h+1}) = \tilde{f}(Y_{j+1,h})$, for all $j \in \mathbb{Z}$, then*

$$\begin{aligned} n^{-2d^*} (C_2)^{-\frac{1}{2}} \sum_{j=0}^{n-1} & \left(f(X_{j,h+1}) - \mathbb{E}\left(f(X_{j,h+1}) \right) \right) \\ & \xrightarrow{\mathcal{D}} \delta_0. \end{aligned}$$

Remark 3.20 *Note that this result is consistent with the statement of Theorem 3.13 that is derived with additional assumptions and which would also yield a degenerate limit distribution in this special case.*

We close this section on limit theorems for functionals of long-range dependent multivariate Gaussian time series that can be expressed in terms of the increment processes with the justification of the degenerated limit distributions above, that can be found in [10], Remark 5.8. The crucial point here is that the long-range dependence of the process $(X_{j,h+1})_{j\in\mathbb{Z}}$ is not passed on the increment process $(Y_{j,h})_{j\in\mathbb{Z}}$, since for all $p, q = 1, \dots, d$

$$\begin{aligned} \mathbb{E}\left(Y_j^{(p)} Y_{j+k}^{(q)} \right) &= \mathbb{E}\left(\left(X_j^{(p)} - X_{j-1}^{(p)} \right) \left(X_{j+k}^{(q)} - X_{j+k-1}^{(q)} \right) \right) \\ &= 2r^{(p,q)}(k) - r^{(p,q)}(k-1) - r^{(p,q)}(k+1) \end{aligned}$$

and therefore

$$\sum_{k=1}^{\infty} \mathbb{E}\left(Y_j^{(p)} Y_{j+k}^{(q)} \right) = \sum_{k=1}^{\infty} 2r^{(p,q)}(k) - r^{(p,q)}(k-1) - r^{(p,q)}(k+1) = 0, \tag{3.51}$$

since $r^{(p,q)}(k) \to 0$ for $k \to \infty$. Hence, the increment process is antipersistent.
An illustrative example that the increment process is short-range dependent is given by the relation to the correlation function of fractional Gaussian noise, see Definition 2.20. The asymptotic property of the correlation function yields

$$\mathbb{E}\left(Y_j^{(p)} Y_{j+k}^{(q)} \right) \simeq -2L_{p,q} \left(\frac{d_p + d_q - 1}{2} \right) (d_p + d_q - 2)\, k^{d_p+d_q-3}, \quad (k \to \infty), \tag{3.52}$$

and, therefore, since $d_p + d_q \in \left(0, \frac{1}{2}\right)$, the cross-correlations of this process are absolutely summable and the increment process short-range dependent.
Since we are considering a functional of a short-range dependent Gaussian process in this special case, a Gaussian limit distribution is obtained by applying Theorem 5.4.1 in [54] using a the classical normalization constant $n^{-1/2}$. For more details, see [10], Remark 5.8.

We conclude this chapter with a summary of the results. We built up the main mathematical theory in terms of limit theorems for functionals of multivariate long-range dependent Gaussian processes with Hermite rank 1 in Section 3.1 and Hermite rank 2 in Section 3.2. A special case, which considers functionals of long-range dependent multivariate Gaussian time series that only depend on the underlying increment process was considered and degenerated limit distributions were obtained. However, the reason of this problem was figured out and a solution is given by using limit theorems for short-range dependent processes instead. In the next chapters, we continue to apply these theoretical results to estimators of ordinal pattern probabilities and ordinal pattern dependence and, therefore, leave the theoretical framework and turn to more practical and statistical matters in order to get more information on the parameters determining the limit distributions depending on the Hermite rank of the respective estimator under consideration.

4 Ordinal pattern analysis

We turn to the application of the limit theorems derived in the previous chapter to ordinal pattern analysis. This chapter is structured as follows: First, we give an introduction to ordinal patterns and provide necessary background information and relations. Subsequently, in Section 4.2, we apply the limit theorems to ordinal pattern probabilities, which correspond to the case $d = 1$ in Chapter 3. In Section 4.3 we focus on the estimation of ordinal pattern dependence and, hence, to the multivariate case with $d = 2$. We study asymptotic properties of different estimators in the context of ordinal pattern dependence. Finally, in Section 4.4, we discuss modifications and generalizations of standard ordinal pattern dependence and the corresponding estimation methods.

4.1 Ordinal patterns

As already illustrated in the introduction of this thesis, the general idea of ordinal analysis is to reduce the information given in a data set to the ordinal one, so that we only compare the relative position of the data points and, therefore, may ignore the absolute values. For an overview of ordinal analysis in general, see [4]. Note that this approach lowers the complexity of a data set significantly. This yields a large simplification in the era of digitalization since we are dealing with huge amounts of data points and need a large computational effort to analyze their dependence structures with standard dependence measures. However, using the ordinal approach, we are still extracting the relevant information that we need to gain results concerning for example the dependence structure within a certain data set or even between several ones. A big advantage of the ordinal point of view is that this method is invariant under monotone transformations and, therefore, is independent of the scaling of the data sets. Since some disturbances in the data as for example outliers due to extreme events do not affect this approach, the model is sufficiently robust, which is a great benefit from the user's perspective. As already discussed in the beginning, the methods introduced in the following, have a strong practical impact and are used in various fields such as hydrology, neurological science, medicine and artificial intelligence.

We start with the definition of an ordinal pattern and the basic mathematical framework that we need to build up the ordinal model.
Let S_h denote the set of permutations in $\{0, \ldots, h\}$, $h \in \mathbb{N}_0$ that we express as $(h+1)$-dimensional tuples, assuring that each tuple contains each of the numbers above exactly once. In mathematical terms, this yields

$$S_h = \left\{ \pi \in \mathbb{N}_0^{h+1} : \quad 0 \leq \pi_i \leq h, \text{ and } \pi_i \neq \pi_k, \text{ whenever } i \neq k, \quad i, k = 0, \ldots, h \right\},$$

see [58], Sec. 2.1.
The number of permutations in S_h is given by $\#S_h = (h+1)!$. In order to get a better intuitive understanding of the concept of ordinal patterns, we have a closer look at the following example, before turning to the formal definition.

Example 4.1 *Figure 4.1 provides an illustrative understanding of the extraction of an ordinal pattern from a data set. The data points of interest are colored in red and we consider a pattern of length $h = 3$, which means we have to take $n = 4$ data points into consideration. We fix the points in time t_0, t_1, t_2 and t_3 and extract the data points from the time series. Then, we search for the point in time which exhibits the largest value in the resulting data and write down the corresponding time index. In this example it is given by $t = t_1$. We order the data points by writing the time position of the largest value as first entry, the time position of the second largest as second entry and so on. Hence, the absolute values are ordered from largest to smallest and the ordinal pattern $(1, 0, 3, 2) \in S_3$ is obtained for the considered data points.*

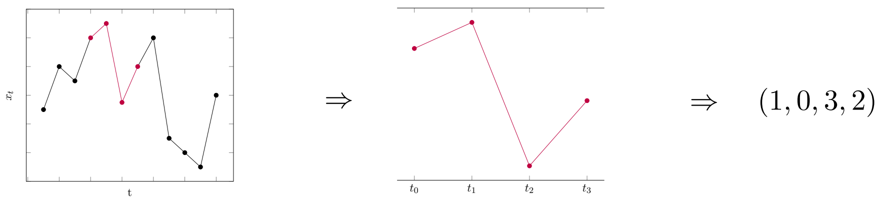

Figure 4.1: Example of the extraction of an ordinal pattern of a given data set.

Formally, the aforementioned procedure can be defined as follows, see [58], Section 2.1.

Definition 4.2 *As ordinal pattern of a vector $x = (x_0, \ldots, x_h) \in \mathbb{R}^{h+1}$, we define the unique permutation $\pi = (\pi_0, \ldots, \pi_h) \in S_h$,*

$$\Pi(x) = \Pi(x_0, \ldots, x_h) = (\pi_0, \ldots, \pi_h),$$

such that

$$x_{\pi_0} \geq \ldots \geq x_{\pi_h},$$

with $\pi_{i-1} < \pi_i$ if $x_{\pi_{i-1}} = x_{\pi_i}$, $i = 1, \ldots, h$.

The last condition assures the uniqueness of π if there are ties in the data sets. In particular, this condition is necessary to remark if real-world data is considered.

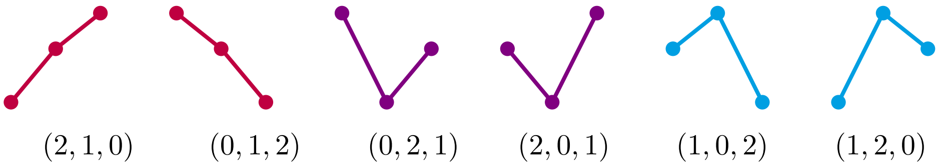

Figure 4.2: Ordinal patterns for $h = 2$.

In Figure 4.2, all ordinal patterns of length $h = 2$ are shown. As already mentioned in the

introduction, from the practical point of view, a highly desirable property of ordinal patterns is that they are not affected by monotone transformations, see [60], p. 1783.
Mathematically, this means: if $f : \mathbb{R} \to \mathbb{R}$ is strictly montone, then

$$\Pi(x_0, \ldots, x_h) = \Pi(f(x_0), \ldots, f(x_h)). \tag{4.1}$$

In particular, this includes linear transformations $f(x) = ax + b$, with $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$.

Following [58], Sec. 1, the minimal requirement of the data sets we use for ordinal analysis in the time series context, i.e., for ordinal pattern probabilities as well as for ordinal pattern dependence later on, is *ordinal pattern stationarity (of order h).* This property implies that the probability of observing a certain ordinal pattern of length h remains the same when shifting the moving window of length h through the entire time series and is not depending on the specific points in time. In the course of this work, the time series, in which the ordinal patterns occur, always have either stationary increments or are even stationary themselves. Note that both properties imply ordinal pattern stationarity. The reason why requiring stationary increments is a sufficient condition is given in the following explanation.

One fundamental property of ordinal patterns is that they are uniquely determined by the increments of the considered time series. As one can imagine in Example 4.1, the knowledge of the increments between the data points is sufficient to obtain the corresponding ordinal pattern. In mathematical terms, we can define another mapping $\tilde{\Pi}$, that assigns the corresponding ordinal pattern to each vector of increments, see [60], p. 1783.

Definition 4.3 *We define for $y = (y_1, \ldots, y_h) \in \mathbb{R}^h$ the mapping $\tilde{\Pi} : \mathbb{R}^h \to S_h$,*

$$\tilde{\Pi}(y_1, \ldots, y_h) := \Pi(0, y_1, y_1 + y_2, \ldots, y_1 + \ldots + y_h),$$

such that for $y_i = x_i - x_{i-1}$, $i = 1, \ldots, h$, we obtain

$$\begin{aligned}\tilde{\Pi}(y_1, \ldots, y_h) &= \Pi(0, y_1, y_1 + y_2, \ldots, y_1 + \ldots + y_h) \\ &= \Pi(0, x_1 - x_0, x_2 - x_0, \ldots, x_h - x_0) \\ &= \Pi(x_0, x_1, x_2, \ldots, x_h).\end{aligned}$$

A very useful tool that is self-explanatory thanks to the calculation of the asymptotic variances of the estimators later on is the symmetry property of the multivariate Gaussian distribution. We recall from (2.2) and (2.3) that for a normally distributed random vector $(Y_1, \ldots, Y_h)^t$ with componentwise equal variances, we have

$$(Y_1, \ldots, Y_h)^t \overset{\mathcal{D}}{=} (-Y_1, \ldots, -Y_h)^t \tag{4.2}$$

$$(Y_1, \ldots, Y_h)^t \overset{\mathcal{D}}{=} (Y_h, \ldots, Y_1)^t, \tag{4.3}$$

where (4.2) is called space symmetry and (4.3) time symmetry.

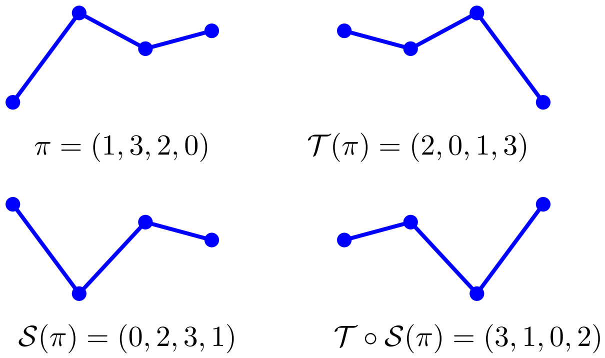

Figure 4.3: Space and time reversion of the pattern $\pi = (1,3,2,0)$.

Based on the symmetry properties above, we define the two mappings, following [60], p. 1784:

$$\mathcal{S} : S_h \to S_h, (\pi_0, \ldots, \pi_h) \to (\pi_h, \ldots, \pi_0),$$
$$\mathcal{T} : S_h \to S_h, (\pi_0, \ldots, \pi_h) \to (h - \pi_0, \ldots, h - \pi_h).$$

An illustrative understanding of these mappings is given as follows. The mapping $\mathcal{S}(\pi)$, which is the spatial reversion of the pattern π, is the reflection of π on a horizontal line, while $\mathcal{T}(\pi)$, the time reversal of π, is its reflection on a vertical line, as one can observe in Figure 4.3.

We obtain a partition of S_h by dividing it into sets $\bar{\pi}$, which are defined by

$$\bar{\pi} := \{\pi, \mathcal{S}(\pi), \mathcal{T}(\pi), \mathcal{T} \circ \mathcal{S}(\pi)\}.$$

This set is closed with respect to $\mathcal{S}$ and $\mathcal{T}$ since $\mathcal{S} \circ \mathcal{S}(\pi) = \pi$ and $\mathcal{T} \circ \mathcal{T}(\pi) = \pi$. If the considered permutation π fulfills $\mathcal{T}(\pi) = \mathcal{S}(\pi)$ the number of elements in $\bar{\pi}$ is two, otherwise it is four. In Example 4.4 this concept is shown for pattern of length $h = 2$.

Example 4.4 *For $h = 2$ all possible patterns are visualized in Figure 4.2. We can divide them into the two sets*

$$\overline{(2,1,0)} = \{(2,1,0), (0,1,2)\}, \qquad \overline{(2,0,1)} = \{(2,0,1), (1,0,2), (0,2,1), (1,2,0)\},$$

such that we have $S_3 = \overline{(2,1,0)} \cup \overline{(2,0,1)}$. Note that the first set only has two elements since $\mathcal{T}(2,1,0) = \mathcal{S}(2,1,0) = (0,1,2)$.

We remark that the pattern that is generating the reversion group is not unique and can be replaced by any other pattern that belongs to the same reversion group.
We introduce sets that include a generating pattern for each reversion group defined above. Hence, these sets are not uniquely determined either.

Definition 4.5 *We denote by $S_{[2],h}$ a set of permutations in S_h, such that each permutation π generates a different reversion class $\bar{\pi}$ with $\#\bar{\pi} = 2$. Correspondingly, $S_{[4],h}$ is the set of* *permutations, such that each* $\pi \in S_{[4],h}$ *generates a different reversion class* $\bar{\pi}$ *with* $\#\bar{\pi} = 4$. *Hence, we obtain*

$$S_h = \bigcup_{\pi \in S_{[2],h}} \{\pi, \mathcal{S}(\pi)\} \cup \bigcup_{\pi \in S_{[2],h}} \{\pi, \mathcal{S}(\pi), \mathcal{T}(\pi), \mathcal{T} \circ \mathcal{S}(\pi))\}.$$

Note that $S_h \neq S_{[2],h} \cup S_{[4],h}$.

We present examples of generating sets $S_{[2],h}$ and $S_{[4],h}$ for $h = 1$, $h = 2$ and $h = 3$.

Example 4.6 *Since* S_1 *only consists of two permutations, we have* $S_{[4],1} = \emptyset$. *Thus, a possible choice is* $S_{[2],1} = \{(1, 0)\}$.

For $h = 2$, *we choose* $S_{[2],2} = \{(2, 1, 0)\}$ *and* $S_{[4],2} = \{(2, 0, 1)\}$.

For $h = 3$, *a possible choice is given by* $S_{[2],3} = \{(3, 2, 1, 0), (3, 1, 2, 0), (2, 3, 0, 1), (2, 0, 3, 1)\}$ *and* $S_{[4],3} = \{(3, 2, 0, 1), (3, 1, 0, 2), (3, 0, 1, 2), (3, 0, 2, 1)\}$.

Based on the spatial reversion, we define another possibility to divide S_h into two disjoint sets besides $\bar{\pi}$.

Definition 4.7 *We define* S_h^* *as the set of permutations that is generated such that* $\pi \in S_h$ *is only allowed to be an element of* S_h^* *if its spatial reversal is not already contained in* S_h^*.

Note that this definition does not yield uniqueness of S_h^*.

Example 4.8 *We consider the case* $h = 2$ *again and we want to divide* S_2 *into a possible choice of* S_2^* *and the corresponding spatial reversal. We choose* $S_2^* = \{(2, 1, 0), (2, 0, 1), (1, 2, 0)\}$ *and, therefore,* $S_2 \setminus S_2^* = \{(0, 1, 2), (1, 0, 2), (0, 2, 1)\}$. *Remark that* $S_2^* = \{(0, 1, 2), (2, 0, 1), (1, 2, 0)\}$ *is also a possible choice. The only condition that has to be satisfied is that if one permutation is chosen for* S_2^*, *then its spatial reverse must not be an element of this set.*

Recall the set $\bar{\pi}$. For $h \in \mathbb{N}$, it is possible to determine the number of reversion groups $\bar{\pi}$ that have two elements. We denote this number by a_h such that $a_h = \#\{\bar{\pi} : \#\bar{\pi} = 2\}$, $\pi \in S_h$. The condition that has to be fulfilled for the elements of such a reversion group is given by $\mathcal{T} \circ \mathcal{S}(\pi) = \pi$ or in other words

$$(\pi_0, \ldots, \pi_h) = (h - \pi_h, \ldots, h - \pi_0). \tag{4.4}$$

Alternatively, $\pi_0 + \pi_h = h$, $\pi_1 + \pi_{h-1} = h$, $\ldots$, $\pi_{\frac{h-1}{2}} = \pi_{\frac{h+1}{2}}$ if h is an odd number.
We may restrict ourselves to the case that $h \in 2\mathbb{N} - 1$, as the following example shows:

Example 4.9 *We are interested in the number of reversion groups for the cases* $h = 1$ *and* $h = 2$. *For* $h = 1$ *we consider the group* $S_1 = \{(0, 1), (1, 0)\}$, $\#S_1 = 2$ *and hence* $a_1 = 1$.
In the case $h = 2$, *we obtain the same number of reversion groups, namely one, since*

$$S_2 = \{(0, 1, 2), (2, 1, 0), (2, 0, 1), (1, 0, 2), (0, 2, 1), (1, 2, 0)\},$$

and, therefore, exhibits one reversion group with two elements and one with four elements. *The second entry of the permutation is not relevant in testing condition* (4.4) *because if* $\pi_0 + \pi_2 = 2$*, it automatically yields* $\pi_1 = 1$*. This can be extended to the general case that h is an even number since* $\pi_{h/2} = h/2$ *is always fulfilled, because it is not possible to order the entries of the permutation in the way described in* (4.4) *with* $\pi_{h/2} \neq h/2$ *due to* $\pi_i \neq \pi_j$ *if* $i \neq j$*.*
Therefore, this is no constraint to the combinatorial problem we are considering and we have $a_h = a_{h+1}$*, for* $h \in 2\mathbb{N} - 1$*.*

Hence, w.l.o.g., we concentrate on the case that $h \in 2\mathbb{N} - 1$ is an odd number.

Example 4.10 *For* $h = 3$ *we have the following patterns that each generate a reversion group with two elements. For each group only one pattern is considered.*

$$\begin{aligned}&(0, 1, 2, 3),\\&(0, 2, 1, 3),\\&(1, 3, 0, 2),\\&(1, 0, 3, 2).\end{aligned}$$

We recognize the following system: the value of π_0 *can be chosen from* $\{0, 1\}$ *and uniquely determines* π_3*. Then, the number of possible orders for* π_1 *and* π_2 *is equal to the number of reversion groups* a_1 *multiplied by the number of elements in each reversion group which is by assumption given by* 2*.*

It is straightforward to extend this concept: the number of possibilities to choose π_0 from $\left\{0, \ldots, \frac{h-1}{2}\right\}$ is given by $\frac{h+1}{2}$. Therefore, we find the recursive relation

$$a_h = \frac{h+1}{2} 2a_{h-1}.$$

This can be extended to

$$a_h = \left(\prod_{j=0}^{\frac{h-1}{2}} \frac{h - 2j + 1}{2} \right) 2^{\frac{h-1}{2}} = \left(\frac{h+1}{2} \right)! \; 2^{\frac{h-1}{2}}.$$

This yields that the number of reversion groups in S_h that have four elements is given by

$$b_h := \frac{(h+1)! - 2a_h}{4} = \frac{(h+1)! - 2\left(\frac{h+1}{2}\right)! \; 2^{\frac{h-1}{2}}}{4}$$

since $\#S_h = (h + 1)!$.
In case that h is an even number, one needs to consider $b_h := \frac{(h+1)! - 2a_{h-1}}{4} = \frac{(h+1)! - 2\left(\frac{h}{2}\right)! 2^{\frac{h-2}{2}}}{4}$.

4.2 Ordinal pattern probabilities

In this section we focus on limit theorems for estimators of ordinal pattern probabilities in the long-range dependent Gaussian case. For the short-range dependent setting, results are provided in [58], Theorem 3. The transformed short-range dependent setting with a long-range dependent underlying Gaussian process is discussed in [60], Theorem 7. Therefore, in [10] we complemented these results and study the case that the long-range dependence is transmitted via the transformation by the estimator. The following results in this section are achieved in joint work with Annika Betken, Jannis Buchsteiner, Herold Dehling, Alexander Schnurr and Jeannette H.C. Woerner and can be found in [10], Section 3 and Section 5. Note that in this section we always consider a one-dimensional Gaussian process and, therefore, $d = 1$ in view of Chapter 3. Hence we do neither have any mixed cases concerning short- and long-range dependence nor have to differentiate between different long-range dependence parameters. For more details on the statistical properties of the considered estimators, we refer to [10] and [60].

In the course of this chapter the Gaussian process $(X_j)_{j\in\mathbb{Z}}$ is **always** the process in which the ordinal patterns occur. The corresponding increment process is denoted by $(Y_j)_{j\in\mathbb{Z}}$, $Y_j := X_j - X_{j-1}$, $j \in \mathbb{Z}$. Since the increment process provides all information needed to study the asymptotic distribution of the estimator of ordinal pattern probabilities, we concentrate on the necessary requirements: we assume $(Y_j)_{j\in\mathbb{Z}}$ is a univariate stationary Gaussian process as defined in (3.1), that is, $Y_j \sim \mathcal{N}(0,1)$ for all $j \in \mathbb{Z}$ and for the correlation function $r(k)$ of $(Y_j)_{j\in\mathbb{Z}}$ it holds that $|r(k)| < 1$ for $k \geq 1$ to exclude ties. Further let $(Y_j)_{j\in\mathbb{Z}}$ be long-range dependent in the sense of Definition 2.6. **Only** in Remark 4.22, we consider the case that $(X_j)_{j\in\mathbb{Z}}$ itself fulfills these properties. We recall the definition of the process $(Y_{j,h})_{j\in\mathbb{Z}}$ given by

$$Y_{j,h} = (Y_j, \ldots, Y_{j+h-1})^t$$

such that $Y_{j,h} \sim \mathcal{N}(0, \Sigma_{1,h})$ with $\Sigma_{1,h}$ as defined in (3.6). We start with the definition of the probability, which we focus on in this section:

Definition 4.11 *For $h \in \mathbb{N}$, $\pi \in S_h$ and a time series $(X_j)_{j\in\mathbb{Z}}$ we define*

$$q(\pi) := \mathbb{P}(\Pi(X_0, \ldots, X_h) = \pi)$$

as the probability of the ordinal pattern $\pi \in S_h$.

Note that, using Definition 4.3, we can represent this probability as

$$q(\pi) := q_Y(\pi) := \mathbb{P}(\Pi(X_0, \ldots, X_h) = \pi) = \mathbb{P}\left(\tilde{\Pi}(Y_1, \ldots, Y_h) = \pi\right).$$

In order to estimate this probability, we introduce the heuristic estimator $\hat{q}_{Y,n}(\pi)$ given by

$$\begin{aligned}\hat{q}_{Y,n}(\pi) := \hat{q}_n(\pi) : &= \frac{1}{n-h} \sum_{i=0}^{n-h-1} 1_{\{\Pi(X_i, X_{i+1}, \ldots, X_{i+h}) = \pi\}} \\ &= \frac{1}{n-h} \sum_{i=1}^{n-h} 1_{\{\Pi(Y_i, Y_{i+1}, \ldots, Y_{i+h-1}) = \pi\}}.\end{aligned}$$

In the following, we omit the index Y, since we are only considering a one-dimensional Gaussian process. This index is used in the next chapter, when dealing with multivariate Gaussian processes.

We turn to an improvement of the estimator $\hat{q}_n(\pi)$. In [60], p.1784 ff., it is shown that, using a Rao-Blackwellization, see [14], it is possible to modify the estimator by averaging the estimates of ordinal pattern probabilities for pattern in the same reversion groups. Essentially for the application of this method is the fact that the ordinal pattern probability of each pattern belongig to the same reversion group is equal.
We define the improved estimator by

$$
\begin{aligned}
\hat{q}_{n,Y,\mathrm{imp}}(\pi) := \hat{q}_{n,\mathrm{imp}}(\pi) :&= \frac{1}{n-h}\sum_{i=0}^{n-h-1}\frac{1}{\#\bar{\pi}}1_{\{\Pi(X_i,X_{i+1},\dots,X_{i+h})\in\bar{\pi}\}} \\
&= \frac{1}{n-h}\sum_{i=1}^{n-h}\frac{1}{\#\bar{\pi}}1_{\left\{\tilde{\Pi}(Y_i,Y_{i+1},\dots,Y_{i+h-1})\in\bar{\pi}\right\}}.
\end{aligned}
$$

This estimator has better statistical properties than the heuristic one, for example a smaller variance, as it is shown in [60], Theorem 2. For details as strong consistency and unbiasedness, we refer to [60], Theorem 2.

4.2.1 Limit theorems for estimators of $q(\pi)$

In order to get an asymptotic result for the heuristic estimator $\hat{q}_n$, we have to calculate the Hermite rank.

Lemma 4.12 *Let $(Y_j)_{j\in\mathbb{Z}}$ be a stationary Gaussian process as defined in* (3.1) *for $d = 1$. Furthermore, we fix $h \in \mathbb{N}$. Then, for any $\pi \in S_h$, the Hermite rank of*

$$
1_{\left\{\Pi(Y_j,Y_{j+1},\dots,Y_{j+h-1})=\pi\right\}} - \mathbb{P}\left(\tilde{\Pi}\left(Y_1,\dots,Y_h\right)=\pi\right)
$$

is equal to 1.

Proof. Recall that $I_{1,h}$ denotes the h-dimensional identity matrix. Since ordinal patterns are not affected by scaling, we may assume that $(\Sigma_{1,h}^{-1} - I_{1,h})$ is positive semidefinite, for details see [10], Remark 5.5. According to Lemma 2.33 it suffices to show $\mathbb{E}\left(U_k\mathbf{1}_{\{\tilde{\Pi}(U_1,\dots,U_h)=\pi\}}\right) \neq 0$ for some independent standard normal random variables $U_1,\dots,U_h$ and some $1 \le k \le h$. For simplicity, we consider the pattern $\pi = (h,\dots,0)$ that corresponds to the event

$$
\{U_i \ge 0,\ i = 1,\dots,h\}.
$$

Hence, we arrive at

$$
\begin{aligned}
\mathbb{E}\left(U_1\mathbf{1}_{\{U_1\ge 0,\dots,U_h\ge 0\}}\right) &= \int_0^\infty\int_0^\infty\dots\int_0^\infty u_1\varphi(u_1)\varphi(u_2)\cdots\varphi(u_h)\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_h \\
&= \left(\frac{1}{2}\right)^{h-1}\varphi(0) \neq 0.
\end{aligned}
$$

It follows by the same reasoning that none of the expected values that correspond to the other ordinal patterns equals zero. □

We observe, that for the set of discontinuity points D_π of

$$1_{\{\Pi(Y_j,Y_{j+1},\ldots,Y_{j+h-1})=\pi\}} - \mathbb{P}\left(\tilde{\Pi}\left(Y_1,\ldots,Y_h\right)=\pi\right)$$

it holds that $\mathbb{P}\left(Y_{j,h}\in D_\pi\right)=0$. We exemplarily show this for $\pi=(h,\ldots,0)$. Then, we have

$$D_\pi = \{\{0\}\times\mathbb{R}^{h-1}\}\cup\ldots\cup\{\mathbb{R}^{h-1}\times\{0\}\}. \tag{4.5}$$

We observe that each of the sets on the right-hand side in (4.5) is a Null set with respect to the h-dimensional Lebesgue measure. The other ordinal patterns $\pi\in S_h$ can be treated analogously, since we explicitly excluded $|r(k)|=1$ for $k\geq 1$. With this knowledge, we can use Theorem 3.3 for $d=1$ and $h\in\mathbb{N}$ and derive the corresponding limit theorem.

Theorem 4.13 *Let $(Y_j)_{j\in\mathbb{Z}}$ be a stationary Gaussian process as defined in (3.1) for $d=1$ that is long-range dependent in the sense of Definition 2.6. Then, we obtain*

$$n^{-d_1+\frac{1}{2}}C_1^{-\frac{1}{2}}L_{1,1}^{-\frac{1}{2}}(n)\left(\hat{q}_n(\pi)-q(\pi)\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\left(\sum_{j=1}^{h}\alpha_j\right)^2\right),$$

with $C_1=\frac{1}{d_1(2d_1+1)}$, $\alpha=(\alpha_1,\ldots,\alpha_h)^t=\Sigma_{1,h}^{-1}c$ and $c=\mathbb{E}\left(1_{\{\Pi(Y_1,Y_2,\ldots,Y_h)=\pi\}}\left(Y_1,\ldots,Y_h\right)^t\right)$.

Remark 4.14 *As already mentioned in the introduction of this section, the corresponding limit theorem of this estimator in the case of short-range dependence is given in [58], Theorem 3.*

We want to take a closer look at the limit variance of the asymptotic distribution and hence need to compute α and c. In [60], p.1786 and Lemma 1 it is shown that, based on the space and time symmetry in (4.2) and (4.3), it holds that

$$\mathbb{E}\left(Y_i 1_{\{\Pi(Y_1,Y_2,\ldots,Y_h)=\pi\}}\right) = -\mathbb{E}\left(Y_i 1_{\{\Pi(Y_1,Y_2,\ldots,Y_h)=\mathcal{S}(\pi)\}}\right), \qquad i=1,\ldots,h, \tag{4.6}$$

$$\mathbb{E}\left(Y_i 1_{\{\Pi(Y_1,Y_2,\ldots,Y_h)=\pi\}}\right) = -\mathbb{E}\left(Y_{h+1-i} 1_{\{\Pi(Y_1,Y_2,\ldots,Y_h)=\mathcal{T}(\pi)\}}\right), \qquad i=1,\ldots,h. \tag{4.7}$$

The justification for the second equation is given in the following calculation since ordinal patterns are not affected by monotone transformations: for $\pi\in S_h$, we obtain

$$\begin{aligned}\left\{\tilde{\Pi}\left(Y_1,...,Y_h\right)=\mathcal{T}(\pi)\right\} &= \left\{\mathcal{T}\left(\tilde{\Pi}(Y_1,\ldots,Y_h)\right)=\pi\right\}\\ &= \left\{\mathcal{T}\left(\Pi(0,Y_1,Y_1+Y_2,\ldots,Y_1+\ldots+Y_h)\right)=\pi\right\}\\ &= \left\{\Pi(Y_1+\ldots+Y_h,\ldots,Y_1+Y_2,Y_1,0)=\pi\right\}\\ &= \left\{\Pi(0,-Y_h,-(Y_h+Y_{h-1}),\ldots,-(Y_1+...+Y_h))=\pi\right\}\\ &= \left\{\tilde{\Pi}(-Y_h,-Y_{h-1},\ldots,-Y_2,-Y_1)=\pi\right\}.\end{aligned} \tag{4.8}$$

Given these symmetry properties it is sufficient to determine the Hermite coefficients for one pattern π of each reversion group. For $h=1$ we, therefore, consider $\pi=(1,0)$, since only one reversion group appears in this case, and following Example 4.4 for $h=2$ we can restrict ourselves to the pattern $\pi=(2,1,0)$ and $\pi=(2,0,1)$. Note that larger values of h exceed the computing capacity of Mathematica.

Example 4.15 *In the case $h=1$ there are only two possible patterns: $\pi=(0,1)$ and the corresponding spatial (or time) reverse $\pi=(1,0)$. We focus on $\pi=(1,0)$. This pattern corresponds to the event $\{\Pi(X_0,X_1)=(1,0)\}=\{X_1\geq X_0\}=\{Y_1\geq 0\}$. Hence, we consider*

$$c_1=\mathbb{E}\left(Y_1 1_{\{Y_1\geq 0\}}\right)=\int_0^\infty y_1\varphi(y_1)\mathrm{d}y_1=\varphi(0).$$

Correspondingly, we obtain $c_1=-\varphi(0)$ for $\pi=(0,1)$, since this is the spatial reversion of $(1,0)$. Thus, for these two ordinal patterns we arrive at a limit distribution of $q_n(\pi)$ given by $\mathcal{N}\left(0,\varphi^2(0)\right)$.

Example 4.16 *First, we study the limit variance for $\pi=(2,1,0)$. In this case, $\bar{\pi}$ has two elements. Note that $\{\Pi(X_0,X_1,X_2)=(2,1,0)\}=\{X_2\geq X_1\geq X_0\}=\{Y_2\geq 0,Y_1\geq 0\}$. Due to the symmetry of the bivariate normal distribution, we obtain $c_1=c_2$ so that we only need to calculate*

$$c_1=\mathbb{E}\left(Y_1 1_{\{Y_2\geq 0,Y_1\geq 0\}}\right)=\int_0^\infty\int_0^\infty y_1\varphi_{(Y_1,Y_2)}(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2=\frac{\varphi(0)}{2}\left(1+r^{(1,1)}(1)\right),$$

where $\varphi_{(Y_1,Y_2)}$ denotes the joint density of (Y_1,Y_2). Hence,

$$\sum_{j=1}^{2}\alpha_j=2c_1\left(g_{1,1}+g_{2,1}\right)=2c_1\frac{1-r^{(1,1)}(1)}{1-\left(r^{(1,1)}(1)\right)^2}=\varphi(0),$$

where $g_{i,j}$ are the entries of $\Sigma_{1,2}^{-1}$ given by

$$\Sigma_2^{-1}=\frac{1}{1-(r^{(1,1)}(1))^2}\left(\begin{pmatrix}1 & -r^{(1,1)}(1)\\ -r^{(1,1)}(1) & 1\end{pmatrix}\right).$$

Again, we obtain the limit variance $\varphi^2(0)$, which for $h=2$ is more surprising than in the case $h=1$, because the result is independent of $r^{(1,1)}(1)$. For the space reverse pattern $\pi_2=(0,1,2)$ we apply (4.6) and obtain $c_1=-\varphi(0)$, leading to the same limit variance. Note that another approach to derive this limit variance using the Cholesky decomposition is given in Section A.1. It is an interesting question, whether it is just a coincidence that this variance is independent of the covariance between the increments. The answer turns out to be yes, since the dependence is reflected in the limit variance of the pattern $\pi=(2,0,1)$.
Note that

$$\begin{aligned}\{\Pi(X_0,X_1,X_2)=(2,0,1)\}&=\{X_1\leq X_0\leq X_2\}\\&=\{Y_1\leq 0,Y_1+Y_2\geq 0\}\\&=\{Y_1\leq 0,Y_2\geq -Y_1\}.\end{aligned}$$

As a result, we have

$$c_1 = \mathbb{E}\left(Y_1 1_{\{Y_1 \le 0, Y_2 \ge -Y_1\}}\right) = \int_{-\infty}^{0}\int_{-y_1}^{\infty} y_1 \varphi_{(Y_1,Y_2)}(y_1, y_2)\mathrm{d}y_2\mathrm{d}y_1$$
$$= \frac{\varphi(0)}{2}\left(\frac{\sqrt{1 + r^{(1,1)}(1)}}{\sqrt{2}} - 1\right)$$

and

$$c_2 = \mathbb{E}\left(Y_2 1_{\{Y_1 \le 0, Y_2 \ge -Y_1\}}\right) = \int_{-\infty}^{0}\int_{-y_1}^{\infty} y_2 \varphi_{(Y_1,Y_2)}(y_1, y_2)\mathrm{d}y_2\mathrm{d}y_1$$
$$= \frac{\varphi(0)}{2}\left(\frac{\sqrt{1 + r^{(1,1)}(1)}}{\sqrt{2}} - r^{(1,1)}(1)\right),$$

where $\varphi_{(Y_1,Y_2)}$ *denotes the joint density of* (Y_1, Y_2). *Finally, we obtain*

$$\sum_{j=1}^{2} \alpha_j = (c_1 + c_2)(g_{1,1} + g_{2,1}) = \frac{\varphi(0)}{2}\left(\frac{\sqrt{2(1 + r^{(1,1)}(1))} - (1 + r^{(1,1)}(1))}{1 + r^{(1,1)}(1)}\right)$$
$$= \frac{\varphi(0)}{2}\left(\frac{\sqrt{2}}{\sqrt{1 + r^{(1,1)}(1)}} - 1\right).$$

The above expression depends on $r^{(1,1)}(1)$. *Due to space and time symmetry discussed in* (4.6) *and* (4.7) *all permutations that belong to the reversion group of* $\pi = (2, 0, 1)$, *i.e.,* $(1, 0, 2), (0, 2, 1)$ *and* $(1, 2, 0)$, *lead to the same limit variance for* $\hat{q}_n(\pi)$, *namely* $\mathcal{N}\left(0, \left(\frac{\varphi(0)}{2}\left(\frac{\sqrt{2}}{\sqrt{1+r^{(1,1)}(1)}} - 1\right)\right)^2\right)$.

By now, we have investigated the heuristic estimator for $q(\pi)$ based on the relative frequency of the ordinal pattern and have derived an asymptotic distribution whose variance can be explicitly stated for the cases $h = 1$ and $h = 2$ in terms of the correlation function of the underlying long-range dependent Gaussian process.

As already mentioned at the beginning of this section, another estimator $\hat{q}_{n,\mathrm{imp}}(\pi)$ was proposed in [60] by using the statistical method of Rao-Blackwellization. We turn to this estimator in the following. Since we are interested in the asymptotic distribution, we need to do some preliminary work in advance and calculate the Hermite rank of $\hat{q}_{n,\mathrm{imp}}(\pi)$. Since it was already proved in [60], p.1786, that the Hermite rank of this estimator is larger or equal to 2 with respect to $\Sigma_{1,h}$, we can restrict our considerations to verifying that the Hermite rank $m\left(\hat{q}_{n,\mathrm{imp}}(\pi), I_{1,h}\right) \le 2$ in order to show $m\left(\hat{q}_{n,\mathrm{imp}}(\pi), \Sigma_{1,h}\right) = 2$.

Lemma 4.17 *Let* $(Y_j)_{j\in\mathbb{Z}}$ *be a stationary Gaussian process as defined in* (3.1) *for* $d = 1$. *Then, for any* $\pi \in S_h$ *the Hermite rank of*

$$\frac{1}{\#\bar{\pi}} 1_{\{\tilde{\Pi}(Y_i, Y_{i+1}, \ldots, Y_{i+h-1}) \in \bar{\pi}\}} - \frac{1}{\#\bar{\pi}}\mathbb{P}\left(\tilde{\Pi}\left(Y_1, \ldots, Y_h\right) \in \bar{\pi}\right)$$

with respect to $\Sigma_{1,h}$ *equals* 2.

Proof. In order to proof the result, we start with the investigation of the second order Hermite coefficients of the original estimator $\hat{q}_n(\pi)$, $\pi \in S_h$. For a multivariate random vector $(Y_1, ..., Y_h) \sim \mathcal{N}(0, \Sigma_{1,h})$ define

$$\begin{aligned} c_{i,i}^{\pi} &:= \mathbb{E}\left[\left(Y_i^2 - 1\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\pi\right\}}\right] = \mathbb{E}\left[\left(Y_i^2 - 1\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\mathcal{S}(\pi)\right\}}\right] \\ &= \mathbb{E}\left[\left(Y_{h+1-i}^2 - 1\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\mathcal{T}(\pi)\right\}}\right], \ i = 1, ..., h. \end{aligned}$$

Analogously, we define

$$\begin{aligned} c_{i,j}^{\pi} :&= \mathbb{E}\left[\left(Y_i Y_j - \mathbb{E}\left(Y_i Y_j\right)\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\pi\right\}}\right] \\ &= \mathbb{E}\left[\left(Y_i Y_j - \mathbb{E}\left(Y_i Y_j\right)\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\mathcal{S}(\pi)\right\}}\right] \\ &= \mathbb{E}\left[\left(Y_{h+1-i} Y_{h+1-j} - \mathbb{E}\left(Y_{h+1-i} Y_{h+1-j}\right)\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\mathcal{T}(\pi)\right\}}\right], \end{aligned}$$

with $i, j = 1, ..., h, \ i \neq j$, so that altogether we obtain

$$c_{i,j}^{\pi} = c_{i,j}^{\mathcal{S}(\pi)} = c_{h+1-i,h+1-j}^{\mathcal{T}(\pi)} = c_{h+1-i,h+1-j}^{\mathcal{T}\circ\mathcal{S}(\pi)}, \ i, j = 1, ..., h. \tag{4.9}$$

With this result we can simplify the second order Hermite coefficients for the improved estimator by

$$\begin{aligned} c_{i,j} := \mathbb{E}\left[\left(Y_i Y_j - \mathbb{E}\left(Y_i Y_j\right)\right) f\left(Y_1, ..., Y_h\right)\right] &= \frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} \mathbb{E}\left[\left(Y_i Y_j - \mathbb{E}\left(Y_i Y_j\right)\right) 1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)=\pi\right\}}\right] \\ &= \frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} c_{i,j}^{\pi} \\ &= \frac{1}{2}\left(c_{i,j}^{\pi} + c_{h+1-i,h+1-j}^{\pi}\right). \end{aligned}$$

In a similar way, we obtain

$$c_{i,i} = \frac{1}{2}\left(c_{i,i}^{\pi} + c_{h+1-i,h+1-i}^{\pi}\right).$$

Hence, we can uniquely determine the second order Hermite coefficients of the improved estimator by calculating the second order Hermite coefficients for only one pattern π that belongs to the considered reversion group $\bar{\pi}$. By following the symmetry properties discussed in (4.9), we derive for the special case $\mathcal{T} \circ \mathcal{S}(\pi) = \pi$ that

$$c_{i,j} = c_{i,j}^{\pi}, \quad \pi \in \bar{\pi}$$

for all $i, j = 1, ..., h$. Hence, the second order Hermite coefficients of the improved estimator $\hat{q}_{n,\text{imp}}(\pi)$ are equal to the second order Hermite coefficients of $\hat{q}_n(\pi)$.

We use this result to determine the Hermite rank of $\hat{q}_{n,\text{imp}}(\pi)$. Let $U = (U_1, \ldots, U_h)^t$ be a standard Gaussian random vector, i.e., with covariance matrix $I_{1,h}$. Following the arguments above, we only need to consider the second order Hermite coefficients of $\hat{q}_n(\pi)$ for a fixed pattern $\pi \in \bar{\pi}$:

$$\begin{aligned} b_{jk}^{\pi} &= \mathbb{E}\left(U_k U_j 1_{\left\{\tilde{\Pi}(U_1,...,U_h)=\pi\right\}}\right), & 1 \leq k < j \leq h, \text{ and} \\ b_{jj}^{\pi} &= \mathbb{E}\left(\left(U_j^2 - 1\right) 1_{\left\{\tilde{\Pi}(U_1,...,U_h)=\pi\right\}}\right), & j = 1, ..., h. \end{aligned}$$

For simplicity we regard $\pi = (h, h-1, ..., 0)$. Note that for this pattern it suffices to show that $b_{jk}^{\pi} \neq 0$, since in this case $c_{i,j} = c_{i,j}^{\pi}$ for $i, j = 1, ..., h$. For $j \neq k$, it holds

$$\begin{aligned} b_{jk}^{\pi} &= \int_0^{\infty} ... \int_0^{\infty} u_j u_k \varphi(u_1)...\varphi(u_h) \mathrm{d}u_1 ... \mathrm{d}u_h \\ &= \frac{1}{2^{h-2}} \int_0^{\infty} u_j \varphi(u_j) \mathrm{d}u_j \int_0^{\infty} u_k \varphi(u_k) \mathrm{d}u_k \\ &= \frac{\varphi^2(0)}{2^{h-2}}. \end{aligned}$$

□

In contrast to the heuristic estimator $\hat{q}_n(\pi)$, we are dealing with an underlying long-range dependent process $(Y_j)_{j\in\mathbb{Z}}$ and an estimator with Hermite rank equal to 2 in case of the improved estimator $\hat{q}_{n,\mathrm{imp}}$. Therefore, a limit theorem with a different normalization and a different asymptotic distribution holds, as we can observe applying Theorem 3.10 for $d = 1$ and $h \in \mathbb{N}$. Note that we may apply Theorem 3.10 as for each π the set of discontinuity points of

$$\frac{1}{\#\bar{\pi}} 1_{\left\{\tilde{\Pi}(Y_i, Y_{i+1}, ..., Y_{i+h-1}) \in \bar{\pi}\right\}} - \frac{1}{\#\bar{\pi}} \mathbb{P}\left(\tilde{\Pi}\left(Y_1, \ldots, Y_h\right) \in \bar{\pi}\right)$$

is a Null set with respect to the h-dimensional Lebesgue measure due to the considerations in (4.5).

Theorem 4.18 *Let $(Y_j)_{j\in\mathbb{Z}}$ be a stationary Gaussian process as defined in (3.1) for $d = 1$ that is long-range dependent in the sense of Definition 2.6. Then, if $d_1 \in \left(\frac{1}{4}, \frac{1}{2}\right)$, we obtain*

$$n^{1-2d_1} (2C_2)^{-\frac{1}{2}} L^{-1}(n) \left(\hat{q}_{n,\mathrm{imp}}(\pi) - q(\pi)\right) \xrightarrow{\mathcal{D}} Z_{2,d_1+\frac{1}{2}}(1) \sum_{j=1}^{h} \sum_{k=1}^{h} \alpha_{j,k}, \tag{4.10}$$

where $Z_{2,d_1+\frac{1}{2}}(1)$ is a standard Rosenblatt random variable and $C_2 := \frac{1}{2d^*(4d^*-1)}$,

$$(\alpha_{l,k})_{1\leq l,k\leq h} = \Sigma_{1,h}^{-1} C \Sigma_{1,h}^{-1}$$

and

$$C = \mathbb{E}\left((Y_1, ..., Y_h) \frac{1}{\#\bar{\pi}} \left[1_{\left\{\tilde{\Pi}(Y_1,...,Y_h)\in\bar{\pi}\right\}} - \mathbb{P}\left(\tilde{\Pi}\left(Y_1, ..., Y_h\right) \in \bar{\pi}\right)\right] (Y_1, ..., Y_h)^t \right).$$

Proof. The proof is an immediate consequence of Theorem 3.10 for $d = 1$ and $h \in \mathbb{N}$. □

Remark 4.19 *As already mentioned in the introduction of this section, the corresponding result for $d_1 \in \left(0, \frac{1}{4}\right)$ can be found in [60], Theorem 7. Note that the asymptotic distribution is Gaussian in this case.*

Analogously to the asymptotic distribution of the heuristic estimator we want to get more information on the asymptotic variance here and, therefore, need to calculate α and C again. We also restrict ourselves to the cases $h = 1$ and $h = 2$ since they have the most practical impact and yield closed formulas for the asymptotic variance.

Example 4.20 (The case h = 1) *For* $\pi = (1, 0)$ *we derive*

$$\{\Pi(X_0, X_1) = (1, 0)\} = \{X_0 \leq X_1\} = \{Y_1 \geq 0\}$$

and, therefore,

$$c_{1,1} = \mathbb{E}\left[\left(Y_1^2 - 1\right) 1_{\{Y_1 \geq 0\}}\right] = \int_0^\infty \left(y_1^2 - 1\right) \varphi(y_1)\mathrm{d}y_1 = 0.$$

So in the trivial case, that is only one increment variable, we derive a degenerate limit distribution again.

For increments of length $h = 2$ we used Mathematica to calculate the integrals in the Hermite coefficients.

Example 4.21 (The case $h = 2$**)** *First, we consider the pattern* $\pi = (2, 1, 0)$ *and the corresponding event* $\{\Pi(X_0, X_1, X_2) = (2, 1, 0)\} = \{X_2 \geq X_1 \geq X_0\} = \{Y_1 \geq 0, Y_2 \geq 0\}$*. We know that* $c_{i,j} = c^\pi_{i,j}$*,* $i, j = 1, 2$*, by the calculations done in the proof of Lemma 4.17. Moreover by* (4.9)*, we derive that* $c_{1,1} = c_{2,2}$ *since* $\mathcal{T} \circ \mathcal{S}(2, 1, 0) = (2, 1, 0)$*. We have*

$$\begin{aligned} c_{1,1} = \mathbb{E}\left[\left(Y_1^2 - 1\right) 1_{\{Y_1 \geq 0, Y_2 \geq 0\}}\right] &= \int_0^\infty \int_0^\infty \left(y_1^2 - 1\right) \varphi_{(Y_1,Y_2)}(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &= \varphi^2(0) r^{(1,1)}(1)\sqrt{1 - \left(r^{(1,1)}(1)\right)^2} \end{aligned}$$

and

$$\begin{aligned} c_{1,2} &= \mathbb{E}\left[\left(Y_1Y_2 - \mathbb{E}\left(Y_1Y_2\right)\right) 1_{\{Y_1 \geq 0, Y_2 \geq 0\}}\right] \\ &= \int_0^\infty \int_0^\infty y_1 y_2 \varphi_{(Y_1,Y_2)}(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 - r^{(1,1)}(1) \int_0^\infty \int_0^\infty \varphi_{(Y_1,Y_2)}(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &= \varphi^2(0)\sqrt{1 - \left(r^{(1,1)}(1)\right)^2}. \end{aligned}$$

This yields

$$\begin{aligned}\sum_{i,j=1}^{2} \alpha_{i,j} &= 2\left(g_{1,2}+g_{2,2}\right)^2\left(c_{1,1}+c_{1,2}\right)\\ &= 2\frac{c_{1,1}+c_{1,2}}{\left(1+r^{(1,1)}(1)\right)^2}\\ &= 2\varphi^2(0)\sqrt{\frac{1-r^{(1,1)}(1)}{1+r^{(1,1)}(1)}}.\end{aligned}$$

For $\pi=(2,1,0)$ *the left-hand side in* (4.10) *converges in distribution to* $2\varphi^2(0)\sqrt{\frac{1-r^{(1,1)}(1)}{1+r^{(1,1)}(1)}}Z_{2,d_1+\frac{1}{2}}(1)$.

Consider the pattern $\pi=(2,0,1)$ *and the corresponding event*

$$\begin{aligned}\{\Pi(X_0,X_1,X_2)=(2,0,1)\} &= \{X_1\le X_0\le X_2\}\\ &= \{Y_1\le 0, Y_1+Y_2\ge 0\}.\end{aligned}$$

It holds that

$$\begin{aligned}c^{\pi}_{1,1} = \mathbb{E}\left[\left(Y_1^2-1\right)1_{\{Y_1\le 0, Y_2\ge -Y_1\}}\right] &= \int_{-\infty}^{0}\int_{-y_1}^{\infty}\left(y_1^2-1\right)\varphi_{(Y_1,Y_2)}(y_1,y_2)\mathrm{d}y_2\mathrm{d}y_1\\ &= -\varphi^2(0)\frac{\sqrt{1-\left(r^{(1,1)}(1)\right)^2}}{2}\end{aligned}$$

and

$$\begin{aligned}c^{\pi}_{1,2} &= \mathbb{E}\left[\left(Y_1Y_2-\mathbb{E}\left(Y_1Y_2\right)\right)1_{\{Y_1\le 0, Y_2\ge -Y_1\}}\right]\\ &= \int_{-\infty}^{0}\int_{-y_1}^{\infty} y_1y_2\varphi_{(Y_1,Y_2)}(y_1,y_2)\mathrm{d}y_2\mathrm{d}y_1 - r^{(1,1)}(1)\int_{-\infty}^{0}\int_{-y_1}^{\infty}\varphi_{(Y_1,Y_2)}(y_1,y_2)\mathrm{d}y_2\mathrm{d}y_1\\ &= -\varphi^2(0)\frac{\sqrt{1-\left(r^{(1,1)}(1)\right)^2}}{2}.\end{aligned}$$

Since the reversion group of this pattern has four elements we also need to calculate

$$\begin{aligned}c^{\pi}_{2,2} = \mathbb{E}\left[\left(Y_2^2-1\right)1_{\{Y_1\le 0, Y_2\ge -Y_1\}}\right] &= \int_{-\infty}^{0}\int_{-y_1}^{\infty}\left(y_2^2-1\right)\varphi_{(Y_1,Y_2)}(y_1,y_2)\mathrm{d}y_2\mathrm{d}y_1\\ &= -\varphi^2(0)\frac{\sqrt{1-\left(r^{(1,1)}(1)\right)^2}(2r^{(1,1)}(1)-1)}{2}.\end{aligned}$$

Altogether we arrive at

$$\begin{aligned}\sum_{i,j=1}^{2}\alpha_{i,j} &= \frac{1}{\left(1+r^{(1,1)}(1)\right)^2}\left(c_{1,1}+2c_{1,2}+c_{2,2}\right)\\ &= \frac{1}{\left(1+r^{(1,1)}(1)\right)^2}\left(c^{\pi}_{1,1}+2c^{\pi}_{1,2}+c^{\pi}_{2,2}\right)\\ &= -\varphi^2(0)\frac{\sqrt{1-\left(r^{(1,1)}(1)\right)^2}}{\left(1+r^{(1,1)}(1)\right)^2}\left(r^{(1,1)}(1)+1\right)\\ &= -\varphi^2(0)\sqrt{\frac{1-r^{(1,1)}(1)}{1+r^{(1,1)}(1)}}.\end{aligned}$$

For $\pi=(2,0,1)$ *the left-hand side in* (4.10) *converges in distribution to* $-\varphi^2(0)\sqrt{\frac{1-r^{(1,1)}(1)}{1+r^{(1,1)}(1)}}Z_{2,H}(1)$.

Remark 4.22 *One might wonder which limit theorems can be derived in the special case that it is not only the increment process* $(Y_j)_{j\in\mathbb{Z}}$ *but the time series* $(X_j)_{j\in\mathbb{Z}}$ *itself which is stationary and long-range dependent.*
We have to determine the Hermite rank of the estimator $\hat{q}_n(\pi)$ *in this setting.*
Let $(X_k)_{k\in\mathbb{Z}}$ *be a stationary, long-range dependent, standard normal Gaussian process and let* $h\in\mathbb{N}$. *By Lemma 2.33 it is enough to show that* $\mathbb{E}(U_k 1_{\{\Pi(U_0,\ldots,U_h)=\pi\}})\neq 0$ *for some independent standard normal random variables* $U_0,\ldots,U_h$ *and some* $0\leq k\leq h$. *Without loss of generality let* $\pi=\mathrm{id}$ *and set* $k=0$. *This yields*

$$\begin{aligned}\mathbb{E}(U_0 1_{\{U_0\leq\ldots\leq U_h\}}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{u_{h+1}}\int_{-\infty}^{u_h}\cdots\int_{-\infty}^{u_2} u_1\varphi(u_1)\varphi(u_2)\cdots\varphi(u_{h+1})du_1du_2\cdots du_{h+1}\\ &= -\int_{-\infty}^{\infty}\int_{-\infty}^{u_{h+1}}\cdots\int_{-\infty}^{u_3}\varphi(u_2)^2\varphi(u_3)\cdots\varphi(u_{h+1})du_2\cdots du_{h+1}\\ &\neq 0\end{aligned}$$

since we integrate a strictly positive function. Hence, for any $\pi\in S_h$ *the Hermite rank of the function* $f:\mathbb{R}^{h+1}\longrightarrow\mathbb{R}$, *defined by*

$$f(x_0,x_1,\ldots,x_h):=1_{\{\Pi(x_0,\ldots,x_h)=\pi\}}-\mathbb{P}(\Pi(x_0,\ldots,x_h)=\pi)$$

is equal to 1.

We get the following asymptotic result concerning the ordinal pattern probability estimator $\hat{q}_n(\pi)$ *in this modified setting by applying Corollary 3.18 for* $d=1$:

$$n^{-d_1+\frac{1}{2}}L_{1,1}^{-1/2}(n)\left(\hat{q}_n(\pi)-\mathbb{P}\left(\Pi(X_0,\ldots,X_h)=\pi\right)\right)\xrightarrow{\mathcal{D}}\delta_0,$$

where δ_0 *denotes the Dirac measure in* 0. *In this special case, the limit distribution for* $\hat{q}_n(\pi)$ *is trivial.*

However, taking the classical rate of convergence $n^{1/2}$, we get a non-trivial Gaussian central limit theorem, see [54], Theorem 5.4.1, since the increment process of $(X_j)_{j\in\mathbb{Z}}$ is short-range dependent, as it is explained at the end of Section 3.3.

We have proved two limit theorems for estimators of ordinal pattern probabilities in Theorem 4.13 and in Theorem 4.18. In the latter result a non-Gaussian limit distribution occurs, called the Rosenblatt distribution. We turn to a simulation study to verify the theoretical results and compare the simulated limit distributions to the standard normal distribution as well as to the Rosenblatt distribution as described in Section 2.7.3.

4.2.2 Simulation studies

We simulate $N = 10\,000$ paths of fractional Gaussian noise by the command "simFGN0" from the RPackage "longmemo", see [34] with sample size $n = 10^6$ for different values of $H = d_1 + \frac{1}{2}$, namely $d_1 = 0,3$ and $d_1 = 0.4$, to compare the asymptotic distribution of the estimators $\hat{q}_n(\pi)$ and $\hat{q}_{n,\text{imp}}(\pi)$ with the theoretical results derived above.

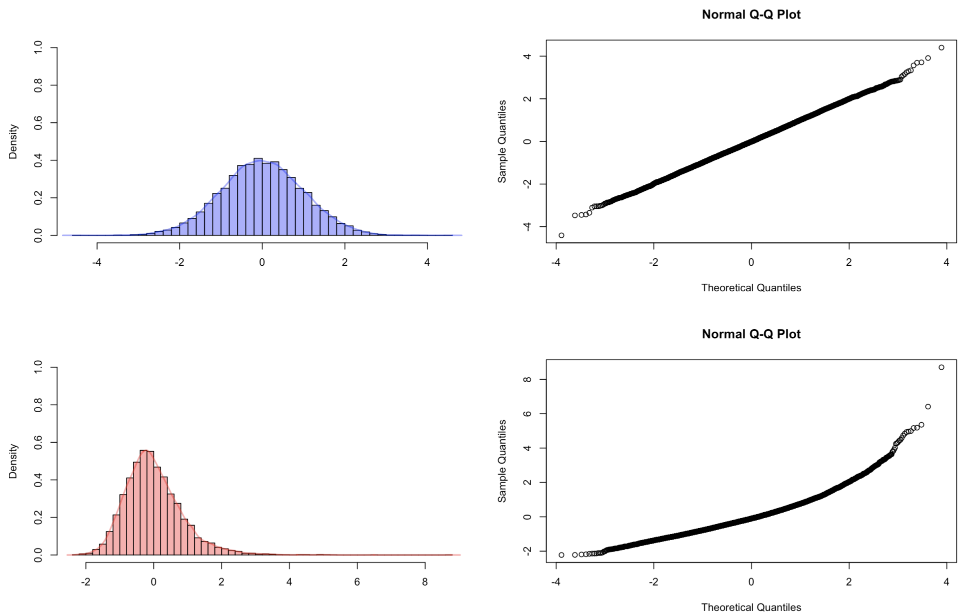

Figure 4.4: Histogram, kernel density estimation and Q-Q plot of the asymptotic distribution of the estimators $\hat{q}_n(\pi)$ (blue) and $\hat{q}_{n,\text{imp}}(\pi)$ (red) for $n = 10^6$ and $\pi = (2,1,0)$ in the case $H = 0.8$.

We standardized the estimators following the normalization constants given in Theorem 4.13 and Theorem 4.18. The results depending on the respective long-range dependence parameter d_1 are displayed in Figure 4.4 and in Figure 4.5. We observe the standard normal distribution of the limit of $\hat{q}_n(\pi)$ for $H = 0.8$ and $H = 0.9$ as shown in Theorem 4.13.
For the improved estimator $\hat{q}_{n,\text{imp}}(\pi)$ a non-Gaussian limit distribution is verified in the Q-Q plots. We observe the densities of the standard Rosenblatt distribution for the respective parameter H as shown in Figure 2.2.

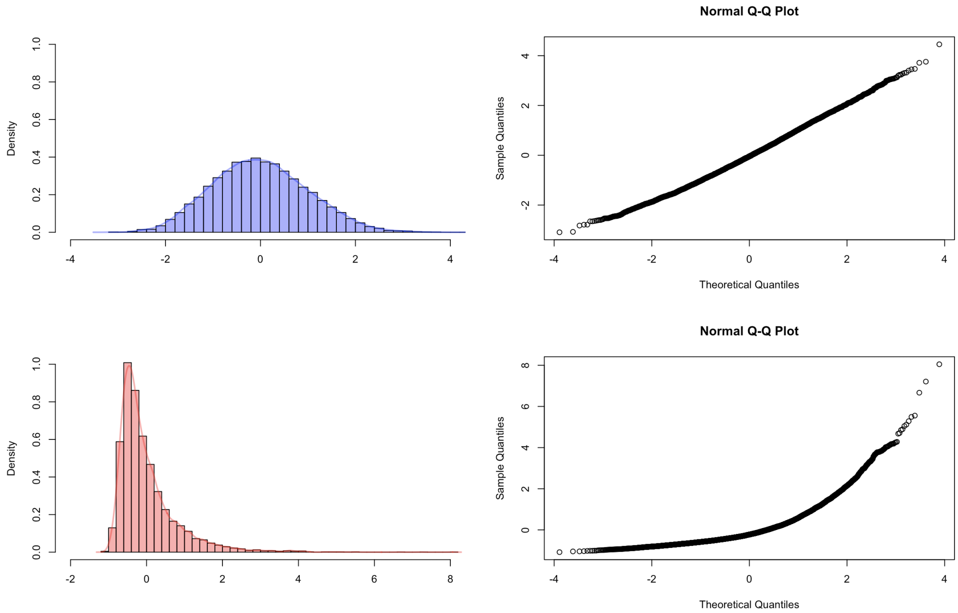

Figure 4.5: Histogram, kernel density estimation and Q-Q plot of the asymptotic distribution of the estimators $\hat{q}_n(\pi)$ (blue) and $\hat{q}_{n,\text{imp}}(\pi)$ (red) for $n = 10^6$ and $\pi = (2, 1, 0)$ in the case $H = 0.9$.

4.3 Ordinal pattern dependence

Ordinal pattern dependence is generally speaking a measure of dependence between time series $\left(X_j^{(1)}\right)_{j\in\mathbb{Z}}$ and $\left(X_j^{(2)}\right)_{j\in\mathbb{Z}}$ that is only based on the ordinal information given in a moving window of a fixed length. First introduced in [57], it was used to model dependence of financial data sets. With [58], the first work was published that deals with ordinal pattern dependence in a theoretical context, i.a., with limit theorems for the underlying bivariate time series being a 1-approximating functional of an absolutely regular process. For details on the aforementioned, we refer to the article. The authors derived a Gaussian limit distribution for an estimator of ordinal pattern dependence. Lately, ordinal pattern dependence was used in various applications, see [24] for an analysis of hydrological data sets, [52] and [50] for applications to extreme value theory and [49] and [11] for an integration of ordinal pattern dependence in the context of other univariate and multivariate dependence measures. In the latter field of research, it is important to mention that ordinal pattern dependence is defined in a time series context in the first place, i.e., to detect dependence structures between time series. However, multivariate dependence measures are often defined to describe dependencies between two random vectors or even within one random vector. Having this in mind, we present the comparison to several other measures in more detail in Chapter 5.

We stick to the formal definition of ordinal pattern dependence, as it is proposed in [58], Sec. 2.1. The considered moving window consists of $h + 1$ data points and, hence, h increments.

We define

$$p := p_{X^{(1)},X^{(2)},h} := \mathbb{P}\left(\Pi\left(X_0^{(1)},\ldots,X_h^{(1)}\right) = \Pi\left(X_0^{(2)},\ldots,X_h^{(2)}\right)\right) \tag{4.11}$$

and

$$q := q_{X^{(1)},X^{(2)},h} := \sum_{\pi\in S_h} \mathbb{P}\left(\Pi\left(X_0^{(1)},\ldots,X_h^{(1)}\right) = \pi\right)\mathbb{P}\left(\Pi\left(X_0^{(2)},\ldots,X_h^{(2)}\right) = \pi\right). \tag{4.12}$$

Then, we define ordinal pattern dependence OPD as

$$OPD := OPD_{X^{(1)},X^{(2)},h} := \frac{p-q}{1-q}. \tag{4.13}$$

The parameter q represents the hypothetical case of independence between the two time series. In this case p and q would obtain equal values and, therefore, OPD would equal zero. Regarding the other extreme, the case in which both processes coincide, we obtain the value 1. However, in the following, we assume $p \in (0,1)$ and $q \in (0,1)$.
Note that the definition of ordinal pattern dependence in (4.13) only measures positive dependence. This is no restriction in practice, because negative dependence can be investigated in an analogous way, by considering $OPD_{X^{(1)},-X^{(2)}}$. If one is interested in both types of dependence simultaneously, in [58] the authors propose to use $\left(OPD_{X^{(1)},X^{(2)}}\right)_+ - \left(OPD_{X^{(1)},-X^{(2)}}\right)_+$. To keep the notation simple, we focus on OPD as it is defined in (4.13).

Remark 4.23 *It is possible to extend the concept of ordinal pattern dependence for more than two time series, by investigating the probability*

$$p_{X^{(1)},\ldots,X^{(d)}} := \mathbb{P}\left(\Pi\left(X_0^{(1)},\ldots,X_h^{(1)}\right) = \Pi\left(X_0^{(2)},\ldots,X_h^{(2)}\right) = \ldots = \Pi\left(X_0^{(d)},\ldots,X_h^{(d)}\right)\right).$$

However, this approach requires a very strong condition, namely the same co-movement in all d time series. In this case, no information would be given on any pairwise dependencies for $d \geq 3$.

For the sake of simplicity, we restrict ourselves to the case $d = 2$ in the following. However, note that in several of the limit theorems in this chapter can be extended to the d-dimensional case in a straightforward way. Note that the calculations concerning the asymptotic variances would need to be adapted in this case.

We compare whether the ordinal patterns in $\left(X_j^{(1)}\right)_{j\in\mathbb{Z}}$ coincide with the ones in $\left(X_j^{(2)}\right)_{j\in\mathbb{Z}}$. Recall that it is an essential property of ordinal patterns that they are uniquely determined by the increment process. Therefore, we have to consider the increment processes $(Y_j)_{j\in\mathbb{Z}} = \left(\left(Y_j^{(1)}, Y_j^{(2)}\right)\right)_{j\in\mathbb{Z}}$ as defined in (3.1) for $d = 2$, where $Y_j^{(p)} = X_j^{(p)} - X_{j-1}^{(p)}$, $p = 1,2$. Hence, we can also express p and q (and consequently OPD) as a probability that only depends on the increments of the considered vectors of the time series. Recall the definition of $(Y_{j,h})_{j\in\mathbb{Z}}$ for $d = 2$, given by

$$Y_{j,h} = \left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)},Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)^t,$$

such that $Y_{j,h} \sim \mathcal{N}(0,\Sigma_{2,h})$ with $\Sigma_{2,h}$ as given in (3.6).

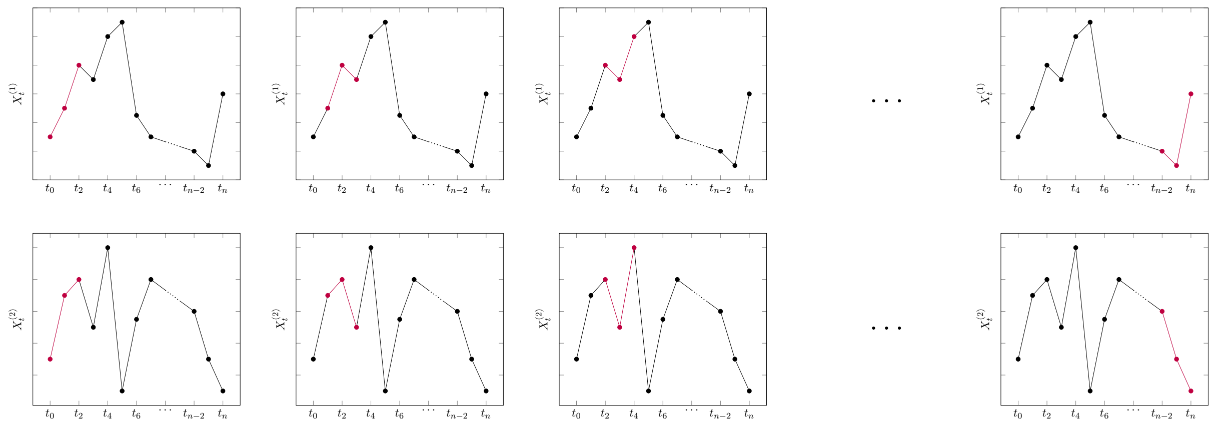

Figure 4.6: Illustration of estimation of ordinal pattern dependence.

In the course of this section, we use different approaches to estimate p and q and finally, by combining these two estimators, to estimate OPD. For all estimators the assumption on the corresponding function f in Theorem 3.3 and Theorem 3.10 holds, namely that it is square integrable with respect to $Y_{j,h}$ and that the set of discontinuity points is a Null set with respect to the $2h$-dimensional Lebesgue measure. This is shown in (4.5). Hence, we omit to emphasize this in the following.

We define the estimator of p, the probability of coincident patterns in both time series in a moving window of fixed length, by

$$\begin{aligned}\hat{p}_n &= \frac{1}{n-h}\sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)\right\}} \\ &= \frac{1}{n-h}\sum_{j=1}^{n-h} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)\right\}},\end{aligned}$$

where

$$\begin{aligned}\tilde{\Pi}\left(Y_1,\ldots,Y_h\right) &:= \Pi\left(0, Y_1, Y_1+Y_2,\ldots,Y_1+\ldots+Y_h\right) \\ &= \Pi\left(0, X_1-X_0,\ldots,X_h-X_0\right) \\ &= \Pi\left(X_0, X_1,\ldots,X_h\right).\end{aligned}$$

Figure 4.6 pictures the way ordinal pattern dependence is estimated by $\hat{p}_n$. The patterns of interest that are compared in each moving window are colored in red.

Having emphasized the crucial importance of the increments, we define the following conditions on the increment process $(Y_j)_{j\in\mathbb{Z}}$. For the reader's convenience, we provide an overview of these conditions.

Overview 4.24

(L1) *We assume* $(Y_j)_{j\in\mathbb{Z}}$ *is a bivariate, stationary Gaussian process with* $Y_j^{(p)} \sim \mathcal{N}(0,1)$*,* $p = 1,2$ *that fulfills* (3.2)*. Furthermore, it holds that* $\left|r^{(p,q)}(k)\right| < 1$ *for* $p,q = 1,2$ *and* $k \geq 1$ *to exclude ties.*

(L1s) *We soften* ***(L1)*** *by allowing* $\min\{d_1, d_2\}$ *to be in the range* $(-\infty, 0)$*, such that the corresponding univariate Gaussian process is short-range dependent.*

(L2) *We strengthen* ***(L1)*** *such that* (3.2) *is fulfilled with* d_1 *and* d_2 *in* $\left(\frac{1}{4}, \frac{1}{2}\right)$*.*

(L2s) *We soften* ***(L2)*** *by allowing* $\min\{d_1, d_2\}$ *to be in the range* $(-\infty, 0) \cup \left(0, \frac{1}{4}\right]$*.*

(S) *We assume* $d_1, d_2 \in (-\infty, 0) \cup \left(0, \frac{1}{4}\right)$ *such that for* $p,q = 1,2$

$$r^{(p,q)}(k) = k^{d_p+d_q-1} L_{p,q}(k) \quad (k \to \infty)$$

with $L_{p,q}(k) \to L_{p,q}$ *and* $L_{p,q} \in \mathbb{R}$ *holds.*

The last condition **(S)** implies absolute summability of $\left(r^{(p,q)}(k)\right)^2$ for $p,q = 1,2$ and represents the case of short-range dependent components for estimators with Hermite rank 2.

Note that the multivariate analogue to the case in which we compare the coincident patterns in the two marginal processes of the bivariate stationary process itself is presented in Section 4.4.3.

Before turning to the investigation of the estimator of coincident patterns, we first have to consider the symmetry properties of the multivariate normal distribution, as they differ from the univariate case given in Section 4.2.
As we have seen in Section 4.2, the symmetry properties of the multivariate normal distribution essentially help to determine the Hermite coefficients needed in Theorem 3.3 and Theorem 3.10. Using these symmetries, the Hermite coefficients can be computed at low computational cost. Since we are considering a bivariate time series here, the symmetry properties slightly change compared to the univariate case. To be more precise, the general case only allows for space symmetry and we need some mild assumptions on the entries of the correlation matrix to get further symmetries and less computational cost having the calculations of the asymptotic variances in mind. These assumptions are often fulfilled in practice.
First note that for a bivariate stationary Gaussian process $\left(Y_j^{(1)}, Y_j^{(2)}\right)_{j\in\mathbb{Z}}$ with $Y_j^{(p)} \sim \mathcal{N}(0,1)$, $p = 1,2$ in general we have

$$(I) \quad \left(Y_1^{(1)}, \ldots, Y_h^{(1)}, Y_1^{(2)}, \ldots, Y_h^{(2)}\right)^t \stackrel{\mathcal{D}}{=} \left(-Y_1^{(1)}, \ldots, -Y_h^{(1)}, -Y_1^{(2)}, \ldots, -Y_h^{(2)}\right)^t. \qquad (4.14)$$

This property was already introduced as *space symmetry*.

In contrast to the univariate case, we do no longer obtain *time symmetry* in the general case. Adding the assumption $r^{(1,1)}(k) = r^{(2,2)}(k)$, $k = 0, \ldots, h-1$ we get

$$(II) \quad \left(Y_1^{(1)}, \ldots, Y_h^{(1)}, Y_1^{(2)}, \ldots, Y_h^{(2)}\right)^t \stackrel{\mathcal{D}}{=} \left(Y_h^{(2)}, \ldots, Y_1^{(2)}, Y_h^{(1)}, \ldots, Y_1^{(1)}\right)^t. \qquad (4.15)$$

Note that in this case the correlation matrix of the vector is persymmetric.

Another condition, which is often fulfilled in practical contexts, is to assume that the cross-correlation function $r^{(1,2)}(k)$ is symmetric, e.g., $r^{(1,2)}(k) = r^{(2,1)}(k)$, for each $k = 0, \ldots, h-1$. Then, we obtain

$$(III) \quad \left(Y_1^{(1)}, \ldots, Y_h^{(1)}, Y_1^{(2)}, \ldots, Y_h^{(2)}\right)^t \stackrel{\mathcal{D}}{=} \left(Y_h^{(1)}, \ldots, Y_1^{(1)}, Y_h^{(2)}, \ldots, Y_1^{(2)}\right)^t. \tag{4.16}$$

In this case the correlation matrix has equal entries on the anti-diagonal for $h = 2$.

Finally, combining the two assumptions, we have

$$(IV) \quad \left(Y_1^{(1)}, \ldots, Y_h^{(1)}, Y_1^{(2)}, \ldots, Y_h^{(2)}\right)^t \stackrel{\mathcal{D}}{=} \left(Y_1^{(2)}, \ldots, Y_h^{(2)}, Y_1^{(1)}, \ldots, Y_h^{(1)}\right)^t. \tag{4.17}$$

Note that even case (IV), as the strongest one considered here, has not the same properties as the univariate case, in which the correlation matrix is a Toeplitz matrix, without any further assumptions but stationarity on the underlying process. This result could be obtained under further and stronger assumptions, but would not yield a simplification concerning our problem and is, therefore, omitted.
It is necessary to emphasize that property (I) always holds for multivariate Gaussian vectors, whereas Conditions (II), (III) or (IV) are **additional** assumptions that have to be verified before the advantages of these symmetries can be used in practical applications.

We begin with the investigation of the asymptotics of $\hat{p}_n$. First, we calculate the Hermite rank of $\hat{p}_n$, since the Hermite rank determines for which ranges of d^* the estimator $\hat{p}_n$ is still long-range dependent. Depending on this range, different limit theorems may hold.

Lemma 4.25 *The Hermite rank of* $f(Y_{j,h}) = \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{j+1}^{(1)}, \ldots, Y_{j+h}^{(1)}\right) = \tilde{\Pi}\left(Y_{j+1}^{(2)}, \ldots, Y_{j+h}^{(2)}\right)\right\}}$ *with respect to* $\Sigma_{2,h}$ *is equal to* 2.

Proof. Following Lemma 2.33 it is sufficient to show the following two properties:

(i) $m(f, \Sigma_{2,h}) \geq 2$,

(ii) $m(f, I_{2,h}) \leq 2$.

Lemma 2.33 can be applied due to the following reasoning. Ordinal patterns are not affected by scaling, therefore, the technical condition that $\Sigma_{2,h}^{-1} - I_{2,h}$ is positive semidefinite is fulfilled in our case. We can scale the standard deviation of the random vector $Y_{j,h}$ by any positive real number $\sigma > 0$ since for all $j \in \mathbb{Z}$ we have

$$\begin{aligned} &\left\{\tilde{\Pi}\left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}\right) = \tilde{\Pi}\left(Y_j^{(2)}, \ldots, Y_{j+h-1}^{(2)}\right)\right\} \\ = &\left\{\tilde{\Pi}\left(\sigma Y_j^{(1)}, \ldots, \sigma Y_{j+h-1}^{(1)}\right) = \tilde{\Pi}\left(\sigma Y_j^{(2)}, \ldots, \sigma Y_{j+h-1}^{(2)}\right)\right\}. \end{aligned}$$

To show property (i), we need to consider a multivariate random vector

$$Y_{1,h} := \left(Y_1^{(1)}, \ldots, Y_h^{(1)}, Y_1^{(2)}, \ldots, Y_h^{(2)}\right)^t$$

with covariance matrix $\Sigma_{2,h}$. We fix $i = 1, \dots, 2h$. We divide the set S_h into disjoint sets, namely into S_h^*, as defined in Definition 4.7 and the complimentary set $S_h \setminus S_h^*$. The symmetry property (I) of the multivariate normal distribution implies

$$\mathbb{E}\left(Y_{j,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right\}}\right) = -\mathbb{E}\left(Y_{j,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\mathcal{S}(\pi)\right\}}\right)$$

for $\pi \in S_h$. Hence, we arrive at:

$$\begin{aligned}
\mathbb{E}\left(Y_{j,h}^{(i)} f(Y_{j,h})\right) &= \mathbb{E}\left(Y_{j,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)\right\}}\right)\\
&= \sum_{\pi\in S_h}\mathbb{E}\left(Y_{j,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right\}}\right)\\
&= \sum_{\pi\in S_h^*}\mathbb{E}\left(Y_{j,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right\}}\right)\\
&\qquad - \sum_{\pi\in S_h\setminus S_h^*}\mathbb{E}\left(Y_{j,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\mathcal{S}(\pi)\right\}}\right)\\
&= 0
\end{aligned}$$

for $i = 1, \dots, 2h$.
Consequently $m(f, \Sigma_{2,h}) \geq 2$.

In order to proof (ii), we consider

$$U_{1,h} := \left(U_1^{(1)}, \dots, U_h^{(1)}, U_1^{(2)}, \dots, U_h^{(2)}\right)^t$$

to be a random vector with independent $\mathcal{N}(0,1)$ distributed entries. For $i = 1, \dots, h$ and $k = h+1, \dots, 2h$ such that $k - h = i$, we obtain

$$\begin{aligned}
\mathbb{E}\left(U_{1,h}^{(i)}U_{1,h}^{(k)} f\left(U_{1,h}\right)\right) &= \mathbb{E}\left(U_i^{(1)}U_{k-h}^{(2)}\mathbf{1}_{\left\{\tilde{\Pi}\left(U_1^{(1)},\dots,U_h^{(1)}\right)=\tilde{\Pi}\left(U_1^{(2)},\dots,U_h^{(2)}\right)\right\}}\right)\\
&= \sum_{\pi\in S_h}\mathbb{E}\left(U_i^{(1)}U_i^{(2)}\mathbf{1}_{\left\{\tilde{\Pi}\left(U_1^{(1)},\dots,U_h^{(1)}\right)=\tilde{\Pi}\left(U_1^{(2)},\dots,U_h^{(2)}\right)=\pi\right\}}\right)\\
&= \sum_{\pi\in S_h}\left(\mathbb{E}\left(U_i^{(1)}\mathbf{1}_{\left\{\tilde{\Pi}\left(U_1^{(1)},\dots,U_h^{(1)}\right)=\pi\right\}}\right)\right)^2\\
&\neq 0,
\end{aligned}$$

since $\mathbb{E}\left(U_i^{(1)}\mathbf{1}_{\left\{\tilde{\Pi}\left(U_1^{(1)},\dots,U_h^{(1)}\right)=\pi\right\}}\right) \neq 0$ for all $\pi \in S_h$. This was shown in the proof of Lemma 4.12.
All in all, we derive $m(f, \Sigma_{2,h}) = 2$ and, hence, have proven the lemma. □

The case $m(f, \Sigma_{2,h}) = 2$ exhibits the property that the standard range of the long-range dependence parameter $d^* \in \left(0, \frac{1}{2}\right)$ has to be divided into two different sets. If $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$, the transformed process $f\left(Y_{j,h}\right)_{j\in\mathbb{Z}}$ is still long-range dependent, see Table 2.1. If $d^* \in \left(0, \frac{1}{4}\right)$, the transformed process is short-range dependent, which means by definition that the autocorrelations of the transformed process are summable, see Definition 2.14. Therefore, we have two different asymptotic distributions that have to be considered for the estimator $\hat{p}_n$ of coincident patterns.

4.3.1 Limit theorem for the estimator of p in case of long-range dependence

First, we restrict ourselves to the case that at least one of the two parameters d_1 and d_2 is in $\left(\frac{1}{4}, \frac{1}{2}\right)$. This assures $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$. We explicitly include mixing cases where the process corresponding to $\min\{d_1, d_2\}$ is allowed to be long-range as well as short-range dependent.
Note that this setting includes the pure long-range dependence case, which means that for $p = 1, 2$, we have $d_p \in \left(\frac{1}{4}, \frac{1}{2}\right)$, or even $d_1 = d_2 = d^*$. However, in general the assumptions are lower, such that we only require $d_p \in \left(\frac{1}{4}, \frac{1}{2}\right)$ for either $p = 1$ or $p = 2$ and the other parameter is allowed to be in $(-\infty, 0)$ or $\left(0, \frac{1}{4}\right)$, too.

We can, therefore, apply the results of Corollary 3.11 and obtain the following asymptotic distribution for $\hat{p}_n$:

Theorem 4.26 *Under the assumptions of Corollary 3.11 in the case $d = 2$, corresponding to **(L2s)** in Overview 4.24, we obtain*

$$n^{1-2d^*} (C_2)^{-\frac{1}{2}} (\hat{p}_n - p) \xrightarrow{\mathcal{D}} \sum_{p,q\in P^*} \tilde{\alpha}^{(p,q)} Z^{(p,q)}_{2,d^*+1/2}(1) \tag{4.18}$$

with $Z^{(p,q)}_{2,d^*+1/2}(1)$ as given in Theorem 3.10 for $p, q \in P^*$ and $C_2 := \frac{1}{2d^*(4d^*-1)}$ being a normalizing constant. We have

$$\tilde{\alpha}^{(p,q)} := \sum_{i,k=1}^{h} \alpha^{(p,q)}_{i,k}, \text{ where } \alpha^{(p,q)}_{i,k} = \alpha_{i+(p-1)h,k+(q-1)h},$$

for each $p, q \in P^*$ and $i, k = 1, \ldots, h$ and $(\alpha_{i,k})_{1\leq i,k\leq dh} = \Sigma^{-1}_{2,h} C \Sigma^{-1}_{2,h}$, where the variable

$$C = (c_{i,k})_{1\leq i,k\leq 2h} = \mathbb{E}\left(Y_{1,h}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y^{(1)}_h\right)=\tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y^{(2)}_h\right)\right\}} - p\right) Y^t_{1,h}\right)$$

denotes the matrix of second order Hermite coefficients.

Proof. The proof of this theorem is an immediate application of Corollary 3.11 following Lemma 4.25. □

Following Theorem 3.13, we are also able to express the limit distribution above in terms of two standard Rosenblatt random variables by modifying the weighting factors in the limit distribution. Note that this requires slightly stronger assumptions as in Theorem 3.10.

Theorem 4.27 *Let **(L2)** hold with $d_1 = d_2$. Additionally we assume that $r^{(1,1)}(l) = r^{(2,2)}(l)$, for $l = 0, \ldots, h-1$, and $L_{1,1} + L_{2,2} \neq L_{1,2} + L_{2,1}$. Then we obtain*

$$n^{1-2d^*} (C_2)^{-\frac{1}{2}} (\hat{p}_n - p) \xrightarrow{\mathcal{D}} \left(\tilde{\alpha}^{(1,1)} - \tilde{\alpha}^{(1,2)}\right) \frac{L_{2,2} - L_{2,1} - L_{1,2} + L_{1,1}}{2} Z^*_{2,d^*+1/2}(1)$$
$$+ \left(\tilde{\alpha}^{(1,1)} + \tilde{\alpha}^{(1,2)}\right) \frac{L_{2,2} + L_{2,1} + L_{1,2} + L_{1,1}}{2} Z^{**}_{2,d^*+1/2}(1),$$

with C_2 and $\tilde{\alpha}^{(p,q)}$ as given in Theorem 4.26. Note that $Z^*_{2,d^*+1/2}(1)$ and $Z^{**}_{2,d^*+1/2}(1)$ are both standard Rosenblatt random variables whose covariance is given by

$$Cov\left(Z^*_{2,d^*+1/2}(1), Z^{**}_{2,d^*+1/2}(1)\right) = \frac{\left(L_{2,2} - L_{1,1}\right)^2}{\left(L_{1,1} + L_{2,2}\right)^2 - \left(L_{1,2} + L_{2,1}\right)^2}. \tag{4.19}$$

Remark 4.28 *Following Corollary 3.14, if additionally $r^{(1,1)}(k) = r^{(2,2)}(k)$ and $r^{(1,2)}(k) = r^{(2,1)}(k)$ is fulfilled for all $k \in \mathbb{Z}$, then, the two limit random variables following a standard Rosenblatt distribution in Theorem 4.27 are independent. Note that due to the considerations in* (2.17)*, we know that the distribution of the sum of two independent standard Rosenblatt random variables is not standard Rosenblatt. However, this yields a computational benefit, as it is possible to efficiently simulate the standard Rosenblatt distribution, for details, see [66].*

Calculation of the asymptotic variance

The covariance structure of $Z^{(p,q)}_{2,d^*+1/2}(1)$, $p, q \in P^*$, is given in Corollary 3.6. Hence, it is possible to determine the variance of the limit distribution, or the variance of each summand, in Theorems 4.26 and 4.27 if we are able to calculate C and accordingly α.

To this end, we show that it is significantly helpful to incorporate the symmetry properties of the multivariate normal distribution, as given in (I) to (IV) in (4.14) to (4.17). We have

$$\begin{aligned} c_{i,k} &= \mathbb{E}\left(Y^{(i)}_{1,2}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y^{(1)}_h\right) = \tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y^{(2)}_h\right)\right\}} - p\right) Y^{(k)}_{1,2}\right) \\ &= \mathbb{E}\left(\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2} - \mathbb{E}\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y^{(1)}_h\right) = \tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y^{(2)}_h\right)\right\}}\right) \\ &= \sum_{\pi \in S_h} \mathbb{E}\left(\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2} - \mathbb{E}\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y^{(1)}_h\right) = \tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y^{(2)}_h\right) = \pi\right\}}\right) \end{aligned}$$

and we define

$$c^{\pi,2}_{i,k} := \mathbb{E}\left(\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2} - \mathbb{E}\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y^{(1)}_h\right) = \tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y^{(2)}_h\right) = \pi\right\}}\right)$$

as the second order Hermite coefficients of the function $f\left(Y_{1,2}\right) = \mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y^{(1)}_h\right) = \tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y^{(2)}_h\right) = \pi\right\}}$.

Using the symmetry properties (I) to (IV), we obtain the following equalities for these second order Hermite coefficients, in terms of the time and spatial reflection of the considered ordinal pattern, where the symmetry property additionally assumed on the underlying process $\left(Y_{j,h}\right)_{j\in\mathbb{Z}}$ is denoted behind each equality:

Lemma 4.29 *For $\pi \in S_h$ and $i,k \in \{1,\dots,2h\}$, it holds*

$$\begin{aligned}
c^{\pi,2}_{i,k} &= c^{\mathcal{S}(\pi),2}_{i,k}, \\
c^{\pi,2}_{i,k} &= c^{\mathcal{T}(\pi),2}_{2h+1-i,2h+1-k}, && (II)\\
c^{\pi,2}_{i,k} &= c^{\mathcal{T}(\pi),2}_{\lfloor\frac{i-1}{h}\rfloor 2h+h+1-i,\lfloor\frac{k-1}{h}\rfloor 2h+h+1-k}, && (III)\\
c^{\pi,2}_{i,k} &= c^{\pi,2}_{i+(-1)^{\lfloor\frac{i-1}{h}\rfloor}h,k+(-1)^{\lfloor\frac{k-1}{h}\rfloor}h}. && (IV)
\end{aligned} \tag{4.20}$$

Proof. We fix $\pi \in S_h$ and $i,k = 1,\dots,2h$. The first equality is derived by using (I):

$$\begin{aligned}
c^{\pi,2}_{i,k} &= \mathbb{E}\left(\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2} - \mathbb{E}\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\dots,Y^{(1)}_h\right)=\tilde{\Pi}\left(Y^{(2)}_1,\dots,Y^{(2)}_h\right)=\pi\right\}}\right)\\
&= \mathbb{E}\left(\left(\left(-Y^{(i)}_{1,2}\right)\left(-Y^{(k)}_{1,2}\right) - \mathbb{E}\left(\left(-Y^{(i)}_{1,2}\right)\left(-Y^{(k)}_{1,2}\right)\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y^{(1)}_1,\dots,-Y^{(1)}_h\right)=\tilde{\Pi}\left(-Y^{(2)}_1,\dots,-Y^{(2)}_h\right)=\pi\right\}}\right)\\
&= \mathbb{E}\left(\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2} - \mathbb{E}\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\dots,Y^{(1)}_h\right)=\tilde{\Pi}\left(Y^{(2)}_1,\dots,Y^{(2)}_h\right)=\mathcal{S}(\pi)\right\}}\right)\\
&= c^{\mathcal{S}(\pi),2}_{i,k}.
\end{aligned}$$

The second equality is obtained with Condition (I) and (II):

$$\begin{aligned}
c^{\pi,2}_{i,k} &= \mathbb{E}\left(\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2} - \mathbb{E}\left(Y^{(i)}_{1,2}Y^{(k)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\dots,Y^{(1)}_h\right)=\tilde{\Pi}\left(Y^{(2)}_1,\dots,Y^{(2)}_h\right)=\pi\right\}}\right)\\
&= \mathbb{E}\Bigg(\left(\left(-Y^{(2h+i-1)}_{1,2}\right)\left(-Y^{(2h+k-1)}_{1,2}\right) - \mathbb{E}\left(\left(-Y^{(2h+i-1)}_{1,2}\right)\left(-Y^{(2h+k-1)}_{1,2}\right)\right)\right)\\
&\qquad \times \mathbf{1}_{\left\{\tilde{\Pi}\left(-Y^{(2)}_h,\dots,-Y^{(2)}_1\right)=\tilde{\Pi}\left(-Y^{(1)}_h,\dots,-Y^{(1)}_1\right)=\pi\right\}}\Bigg)\\
&\overset{(4.8)}{=} \mathbb{E}\left(\left(Y^{(2h+i-1)}_{1,2}Y^{(2h+k-1)}_{1,2} - \mathbb{E}\left(Y^{(2h+i-1)}_{1,2}Y^{(2h+k-1)}_{1,2}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(2)}_1,\dots,Y^{(2)}_h\right)=\tilde{\Pi}\left(Y^{(1)}_1,\dots,Y^{(1)}_h\right)=\mathcal{T}(\pi)\right\}}\right)\\
&= c^{\mathcal{T}(\pi),2}_{2h+1-i,2h+1-k}.
\end{aligned}$$

The third relation is given by the following calculation. It holds

$$c_{i,k}^{\pi,2} = \mathbb{E}\left(\left(Y_{1,2}^{(i)}Y_{1,2}^{(k)} - \mathbb{E}\left(Y_{1,2}^{(i)}Y_{1,2}^{(k)}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right\}}\right)$$

$$= \begin{cases} \mathbb{E}\Bigg(\left(Y_{1,2}^{(h+1-i)}Y_{1,2}^{(h+1-k)} - \mathbb{E}\left(Y_{1,2}^{(h+1-i)}Y_{1,2}^{(h+1-k)}\right)\right) \\ \qquad \times\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y_h^{(1)},\dots,-Y_1^{(1)}\right)=\tilde{\Pi}\left(-Y_h^{(2)},\dots,-Y_1^{(2)}\right)=\pi\right\}}\Bigg), & \text{if } i,k=1,\dots,h, \\ \mathbb{E}\Bigg(\left(Y_{1,2}^{(h+1-i)}Y_{1,2}^{(3h+1-k)} - \mathbb{E}\left(Y_{1,2}^{(h+1-i)}Y_{1,2}^{(3h+1-k)}\right)\right) \\ \qquad \times\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y_h^{(1)},\dots,-Y_1^{(1)}\right)=\tilde{\Pi}\left(-Y_h^{(2)},\dots,-Y_1^{(2)}\right)=\pi\right\}}\Bigg), & \text{if } i=1,\dots,h,\ k=h+1,\dots,2h, \\ \mathbb{E}\Bigg(\left(Y_{1,2}^{(3h+1-i)}Y_{1,2}^{(h+1-k)} - \mathbb{E}\left(Y_{1,2}^{(3h+1-i)}Y_{1,2}^{(h+1-k)}\right)\right) \\ \qquad \times\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y_h^{(1)},\dots,-Y_1^{(1)}\right)=\tilde{\Pi}\left(-Y_h^{(2)},\dots,-Y_1^{(2)}\right)=\pi\right\}}\Bigg), & \text{if } i=h+1,\dots,2h,\ k=1,\dots,h, \\ \mathbb{E}\Bigg(\left(Y_{1,2}^{(3h+1-i)}Y_{1,2}^{(3h+1-k)} - \mathbb{E}\left(Y_{1,2}^{(3h+1-i)}Y_{1,2}^{(3h+1-k)}\right)\right) \\ \qquad \times\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y_h^{(1)},\dots,-Y_1^{(1)}\right)=\tilde{\Pi}\left(-Y_h^{(2)},\dots,-Y_1^{(2)}\right)=\pi\right\}}\Bigg), & \text{if } i,k=h+1,\dots,2h \end{cases}$$

$$= \mathbb{E}\Bigg(\left(Y_{1,2}^{(\lfloor\frac{i-1}{h}\rfloor 2h+h+1-i)}Y_{1,2}^{(\lfloor\frac{k-1}{h}\rfloor 2h+h+1-k)} - \mathbb{E}\left(Y_{1,2}^{(\lfloor\frac{i-1}{h}\rfloor 2h+h+1-i)}Y_{1,2}^{(\lfloor\frac{k-1}{h}\rfloor 2h+h+1-k)}\right)\right)$$

$$\times\,\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\mathcal{T}(\pi)\right\}}\Bigg)$$

$$= c^{\mathcal{T}(\pi),2}_{\lfloor\frac{i-1}{h}\rfloor 2h+h+1-i,\lfloor\frac{k-1}{h}\rfloor 2h+h+1-k},$$

using (I) and (III).

Finally, the last equation is proven in the following way: recall that in (IV) only the order of the two h-dimensional vectors is changed, but no changes happen to the indices of the random variables within each vector.

It holds

$$
\begin{aligned}
c_{i,k}^{\pi,2} &= \mathbb{E}\left(\left(Y_{1,2}^{(i)}Y_{1,2}^{(k)} - \mathbb{E}\left(Y_{1,2}^{(i)}Y_{1,2}^{(k)}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right\}}\right) \\
&= \begin{cases}
\mathbb{E}\Bigg(\left(Y_{1,2}^{(i+h)}Y_{1,2}^{(k+h)} - \mathbb{E}\left(Y_{1,2}^{(i+h)}Y_{1,2}^{(k+h)}\right)\right) \\
\qquad \times \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\pi\right\}}\Bigg), & \text{if } i,k = 1,\dots,h, \\
\mathbb{E}\Bigg(\left(Y_{1,2}^{(i+h)}Y_{1,2}^{(k-h)} - \mathbb{E}\left(Y_{1,2}^{(i+h)}Y_{1,2}^{(k-h)}\right)\right) \\
\qquad \times \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\pi\right\}}\Bigg), & \text{if } i = 1,\dots,h \text{ and } k = h+1,\dots,2h, \\
\mathbb{E}\Bigg(\left(Y_{1,2}^{(i-h)}Y_{1,2}^{(k+h)} - \mathbb{E}\left(Y_{1,2}^{(i-h)}Y_{1,2}^{(k+h)}\right)\right) \\
\qquad \times \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\pi\right\}}\Bigg), & \text{if } i = h+1,\dots,2h \text{ and } k = 1,\dots,h, \\
\mathbb{E}\Bigg(\left(Y_{1,2}^{(i-h)}Y_{1,2}^{(k-h)} - \mathbb{E}\left(Y_{1,2}^{(i-h)}Y_{1,2}^{(k-h)}\right)\right) \\
\qquad \times \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\pi\right\}}\Bigg), & \text{if } i,k = h+1,\dots,2h,
\end{cases} \\
&= \mathbb{E}\Bigg(\left(Y_{1,2}^{(i+(-1)^{\lfloor \frac{i-1}{h}\rfloor}h)}Y_{1,2}^{(k+(-1)^{\lfloor \frac{k-1}{h}\rfloor}h)} - \mathbb{E}\left(Y_{1,2}^{(i+(-1)^{\lfloor \frac{i-1}{h}\rfloor}h)}Y_{1,2}^{(k+(-1)^{\lfloor \frac{k-1}{h}\rfloor}h)}\right)\right) \\
&\qquad \times \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right\}}\Bigg) \\
&= c^{\pi,2}_{i+(-1)^{\lfloor \frac{i-1}{h}\rfloor}h,\,k+(-1)^{\lfloor \frac{k-1}{h}\rfloor}h}.
\end{aligned}
$$

□

We apply these results to $c_{i,k}$ in the following. In order to do so, we recall Definition 4.5: $S_{[2],h}$ is the set of permutations, where each permutation is generating a different reversion group with two elements. Analogously, $S_{[4],h}$ is the set of permutations where each element generates reversion groups having the cardinality four. Then,

$$S_h = \bigcup_{\pi \in S_{[2],h}} \{\pi, \mathcal{S}(\pi)\} \cup \bigcup_{\pi \in S_{[4],h}} \{\pi, \mathcal{S}(\pi), \mathcal{T}(\pi), \mathcal{T}\circ\mathcal{S}(\pi)\}.$$

Recall that $S_h \neq S_{[2],h} \cup S_{[4],h}$. In the general case we obtain

$$
\begin{aligned}
c_{i,k} &= \sum_{\pi \in S_h} c_{i,k}^{\pi,2} \\
&= \sum_{\pi \in S_{[2],h}} c_{i,k}^{\pi,2} + c_{i,k}^{\mathcal{S}(\pi),2} + \sum_{\pi \in S_{[4],h}} c_{i,k}^{\pi,2} + c_{i,k}^{\mathcal{S}(\pi),2} + c_{i,k}^{\mathcal{T}(\pi),2} + c_{i,k}^{\mathcal{T}\circ\mathcal{S}(\pi),2} \\
&= 2\sum_{\pi \in S_{[2],h}} c_{i,k}^{\pi,2} + 2\sum_{\pi \in S_{[4],h}} c_{i,k}^{\pi,2} + c_{i,k}^{\mathcal{T}(\pi),2}.
\end{aligned}
$$

Using Condition (II) in (4.15), it holds

$$c_{i,k} = 2 \sum_{\pi \in S_{[2],h}} c_{i,k}^{\pi,2} + 2 \sum_{\pi \in S_{[4],h}} c_{i,k}^{\pi,2} + c_{2h+1-i,2h+1-k}^{\pi,2}.$$

With Condition (III) in (4.16), we get

$$c_{i,k} = 2 \sum_{\pi \in S_{[2],h}} c_{i,k}^{\pi,2} + 2 \sum_{\pi \in S_{[4],h}} c_{i,k}^{\pi,2} + c^{\pi,2}_{\lfloor \frac{i-1}{h} \rfloor 2h+h+1-i, \lfloor \frac{k-1}{h} \rfloor 2h+h+1-k}$$

for $i, k = 1, \ldots, 2h$.

Combining conditions (II) and (III) does not yield a further simplification in terms of the reversion groups and is, therefore, omitted. However, since we have seen that in the cases (II) and/or (III) determining one matrix of second-order Hermite coefficients for one fixed pattern of each reversion group is sufficient to obtain the limit variance, we turn to the calculation of this explicit matrix. We figure out under which of the above used assumptions we do not need to determine all $(2h)^2$-entries of the matrix $C^{\pi,2}$, but can reduce this number. This yields a reduction of the computational effort, since each entry of the matrix is a $(2h)$-dimensional integral. Evaluating such an integral has a certain computational cost that increases with increasing h.

We turn to the calculcation of the second-order Hermite coefficient matrix $C^{\pi,2}$:

Example 4.30 *For the reader's convenience, we consider the case $h = 2$. In the most general setting, we have the following representation of the matrix $C^{\pi,2}$:*

$$C^{\pi,2} = \begin{pmatrix} c_{1,1}^{\pi,2} & c_{1,2}^{\pi,2} & c_{1,3}^{\pi,2} & c_{1,4}^{\pi,2} \\ c_{1,2}^{\pi,2} & c_{2,2}^{\pi,2} & c_{2,3}^{\pi,2} & c_{2,4}^{\pi,2} \\ c_{1,3}^{\pi,2} & c_{2,3}^{\pi,2} & c_{3,3}^{\pi,2} & c_{3,4}^{\pi,2} \\ c_{1,4}^{\pi,2} & c_{2,4}^{\pi,2} & c_{3,4}^{\pi,2} & c_{4,4}^{\pi,2} \end{pmatrix},$$

where the symmetry of the matrix follows from

$$\begin{aligned} c_{i,k}^{\pi,2} &= \mathbb{E}\left(\left(Y_{1,2}^{(i)} Y_{1,2}^{(k)} - \mathbb{E}\left(Y_{1,2}^{(i)} Y_{1,2}^{(k)}\right)\right) \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)}, \ldots, Y_h^{(1)}\right) = \tilde{\Pi}\left(Y_1^{(2)}, \ldots, Y_h^{(2)}\right) = \pi\right\}}\right) \\ &= \mathbb{E}\left(\left(Y_{1,2}^{(k)} Y_{1,2}^{(i)} - \mathbb{E}\left(Y_{1,2}^{(i)} Y_{1,2}^{(k)}\right)\right) \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)}, \ldots, Y_h^{(1)}\right) = \tilde{\Pi}\left(Y_1^{(2)}, \ldots, Y_h^{(2)}\right) = \pi\right\}}\right) \\ &= c_{k,i}^{\pi,2}. \end{aligned}$$

The number of different entries of this matrix can be reduced in the case $\#\bar{\pi} = 4$ if Conditions (II) and (III) are fulfilled. Following (4.20) *we obtain*

$$C^{\pi,2} = \begin{pmatrix} c_{1,1}^{\pi,2} & c_{1,2}^{\pi,2} & c_{1,3}^{\pi,2} & c_{1,4}^{\pi,2} \\ c_{1,2}^{\pi,2} & c_{2,2}^{\pi,2} & c_{1,4}^{\pi,2} & c_{2,4}^{\pi,2} \\ c_{1,3}^{\pi,2} & c_{1,4}^{\pi,2} & c_{1,1}^{\pi,2} & c_{1,2}^{\pi,2} \\ c_{1,4}^{\pi,2} & c_{2,4}^{\pi,2} & c_{1,2}^{\pi,2} & c_{2,2}^{\pi,2} \end{pmatrix}.$$

Hence, the number of different entries was reduced from 10 *to* 6.
A different simplification can be obtained in the case that $\#\bar{\pi} = 2$*, where* $\mathcal{T} \circ \mathcal{S}(\pi) = \pi$*.*
By Condition (II) we get

$$C^{\pi,2} = \begin{pmatrix} c^{\pi,2}_{1,1} & c^{\pi,2}_{1,2} & c^{\pi,2}_{1,3} & c^{\pi,2}_{1,4} \\ c^{\pi,2}_{1,2} & c^{\pi,2}_{2,2} & c^{\pi,2}_{2,3} & c^{\pi,2}_{1,3} \\ c^{\pi,2}_{1,3} & c^{\pi,2}_{2,3} & c^{\pi,2}_{2,2} & c^{\pi,2}_{1,2} \\ c^{\pi,2}_{1,4} & c^{\pi,2}_{1,3} & c^{\pi,2}_{1,2} & c^{\pi,2}_{1,1} \end{pmatrix},$$

while, if Condition (III) holds, we obtain

$$C^{\pi,2} = \begin{pmatrix} c^{\pi,2}_{1,1} & c^{\pi,2}_{1,2} & c^{\pi,2}_{1,3} & c^{\pi,2}_{1,4} \\ c^{\pi,2}_{1,2} & c^{\pi,2}_{1,1} & c^{\pi,2}_{1,4} & c^{\pi,2}_{1,3} \\ c^{\pi,2}_{1,3} & c^{\pi,2}_{1,4} & c^{\pi,2}_{3,3} & c^{\pi,2}_{3,4} \\ c^{\pi,2}_{1,4} & c^{\pi,2}_{1,3} & c^{\pi,2}_{3,4} & c^{\pi,2}_{3,3} \end{pmatrix}.$$

Hence, in both cases for $\#\bar{\pi} = 2$*, we reduce the number of different matrix entries from* 10 *to* 6*.*
By combining conditions (II) and (III), we get the most convenient (and most restrictive) case, where

$$C^{\pi,2} = \begin{pmatrix} c^{\pi,2}_{1,1} & c^{\pi,2}_{1,2} & c^{\pi,2}_{1,3} & c^{\pi,2}_{1,4} \\ c^{\pi,2}_{1,2} & c^{\pi,2}_{1,1} & c^{\pi,2}_{1,4} & c^{\pi,2}_{1,3} \\ c^{\pi,2}_{1,3} & c^{\pi,2}_{1,4} & c^{\pi,2}_{1,1} & c^{\pi,2}_{1,2} \\ c^{\pi,2}_{1,4} & c^{\pi,2}_{1,3} & c^{\pi,2}_{1,2} & c^{\pi,2}_{1,1} \end{pmatrix}.$$

This improves the reduction from 10 *to* 4 *different matrix entries.*

It is possible to generalize the reductions to $h \in \mathbb{N}$. First, note that the number of entries in the matrix $C^{\pi,2}$ is given by $(2h)^2$. Using the general symmetry of the matrix, we need to determine $\frac{2h(2h+1)}{2} = h(2h+1)$ entries. We imagine that we divide $C^{\pi,h}$ into four $h \times h$ block matrices, namely

$$C^{\pi,h} = \begin{pmatrix} C^{\pi,h,1} & C^{\pi,h,2} \\ C^{\pi,h,3} & C^{\pi,h,4} \end{pmatrix},$$

such that $C^{\pi,h,p} = (c_{i,j})_{i=\lfloor \frac{p}{3} \rfloor h+1,\ldots,\lfloor \frac{p}{3} \rfloor h+h, j=((p+1) \mod 2)h+1,\ldots,((p+1) \mod 2)h+h}$, $p = 1,2,3,4$.
If (II) and (III) are fulfilled and we have $\#\bar{\pi} = 4$, we see that each of these block matrices is symmetric and that we only need to determine the entries of two of them. Therefore, we end up with $2\frac{h(h+1)}{2} = h(h+1)$ entries to calculate.

If $\#\bar{\pi} = 2$ the same idea can be applied. If Condition (II) holds, the matrix C is additionally persymmetric as it is described in (4.20). Therefore, it is again sufficient to determine two of the block matrices above, namely $C^{\pi,h,1}$ and $C^{\pi,h,2}$. The block matrix $C^{\pi,h,1}$ with entries $c_{i,j}$, $i,j = 1,\ldots,h$, is symmetric. However, $C^{\pi,h,2}$ is still persymmetric. Therefore, we need to identify $2\frac{h(h+1)}{2} = h(h+1)$ different entries.

Assuming Condition (III) holds, we need to distinguish between the block matrices. For $C^{\pi,h,1}$

with entries $c_{i,j}$, $i,j = 1,\dots,h$, we have to calculate $1+\ldots+h$ entries if h is odd and $2+\ldots+h$ if h is even by (4.20), since this matrix is both symmetric and persymmetric. This yields $\frac{h}{2}\left(\frac{h}{2}+1\right)$ calculations if h is even and $\left(\frac{h+1}{2}\right)^2$ if h is odd. Note that this is the number of calculations to obtain $C^{\pi,h,1}$ as well as $C^{\pi,h,4}$ with entries $c_{i,j}$, $i,j = h+1,\dots,2h$. Due to symmetry, the last block matrix we need to study is $C^{\pi,h,2}$ with entries $c_{i,j}$, $i = 1,\dots,h$, $j = h+1,\dots,2h$. This case is different from $C^{\pi,h,1}$, since we have no symmetry property of $c_{i,j}$ within the block matrix. However, the property derived in (4.20) still holds and we arrive at $\frac{h}{2}h$ entries to be determined if h is even and at $\frac{h-1}{2}(h+1)+1$ if h is odd. Taking all these formulas into consideration, in case Condition (III) holds, we arrive at $h\left(\frac{h}{2}+1\right)+\frac{h}{2}h = h(h+1)$ entries to identify if h is even and at $2\left(\frac{h+1}{2}\right)^2+\frac{h-1}{2}(h+1)+1 = h^2+h+1$ if h is odd.

In the last and most restrictive case conditions (II) and (III) hold. Here, we only need to distinguish between the first two block matrices $C^{\pi,h,1}$ and $C^{\pi,h,2}$, since $C^{\pi,h,4}$, given by $(c_{i,j})_{i,j=h+1\dots,2h}$, coincides with $C^{\pi,h,1}$. Therefore, we arrive at $\frac{3h^2+2h+3}{4}$ matrix entries to be calculated if h is odd and $\frac{3h^2+2h}{4}$ entries if h is even.

Table 4.1 gives an overview of the number of $(2h)$-dimensional integrals that we need to calculate in each of the considered cases in order to determine the matrix C of second order Hermite coefficients for the estimator of coincident patterns.

Recall that a_h denotes the number of reversion groups $\bar{\pi}$ in S_h with two elements and, correspondingly, b_h is the number of reversion groups with four elements, as introduced in Section 4.1. If h is odd, we have

$$\begin{aligned} a_h &= \left(\frac{h+1}{2}\right)!2^{\frac{h-1}{2}}, \\ b_h &= \frac{(h+1)! - 2\left(\frac{h+1}{2}\right)!2^{\frac{h-1}{2}}}{4}, \end{aligned}$$

and $a_h = a_{h-1}$, $b_h = \frac{(h+1)!-2a_{h-1}}{4} = \frac{(h+1)!-2\left(\frac{h}{2}\right)!2^{\frac{h-2}{2}}}{4}$ if h is even. Without requiring any of the symmetry properties, the number of integrals needed to calculate the matrix C is given by $(h+1)!(2h)^2$, since $\#S_h = (h+1)!$ and each matrix $C^{\pi,2}$ has $(2h)^2$ entries. The following example shows the reduction possibilities by taking the symmetry properties into account.

Example 4.31 *If $h = 2$ and conditions (II) and (III) hold, we need to calculate 10 four-dimensional integrals. In the case without any symmetries considered, we would have needed 96. If $h = 3$ and conditions (II) and (III) hold, we have to evaluate 84 six-dimensional integrals, while the most general case requires 864.*

	(I)
Formula for $i, k = 1, \ldots, 2h$	$c_{i,k} = 2\sum_{\pi\in S_{[2],h}} c_{i,k}^{\pi,2} + 2\sum_{\pi\in S_{[4],h}} c_{i,k}^{\pi,2} + c_{i,k}^{\mathcal{T}(\pi),2}$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[2],h}$	$h(2h+1)$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[4],h}$	$h(2h+1)$
Number of integrals needed to determine C	$(a_h + 2b_h)\, h(2h+1)$
	(II)
Formula for $i, k = 1, \ldots, 2h$	$c_{i,k} = 2\sum_{\pi\in S_{[2],h}} c_{i,k}^{\pi,2} + 2\sum_{\pi\in S_{[4],h}} c_{i,k}^{\pi,2} + c_{2h+1-i,2h+1-k}^{\pi,2}$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[2],h}$	$h(h+1)$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[4],h}$	$h(2h+1)$
Number of integrals needed to determine C	$a_h h(h+1) + b_h h(2h+1)$
	(III)
Formula for $i, k = 1, \ldots, 2h$	$c_{i,k} = 2\sum_{\pi\in S_{[2],h}} c_{i,k}^{\pi,2} + 2\sum_{\pi\in S_{[4],h}} c_{i,k}^{\pi,2} + c_{\lfloor \frac{i-1}{h} \rfloor 2h+h+1-i, \lfloor \frac{k-1}{h} \rfloor 2h+h+1-k}^{\pi,2}$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[2],h}$	$h^2 + h + (h \mod 2)$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[4],h}$	$h(2h+1)$
Number of integrals needed to determine C	$a_h \left(h^2 + h + (h \mod 2)\right) + b_h h(2h+1)$
	(IV)
Formula for $i, k = 1, \ldots, 2h$	same as in (III)
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[2],h}$	$\frac{3h^2+2h+3(h \mod 2)}{4}$
Calculations needed for $C^{\pi,2}$ if $\pi \in S_{[4],h}$	$h(h+1)$
Number of integrals needed to determine C	$a_h \frac{3h^2+2h+3(h \mod 2)}{4} + b_h h(h+1)$

Table 4.1: Number of $(2h)$-dimensional integrals needed to determine the matrix of second order Hermite coefficients of the estimator of ordinal pattern probabilities.

Comparing the number of integrals for general $h \in \mathbb{N}$, we see that if we exploit the symmetry of C and property (I) (which is no additional assumption) we already obtain

$$\frac{(a_h + 2b_h)\, h(2h+1)}{(h+1)!(2h)^2} = \frac{1}{4} + \frac{1}{8h}.$$

Therefore, the computational effort is approximately reduced by $\frac{1}{4}$. In the optimized case, where conditions (II) and (III) hold, we even reduce the number of integrals by a considerable factor as shown in Figure 4.7. In this figure, the reduction factor is provided for increasing $h \in \mathbb{N}$.

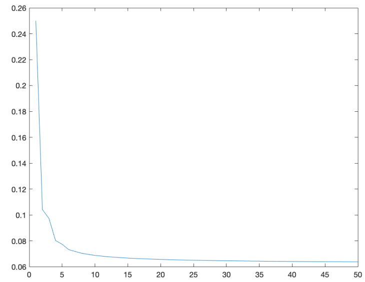

Figure 4.7: Factor of reduced computational time for the second order Hermite coefficients of $\hat{p}_n$.

We are now able to compute C, and hence turn to the computation of α. Note that the symmetries described in (4.20) combined with the formulas given in Table 4.1 imply that these symmetries can be transfered to $c_{i,k}$, $i,k = 1,\dots,2h$, such that we obtain the following results for $i,k = 1,\dots,2h$:

$$\begin{aligned}
c_{i,k} &= c_{2h+1-i,2h+1-k}, && (II),\\
c_{i,k} &= c_{\lfloor\frac{i-1}{h}\rfloor 2h+h+1-i,\lfloor\frac{k-1}{h}\rfloor 2h+h+1-k}, && (III),\\
c_{i,k} &= c_{i+(-1)^{\lfloor\frac{i-1}{h}\rfloor}h,k+(-1)^{\lfloor\frac{k-1}{h}\rfloor}h}, && (II),(III).
\end{aligned} \tag{4.21}$$

Recall the last three matrices in Example 4.30, which display the entries of $C^{\pi,2}$ in the case $\pi = \mathcal{T} \circ \mathcal{S}(\pi)$. Following the symmetries (4.21), we are in the same situation. The last three matrices in Example 4.30 without π as index display the entries of the matrix C in case $h = 2$ under the given assumptions (II), (III) or (II) combined with (III), respectively. The symmetries in (4.21) have no impact on the number of integrals needed to determine the matrix C, since they were already included in the results above. However, while calculating the matrix α, they appear to be helpful. Therefore, we have to take a closer look at the matrix α, and want to deduce the values of $\tilde{\alpha}^{(p,q)}$ for $p,q = 1,2$. Note that the following holds:

$$\begin{aligned}
\alpha &= \Sigma_{2,h}^{-1} C \Sigma_{2,h}^{-1}\\
&= \Sigma_{2,h}^{-1}\left(2\sum_{\pi\in S_{[2],h}} C^{\pi,2} + 2\sum_{\pi\in S_{[4],h}}\left(C^{\pi,2} + C^{\mathcal{T}(\pi),2}\right)\right)\Sigma_{2,h}^{-1}\\
&= 2\sum_{\pi\in S_{[2],h}} \Sigma_{2,h}^{-1} C^{\pi,2}\Sigma_{2,h}^{-1} + 2\sum_{\pi\in S_{[4],h}} \Sigma_{2,h}^{-1} C^{\pi,2}\Sigma_{2,h}^{-1} + \Sigma_{2,h}^{-1} C^{\mathcal{T}(\pi),2}\Sigma_{2,h}^{-1}\\
&= 2\sum_{\pi\in S_{[2],h}} \alpha^{\pi,2} + 2\sum_{\pi\in S_{[4],h}} \alpha^{\pi,2} + \alpha^{\mathcal{T}(\pi),2},
\end{aligned} \tag{4.22}$$

where

$$\alpha^{\pi,2} := \Sigma_{2,h}^{-1} C^{\pi,2} \Sigma_{2,h}^{-1}$$

is the matrix of second order Hermite coefficients for one fixed π multiplied with the inverse of the covariance matrix $\Sigma_{2,h}^{-1}$ from both sides.
We concentrate on $\alpha^{\pi,2}$ and derive relations similar to those in Lemma 4.29.

Lemma 4.32 *For the values of*

$$\tilde{\alpha}_{\pi,2}^{(p,q)} = \sum_{i,k=1}^{h} \alpha_{(p-1)h+i,(q-1)h+k}^{\pi,2}, \quad p,q = 1,2,$$

we can deduce the following equalities:

$$\begin{aligned}
\tilde{\alpha}_{\pi,2}^{(p,q)} &= \tilde{\alpha}_{\mathcal{T}(\pi),2}^{(3-p,3-q)}, && (II)\\
\tilde{\alpha}_{\pi,2}^{(p,q)} &= \tilde{\alpha}_{\mathcal{T}(\pi),2}^{(p,q)}, && (III)\\
\tilde{\alpha}_{\pi,2}^{(1,1)} &= \tilde{\alpha}_{\pi,2}^{(2,2)}. && (II)+(III)
\end{aligned}$$

Proof. Recall that $\Sigma_{2,h}^{-1} = (g_{i,j})_{i,j=1,\dots,2h}$. In general, the following holds for $p,q = 1,2$:

$$\begin{aligned}
\tilde{\alpha}_{\pi,2}^{(p,q)} &= \sum_{i,k=1}^{h} \alpha_{(p-1)h+i,(q-1)h+k)}^{\pi,2}\\
&= \sum_{i,k=1}^{h} \sum_{j,l=1}^{2h} g_{(p-1)h+i,l}\, g_{j,(q-1)h+k}\, c_{j,l}^{\pi,2}\\
&= \sum_{j,l=1}^{2h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l} \right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k} \right) c_{j,l}^{\pi,2}. \qquad (4.23)
\end{aligned}$$

The first equality follows, since Condition (II) yields $c_{i,k}^{\pi,2} = c_{2h+1-i,2h+1-k}^{\mathcal{T}(\pi),2}$ and $\Sigma_{2,h}^{-1}$ is persymmetric such that $g_{i,j} = g_{2h+1-i,2h+1-j}$:

$$\begin{aligned}
\tilde{\alpha}_{\mathcal{T}(\pi),2}^{(p,q)} &= \sum_{j,l=1}^{2h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l} \right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k} \right) c_{j,l}^{\mathcal{T}(\pi),2}\\
&= \sum_{j,l=1}^{2h} \left(\sum_{i=1}^{h} g_{2h+1-i-(p-1)h,2h+1-l} \right) \left(\sum_{k=1}^{h} g_{2h+1-j,2h+1-k-(q-1)h} \right) c_{2h+1-j,2h+1-l}^{\pi,2}\\
&= \sum_{j,l=1}^{2h} \left(\sum_{i=1}^{h} g_{(3-p)h+1-i,l} \right) \left(\sum_{k=1}^{h} g_{j,(3-q)h+1-k} \right) c_{j,l}^{\pi,2}\\
&= \sum_{j,l=1}^{2h} \left(\sum_{i=1}^{h} g_{(2-p)h+i,l} \right) \left(\sum_{k=1}^{h} g_{j,(2-q)h+k} \right) c_{j,l}^{\pi,2}\\
&= \tilde{\alpha}_{\pi,2}^{(3-p,3-q)}.
\end{aligned}$$

By Condition (III) it holds that $g_{i+(p-1)h,j} = g_{h+1-i+(p-1)h,h+1-j}$ if $j = 1,\dots,h$ and $g_{i+(p-1)h,j} = g_{h+1-i+(p-1)h,3h+1-j}$ if $j = h+1,\dots,2h$. We get

$$\tilde{\alpha}_{\mathcal{T}(\pi),2}^{(p,q)} = \sum_{j,l=1}^{2h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l} \right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k} \right) c_{j,l}^{\mathcal{T}(\pi),2}$$

$$
\begin{aligned}
&= \sum_{j,l=1}^{h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l}\right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k}\right) c_{j,l}^{\mathcal{T}(\pi),2} \\
&\quad + \sum_{j=1}^{h} \sum_{l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l}\right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k}\right) c_{j,l}^{\mathcal{T}(\pi),2} \\
&\quad + \sum_{j=h+1}^{2h} \sum_{l=1}^{h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l}\right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k}\right) c_{j,l}^{\mathcal{T}(\pi),2} \\
&\quad + \sum_{j,l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{(p-1)h+i,l}\right) \left(\sum_{k=1}^{h} g_{j,(q-1)h+k}\right) c_{j,l}^{\mathcal{T}(\pi),2} \\
&= \sum_{j,l=1}^{h} \left(\sum_{i=1}^{h} g_{h+1-i+(p-1)h,h+1-l}\right) \left(\sum_{k=1}^{h} g_{h+1-j,h+1-k+(q-1)h}\right) c_{h+1-j,h+1-l}^{\pi,2} \\
&\quad + \sum_{j=1}^{h} \sum_{l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{h+1-i+(p-1)h,3h+1-l}\right) \left(\sum_{k=1}^{h} g_{h+1-j,h+1-k+(q-1)h}\right) c_{h+1-j,3h+1-l}^{\pi,2} \\
&\quad + \sum_{j=h+1}^{2h} \sum_{l=1}^{h} \left(\sum_{i=1}^{h} g_{h+1-i+(p-1)h,h+1-l}\right) \left(\sum_{k=1}^{h} g_{3h+1-j,h+1-k+(q-1)h}\right) c_{3h+1-j,h+1-l}^{\pi,2} \\
&\quad + \sum_{j,l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{h+1-i+(p-1)h,3h+1-l}\right) \left(\sum_{k=1}^{h} g_{3h+1-j,h+1-k+(q-1)h}\right) c_{3h+1-j,3h+1-l}^{\pi,2} \\
&= \sum_{j,l=1}^{h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,h+1-l}\right) \left(\sum_{k=1}^{h} g_{h+1-j,k+(q-1)h}\right) c_{h+1-j,h+1-l}^{\pi,2} \\
&\quad + \sum_{j=1}^{h} \sum_{l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,3h+1-l}\right) \left(\sum_{k=1}^{h} g_{h+1-j,k+(q-1)h}\right) c_{h+1-j,3h+1-l}^{\pi,2} \\
&\quad + \sum_{j=h+1}^{2h} \sum_{l=1}^{h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,h+1-l}\right) \left(\sum_{k=1}^{h} g_{3h+1-j,k+(q-1)h}\right) c_{3h+1-j,h+1-l}^{\pi,2} \\
&\quad + \sum_{j,l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,3h+1-l}\right) \left(\sum_{k=1}^{h} g_{3h+1-j,k+(q-1)h}\right) c_{3h+1-j,3h+1-l}^{\pi,2} \\
&= \sum_{j,l=1}^{h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,l}\right) \left(\sum_{k=1}^{h} g_{j,k+(q-1)h}\right) c_{j,l}^{\pi,2} \\
&\quad + \sum_{j=1}^{h} \sum_{l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,l}\right) \left(\sum_{k=1}^{h} g_{j,k+(q-1)h}\right) c_{j,l}^{\pi,2} \\
&\quad + \sum_{j=h+1}^{2h} \sum_{l=1}^{h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,l}\right) \left(\sum_{k=1}^{h} g_{j,k+(q-1)h}\right) c_{j,l}^{\pi,2} \\
&\quad + \sum_{j,l=h+1}^{2h} \left(\sum_{i=1}^{h} g_{i+(p-1)h,l}\right) \left(\sum_{k=1}^{h} g_{j,k+(q-1)h}\right) c_{j,l}^{\pi,2} \\
&= \tilde{\alpha}_{\pi,2}^{(p,q)}.
\end{aligned}
$$

Combining both results, we obtain $\tilde{\alpha}_{\pi,2}^{(1,1)} = \tilde{\alpha}_{\pi,2}^{(2,2)}$ if (II) and (III) are fulfilled. □

Finally, we apply the previous results to compute $\tilde{\alpha}^{(p,q)}$ for $p, q = 1, 2$.
In general, we have

$$\tilde{\alpha}^{(p,q)} = 2 \sum_{\pi \in S_{[2],h}} \tilde{\alpha}^{(p,q)}_{\pi,2} + 2 \sum_{\pi \in S_{[4],h}} \tilde{\alpha}^{(p,q)}_{\pi,2} + \tilde{\alpha}^{(p,q)}_{\mathcal{T}(\pi),2}. \tag{4.24}$$

If additionally (II) holds, we obtain

$$\tilde{\alpha}^{(p,q)} = 2 \sum_{\pi \in S_{[2],h}} \tilde{\alpha}^{(p,q)}_{\pi,2} + 2 \sum_{\pi \in S_{[4],h}} \tilde{\alpha}^{(p,q)}_{\pi,2} + \tilde{\alpha}^{(3-p,3-q)}_{\pi,2}. \tag{4.25}$$

Therefore,

$$\tilde{\alpha}^{(1,1)} = \tilde{\alpha}^{(2,2)}.$$

In case (III) holds, we have

$$\tilde{\alpha}^{(p,q)} = 2 \sum_{\pi \in S_{[2],h}} \tilde{\alpha}^{(p,q)}_{\pi,2} + 4 \sum_{\pi \in S_{[4],h}} \tilde{\alpha}^{(p,q)}_{\pi,2}. \tag{4.26}$$

Finally, in case (IV) is fulfilled the last two properties are combined.
Note that symmetry properties analogous to those for the last three matrices in Example 4.30 hold for α. In case $h = 2$ and Condition (III) holds, it is even sufficient to calculate two entries of each block matrix, respectively, to determine the limit variance. This reduction property is already displayed in Table 4.1.

It is possible to obtain closed formulas for the asymptotic variance for $h = 1$ as the following example shows:

Example 4.33 *We focus on the case $h = 1$ and consider the underlying process $(Y_{j,1})_{j\in\mathbb{Z}} = \left(Y_j^{(1)}, Y_j^{(2)}\right)_{j\in\mathbb{Z}}$. It is possible to determine the asymptotic variance depending on the correlation $r^{(1,2)}(0)$ between these two increment variables.*
We start with the calculation of the second order Hermite coefficients in the case $\pi = (1, 0)$. This corresponds to the event $\left\{Y_j^{(1)} \geq 0, Y_j^{(2)} \geq 0\right\}$, which yields

$$c^{\pi,2}_{1,1} = \mathbb{E}\left(\left(\left(Y_j^{(1)}\right)^2 - 1\right) \mathbf{1}_{\left\{Y_j^{(1)} \geq 0, Y_j^{(2)} \geq 0\right\}}\right)$$

and

$$c^{\pi,2}_{1,2} = \mathbb{E}\left(\left(Y_j^{(1)} Y_j^{(2)}\right) \mathbf{1}_{\left\{Y_j^{(1)} \geq 0, Y_j^{(2)} \geq 0\right\}}\right).$$

Due to $r^{(1,2)}(0) = r^{(2,1)}(0)$, we have $\left(Y_j^{(1)}, Y_j^{(2)}\right) \stackrel{\mathcal{D}}{=} \left(Y_j^{(2)}, Y_j^{(1)}\right)$ and, therefore, $c^{\pi,2}_{1,1} = c^{\pi,2}_{2,2}$. We identify the second order Hermite coeficients as the ones already calculated in Example 4.21, although we are considering two consecutive increments of a univariate Gaussian process there. However, since the corresponding values are only determined by the correlation between the

Gaussian variables, we can simply replace the autocorrelation at lag 1 *by the cross-correlation at lag* 0*. Hence, we obtain*

$$c_{1,1}^{\pi,2} = \varphi^2(0) r^{(1,2)}(0) \sqrt{1 - \left(r^{(1,2)}(0)\right)^2},$$
$$c_{1,2}^{\pi,2} = \varphi^2(0) \sqrt{1 - \left(r^{(1,2)}(0)\right)^2}.$$

Recall that the inverse $\Sigma_{2,1}^{-1} = (g_{i,j})_{i,j=1,2}$ *of the correlation matrix of* $\left(Y_j^{(1)}, Y_j^{(2)}\right)$ *is given by*

$$\Sigma_{2,1}^{-1} = \frac{1}{1 - \left(r^{(1,2)}(0)\right)^2} \begin{pmatrix} 1 & -r^{(1,2)}(0) \\ -r^{(1,2)}(0) & 1 \end{pmatrix}.$$

By using the formula for $\tilde{\alpha}^{(p,q)}$ *obtained in* (4.23)*, we derive*

$$\tilde{\alpha}_{\pi,2}^{(1,1)} = \alpha_{1,1}^{\pi,2} = \left(g_{1,1}^2 + g_{1,2}^2\right) c_{1,1}^{\pi,2} + 2 g_{1,1} g_{1,2} c_{1,2}^{\pi,2},$$
$$\tilde{\alpha}_{\pi,2}^{(1,2)} = \alpha_{1,2}^{\pi,2} = \left(g_{1,1}^2 + g_{1,2}^2\right) c_{1,2}^{\pi,2} + 2 g_{1,1} g_{1,2} c_{1,1}^{\pi,2}.$$

Plugging the second order Hermite cofficients and the entries of the inverse of the covariance matrix depending on $r^{(1,2)}(0)$ *into the formulas, we arrive at*

$$\tilde{\alpha}_{\pi,2}^{(1,1)} = \frac{-\varphi^2(0) r^{(1,2)}(0)}{\left(1 - \left(r^{(1,2)}(0)\right)^2\right)^{1/2}}$$

and

$$\tilde{\alpha}_{\pi,2}^{(1,2)} = \frac{\varphi^2(0)}{\left(1 - \left(r^{(1,2)}(0)\right)^2\right)^{1/2}}.$$

Therefore, in the case $h = 1$*, we obtain the following factors in the limit variance in Theorem 4.26:*

$$\tilde{\alpha}^{(1,1)} = \tilde{\alpha}^{(2,2)} = \frac{-2\varphi^2(0) r^{(1,2)}(0)}{\left(1 - \left(r^{(1,2)}(0)\right)^2\right)^{1/2}}$$
$$\tilde{\alpha}^{(1,2)} = \tilde{\alpha}^{(2,1)} = \frac{2\varphi^2(0)}{\left(1 - \left(r^{(1,2)}(0)\right)^2\right)^{1/2}}.$$

Remark 4.34 *In the case* $h = 2$*, it is not possible to derive closed formulas since the second order Hermite coefficients are integrals whose solutions need orthant probabilities of the four-dimensional Gaussian distribution. Unfortunately, no closed formulas are available for this dimension, see [1]. Nonetheless, we developed Matlab algorithms cHR2 and HR2LimVar, that yield the limit variance in Theorem 4.26, following the efficiency criterions we derived in this section to keep the computational cost low. The Matlab algorithms are described in Section A.3. Exemplarily, assuming that (II) and (III) hold, we have a computational cost of 38 seconds, while in the case where no symmetry properties at all are taken into consideration, we need 8 minutes and 24 seconds.*

4.3.2 Limit theorem for the estimator of p in case of short-range dependence

In this section, we focus on the case that $d^* \in (-\infty, 0) \cup \left(0, \frac{1}{4}\right)$. If $d^* \in \left(0, \frac{1}{4}\right)$, we are still dealing with a long-range dependent multivariate Gaussian process $(Y_{j,h})_{j\in\mathbb{Z}}$. However, the transformed process $\hat{p}_n - p$ is no longer long-range dependent, since we are considering a function with Hermite rank 2, see also Table 2.1. Otherwise, if $d^* \in (-\infty, 0)$, the process $(Y_{j,h})_{j\in\mathbb{Z}}$ itself is already short-range dependent, since the cross-correlations are summable. Therefore, we obtain the following central limit theorem by applying Remark 3.15.

Theorem 4.35 *Under the assumptions in **(S)** in Overview 4.24, we obtain*

$$n^{\frac{1}{2}} (\hat{p}_n - p) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2\right)$$

with

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}\Bigg[\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)\right\}} - p\right) \times \left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{1+k}^{(1)},\dots,Y_{h+k}^{(1)}\right)=\tilde{\Pi}\left(Y_{1+k}^{(2)},\dots,Y_{h+k}^{(2)}\right)\right\}} - p\right)\Bigg].$$

We already proved limit theorems for the estimator $\hat{p}_n$ of coincident patterns for multivariate Gaussian processes in the long-range dependent, the short-range dependent, as well as in the mixed setting. What remains to do is the estimation of the other parameters q and subsequently OPD as given in the definition of ordinal pattern dependence in (4.13), to get asymptotic results for estimators of OPD in the end.

4.3.3 Limit theorems for estimators of q

In this part of the thesis, we concentrate on q as defined in (4.12) that represents the hypothetical case of independence between the time series $\left(Y_j^{(1)}\right)_{j\in\mathbb{Z}}$ and $\left(Y_j^{(2)}\right)_{j\in\mathbb{Z}}$ in the context of ordinal pattern dependence. However, of in our setting, these time series are not independent. We take a closer look at the definition of q:

$$q = \sum_{\pi\in S_h} \mathbb{P}\left(\Pi\left(X_0^{(1)}, \dots, X_h^{(1)}\right) = \pi\right) \mathbb{P}\left(\Pi\left(X_0^{(2)}, \dots, X_h^{(2)}\right) = \pi\right) \tag{4.27}$$

$$= \sum_{\pi\in S_h} \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)}, \dots, Y_h^{(1)}\right) = \pi\right) \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(2)}, \dots, Y_h^{(2)}\right) = \pi\right) \tag{4.28}$$

$$= \sum_{\pi\in S_h} q_{Y^{(1)}}(\pi) q_{Y^{(2)}}(\pi), \tag{4.29}$$

where $q_{Y^{(p)}}(\pi) = \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(p)}, \dots, Y_h^{(p)}\right) = \pi\right)$, $p = 1, 2$.
We already derived asymptotic results for the estimators $\hat{q}_{n,Y^{(p)}}(\pi)$ of $q_{Y^{(p)}}(\pi)$, for $p = 1, 2$, in Section 4.2.1. These estimators only rely on the univariate case. In order to determine the asymptotic distribution of an estimator of q, we take advantage of the approach used in [58], Section 2.4 and use the plug-in estimator of q, given by

$$\hat{q}_n = \sum_{\pi\in S_h} \hat{q}_{Y^{(1)},n}(\pi) \hat{q}_{Y^{(2)},n}(\pi).$$

In order to derive asymptotic results for this estimator, it is not sufficient to know the marginal asymptotic distributions of $\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi)$ and $\hat{q}_{Y^{(2)},n}(\pi) - q_{Y^{(2)}}(\pi)$ as given in Theorem 4.13. Instead we need to determine the limit behaviour of the joint vector

$$\left(\left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}\right)^t.$$

Before turning to the investigation of this joint convergence, we justify, why we may apply the techniques in the proofs of Theorem 3.3 and Theorem 3.10 for vectors of several functions $f_1, \dots, f_k$, $k \in \mathbb{N}$ all with equal Hermite rank. We define

$$g\left(Y_{j,h}\right) := \left(\sum_{j=1}^{n} f_1\left(Y_{j,h}\right), \dots, \sum_{j=1}^{n} f_k\left(Y_{j,h}\right)\right)^t$$

and by applying the continuous mapping theorem to (3.14), we derive

$$\left(\sum_{j=1}^{n} f_1\left(Y_{j,h}\right), \dots, \sum_{j=1}^{n} f_k\left(Y_{j,h}\right)\right)^t \overset{\mathcal{D}}{=} \left(\sum_{j=1}^{n} f_1\left(AU_{j,h}\right), \dots, \sum_{j=1}^{n} f_k\left(AU_{j,h}\right)\right)^t. \tag{4.30}$$

Hence, the equalities in distribution in the proofs hold componentwisely. We turn to the joint convergence result.

Theorem 4.36 *We assume **(L1)** in Overview 4.24 holds. Let $A_n = \left(a_{n,i,i}\right)_{i=1,\dots,2(h+1)!}$ be a diagonal matrix with entries*

$$a_{n,i,i} = n^{-d_1+\frac{1}{2}}, \text{ for } i = 1, \dots, (h+1)!,$$

and

$$a_{n,i,i} = n^{-d_2+\frac{1}{2}}, \text{ for } i = (h+1)!+1, \dots, 2(h+1)!.$$

Then,

$$A_n \left(\left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(\hat{q}_{Y^{(2)},n}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right)^t$$
$$\xrightarrow{\mathcal{D}} \left(\left(\tilde{\alpha}_\pi^{(1)} Y^{(1)}\right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2)} Y^{(2)}\right)_{\pi \in S_h}\right)^t \tag{4.31}$$

with

$$\left(\left(\tilde{\alpha}_\pi^{(1)} Y^{(1)}\right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2)} Y^{(2)}\right)_{\pi \in S_h}\right)^t \sim \mathcal{N}\left(0, \Sigma\right).$$

Furthermore,

$$\Sigma = A \left(\frac{L_{p,q} + L_{q,p}}{(d_p+d_q)(1+d_p+d_q)}\right)_{p,q=1,2} A^t,$$

where $A = \begin{pmatrix} \tilde{\alpha}_{\pi_1}^{(1)} & 0 \\ \vdots & \vdots \\ \tilde{\alpha}_{\pi_{(h+1)!}}^{(1)} & 0 \\ 0 & \tilde{\alpha}_{\pi_1}^{(2)} \\ \vdots & \vdots \\ 0 & \tilde{\alpha}_{\pi_{(h+1)!}}^{(2)} \end{pmatrix}$. *The values of $\tilde{\alpha}_\pi^{(p)}$, $\pi \in S_h$ and $p = 1, 2$, are given in Theorem 4.13.*

Proof. Note that for each entry of the vector on the left-hand side of (4.31), we use (3.20) in the case $d = 1$ and derive

$$
\begin{aligned}
&\left(n \left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi) \right)_{\pi \in S_h}, n \left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi) \right)_{\pi \in S_h} \right)^t \\
\overset{\mathcal{D}}{=} &\left(\left(\tilde{\alpha}_\pi^{(1)} \sum_{j=1}^{n} Y_j^{(1)} + o_{\mathbb{P}} \left(n^{d_1+1/2} \right) \right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2)} \sum_{j=1}^{n} Y_j^{(2)} + o_{\mathbb{P}} \left(n^{d_2+1/2} \right) \right)_{\pi \in S_h} \right)^t \\
= &A \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)} \right)^t + R_n,
\end{aligned}
$$

where

$$
A = \begin{pmatrix}
\tilde{\alpha}_{\pi_1}^{(1)} & 0 \\
\vdots & \vdots \\
\tilde{\alpha}_{\pi_{(h+1)!}}^{(1)} & 0 \\
0 & \tilde{\alpha}_{\pi_1}^{(2)} \\
\vdots & \vdots \\
0 & \tilde{\alpha}_{\pi_{(h+1)!}}^{(2)}
\end{pmatrix}
$$

and R_n is the vector of remainders. Since for each entry $R_{n,i}$ it holds that $R_i = o_{\mathbb{P}} \left(n^{d_1+1/2} \right)$ if $i = 1, \ldots, (h+1)!$ and $R_i = o_{\mathbb{P}} \left(n^{d_2+1/2} \right)$ if $i = 1 + (h+1)!, \ldots, 2(h+1)!$, we have

$$
A_n n^{-1} R_n \xrightarrow{\mathbb{P}} (0, \ldots, 0)^t .
$$

Following [31], Theorem 2, we have

$$
\begin{pmatrix} n^{-d_1-\frac{1}{2}} & 0 \\ 0 & n^{-d_2-\frac{1}{2}} \end{pmatrix} \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)} \right)^t \xrightarrow{\mathcal{D}} \left(Y^{(1)}, Y^{(2)} \right)^t ,
$$

where $\left(Y^{(1)}, Y^{(2)} \right)^t \sim \mathcal{N} \left(0, \tilde{\Sigma} \right)$ and

$$
\tilde{\Sigma} = \left(\frac{L_{p,q} + L_{q,p}}{(d_p + d_q)(1 + d_p + d_q)} \right)_{p,q=1,2}
$$

Note that $\tilde{\Sigma}$ is positive definite, because L is positive definite. We can apply the multivariate version of Slutzsky's Theorem, see Lemma 6.3 in [13]. Hence, we obtain

$$
A \begin{pmatrix} n^{-d_1-\frac{1}{2}} & 0 \\ 0 & n^{-d_2-\frac{1}{2}} \end{pmatrix} \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)} \right)^t + R_n \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \Sigma \right),
$$

with $\Sigma = A \tilde{\Sigma} A^t$.

Combining all results, we arrive at

$$\begin{aligned}
&A_n\left(\left(\hat{q}_{Y^{(1)},n}(\pi)-q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h},\left(\hat{q}_{Y^{(2)},n}(\pi)-q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)^t\\
&=A_n n^{-1}A\left(\sum_{j=1}^n Y_j^{(1)},\sum_{j=1}^n Y_j^{(2)}\right)^t+R_n\\
&=A\begin{pmatrix} n^{-d_1-\frac{1}{2}} & 0\\ 0 & n^{-d_2-\frac{1}{2}}\end{pmatrix}\left(\sum_{j=1}^n Y_j^{(1)},\sum_{j=1}^n Y_j^{(2)}\right)^t+R_n\\
&\xrightarrow{\mathcal{D}}\ \mathcal{N}(0,\Sigma),
\end{aligned}$$

since $A_n n^{-1}A=A\begin{pmatrix} n^{-d_1-\frac{1}{2}} & 0\\ 0 & n^{-d_2-\frac{1}{2}}\end{pmatrix}$. □

To get a limit distribution for $\hat{q}_n$ by using the multivariate delta method, we may not allow for a normalization depending on $p=1,2$ as it is given in the matrix A_n in Theorem 4.36, since we need to apply a scalar-valued function to the vector on the left-hand side in (4.31) before normalization. Therefore, we need the same normalization value for each vector entry and recall P^* and d^* as given in (3.10).

Corollary 4.37 *The following statements hold:*

*(i) Let **(L1)** in Overview 4.24 hold with $d_1=d_2$. Then,*

$$n^{-d_1+\frac{1}{2}}\left(\left(\hat{q}_{Y^{(1)},n}(\pi)-q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h},\left(\hat{q}_{Y^{(2)},n}(\pi)-q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)^t \xrightarrow{\mathcal{D}}\left(\left(\tilde{\alpha}_\pi^{(1)}Y^{(1)}\right)_{\pi\in S_h},\left(\tilde{\alpha}_\pi^{(2)}Y^{(2)}\right)_{\pi\in S_h}\right)^t,$$

with $\left(\left(\tilde{\alpha}_\pi^{(1)}Y^{(1)}\right)_{\pi\in S_h},\left(\tilde{\alpha}_\pi^{(2)}Y^{(2)}\right)_{\pi\in S_h}\right)^t\sim\mathcal{N}(0,\Sigma)$ and Σ as given in Theorem 4.36.

*(ii) Under the assumptions in **(L1s)** with $d^*=d_1$, it holds that*

$$n^{-d^*+\frac{1}{2}}\left(\left(\hat{q}_{Y^{(1)},n}(\pi)-q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h},\left(\hat{q}_{Y^{(2)},n}(\pi)-q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)^t \xrightarrow{\mathcal{D}}\left(\left(\tilde{\alpha}_\pi^{(1)}Y\right)_{\pi\in S_h},0,\dots,0\right),$$

where $Y\sim\mathcal{N}(0,\sigma^2)$ with $\sigma^2=\frac{L_{1,1}}{d_1(1+2d_1)}$ and $\tilde{\alpha}_\pi^{(1)}$, $\pi\in S_h$, as given in Theorem 4.13.

Proof. Statement (i) is an immediate consequence of Theorem 4.36, since the matrix A_n is a diagonal matrix with identical entries.
In Setting (ii), we can conclude from

$$\begin{pmatrix} n^{-d_1-\frac{1}{2}} & 0\\ 0 & n^{-d_2-\frac{1}{2}}\end{pmatrix}\left(\sum_{j=1}^n Y_j^{(1)},\sum_{j=1}^n Y_j^{(2)}\right)^t\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\tilde{\Sigma}\right),$$

with $\tilde{\Sigma} = \left(\frac{L_{p,q}+L_{q,p}}{(d_p+d_q)(1+d_p+d_q)}\right)_{p,q=1,2}$, that for $d_2 < d_1$ but $d_2 \in \left(0, \frac{1}{2}\right)$, we have

$$n^{-d_1-\frac{1}{2}} \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)}\right)^t \xrightarrow{\mathcal{D}} \left(\mathcal{N}\left(0, \sigma^2\right), 0\right)^t,$$

with $\sigma^2 = \frac{L_{1,1}}{d_1(1+2d_1)}$.
Note that in the case $d_2 < 0$, the term $n^{-d_2-1/2}$ is not the correct normalization factor, but taking the classical convergence rate $n^{-\frac{1}{2}}$ as normalization yields a non-trivial result. However, $\left(d_1 + \frac{1}{2}\right) > \frac{1}{2}$ and, therefore, the above result holds if $\left(Y_j^{(2)}\right)_{j\in\mathbb{Z}}$ is short-range dependent. Note that the joint convergence of the sample means in mixed cases of long- and short-range dependence (with individual normalization for each of the components) is proved in [23], Theorem 3.3.
Finally, the result follows by adapting the proof of Theorem 4.36.
If $d^* = d_2$ the order of the entries of the vector in the limit distribution has to be changed, $\tilde{\alpha}_\pi^{(1)}$ has to be replaced by $\tilde{\alpha}_\pi^{(2)}$ for each $\pi \in S_h$, and the limit variance is given by $\sigma^2 = \frac{L_{2,2}}{d_2(1+2d_2)}$. □

The next step is to use the idea of [58], Theorem 4, and obtain the limit distribution for the plug-in estimator $\hat{q}_n$ of q by applying the multivariate delta method, see [42], Appendix A.

Theorem 4.38 *Let **(L1s)** in Overview 4.24 be fulfilled. Then, it holds that*

$$n^{-d^*+\frac{1}{2}} \left(\hat{q}_n - q\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

with

$$\sigma^2 = \begin{cases} \frac{L_{p,p}}{d^*(1+2d^*)} \sum_{\pi,\tilde{\pi}\in S_h} q_{Y^{(3-p)}}(\pi)\tilde{\alpha}_\pi^{(p)}\tilde{\alpha}_{\tilde{\pi}}^{(p)} q_{Y^{(3-p)}}(\tilde{\pi}), & \text{if } d^* = d_p, \text{ for exactly one } p = 1,2, \\ \\ \frac{L_{1,1}}{d^*(1+2d^*)} \sum_{\pi,\tilde{\pi}\in S_h} q_{Y^{(2)}}(\pi)\tilde{\alpha}_\pi^{(1)}\tilde{\alpha}_{\tilde{\pi}}^{(1)} q_{Y^{(2)}}(\tilde{\pi}) & \\ +2\frac{L_{1,2}}{d^*(1+2d^*)} \sum_{\pi,\tilde{\pi}\in S_h} q_{Y^{(2)}}(\pi)\tilde{\alpha}_\pi^{(1)}\tilde{\alpha}_{\tilde{\pi}}^{(2)} q_{Y^{(1)}}(\tilde{\pi}), & \text{if } d^* = d_1 = d_2 \\ +\frac{L_{2,2}}{d^*(1+2d^*)} \sum_{\pi,\tilde{\pi}\in S_h} q_{Y^{(1)}}(\pi)\tilde{\alpha}_\pi^{(2)}\tilde{\alpha}_{\tilde{\pi}}^{(2)} q_{Y^{(1)}}(\tilde{\pi}). & \end{cases}$$

Proof. The proof works exactly in the same way as the proof of Theorem 4 in [58] by applying the multivariate delta method. The function

$$f: \mathbb{R}^{2(h+1)!} \to \mathbb{R}, \quad f(x,y) = \sum_{i=1}^{(h+1)!} x_i y_i$$

is everywhere differentiable with $\nabla f(x,y) = (y,x)^t$. Furthermore it holds that

$$\hat{q}_n = f\left(\left(\hat{q}_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(\hat{q}_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)$$

as well as

$$q = f\left(\left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right).$$

Consequently, the formulas of the limit variance are given by

$$\left(\nabla f\left(\left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)\right)^t \Sigma \nabla f\left(\left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right),$$

where Σ is the variance of the limit vector given in Corollary 4.37. For details see [58], p. 719. $\square$

Finally, we proved limit theorems for all components of the estimator of ordinal pattern dependence. It remains to study the joint convergence behaviour of these components in order to achieve a limit theorem for the estimator of ordinal pattern dependence. However, for reasons that becomes clear in the next section, we continue with two alternative approaches to estimate q before turning to the joint convergence behaviour. First we propose an alternative estimator that is only slightly different from $\hat{q}_n$.
Define

$$\hat{q}_{n,\text{alt}} := \sum_{\pi \in S_h} \left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi)\right) \left(\hat{q}_{Y^{(2)},n}(\pi) - q_{Y^{(2)}}(\pi)\right) + q.$$

We can propose another limit theorem for the asymptotics of this estimator, directly following from Theorem 4.36.

Theorem 4.39 *Under the assumptions of **(L1)** in Overview 4.24, it holds*

$$n^{1-d_1-d_2} \left(\hat{q}_{n,\text{alt}} - q\right) \xrightarrow{\mathcal{D}} \left(\sum_{\pi \in S_h} \tilde{\alpha}_\pi^{(1)} \tilde{\alpha}_\pi^{(2)}\right) Y^{(1)} Y^{(2)},$$

where $\left(Y^{(1)}, Y^{(2)}\right)^t \sim \mathcal{N}\left(0, \tilde{\Sigma}\right)$ *and*

$$\tilde{\Sigma} = \left(\frac{L_{p,q} + L_{q,p}}{(d_p + d_q)(1 + d_p + d_q)}\right)_{p,q=1,2}$$

with $\tilde{\alpha}_\pi^{(p)}$*,* $p = 1, 2$ *as given in Theorem 4.13.*

Proof. Consider the function $f : \mathbb{R}^{2(h+1)!} \to \mathbb{R}$, $f(x, y) = \sum_{i=1}^{(h+1)!} x_i y_i$ and note that

$$\hat{q}_{n,\text{alt}} - q = f\left(\left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(\hat{q}_{Y^{(2)},n}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right).$$

Therefore, we can apply the continuous mapping theorem on the result of Theorem 4.36. Then, we obtain

$$\begin{aligned} f\left(\left(\tilde{\alpha}_\pi^{(1)} Y^{(1)}\right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2)} Y^{(2)}\right)_{\pi \in S_h}\right) &= \sum_{\pi \in S_h} \tilde{\alpha}_\pi^{(1)} Y^{(1)} \tilde{\alpha}_\pi^{(2)} Y^{(2)} \\ &= \left(\sum_{\pi \in S_h} \tilde{\alpha}_\pi^{(1)} \tilde{\alpha}_\pi^{(2)}\right) Y^{(1)} Y^{(2)}. \end{aligned}$$

$\square$

Corollary 4.40 *Under the conditions in **(L1s)** in Overview 4.24, the following statements are true:*

(i) if $d_1 = d_2$ and, consequently, $d^* = d_1$, it holds

$$n^{1-2d^*}\left(\hat{q}_{n,\text{alt}} - q\right) \xrightarrow{\mathcal{D}} \left(\sum_{\pi \in S_h} \tilde{\alpha}_\pi^{(1)} \tilde{\alpha}_\pi^{(2)}\right) Y^{(1)} Y^{(2)},$$

where $\tilde{\alpha}_\pi^{(p)}$, $p = 1, 2$ and $Y^{(1)}, Y^{(2)}$ as given in Theorem 4.39.

(ii) If $d_1 \neq d_2$ and, hence, $d_1 = d^*$ or $d_2 = d^*$, it holds

$$n^{1-2d^*}\left(\hat{q}_{n,\text{alt}} - q\right) \xrightarrow{\mathcal{D}} 0.$$

Proof. The first result is an immediate consequence of Theorem 4.39. The second result is obtained by applying the continuous mapping theorem to the second case of Corollary 4.37 as done in the proof of Theorem 4.39. □

A third approach to estimate q turns out to be very helpful in order to determine a reasonable asymptotic distribution of an estimator of ordinal pattern dependence: The estimation of q in terms of the improved estimator $\hat{q}_{n,\text{imp}}$ as investigated in Section 4.2.1.
We define

$$\hat{q}_{n,\text{imp}} := \sum_{\pi \in S_h} \hat{q}_{n,Y^{(1)},\text{imp}}(\pi) \hat{q}_{n,Y^{(2)},\text{imp}}(\pi).$$

From the mathematical point of view, the interesting difference of $\hat{q}_{n,\text{imp}}$ to the heuristic estimator of ordinal pattern probabilities for each marginal time series $\hat{q}_{n,Y^{(p)}}$, $p = 1, 2$, is that this improved estimator has Hermite rank 2. Therefore, it needs a different normalization and possesses another asymptotic structure. We also need to distinguish the case $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$, where the process transformed by the estimator is still long-range dependent from the case $d^* \in \left(0, \frac{1}{4}\right)$, where the transformed process is short-range dependent.

Lemma 4.41 *We assume **(L2)** in Overview 4.24. Let $A_n = (a_{n,i,i})_{i=1,\ldots,2(h+1)!}$ be a diagonal matrix with entries*

$$a_{n,i,i} = n^{1-2d_1}\left(2!C_2\right)^{-\frac{1}{2}} L_{1,1}^{-1}(n), \text{ for } i = 1, \ldots, (h+1)!$$

and

$$a_{n,i,i} = n^{1-2d_2}\left(2!C_2\right)^{-\frac{1}{2}} L_{2,2}^{-1}(n), \text{ for } i = (h+1)! + 1, \ldots, 2(h+1)!.$$

Then, it holds

$$A_n \left(\left(\hat{q}_{n,Y^{(1)},\text{imp}}(\pi) - q_{Y^{(1)}}(\pi) \right)_{\pi \in S_h}, \left(\hat{q}_{n,Y^{(2)},\text{imp}}(\pi) - q_{Y^{(2)}}(\pi) \right)_{\pi \in S_h} \right)^t$$
$$\xrightarrow{\mathcal{D}} \left(\left(\tilde{\alpha}_\pi^{(1,1)} Z^{(1,1)}_{2,d_1+\frac{1}{2}}(1) \right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2,2)} Z^{(2,2)}_{2,d_2+\frac{1}{2}}(1) \right)_{\pi \in S_h} \right)^t,$$

where $Z^{(1,1)}_{2,d_1+\frac{1}{2}}(1)$ and $Z^{(2,2)}_{2,d_2+\frac{1}{2}}(1)$ are standard Rosenblatt random variables. The scaling factor C_2, as well as $\tilde{\alpha}_\pi^{(p,p)}$, $p = 1, 2$, are given in Theorem 4.18.

Proof. We may use the proof of Theorem 3.10 for $d = 1$ as justified in (4.30) and we derive

$$A_n \left(\left(\hat{q}_{n,Y^{(1)},\text{imp}}(\pi) - q_{Y^{(1)}}(\pi) \right)_{\pi \in S_h}, \left(\hat{q}_{n,Y^{(2)},\text{imp}}(\pi) - q_{Y^{(2)}}(\pi) \right)_{\pi \in S_h} \right)^t$$
$$\stackrel{\mathcal{D}}{=} A_n \Bigg(\Bigg(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha^{(1,1)}_{i,i+l,\pi} \left(\hat{r}_n^{(1,1)}(l) - r^{(1,1)}(l) \right)$$
$$+ \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha^{(1,1)}_{i+l,i,\pi} \left(\hat{r}_n^{(1,1)}(l) - r^{(1,1)}(l) \right) + o_{\mathbb{P}}(n^{2d^*}) \Bigg)_{\pi \in S_h},$$
$$\Bigg(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha^{(1,1)}_{i,i+l,\pi} \left(\hat{r}_n^{(2,2)}(l) - r^{(2,2)}(l) \right)$$
$$+ \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha^{(1,1)}_{i+l,i,\pi} \left(\hat{r}_n^{(2,2)}(l) - r^{(2,2)}(l) \right) + o_{\mathbb{P}}(n^{2d^*}) \Bigg)_{\pi \in S_h} \Bigg)^t$$
$$\xrightarrow{\mathcal{D}} \left(\left(\tilde{\alpha}_\pi^{(1,1)} Z^{(1,1)}_{2,d_1+\frac{1}{2}}(1) \right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2,2)} Z^{(2,2)}_{2,d_2+\frac{1}{2}}(1) \right)_{\pi \in S_h} \right)^t,$$

by applying the continuous mapping theorem to Theorem 4.1 of [23].
Similar to the proof of Theorem 4.36, the dependence on π is only apparent in the coefficients $\tilde{\alpha}_\pi^{(p,p)}$, $p = 1, 2$, while the asymptotic distribution is determined by the estimator of the autocorrelations, which is independent of π. Therefore, it yields the same limit random variable (within each time series). □

We have derived the asymptotic distributions for the estimator of p in Section 4.3.1 and in Section 4.3.2 and for estimators of q in Section 4.3.3. The final step is to combine these results in order to obtain the asymptotic distribution of the estimator of ordinal pattern dependence, which is introduced in the next section.

4.3.4 Limit theorems for estimators of ordinal pattern dependence

We turn to the estimation of ordinal pattern dependence as it is defined in (4.13). The approach used in [58] is to show joint convergence of the heuristic estimators of the respective parameters p and q by using the multivariate delta method. We start with this idea, but we observe that a problem occurs in the considered case, in which the underlying time series are exhibiting long-range dependence. However, in the course of this section, we figure out a way to solve this problem and provide asymptotic distributions of ordinal pattern dependence for long-range dependent time series.

Lemma 4.42 *Under the assumptions* ***(L2s)*** *in Overview 4.24 with the addition that in the case* $d_1 = d_2$ *the linear representation of* $(Y_j)_{j\in\mathbb{Z}}$ *fulfills* (2.6) *the convergence results in Theorem 4.26 and Corollary 4.37 hold jointly.*

Proof. We restrict ourselves to the case that $d_1 = d_2 = d^*$. The other case, in which either $d^* = d_1$ or $d^* = d_2$, can be proved analogously.
Following (3.42), we know that

$$
\begin{aligned}
&(\hat{p}_n - p)\\
&= \sum_{p,q=1}^{d} \left(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right) + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q)} \left(\hat{r}_n^{(q,p)}(l) - r^{(q,p)}(l) \right) \right) + o_{\mathbb{P}}(n^{2d^*-1})
\end{aligned}
$$

and following the proof of Theorem 4.36,

$$
\begin{aligned}
&\left(n \left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi) \right)_{\pi\in S_h}, n \left(\hat{q}_{Y^{(1)},n}(\pi) - q_{Y^{(1)}}(\pi) \right)_{\pi\in S_h} \right)^t\\
=&A \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)} \right)^t + R_n.
\end{aligned}
$$

Applying the results of [18], p. 76, where joint convergence of the sample means and the sample autocorrelations independent of the lag was proved and recalling (4.30) we arrive at

$$
\begin{aligned}
&\Bigg(n^{1-2d^*} \sum_{p,q=1}^{d} \Bigg(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q)} \left(\hat{r}_n^{(p,q)}(l) - r^{(p,q)}(l) \right)\\
&\qquad + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q)} \left(\hat{r}_n^{(q,p)}(l) - r^{(q,p)}(l) \right) \Bigg) + o_{\mathbb{P}}(n^{2d^*-1}),\\
&\qquad n^{-d^*+\frac{1}{2}} A \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)} \right)^t + R_n \Bigg)^t\\
=& \left(n^{1-2d^*} \sum_{p,q=1}^{d} \tilde{\alpha}^{(p,q)} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0) \right), n^{-d^*+\frac{1}{2}} A \left(\sum_{j=1}^{n} Y_j^{(1)}, \sum_{j=1}^{n} Y_j^{(2)} \right)^t \right)^t + o_{\mathbb{P}}(1)\\
\xrightarrow{\mathcal{D}}& \left(C_2 \sum_{p,q\in P^*} \tilde{\alpha}^{(p,q)} Z_{2,d^*+1/2}^{(p,q)}(1), A \left(Y^{(1)}, Y^{(2)} \right)^t \right)^t,
\end{aligned}
$$

where $Z_{2,d^*+1/2}^{(p,q)}(1)$ and $\tilde{\alpha}^{(p,q)}$ are as given in Theorem 4.26 and $A, Y^{(1)}, Y^{(2)}$ as given in Theorem 4.36.
In the case $d_1 \neq d_2$, we assume w.l.o.g. that $d_1 > d_2$, the joint convergence of the sample mean and the sample-correlations of $(Y_j)_{j\in\mathbb{Z}}$ reduce to the joint convergence of

$$
\left(\sum_{j=1}^{n} H_1 \left(Y_j^{(1)} \right), \sum_{j=1}^{n} H_2 \left(Y_j^{(1)} \right) \right)^t.
$$

The joint convergence to the expected asymptotic distributions as given in the univariate case, respectively, is proved in [3], Theorem 3.3. Hence, the result follows. □

Remark 4.43 *Let us shortly comment on the dependence structure of*

$$\left(Z^{(p,q)}_{2,d^*+1/2}(1),\ p,q=1,2\right) \text{ and } \left(Y^{(1)}, Y^{(2)}\right)^t.$$

Note that both random vectors are components of the multivariate Hermite processes for $m = 2$ and $m = 1$, respectively. Therefore, the underlying Gaussian measure $\tilde{B}(\mathrm{d}x)$, described in Theorem 3.10, is the same for the time representation of both processes, as we have already seen in Section 2.7. For details, see for example [23], p. 5398, where $Y^{(1)}$ and $Y^{(2)}$ are denoted by $\mathcal{B}^{(1)}_{m,H}(1)$ and $\mathcal{B}^{(2)}_{m,H}(1)$, respectively, and $\mathcal{B}_{m,H}(t)$ is an operator fractional Brownian motion for $t \in [0,1]$. By the properties of $\tilde{B}(\mathrm{d}x)$, we know that the respective univariate components of these processes, e.g. $Z^{(1,1)}_{2,d^*+1/2}(1)$ and $Y^{(1)}$, are uncorrelated but not independent; see [3], Proposition 3.5 for details on the dependence properties. It is possible to extend this result to the multivariate case. The first and second order multivariate Hermite processes are uncorrelated, which is an immediate consequence of the orthogonality of Hermite polynomials, as described in Lemma 2.35, and of McKean's representation of Hermite processes [47]. For details, see [9], p. 196 f. The non-linear dependence structure of both processes in the multivariate case is a part of future research.

Corollary 4.44 *Under the assumptions in **(L2s)** in Overview 4.24 with the addition that in the case $d_1 = d_2$ the linear representation of $(Y_j)_{j\in\mathbb{Z}}$ fulfills (2.6) the convergence results in Theorem 4.27 and Corollary 4.37 hold jointly.*

Proof. We have joint convergence of the sample mean and the sample autocorrelations as given in [18], p. 76. The limit random variables in Theorem 4.27 have the same distribution as a linear combination of $Z^{(p,q)}_{2,d^*+1/2}(1)$, $p,q = 1,2$, which results from an application of the continuous mapping theorem to Corollary 3.6. Hence, the result follows. □

Due to the arguments in the proof of Corollary 4.44, it is clear that all joint convergence results in this section that hold for Theorem 4.26 remain true for Theorem 4.27. We, therefore, omit to emphasize this in the following. However, note that in the limit distributions of Theorem 4.45 and Theorem 4.47 the limit random variables corresponding to $\hat{p}_n$ could be replaced by those in Theorem 4.27.

Having Lemma 4.42 in mind, we proceed to derive asymptotic results for an estimator of ordinal pattern dependence. The function

$$f : \mathbb{R}^{2(h+1)!+1} \to \mathbb{R}, \quad f\left(u, (v_\pi)_{\pi\in S_h}, (w_\pi)_{\pi\in S_h}\right) = \frac{u - \sum_{\pi\in S_h} v_\pi w_\pi}{1 - \sum_{\pi\in S_h} v_\pi w_\pi}$$

used in [58], proof of Theorem 6, combined with the multivariate delta method, is a possible solution to the convergence problem of $\widehat{OPD} - OPD := \frac{\hat{p}_n - \hat{q}_n}{1-\hat{q}_n} - \frac{p-q}{1-q}$ as plug-in estimator.

Following Lemma 4.42, we indeed have joint convergence of the input estimators needed in f, but, unfortunately, we do not have the same normalizing factor, as it is required in the (multivariate) delta method.
Proving joint convergence of $\hat{p}_n$ and $\hat{q}_n$ would not be constructive either. Using the continuous mapping theorem yields the term $\frac{(\hat{p}_n-p)-(\hat{q}_n-q)}{1-(\hat{q}_n-q)}$, which does not coincide with the term whose convergence we are interested in, namely $\widehat{OPD}-OPD$. The other reason why we have to reject this approach is that each summand in $\frac{(\hat{p}_n-p)-(\hat{q}_n-q)}{1-(\hat{q}_n-q)}$ would need an own normalization constant, which reflects exactly the same problem due to which we had to refuse the delta method. Using the same normalization constant would either yield a convergence to infinity for one of the two estimators or one estimator would be forced to converge to zero. Either way, this does not seem a reasonable approach to estimate the value of ordinal pattern dependence.

Fortunately, there are two possible ways to solve this problem. The first one is the investigation of the alternative estimator $\hat{q}_{n,\text{alt}}$ of q, whose limit distribution is given in Theorem 4.39. For this reason, we define $\widehat{OPD}_{n,\text{alt}} := \frac{\hat{p}_n-\hat{q}_{n,\text{alt}}}{1-\hat{q}_{n,\text{alt}}}$. Note that the case $d_1 = d_2$ is the one more intriguing here, since otherwise, following the second part of Corollary 4.40, $\hat{q}_{n,\text{alt}}$ converges to zero and the convergence behaviour of $\hat{q}_{n,\text{alt}}$ of q, whose limit distribution is given in Theorem 4.39. Then, the convergence behaviour of $\widehat{OPD}_{n,\text{alt}}$ is reduced to the one of $\hat{p}_n - p$.

Theorem 4.45 *(i) Under the assumptions of **(L2)** in Overview 4.24, with $d^* = d_1 = d_2$ with the addition that in the case $d_1 = d_2$ the linear representation of $(Y_j)_{j\in\mathbb{Z}}$ fulfills (2.6) we have*

$$\begin{aligned}
&n^{1-2d^*}\left(\widehat{OPD}_{n,\text{alt}}-OPD\right)\\
&\xrightarrow{\mathcal{D}} \frac{1}{1-q}C_2^{\frac{1}{2}}\sum_{p,q=1}^{2}\tilde{\alpha}^{(p,q)}Z_{2,d^*+1/2}^{(p,q)}(1)\\
&\qquad+\frac{p-1}{(1-q)^2}\left(\sum_{\pi\in S_h}\tilde{\alpha}_\pi^{(1)}\tilde{\alpha}_\pi^{(2)}\right)C_1\left(L_{1,1}L_{2,2}\right)^{\frac{1}{2}}\tilde{Y}^{(1)}\tilde{Y}^{(2)},
\end{aligned}$$

such that

$$\left(\tilde{Y}^{(1)},\tilde{Y}^{(2)}\right)\sim\mathcal{N}\left(0,\begin{pmatrix}1 & \frac{1}{2}\left(L_{1,2}+L_{2,1}\right)\left(L_{1,1}L_{2,2}\right)^{-\frac{1}{2}}\\ \frac{1}{2}\left(L_{1,2}+L_{2,1}\right)\left(L_{1,1}L_{2,2}\right)^{-\frac{1}{2}} & 1\end{pmatrix}\right),$$

where $\tilde{\alpha}^{(p,q)}$, $Z_{2,d^*+1/2}^{(p,q)}(1)$, $p,q=1,2$ are as given in Theorem 4.26 and $\tilde{\alpha}_\pi^{(p)}$, $p=1,2$ are as given in Theorem 4.13.

*(ii) Under the assumptions of **(L2s)** in Overview 4.24, but with $d_1 \neq d_2$, we obtain*

$$\begin{aligned}
&n^{1-2d^*}C_2^{-\frac{1}{2}}(2)^{-\frac{1}{2}}L_{pp}^{-1}\left(\widehat{OPD}_{n,\text{alt}}-OPD\right)\\
&\xrightarrow{\mathcal{D}}\frac{1}{1-q}\tilde{\alpha}^{(p,p)}Z_{2,d^*+1/2}^{(p,p)}(1),
\end{aligned}$$

with $P^* = \{p\}$, $p=1$ or $p=2$. The limit random variable $Z_{2,d^*+1/2}^{(p,p)}(1)$ follows a standard Rosenblatt distribution and $\tilde{\alpha}^{(p,p)}$ are as given in Theorem 4.26.

Proof. We make use of Lemma 4.42. By applying the continuous mapping theorem as in the proof of Theorem 4.39, but for $f_2\left(u, (v_\pi)_{\pi \in S_h}, (w_\pi)_{\pi \in S_h}\right) = \left(u, \sum_{\pi \in S_h} v_\pi w_\pi\right)$, we obtain joint convergence in distribution of $(\hat{p}_n - p)$ and $(\hat{q}_{n,\text{alt}} - q)$. We can now apply the multivariate delta method, since we have the same normalizing constants for both estimators. We consider the function $f_3(u, v) := \frac{u-v}{1-v}$, $v \neq 1$. The partial derivates are given by $\frac{\partial f_3}{\partial u} = \frac{1}{1-v}$, and $\frac{\partial f_3}{\partial v} = \frac{u-1}{(1-v)^2}$. Note that $\widehat{OPD}_{n,\text{alt}} = f_3(\hat{p}_n, \hat{q}_{n,\text{alt}})$ and $OPD = f_3(p, q)$. Since $p \neq 1$ and $q \neq 1$, the partial derivates of f_3 with respect to p and q, respectively, exist and do not equal zero. Therefore, we obtain

$$
\begin{aligned}
&n^{1-2d^*}\left(\widehat{OPD}_{n,\text{alt}} - OPD\right) \\
&\xrightarrow{\mathcal{D}} \nabla f(p,q) \left(C_2^{\frac{1}{2}} \sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} Z^{(p,q)}_{2,d^*+1/2}(1), \left(\sum_{\pi \in S_h} \tilde{\alpha}^{(1)}_\pi \tilde{\alpha}^{(2)}_\pi\right) Y^{(1)} Y^{(2)} \right) \\
&\quad = \frac{1}{1-q} C_2^{\frac{1}{2}} \sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} Z^{(p,q)}_{2,d^*+1/2}(1) \\
&\qquad + \frac{p-1}{(1-q)^2} \left(\sum_{\pi \in S_h} \tilde{\alpha}^{(1)}_\pi \tilde{\alpha}^{(2)}_\pi\right) C_1 \left(L_{1,1} L_{2,2}\right)^{\frac{1}{2}} \tilde{Y}^{(1)} \tilde{Y}^{(2)},
\end{aligned}
$$

such that $\left(\tilde{Y}^{(1)}, \tilde{Y}^{(2)}\right) \sim \mathcal{N}\left(0, \begin{pmatrix} 1 & \frac{1}{2}\left(L_{1,2}+L_{2,1}\right)\left(L_{1,1}L_{2,2}\right)^{-\frac{1}{2}} \\ \frac{1}{2}\left(L_{1,2}+L_{2,1}\right)\left(L_{1,1}L_{2,2}\right)^{-\frac{1}{2}} & 1 \end{pmatrix}\right)$, $\tilde{\alpha}^{(p,q)}$, $Z^{(p,q)}_{2,d^*+1/2}(1)$, $p, q = 1, 2$ as given in Theorem 4.26 and $\tilde{\alpha}^{(p)}_\pi$, $p = 1, 2$ as given in Theorem 4.13.

In case (ii) is an immediate consequence of Corollary 4.40 that the limit distribution is only depending on the asymptotic behaviour of $\hat{p}_n - p$. □

Concerning the asymptotics of the alternative estimator of ordinal pattern dependence, we so far considered the case that at least one of the two parameters d_1 and d_2 is in $\left(\frac{1}{4}, \frac{1}{2}\right)$. By this assumption, we assured that the process transformed by the estimator remains long-range dependent. However, the natural question arises what happens if $d^* \in (-\infty, 0) \cup \left(0, \frac{1}{4}\right)$, such that the transformed process does no longer exhibit long-range dependence. For an estimator with Hermite rank 2, a central limit theorem holds in this case with the standard normalization $n^{\frac{1}{2}}$, as we have already seen in Theorem 4.35 for $\hat{p}_n$. If $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$ the joint convergence of the estimator of p and the alternative estimator of q was justified by the ability to reduce the convergence behaviour of these functionals to the joint convergence of the sample mean and the sample cross-correlations of the underlying time series. Considering functionals with Hermite rank 2, the main reason for this reduction possibility is given in the Reduction Theorem, as introduced in Theorem 2.30 and proved in [2] in the multivariate case. However, the result of this theorem only remains valid in case of Hermite rank 2 if $d^* \left(\frac{1}{4}, \frac{1}{2}\right)$. Therefore, we cannot exploit this reduction possibility anymore if $d^* \in \left(0, \frac{1}{4}\right)$. Hence, we need a different argumentation for the convergence behaviour of $\widehat{OPD}_{n,\text{alt}}$ in the short-range dependent setting. The setting we consider in $\widehat{OPD}_{n,\text{alt}}$ is interesting, since we are having a functional with Hermite rank 2 given by $\hat{p}_n$ and a product of functions with Hermite rank 1 concerning $\hat{q}_{n,\text{alt}}$. Since we assumed that

$d^* \in \left(0, \frac{1}{4}\right)$, we know that $\hat{p}_n$ is already short-range dependent and needs to be normed by $n^{\frac{1}{2}}$ to get a non-trivial asymptotic result. However, $\hat{q}_{n,\text{alt}}$ consists of the product of estimators that are still exhibiting long-range dependence. Hence, we note that the statement of Corollary 4.40 remains valid in this case (by now only the statement itself and not any joint convergence with Theorem 4.35).

Theorem 4.46 *Let (S) in Overview 4.24 hold, with $d^* \in \left(0, \frac{1}{4}\right)$. We obtain*

$$n^{\frac{1}{2}}\left(\widehat{OPD}_{n,\text{alt}} - OPD\right) \xrightarrow{\mathcal{D}} \frac{1}{1-q}\mathcal{N}\left(0, \sigma^2\right),$$

with σ^2 as given in Theorem 4.35.

Proof. For $d^* \in \left(0, \frac{1}{4}\right)$, we have $1 - 2d^* \in \left(\frac{1}{2}, 1\right)$. Hence, $n^{1-2d^*} > n^{\frac{1}{2}}$. Therefore, it holds that

$$\begin{aligned}
\mathbb{E}\left(n^{\frac{1}{2}}\left(\hat{q}_{n,\text{alt}} - q\right)\right)^2 &= \mathbb{E}\left(n^{\frac{1}{2}-1+2d^*}n^{1-2d^*}\left(\hat{q}_{n,\text{alt}} - q\right)\right)^2 \\
&= n^{4d^*-1}\mathbb{E}\left(n^{1-2d^*}\left(\hat{q}_{n,\text{alt}} - q\right)\right)^2 \\
&\to 0,
\end{aligned}$$

since $4d^* - 1 < 0$.
Hence, we proved that $n^{\frac{1}{2}}\left(\hat{q}_{n,\text{alt}} - q\right) \xrightarrow{L_2} 0$ and, therefore, $n^{\frac{1}{2}}\left(\hat{q}_{n,\text{alt}} - q\right) \xrightarrow{\mathbb{P}} 0$. Consequently, we obtain joint convergence in distribution of $\hat{p}_n$ and $\hat{q}_{n,\text{alt}}$ using Slutsky's theorem and the Cramer-Wold device. We apply the multivariate delta method and obtain, analogously to the result in the long-range dependent case with $d_1 \neq d_2$ given in Theorem 4.45 (*ii*) that the limit distribution is only determined by the asymptotic behaviour of $\hat{p}_n$. Therefore, in this transformed short-range dependent setting, even the case $d_1 = d_2 = d^*$ does not factor in the limit distribution of $\hat{q}_{n,\text{alt}}$ to the limit distribution of $\widehat{OPD}_{n,\text{alt}}$. □

Some of the last results may seem unsatisfactory, as they do not involve the asymptotics of the estimator q_{alt}. The reason for this is given by the Hermite rank 1 of the estimators $\hat{q}_{n,Y^{(p)}}(\pi)$, for $\pi \in S_h$ and $p = 1, 2$ and the resulting different normalization, depending on the strongest long-range dependence parameter, and albeit $n^{\frac{1}{2}}$. However, this variation of normalization constants can be avoided by using the estimator $\hat{q}_{n,\text{imp}}$, which consists of the sum of products of estimators exhibiting Hermite rank 2. Hence, another approach is to study the limit distribution of $\widehat{OPD}_{n,\text{imp}} := \frac{\hat{p}_n - \hat{q}_{n,\text{imp}}}{1-\hat{q}_{n,\text{imp}}}$. First we focus on the long-range dependent case, with $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$. This turns out to be interesting, since the limit random variables of $\hat{q}_{n,\text{imp}}$ can be found again in the limit random variables of $\hat{p}_n$, due to the asymptotic equality to the respective sample cross-correlations. Therefore, we only get different weighting factors in the linear combination of the components of the 2×2-dimensional Rosenblatt process.

Theorem 4.47 *We assume **(L2s)** is fulfilled. Then, the following statements are true:*

(i) If $d_1 = d_2 = d^*$, it holds that

$$
\begin{aligned}
&n^{1-2d^*} (C_2)^{-1/2} \left(\widehat{OPD}_{n,\mathrm{imp}} - OPD\right) \\
&\xrightarrow{\mathcal{D}} \left(\frac{\tilde{\alpha}^{(1,1)}}{1-q} - \frac{1-p}{(1-q)^2} \sum_{\pi \in S_h} q_{Y^{(2)}}(\pi) \tilde{\alpha}_\pi^{(1,1)}\right) Z_{2,d^*+\frac{1}{2}}^{(1,1)}(1) \\
&\quad + \left(\frac{\tilde{\alpha}^{(2,2)}}{1-q} - \frac{1-p}{(1-q)^2} \sum_{\pi \in S_h} q_{Y^{(1)}}(\pi) \tilde{\alpha}_\pi^{(2,2)}\right) Z_{2,d^*+\frac{1}{2}}^{(2,2)}(1) \\
&\quad + 2 \frac{\tilde{\alpha}^{(1,2)}}{1-q} Z_{2,d^*+\frac{1}{2}}^{(1,2)}(1).
\end{aligned}
$$

(ii) If $d_1 \neq d_2$, such that $d_p = d^*$ for exactly one $p = 1, 2$, we obtain

$$
\begin{aligned}
&n^{1-2d^*} (C_2)^{-1/2} \left(\widehat{OPD}_{n,\mathrm{imp}} - OPD\right) \\
&\xrightarrow{\mathcal{D}} \left(\frac{\tilde{\alpha}^{(p,p)}}{1-q} - \frac{1-p}{(1-q)^2} \sum_{\pi \in S_h} q_{Y^{(3-p)}}(\pi) \tilde{\alpha}_\pi^{(p,p)}\right) Z_{2,d^*+\frac{1}{2}}^{(p,p)}(1),
\end{aligned}
$$

with $\tilde{\alpha}^{(p,p)}$, $Z_{2,d^*+1/2}^{(p,p)}(1)$, $p, q = 1, 2$ as given in Theorem 4.26 and $\tilde{\alpha}_\pi^{(p,p)}$, $p = 1, 2$ as given in Theorem 4.18.

Proof. The convergence results in Theorem 4.26 and Lemma 4.41 hold jointly: by (3.44), (4.30) and the proof of Lemma 4.41, the crucial terms that determine the limit distribution are given by the sample cross-correlations, independent of the considered function, whereas the weighting factors given by $\tilde{\alpha}^{(p,q)}$ are highly dependent on the underlying function. Hence, an application of the continuous mapping theorem to Corollary 3.6 yields the joint convergence.
We distinguish between the cases $d_1 = d_2$ and $d_1 \neq d_2$.
In the first case, we have $d^* = d_1 = d_2$ and, therefore, by an application of the multivariate delta method, using the function

$$
f : \mathbb{R}^{2(h+1)!+1} \to \mathbb{R}, \quad f\left(u, (v_\pi)_{\pi \in S_h}, (w_\pi)_{\pi \in S_h}\right) = \frac{u - \sum_{\pi \in S_h} v_\pi w_\pi}{1 - \sum_{\pi \in S_h} v_\pi w_\pi},
$$

we arrive at

$$
\begin{aligned}
&n^{1-2d^*} (C_2)^{-1/2} \left(\widehat{OPD}_{n,\mathrm{imp}} - OPD\right) \\
&\xrightarrow{\mathcal{D}} \nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right) \\
&\quad \times \left(\sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} Z_{2,d^*+1/2}^{(p,q)}(1),\right. \\
&\qquad \left.\left(\tilde{\alpha}_\pi^{(1,1)} Z_{2,d_1+\frac{1}{2}}^{(1,1)}(1)\right)_{\pi \in S_h}, \left(\tilde{\alpha}_\pi^{(2,2)} Z_{2,d_2+\frac{1}{2}}^{(2,2)}(1)\right)_{\pi \in S_h}\right)^t
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} Z_{2,d^*+1/2}^{(p,q)}(1)}{1-q} \\
&\quad + \sum_{\pi \in S_h} \frac{(p-1) q_{Y^{(2)}}(\pi)}{(1-q)^2} \tilde{\alpha}_{\pi}^{(1,1)} Z_{2,d^*+\frac{1}{2}}^{(1,1)}(1) \\
&\quad + \frac{(p-1) q_{Y^{(1)}}(\pi)}{(1-q)^2} \tilde{\alpha}_{\pi}^{(2,2)} Z_{2,d^*+\frac{1}{2}}^{(2,2)}(1) \\
&= \left(\frac{\tilde{\alpha}^{(1,1)}}{1-q} - \sum_{\pi \in S_h} \frac{(1-p) q_{Y^{(2)}}(\pi)}{(1-q)^2} \tilde{\alpha}_{\pi}^{(1,1)} \right) Z_{2,d^*+\frac{1}{2}}^{(1,1)}(1) \\
&\quad + \left(\frac{\tilde{\alpha}^{(2,2)}}{1-q} - \sum_{\pi \in S_h} \frac{(1-p) q_{Y^{(1)}}(\pi)}{(1-q)^2} \tilde{\alpha}_{\pi}^{(2,2)} \right) Z_{2,d^*+\frac{1}{2}}^{(2,2)}(1) \\
&\quad + 2 \frac{\tilde{\alpha}^{(1,2)}}{1-q} Z_{2,d^*+\frac{1}{2}}^{(1,2)}(1)
\end{aligned}
$$

with

$$
\begin{aligned}
&\nabla f \left(u, (v_\pi)_{\pi \in S_h}, (w_\pi)_{\pi \in S_h} \right) \\
&= \left(\frac{1}{1 - \sum_{\pi \in S_h} v_\pi w_\pi}, \left(\frac{(u-1) w_\pi}{\left(1 - \sum_{\pi \in S_h} v_\pi w_\pi\right)^2} \right)_{\pi \in S_h}, \left(\frac{(u-1) v_\pi}{\left(1 - \sum_{\pi \in S_h} v_\pi w_\pi\right)^2} \right)_{\pi \in S_h} \right).
\end{aligned}
$$

Note that no entry of $\nabla f \left(p, (q_{Y^{(1)}}(\pi))_{\pi \in S_h}, (q_{Y^{(2)}}(\pi))_{\pi \in S_h} \right)$ is equal to zero, since we assumed $p \neq 1$ and $q \neq 1$.
In case $d_1 \neq d_2$ only the convergence behaviour of the marginal process with the larger long-range dependence parameter would be taken into consideration as it was proved in Theorem 3.10, (3.43). Therefore, the formula in the case $d_1 = d_2$ is reduced, respectively. □

Remark 4.48 *We have a closer look at the weighting factors of the limit distribution of $\widehat{OPD}_{n,\text{imp}}$. First we briefly illustrate the hypothetical case of independence. In this case, the covariance matrix $\Sigma_{2,h}$ would consist of four block matrices of size $h \times h$, such that all entries of the off-diagonal blocks would be equal to zero. More precisely,*

$$
\Sigma_{2,h} = \begin{pmatrix} \Sigma_{1,h} & 0 \\ 0 & \tilde{\Sigma}_{1,h}, \end{pmatrix}
$$

such that $\left(Y_1^{(1)}, \ldots, Y_h^{(1)} \right)^t \sim \mathcal{N}(0, \Sigma_{1,h})$ and correspondingly $\left(Y_1^{(2)}, \ldots, Y_h^{(2)} \right)^t \sim \mathcal{N}\left(0, \tilde{\Sigma}_{1,h}\right)$. Since we are considering a block-diagonal matrix, the inverse of this matrix is given as a block-diagonal matrix, too, where each block matrix gets inverted.

Concerning the vector of second order Hermite coefficients of $\hat{p}_n - p$, *we have*

$$c_{i,k} = \sum_{\pi} \mathbb{E}\left[\left(Y^{(i)}_{1,h}Y^{(k)}_{1,h} - \mathbb{E}\left(Y^{(i)}_{1,h}Y^{(k)}_{1,h}\right)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_1,\ldots,Y_h 1^{(1)}\right)=\tilde{\Pi}\left(Y^{(2)}_1,\ldots,Y_h 1^{(2)}\right)=\pi\right\}}\right]$$
$$= \begin{cases} \sum_\pi c^{\pi,\text{imp}}_{i,k,Y^{(1)}} q_{Y^{(2)}}(\pi), & \text{if } i,k = 1,\ldots,h, \\ \sum_\pi c^{\pi,\text{imp}}_{i,k,Y^{(1)}} c^{\pi,\text{imp}}_{i,k,Y^{(2)}}, & \text{if } i = 1,\ldots,h; k = h+1,\ldots,2h \text{ or vice versa}, \\ \sum_\pi c^{\pi,\text{imp}}_{i,k,Y^{(2)}} q_{Y^{(1)}}(\pi), & \text{if } i,k = h+1,\ldots,2h, \end{cases}$$

since we consider the hypothetical case of independence. We define

$$\tilde{C}_1 := (c_{i,k})_{1\leq i,k\leq h},$$
$$\tilde{C}_2 := (c_{i,k})_{1\leq i\leq h, h+1\leq k\leq 2h}$$

and

$$\tilde{C}_3 := (c_{i,k})_{h+1\leq i,k\leq 2h},$$

such that $C = \begin{pmatrix} \tilde{C}_1 & \tilde{C}_2 \\ \tilde{C}_2 & \tilde{C}_3. \end{pmatrix}$.
Concerning $\alpha = \Sigma^{-1}_{2,h} C \Sigma^{-1}_{2,h}$, *we obtain*

$$\alpha = \Sigma^{-1}_{2,h} C \Sigma^{-1}_{2,h} = \begin{pmatrix} \Sigma^{-1}_{1,h} & 0 \\ 0 & \Sigma^{-1}_{2,h} \end{pmatrix} \begin{pmatrix} \tilde{C}_1 & \tilde{C}_2 \\ \tilde{C}_2 & \tilde{C}_3 \end{pmatrix} \begin{pmatrix} \Sigma^{-1}_{1,h} & 0 \\ 0 & \Sigma^{-1}_{2,h} \end{pmatrix}$$
$$= \begin{pmatrix} \Sigma^{-1}_{1,h}\tilde{C}_1\Sigma^{-1}_{1,h} & \Sigma^{-1}_{1,h}\tilde{C}_2\Sigma^{-1}_{2,h} \\ \Sigma^{-1}_{1,h}\tilde{C}_2\Sigma^{-1}_{2,h} & \Sigma^{-1}_{2,h}\tilde{C}_3\Sigma^{-1}_{2,h} \end{pmatrix},$$

since we are dealing with symmetric matrices. Recall that for $p = 1, 2$, *we have*

$$\tilde{\alpha}^{(p,p)} = \sum_{i,k=1}^{h} \alpha_{(p-1)h+i,(p-1)h+k}.$$

Hence, we obtain for $p = 1$ *(*$p = 2$ *can be treated analogously)*

$$\Sigma^{-1}_{1,h}\tilde{C}_1\Sigma^{-1}_{1,h} = \sum_\pi \Sigma^{-1}_{1,h}\left(c^{\pi,\text{imp}}_{i,k,Y^{(1)}} q_{Y^{(2)}}(\pi)\right)_{1\leq i,k\leq h}\Sigma^{-1}_{1,h}$$
$$= \sum_\pi q_{Y^{(2)}}(\pi)\Sigma^{-1}_{1,h}\left(c^{\pi,\text{imp}}_{i,k,Y^{(1)}}\right)_{1\leq i,k\leq h}\Sigma^{-1}_{1,h}$$
$$= \sum_\pi q_{Y^{(2)}}(\pi)\alpha^{(1,1)}_\pi.$$

Therefore, the terms $\sum_\pi q_{Y^{(2)}}(\pi)\tilde{\alpha}^{(1,1)}_\pi$ *and* $\sum_\pi q_{Y^{(1)}}(\pi)\tilde{\alpha}^{(2,2)}_\pi$ *in the limit distribution of Theorem 4.47 correspond to those of* $\tilde{\alpha}^{(1,1)}$ *and* $\tilde{\alpha}^{(2,2)}$ *in the hypothetical case of independence. Furthermore* $p = q$ *and, therefore, the factors in the parentheses in the limit distribution in Theorem 4.47 would be equal to zero.*

For $d^* \in (-\infty, 0) \cup \left(0, \frac{1}{4}\right)$, representing the case of short-range dependence, the setting differs considerably from those discussed before. We want to apply the multivariate delta method. Therefore, we need joint convergence of $\hat{p}_n$ and

$$\left(\left(\hat{q}_{n,Y^{(1)},\text{imp}}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(\hat{q}_{n,Y^{(2)},\text{imp}}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right)^t$$

to a multivariate normal distribution. However, the reduction theorem and the resulting reduction to the convergence of the sample cross-correlations does not hold anymore. The framework needed in this case is given in the proofs of Theorem 2 and Theorem 4 in [2], which require techniques different to those we used in Chapter 3, consisting of the diagram formula and graph theory. For a detailed version of these proofs, we refer to [48], Chapter 5.

Theorem 4.49 *Let **(S)** in Overview 4.24 be fulfilled. We define*

$$\left(f_1, \ldots, f_{(2h+1)!+1}\right)^t := \begin{pmatrix} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}\right) = \tilde{\Pi}\left(Y_j^{(2)}, \ldots, Y_{j+h-1}^{(2)}\right)\right\}} - p \\ \left(\frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}\right) = \pi\right\}} - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h} \\ \left(\frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)}, \ldots, Y_{j+h-1}^{(2)}\right) = \pi\right\}} - q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h} \end{pmatrix}.$$

The corresponding Hermite coefficients depending on $(l_1, \ldots, l_h) \in \mathbb{N}_0^h \setminus \{0, \ldots, 0\}$ are given by

$$c_{l_1, \ldots, l_{2h}, r} := \mathbb{E}\left(f_r\left(Y_{j,h}\right) \prod_{k=1}^{2h} H_{l_k}\left(Y_{j,h}^{(k)}\right)\right).$$

It holds that

$$n^{\frac{1}{2}}\left(\widehat{OPD}_{n,\text{imp}} - OPD\right) \xrightarrow{D} \mathcal{N}\left(0, \sigma_{opd}^2\right),$$

with

$$\sigma_{opd}^2 = \nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right) \Sigma \left(\nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right)\right)^t,$$

where

$$\nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right) = \left(\frac{1}{1-q}, \left(\frac{(p-1) q_{Y^{(2)}(\pi)}}{(1-q)^2}\right)_{\pi \in S_h}, \left(\frac{(p-1) q_{Y^{(1)}(\pi)}}{(1-q)^2}\right)_{\pi \in S_h}\right)$$

and $\Sigma = \left(\sigma_{r,s}\right)_{r,s=1,\ldots,2(h+1)!+1}$ is given by

$$\sigma_{r,s} = \sum_{l_1, \ldots, l_{2h}, m_1, \ldots, m_{2h} = 2}^{\infty} c_{l_1, \ldots, l_{2h}, r} c_{m_1, \ldots, m_{2h}, s} \lim_{n \to \infty} n^{-1} \sum_{j_1, j_2 = 1}^{n} \mathbb{E}\left(\prod_{p=1}^{2h} H_{l_p}\left(Y_{j_1,h}^{(p)}\right) H_{m_p}\left(Y_{j_2,h}^{(p)}\right)\right).$$

Proof. We examine the joint convergence behaviour of

$$\left(\hat{p}_n - p, \left(\hat{q}_{n,Y^{(1)},\text{imp}}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi \in S_h}, \left(\hat{q}_{n,Y^{(2)},\text{imp}}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi \in S_h}\right)^t$$

in the case that $d^* \in \left(0, \frac{1}{4}\right)$. Note that in this case it holds

$$\sum_{k=1}^{\infty} \left(r^{(p,q)}(k)\right)^2 \leq \sum_{k=1}^{\infty} L_{p,q}(k)^2 k^{4d^*-2} < \infty,$$

since $4d^* - 2 \in (-2, -1)$.
Hence, the assumptions of Theorem 4 in [2] are fulfilled. To emphasize the crucial argument needed in the following, we use the proof of Theorem 2 in [2]. There the method of moments is applied, see [30], Theorem 8.6 and also [30] Section 4.10, to prove convergence to the normal distribution. It is shown in [2], p. 2250, that by using the multivariate Hermite expansion of the square-integrable function f with $\mathbb{E}f\left(Y_{j,h}\right) = 0$ we obtain

$$n^{-\frac{1}{2}} \sum_{j=1}^{n} \sum_{l_1,\ldots,l_{2h}=2}^{t} c_{l_1,\ldots,l_{2h}} \left(l_1 \cdot \ldots \cdot l_{2h}\right)^{-1} \prod_{p=1}^{2h} H_{l_p}\left(Y_{j,h}^{(p)}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_t^2\right),$$

where

$$\sigma_t^2 = \sum_{l_1,\ldots,l_{2h},m_1,\ldots,m_{2h}=2}^{t} c_{l_1,\ldots,l_{2h}} c_{m_1,\ldots,m_{2h}} \times \lim_{n\to\infty} n^{-1} \sum_{j_1,j_2=1}^{n} \mathbb{E}\left(\prod_{p=1}^{2h} H_{l_p}\left(Y_{j_1,h}^{(p)}\right) H_{m_p}\left(Y_{j_2,h}^{(p)}\right)\right). \tag{4.32}$$

The variable

$$c_{l_1,\ldots,l_{2h}} := \mathbb{E}\left(f\left(Y_{j,h}\right) \prod_{k=1}^{2h} H_{l_k}\left(Y_{j,h}^{(k)}\right)\right)$$

denotes the Hermite coefficients corresponding to $l_1, \ldots, l_{2h}$.
Note that the expected value in this formula has a closed form for $d = 1$, by Lemma 2.35, see also [54], Theorem 5.4.1 Therefore, it is possible to obtain a closed form of the variance in the univariate case, depending on the covariance function of the corresponding underlying process. Unfortunately, in dimensions $d \geq 2$, we do not have a closed formula and, therefore, need to apply the diagram formula. This formula states a relation between the expected value of interest and the cross-correlation functions of the underlying multivariate Gaussian process. They are linked via edges of so called diagrams. These edges are used as input parameters to determine the cross-correlation function. For details on this framework see [48], Section 4.3, as well as Section 5 for the proofs of limit theorems that make use of these mathematical tools.
We show that

$$\left(\hat{p}_n - p, \left(\hat{q}_{n,Y^{(1)},\mathrm{imp}}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(\hat{q}_{n,Y^{(2)},\mathrm{imp}}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)^t$$

jointly converge to a multivariate normal distribution with covariance matrix

$$\Sigma = \left(\sigma_{r,s}\right)_{r,s=1,\ldots,2(h+1)!+1}.$$

The entries of the covariance matrix Σ are given by

$$\sigma_{r,s} = \sum_{l_1,\dots,l_{2h},m_1,\dots,m_{2h}=2}^{\infty} c_{l_1,\dots,l_{2h},r} c_{m_1,\dots,m_{2h},s} \times \lim_{n\to\infty} n^{-1} \sum_{j_1,j_2=1}^{n} \mathbb{E}\left(\prod_{p=1}^{2h} H_{l_p}\left(Y_{j_1,h}^{(p)}\right) H_{m_p}\left(Y_{j_2,h}^{(p)}\right)\right), \tag{4.33}$$

and

$$c_{l_1,\dots,l_{2h},r} := \mathbb{E}\left(f_r\left(Y_{j,h}\right)\prod_{k=1}^{2h} H_{l_k}\left(Y_{j,h}^{(k)}\right)\right).$$

We employ the Cramer-Wold device to show that for $a := \left(a_1,\dots,a_{2(h+1)!+1}\right)^t \in \mathbb{R}^{2(h+1)!+1}$

$$n^{1/2}\left(a_1,\dots,a_{2(h+1)!+1}\right)\left(\hat{p}_n - p, \left(\hat{q}_{n,Y^{(1)},\text{imp}}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(\hat{q}_{n,Y^{(2)},\text{imp}}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)^t \xrightarrow{\mathcal{D}} \mathcal{N}(0,\sigma^2),$$

where $\sigma^2 = a\Sigma a^t$.

Each of the estimators above are consisting of functionals that are square integrable with respect to $Y_{j,h}$. We know that each of them can be expressed via the Hermite expansion. We show this exemplarily for $\hat{p}_n - p$, where the corresponding function is given by $f_1\left(Y_{j,h}\right) = \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\dots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_j^{(2)},\dots,Y_{j+h-1}^{(2)}\right)\right\}}$. Then, we obtain as Hermite expansion of f_1:

$$\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\dots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_j^{(2)},\dots,Y_{j+h-1}^{(2)}\right)\right\}} - p = \sum_{l_1,\dots,l_{2h}=2}^{\infty} \frac{c_{l_1,\dots,l_{2h},1}}{l_1!\dots l_{2h}!} \prod_{p=1}^{2h} H_{l_p}\left(Y_{j,h}^{(p)}\right), \tag{4.34}$$

where $c_{l_1,\dots,l_{2h},1}$ are the Hermite coefficients corresponding to f_1. Note that only the Hermite coefficients depend on the f_1 and hence are indexed by 1 as well. The right-hand side in (4.34) is equal for any function f that is square integrable with respect to $Y_{j,h}$, albeit the Hermite coefficients. The most technical part of the proof in Theorem 4, [2] is the calculation of the expected value in (4.32). In this calculation the Hermite coefficients are only used as scaling factors. The crucial point is the application of graph theory and using the diagram formula. Using the Hermite expansion we obtain

$$a\left(f_1,\dots,f_{2(h+1)!+1}\right)^t = \left(a_1,\dots,a_{2(h+1)!+1}\right)\begin{pmatrix} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\dots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_j^{(2)},\dots,Y_{j+h-1}^{(2)}\right)\right\}} - p \\ \left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\dots,Y_{j+h-1}^{(1)}\right)=\pi\right\}} - q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h} \\ \left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\dots,Y_{j+h-1}^{(2)}\right)=\pi\right\}} - q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h} \end{pmatrix}$$

$$
\begin{aligned}
&= \sum_{r=1}^{2(h+1)!+1} a_r \sum_{l_1,\ldots,l_{2h}=2}^{\infty} \frac{c_{l_1,\ldots,l_{2h},r}}{l_1!\ldots l_{2h}!} \prod_{p=1}^{2h} H_{l_p}\left(Y_{j,h}^{(p)}\right) \\
&= \sum_{l_1,\ldots,l_{2h}=2}^{\infty} \frac{\sum_{r=1}^{2(h+1)!+1} a_r c_{l_1,\ldots,l_{2h},r}}{l_1!\ldots l_{2h}!} \prod_{p=1}^{2h} H_{l_p}\left(Y_{j,h}^{(p)}\right).
\end{aligned}
$$

We define $\tilde{c}_{l_1,\ldots,l_{2h}} := \sum_{r=1}^{2(h+1)!+1} a_r c_{l_1,\ldots,l_{2h},r}$ and follow the proof of Theorem 4. We obtain

$$
\begin{aligned}
\sigma_t^2 = &\sum_{l_1,\ldots,l_{2h},m_1,\ldots,m_{2h}=2}^{t} \tilde{c}_{l_1,\ldots,l_{2h}} \tilde{c}_{m_1,\ldots,m_{2h}} \\
&\qquad \times \lim_{n\to\infty} n^{-1} \sum_{j_1,j_2=1}^{n} \mathbb{E}\left(\prod_{p=1}^{2h} H_{l_p}\left(Y_{j_1,h}^{(p)}\right) H_{m_p}\left(Y_{j_2,h}^{(p)}\right)\right) \\
= &\sum_{s,r=1}^{2(h+1)!+1} a_r a_s \sum_{l_1,\ldots,l_{2h},m_1,\ldots,m_{2h}=2}^{t} c_{l_1,\ldots,l_{2h},r} c_{m_1,\ldots,m_{2h},s} \\
&\qquad \times \lim_{n\to\infty} n^{-1} \sum_{j_1,j_2=1}^{n} \mathbb{E}\left(\prod_{p=1}^{2h} H_{l_p}\left(Y_{j_1,h}^{(p)}\right) H_{m_p}\left(Y_{j_2,h}^{(p)}\right)\right).
\end{aligned}
$$

Note that the redefined Hermite coefficients do not affect the proof, since they are only used as scaling factors. There is no problem regarding the finiteness of the Hermite expansion since they are only a linear combination of the original Hermite coefficients of the square integrable functions corresponding to the respective estimators. For more details, we refer to [54], proof of Theorem 5.7.1.
Therefore,

$$
\left(\hat{p}_n - p, \left(\hat{q}_{n,Y^{(1)},\mathrm{imp}}(\pi) - q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(\hat{q}_{n,Y^{(2)},\mathrm{imp}}(\pi) - q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)^t \xrightarrow{\mathcal{D}} \mathcal{N}(0,\Sigma),
$$

with Σ as given in (4.33).
We follow the proof of Theorem 4.47, and obtain

$$
n^{\frac{1}{2}}\left(\widehat{OPD}_{n,\mathrm{imp}} - OPD\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_{opd}^2\right),
$$

where $\sigma_{opd}^2 = \nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right) \Sigma \left(\nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)\right)^t$, with $\nabla f\left(p, \left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)$ as given in the proof of Theorem 4.47. □

Summarizing, in this section, we derived limit theorems for the plug-in estimator of ordinal pattern dependence. As we hoped, based on the results for the estimators of p and q, joint convergence results were obtained. To get a more illustrative impression of the asymptotic distributions, we conduct a simulation study in the following section.

4.3.5 Simulation studies

We begin with the generation of a bivariate long-range dependent fractional Gaussian noise series $\left(Y_j^{(1)}, Y_j^{(2)}\right)_{j=1,\dots,n}$. First, we simulate two independent fractional Gaussian noise processes $\left(U_j^{(1)}\right)_{j=1,\dots,n}$ and $\left(U_j^{(2)}\right)_{j=1,\dots,n}$ derived by the R-package "longmemo", see [34], for a fixed parameter $H \in \left(\frac{1}{2}, 1\right)$ in both time series. For the reader's convenience, we denote the long-range dependence parameter d by $H = d + \frac{1}{2}$ as it is common when dealing with fractional Gaussian noise and fractional Brownian motion, see Section 2.3. We refer to H as *Hurst parameter*, tracing back to the work of [33].

For $H = 0.7$ and $H = 0.8$ we generate $n = 10^6$ samples, for $H = 0.9$, we choose $n = 2 \cdot 10^6$. We denote the correlation function of univariate fractional Gaussian noise by $r_H^{(1,1)}(k)$, $k \geq 0$. Then, we obtain $\left(Y_j^{(1)}, Y_j^{(2)}\right)_j$ for $j = 1, \dots, n$:

$$\begin{aligned} Y_j^{(1)} &= U_j^{(1)}, \\ Y_j^{(2)} &= \psi U_j^{(1)} + \phi U_j^{(2)}, \end{aligned} \tag{4.35}$$

for $\psi, \phi \in \mathbb{R}$.

Note that this yields the following properties for the cross-correlations of the two processes for $k \geq 0$:

$$\begin{aligned} r_H^{(1,2)}(k) &= \mathbb{E}\left(Y_j^{(1)} Y_{j+k}^{(2)}\right) = \psi r_H^{(1,1)}(k) \\ r_H^{(2,1)}(k) &= r^{(1,2)}(-k) = \psi r_H^{(1,1)}(k) \\ r_H^{(2,2)}(k) &= \mathbb{E}\left(Y_j^{(2)} Y_{j+k}^{(2)}\right) = \left(\psi^2 + \phi^2\right) r_H^{(1,1)}(k). \end{aligned}$$

We use $\psi = 0.6$ and $\phi = 0.8$ to get unit variance in the second process. This yields a statistical setting with favourable symmetry properties of the multivariate normal distribution, since (IV) in (4.17) is fulfilled. Note that we choose the same Hurst parameter in both processes to get a better simulation result.

We restrict ourselves to the pattern length $h = 2$. Then, we need to determine $\Sigma_{2,2,H}$, depending on the value of the Hurst parameter. Given $r_H^{(1,1)}(1) = 2^{2H-1} - 1$ for fractional Gaussian noise, we obtain

$$\Sigma_{2,2,H} = \begin{pmatrix} 1 & 2^{2H-1} - 1 & \psi & \psi\left(2^{2H-1} - 1\right) \\ 2^{2H-1} - 1 & 1 & \psi\left(2^{2H-1} - 1\right) & \psi \\ \psi & \psi\left(2^{2H-1} - 1\right) & 1 & 2^{2H-1} - 1 \\ \psi\left(2^{2H-1} - 1\right) & \psi & 2^{2H-1} - 1 & 1 \end{pmatrix}.$$

To calculate the limit variance by the Matlab function $HR2LimVar$ we developed we need to determine the limits of the slowly varsing functions in the matrix L_H. With $r_H^{(1,1)}(k) \simeq H(2H-1)k^{2H-2}$ we obtain that

$$L_H = \begin{pmatrix} H(2H-1) & \psi H(2H-1) \\ \psi H(2H-1) & H(2H-1) \end{pmatrix}.$$

The simulations of the processes $\left(Y_j^{(1)}\right)_{j\in\mathbb{Z}}$ and $\left(Y_j^{(2)}\right)_{j\in\mathbb{Z}}$ are visualized in Figure 4.8. On the left-hand side the different fractional Gaussian noises depending on the Hurst parameter H are displayed. They represent the stationary long-range dependent Gaussian *increment* processes we need in the view of the limit theorems we derived in Section 4.3. The processes in which we are comparing the coincident ordinal patterns, namely $\left(X_j^{(1)}\right)_{j\in\mathbb{Z}}$ and $\left(X_j^{(2)}\right)_{j\in\mathbb{Z}}$, are shown on the right-hand side in Figure 4.8. The long-range dependent behaviour of the increment processes is very illustrative in these processes: roughly speaking they get smoother the larger the Hurst parameter gets.

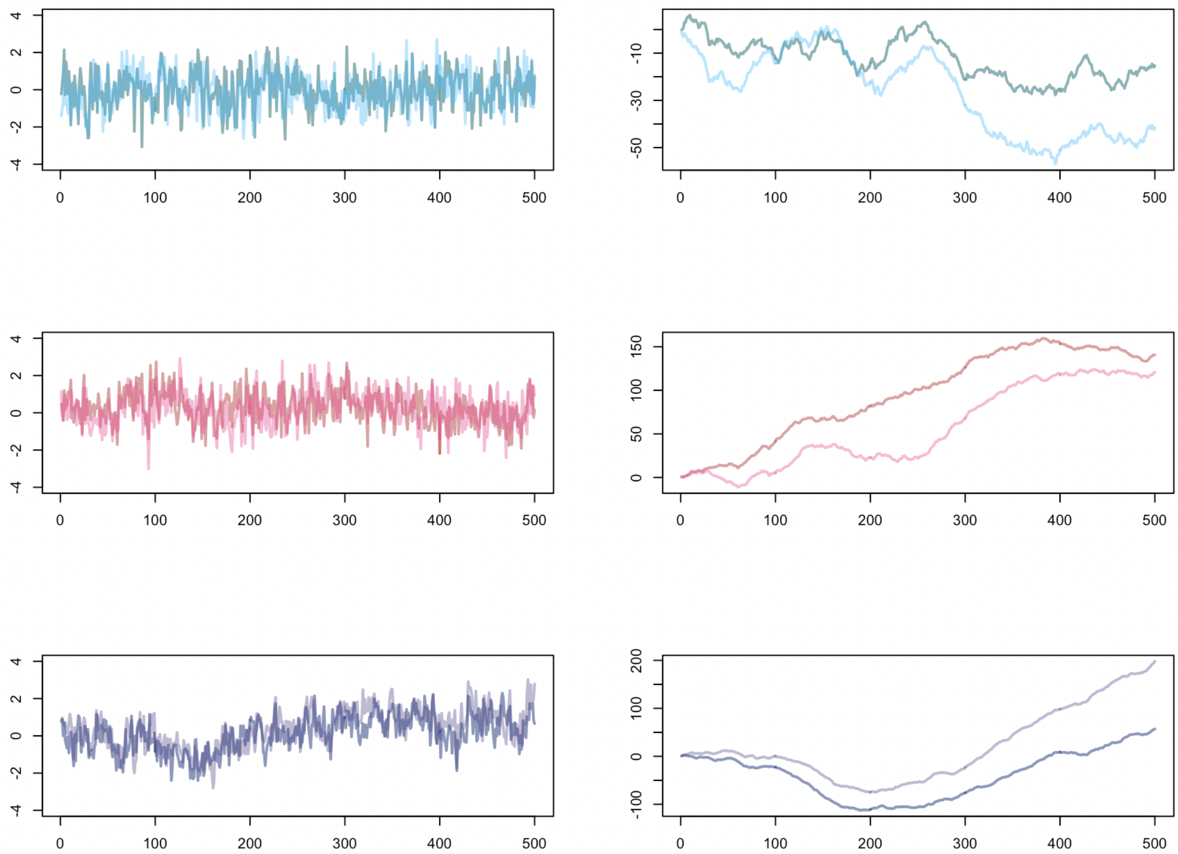

Figure 4.8: Plots of 500 data points of one path of two dependent fractional Gaussian noise processes (left) and the paths of the corresponding fractional Brownian motions (right) for different Hurst parameters: $H = 0.7$ (top), $H = 0.8$ (middle), $H = 0.9$ (bottom).

We turn to the outputs of the simulations concerning our limit theorems. We simulate $N = 10\,000$ paths of each increment process.

In order to get a better understanding of the distributions that play a key role in the determination of the asymptotic distribution of functionals with Hermite rank 2, see (3.44), we need to take a closer look on the convergence results of the left-hand side in Corollary 3.6. Therefore, we consider $H \in \left(\frac{3}{4}, 1\right)$ for the asymptotics of the sample cross-correlations. Since we restrict ourselves to the case $d = 2$, we need to consider the sample (cross-)correlations:

$$\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0) = \frac{1}{n}\sum_{j=1}^{n} H_2\left(Y_j^{(1)}\right), \tag{4.36}$$

$$\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0) = \frac{1}{n}\sum_{j=1}^{n} Y_j^{(1)}Y_j^{(2)} - \mathbb{E}\left(Y_j^{(1)}Y_j^{(2)}\right) \tag{4.37}$$

$$\hat{r}_n^{(2,2)}(0) - r_H^{(2,2)}(0) = \frac{1}{n}\sum_{j=1}^{n} H_2\left(Y_j^{(2)}\right). \tag{4.38}$$

In (4.36) and (4.38), we expect convergence to the standard Rosenblatt distribution, as we have already seen in Theorem 4.18 and in the corresponding simulations in Section 4.2.2. However, the case in (4.37) is novel, since we did not consider a bivariate process so far.

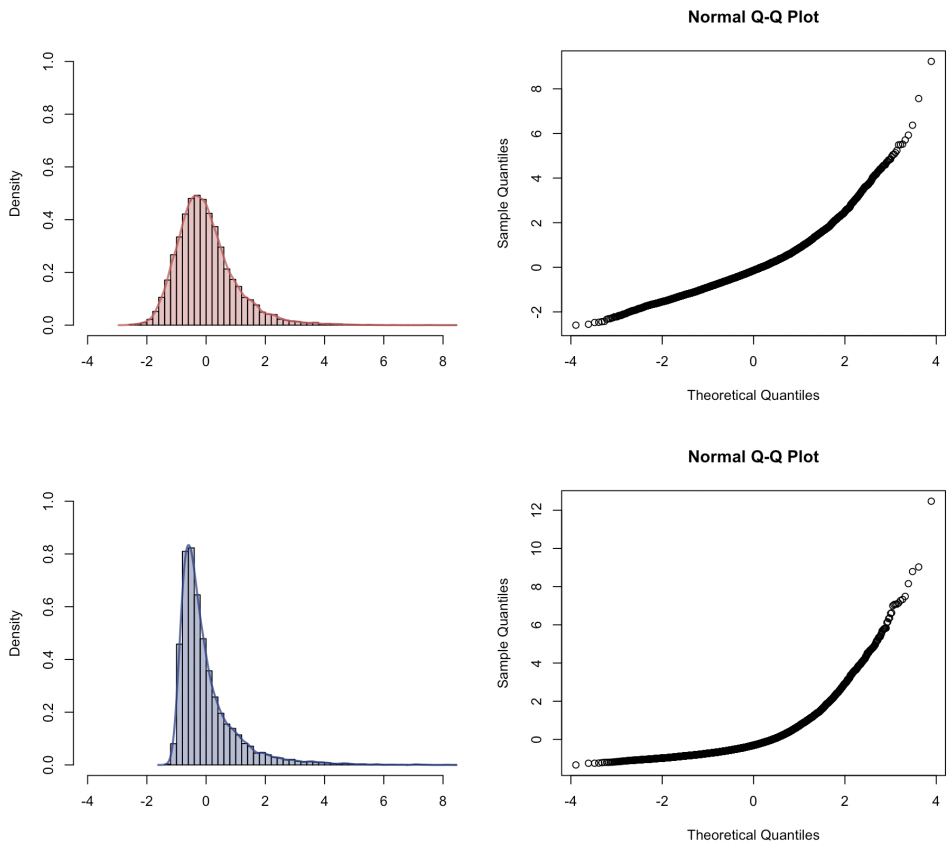

Figure 4.9: Histogram, kernel density estimation and normal Q-Q plot of $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0)$ for different Hurst parameters: $H = 0.8$ (top), $H = 0.9$ (bottom).

Using the normalization as described in Corollary 3.6 in order to get unit variance, the histograms and the sample densities of $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0)$ are displayed in Figure 4.9 for the respective Hurst parameters. A Q-Q plot with respect to the standard normal distribution is also provided to emphasize the deviation between the Rosenblatt distribution and the Gaussian distribution. These results coincide with the ones in Section 4.2.2. We use this simulation result as reference

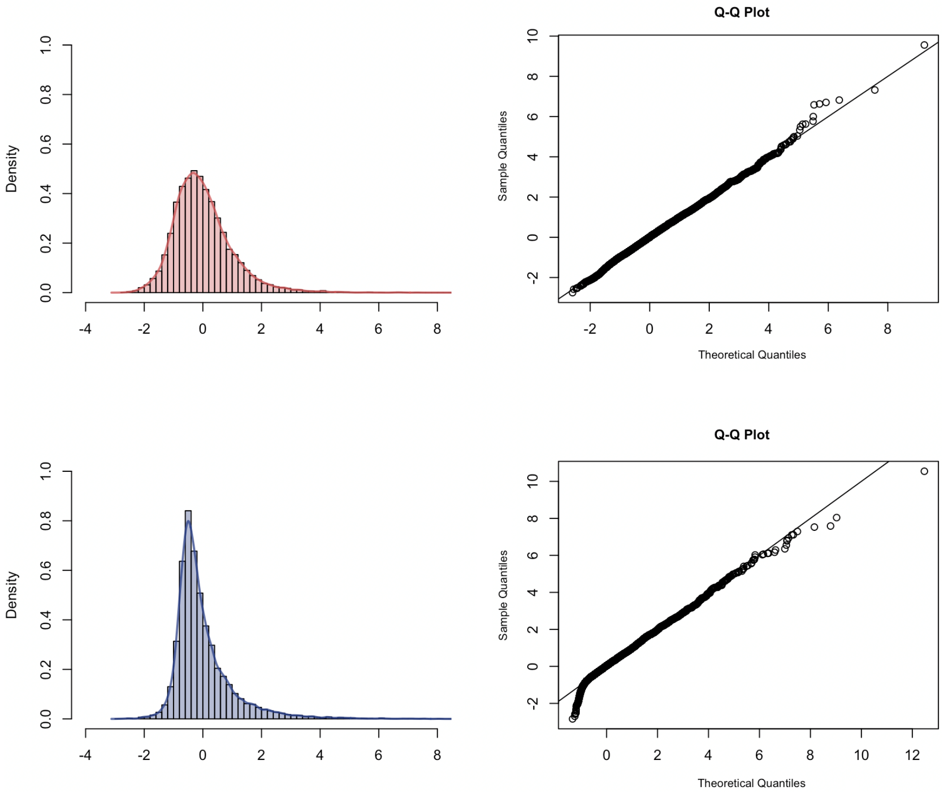

Figure 4.10: Histogram, kernel density estimation and Q-Q plot with respect to the Rosenblatt distribution of $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ for different Hurst parameters: $H = 0.8$ (top) and $H = 0.9$ (bottom).

data set for Q-Q plots with respect to the Rosenblatt distribution with corresponding parameter H in the following.

The simulated limit distribution of appropriately normalized $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ is given in Figure 4.10. The histograms and the sample densities are provided as well as the Q-Q plot with respect to the Rosenblatt distribution. We know from Remark 3.9 that the limit distribution of appropriately normalized $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ does not follow a Rosenblatt distribution, however, we observe a very small deviation. There are two parameters that are responsible for this result: first, we know from Remark 3.9 that for a correlation close to 1 in the entries beside the diagonals in the long-range dependence matrix L the asymptotic distribution of $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ approaches the Rosenblatt distribution with parameter H. The second reason is that the marginal asymptotic distributions of $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ and $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0)$ are equal for $H < \frac{3}{4}$. Therefore, for $H = 0.8$ the deviation between these two distributions is very small and difficult to detect. This behavior is reflected in Figure 4.10 for $H = 0.9$, however, for larger H the asymptotic simulation results get very unstable due to the small rate of convergence. To illustrate this reasoning, we change the parameters in (4.35) to $\psi = 0$ and $\phi = 1$, such that the processes $\left(Y_j^{(1)}\right)_{j=1,\dots,n}$ and $\left(Y_j^{(2)}\right)_{j=1,\dots,n}$ are independent. We simulate $n = 10^6$ data points and $N = 10\,000$ repetitions again. We observe the impact of the independence in Figure 4.11. The asymptotic distribution of $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0)$ still follows a Rosenblatt distribution as shown in Figure 4.9 whereas the the asymptotic distribution of $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ shows a large deviation from the Rosenblatt distribution now.

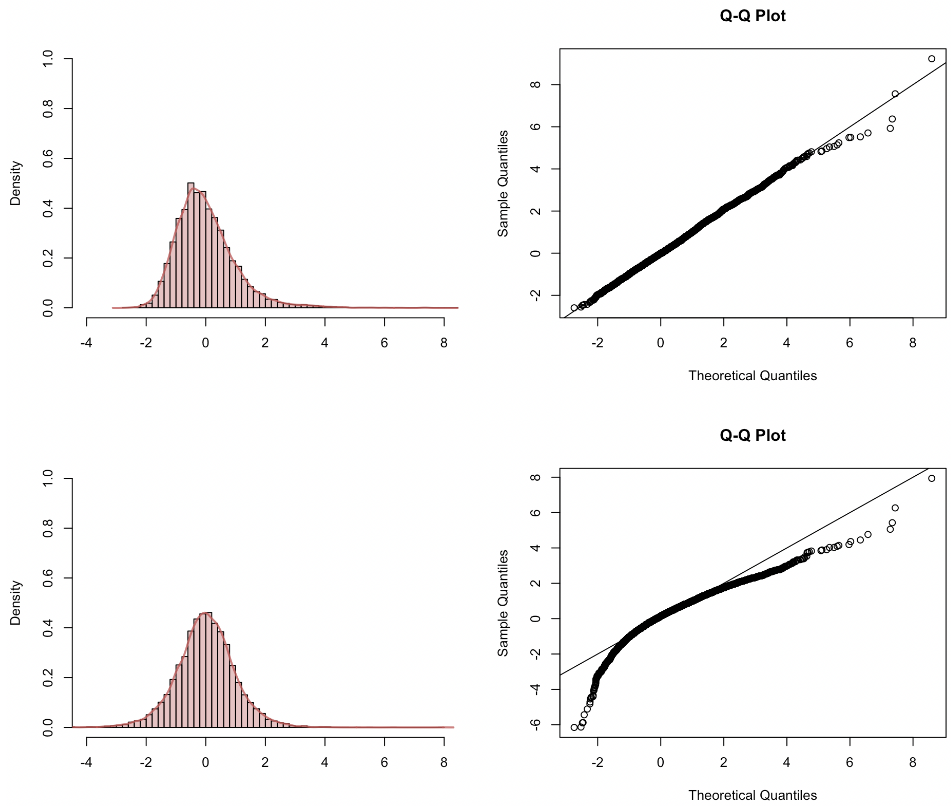

Figure 4.11: Histogram, kernel density estimation and Q-Q plot of $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0)$ (top) and of $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ (bottom) for two independent fractional Gaussian noise processes $\left(Y_j^{(1)}\right)_{j=1,\dots,n}$ and $\left(Y_j^{(2)}\right)_{j=1,\dots,n}$ for $H = 0.8$.

The limit case of this observation is to consider $H \to 1$ in which the Rosenblatt distribution approaches the χ^2-distribution with mean 0 and variance 1, following [66], p. 983. The corresponding simulation result is shown in Figure A.1 in the Appendix.

We return to to the asymptotics of (4.36), (4.37) and (4.38) for $\psi = 0.6$ and $\phi = 0.8$ in (4.35). The covariance structure described in Corollary 3.6 is confirmed in the simulations. The theoretical values of the covariances are

$$\begin{aligned}\operatorname{Cor}\left(\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0), \hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)\right) &= \operatorname{Cor}\left(\hat{r}_n^{(2,2)}(0) - r_H^{(1,1)}(0), \hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)\right) \\ &= \frac{\sqrt{2} L_{H,1,1} L_{H,1,2}}{\left(L_{H,1,1}^2 + L_{H,1,2}^2\right)^{1/2}} = 0.7276,\end{aligned}$$

and

$$\operatorname{Cor}\left(\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0), \hat{r}_n^{(2,2)}(0) - r_H^{(2,2)}(0)\right) = \frac{2L_{H,1,2}^2}{2L_{H,1,1}^2} = \left(\frac{L_{H,1,2}}{L_{H,1,1}}\right)^2 = 0.36.$$

Note that these values hold for $H = 0.8$ as well as $H = 0.9$, since $L_{0.9} = 1.5L_{0.8}$. The simulated correlation values are given in Table 4.2. In Figure 4.12 the bivariate densities of each two entries

	$\operatorname{Cor}\left(\hat{r}_n^{(1,1)}(0), \hat{r}_n^{(1,2)}(0)\right)$	$\operatorname{Cor}\left(\hat{r}_n^{(2,2)}(0), \hat{r}_n^{(1,2)}(0)\right)$	$\operatorname{Cor}\left(\hat{r}_n^{(1,1)}(0), \hat{r}_n^{(2,2)}(0)\right)$
$H = 0.8$	0.7331	0.7359	0.3729
$H = 0.9$	0.7341	0.7363	0.3767

Table 4.2: Simulated correlation between $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0), \hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ and $\hat{r}_n^{(2,2)}(0) - r_H^{(2,2)}(0)$ for different values of H.

of the sample covariance matrix are shown for $H = 0.8$.

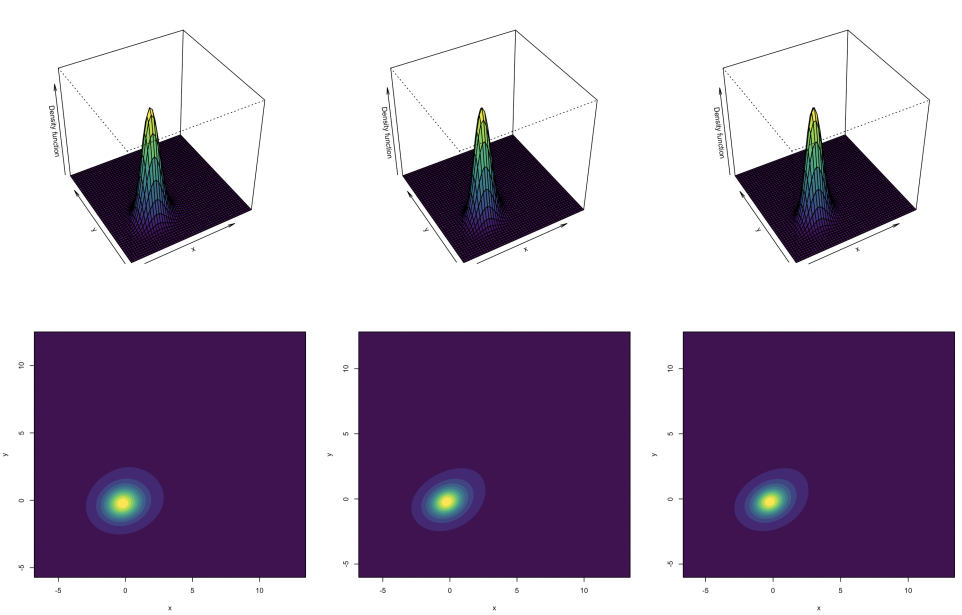

Figure 4.12: Probability density function and corresponding contour lines of $\left(Z_H^{(1,1)}(1), Z_H^{(2,2)}(1)\right)^t$, $\left(Z_H^{(1,1)}(1), Z_H^{(1,2)}(1)\right)^t$ and $\left(Z_H^{(1,2)}(1), Z_H^{(2,2)}(1)\right)^t$ for $H = 0.8$, created by using the R-packages "ks" and "viridis" and the results of [26], Sec. 3.1.

We know that the asymptotics of a $\mathbb{P}$-almost everywhere continuous, square-integrable function f with Hermite rank 2 for $H > \frac{3}{4}$ can be expressed by the three sample (cross-)correlations weighted by the corresponding Hermite coefficients transformed via the inverse of the covariance matrix, see (3.44). The corresponding limit theorem is given in Theorem 4.26. In the case $H = 0.7$ a different limit theorem holds, see Theorem 4.35. Therefore, we turn to the simulation results of the asymptotic distribution of the estimator of standard ordinal pattern dependence $\hat{p}_n$, as shown in Figure 4.13 for pattern length $h = 2$. The asymptotic normality in case $H = 0.7$ can be clearly

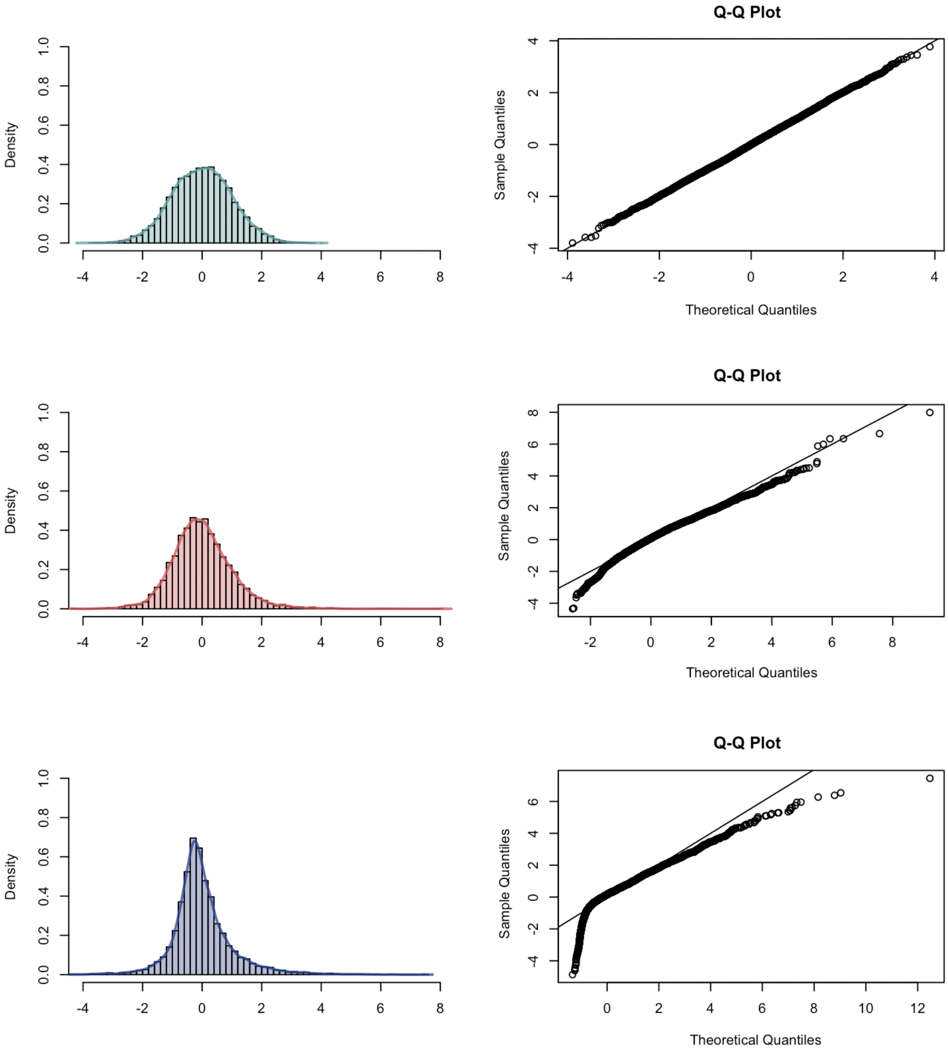

Figure 4.13: Histogram, kernel density estimation and Q-Q plot with respect to the normal distribution ($H = 0.7$) or to the Rosenblatt distribution of $\hat{p}_n - p$ with $h = 2$ for different Hurst parameters: $H = 0.7$ (top), $H = 0.8$ (middle), $H = 0.9$ (bottom).

observed. We turn to the interpretation of the simulation results of the distribution of $\hat{p}_n - p$ for $H = 0.8$ and $H = 0.9$ as weighted sum of the sample (cross-)correlations: we observe in the Q-Q plot for $H = 0.8$ that the samples in the upper and lower tail deviate from the reference line. For $H = 0.9$ a similar behaviour in the Q-Q plot is observed. We are able to explicitly calculate the weighting coefficients $\tilde{\alpha}^{(1,1)} = \tilde{\alpha}^{(2,2)}$ and $\tilde{\alpha}^{(1,2)} = \tilde{\alpha}^{(2,1)}$ here. The simulations confirm the results in Theorem 4.26 that the convergence behaviour of the distribution of $\hat{p}_n - p$ is equal to the convergence behaviour of the distribution of a linear combination of each term in (4.36), (4.37) and (4.38), weighted by the corresponding $\alpha^{(p,q)}$, $p = 1, 2$. This can be observed in Figure 4.14. Further, it is possible to determine the limit variance using the Matlab functions *cHR2* and *HR2LimVar* as described in Section A.3. The results are displayed in Table 4.3. The simulated limit variances of $\hat{p}_n - p$ normalized as described in Theorem 4.26 are 0.034 for $H = 0.8$ and 0.045 for $H = 0.9$. The deviation of these values to the theoretical values in Table 4.3 can be explained

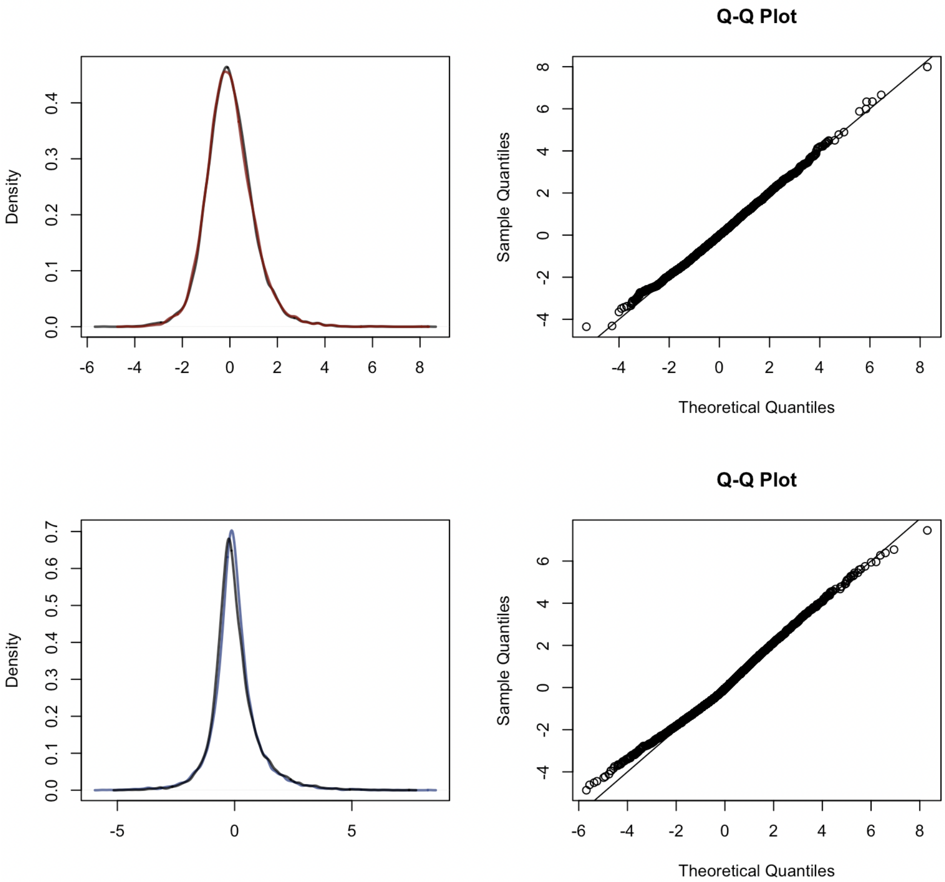

Figure 4.14: Kernel density estimation and Q-Q plot of $\tilde{\alpha}^{(1,1)}\left(\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0) + \hat{r}_n^{(2,2)}(0) - r_H^{(2,2)}(0)\right) + 2\tilde{\alpha}^{(1,2)}\left(\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)\right)$ and $\hat{p}_n - p$ for $H = 0.8$ (top) and $H = 0.9$ (bottom).

	$\tilde{\alpha}^{(1,1)}$	$\tilde{\alpha}^{(1,2)}$	Limit variance
$H = 0.8$	-0.2032	0.3413	0.0443
$H = 0.9$	-0.1876	0.2993	0.0724

Table 4.3: Theoretical values of $\tilde{\alpha}^{(1,1)}, \tilde{\alpha}^{(1,2)}$ and the limit variance of appropriately normalized $\hat{p}_n - p$ for $H = 0.8$ and $H = 0.9$.

by the slow convergence rate that we have in this setting. The marked terms in (3.39) only vanish asymptotically and, therefore, bias the variance. The very small absolute value of the variances above promote the instability of the convergence results. In the next section, we see that in the case of a function with Hermite rank 1 the results concerning the asymptotic variance get more stable due to a faster convergence rate. We want to verify the result in Theorem 4.27, that it is possible by a different weighting, to express the limit distribution of $\hat{p}_n - p$ as the distribution of the sum of two independent standard Rosenblatt random variables in the case considered here, see Corollary 3.14. Following (3.48), we have to consider the convergence behaviour of the second order Hermite polynomial applied to $Y_j^* := \frac{Y_j^{(2)} - Y_j^{(1)}}{\sqrt{2 - 2r_H^{(1,2)}(0)}}$ and $Y_j^{**} := \frac{Y_j^{(1)} + Y_j^{(2)}}{\sqrt{2 + 2r_H^{(1,2)}(0)}}$. We show exemplarily that

$$\frac{1}{n}\sum_{j=1}^{n} H_2\left(Y_j^*\right) = \frac{1}{n}\sum_{j=1}^{n} \frac{H_2\left(Y_j^{(1)}\right) + H_2\left(Y_j^{(1)}\right) - 2\left(Y_j^{(1)}Y_j^{(2)} - r_H^{(1,2)}(0)\right)}{2 - 2r_H^{(1,2)}(0)}$$
$$= \frac{\hat{r}_n^{(1,1)} - r_H^{(1,1)}(0) + \hat{r}_n^{(2,2)} - r_H^{(2,2)}(0) - 2\left(\hat{r}_n^{(1,2)} - r_H^{(1,2)}(0)\right)}{2 - 2r_H^{(1,2)}(0)}.$$

Using this formula the simulated convergence result for appropriately normalized $\frac{1}{n}\sum_{j=1}^{n} H_2\left(Y_j^*\right)$ is provided in Figure 4.15.

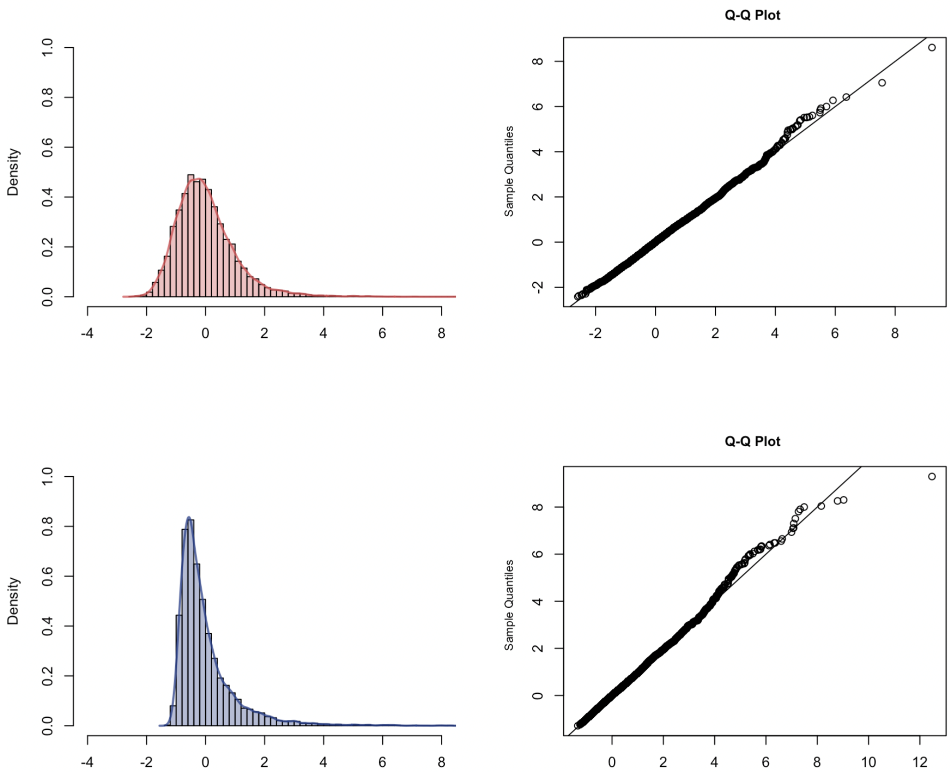

Figure 4.15: Histogram, kernel density estimation and Q-Q plot with respect to the Rosenblatt distribution of $\frac{1}{n}\sum_{j=1}^{n} H_2\left(Y_j^*\right)$ for different Hurst parameters: $H = 0.8$ (top), $H = 0.9$ (bottom).

We observe the standard Rosenblatt distribution. The results for $\frac{1}{n}\sum_{j=1}^{n} H_2\left(Y_j^{**}\right)$ yield similar results.

Due to the reasons described above, it is in general difficult to obtain satisfying simulation results in the case $H = 0.9$ (or H even closer to 1) since the simulations turn out to be very instable. We used the sample size $n = 2000000$. However, the computational effort is very large, for $N = 1000$ simulations of these two paths of fractional Gaussian noise with $H = 0.9$ and their application to the functions under consideration using parallel computing on all kernels around 40 hours of computational time is needed.

We also omitted the case to consider two different Hurst parameters for the two fractional Gaussian noises. We know that the influence of the smaller parameter vanishes asymptotically. However, as already seen above in the considerations to the second order Hermite coefficients, the influence vanishes very slowly and, therefore, the simulated limit distribution would be biased.

We illustrated our theoretical results by providing the sample densities of univariate random variables arising as limit random variables of functionals with Hermite rank 2. We observed sample densities equal to those in the univariate case as considered in Section 4.2.2 representing the standard Rosenblatt distribution. Taking $\hat{p}_n - p$ as an example of an underlying function with Hermite rank 2, we have seen in Figure 4.13 that the different ranges of H are justified in the simulations: for $H = 0.7$, asymptotic normality as proved in Theorem 4.35 is validated. For $H = 0.8$ and $H = 0.9$ we observe in Figure 4.14 that the asymptotic behaviour of the distribution of appropriately normalized $\hat{p}_n - p$ and the linear combination of the sample cross-correlation is almost equal. This confirms our results in Theorem 4.26. In general, the theoretical investigation of the distribution of a linear combination of the entries of the limit distribution of the sample covariance matrix $\hat{\Gamma}_{Y,n}(0) - \Gamma_Y(0)$ of (Y_j) is an interesting open research challenge.

4.4 Adapted and generalized concepts of ordinal pattern dependence

In this section, we enlarge the applicability of ordinal pattern dependence. We allow for certain changes of the standard setting, which we have investigated in details in the previous section, to be able to analyze a wider range of data sets in a more efficient way. As in the previous section, we denote by $(X_j)_{j\in\mathbb{Z}}$ with $X_j = \left(X_j^{(1)}, X_j^{(2)}\right)^t$ the bivariate Gaussian process in which we compare the ordinal patterns and by $(Y_j)_{j\in\mathbb{Z}}$ with $Y_j = \left(Y_j^{(1)}, Y_j^{(2)}\right)^t$ the corresponding increment process.

4.4.1 Estimator of ordinal pattern dependence for a single fixed pattern

First, we examine the case in which only one certain ordinal pattern $\pi \in S_h$ in coincident moving windows of the two time series under consideration is of interest. We focus on the probability

$$\begin{aligned} p(\pi) :&= \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right) = \pi\right) \\ &= \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)}, \ldots, Y_h^{(1)}\right) = \tilde{\Pi}\left(Y_1^{(2)}, \ldots, Y_h^{(2)}\right) = \pi\right), \end{aligned} \tag{4.39}$$

for $\pi \in S_h$.
Hence, the case of hypothetical independence is given by $q_{n,Y^{(1)}}(\pi)q_{n,Y^{(2)}(\pi)}$ and is, therefore, a special case of the considerations in Section 4.3.3. We omit the investigation of the quotient analog to the definition of standard ordinal pattern dependence in(4.13), as the argumentation works similar to the one in Section 4.3.4.
Note that the probability in (4.39) is one summand in the probability of coincident patterns in

(4.11), but also reminds of the already considered univariate setting in Section 4.2. We define the corresponding estimator based on the relative frequency as

$$\hat{p}_n(\pi) := \frac{1}{n-h} \sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)=\pi\right\}} \tag{4.40}$$
$$= \frac{1}{n-h} \sum_{j=1}^{n-h+1} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\pi\right\}}.$$

Note that the main difference to the setting considered in Section 4.2 is given by the fact that we are no longer considering consecutive random variables within one time series but within two dependent time series respectively.

Remark 4.50 *It is indeed possible to express the event described in* (4.40) *in terms of* $\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)},X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)$. *Note that there exists a possibility to capture the event*

$$\frac{1}{n-h} \sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_h^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_h^{(2)}\right)\right\}}$$
$$= \frac{1}{n-h} \sum_{j=0}^{n-h-1} \sum_{\pi\in S_h} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)=\pi\right\}}$$
$$= \frac{1}{n-h} \sum_{j=0}^{n-h-1} \sum_{\pi\in S_h} \sum_{\sigma(\pi)\in\tilde{S}_{2h+1}} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)},X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)=\sigma(\pi)\right\}}. \tag{4.41}$$

Hence, there exists a way to capture the event of a coincident patterns $\pi \in S_h$ *of the random vectors* $\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)$ *and* $\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)$ *in one single pattern* $\sigma(\pi)$, *by plugging the joint random vector* $\left(X_j^{(1)},\ldots,X_{j+h}^{(1)},X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)$ *into the map* Π. *Now, we need to figure out which pattern* $\sigma(\pi)$ *are included in* $\tilde{S}_{2h+1}$.
We consider the following example. We choose $h = 2$ *and the pattern* $\pi = (0,1,2)$. *Then, it is easy to see that the event*

$$\left\{\Pi\left(X_0^{(1)},X_1^{(1)},X_2^{(1)}\right) = \Pi\left(X_0^{(2)},X_1^{(2)},X_2^{(2)}\right) = (0,1,2)\right\},$$

which coincides with

$$\left\{X_2^{(1)} \leq X_1^{(1)} \leq X_0^{(1)}\right\} \cap \left\{X_2^{(2)} \leq X_1^{(2)} \leq X_0^{(2)}\right\} \tag{4.42}$$

contains for example the event

$$\left\{\Pi\left(X_0^{(1)},X_1^{(1)},X_2^{(1)},X_0^{(2)},X_1^{(2)},X_2^{(2)}\right) = (0,1,2,3,4,5)\right\}$$

for a permutation $\sigma(\pi) = (0,1,2,3,4,5) \in S_5$. *This event is equivalent to the inequality*

$$\left\{X_2^{(1)} \leq X_1^{(1)} \leq X_0^{(1)} \leq X_2^{(2)} \leq X_1^{(2)} \leq X_0^{(2)}\right\}.$$

Of course, there are more options to choose $\sigma(\pi)$ *for fixed* π. *We could also use for example* $\sigma(\pi) = (0,1,3,4,5,2)$ *or* $\sigma(\pi) = (0,3,1,4,2,5)$. *These permutations yield the events*

$$\left\{X_2^{(1)} \leq X_2^{(2)} \leq X_1^{(2)} \leq X_0^{(2)} \leq X_1^{(1)} \leq X_0^{(1)}\right\}$$

and

$$\left\{X_2^{(1)} \leq X_2^{(2)} \leq X_1^{(1)} \leq X_1^{(2)} \leq X_0^{(1)} \leq X_0^{(2)}\right\},$$

respectively. The inequalities also fulfill the condition above, namely $\left\{X_2^{(1)} \leq X_1^{(1)} \leq X_0^{(1)}\right\}$ *and* $\left\{X_2^{(2)} \leq X_1^{(2)} \leq X_0^{(2)}\right\}$.

The first important insight is that the original event in (4.42) *does not provide us any ordinal information **between** the random vectors* $\left(X_j^{(1)}, \ldots, X_{j+h}^{(1)}\right)$ *and* $\left(X_j^{(2)}, \ldots, X_{j+h}^{(2)}\right)$. *By using the joint random vector we are artificially generating an ordinal relation between them, although actually there is none. The summation of all* $\sigma(\pi) \in \tilde{S}_{2h+1}$ *can be seen as a way of computing the total probability since we are eliminating the additional constraints on the ordinal information between the two random vectors by adding all possibilities up.*
We need to determine the set of permutations that generates $\tilde{S}_{2h+1}$*, namely all permutations* $\sigma(\pi)$ *that fulfill the two inequalities in* (4.42). *In our example, roughly speaking, we have to assure that reading the permutation* $\sigma(\pi)$ *from left to right, the only two constraints are that* 0 *is before* 1 *is before* 2 *and that* 3 *is before* 4 *is before* 5. *So this is exactly displaying the nonexistent ordinal information between the two random vectors since there are no restrictions between for example* 1 *and* 5.

In general, for a permutation $\pi \in S_h$*, we can formulate the following restrictions for a pattern* $\sigma(\pi)$ *belonging to* $\tilde{S}_{2h+1}$. *For* $\pi = (\pi_0, \pi_1, \ldots, \pi_h)$ *and* $\sigma(\pi) = (\sigma_0, \sigma_1, \ldots, \sigma_{2h+1})$ *denote the index of the permutation* $\sigma(\pi)$ *by* $k(i)$ *such that* $\sigma_{k(i)}(\pi) = \pi_i$ *for* $i = 0, \ldots, h$ *and* $\sigma_{k(i)}(\pi) = \pi_i + h + 1$ *for* $i = h+1, \ldots, 2h+1$. *Then,* $\sigma(\pi)$ *is in the set of permutations* $\tilde{S}_{2h+1}$ *if and only if* $k(0) < k(1) < \ldots < k(h)$ *and* $k(h+1) < k(h+2) < \ldots < k(2h+2)$.
In this case, the events

$$\left\{\Pi\left(X_j^{(1)}, \ldots, X_{j+h}^{(1)}\right) = \Pi\left(X_j^{(2)}, \ldots, X_{j+h}^{(2)}\right) = \pi\right\}$$

and

$$\bigcup_{\sigma(\pi) \in \tilde{S}_{2h+1}} \left\{\Pi\left(X_j^{(1)}, \ldots, X_{j+h}^{(1)}, X_j^{(2)}, \ldots, X_{j+h}^{(2)}\right) = \sigma(\pi)\right\}$$

are equivalent.

It is possible to determine the number of permutations in $\tilde{S}_{2h+1}$ *by the following considerations. If we have allocated the* $h+1$ *entries of* $\pi \in S_h$ *to the permutation* $\sigma(\pi) \in S_{2h+1}$*, then the remaining entries of the permutation at the positions* $\pi_0 + h + 1, \ldots, \pi_h + h + 1$ *are fixed, since we may not disturb the given order. Then, we need to determine the number of possibilities to allocate* $h+1$ *numbers of a fixed order to* $2h+2$ *places. This corresponds exactly to the set* $\mathcal{K}_{h+1}^{\{0,1,\ldots,2h+1\}}$*, which denotes the set of* $h+1$*-combinations of* $\{0, 1, 2, \ldots, 2h+1\}$ *without repetition. The cardinality of this set is given by* $\#\mathcal{K}_{h+1}^{\{0,1,\ldots,2h+1\}} = \binom{2h+2}{h+1}$. *It is well known that the cardinality of* S_h *is given by* $\#S_h = (h+1)!$ *and hence the cardinality of the subset of* S_{2h+1} *that fulfills the conditions above and forms the set* $\tilde{S}_{2h+1}$ *is given by* $\frac{\#\mathcal{K}_{h+1}^{\{1,\ldots,2h+1\}}}{(2h+2)!} = \frac{1}{(h+1)!(h+1)!}$.
However, in order to determine the limit distribution of $\hat{p}_n(\pi) - p(\pi)$, $\pi \in S_h$ *these considerations*

show that the assumptions on the underlying time series in Theorem 3.3 or in Theorem 3.10 are only fulfilled for $d = 2$ but not for $d = 1$. In (4.41) the underyling random vector do not fulfill univariate stationarity as given in Definition 2.3. Therefore, we may not apply the results of Section 4.2.

We have shown that it is not possible to represent the parameter of interest in terms of a modified univariate stationary Gaussian random vector. In the following, we turn to the bivariate case.

Lemma 4.51 *The Hermite rank $m(f_\pi, \Sigma_{2,h})$, of*

$$f_\pi(Y_{j,h}) = \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\dots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_j^{(2)},\dots,Y_{j+h-1}^{(2)}\right)=\pi\right\}} - \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)},\dots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\dots,Y_h^{(2)}\right)=\pi\right)$$

with respect to $\Sigma_{2,h}$ is equal to 1.

Proof. Following Lemma 2.33 we need to show that $m(f_\pi, I_{2,h}) = 1$. Define $U_{1,h} := \left(U_1^{(1)},\dots,U_h^{(1)},U_1^{(2)},\dots,U_h^{(2)}\right)^t$ with $U_{1,h} \sim \mathcal{N}(0, I_{2,h})$ Let without loss of generality be $i = 1,\dots,h$. It holds that

$$\begin{aligned}\mathbb{E}\left(U_{1,h}^{(i)} f_\pi(U_{1,h})\right) &= \mathbb{E}\left(U_{1,h}^{(i)}\mathbf{1}_{\left\{\tilde{\Pi}\left(U_1^{(1)},\dots,U_h^{(1)}\right)=\tilde{\Pi}\left(U_1^{(2)},\dots,U_h^{(2)}\right)=\pi\right\}}\right)\\ &= \mathbb{E}\left(U_i^{(1)}\mathbf{1}_{\left\{\tilde{\Pi}\left(U_1^{(1)},\dots,U_h^{(1)}\right)=\pi\right\}}\right)\mathbb{P}\left(\tilde{\Pi}\left(U_1^{(2)},\dots,U_h^{(2)}\right)=\pi\right)\\ &\neq 0,\end{aligned}$$

by the same argument as used in the proof of Lemma 4.12. □

We now turn to the asymptotic properties of the estimator of ordinal pattern dependence for one fixed pattern. We consider the dependence structure of the underlying bivariate Gaussian time series for the following cases: We examine the pure long-range dependence case and the mixed case, where at least one of the two considered increment processes is long-range dependent, but the other one is allowed to be short-range dependent, as given in **(L1s)**. Both can be handled with an application based on Theorem 3.3, namely by using its extended version given in Corollary 3.4.

Theorem 4.52 *Under the assumptions of **(L1s)** in Overview 4.24, it holds*

$$n^{-d^*+1/2} C_1^{-1/2}\left(\hat{p}_n(\pi) - p(\pi)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \left(\sum_{p,q\in P^*} \tilde{\alpha}_{\pi,2}^{(p)}\tilde{\alpha}_{\pi,2}^{(q)} L_{p,q}\right)\right)$$

with $C_1 := \frac{1}{d^*(2d^*+1)}$, $\tilde{\alpha}_{\pi,2}^{(p)} = \sum_{j=1}^{h}\alpha_{j+(p-1)h}^{\pi,2}$ for $p = 1, 2$. Further, we have

$$\alpha^{\pi,2} = \left(\alpha_1^{\pi,2},\dots,\alpha_{2h}^{\pi,2}\right)^t = \Sigma_{2,h}^{-1} c^{\pi,2}$$

and

$$c^{\pi,2} := \mathbb{E}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\ldots,Y_h^{(2)}\right)=\pi\right\}}\left(Y_1^{(1)},\ldots,Y_h^{(1)},Y_1^{(2)},\ldots,Y_h^{(2)}\right)^t\right)$$

and P^*, d^* *as defined in* (3.10).

Proof. We already determined the Hermite rank of the function $f_\pi\left(Y_{j,h}\right)$ with respect to $\Sigma_{2,h}$ in Lemma 4.51. Since $\mathbb{E}\left(f_\pi\left(Y_{j,h}\right)^2\right) = p(\pi)(1-p(\pi)) < \infty$, the proof follows by an application of Corollary 3.4. □

Asymptotic variance

We need to determine c, the vector of first order Hermite coefficients of the function f_π. Then, we are able to explicitly calculate the limit variance.
By the symmetry properties (I)-(IV) in (4.14) to (4.17), we obtain the following equalities in terms of the Hermite coefficients. Recall that

$$c^{\pi,2} = \mathbb{E}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\ldots,Y_h^{(2)}\right)=\pi\right\}}\left(Y_1^{(1)},\ldots,Y_h^{(1)},Y_1^{(2)},\ldots,Y_h^{(2)}\right)^t\right),$$

such that

$$c_i^{\pi,2} = \mathbb{E}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\ldots,Y_h^{(2)}\right)=\pi\right\}} Y_{1,h}^{(i)}\right), \quad i = 1,\ldots,2h.$$

Then, we obtain the following equalities for $i = 1,\ldots,2h$, $k = 1,\ldots,h$ and $l = h+1,\ldots,2h$:

$$\begin{aligned}
c_i^{\pi,2} &= -c_i^{\mathcal{S}(\pi),2} && (I)\\
c_i^{\pi,2} &= c_{2h+1-i}^{\mathcal{T}\circ\mathcal{S}(\pi),2} && (I),(II)\\
c_k^{\pi,2} &= c_{h+1-k}^{\mathcal{T}\circ\mathcal{S}(\pi),2}, \quad c_l^{\pi,2} = c_{3h+1-l}^{\mathcal{T}\circ\mathcal{S}(\pi),2} && (I),(III)\\
c_k^{\pi,2} &= c_{k+h}^{\pi,2} && (I),(II),(III).
\end{aligned} \tag{4.43}$$

The condition that needs to hold for each equality is always denoted behind each equality. The calculations to obtain these relations are similar to those provided in (4.20). For an illustrative comparison how these relations can help to significantly reduce the number of integrals one has to calculate in order to determine the vector of Hermite coefficients in the case $d = 2$, $h = 2$, see Table A.1 and Table A.2 in Section A.2. Note that each entry of the vector of Hermite coefficients is a four-dimensional integral. The tables mentioned above deal with determining the Hermite coefficients for each pattern π in the reversion group $\bar{\pi} = \{\pi, \mathcal{S}(\pi), \mathcal{T}(\pi), \mathcal{T}\circ\mathcal{S}(\pi)\}$ and the reduction possibilities, if one wants to classify the vector of Hermite coefficients for all of the patterns in this group. If any of the assumptions (II) or/and (III) is fulfilled, it is possible to reduce the number of four-dimensional integrals we have to calculate from 16 to 8 in general, and even to 4 (if Condition (II) or (III) holds) or 2 integrals (if Conditions (II) and (III) hold). In the special case that $\pi = \mathcal{T}\circ\mathcal{S}(\pi)$ and Conditions (II) and (III) hold, it is even possible to obtain all Hermite coefficients needed by considering only one integral. If we want to determine the vector of Hermite coefficients for all patterns in S_2 and if both Condition (II) and (III) are fulfilled it is possible to reduce the number of integrals we have to calculate from 24 to 3 and, therefore, obtain a huge reduction of the computational cost.

In the following, we study the vector $\left(\tilde{\alpha}^{(1)}_{\pi,2}, \tilde{\alpha}^{(2)}_{\pi,2}\right)^t$ for $\pi \in S_2$, where the entries are given by $\tilde{\alpha}^{(p)}_{\pi,2} = \alpha^{\pi,2,\star}_{1+(p-1)h} + \alpha^{\pi,2,\star}_{2+(p-1)h}$, $p = 1,2$ and $\alpha^{\pi,2,\star} = \Sigma^{-1}_{2,2} c^{\pi,2}$. This vector is crucial to determine the limit variance in Theorem 4.52. By the symmetries of the vector $c^{\pi,2}$ within the reversion group $\bar{\pi}$ given in (4.43) it is possible to show that the number of calculations can be reduced. Under assumptions (II) and (III) it is even possible to show that knowing only three certain Hermite coefficients suffices to determine the limit variance in 4.52 for all patterns in S_2. We denote $\Sigma^{-1}_{2,h} = (g_{i,j})_{i,j=1,\dots,2h}$.
The limit variance in Theorem 4.52 is given by the formula

$$\sum_{p,q \in P^*} \tilde{\alpha}^{(p)}_{\pi,2} \tilde{\alpha}^{(q)}_{\pi,2} L_{p,q}. \tag{4.44}$$

To be able to handle the most general setting, we assume $P^* = \{1,2\}$. Otherwise the coefficients needed to determine the limit variance reduce corresponding to P^*.
Note that in general for each $\pi \in S_2$, we know due to (I) that for $p = 1,2$:

$$\begin{aligned}
\tilde{\alpha}^{(p)}_{\pi,2} &= \sum_{i=1}^{h} \alpha^{\pi,2}_{i+(p-1)h} \\
&= \sum_{i=1}^{h} \sum_{j=1}^{2h} g_{i+(p-1)h,j} c^{\pi,2}_j \\
&= \sum_{i=1}^{h} \sum_{j=1}^{2h} g_{i+(p-1)h,j} - c^{\mathcal{S}(\pi),2}_j \\
&= -\tilde{\alpha}^{(p)}_{\mathcal{S}(\pi),2}.
\end{aligned}$$

Therefore, according to (4.44), $\hat{p}_n(\pi)$ and $\hat{p}_n(\mathcal{S}(\pi))$ always have the same limit variance.
In the most general setting, in which only (I) is fulfilled, one has to calculate $c^{\pi,2}$ if $\#\bar{\pi} = 2$ and $c^{\pi,2}$ and $c^{\mathcal{T}(\pi),2}$ if $\#\bar{\pi} = 4$ in order to determine the limit variance for each pattern in the reversion group $\bar{\pi}$. This corresponds to a number of $2h$ or $4h$ of $2h$-dimensional integrals, respectively.
This number of calculations can be reduced again by checking whether conditions (II) and/or (III) are satisfied.
If Condition (II) holds, we know that $\Sigma^{-1}_{2,h}$ is persymmetric, such that $g_{i,j} = g_{2h+1-i,2h+1-j}$.
Then, we obtain for $p = 1,2$, that

$$\begin{aligned}
\tilde{\alpha}^{(p)}_{\mathcal{T}\circ\mathcal{S}(\pi),2} &= \sum_{j=1}^{2h} \sum_{i=1}^{h} g_{i+(p-1)h,j} c^{\mathcal{T}\circ\mathcal{S}(\pi),2}_j \\
&= \sum_{j=1}^{2h} \sum_{i=1}^{h} g_{2h+1-i-(p-1)h,2h+1-j} c^{\pi,2}_{2h+1-j}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{2h}\sum_{i=1}^{h} g_{(3-p)h+1-i,j}c_j^{\pi,2}\\
&= \sum_{j=1}^{2h}\sum_{i=1}^{h} g_{(2-p)h+i,j}c_j^{\pi,2}\\
&= \tilde{\alpha}_{\pi,2}^{(3-p)}.
\end{aligned}
$$

The limit variance of $\hat{p}_n(\pi)$ for each $\pi \in \bar{\pi}$ given in (4.44) is uniquely determined by the vector $\tilde{\alpha}_{\pi,2}^{(1)}, \tilde{\alpha}_{\pi,2}^{(2)}$, for one fixed π, although it does not necessarily have the same value for each pattern in this reversion group. This only occurs if $L_{1,1} = L_{2,2}$ (if $\tilde{\alpha}_{\pi,2}^{(1)} \neq \tilde{\alpha}_{\pi,2}^{(2)}$).
In the special case $\mathcal{T} \circ \mathcal{S}(\pi) = \pi$, we even have $\tilde{\alpha}_{\pi,2}^{(1)} = \tilde{\alpha}_{\pi,2}^{(2)}$.

If Condition (III) holds, due to the symmetry of the autocorrelation-, as well as the cross-correlation functions, we obtain for all $i = 1, \ldots, h$ that $g_{i+(p-1)h,j} = g_{h+1-i+(p-1)h,h+1-j}$ if $j = 1, \ldots, h$, and $g_{i+(p-1)h,j} = g_{h+1-i+(p-1)h,3h+1-j}$ if $j = h+1, \ldots, 2h$. We arrive at

$$
\begin{aligned}
\tilde{\alpha}_{\mathcal{T}\circ\mathcal{S}(\pi),2}^{(p)} &= \sum_{j=1}^{2h}\sum_{i=1}^{h} g_{i+(p-1)h,j}c_j^{\mathcal{T}\circ\mathcal{S}(\pi),2}\\
&= \sum_{j=1}^{h}\sum_{i=1}^{h} g_{i+(p-1)h,j}c_{h+1-j}^{\pi,2} + \sum_{j=h+1}^{2h}\sum_{i=1}^{h} g_{i+(p-1)h,j}c_{3h+1-j}^{\pi,2}\\
&= \sum_{j=1}^{h}\sum_{i=1}^{h} g_{h+1-i+(p-1)h,h+1-j}c_{h+1-j}^{\pi,2} + \sum_{j=h+1}^{2h}\sum_{i=1}^{h} g_{h+1-i+(p-1)h,3h+1-j}c_{3h+1-j}^{\pi,2}\\
&= \sum_{j=1}^{h}\sum_{i=1}^{h} g_{i+(p-1)h,h+1-j}c_{h+1-j}^{\pi,2} + \sum_{j=h+1}^{2h}\sum_{i=1}^{h} g_{i+(p-1)h,3h+1-j}c_{3h+1-j}^{\pi,2}\\
&= \sum_{j=1}^{h}\sum_{i=1}^{h} g_{i+(p-1)h,j}c_j^{\pi,2} + \sum_{j=h+1}^{2h}\sum_{i=1}^{h} g_{i+(p-1)h,j}c_j^{\pi,2}\\
&= \tilde{\alpha}_{\pi,2}^{(p)}.
\end{aligned}
$$

Therefore, the term in (4.44) attains the same value for each $\pi \in \bar{\pi}$.
Combining the last two results, we see that in the most special case where Conditions (II) and (III) hold, we obtain that the limit variance in (4.44) has the same value for each $\pi \in \bar{\pi}$. In addition, we also have $\tilde{\alpha}_{\pi,2}^{(1)} = \tilde{\alpha}_{\pi,2}^{(2)}$. Therefore, we can identify the limit variance to the term $\left(\tilde{\alpha}_{\pi,2}^{(1)}\right)^2 \sum_{p,q=1}^{2} L_{p,q}$.
Summing up these results, the advantages of the symmetries studied can be emphasized by considering two different settings seperately.
The first setting is to calculate the limit variance for one explicit pattern $\pi \in S_h$ only. Without any symmetry properties we need to determine all $2h$ entries of the vector of Hermite coefficients $c^{\pi,2}$, while if Conditions (II) and (III) are fulfilled, we can reduce the number of entries we need to determine to h if the pattern we are considering belong to a reversion group with 4 elements, or even to $\left\lfloor \frac{h+1}{2} \right\rfloor$ if $\pi = \mathcal{T} \circ \mathcal{S}(\pi)$.
However, the most efficient improvement is not achieved for the asymptotic variance for the estimator of one fixed pattern, but if we want to determine this variance for all estimators of the

patterns in the corresponding reversion class, or even in the entire permutation group S_h. In the most general setting, by only using symmetry property (I), it is already possible to reduce the number of calculations by a factor of 2. If additionally any of the conditions (II) or (III) is fulfilled, we get the asymptotic variance of the whole reversion group by only calculating it for one fixed pattern of this group. If we consider a reversion group with only two elements, the number of integrals further reduces due to the symmetry properties *within* the vector $c^{\pi,2}$ of Hermite coefficients, see (4.43). In the best (and most restrictive) case considered, namely Conditions (II) and (III) hold, and we are interested in the asymptotic variances for all estimators $\hat{p}_n(\pi)$ for $\pi \in S_2$, we only need to determine three entries of certain Hermite coefficients, namely $c_1^{(2,1,0),2}$, $c_1^{(2,0,1),2}$ and $c_2^{(2,0,1),2}$. Without exploiting these properties, each Hermite coefficient would have to be determined separately. Therefore, since $\#S_2 = 6$, we have a reduction from 24 to 3 four-dimensional integrals that need to be determined. In general, with $h \in \mathbb{N}$, we know that the number of reversion groups with 2 elements is given by $a_h = \left(\frac{h+1}{2}\right)!2^{\frac{h-1}{2}}$, for $h \in 2\mathbb{N} - 1$, and $a_h = a_{h+1}$. Without loss of generality, we assume that h is an odd number. Then, the number of integrals using the symmetries, divided by the number of integrals needed without optimization, is given by

$$\frac{\frac{h+1}{2}a_h + h\frac{(h+1)!-2a_h}{4}}{(h+1)!2h} = \frac{x!2^{x-1}}{4(2x)!(2x-1)} + \frac{1}{8} \begin{cases} = \frac{1}{4}, \ h = 1, \\ \approx \frac{1}{8}, h = 3, 5, 7, \dots \end{cases},$$

with $x := \frac{h+1}{2}$.

Hence, we can reduce the computational cost by $\frac{1}{8}$ if $h \neq 1$. Since the dimension of the integrals considered is given by $2h$, a reasonable computational time is only obtained in the case where h is small. We give an example for the case $h = 1$, which is the only one possible to derive closed formulas for $\tilde{\alpha}^{(p)}_{\pi,2}$, $p = 1, 2$.

Example 4.53 *We consider the case* $h = 2$ *and we observe that we are in a similar situation as in Example 4.21. Therefore, we obtain for* $\pi = (1, 0)$*:*

$$c_1^{\pi,2} = \mathbb{E}\left(Y_1^{(1)}\mathbf{1}_{\left\{Y_1^{(1)}\geq 0, Y_1^{(2)}\geq 0\right\}}\right) = \frac{\varphi(0)}{2}\left(1 + r^{(1,2)}(0)\right).$$

Since Condition (II) in (4.15) *holds due to the symmetry of the cross-correlation function in* 0*, we know that* $c_1^{\pi,2} = c_2^{\pi,2}$ *and hence derive*

$$\tilde{\alpha}^{(1)}_{\pi,2} = \frac{\varphi(0)}{2}$$

and $\tilde{\alpha}^{(1)}_{\pi,2} = \tilde{\alpha}^{(2)}_{\pi,2}$*. Therefore, the asymptotic variance in Theorem 4.52 in the case* $h = 1$ *is given by* $\frac{\varphi^2(0)}{4}\sum_{p,q\in P^*} L_{p,q}$*. Surprisingly it is independent of* $r^{(1,2)}(0)$*.*

In the case $h = 2$ one can obtain the limit variance of $\hat{p}_n(\pi) - p(\pi)$ by using the Matlab function *cHR1*. Note that we provide another Matlab function *HR1allpattern*, if one is interested in the limit variances for all permutations in S_2, since it is possible to significantly reduce the computational cost compared to using the *cHR1* for all $\#S_2 = 6$ pattern. Both algorithms are described in Section A.3.

However, it is possible to improve the estimator of ordinal pattern dependence for one certain pattern based on a Rao-Blackwellization as described in [60], p.1784ff. We have to add the constraint that $r^{(1,1)}(k) = r^{(2,2)}(k)$, $k = 0, 1, \ldots, h-1$, such that (III) in (4.16) holds. This condition assures that

$$p(\pi) = p(\mathcal{S}(\pi)) = p(\mathcal{T}(\pi)) = p(\mathcal{T} \circ \mathcal{S}(\pi)).$$

We define

$$\hat{p}_{n,\text{imp}}(\pi) := \frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} \hat{p}_n(\pi)$$

as improved estimator for $p(\pi)$. We now benefit from the considerations in (4.14) - (4.17) and Section 4.3.1 concerning the Hermite rank and the symmetry properties for the asymptotic variance of the estimator of p.
The Hermite rank of this estimator is an immediate consequence of the proof of Lemma 4.25.

Lemma 4.54 *The Hermite rank of*

$$f(Y_{j,h}) = \frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{j+1}^{(1)},\ldots,Y_{j+h}^{(1)}\right) = \tilde{\Pi}\left(Y_{j+1}^{(2)},\ldots,Y_{j+h}^{(2)}\right) = \pi\right\}} - p(\pi)$$

with respect to $\Sigma_{2,h}$ is equal to 2.

Hence, we can state the following result as an immediate applications of Corollary 3.11 and Theorem 4 of [2].

Theorem 4.55 *The following statements are true:*

(i) Under the assumptions in ***(L2s)*** *in Overview 4.24, it holds for each $\pi \in S_h$, that*

$$n^{1-2d^*} (C_2)^{-\frac{1}{2}} \left(\hat{p}_{n,\text{imp}}(\pi) - p(\pi)\right) \xrightarrow{\mathcal{D}} \sum_{p,q \in P^*} \tilde{\alpha}^{(p,q)} Z_{2,d^*+1/2}^{(p,q)}(1),$$

with $Z_{2,d^*+1/2}^{(p,q)}(1)$ as given in Theorem 3.10 and $C_2 := \frac{1}{2d^*(4d^*-1)}$ being a normalizing constant.
We have $\tilde{\alpha}^{(p,q)} := \sum_{i,k=1}^{h} \alpha_{i,k}^{(p,q)}$, where $\alpha_{i,k}^{(p,q)} = \alpha_{i+(p-1)h, k+(q-1)h}$ for each $p, q \in P^*$ and $i, k = 1, \ldots, h$ and $(\alpha_{i,k})_{1 \leq i,k \leq dh} = \Sigma_{2,h}^{-1} C \Sigma_{2,h}^{-1}$ where

$$C = (c_{i,k})_{1 \leq i,k \leq 2h} = \mathbb{E}\left(Y_{1,h}\left(\frac{1}{\#\bar{\pi}} \sum_{\pi \in \bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{j+1}^{(1)},\ldots,Y_{j+h}^{(1)}\right) = \tilde{\Pi}\left(Y_{j+1}^{(2)},\ldots,Y_{j+h}^{(2)}\right) = \pi\right\}} - p(\pi)\right) Y_{1,h}^{t}\right)$$

denotes the matrix of second order Hermite coefficients.

*(ii) Under the assumptions in **(S)** in Overview 4.24, it holds that*

$$n^{\frac{1}{2}}\left(\hat{p}_{n,\mathrm{imp}}(\pi)-p(\pi)\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\sigma^2\right),$$

with

$$\sigma^2=\sum_{k=-\infty}^{\infty}\mathbb{E}\Bigg[\left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_1^{(2)},\ldots,Y_h^{(2)}\right)=\pi\right\}}-p(\pi)\right)$$
$$\times\left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{1+k}^{(1)},\ldots,Y_{h+k}^{(1)}\right)=\tilde{\Pi}\left(Y_{1+k}^{(2)},\ldots,Y_{h+k}^{(2)}\right)=\pi\right\}}-p(\pi)\right)\Bigg].$$

Note that the asymptotic variance in (i) can easily be deduced from the considerations concerning the asymptotic variance in Section 4.3.1 of $\hat{p}_n$, since

$$\hat{p}_n=2\sum_{\pi\in S_{[2,h]}}\hat{p}_{n,\mathrm{imp}}(\pi)+4\sum_{\pi\in S_{[4,h]}}\hat{p}_{n,\mathrm{imp}}(\pi).$$

Using this relation, the expression of the second order Hermite coefficients for the considered estimator can be obtained and are given in Table 4.1. For an overview of the second order Hermite coefficients for $h=2$ see Example 4.30. Concerning the asymptotic variance determined by $\tilde{\alpha}^{(p,q)}$ as given in Theorem 4.55, using relation (4.22) shows that the considerations in Section 4.3.1 are sufficient to fully determine the asymptotic variance of the estimator $\hat{p}_{n,\mathrm{imp}}(\pi)$, as well. We give a brief example in the case $h=1$, which is directly deduced from Example 4.33.

Example 4.56 *In the case $h=1$, the factors $\tilde{\alpha}^{(p,q)}$, $p,q=1,2$, of the asymptotic variance of the estimator $\hat{p}_{n,\mathrm{imp}}$ in Theorem 4.55 (i) are given by*

$$\tilde{\alpha}_{\pi,2}^{(1,1)}=\tilde{\alpha}_{\pi,2}^{(2,2)}=\frac{-\varphi^2(0)r^{(1,2)}(0)}{\left(1-\left(r^{(1,2)}(0)\right)^2\right)^{1/2}}$$

and

$$\tilde{\alpha}_{\pi,2}^{(1,2)}=\tilde{\alpha}_{\pi,2}^{(2,1)}=\frac{\varphi^2(0)}{\left(1-\left(r^{(1,2)}(0)\right)^2\right)^{1/2}},$$

for $\pi\in S_1$.

In the case $h=2$, the factors $\tilde{\alpha}^{(p,q)}$, $p,q=1,2$, can be obtained by using the Matlab function *cHR2*, which is described in Section A.3. In order to get $\tilde{\alpha}^{(p,q)}$, $p,q=1,2$, one follows the considerations in (4.24) to (4.26). Depending on the permutation $\pi\in S_2$, one needs to focus on the first sum on the right-hand side in (4.24) to (4.26), if $\#\bar{\pi}=2$, or on the second one if $\#\bar{\pi}=4$. In both cases it is important to divide each sum by 2 or 4, respectively, to get the appropriate factors for the estimator considered here. One can see that only if neither Condition (II) given in (4.15) nor Condition (III) in (4.16) hold, one has to use (4.24) and, therefore, the Matlab function *cHR2* needs to be applied for e.g. $\pi=(2,0,1)$ as well as for $\pi=(0,2,1)$. In all other cases one run of the Matlab function for $\pi=(2,0,1)$ is sufficient to determine $\tilde{\alpha}^{(p,q)}$, $p,q=1,2$, from the result.

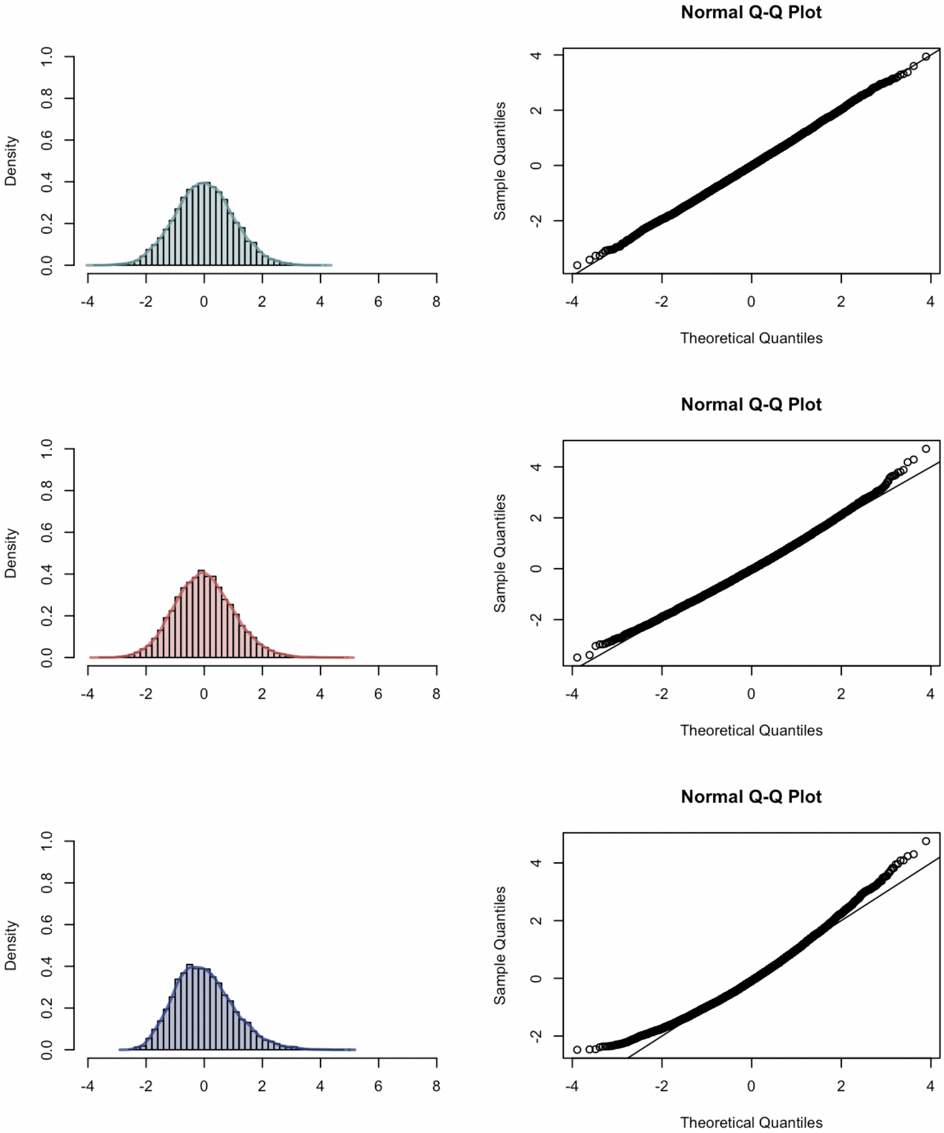

Figure 4.16: Histogram, kernel density estimation and Q-Q plot of suitably normalized $\hat{p}_n((2,1,0)) - p((2,1,0))$ for different Hurst parameters: $H = 0.7$ (top), $H = 0.8$ (middle), $H = 0.9$ (bottom).

Finally we complement the simulation study in Section 4.3.5, such that all assumptions of this section hold as well. By choosing the estimator of ordinal pattern dependence for one fixed pattern $\hat{p}_n(\pi)$, we provide an example of a function with Hermite rank 1. Note that we omit the improved estimator $\hat{p}_{n,\text{imp}}(\pi)$ here, as simulation results for functionals with Hermite rank 2 were already presented in Section 4.3.5. Based on the results in Theorem 4.52, we expect asymptotic normality for all Hurst parameters H considered here. The simulation results of the asymptotic distribution are given in Figure 4.16. We observe the density of the standard normal distribution for each parameter H, so Theorem 4.52 is validated. Concerning the asymptotic variance of the limit distributions as given in Theorem 4.52, we can use the Matlab function *cHR1*, see Section A.3, to obtain the quantities displayed in Table 4.4. Note that due to the symmetry property (IV) in (4.17), we have $\tilde{\alpha}^{(1)} = \tilde{\alpha}^{(2)}$. Concerning the deviation between the theoretical values

	$\tilde{\alpha}^{(1)}$	Limit variance	Simulated limit variance
$H = 0.7$	0.1557	0.0217	0.0218
$H = 0.8$	0.1645	0.0416	0.0425
$H = 0.9$	0.1753	0.0708	0.0624

Table 4.4: Theoretical values of $\tilde{\alpha}^{(1)}$ and the limit variance as well as the simulated limit variance of suitably normalized $\hat{p}_n((2,1,0)) - p((2,1,0))$ for $H = 0.7$, $H = 0.8$ and $H = 0.9$.

of the limit variance and the results in the simulations, it is clearly observable that the larger the Hurst parameter gets, the larger the deviation becomes. The reason for this is, as already mentioned concerning the limit variances of the estimator $\hat{p}_n$, that the convergence rate of the cross-correlations in (3.2) gets slower the larger the Hurst parameter is. Therefore, the simulated results get more biased the larger the Hurst parameter gets. However, in the case of Hermite rank 1, we still obtain a faster convergence rate as in the case of Hermite rank 2, see Theorem 3.3 and Theorem 3.10. Therefore, we obtain more stable simulation results for the limit variances here.

4.4.2 Estimating the Hurst parameters of vector fractional Gaussian noise based on ordinal pattern analysis

As proposed in [60], Section 3, in the univariate case, it is possible to estimate the Hurst parameter of fractional Gaussian noise based on the improved estimator of ordinal pattern probabilities $\hat{q}_{n,Y^{(p)},\text{imp}}(\pi)$, $p = 1, 2$, as introduced in Section 4.2. This estimator is called *zero-crossing estimator* of the Hurst parameter. Note that the Hurst parameter H_p of the respective long-range dependent time series $Y_j^{(p)}$, $p = 1, 2$, is related to the long-range dependence parameter d_p in Definition 2.13 by $H_p = d_p + \frac{1}{2}$. In [60], Corollary 5, it is proved that this estimator of the Hurst parameter is strongly consistent and asymptotically unbiased. Asymptotic normality is also derived for $H < \frac{3}{4}$, the transformed short-range dependence case. This result was complemented in [10], Theorem 4.1, for $H > \frac{3}{4}$, that is the long-range dependent case, with a Rosenblatt distributed limit random variable.

We want to give an extension to the estimation of the Hurst parameter of each (cross-)correlation-function for multivariate vector fractional Gaussian noise. This is the increment process of vector fractional Brownian motion and was introduced in Definition 2.23. Note that we restrict ourselves to the bivariate case, but the results can be extended straightforward. We start with a brief description of the setting used in [60], Section 3, to which we adapt our multivariate generalization.

As their method needs closed formulas of orthant probabilities of the multivariate normal distribution, they restricted themselves to the case $h = 2$ and considered the probability of changes of the signs of the respective increments of the one dimensional underlying process. This probability is equal to the probability of changes in the *up-and-down* behaviour of the process $\left(X_j^{(p)}\right)_{j\in\mathbb{Z}}$. In mathematical terms this probability can be described for $p = 1, 2$ as

$$\begin{aligned}\theta^{(p,p)} &:= \mathbb{P}\left(W^{(p,p)} = 1\right)\\ &= 2\mathbb{P}\left(Y_{i+1}^{(p)} \geq 0, Y_{i+2}^{(p)} \leq 0\right)\\ &= \frac{1}{2} - \frac{1}{\pi}\arcsin\left(r^{(p,p)}(1)\right),\end{aligned}$$

with $W^{(p,p)} := \mathbf{1}_{\{Y_1^{(p)}\geq 0, Y_2^{(p)}\leq 0\}} + \mathbf{1}_{\{Y_1^{(p)}\leq 0, Y_2^{(p)}\geq 0\}}$.
In the context of a bivariate process, in order to estimate the long-range dependence parameter of the cross-correlation function, one needs to slightly modify this approach. We consider the *up-and-down* behaviour between the processes $\left(X_j^{(p)}\right)_{j\in\mathbb{Z}}$ and $\left(X_j^{(q)}\right)_{j\in\mathbb{Z}}$, described by

$$\begin{aligned}\theta^{(p,q)} &:= \mathbb{P}\left(W^{(p,q)} = 1\right)\\ &= 2\mathbb{P}\left(Y_{i+1}^{(p)} \geq 0, Y_{i+2}^{(q)} \leq 0\right) = \frac{1}{2} - \frac{1}{\pi}\arcsin\left(r^{(p,q)}(1)\right),\end{aligned}$$

with $W^{(p,q)} := \mathbf{1}_{\{Y_1^{(p)}\geq 0, Y_2^{(q)}\leq 0\}} + \mathbf{1}_{\{Y_1^{(p)}\leq 0, Y_2^{(q)}\geq 0\}}$.
As estimator of $\theta^{(p,q)}$, $p, q = 1, 2$, in the case $p = q$ it is natural to consider:

$$\hat{\theta}_n^{(p,p)} := 4\hat{q}_{n,Y^{(p)},\mathrm{imp}}(\pi)$$

for $\pi \in \{(2,0,1),(1,0,2),(1,2,0),(0,2,1)\}$, because these ordinal patterns express a change of the sign in two consecutive increments.
In the case $p \neq q$, we need to define a different estimator. We obtain

$$\begin{aligned}\mathbb{P}\left(W^{(p,q)} = 1\right) &= \mathbb{P}\left(\mathbf{1}_{\{Y_1^{(1)}\geq 0, Y_2^{(2)}\leq 0\}} + \mathbf{1}_{\{Y_1^{(1)}\leq 0, Y_2^{(2)}\geq 0\}} = 1\right)\\ &= \mathbb{P}\left(Y_1^{(1)} \geq 0, Y_2^{(2)} \leq 0\right) + \mathbb{P}\left(Y_1^{(1)} \leq 0, Y_2^{(2)} \geq 0\right)\\ &= \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)}\right) = (1,0), \tilde{\Pi}\left(Y_2^{(2)}\right) = (0,1)\right)\\ &\qquad + \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)}\right) = (0,1), \tilde{\Pi}\left(Y_2^{(2)}\right) = (1,0)\right)\\ &= \sum_{\pi\in\bar{\pi}} \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)}\right) = \pi, \tilde{\Pi}\left(Y_2^{(2)}\right) = \mathcal{S}(\pi)\right)\\ &= \sum_{\pi\in\bar{\pi}} \mathbb{P}\left(\tilde{\Pi}\left(Y_1^{(1)}\right) = \pi, \tilde{\Pi}\left(-Y_2^{(2)}\right) = \pi\right),\end{aligned}$$

for $\pi \in \{(0,1),(1,0)\} = S_1$. Therefore, it is sufficient to restrict ourselves to the case $h = 1$ in a modified ordinal pattern dependence setting, since we are now facing the challenge that we are not comparing coincident patterns, but space-reflected patterns. Hence, we have to consider a maximal distance between the patterns. This corresponds to negative dependence between the increment variables and we exploit the possibility to change the sign of the second increment variable in order to measure ordinal pattern dependence, as it was proposed at the beginning of Section 4.3. Additionally, we are still able to plug in the well-known improved estimator of ordinal pattern dependence for one certain pattern in order to estimate $\hat{\theta}^{(p,q)}$. Compared to the univariate case, note that

$$\mathbb{P}\left(Y_j^{(1)} \geq 0, Y_{j+1}^{(2)} \leq 0\right) + \mathbb{P}\left(Y_j^{(1)} \leq 0, Y_{j+1}^{(2)} \geq 0\right) = \mathbb{P}\left(\tilde{\Pi}\left(Y_j^{(1)}, Y_{j+1}^{(2)}\right) \in \bar{\pi}\right)$$

for $\pi \in \{(2,0,1),(1,0,2),(1,2,0),(0,2,1)\}$. However, this case is not helpful here, since we do not have any closed formulas for the limit variance in the case $h = 2$.
Hence, we define the estimator of $\theta^{(p,q)}$ for $p \neq q$ by

$$\hat{\theta}_n^{(p,q)} := \hat{p}_n = 2\hat{p}_n(\pi)$$

for $\pi \in \{(0,1),(1,0)\}$. Note that $\hat{p}_n$ and $\hat{p}_n(\pi)$ are applied to a bivariate Gaussian process for $h = 1$ given by $\left(\left(Y_j^{(1)}, -Y_{j+1}^{(2)}\right)\right)_{j \in \mathbb{Z}}$.

Since we are considering vector valued fractional Gaussian noise, we know by Definition 2.23 that the cross-correlation function of $\left(Y_j^{(1)}, Y_j^{(2)}\right)_{j \in \mathbb{Z}}$ is given by

$$r^{(p,q)}(k) = \frac{1}{2}\left(w_{p,q}(k+1)|k+1|^{(H_p+H_q)} + w_{p,q}(k-1)|k-1|^{(H_p+H_q)} - 2w_{p,q}(k)|k|^{(H_p+H_q)}\right),$$

with $w_{p,q}(k) = \rho_{p,q} - \eta_{p,q} sign(k)$, where $\rho_{p,q} = \mathrm{Cor}\left(B_H^{(p)}(1), B_H^{(q)}(1)\right)$. The random variable $B_H(1)$ denotes the corresponding vector fractional Brownian motion evaluated in 1. We have $\eta_{p,q} \in \mathbb{R}$ and $\eta_{p,q} = -\eta_{q,p}$. For details on $\eta_{p,q}$, see [54], Proposition 9.3.19 and Section 9.4.3. For technical reasons, we assume in the following, that $\rho_{p,q} \pm \eta_{p,q} \neq \cos\left(\pi \frac{H_p+H_q-1}{2}\right)$.
Note that if $p = q$, $\rho_{p,p} = 1$ and $\eta_{p,p} = 0$. Recall from (2.4) that

$$r^{(p,q)}(k) \simeq \frac{1}{2}\left(\rho_{p,q} - \eta_{p,q} sign(k)\right)\left(H_p + H_q\right)\left(H_p + H_q - 1\right)|k|^{H_p+H_q-2}. \tag{4.45}$$

Therefore, we obtain for all $p, q = 1, 2$, that

$$r^{(p,q)}(1) = \left(2^{H_p+H_q-1} - 1\right)\left(\rho_{p,q} - \eta_{p,q}\right).$$

Hence, we can write

$$\theta^{(p,q)}\left(\frac{H_p + H_q}{2}\right) = \frac{1}{2} - \frac{1}{\pi}\arcsin\left(\left(2^{2\frac{H_p+H_q}{2}-1} - 1\right)\left(\rho_{p,q} - \eta_{p,q}\right)\right)$$

for $p, q = 1, 2$.
Now we determine a function $g^{p,q}$, such that $g^{p,q}\left(\theta^{(p,q)}\left(\frac{H_p+H_q}{2}\right)\right) = \frac{H_p+H_q}{2}$, to be able to apply the delta method later on. For $x \in [0,1]$ we obtain

$$g^{p,q}(x) := \max\left\{0, \frac{1}{2}\left(\log_2\left(\frac{\sin\left(\pi\left(\frac{1}{2} - x\right)\right)}{\rho_{p,q} - \eta_{p,q}} + 1\right) + 1\right)\right\}$$

with

$$\left(g^{p,q}\right)'(x) = -\frac{\pi\sin(\pi x)}{\log(4)\left(\left(\rho_{p,q} - \eta_{p,q}\right) + \cos(\pi x)\right)}.$$

Note that for all $p, q = 1, 2$ we have $\theta^{(p,q)}\left(\frac{H_p+H_q}{2}\right) \in (0,1)$, since $H_p, H_q \in \left(\frac{1}{2}, 1\right)$, with $H_p + H_q \neq 1$. Therefore, $(g^{p,q})'\left(\frac{H_p+H_q}{2}\right)$ exists, since we assumed $\rho_{p,q} \pm \eta_{p,q} \neq \cos\left(\pi \frac{H_p+H_q-1}{2}\right)$ and $(g^{p,q})'\left(\frac{H_p+H_q}{2}\right)$ is not equal to zero, because $\sin\left(\pi \frac{H_p+H_q-1}{2}\right) \neq 0$ for $H_p, H_q \in \left(\frac{1}{2}, 1\right)$. We define

$$\hat{H}_n^{(p,q)} := \widehat{\frac{H_p + H_q}{2}} = g\left(\hat{\theta}_n^{(p,q)}\right)$$

and arrive at the following result.

Theorem 4.57 *Let $(Y_j)_{j\in\mathbb{Z}}$ be a bivariate vector fractional Gaussian noise that fulfills **(L1)** in Overview 4.24. Therefore, $H_p \in \left(\frac{1}{2}, 1\right)$, for $p = 1, 2$. Further, let $Z^{(p,q)}_{2,\frac{H_p+H_q}{2}}$, $p, q = 1, 2$, be the random variables as given in Theorem 4.55 but scaled to unit variance. They exhibit the covariance matrix*

$$\mathbb{E}\left(\operatorname{vec}\left(Z^{(p,q)}_{2,\frac{H_p+H_q}{2}},\ p,q=1,2\right)\operatorname{vec}\left(Z^{(p,q)}_{2,\frac{H_p+H_q}{2}},\ p,q=1,2\right)^t\right) = \left(I_{d^2} + K_{d^2}\right) D_L^{-1} L \otimes L, \quad (4.46)$$

with K_{d^2}, L as given in Lemma 3.5 and D_L is a diagonal matrix

$$D_L = \operatorname{diag}\left(2L_{1,1}^2, L_{1,1}L_{2,2} + L_{1,2}L_{2,1}, L_{1,1}L_{2,2} + L_{1,2}L_{2,1}, 2L_{2,2}^2\right).$$

assuring that the diagonal entries of the covariance matrix in (4.46) are all equal to 1. Then, the following statements are true for $p, q = 1, 2$:

(i) If $p \neq q$, $H_p > H_q$ and $H_p \in \left(\frac{3}{4}, 1\right)$, then

$$n^{2-2H_p}\sqrt{\frac{4H_p - 3}{8\left(H_p\right)^2\left(2H_p - 1\right)}}\left(\hat{H}_n^{(p,q)} - \frac{H_p + H_q}{2}\right)$$
$$\xrightarrow{\mathcal{D}} -\frac{1}{2}\frac{\sin\left(\pi\theta^{(p,q)}\left(\frac{H_p+H_q}{2}\right)\right)}{\log(4)\left(\left(\rho_{p,q} - \eta_{p,q}\right) + \cos\left(\pi\theta^{(p,q)}\left(\frac{H_p+H_q}{2}\right)\right)\right)}$$
$$\times \frac{2^{H_p+H_q-1} - 1}{\left(\left(2 - 2^{H_p+H_q-1}\right)2^{H_p+H_q-1}\right)^{1/2}} Z^{(p,p)}_{2,H_p}(1).$$

(ii) If $p \neq q$, $H_p = H_q$ and $H_p \in \left(\frac{3}{4}, 1\right)$, then

$$n^{2-2H_p}\sqrt{\frac{4H_p - 3}{4\left(H_p\right)^2\left(2H_p - 1\right)}}\left(\hat{H}_n^{(p,q)} - \frac{H_p + H_q}{2}\right)$$
$$\xrightarrow{\mathcal{D}} \frac{-\sin\left(\pi\theta^{(p,q)}\left(\frac{H_p+H_q}{2}\right)\right)}{2\log(4)\left(\left(\rho_{p,q} - \eta_{p,q}\right) + \cos(\pi\theta^{(p,q)}\left(\frac{H_p+H_q}{2}\right))\right)}\frac{1}{\left(\left(2 - 2^{2H_p-1}\right)2^{2H_p-1}\right)^{1/2}}$$
$$\times \Bigg(2^{1/2}\left(2^{2H_p-1} - 1\right)\left(\rho_{1,2} - \eta_{1,2}\right)\left(Z^{(1,1)}_{2,H_p}(1) + Z^{(2,2)}_{2,H_p}(1)\right)$$
$$- 2\left(1 + \rho_{1,2}^2 - \eta_{1,2}^2\right)^{1/2} Z^{(1,2)}_{2,H_p}(1)\Bigg).$$

(iii) If $p \neq q$ and $H_p, H_q \in \left(\frac{1}{2}, \frac{3}{4}\right)$, then

$$n^{\frac{1}{2}}\left(\hat{H}_n^{(p,q)} - \frac{H_p + H_q}{2}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 4\sigma^2\left(\left(g^{p,q}\right)'\left(\theta^{(p,q)}\left(\frac{H_p + H_q}{2}\right)\right)\right)^2\right),$$

with

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}\Bigg[\left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)}\right)=\tilde{\Pi}\left(-Y_2^{(2)}\right)=\pi\right\}} - p(\pi)\right)$$
$$\times\left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{1+k}^{(1)}\right)=\tilde{\Pi}\left(-Y_{2+k}^{(2)}\right)=\pi\right\}} - p(\pi)\right)\Bigg], \qquad \pi \in S_1.$$

(iv) If $p = q$, $p = 1, 2$ and $H_p \in \left(\frac{3}{4}, 1\right)$, then

$$n^{2H_p-2}\sqrt{\frac{4H_p - 3}{4H_p^2\left(2H_p - 1\right)}}\left(\hat{H}_n^{(p,p)} - H_p\right)$$
$$\xrightarrow{\mathcal{D}} Z_{2,H_p}^{(p,p)}(1)\left(\frac{1}{\log(2)}\tan\left(\frac{\pi\theta^{(p,p)}(H_p)}{2}\right)\sqrt{2^{2-2H_p} - 1}\right).$$

Remark 4.58 *Note that the case if $p = q$ and $H_p \in \left(\frac{1}{2}, \frac{3}{4}\right)$ is given in [60], Corollary 11: In this case, we have*

$$n^{\frac{1}{2}}\left(\hat{H}_n^{(p,p)} - H_p\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \left(\frac{\pi}{2\log 2}\tan\left(\frac{\pi\theta^{(p,p)}(H_p)}{2}\right)\right)^2\sigma^2\right),$$

with

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}\left[\left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},Y_2^{(1)}\right)=\pi\right\}} - p(\pi)\right)\left(\frac{1}{\#\bar{\pi}}\sum_{\pi\in\bar{\pi}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{1+k}^{(1)},Y_{2+k}^{(1)}\right)=\pi\right\}} - p(\pi)\right)\right].$$

Proof. We can deduce from Theorem 4.55 that for $\pi = (1, 0)$

$$n^{1-2d^*}\left(C_2\right)^{-\frac{1}{2}} 2\left(\hat{p}_{n,\text{imp}}(\pi) - p(\pi)\right) \tag{4.47}$$
$$\xrightarrow{\mathcal{D}} \sum_{p^*,q^*\in P^*} 2\tilde{\alpha}^{(p^*,q^*)}\left(L_{p^*p^*}L_{q^*q^*} + L_{p^*q^*}L_{q^*p^*}\right)^{\frac{1}{2}}(-1)^{\delta_{|p^*-q^*|,1}} Z_{2,H_{p^*}}^{(p^*,q^*)}(1), \tag{4.48}$$

with $Z_{2,H_{p^*}}^{(p^*,q^*)}(1)$, such that $\text{Var}\left(Z_{2,d^*+1/2}^{(p^*,q^*)}(1)\right) = 1$ for all $p^*, q^* \in P^*$. The unit variance is derived from Theorem 4.26 by factoring out $\left(L_{p^*p^*}L_{q^*q^*} + L_{p^*q^*}L_{q^*p^*}\right)^{\frac{1}{2}}$, which represents square root of the diagonal entry for p^*, q^* in the covariance matrix of the sample covariances in (3.27). Further, $C_2 := \frac{1}{2d^*(4d^*-1)}$ is a normalizing constant. Since we are considering $\left(\left(Y_j^{(1)}, -Y_{j+1}^{(2)}\right)^t\right)_{j\in\mathbb{Z}}$ as underlying bivariate Gaussian process, the sign of the limit random variables depends on whether $p^* = q^*$ or $p^* \neq q^*$. In the latter case, the sample cross-correlations in (3.42) for the process considered here can be expressed as those of the process $\left(\left(Y_j^{(1)}, Y_j^{(2)}\right)^t\right)_{j\in\mathbb{Z}}$, but with a negative sign. The second order Hermite coefficients are given in Example 4.33. However, we have to be careful, since $\mathbb{E}\left(-Y_j^{(1)}Y_{j+1}^{(2)}\right) = -r^{(1,2)}(1)$. Using the relation in (4.45), we can further determine

$L_{1,1}, L_{1,2}, L_{2,1}$ and $L_{2,2}$. We would actually have to change the sign in $L_{1,2}$ and $L_{2,1}$ as well, but since exactly these two values are multiplicated in the limit distribution, we can omit this operation.
We obtain

$$\tilde{\alpha}^{(1,1)} = \tilde{\alpha}^{(2,2)} = \frac{\varphi^2(0) r^{(1,2)}(1)}{\left(1 - \left(r^{(1,2)}(1)\right)^2\right)^{1/2}}$$
$$\tilde{\alpha}^{(1,2)} = \tilde{\alpha}^{(2,1)} = \frac{\varphi^2(0)}{\left(1 - \left(r^{(1,2)}(1)\right)^2\right)^{1/2}}$$

and for $p, q = 1, 2$ we have

$$L_{p,q} = \frac{1}{2} \left(\rho_{p,q} - \eta_{p,q}\right) \left(H_p + H_q\right) \left(H_p + H_q - 1\right).$$

Concerning (i), we turn to $\max\{H_1, H_2\} = H_p$ for either $p = 1$ or $p = 2$. Plugging the values into the formula in (4.48) we get

$$n^{2-2H_p} \sqrt{\frac{4H_p - 3}{8 \left(H_p\right)^2 \left(2H_p - 1\right)}} \left(\hat{\theta}_n^{(p,q)} - \theta^{(p,q)} \left(\frac{H_p + H_q}{2}\right)\right)$$
$$\xrightarrow{\mathcal{D}} \varphi^2(0) \frac{2^{H_p + H_q - 1} - 1}{\left(\left(2 - 2^{H_p + H_q - 1}\right) 2^{H_p + H_q - 1}\right)^{1/2}} Z_{2,H_p}^{(p,p)}(1).$$

Now we can apply the delta method since $\hat{H}_n^{(p,q)} = g\left(\theta_n^{(p,q)}\right)$ and $\frac{H_p + H_q}{2} = g\left(\theta^{(p,q)}\left(\frac{H_p + H_q}{2}\right)\right)$.
Analogously, we obtain the result in case $H_p = H_q$ as described in (ii). We have

$$\begin{aligned}
&n^{2-2H_p} \sqrt{\frac{4H_p - 3}{4 \left(H_p\right)^2 \left(2H_p - 1\right)}} \left(\hat{\theta}_n^{(p,q)} - \theta^{(p,q)} \left(H_p\right)\right) \\
&\xrightarrow{\mathcal{D}} \sum_{p,q=1}^{2} \tilde{\alpha}^{(p,q)} \left(1 + \rho_{p,q}^2 - \eta_{p,q}^2\right)^{1/2} Z_{2,H_p}^{(p,q)}(1) \\
&= \tilde{\alpha}^{(1,2)} \left(2^{1/2} r^{(1,2)}(1) \left(Z_{2,H_p}^{(1,1)}(1) + Z_{2,H_p}^{(2,2)}(1)\right) + 2 \left(1 + \rho_{1,2}^2 - \eta_{1,2}^2\right)^{1/2} Z_{2,H_p}^{(1,2)}(1)\right) \\
&= \frac{\varphi^2(0)}{\left(\left(2 - 2^{2H_p - 1}\right) 2^{2H_p - 1}\right)^{1/2}} \Bigg(2^{1/2} \left(2^{2H_p - 1} - 1\right) \left(\rho_{1,2} - \eta_{1,2}\right) \left(Z_{2,H_p}^{(1,1)}(1) + Z_{2,H_p}^{(2,2)}(1)\right) \\
&\qquad - 2 \left(1 + \rho_{1,2}^2 - \eta_{1,2}^2\right)^{1/2} Z_{2,H_p}^{(1,2)}(1)\Bigg),
\end{aligned}$$

and by applying the delta method, we have finally shown the result.
Part (iii) of the theorem is an immediate consequence of Theorem 4.55 (ii), with modified entries of the limit variance concerning the underlying bivariate process of the estimator $\hat{p}_{n,\text{imp}}(\pi)$ and the application of the delta method as described before.
In part (iv), we focus on the case $p = q$ and, therefore, on the univariate version. Consequently, it is sufficient to consider the respective marginal process $\left(Y_j^{(p)}\right)_{j \in \mathbb{Z}}$, for $p = 1, 2$. Hence, the estimation of the corresponding Hurst parameter H_p does not rely on an estimator in the context of ordinal pattern dependence but can be derived in terms of the estimation of ordinal pattern

probabilities, as already mentioned above. This is considered in [10], Theorem 4.1 and is briefly introduced in the following. We use

$$\hat{\theta}_n^{(p,p)} := 4\hat{q}_{n,Y^{(p)},\mathrm{imp}}(\pi)$$

as estimator of $\theta^{(p,p)}$, $p = 1, 2$, for $\pi \in \{(2,0,1),(1,0,2),(1,2,0),(0,2,1)\}$. Therefore, the limit theorem can be derived by Theorem 4.18. We obtain

$$n^{2H_p-2}\sqrt{\frac{4H_p-3}{4H_p^2\left(2H_p-1\right)}}\left(\hat{\theta}_n^{(p,p)}-\theta^{(p,p)}\left(H_p\right)\right)\xrightarrow{\mathcal{D}} Z_{2,H_p}^{(p,p)}(1)\left(-4\varphi^2(0)\sqrt{2^{2-2H_p}-1}\right), \qquad (4.49)$$

since, following Example 4.21, we have $\sum_{k,l=1}^{2}\alpha_{l,k}^{\pi} = -\varphi^2(0)\sqrt{2^{2-2H_p}-1}$. It is possible to simplify $g^{(p,p)}$ and, hence, $\left(g^{(p,p)}\right)'$ for $x \in \left(0, \frac{2}{3}\right)$ using trigonometrical relations, since we have $\rho_{p,p} = 1$ and $\eta_{p,p} = 0$. For details see [10], proof of Theorem 4.1. We obtain $g(x) := \max\left\{0, \log_2\left(\cos\left(\frac{\pi x}{2}\right)\right)+1\right\}$ and $g'(x) = -\frac{\pi}{2\log 2}\tan\left(\frac{\pi x}{2}\right)$ and note that these functions coincide with $g^{(p,p)}(x)$ and $\left(g^{(p,p)}\right)'(x)$ for $x \in \left(0, \frac{1}{2}\right)$, which is the interval in which $\theta^{(p,p)}\left(H_p\right)$ takes values for $H_p \in \left(\frac{1}{2}, 1\right)$. We get sharper bounds in the univariate case, since $\rho_{p,p} = 1$ and $\eta_{p,p} = 0$. Therefore, we apply the delta method with these simplified functions and hence get the result in a tighter formula as in the bivariate case. □

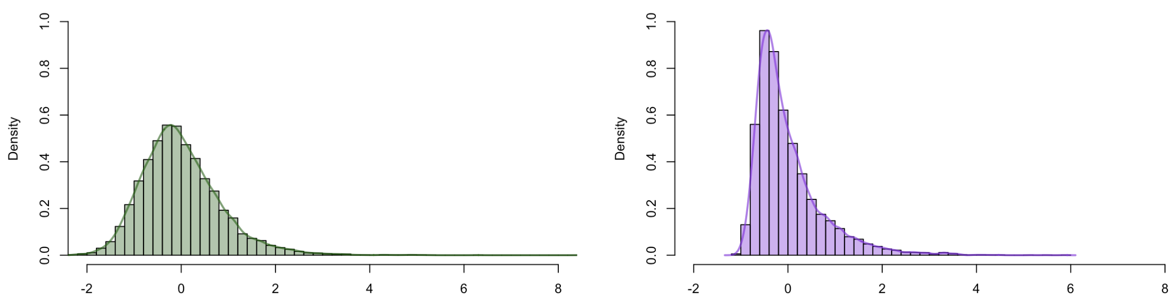

Figure 4.17: Histogram and density of the Hurst parameter estimator for $H = 0.8$ (green) and $H = 0.9$ (violet).

We close this section with the simulation results derived in [10], Section 6, that confirm the statement in Theorem 4.57 (iv). We simulate $N = 10\,000$ paths of fractional Gaussian noise created by the command "simFGN0" from the R-Package "longmemo", see [34] with sample size $n = 10^6$ for different values of H. In Figure 4.17, the histograms and kernel density estimations of the estimator of the Hurst parameter are given, using the normalization we obtain from (4.49). We observe that the simulated densities coincide with the densities in Figure 4.9 that characterize the Rosenblatt distribution for $H = 0.8$ and $H = 0.9$, respectively.

Hence, the multivariate extension of the Hurst parameter estimation is an application of the theoretical considerations for $\hat{p}_n$ and $\hat{p}_n(\pi)$, in contrast to the univariate case, in which the Hurst parameter estimation only relies on estimators of ordinal pattern probabilities.

4.4.3 Asymptotics of the estimators of ordinal pattern dependence in case of a stationary time series

In Section 4.2 and Section 4.3, we have proved limit theorems for estimators in the context of ordinal pattern dependence using that the increment process $(Y_j)_{j\in\mathbb{Z}}$ fulfills a certain assumption in Overview 4.24. However, this assumption does not require the original bivariate Gaussian process $(X_j)_{j\in\mathbb{Z}}$, in which we actually compare the patterns, to be stationary or even long-range dependent. This is an additional assumption and hence a special case of the setting we already considered. However, it is worth to take a closer look at this setting, since the property of ordinal patterns to rely on the increments leads to an unexpected result.
We remember the estimator of ordinal pattern dependence given by

$$\hat{p}_n := \hat{p}\left(\left(X_t^{(1)}\right), \left(X_t^{(2)}\right)\right) := \frac{1}{n-h} \sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)\right\}}.$$

Now we assume the process $(X_j)_{j\in\mathbb{Z}} := \left\{\left(X_j^{(1)}, X_j^{(2)}\right)^t\right\}_{j\in\mathbb{Z}}$ to fulfill **(L1s)** in Overview 4.24. First, we define $X_{j,h+1} := \left(X_j^{(1)}, \ldots, X_{j+h}^{(1)}, X_j^{(2)}, \ldots, X_{j+h}^{(2)}\right)^t$. Note that $X_{j,h+1} \in \mathbb{R}^{2(h+1)}$. We recognize the notation used in Section 3.3 in the case $d = 2$. In order to apply any of the limit theorems proved in Chapter 3, we need to determine the Hermite rank of the estimator of ordinal pattern dependence, now with respect to $X_{0,h+1}$.

Lemma 4.59 *The Hermite rank $m(f, \Sigma_{2,h+1})$ of the function*

$$f(X_{j,h+1}) = \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)\right\}} - \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right)\right)$$

is equal to 2.

Proof. The proof follows the idea of the proof of Lemma 4.25. Since $\{X_j\}_{j\in\mathbb{Z}}$ is a bivariate Gaussian process, $X_{j,h+1}$ is a $(2h+2)$-dimensional Gaussian vector. We show that the Hermite rank of f with respect to $\Sigma_{2,h+1}$ is equal to 2. Hence, it is sufficient again to verify the following two properties:

$$(i) \quad m(f, \Sigma_{2h+2}) \geq 2,$$
$$(ii) \quad m(f, I_{2h+2}) \leq 2.$$

We start with the proof of property (i):
We have to proof that $m(f, \Sigma_{2h+2}) \neq 1$, which is equivalent to $\mathbb{E}\left(X_{j,h+1}^{(i)} f(X_{j,h+1})\right) = 0$, for all $i = 1, \ldots, h+1$. It holds

$$\begin{aligned}
\mathbb{E}\left(X_{j,h+1}^{(i)} f(X_{j,h+1})\right) &= \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_j^{(1)},\ldots,X_{j+h}^{(1)}\right)=\Pi\left(X_j^{(2)},\ldots,X_{j+h}^{(2)}\right)\right\}}\right) \\
&\quad - \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right)\right)\right) \\
&= \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\ldots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\ldots,X_h^{(2)}\right)\right\}}\right)
\end{aligned}$$

$$= \sum_{\pi \in S_h} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right).$$

Again, we split the set S_h into disjoint sets, namely into S_h^* and the complimentary set $S_h \setminus S_h^*$. With the property of the multivariate normal distribution that $X_{0,h+1} \overset{\mathcal{D}}{=} -X_{0,h+1}$ we arrive at:

$$\begin{aligned}
&\sum_{\pi \in S_h} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right) \\
=&\sum_{\pi \in S_h^*} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right) \\
&\quad + \sum_{\pi \in S_h \setminus S_h^*} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right) \\
=&\sum_{\pi \in S_h^*} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right) \\
&\quad + \sum_{\pi \in S_h \setminus S_h^*} \mathbb{E}\left(-X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(-X_0^{(1)},\dots,-X_h^{(1)}\right)=\Pi\left(-X_0^{(2)},\dots,-X_h^{(2)}\right)=\pi\right\}} \right) \\
=&\sum_{\pi \in S_h^*} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right) \\
&\quad - \sum_{\pi \in S_h \setminus S_h^*} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\mathcal{S}(\pi)\right\}} \right) \\
=&\sum_{\pi \in S_h^*} \mathbb{E}\left(X_{j,h+1}^{(i)} \mathbf{1}_{\left\{\Pi\left(X_0^{(1)},\dots,X_h^{(1)}\right)=\Pi\left(X_0^{(2)},\dots,X_h^{(2)}\right)=\pi\right\}} \right) \\
=&0.
\end{aligned}$$

So we have shown that $m(f, \Sigma_{2h+2}) \geq 2$ and hence (i).
We continue with the proof of (ii). We define $U_{1,h+1} := \left(U_1^{(1)}, \dots, U_{h+1}^{(1)}, U_1^{(2)}, \dots, U_{h+1}^{(2)}\right)^t$ analogously. Hence, $U_{1,h+1}$ is multivariate normal with covariance matrix $I_{2,h+1}$. Then, we obtain for $i = 1, \dots, h+1$, $k = h+2 \dots, 2h+2$, such that $k = i+h+1$:

$$\begin{aligned}
\mathbb{E}\left(U_{1,h+1}^{(i)} U_{1,h+1}^{(k)} f\left(U_{1,h+1}\right)\right) &= \mathbb{E}\left(U_{1,h+1}^{(i)} U_{1,h+1}^{(k)} \frac{1}{n-h} \sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{\Pi\left(U_1^{(1)},\dots,U_{h+1}^{(1)}\right)=\Pi\left(U_1^{(2)},\dots,U_{h+1}^{(2)}\right)\right\}} \right) \\
&\quad - \mathbb{E}\left(U_{1,h+1}^{(i)} U_{1,h+1}^{(k)} \mathbb{P}\left(\Pi\left(U_1^{(1)}, \dots, U_{h+1}^{(1)}\right) = \Pi\left(U_1^{(2)}, \dots, U_{h+1}^{(2)}\right)\right)\right) \\
&= \mathbb{E}\left(U_i^{(1)} U_k^{(2)} \mathbf{1}_{\left\{\Pi\left(U_1^{(1)},\dots,U_{h+1}^{(1)}\right)=\Pi\left(U_1^{(2)},\dots,U_{h+1}^{(2)}\right)\right\}} \right) \\
&= \sum_{\pi \in S_h} \mathbb{E}\left(U_i^{(1)} U_k^{(2)} \mathbf{1}_{\left\{\Pi\left(U_1^{(1)},\dots,U_{h+1}^{(1)}\right)=\Pi\left(U_1^{(2)},\dots,U_{h+1}^{(2)}\right)=\pi\right\}} \right)
\end{aligned}$$

$$= \sum_{\pi \in S_h} \left(\mathbb{E} \left(U_i^{(1)} \mathbf{1}_{\left\{ \Pi\left(U_1^{(1)}, \ldots, U_{h+1}^{(1)} \right) = \pi \right\}} \right) \right)^2$$
$$\neq 0,$$

since $\mathbb{E} \left(U_i^{(1)} \mathbf{1}_{\left\{ \Pi\left(U_1^{(1)}, \ldots, U_{h+1}^{(1)} \right) = \pi \right\}} \right) \neq 0$, as it is shown in the proof of Lemma 4.12.
This completes the proof. □

Remark 4.60 *Note that the estimator of the probability that the two time series both contain one fixed pattern $\pi \in S_h$ in the same moving window, namely*

$$g\left(X_{j,h+1}\right) = \frac{1}{n-h} \sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{ \Pi\left(X_j^{(1)}, \ldots, X_{j+h}^{(1)} \right) = \Pi\left(X_j^{(2)}, \ldots, X_{j+h}^{(2)} \right) = \pi \right\}} - \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)} \right) = \Pi\left(X_0^{(2)}, \ldots, X_h^{(2)} \right) = \pi \right)$$

has Hermite rank 1. This can be easily seen, since we only need to show $m\left(g, I_{2h+2}\right) = 1$. With $U_{j,h}$ as in the proof of Lemma 4.59 and $i = 1, ..., h+1$, we have

$$\mathbb{E}\left(U_{j,h}^{(i)} g\left(U_{j,h}\right) \right) = \mathbb{E} \left(U_{j,h}^{(i)} \mathbf{1}_{\left\{ \Pi\left(U_j^{(1)}, \ldots, U_{j+h}^{(1)} \right) = \Pi\left(U_j^{(2)}, \ldots, U_{j+h}^{(2)} \right) = \pi \right\}} \right)$$
$$= \mathbb{E} \left(U_{j+i-1}^{(1)} \mathbf{1}_{\left\{ \Pi\left(U_j^{(1)}, \ldots, U_{j+h}^{(1)} \right) = \pi \right\}} \right) \mathbb{P}\left(\Pi\left(U_j^{(2)}, \ldots, U_{j+h}^{(2)} \right) = \pi \right)$$
$$\neq 0,$$

by the same argument as in the proof of Lemma 4.12.

This concludes the preliminary considerations in order to apply the corresponding limit theorems in Chapter 3. However, in dealing with estimators in the context of ordinal pattern dependence, we are exactly in the special case that is studied in Section 3.3, namely that the functional we are considering is only depending on the increments processes. Therefore, the following two results are immediate consequences of Corollary 3.19 for the estimator of coincident patterns and Corollary 3.18 for the estimator of one fixed coincident patterns over time.

Corollary 4.61 *Under the assumptions of Corollary 3.19, it holds for $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$ that*

$$n^{1-2d^*} \left(\hat{p}_n - p\right) \xrightarrow{\mathcal{D}} \delta_0,$$

where δ_0 denotes the dirac measure in 0.

Similarly, we get the following result:

Corollary 4.62 *Under the assumptions of Corollary 3.18, it holds for $\pi \in S_h$ that*

$$n^{1/2-d^*} \left(\hat{p}_n(\pi) - p(\pi)\right) \xrightarrow{\mathcal{D}} \delta_0,$$

where δ_0 denotes the dirac measure in 0.

These results might be surprising, since we are considering the same estimators as in Section 4.3 and e.g. the result in Theorem 4.26 is expected to be independent of the use of $(X_{j,h})_{j\in\mathbb{Z}}$ or the corresponding increment process $(Y_{j,h})_{j\in\mathbb{Z}}$. The good news is that, of course, the same limit distribution is derived no matter which of the two processes we use to estimate p. The reason is that we are no longer in the setting of Theorem 4.26, as a closer look on the corresponding increment process $(Y_{j,h})_{j\in\mathbb{Z}}$ shows. As mentioned at the end of Section 3.3, in (3.52), the corresponding increment process is short-range dependent and we need to adapt the normalization to obtain a non-degenerated limit distribution. The normalization is taken as the classical convergence rate $n^{1/2}$ as it arises in central limit theorems for short-range dependent Gaussian time series, namely Theorem 4.35 for $\hat{p}_n$ and Theorem 4 in [2] for $\hat{p}_n(\pi)$.

By now, we considered special cases of standard ordinal pattern dependence. In the following three sections, we generalize the concept of standard ordinal pattern dependence to be able to compare a larger class of data sets by using ordinal pattern dependence. In doing so, we soften some assumptions on the parameter under consideration and get a more flexible, but still robust model. A different estimation method, albeit the stepwise approach on which we concentrated by now, is also provided.

4.4.4 Time shifted estimation of ordinal pattern dependence

Standard ordinal pattern dependence only detects coincident patterns in the same moving window, namely at the same points in time. Therefore, this approach would not detect dependence between two time series if the coincident patterns do not occur in the same moving window, but with a certain time shift.

Example 4.63 *Consider an example of one time series $\left(X_j^{(1)}\right)_{j\in\mathbb{Z}}$, which we shift by $a = 1$ to obtain the time series $\left(X_j^{(2)}\right)_{j\in\mathbb{Z}}$, such that $X_j^{(2)} = X_{j+1}^{(1)}$. These two time series are perfectly dependent. However, as it is shown in Figure 4.18, if we compare the patterns of length $h = 2$ marked by red at the same fixed windows of time, we do not recognize any coincident patterns at the considered points in time. There is a solution to this problem: if we shift the considered time window of the second process by 1, we would detect that all patterns are coincident. Therefore, we would get a correct representation of the dependence structure of $\left(X_j^{(1)}\right)_{j\in\mathbb{Z}}$ and $\left(X_j^{(2)}\right)_{j\in\mathbb{Z}}$.*

As proposed in [58], p.713, we define the estimator of time-shifted or time-warped ordinal pattern dependence as

$$\hat{p}_n^{a,b} := \frac{1}{n-a-hb+1} \sum_{j=1}^{n-a-hb+1} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\tilde{\Pi}\left(Y_{a+jb}^{(2)},\ldots,Y_{a+(j+h-1)b}^{(2)}\right)\right\}},$$

with $a, b \in \mathbb{N}_0$. We restrict ourselves to the case of time-shifting with $a \in \{0, \ldots, n-h-1\}$. We only consider $b = 1$, since stationarity is lost for $b \geq 2$, as the following calculation shows:
We take a closer look at the cross-correlation function of the process $\left(Y_{j,h}^{a,b}\right)_{j\in\mathbb{Z}}$, defined as

$$\left(Y_{j,h}^{a,b}\right)_{j\in\mathbb{Z}} := \left(\left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}, Y_{a+bj}^{(2)}, \ldots, Y_{a+(j+h-1)b}^{(2)}\right)^t\right)_{j\in\mathbb{Z}}.$$

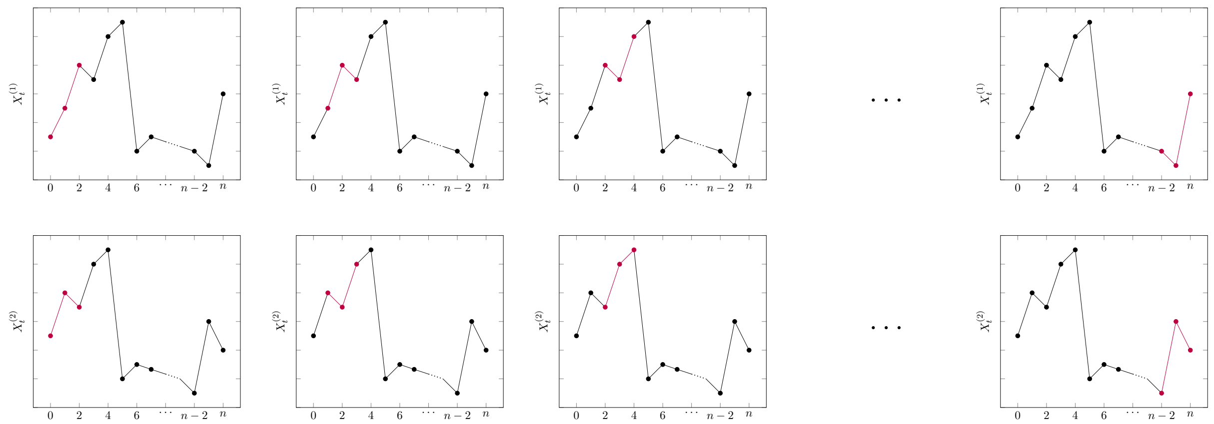

Figure 4.18: Illustration of the estimation of ordinal pattern dependence for $X_j^{(2)} = X_{j+1}^{(1)}$, $j = 0, \dots, n$.

Note that this is a Gaussian process, which is an immediate consequence of $(Y_j)_{j\in\mathbb{Z}}$ being a bivariate Gaussian process. The entries of the covariance matrix $\Sigma_{2,h}^{a,b}$ of

$$\left(Y_j^{(1)}, \dots, Y_{j+h-1}^{(1)}, Y_{a+bj}^{(2)}, \dots, Y_{a+(j+h)b-1}^{(2)}\right)^t$$

depend on a, b and the cross-correlation functions of $(Y_j)_{j\in\mathbb{Z}}$.
We assume for a moment that $p = 1, \dots, h$ and $q = h+1, \dots, 2h$. It holds that

$$\begin{aligned}\mathbb{E}\left(\left(Y_j^{a,b}\right)^{(p)}\left(Y_{j+k}^{a,b}\right)^{(q)}\right) &= \mathbb{E}\left(Y_{j+p-1}^{(1)} Y_{a+(j+q-h-1)b}^{(2)}\right)) \\ &= r^{(1,2)}(a + (j+k+q-h-1)b - j - p + 1) \\ &= r^{(1,2)}(a + j(b-1) + b(k+q-h-1) - p + 1). \end{aligned} \tag{4.50}$$

Expression (4.50) is, therefore, depending on j if $b \geq 2$.
Hence, we concentrate on $\hat{p}_n^a := \hat{p}_n^{a,1}$ and $\left(Y_{j,h}^a\right)_{j\in\mathbb{Z}} := \left(Y_{j,h}^{a,1}\right)_{j\in\mathbb{Z}}$.
We want to estimate $\max_a p^a$, with

$$p^a := \mathbb{P}\left(\tilde{\Pi}\left(Y_j^{(1)}, \dots, Y_{j+h-1}^{(1)}\right) = \tilde{\Pi}\left(Y_{j+a}^{(2)}, \dots, Y_{j+a+h-1}^{(2)}\right)\right).$$

An interesting question is the asymptotic distribution of $\max_a \hat{p}_n^a$, which provides the largest value of coincident ordinal patterns for any possible time shift a. Therefore, we obtain a huge amendment and generalization of the model considered. Note that $a = 0$ yields the standard estimator of ordinal pattern dependence.
Similar arguments as in (3.5) concerning the cross-correlation function calculated in (4.50) imply that $\left(Y_{j,h}^a\right)_{j\in\mathbb{Z}}$ is long-range dependent because eqrefmultivariateLRDconditionII holds for all $a \in \{0, \dots, n-h-1\}$. Therefore, as an immediate consequence of Lemma 4.25, we obtain the following result.

Lemma 4.64 *The Hermite rank of*

$$f\left(Y_j^{(1)}, \dots, Y_{j+h-1}^{(1)}, Y_{a+j}^{(2)}, \dots, Y_{a+j+h-1}^{(2)}\right) = \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\dots,Y_{j+h-1}^{(1)}\right) = \tilde{\Pi}\left(Y_{a+j}^{(2)},\dots,Y_{a+j+h-1}^{(2)}\right)\right\}}$$

for $a \in \{0, \dots, n-h-1\}$ *is equal to* 2.

We know that the asymptotic distribution of square-integrable functionals of long-range dependent Gaussian time series with Hermite rank 2 is only affected by the choice of the function f in the second order Hermite coefficients that play the role of scaling parameters, following Theorem 3.10 in (3.42). However, in this case, we additionally need to pay attention to the fact that the underlying process to which we apply the function f is given by $\left(Y_{j,h}^{a,1}\right)_{j\in\mathbb{Z}}$ and no longer by $(Y_{j,h})_{j\in\mathbb{Z}}$ as assumed in Theorem 3.10. Therefore, the term in (3.42) needs to be adapted to further study the asymptotic behaviour. We obtain

$$f\left(Y_{j,h}^{a}\right) \overset{\mathcal{D}}{=} n \sum_{p,q=1}^{2} \left(\sum_{l=0}^{h-1} \sum_{i=1}^{h-l} \alpha_{i,i+l}^{(p,q),a} \left(\hat{r}_n^{(p,q)}(l+\delta_{|p-q|,1}a) - r^{(p,q)}(l+\delta_{|p-q|,1}a)\right) + \sum_{l=1}^{h-1} \sum_{i=1}^{h-l} \alpha_{i+l,i}^{(p,q),a} \left(\hat{r}_n^{(q,p)}(l+\delta_{|p-q|,1}a) - r^{(q,p)}(l+\delta_{|p-q|,1}a)\right)\right) + o_{\mathbb{P}}(n^{2d^*}).$$

Since the convergence result in Lemma 3.5 holds jointly for all lags, we have for $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$ that

$$\hat{p}_n^a - p^a = \sum_{p,q\in P^*} \tilde{\alpha}_a^{(p,q)} \left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) + o_{\mathbb{P}}(n^{2d^*}),$$

with $\tilde{\alpha}_a^{(p,q)} := \sum_{i,k=1}^{h} \alpha_{i,k}^{(p,q),a}$, where $\alpha_{i,k}^{(p,q),a} = \alpha_{i+(p-1)h,k+(q-1)h}^{a}$ for each $p,q \in P^*$ and $i,k = 1,\ldots,h$ and $\left(\alpha_{i,k}^{a}\right)_{1\leq i,k\leq dh} = \left(\Sigma_{2,h}^{a}\right)^{-1} C^a \left(\Sigma_{2,h}^{a}\right)^{-1}$. Here

$$C^a = \left(c_{i,k}^{a}\right)_{1\leq i,k\leq 2h} = \mathbb{E}\left(Y_{1,h}^{a,1}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_{a+1}^{(2)},\ldots,Y_{a+h}^{(2)}\right)\right\}} - p\right)\left(Y_{1,h}^{a,1}\right)^t\right)$$

denotes the matrix of second order Hermite coefficients.

We show that $\hat{p}_n^a - p^a$ as well as $\tilde{\alpha}_a^{(p,q)}$ are bounded by a constant independent of a. Obviously we have that $\hat{p}_n^a, p^a \in [0,1]$ and, therefore, $|\hat{p}_n^a - p^a| \leq 1$. We take a closer look at the second order Hermite coefficients and for simplicity we regard $c_{1,1}^a$ representing the case $c_{i,i}^a$, $i = 1,\ldots,2h$. It holds that

$$\begin{aligned}
\left|c_{1,1}^{a}\right| &= \left|\mathbb{E}\left[\left(\left(Y_j^{(1)}\right)^2 - 1\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_{a+1}^{(2)},\ldots,Y_{a+h}^{(2)}\right)\right\}}\right]\right| \\
&\leq \mathbb{E}\left|\left(Y_j^{(1)}\right)^2 - 1\right|\left|\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_{a+1}^{(2)},\ldots,Y_{a+h}^{(2)}\right)\right\}}\right| \\
&\leq \mathbb{E}\left|\left(Y_j^{(1)}\right)^2 - 1\right| \\
&\leq \mathbb{E}\left(Y_j^{(1)}\right)^2 + 1 \\
&= 2
\end{aligned}$$

using Jensen's inequality and

$$
\begin{aligned}
\left|c^a_{1,h+1}\right| &= \left|\mathbb{E}\left[\left(Y_j^{(1)}Y_{j+a}^{(2)} - r^{(1,2)}(a)\right)\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_{a+1}^{(2)},\ldots,Y_{a+h}^{(2)}\right)\right\}}\right]\right| \\
&\leq \mathbb{E}\left|Y_j^{(1)}Y_{j+a}^{(2)} - r^{(1,2)}(a)\right|\left|\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_1^{(1)},\ldots,Y_h^{(1)}\right)=\tilde{\Pi}\left(Y_{b+1}^{(2)},\ldots,Y_{a+h}^{(2)}\right)\right\}}\right| \\
&\leq \mathbb{E}\left|Y_j^{(1)}Y_{j+a}^{(2)} - r^{(1,2)}(a)\right| \\
&= \mathbb{E}\left|Y_j^{(1)}Y_{j+a}^{(2)}\right| + \left|r^{(1,2)}(a)\right| \\
&= \mathbb{E}\left|U_1\left(r^{(1,2)}(a)U_1 + \sqrt{1-\left(r^{(1,2)}(a)\right)^2}U_2\right)\right| + 1 \\
&\leq \left|r^{(1,2)}(a)\right| + \sqrt{1-\left(r^{(1,2)}(a)\right)^2}\mathbb{E}\left|U_1\right|\mathbb{E}\left|U_2\right| + 1 \\
&\leq 2 + \mathbb{E}\left|U_1\right|\mathbb{E}\left|U_2\right| \\
&= 2 + \frac{2}{\pi}
\end{aligned}
$$

using the Cholesky decomposition of $\Sigma^a_{2,h}$. The Cholesky decomposition is given by $\Sigma^a_{2,h} = A^a\left(A^a\right)^t$ with

$$
A^a = \begin{pmatrix} 1 & 0 \\ r^{(1,2)}(a) & \sqrt{1-\left(r^{(1,2)}(a)\right)^2} \end{pmatrix}
$$

and it yields

$$
\left(Y_j^{(1)}, Y_{j+a}^{(2)}\right)^t \overset{\mathcal{D}}{=} \left(U_1, r^{(1,2)}(a)U_1 + \sqrt{1-\left(r^{(1,2)}(a)\right)^2}U_2\right)^t
$$

for $(U_1, U_2)^t \sim \mathcal{N}(0, I_2)$, where I_2 denotes the two-dimensional identity matrix.
Hence, we know that all entries in C^a are bounded by $2 + \frac{2}{\pi}$. The absolut values of the entries of the correlation matrix are also bounded by 1, independent of a. Therefore, there is a finite constant $K \in \mathbb{N}$, independent of a, such that the absolute values of the entries of the inverse of the correlation matrix are bounded by K. We have proven that all relevant values are bounded by a constant that is independent of a. Therefore, $\max_a \tilde{\alpha}_a^{(p,q)}$ exists for all $p, q = 1, 2$ and we can conclude

$$
\max_a\left(\hat{p}^a_n - p^a\right) = \sum_{p,q\in P^*}\max_a \tilde{\alpha}_a^{(p,q)}\left(\hat{r}_n^{(p,q)}(0) - r^{(p,q)}(0)\right) + o_{\mathbb{P}}(n^{2d^*-1}).
$$

Finally we obtain the following result.

Corollary 4.65 *Under the assumptions* ***(L2s)*** *in Overview 4.24, it holds that*

$$
\begin{aligned}
&n^{1-2d^*}(C_2)^{-\frac{1}{2}}\max_a\left(\hat{p}^a_n - p^a\right) \\
&\qquad \xrightarrow{\mathcal{D}} \sum_{p,q\in P^*}\max_a \tilde{\alpha}_a^{(p,q)} Z^{(p,q)}_{2,d^*+1/2}(1),
\end{aligned}
$$

with $Z^{(p,q)}_{2,d^*+1/2}(1)$ as given in Theorem 4.26 and $C_2 := \frac{1}{2d^*(4d^*-1)}$ being a normalizing constant.

4.4.5 Blockwise estimation of ordinal pattern dependence

We turn to an approach, where not the standard ordinal pattern dependence as underlying model is modified, but the way of estimating it. In certain applications, it might be more useful to change the original idea of using a moving window of a fixed length that is shifted by one for each consecutive summand of the estimator, as done by

$$\hat{p}_n = \frac{1}{n-h} \sum_{j=0}^{n-h-1} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{j+1}^{(1)},\ldots,Y_{j+h}^{(1)}\right)=\tilde{\Pi}\left(Y_{j+1}^{(2)},\ldots,Y_{j+h}^{(2)}\right)\right\}}$$

and to modify it to the so-called *blockwise* estimation of ordinal pattern dependence. As illustrated in Figure 4.19, this means that we are no longer considering overlapping moving windows for the comparison of the ordinal patterns marked by red, in contrast to the stepwise estimation. Instead we shift the moving windows such that only the the last point in time in a certain moving window coincides with the first point in time of the following moving window. This guarantees that for each moving window the set of increments under consideration has no intersections with the set of increments corresponding to the following (or any other) moving window. Therefore,

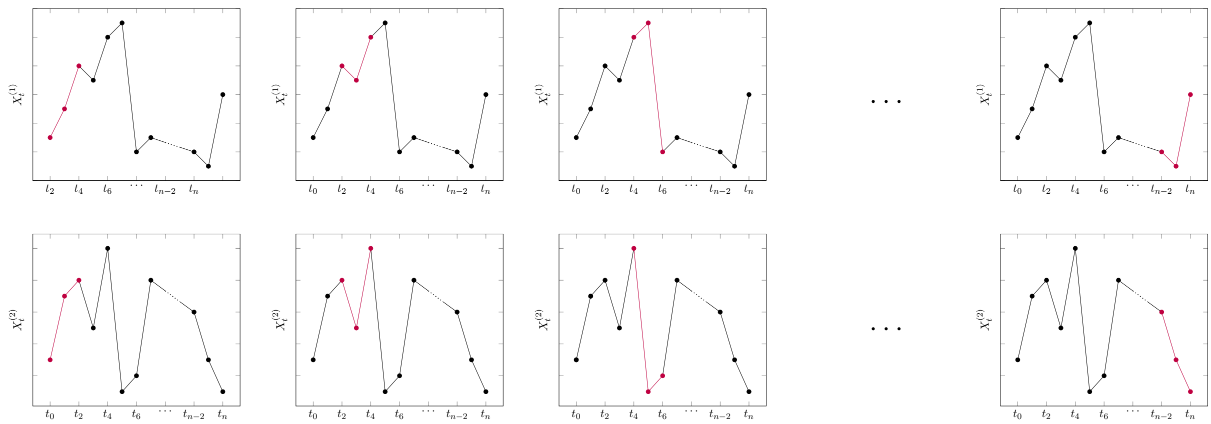

Figure 4.19: Illustration of blockwise estimation of ordinal pattern dependence for $h = 2$.

we are still interested in estimating the parameter p, but in the run of the time series we consider

$$\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{1}^{(1)},\ldots,Y_{h}^{(1)}\right)=\tilde{\Pi}\left(Y_{1}^{(2)},\ldots,Y_{h}^{(2)}\right)\right\}} + \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{h+1}^{(1)},\ldots,Y_{2h}^{(1)}\right)=\tilde{\Pi}\left(Y_{h+1}^{(2)},\ldots,Y_{2h}^{(2)}\right)\right\}} + \ldots$$

and hence obtain the following estimator

$$\hat{p}_n^{block} := \frac{1}{\left\lfloor \frac{n}{h} \right\rfloor} \sum_{j=0}^{\left\lfloor \frac{n}{h} \right\rfloor -1} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_{1+jh}^{(1)},\ldots,Y_{h+jh}^{(1)}\right)=\tilde{\Pi}\left(Y_{1+jh}^{(2)},\ldots,Y_{h+jh}^{(2)}\right)\right\}}.$$

Note that if $(Y_{j,h})_{j\in\mathbb{Z}}$ fulfills **(L1)** in Overview 4.24, the same holds true for the considered process $\left(Y_{j,h}^{block}\right)_{j\in\mathbb{Z}}$ here, with

$$Y_{j,h}^{block} := \left(Y_{1+jh}^{(1)}, \ldots, Y_{h+jh}^{(1)}, Y_{1+jh}^{(2)}, \ldots, Y_{h+jh}^{(2)}\right)^t.$$

The justification is similar to the one in (3.5). For example, consider for $p = 1, \dots, h$ and $q = h+1, \dots, 2h$

$$\begin{aligned}\mathbb{E}\left(\left(Y_{j,h}^{block}\right)^{(p)}\left(Y_{j+k,h}^{block}\right)^{(q)}\right) &= \mathbb{E}\left(Y_{p+jh}^{(1)} Y_{q-h+(j+k)h}^{(2)}\right) \\ &= r^{(1,2)}(q + (k-1)h - p) \\ &\simeq r^{(1,2)}(k), \qquad (k \to \infty).\end{aligned}$$

Note that $Y_{j,h}^{block} \overset{\mathcal{D}}{=} Y_{j,h}$.
However, there is a significant difference in this setting compared to the estimator of coincident patterns $\hat{p}_n$. In the standard setting, we define $(Y_{j,h})$ by plugging h consecutive entries of $Y_j^{(1)}$ and of $Y_j^{(2)}$ into one vector. That was also possible concerning the estimator of time shifts, although we did not use the same starting point j. However, the underyling process $\left(Y_j^{(1)}, Y_{j+a}^{(2)}\right)_{j\in\mathbb{Z}}$ was still a bivariate long-range dependent Gaussian process. Unfortunately, it is not possible to create a bivariate Gaussian process such that $\left(Y_{j,h}^{block}\right)$ consists of h **consecutive** entries of the marginal processes. The reason for this is the special structure of $\left(Y_{j,h}^{block}\right)$. If we raise j by 1, the corresponding univariate components get shifted by h and not only by 1. This is an issue concerning the application of Theorem 3.10, because following the proof of this theorem in this modified setting, we are no longer considering the convergence of the classical sample cross-correlations in (3.42), but need asymptotic results for

$$\sum_{p,q=1}^{2} \sum_{l=0}^{h-1} \sum_{j=1}^{n-l} Y_{1+jh}^{(p)} Y_{1+jh+l}^{(q)} - r^{(p,q)}(l).$$

It is possible to show that for $p, q = 1, 2$ the asymptotics of $Y_{1+jh}^{(p)} Y_{1+jh+l}^{(q)} - r^{(p,q)}(l)$ are independent of $l = 0, \dots, h-1$. To this end, we follow [9], Section 4.4.1.3. Following the linear representation in (2.5), we know that we can express the original process

$$\begin{aligned}Y_t &= \sum_{j=1}^{\infty} \Psi_j \varepsilon_{t-j} \\ &= \sum_{j=1}^{\infty} \begin{pmatrix} \psi_j^{1,1} & \psi_j^{1,2} \\ \psi_j^{2,1} & \psi_j^{2,2} \end{pmatrix} \begin{pmatrix} \varepsilon_{t-j}^{(1)} \\ \varepsilon_{t-j}^{(2)} \end{pmatrix}.\end{aligned}$$

Note that $(\varepsilon_j)_{j\ in\mathbb{Z}}^t$ is a two-dimensional white noise sequence with $\mathbb{E}\left(\varepsilon_1 \varepsilon_1^t\right) = I_2$.
Following [9], p. 225 and Section 4.4.1.1, ε_j has the spectral representation

$$\varepsilon_{t-j} = \int_{-\pi}^{\pi} \mathrm{e}^{i(t-j)\lambda} M_\varepsilon(\mathrm{d}\lambda),$$

with $M_\varepsilon(\mathrm{d}\lambda) = \left(M_\varepsilon^{(1)}(\mathrm{d}\lambda), M_\varepsilon^{(2)}(\mathrm{d}\lambda)\right)^t$. Each $M_\varepsilon^{(p)}(\mathrm{d}\lambda)$, $p = 1, 2$ is an Hermitian-Gaussian random spectral measure on $(-\pi, \pi]$ as discussed in Section 2.6. Furthermore, $M_\varepsilon^{(1)}(\lambda)$ and $M_\varepsilon^{(2)}(\lambda)$ are independent. The spectral measure fulfills $n^{1/2} M_\varepsilon\left(n^{-1}\cdot\right) \overset{\mathcal{D}}{=} M_\varepsilon(\cdot)$, see [9], Sec. 3.7.1.

We can describe the original process Y_t by

$$
\begin{aligned}
Y_t &= \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} M_Y(\mathrm{d}\lambda) \\
&= \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} \sum_{j=-\infty}^{\infty} \begin{pmatrix} \psi_j^{1,1} & \psi_j^{1,2} \\ \psi_j^{2,1} & \psi_j^{2,2} \end{pmatrix} \left(\mathrm{e}^{-i\lambda}\right)^j M_\varepsilon(\mathrm{d}\lambda) \\
&= \int_{-\pi}^{\pi} \mathrm{e}^{it\lambda} \left(\sum_{j=-\infty}^{\infty} \sum_{q=1}^{2} \psi_j^{p,q} \left(\mathrm{e}^{-i\lambda}\right)^j M_\varepsilon^{(q)}(\mathrm{d}\lambda) \right)_{p=1,2}.
\end{aligned}
$$

Therefore, for $p = 1, 2$, we have

$$
\mathrm{d}M_Y^{(p)}(\lambda) = \sum_{q=1}^{2} \underbrace{\sum_{j=-\infty}^{\infty} \psi_j^{p,q} \left(\mathrm{e}^{-i\lambda}\right)^j}_{=:a^{(p,q)}(\lambda)} \mathrm{d}M_\varepsilon^{(q)}(\lambda).
$$

Following [9], Section 4.4.1.3, p. 305, we obtain

$$
\begin{aligned}
&\sum_{j=0}^{\frac{n}{h}-1} Y_{1+jh}^{(p)} Y_{1+jh+l}^{(q)} - r^{(p,q)}(l) \\
&= \sum_{j=0}^{\frac{n}{h}-1} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathrm{e}^{i(1+jh)\lambda_1} \mathrm{e}^{i(1+jh+l)\lambda_2} \; M_Y^{(p)}(\mathrm{d}\lambda_1) \; M_Y^{(q)}(\mathrm{d}\lambda_2) \\
&= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\mathrm{e}^{in(\lambda_1+\lambda_2)} - 1}{\mathrm{e}^{ih(\lambda_1+\lambda_2)} - 1} \mathrm{e}^{i(\lambda_1+\lambda_2)} \mathrm{e}^{il\lambda_2} \; M_Y^{(p)}(\mathrm{d}\lambda_1) \; M_Y^{(q)}(\mathrm{d}\lambda_2) \\
&= \int_{-n\pi}^{n\pi} \int_{-n\pi}^{n\pi} \frac{\mathrm{e}^{i(\lambda_1+\lambda_2)} - 1}{n\left(\mathrm{e}^{i/nh(\lambda_1+\lambda_2)} - 1\right)} \mathbf{1}_{\left\{\left|\frac{\lambda_1}{n}\right|, \left|\frac{\lambda_2}{n}\right| \leq \pi n\right\}} \mathrm{e}^{i\frac{\lambda_1+\lambda_2}{n}} \mathrm{e}^{il\frac{\lambda_2}{n}} n^{1/2} M_Y^{(p)}\left(n^{-1}\mathrm{d}\lambda_1\right) n^{1/2} M_Y^{(q)}\left(n^{-1}\mathrm{d}\lambda_2\right) \\
&= \int_{-n\pi}^{n\pi} \int_{-n\pi}^{n\pi} \frac{\mathrm{e}^{ih(\lambda_1+\lambda_2)} - 1}{n\left(\mathrm{e}^{i/nh(\lambda_1+\lambda_2)} - 1\right)} \mathbf{1}_{\left\{\left|\frac{\lambda_1}{n}\right|, \left|\frac{\lambda_2}{n}\right| \leq \pi n\right\}} \mathrm{e}^{i\frac{\lambda_1+\lambda_2}{n}} \mathrm{e}^{il\frac{\lambda_2}{n}} \left(\sum_{q_1=1}^{2} a^{p,q_1}\left(\frac{\lambda_1}{2}\right) n^{1/2} M_\varepsilon^{(q_1)}\left(n^{-1}\mathrm{d}\lambda_1\right) \right) \\
&\qquad \times \left(\sum_{q_2=1}^{2} a^{q,q_2}\left(\frac{\lambda_2}{2}\right) n^{1/2} M_\varepsilon^{(q_2)}\left(n^{-1}\mathrm{d}\lambda_2\right) \right) \\
&= \int_{-n\pi}^{n\pi} \int_{-n\pi}^{n\pi} \underbrace{\frac{\mathrm{e}^{ih(\lambda_1+\lambda_2)} - 1}{\mathrm{e}^{i/nh(\lambda_1+\lambda_2)} - 1} \mathbf{1}_{\left\{\left|\frac{\lambda_1}{n}\right|, \left|\frac{\lambda_2}{n}\right| \leq \pi n\right\}}}_{\xrightarrow{(n\to\infty)} \frac{\mathrm{e}^{i(\lambda_1+\lambda_2)}}{ih(\lambda_1+\lambda_2)}} \underbrace{\mathrm{e}^{i\frac{\lambda_1+\lambda_2}{n}} \mathrm{e}^{il\frac{\lambda_2}{n}}}_{\xrightarrow{(n\to\infty)} 1} \left(\sum_{q_1=1}^{2} a^{p,q_1}\left(\frac{\lambda_1}{2}\right) M_\varepsilon^{(q_1)}(\mathrm{d}\lambda_1) \right) \\
&\qquad \times \left(\sum_{q_2=1}^{2} a^{q,q_2}\left(\frac{\lambda_2}{2}\right) M_\varepsilon^{(q_2)}(\mathrm{d}\lambda_2) \right) \\
&= \frac{1}{h} \sum_{j=0}^{n-1} Y_j^{(p)} Y_j^{(q)} - r^{(p,q)}(0) + o_{\mathbb{P}}(1). \qquad (4.51)
\end{aligned}
$$

Hence, the limiting behaviour is independent of the considered lag $l = 0, \ldots, h+1$ and it can be expressed in terms of the original bivariate process $(Y_j)_{j\in\mathbb{Z}}$ in our setting.

Finally, by following the proof of Theorem 3.10 until (3.44) we arrive at

$$\begin{aligned} n^{1-2d^*}(C_2)^{-\frac{1}{2}}\left(\hat{p}_n^{block}-p\right) &= h(C_2)^{-\frac{1}{2}}n^{-2d^*}\sum_{j=0}^{\frac{n}{h}-1}\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_{1+jh},\ldots,Y^{(1)}_{h+jh}\right)=\tilde{\Pi}\left(Y^{(2)}_{1+jh},\ldots,Y^{(2)}_{h+jh}\right)\right\}}-p\right) \\ &\overset{\mathcal{D}}{=} h(C_2)^{-\frac{1}{2}}n^{-2d*}\sum_{p,q=1}^{2}\tilde{\alpha}^{(p,q)}_{block}\frac{1}{h}\sum_{j=0}^{n-1}Y^{(p)}_jY^{(q)}_j - r^{(p,q)}(0)+o_{\mathbb{P}}(1) \\ &= n^{-2d^*}(C_2)^{-\frac{1}{2}}\sum_{p,q\in P^*}\tilde{\alpha}^{(p,q)}Y^{(p)}_jY^{(q)}_j - r^{(p,q)}(0)+o_{\mathbb{P}}(1), \end{aligned}$$

with $\tilde{\alpha}^{(p,q)}$ as given in Theorem 4.26, since $Y^{block}_{j,h}\overset{\mathcal{D}}{=}Y_{j,h}$. Therefore, the convergence behaviour of $\hat{p}^{block}_n$ is exactly the same as for the standard estimator $\hat{p}_n$. Hence, we obtain a limit theorem for the considered problem that includes the mixed cases as well as the short-range dependent setting.

Corollary 4.66 *The following statements are true:*

(i) Under the assumptions in ***(L2s)*** *in Overview 4.24, it holds that*

$$\begin{aligned} &n^{1-2d^*}(C_2)^{-\frac{1}{2}}\left(\hat{p}_n^{block}-p\right) \\ &\qquad\overset{\mathcal{D}}{\longrightarrow}\sum_{p,q\in P^*}\tilde{\alpha}^{(p,q)}Z^{(p,q)}_{2,d^*+1/2}(1), \end{aligned}$$

with $Z^{(p,q)}_{2,d^*+1/2}(1)$ as given in Theorem 4.26 and $C_2 := \frac{1}{2d^*(4d^*-1)}$ being a normalizing constant.

(ii) Under the assumptions in ***(S)*** *in Overview 4.24, it holds that*

$$\left(\frac{n}{h}\right)^{\frac{1}{2}}(\hat{p}-p)\overset{\mathcal{D}}{\longrightarrow}\mathcal{N}\left(0,\sigma^2\right),$$

with

$$\begin{aligned} \sigma^2 = \sum_{k=-\infty}^{\infty}\mathbb{E}\Bigg[&\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_{1},\ldots,Y^{(1)}_{h}\right)=\tilde{\Pi}\left(Y^{(2)}_{1},\ldots,Y^{(2)}_{h}\right)\right\}}-p\right) \\ &\times\left(\mathbf{1}_{\left\{\tilde{\Pi}\left(Y^{(1)}_{1+kh},\ldots,Y^{(1)}_{h+kh}\right)=\tilde{\Pi}\left(Y^{(2)}_{1+kh},\ldots,Y^{(2)}_{h+kh}\right)\right\}}-p\right)\Bigg]. \end{aligned}$$

Remark 4.67 *The normalization in part (ii) of Corollary 4.66 depends on h, but it does not in part (i). It is possible to express*

$$\sum_{j=0}^{\frac{n}{h}-1}Y^{(p)}_{1+jh}Y^{(q)}_{1+jh} - r^{(p,q)}(0) = \sum_{j=0}^{\frac{n}{h}-1}\check{Y}^{(p)}_j\check{Y}^{(q)}_j - r^{(p,q)}(0),$$

with $\check{Y}^{(p)}_j := Y^{(p)}_{1+jh}$ for $p=1,2$. Note that the considerations in (4.51) *are already applied here and, therefore, this approach does not replace their necessity. With similar arguments as for*

$\left(Y_{j,h}^{block}\right)_{j\in\mathbb{Z}}$, *we know that* $\left(\check{Y}_j\right)_{j\in\mathbb{Z}}$ *is a bivariate long-range dependent Gaussian process, if* $(Y_j)_{j\in\mathbb{Z}}$ *is one. Hence, it is possible to apply Lemma 3.5 and obtain*

$$\left(\frac{n}{h}\right)^{1-2d^*}(C_2)^{-\frac{1}{2}}\left(\hat{p}_n^{block}-p\right)\overset{\mathcal{D}}{=}\left(\frac{h}{n}\right)^{-2d^*}(C_2)^{-\frac{1}{2}}\sum_{p,q\in P^*}\tilde{\alpha}^{(p,q)}\sum_{j=0}^{\frac{n}{h}-1}\check{Y}_j^{(p)}\check{Y}_j^{(q)}-r^{(p,q)}(0)+o_{\mathbb{P}}(1)$$
$$\xrightarrow{\mathcal{D}}\sum_{p,q\in P^*}\tilde{\alpha}^{(p,q)}Z_{2,d^*+1/2}^{(p,q)}(1).$$

jointly for $p,q=1,2$. *The factor* h^{-2d^*} *is hidden in the slowly varying constants on the right-hand side, because we have for* $p,q=1,2$ *that*

$$\check{r}^{(p,q)}(k):=\mathbb{E}\left(\check{Y}_j^{(p)}\check{Y}_{j+k}^{(q)}\right)=r^{(p,q)}(kh)=\underbrace{L_{p,q}(kh)h^{d_p+d_q-1}}_{=:\check{L}_{p,q}(k)}k^{d_p+d_q-1},$$

such that $\check{L}_{p,q}(k)\to L_{p,q}h^{d_p+d_q-1}$. *Therefore, this approach is consistent with part (i) of Corollary 4.66. However, the steps we used to prove Corollary 4.66, part (i) are more illustrative, since we showed that the convergence behaviour could be expressed in terms of the original process* $(Y_j)_{j\in\mathbb{Z}}$ *again.*

Finally, we provide a generalization of standard ordinal pattern dependence that extends the underlying method of only taking exactly coincident patterns into consideration. Especially for larger h, this is a very strong condition and the dependence structure between data sets is underestimated. Therefore, in the next section, a model is proposed that allows for certain deviations between the patterns occuring.

4.4.6 Average-weighted ordinal pattern dependence

As we have already seen in Section 4.4.2, in some situations it is useful not to compare coincident patterns but, for example, a fixed pattern π in the moving window of the first time series and the spatial reflected pattern $\mathcal{S}(\pi)$ in the corresponding window of the second time series. In the context of the Hurst parameter estimation, we were able to solve that, since by changing the sign of the increments in the second time series, we could boil this problem down to the well-known case of comparing coincident pattern.

In the method introduced in this section, we focus on positive dependence. However, we want to take a closer look at the general idea of softening the condition of coincident patterns in both time series to allowing for patterns with a certain deviation. An illustrative approach is given in [58], Section 3, where *average weighted ordinal pattern dependence* is introduced. Their main idea is, that, since standard ordinal pattern dependence deals with estimating dependence based on the co-movement of the two time series, especially for large h it might happen that the movement of the two time series is very similar, but the pattern are not exactly coincident. Then, standard ordinal pattern dependence does not detect this "almost co-movement", as is illustrated in Figure 4.20. Although the two time series displayed only have a small difference in the value of $X_{t_6}^{(1)}$ and are equal otherwise, the estimator of standard ordinal pattern dependence

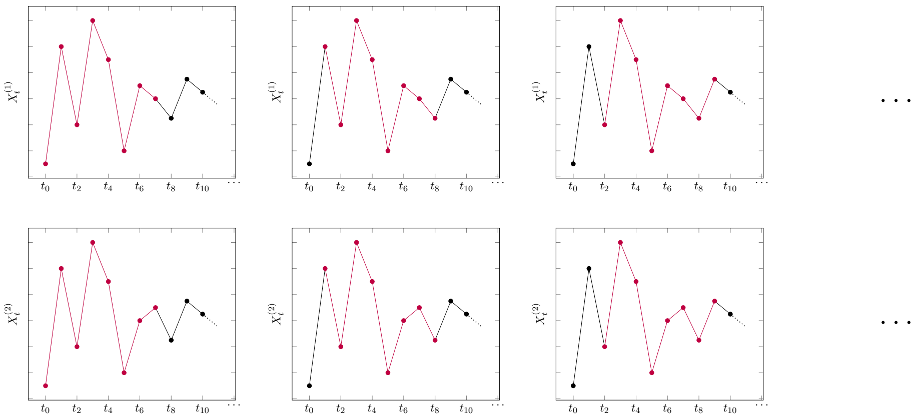

Figure 4.20: Estimation of ordinal pattern dependence for $h = 7$ for two almost similar time series.

for $h = 7$ does not detect one coincident patterns in the first moving windows, in which the patterns under consideration are red colored. Therfore, standard ordinal pattern dependence underestimates the dependence.

	1st moving window	2nd moving window	3rd moving window
$X_j^{(1)}$	$(3,1,4,6,7,2,5,0)$	$(2,0,3,5,6,7,1,4)$	$(1,2,7,4,5,6,0,3)$
$X_j^{(2)}$	$(3,1,4,7,6,2,5,0)$	$(2,0,3,6,5,7,1,4)$	$(1,2,7,5,4,6,0,3)$

Table 4.5: Ordinal patterns arising in the moving windows in Figure 4.20.

Only counting coincident patterns is, therefore, too restrictive and [58] suggest to soften it to the case, that a certain (small) deviation is allowed, concerning the respective pattern each time series exhibits in the same moving window. However, this deviation has to be measured in a reasonable sense. If we consider the patterns $\pi = (3,1,2,4,0)$ and $\sigma = (3,4,2,1,0)$, they seem to be very similar: only two entries are changed. However, as visualized in Figure 4.21, one recognizes a larger difference in their behaviour.

Figure 4.21: Two ordinal patterns of length $h = 4$.

This motivates to capture the expression *small deviation* mathematically to assure plausibility.

Recalling Figure 4.20, we take a closer look at the ordinal patterns displayed. They are listed in Table 4.5. In each moving window, we only have one transposition of two consecutive numbers. Therefore, this kind of difference should be allowed. This small deviation of the patterns is emphasized in Figure 4.22, where the two patterns displayed are close enough to count in the

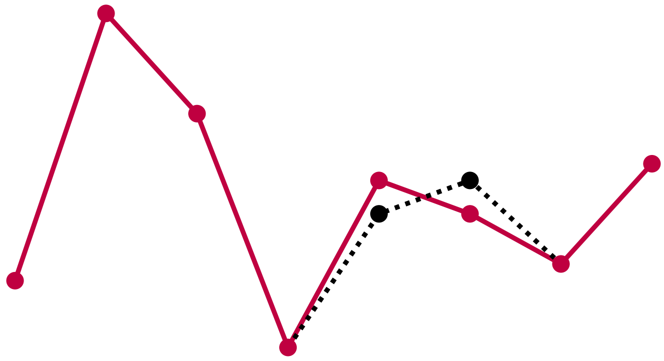

Figure 4.22: Modification of the pattern $(1, 2, 7, 4, 5, 6, 0, 3)$ (red) to $(1, 2, 7, 5, 4, 6, 0, 3)$ (black).

dependence structure. However, we have to differentiate between coincident patterns and two patterns with a tolerated difference.

Therefore, we follow [58], Section 3 and describe the mathematical framework in the following. We need a (pseudo-)metric $d : S_h \times S_h \to \mathbb{R}_+$ to measure the deviation between the patterns and a weight-function $\omega : d(S_h, S_h) \to [0, 1]$ with $\omega(0) = 1$ that assigns a certain value to each deviation and should, therefore, be monotonically decreasing. This assures that a large deviation only yields a small weight. Note that $\omega \circ d$ is a similarity measure. For further mathematical conditions on the possible choices of these two functions, the reader is refered to [58], Section 3.1. Following their approach in (21), we define average weighted ordinal pattern dependence as

$$\begin{aligned} AWOPD :&= \mathbb{E}\left[\omega\left(d\left(\Pi\left(X_j^{(1)}, \ldots X_{j+h}^{(1)}\right), \Pi\left(X_j^{(2)}, \ldots X_{j+h}^{(2)}\right)\right)\right)\right] \\ &\quad - \sum_{\pi,\sigma \in S_h} \omega(d(\pi,\sigma))\mathbb{P}\left(\ \Pi\left(X_j^{(1)}, \ldots X_{j+h}^{(1)}\right) = \pi\right)\mathbb{P}\left(\ \Pi\left(X_j^{(2)}, \ldots X_{j+h}^{(2)}\right) = \sigma\right). \end{aligned}$$

Hence, the estimator of the first summand is given by

$$\begin{aligned} \hat{p}_n^{\text{awopd}} :&= \frac{1}{n-h}\sum_{j=0}^{n-h-1} \omega\left(d\left(\Pi\left(X_j^{(1)}, \ldots X_{j+h}^{(1)}\right), \Pi\left(X_j^{(2)}, \ldots X_{j+h}^{(2)}\right)\right)\right) \\ &= \frac{1}{n-h}\sum_{j=1}^{n-h} \omega\left(d\left(\tilde{\Pi}\left(Y_j^{(1)}, \ldots Y_{j+h-1}^{(1)}\right), \tilde{\Pi}\left(Y_j^{(2)}, \ldots Y_{j+h-1}^{(2)}\right)\right)\right). \end{aligned}$$

Correspondingly we denote the true value by

$$p^{\text{awopd}} := \mathbb{E}\left[\omega\left(d\left(\Pi\left(X_j^{(1)}, \ldots X_{j+h}^{(1)}\right), \Pi\left(X_j^{(2)}, \ldots X_{j+h}^{(2)}\right)\right)\right)\right].$$

We modify the estimator for the hypothetical case of independence, as in Section 4.3.4, to

$$\hat{q}_n^{\text{awopd}} := \sum_{\pi,\sigma \in S_h} \omega(d(\pi,\sigma))\hat{q}_{Y^{(1)},n,\text{imp}}(\pi)\hat{q}_{Y^{(2)},n,\text{imp}}(\sigma).$$

Hence, combining both estimators yields

$$\widehat{AWOPD} := \hat{p}_n^{\text{awopd}} - \hat{q}_n^{\text{awopd}}.$$

Note that it is possible to norm $\widehat{AWOPD}$ the same way as $\widehat{OPD}$ as described in (4.13). However, as it is remarked in [58], p. 715, we omit the normalization here for the practitioner's convenience.

The first difference to the setting considered in [58], Section 3.2, is that the limit behaviour of the estimators in the context of long-range dependence depend on the Hermite rank of the estimators. The Hermite rank highly depends on the choice of the weight-function and the (pseudo-)metric considered.
We focus on the setting described in Example 4 in [58]. They introduced the intuitive l^1-distance $d_{l^1}(\sigma, \pi) := \sum_{j=0}^{h} |\pi_j - \sigma_j|$ to measure the deviation between the two patterns $\pi, \sigma \in S_h$. However, the value of d_{l^1} gets bigger the larger the distance between the two patterns is and is, therefore, exhibiting exactly the contrary behaviour to our aim, namely to reward a small distance and to penalize a large one. This explains the need of the monotonically decreasing weight function. In Example 4 in [58], the l^1-distance gets equipped with the weight function $\omega := \mathbf{1}_{\{0\}} + \frac{3}{4}\mathbf{1}_{\{2\}} + \frac{1}{2}\mathbf{1}_{\{4\}} + \frac{1}{4}\mathbf{1}_{\{6\}}$. Note, that for this choice of the metric, standard ordinal pattern dependence is obtained, if $\omega := \mathbf{1}_{\{0\}}$.
Remark, that this choice of d and w exactly pictures the requirements we made at the beginning of this chapter. In the example given in Figure 4.21, where the two patterns rather exhibit a far-out similar behaviour, the l^1-distance assigns them the value 6. With applying the weight-function above, this only yields a contribution of $\frac{1}{4}$ to the estimation of average-weighted ordinal pattern dependence. However, for the example in Table 4.5, where only a small deviation between the patterns within the same moving windows was recognized, the l^1-distance yields the value 2 for each moving window, which takes the similarity of the pattern, emphasized in Figure 4.22, into account. This yields $\frac{3}{4}$ as contribution to the estimation of average-weighted ordinal pattern dependence for each moving window displayed in the time series in Figure 4.20.
Before investigating the asymptotics of the estimator of average-weighted ordinal pattern dependence, we give an example why this choice of d and w is especially suitable for large h.

Example 4.68 *We take a closer look at the case $h = 2$. In this case, the patterns with the largest deviation to π are given by the time- and space reversion of π. Considering, for example, $\pi = (0, 1, 2)$ and $\mathcal{S}(\pi) = (2, 1, 0)$, we obtain $d(\pi, \mathcal{S}(\pi)) = 4$. This still yields a contribution of $\frac{1}{2}$ to the estimation of average-weighted ordinal pattern dependence, although we are investigating positive dependence and these two patterns clearly exhibit anti-monotonic behaviour. However, this problem vanishes with increasing h: if we have $\pi = (\pi_0, \ldots, \pi_h)$, it holds that $\mathcal{S}(\pi) = (\pi_h, \ldots, \pi_0)$ and $\mathcal{T}(\pi) = (h - \pi_0, \ldots, h - \pi_h)$. Concerning the l^1-distance, this yields*

$$
\begin{aligned}
d(\pi, \mathcal{S}(\pi)) &= \sum_{i=0}^{h} |\pi_i - \pi_{h-i}| \\
&= 2 \sum_{i=0}^{\lfloor \frac{h-1}{2} \rfloor} \underbrace{|\pi_i - \pi_{h-i}|}_{\geq 1} \\
&\geq 2 \left(\left\lfloor \frac{h-1}{2} \right\rfloor + 1 \right)
\end{aligned}
$$

$$= \begin{cases} h+1, & h \text{ odd}, \\ h, & h \text{ even}. \end{cases}$$

Analogously we obtain

$$\begin{aligned} d(\pi, \mathcal{T}(\pi)) &= \sum_{i=0}^{h} |h - 2\pi_i| \\ &= 2 \sum_{i=0}^{\lfloor \frac{h-1}{2} \rfloor} \underbrace{|h - 2\pi_i|}_{\geq 1} \\ &\geq \begin{cases} h+1, & h \text{ odd}, \\ h, & h \text{ even}. \end{cases} \end{aligned}$$

Hence, for $h > 6$, neither a time-reversed nor a space-reversed pattern in the other time series have any impact on the estimation of (positive) average-weighted ordinal pattern dependence.

In practice, the user has to decide which value of h she or he employs, since larger values result in a higher computational effort. Smaller values like $h = 5$ or $h = 6$ yield a good balance, since anti-monotonic patterns get, if at all, at most a small value and the computational effort needed is still feasible. We focus on the asymptotic properties of the estimator of average-weighted ordinal pattern dependence and we begin with the estimator of p^{awopd}. We determine the Hermite rank of $\hat{p}_n^{\text{awopd}}$ in order to get an asymptotic result for $\widehat{AWOPD}$.

Lemma 4.69 *The Hermite rank of*

$$f^{\text{awopd}}(Y_{j,h}) := \omega\left(d\left(\tilde{\Pi}\left(Y_j^{(1)}, \ldots Y_{j+h-1}^{(1)}\right), \tilde{\Pi}\left(Y_j^{(2)}, \ldots Y_{j+h-1}^{(2)}\right)\right)\right),$$

with

$$d(\sigma, \pi) := d_{l^1}(\sigma, \pi) := \sum_{j=0}^{h} |\pi_j - \sigma_j|$$

and

$$\omega := \mathbf{1}_{\{0\}} + \frac{3}{4}\mathbf{1}_{\{2\}} + \frac{1}{2}\mathbf{1}_{\{4\}} + \frac{1}{4}\mathbf{1}_{\{6\}}$$

with respect to $\Sigma_{2,h}$ is equal to 2.

Proof. We have

$$\begin{aligned} f^{\text{awopd}}(Y_{j,h}) &= \sum_{k=0}^{3} \frac{4-k}{4} \mathbf{1}_{\{2k\}}\left(d_{l^1}\left(\tilde{\Pi}\left(Y_j^{(1)}, \ldots Y_{j+h-1}^{(1)}\right), \tilde{\Pi}\left(Y_j^{(2)}, \ldots Y_{j+h-1}^{(2)}\right)\right)\right) \\ &= \sum_{k=0}^{3} \sum_{\pi, \sigma \in S_h} \frac{4-k}{4} \mathbf{1}_{\{2k\}}(d_{l^1}(\sigma, \pi)) \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}\right) = \pi\right\}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)}, \ldots, Y_{j+h-1}^{(2)}\right) = \sigma\right\}} \\ &= \sum_{k=0}^{3} \sum_{\pi, \sigma \in S_h} \frac{4-k}{4} \mathbf{1}_{\{\sum_{j=0}^{h} |\sigma_j - \pi_j| = 2k\}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}\right) = \pi\right\}} \mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)}, \ldots, Y_{j+h-1}^{(2)}\right) = \sigma\right\}}. \end{aligned}$$

We introduce the following partition of $S_h \times S_h$ into four disjoint sets by following the idea in the proof of Lemma 4.25 to divide S_h into one set that contains all patterns such that their spatial reversions are not elements of this set. We adapt the idea and obtain

$$S_h \times S_h = (S_h^* \times S_h^*) \cup ((S_h \setminus S_h^*) \times (S_h \setminus S_h^*)) \cup (S_h^* \times (S_h \setminus S_h^*)) \cup ((S_h \setminus S_h^*) \times S_h^*)\,.$$

Then, we have

$$\begin{aligned}
&\sum_{(\pi,\sigma)\in S_h\times S_h}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\pi\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\sigma\right\}}\\
=&\sum_{(\pi,\sigma)\in S_h^*\times S_h^*}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\pi\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\sigma\right\}}\\
&+\sum_{(\pi,\sigma)\in\left(S_h\setminus S_h^*\right)\times\left(S_h\setminus S_h^*\right)}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\pi\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\sigma\right\}}\\
&+\sum_{(\pi,\sigma)\in S_h^*\times\left(S_h\setminus S_h^*\right)}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\pi\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\sigma\right\}}\\
&+\sum_{(\pi,\sigma)\in\left(S_h\setminus S_h^*\right)\times S_h^*}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\pi\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\sigma\right\}}
\end{aligned} \tag{4.52}$$

Note that if for $(\pi,\sigma)\in S_h\times S_h$, it is fulfilled that $\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k$ for $k=0,\ldots,3$, the same l^1-distance also holds for $(\mathcal{S}(\pi),\mathcal{S}(\sigma))$, because $\mathcal{S}(\pi)=(\pi_h,\ldots,\pi_0)$. Concerning the first order Hermite coefficients, we have for $i=1,\ldots,2h$

$$\begin{aligned}
&\mathbb{E}\left(Y_{j,h}^{(i)}\sum_{(\pi,\sigma)\in S_h^*\times S_h^*}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\pi\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\sigma\right\}}\right)\\
=&\mathbb{E}\Bigg(-Y_{j,h}^{(i)}\sum_{(\pi,\sigma)\in S_h^*\times S_h^*}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y_j^{(1)},\ldots,-Y_{j+h-1}^{(1)}\right)=\pi\right\}}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\mathbf{1}_{\left\{\tilde{\Pi}\left(-Y_j^{(2)},\ldots,-Y_{j+h-1}^{(2)}\right)=\sigma\right\}}\Bigg)\\
=&-\mathbb{E}\Bigg(Y_{j,h}^{(i)}\sum_{(\pi,\sigma)\in S_h^*\times S_h^*}\sum_{k=0}^{3}\frac{4-k}{4}\mathbf{1}_{\left\{\sum_{j=0}^{h}|\sigma_j-\pi_j|=2k\right\}}\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(1)},\ldots,Y_{j+h-1}^{(1)}\right)=\mathcal{S}(\pi)\right\}}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\mathbf{1}_{\left\{\tilde{\Pi}\left(Y_j^{(2)},\ldots,Y_{j+h-1}^{(2)}\right)=\mathcal{S}(\sigma)\right\}}\Bigg),
\end{aligned}$$

for the first two summands in (4.52). A similar result is derived for the last two summands, and, hence, we proved that $m\left(f^{\text{awopd}},\Sigma_{2,h}\right)\geq 2$. Following Lemma 2.33, we have to show that $m\left(f^{\text{awopd}},I_{2,h}\right)\leq 2$. The proof works analogously to the second part of the proof of Lemma 4.25. □

We are now able to provide the asymptotic distribution of $\hat{p}_n^{\text{awopd}}$ for long-range dependence with $d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$, as well as for transformed short-range dependence with $d^* \in \left(0, \frac{1}{4}\right)$. We start with the first case in the following corollary, which is, like Theorem 4.26, an immediate application of Corollary 3.11. This case includes pure long-range dependence with both parameters $d_1, d_2 \in \left(\frac{1}{4}, \frac{1}{2}\right)$ as well as a mixed setting, where only one parameter falls in this range.

Corollary 4.70 *Under the assumptions in **(L2s)** in Overview 4.24, it holds*

$$n^{1-2d^*}(C_2)^{-\frac{1}{2}}\left(\hat{p}_n^{\text{awopd}} - p^{\text{awopd}}\right) \xrightarrow{\mathcal{D}} \sum_{p,q \in P^*} \tilde{\alpha}^{(p,q)} Z_{2,d^*+1/2}^{(p,q)}(1), \tag{4.53}$$

with $Z_{2,d^*+1/2}^{(p,q)}(1)$ as given in Theorem 3.10 and $C_2 := \frac{1}{2d^*(4d^*-1)}$ being a normalizing constant. We have $\tilde{\alpha}^{(p,q)} := \sum_{i,k=1}^{h} \alpha_{i,k}^{(p,q)}$, where $\alpha_{i,k}^{(p,q)} = \alpha_{i+(p-1)h,k+(q-1)h}$ for each $p, q \in P^*$ and $i, k = 1, \ldots, h$ and $(\alpha_{i,k})_{1 \le i,k \le dh} = \Sigma_{2,h}^{-1} C \Sigma_{2,h}^{-1}$ where

$$C = (c_{i,k})_{1 \le i,k \le 2h} = \mathbb{E}\left(Y_{1,h}\left(f^{\text{awopd}}(Y_{1,h}) - p^{\text{awopd}}\right) Y_{1,h}^t\right)$$

denotes the matrix of second order Hermite coefficients.

The second setting we turn to, deals i.a. with the case that at least one of the orginal processes is long-range dependent, however, by transformation with the estimator, the long-range dependence is lost and short-range dependence arises. In Section 4.3.2 this setting is discussed for standard ordinal pattern dependence and the following result works analogously to Theorem 4.35.

Corollary 4.71 *Under the assumptions in **(S)** in Overview 4.24, it holds*

$$n^{\frac{1}{2}}\left(\hat{p}_n^{\text{awopd}} - p^{\text{awopd}}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2\right),$$

with

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \mathbb{E}\left[\left(f^{\text{awopd}}(Y_{1,h}) - \mathbb{E}\left(f^{\text{awopd}}(Y_{1,h})\right)\right)\left(f^{\text{awopd}}(Y_{1+k,h}) - \mathbb{E}\left(f^{\text{awopd}}(Y_{1+k,h})\right)\right)\right].$$

We want to obtain a limit theorem for the estimator $\widehat{AWOPD}$. In general, the argumentation in Section 4.3.3 and Section 4.3.4 can be used to obtain asymptotic results for $\hat{p}_n^{\text{awopd}}$ and for the estimator of the adapted case of hypothetical independence. Note that $\hat{q}_n^{\text{awopd}}$ was already introduced as plug-in estimator in terms of the improved estimators of ordinal pattern probabilities $\hat{q}_{Y^{(1)},n,\text{imp}}(\pi)$ and $\hat{q}_{Y^{(2)},n,\text{imp}}(\pi)$. This has a practical background. As it has turned out in Section 4.3.4, this approach yields an asymptotic distribution that takes the joint consideration of the pattern in the two time series (here given in $\hat{p}_n^{\text{awopd}}$), as well as the hypothetical case of independence (here given in $\hat{q}_n^{\text{awopd}}$) into account and, therefore, is most interesting for applications. Hence, we obtain the following two results for $\widehat{AWOPD}$, adapting Theorem 4.47 in the case of long-range dependence and Theorem 4.49 in the case of transformed short-range dependence.

Corollary 4.72 *Under the assumptions of **(L2s)** in Overview 4.24, the following statements are true.*

(i) If $d_1 = d_2 = d^* \in \left(\frac{1}{4}, \frac{1}{2}\right)$*, it holds that*

$$n^{1-2d^*} (C_2)^{1/2} \left(\widehat{AWOPD}_n - AWOPD\right)$$
$$\xrightarrow{\mathcal{D}} \left(\tilde{\alpha}^{(1,1)} - \sum_{\pi,\sigma \in S_h} w(d(\pi,\sigma)) q_{Y^{(2)}}(\sigma) \tilde{\alpha}_\pi^{(1,1)}\right) Z_{2,d^*+\frac{1}{2}}^{(1,1)}(1)$$
$$+ \left(\tilde{\alpha}^{(2,2)} - \sum_{\pi,\sigma \in S_h} w(d(\pi,\sigma)) q_{Y^{(1)}}(\sigma) \tilde{\alpha}_\pi^{(2,2)}\right) Z_{2,d^*+\frac{1}{2}}^{(2,2)}(1)$$
$$+ 2\tilde{\alpha}^{(1,2)} Z_{2,d^*+\frac{1}{2}}^{(1,2)}(1).$$

(ii) if $d_1 \neq d_2$ *and* $d_p = d^*$ *for exactly one* $p = 1, 2$*, we obtain*

$$n^{1-2d^*} (C_2)^{1/2} L_{pp}^{-1} \left(\widehat{AWOPD}_n - AWOPD\right)$$
$$\xrightarrow{\mathcal{D}} (2C_2)^{\frac{1}{2}} \left(\tilde{\alpha}^{(p,p)} - \sum_{\pi,\sigma \in S_h} w(d(\pi,\sigma)) q_{Y^{3-p}}(\sigma) \tilde{\alpha}_\pi^{(p,p)}\right) Z_{2,d^*+\frac{1}{2}}^{(p,p)}(1),$$

with $\tilde{\alpha}^{(p,q)}$*,* $Z_{2,d^*+1/2}^{(p,q)}(1)$*,* $p, q = 1, 2$*, as given in Theorem 4.26 and* $\tilde{\alpha}_\pi^{(p,p)}$*,* $p = 1, 2$*, as given in Theorem 4.18.*

Proof. We have joint convergence of appropriately normalized

$$\left(\hat{p}_n^{\text{awopd}}, \left(\hat{q}_{Y^{(1)},n}(\pi)\right)_{\pi \in S_h}, \left(\hat{q}_{Y^{(2)},n}(\pi)\right)_{\pi \in S_h}\right).$$

For details, see proof of Theorem 4.47. We modify the considered function f here, as proposed in [58], Theorem 6, by

$$f\left(u, (v_\pi)_{\pi \in S_h}, (v_\pi)_{\pi \in S_h}\right) = u - \sum_{\pi,\sigma} w(d(\pi,\sigma)) v_\pi w_\sigma.$$

The gradient $\nabla\left(p^{\text{awopd}}, (q_{Y^{(1)}}(\pi))_{\pi \in S_h}, (q_{Y^{(2)}}(\pi))_{\pi \in S_h}\right)$ is given by

$$\nabla\left(p^{\text{awopd}}, (q_{Y^{(1)}}(\pi))_{\pi \in S_h}, (q_{Y^{(2)}}(\pi))_{\pi \in S_h}\right)$$
$$= \left(1, -\left(\sum_{\sigma \in S_h} w(d(\pi,\sigma)) q_{Y^{(2)}}(\sigma)\right)_{\pi \in S_h}, -\left(\sum_{\sigma \in S_h} w(d(\pi,\sigma)) q_{Y^{(1)}}(\sigma)\right)_{\pi \in S_h}\right).$$

By applying the multivariate delta method and similar calculations as in the proof of Theorem 4.47 the results are obtained. □

What remains open is to derive a result for transformed short-range dependence. By using the same argumentation as in the proof of Theorem 4.49, we close this gap.

Corollary 4.73 *We define*

$$\sigma^2_{awopd} := \nabla\left(p^{\text{awopd}}, \left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right) \times \Sigma \left(\nabla\left(p^{\text{awopd}}, \left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right)\right)^t,$$

where

$$\nabla\left(p^{\text{awopd}}, \left(q_{Y^{(1)}}(\pi)\right)_{\pi\in S_h}, \left(q_{Y^{(2)}}(\pi)\right)_{\pi\in S_h}\right) = \left(1, -\left(\sum_{\sigma\in S_h} w(d(\pi,\sigma)) q_{Y^{(2)}}(\sigma)\right)_{\pi\in S_h}, -\left(\sum_{\sigma\in S_h} w(d(\pi,\sigma)) q_{Y^{(1)}}(\sigma)\right)_{\pi\in S_h}\right).$$

The matrix $\Sigma = (\sigma_{r,s})_{r,s=1,\ldots,2(h+1)!+1}$ *is given by*

$$\sigma_{r,s} = \sum_{l_1,\ldots,l_{2h},m_1,\ldots,m_{2h}=2}^{\infty} c_{l_1,\ldots,l_{2h},r} c_{m_1,\ldots,m_{2h},s} \lim_{n\to\infty} n^{-1} \sum_{j_1,j_2=1}^{n} \mathbb{E}\left(\prod_{p=1}^{2h} H_{l_p}\left(Y^{(p)}_{j_1,h}\right) H_{m_p}\left(Y^{(p)}_{j_2,h}\right)\right)$$

with

$$c_{l_1,\ldots,l_{2h},r} := \mathbb{E}\left(f_r\left(Y_{j,h}\right) \prod_{k=1}^{2h} H_{l_k}\left(Y^{(k)}_{j,h}\right)\right),$$

where we define

$$\left(f_1,\ldots,f_{2h+1)!+1}\right)^t := \left(\hat{p}^{\text{awopd}}_n - p^{\text{awopd}}, \left(\hat{q}_{Y^{(1)},imp,n} - q_{Y^{(1)}}\right)_{\pi\in S_h}, \left(\hat{q}_{Y^{(2)},imp,n} - q_{Y^{(2)}}\right)_{\pi\in S_h}\right)^t.$$

*Then, under the assumptions in **(S)** in Overview 4.24, it holds that*

$$n^{\frac{1}{2}}\left(\widehat{AWOPD} - AWOPD\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2_{awopd}\right),$$

with

We close this chapter, which deals with ordinal analysis and, hence, represents a main part of this thesis, with a summary of the results. After providing the necessary background and basic information on ordinal patterns in Section 4.1, the two estimators $\hat{q}_n(\pi)$ and $\hat{q}_{n,\text{imp}}(\pi)$ of ordinal pattern probabilities were investigated in Section 4.2. As ordinal pattern probabilities rely on a univariate Gaussian process, we applied the results in Chapter 3 for $d = 1$ and derived limit theorems in the case of long-range dependence. Therefore, we were able to complement the results of [60], Theorem 3.

In Section 4.3, we turned to ordinal pattern dependence as an approach of measuring dependence between two time series based on ordinal information. This method relies on a bivariate Gaussian process and hence, we employed the results in Chapter 3 for $d = 2$. We considered mixed cases of short- and long-range dependent components in the bivariate time series and built an extensive and detailed theoretical framework for standard ordinal pattern dependence. In Section 4.3.1 and 4.3.2, we investigated the limit distribution of the estimator $\hat{p}_n$ of the cruicial parameter p of ordinal pattern dependence. In Section 4.3.1, we proved a limit theorem that allows for pure long-range dependence as well as for mixed cases of short- and long range dependence. The limit variance of the asymptotic distribution was studied in details and we derived possibilities to calculate it analytically for $h = 1$ and presented numerical solutions for $h = 2$. With the help of the symmetry properties of the underlying multivariate Gaussian distribution, it is possible to save much computational cost and to provide an efficient way to determine the quantities of interest. In Section 4.3.2, the case of short-range dependence concerning the parameter p was tackled. A detailed elaboration on the asymptotic properties of three different estimators of the parameter q was given in Section 4.3.3. Finally, in Section 4.3.4, we put the previous results together to obtain a general asymptotic result for estimators of ordinal pattern dependence. Several challenges had to be faced in the course of this last section, mainly occuring due to the necessary joint convergence of certain parameter estimators from the preceeding sections. In the end, all problems were solved, such that it was possible to obtain asymptotic results for estimators of ordinal pattern dependence. In the short range dependent setting, the mathematical theory was supplemented by accessing graph theory and the diagram formula, whereas the case of long-range dependence and the mixed cases of short- and long-range dependence relied on the limit theorems derived in Chapter 3. Finally, a simulation study was provided that illustrated the theoretical results.

In Section 4.4, we introduced possibilities to modify and generalize standard ordinal pattern dependence and the method of estimating this dependence. First, we reduced standard ordinal pattern dependence to the condition that only one fixed pattern occuring in both time series is taken into consideration, as discussed in Section 4.4.1. We benefitted from the detailed calculations on the asymptotic variance in Section 4.3.1, as many considerations could be adopted here. An application of one estimator of ordinal pattern dependence for one fixed pattern was given in Section 4.4.2, to extend the Hurst parameter estimation given in [60], Section 3, to vector-valued fractional Gaussian noise. Moreover, we mentioned the theoretical setting, that we are no longer assuming the increment process of the time series in which we are comparing the coincident patterns to be stationary and fulfill certain dependence conditions. Instead we compare the coincident patterns in the stationary and possibly long- and/or short-range dependent processes themselves. This yields a degenerated limit distribution if we apply the limit theorems arising in the context of long-range dependence, since ordinal patterns are uniquely determined by their increments. This class of functionals was already discussed in Section 3.3, and based on the argumentation there, a detailed discussion on that topic was given. Furthermore, a solution to a non-degenerated limit distribution with a different normalization was provided.
Finally, time-shifted ordinal pattern dependence was presented as a possibility to detect the maximal ordinal pattern dependence between two time series shifted over time. Blockwise estimation of ordinal pattern dependence modifies the estimation method of standard ordinal pattern dependence. Instead of considering stepwise moving windows as it naturally appears in relative frequency estimators, we investigate blockwise estimation, that is, taking disjoint moving windows, such that the increments under consideration never have impact on more than one ordinal pattern. For time-shifted ordinal pattern dependence, as well as for the blockwise

estimation, we had to do a careful preprocessing. The limit theorems to derive asymptotic results for estimators of standard ordinal pattern dependence could not be applied immediately. Nevertheless, we obtained the limit distribution in both modified cases.
In the last section, average-weighted ordinal pattern dependence was introduced as a generalization of standard ordinal pattern dependence, allowing for a certain deviation between the patterns of the two respective time series in the same moving window. After discussing the properties of the weight function and the (pseudo-)metric and providing illustrative examples on how to choose the latter one, we obtained the asymptotic distribution for this extension of ordinal pattern dependence for the long-range dependent case as well as for mixed cases.

5 Ordinal pattern dependence in contrast to other measures of dependence

In the previous sections ordinal pattern dependence was investigated in various ways and modifications, with focus on the asymptotic distribution for short- or long-range dependent Gaussian processes. A natural question that arises is, whether it is possible to integrate the approach of ordinal pattern dependence in the class of (multivariate) dependence measures. In this context, a comparison to other classical statistical dependence measures is also of strong interest. We structure this chapter as follows: We begin with a pilot study, where we compare standard ordinal pattern dependence as well as average-weighted ordinal pattern dependence to one dimensional dependence measures as the classical Pearson's correlation, Spearman's ρ, and Kendall's τ. In order to get a more intuitive understanding of the differences between these measures, we modify the underlying bivariate processes by small disturbances, which are further explained in the following. Note that this first section is based on [49] and is joint work with Alexander Schnurr, where I was responsible for coding and analyzing the pilot study.
The second part of this chapter deals with the mathematical theory behind these ideas. We investigate the connection of ordinal pattern dependence and multivariate dependence measures as defined in [29] and [55] and provide a dissociation from the one-dimensional dependence measures. We keep in mind, that multivariate dependence measures are usually used to detect dependencies between two random vectors or, in some cases, even within one multivariate random vector. On the contrary, ordinal pattern dependence is defined in a time series context to shed light on the intensity of comonotonic behaviour of two time series. The results of this section can be found in [11] and are joint work with Annika Betken, Herold Dehling and Alexander Schnurr. My main contributions in this work can be found in there in Section 5, which deals with the comparison of ordinal pattern dependence to other measures and are provided in the proofs.
Note that we leave the setting of short-/long-range dependence here and in general only assume stationarity, if not declared differently, to guarantee a reasonable estimation of ordinal pattern dependence.

5.1 Approaches of ordinal pattern dependence in comparison to other classical dependence measures: a pilot study

We start with the empirical comparison. We consider 11 data sets, of which 10 are simulated and one is a real world example. Each data set consists of two time series between which we are estimating the respective dependence measures.
The first five data sets contain an underyling bivariate white noise process and we consider

$n = 1000$ data points, so in the standard setting a1) there are two independent white noise processes $\left(X_j^{(1)}\right)_{j=1,\ldots,1000}$, $\left(X_j^{(2)}\right)_{j=1,\ldots,1000}$. Roughly speaking, they represent the archetype of independent time series. They are displayed in Figure 5.1.

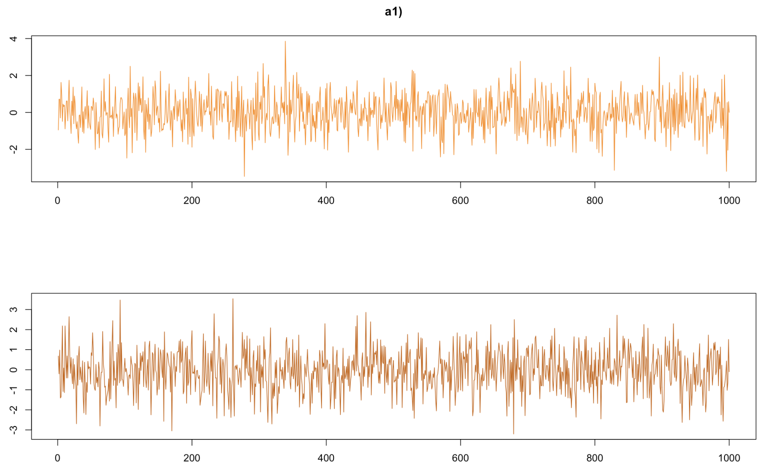

Figure 5.1: Plot of a1): Two independent white noise processes with $n = 1000$ data points.

We modify these two processes by the following disturbances:

- a2): Superposition of setting a1) and 10 shocks with a height of $\pm$ 10, happening randomly between 1 and 1000 with 5 shocks having a negative sign and 5 shocks having a positive sign.
- a3): In both processes happens one large shock of height 1000 at point 500.
- a4): We caused a structural break in the expected value at point 500 of both processes and raised the expected value by 100.
- a5): We generated a structural break in the variance at point 500 and raised the variance by 100.

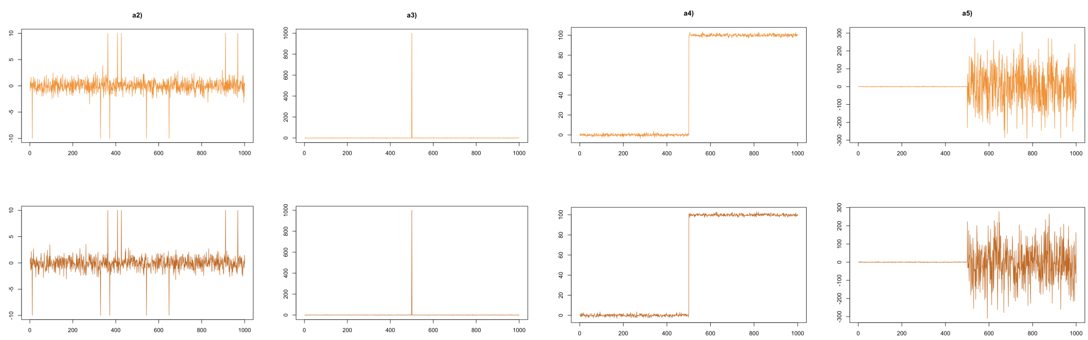

Figure 5.2: Plots of modifications a2), a3), a4) and a5).

All modifications described above are illustrated in Figure 5.2.

The other five simulated data sets are disturbed in the same way as characterized above, but with different underlying processes, namely two correlated AR(1)-processes $\left(X_j^{*(1)}\right)$ and $\left(X_j^{*(2)}\right)$, generated in the following way:

$$X_j^{*(1)} = \psi X_{j-1}^{*(1)} + N_j$$
$$X_j^{*(2)} = \phi X_{j-1}^{*(2)} + M_j,$$

where we choose $\psi = 0.8$, $\phi = 0.1$. The correlation is generated by sampling a bivariate normal distribution with covariance matrix $\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$ and storing the first column of the sample in N and the second one in M. As in a1), we use a sample size of $n = 1000$. A simulation of setting b1) is provided in Figure 5.3.

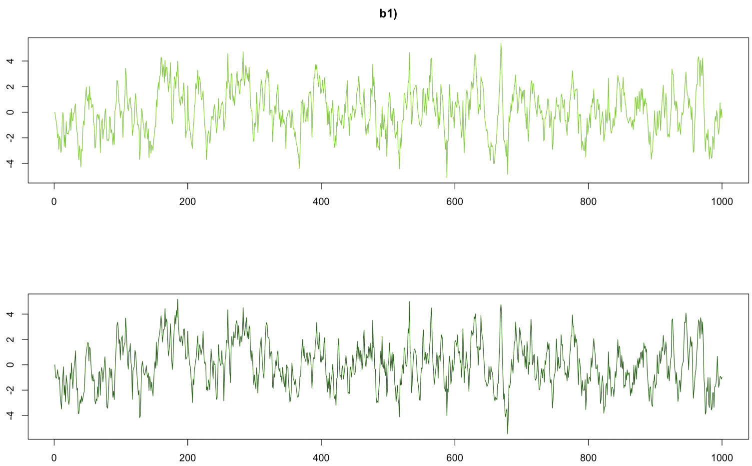

Figure 5.3: Plot of b1): Two correlated AR(1)-processes with $n = 1000$ data points.

As indicated above, the disturbances of b1) denoted by b2) to b5) are obtained in the same way as a2) to a5) described in the listing above and are illustrated in Figure 5.4.

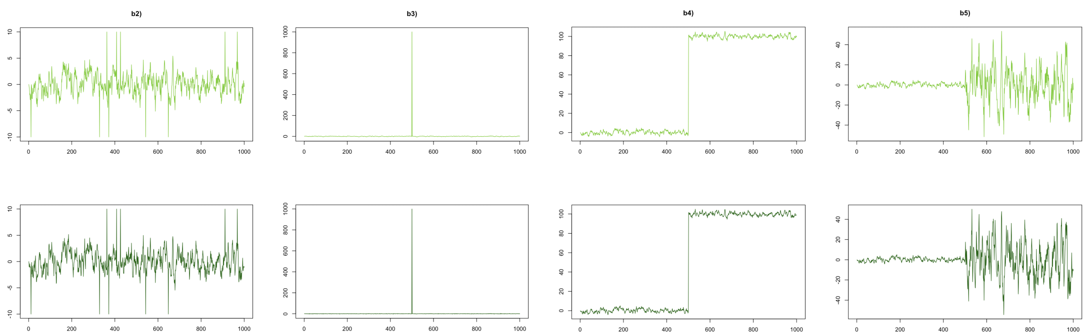

Figure 5.4: Plots of modifications b2), b3), b4) and b5).

We have introduced all 10 simulated data sets now. The last data set we consider is a classical example of mathematical finance, namely the well-known S&P 500 (SPX) and the corresponding Chicago Board Options Exchange Volatility Index (VIX), see Figure 5.5. We analyze close prices from the 20th of Febuary 2013 until 07th of Febuary 2017, which are 1000 data points in total. The data was extracted from finance.yahoo.com.

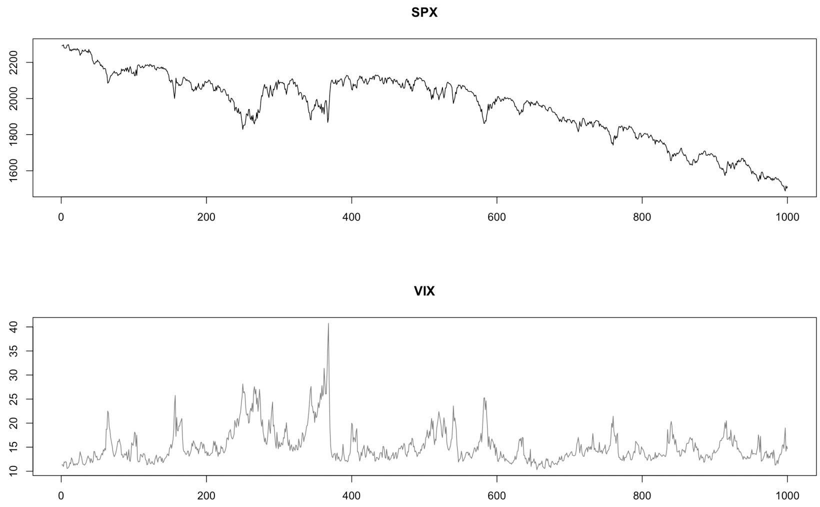

Figure 5.5: Plots of SPX and VIX from 02/20/2013 to 02/07/2017 ($n = 1000$ data points).

As one can already see, regarding Figure 5.5, the two time series show anti-monotonic behaviour. As it is discussed in [57], Section 3, in the standard literature they are known to have a negative correlation. However, in [67], p.7, it is stated, that there is an asymmetric relation between rates of change in the VIX and the SPX. Therefore, due to non-linearity this is not displayable by correlation. Hence, this yields a nice example to test the potential of ordinal pattern dependence. Note that we consider -VIX and, therefore, positive dependence in order to get comparable results to the other 11 settings.

Now we turn to the dependence measures we want to compare, based on the 11 datasets described before. Our new approaches are standard ordinal pattern dependence as introduced in (4.13) and average-weighted ordinal pattern dependence as introduced in Section 4.4.6. We denote the corresponding values of the two dependence measures as *standardized ordinal pattern coefficient* (SOPC) and *standardized average-weighted ordinal pattern coefficient* (SOPC(AW)). The well-known dependence measures we take into consideration are given by Pearson's correlation coefficient, Spearman's ρ and Kendall's τ. The results we obtain are given in Figure 5.6.

We recognize that in both settings a1) and b1) the dependence structure we generated is pictured. It is also possible to observe that the correlation is the most sensitive dependence measure concerning the created disturbances and, therefore, significantly increases in the settings a2) to a4) and b2) to b4). The other two classical dependence measures are not biased by the two "external shock" settings but do rise in the setting with the structural break in the expected value, a4) and b4). A nice but also expectable result is that neither SOPC nor SOPC(AW) is affected by the disturbances. The intuition why this is expectable is given in the definition of ordinal pattern dependence. Since we are ignoring the absolute values of the data considered and only

	correlation	Spearman's rho	Kendall's tau	SOPC	SOPC(AW)
a1)	0.01	0.02	0.01	0.00	0.00
a2)	0.51	0.05	0.04	0.01	0.00
a3)	1.00	0.02	0.02	0.00	0.00
a4)	1.00	0.75	0.50	0.00	0.00
a5)	-0.01	0.04	0.03	0.00	0.00
b1)	0.50	0.47	0.32	0.28	0.17
b2)	0.68	0.48	0.34	0.29	0.18
b3)	1.00	0.47	0.33	0.28	0.17
b4)	1.00	0.87	0.66	0.28	0.17
b5)	0.52	0.40	0.31	0.28	0.17
(SPX,-VIX)	0.11	0.20	0.16	0.50	0.44

Figure 5.6: Empirical dependence measures for 11 datasets.

take the ordinal structure into account, it does not matter, whether the largest point in a certain time window has only a small Euclidean distance to the second largest point or a very large one. As it is possible to observe in this explorative study, ordinal pattern dependence is extremely robust concerning several types of outliers in the considered data sets. A comprehensive way to visualize how the respective modifications of the data sets impact on the dependence measures is given in Figure 5.7. Note that the differences between the values of SOPC and SOPC(AW) can be explained, since we use $h = 2$ (SOPC) and $h = 5$ (SOPC(AW)). One might wonder why we do not choose the same values here. The reason is, that for small values of h, almost every combination of pattern would be taken into account using average-weighted ordinal pattern dependence, see Example 4.68, but for larger h the standard ordinal pattern dependence would yield very small values and underestimate the dependence. Therefore, in pratice, the user has to decide which approach seems to be more reasonable to her or him. Regarding the SPX and

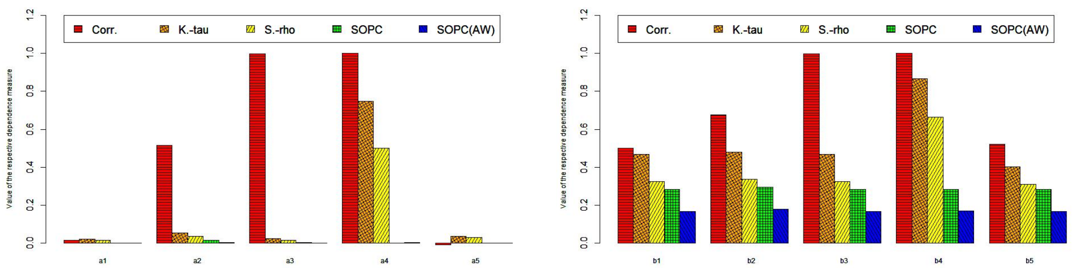

Figure 5.7: Empirical dependence in the settings a1) to a5) (left) and b1) to b5) (right).

VIX, we observe a small correlation, which seems to be negligible, but a strong (negative) ordinal pattern dependence. This strengthens our hope that ordinal pattern dependence is a better approach to measure dependence for data sets exhibiting asymmetric and, therefore, non-linear dependence. However, note that the correlation between the increments of SPX and VIX is much larger.

We close this explorative study with the summary that ordinal pattern dependence as well as average-weighted ordinal pattern dependence exhibits robustness concerning outliers in data

sets and are less sensitive than classical dependence measures like correlation, Spearman's ρ, and Kendall's τ. This yields a great advantage if one is interested in omitting and not further investigating these outliers. One field of application that might benefit is medicine, e.g. the analysis of EEG data, since they can possibly be biased by eye movements.
There are several questions that arise from this study. To start with, there is certainly a need to an explanation of the differences between the classical measures and ordinal pattern dependence. The role of the increments, which seem to play an important role as the SPX-VIX-example indicates, should be investigated, too. In the next chapter, we capture these empirically observed phenomenons more theoretically and seek answers to these questions.

5.2 Comparison: a theoretical approach

In this section, we try to provide answers to the questions that are open regarding the pilot study. First, we investigate, whether it is possible to integrate standard ordinal pattern dependence in the class of multivariate dependence measures as defined in [29], Chapter 3. Subsequently, we compare standard ordinal pattern dependence to other uni- and multivariate dependence measures, show differences and give relations, if possible. The results of this section can be found in [11], with focus on Chapter 5.

First, note that there is a significant difference in the approaches of classical dependence measures and ordinal pattern dependence. Usually Pearson's correlation, Spearman's ρ or Kendall's τ have been developed to describe the behaviour of two one-dimensional random variables. The application to stationary time series boils down to the estimation of the dependence between two random variables at the same points in time and, therefore, does not imply any kind of dynamical dependence, which is crucial in the following. Analogously, the multivariate generalization of those dependence measures deal with the relation between two multivariate random vectors and are not defined in a time series context in general. In contrast, ordinal pattern dependence describes the co-movement of two (ordinal pattern-)stationary time series and does not have an intuitive meaning concerning componentwise dependencies within the multivariate vectors considered in $p = \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right)\right)$. Therefore, we need to take a careful look at what we are comparing and adapt the definitions if necessary.

In the course of this chapter we refer to OPD_h for the value of ordinal pattern dependence defined in (4.13) to emphasize the h increments under consideration. Moreover, we define

$$X_j^{(1,h+1)} := \left(X_j^{(1)}, X_{j+1}^{(1)}, \ldots, X_{j+h}^{(1)}\right)$$

and

$$X_j^{(2,h+1)} := \left(X_j^{(2)}, X_{j+1}^{(2)}, \ldots, X_{j+h}^{(2)}\right).$$

To keep the notation simple, we also fix $X^{(1,h+1)} := X_0^{(1,h+1)}$, such that $X^{(1,h+1)} = \left(X_0^{(1)}, \ldots, X_h^{(1)}\right)^t$ and $X^{(2,h+1)} := X_0^{(2,h+1)}$ with $X^{(2,h+1)} = \left(X_0^{(2)}, \ldots, X_h^{(2)}\right)^t$. The corresponding correlation matrices are denoted by $\Sigma_{1,h+1}$ and $\tilde{\Sigma}_{1,h+1}$, and the cross-covariance matrix is given by $\Sigma_{X^{(1,h+1)},X^{(2,h+1)}}$, such that we have the following relation concerning the already introduced notation in this work:

$$X_{2,h+1} = \left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t, \quad \Sigma_{2,h+1} = \begin{pmatrix} \Sigma_{1,h+1} & \Sigma_{X^{(1,h+1)},X^{(2,h+1)}} \\ \Sigma_{X^{(1,h+1)},X^{(2,h+1)}} & \tilde{\Sigma}_{1,h+1} \end{pmatrix}$$

Remark that we leave the setting of long- and short-range dependence here, as already mentioned in the introduction of this chapter. When dealing with ordinal pattern dependence, for simplicity, we assume stationarity of the underlying bivariate process. In this context, let us recall that stationarity implies ordinal pattern stationarity.
We start with the result dealing with the integration of ordinal pattern dependence in the class of multivariate dependence measures, as introduced in [29], Section 3, p.98. Therefore, we introduce the definition of a multivariate dependence measure first.

Definition 5.1 *(Multivariate dependence measure)*
Let L_0 denote the space of random vectors with values in $\mathbb{R}^d$ on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We call a function $\mu : L_0 \times L_0 \to \mathbb{R}$ an d-dimensional measure of dependence, if

(i) it takes values in $[-1, 1]$,

(ii) it is invariant with respect to simultaneous permutations of the components within two random vectors X and Y,

(iii) it is invariant with respect to increasing transformations of the components within two random vectors X and Y,

(iv) it is zero for two independent random vectors X and Y,

(v) it respects concordance ordering, i.e., for two pairs of random vectors X, Y and X^*, Y^*, it holds that

$$\begin{pmatrix} X \\ Y \end{pmatrix} \preccurlyeq_C \begin{pmatrix} X^* \\ Y^* \end{pmatrix} \Rightarrow \mu(X, Y) \leq \mu(X^*, Y^*).$$

Here, $\preccurlyeq_C$ denotes concordance ordering, i.e.,

$$\begin{pmatrix} X \\ Y \end{pmatrix} \preccurlyeq_C \begin{pmatrix} X^* \\ Y^* \end{pmatrix} \text{ if and only if } F_{\binom{X}{Y}} \leq F_{\binom{X^*}{Y^*}} \text{ and } \bar{F}_{\binom{X}{Y}} \leq \bar{F}_{\binom{X^*}{Y^*}},$$

where $\leq$ is meant pointwise and $\bar{F}$ denotes the survival function.

Theorem 5.2 *Ordinal pattern dependence as defined in* (4.13) *is an* $(h+1)$*-dimensional measure of dependence.*

The proof is given in in [11], proof of Theorem 2.3.

This is an important result for the general standing of ordinal pattern dependence and for the comparability to other measures of dependence within the same class. We now turn to the introduction of these other measures, which are studied in the following. We start with the multivariate extension of Pearson's ρ, introduced in [55], p. 2. We modify the definition of multivariate Pearson's ρ and later on of multivariate Kendall's τ in terms of two $(h+1)$-dimensional random vectors to get comparable results in the time series approach.

Definition 5.3 *For two random vectors $X^{(1,h+1)}$, $X^{(2,h+1)} \in L_2\left(\mathbb{R}^{h+1}\right)$ with invertible correlation matrices $\Sigma_{1,h+1}$ and $\tilde{\Sigma}_{1,h+1}$ and cross-covariance matrix $\Sigma_{X^{(1,h+1)},X^{(2,h+1)}}$, we define Pearson's correlation coefficient by*

$$\rho_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) := \frac{\operatorname{tr}\left(\Sigma_{X^{(1,h+1)},X^{(2,h+1)}}\right)}{\operatorname{tr}\left(\left(\Sigma_{1,h+1}\tilde{\Sigma}_{1,h+1}\right)^{1/2}\right)},$$

where $A^{1/2}$ is the principal square root of the matrix A, such that $A^{1/2}A^{1/2} = A$. Furthermore, $\operatorname{tr}(A)$ *denotes the trace of a matrix A.*

The main issue of the univariate dependence measures considered before, is that they do not take the cross-dependencies of the considered random vectors into account, here displayed by $\operatorname{Cor}\left(X_0^{(1)}, X_k^{(2)}\right)$ and $\operatorname{Cor}\left(X_k^{(1)}, X_0^{(2)}\right)$, $k = 0, \ldots, h+1$. They represent the *dynamical* dependence between the two random vectors arising from the two time series considered, and they are included in ordinal pattern dependence if $h \geq 2$. To classify this into the time series context, univariate dependence measures concentrate on the dependence between $\left(X_j^{(1)}, X_j^{(2)}\right)_{j\in\mathbb{Z}}$ only at the same points in time. For the multivariate generalization of Pearson's rho we obtain

$$\begin{aligned}\rho_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) :&= \frac{\operatorname{tr}\left(\Sigma_{X^{(1,h+1)},X^{(2,h+1)}}\right)}{\operatorname{tr}\left(\left(\Sigma_{1,h+1}\tilde{\Sigma}_{1,h+1}\right)^{1/2}\right)} \\ &= \frac{\operatorname{Cov}\left(X_0^{(1)}, X_0^{(2)}\right) + \ldots + \operatorname{Cov}\left(X_{h+1}^{(1)}, X_{h+1}^{(2)}\right)}{\operatorname{tr}\left(\left(\Sigma_{1,h+1}\tilde{\Sigma}_{1,h+1}\right)^{1/2}\right)}.\end{aligned}$$

Note that in the case that $\left(X_j^{(1)}, X_j^{(2)}\right)_{j\in\mathbb{Z}}$ is stationary and $\Sigma_{1,h+1} = \tilde{\Sigma}_{1,h+1}$, this even yields

$$\rho_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) = \operatorname{Cov}\left(X_0^{(1)}, X_0^{(2)}\right).$$

In both cases, we see that the cross-correlations $r^{(1,2)}(l)$, $l \geq 1$ have no impact on the value of the multivariate ρ_h. This gets explicitly clear in the case where the underlying time series is stationary, since we can easily determine the numerator in terms of $\operatorname{Cov}\left(X_0^{(1)}, X_0^{(2)}\right)$ then. However, even if the time series is not stationary, the numerator only depends on the covariances at equal points in time and therefore, this measure is not appropriate for our approach. Hence, we focus on the multivariate generalization of Kendall's τ, as it is provided in [29], p. 100.

Definition 5.4 *For two random vectors $X^{(1,h+1)}$ and $X^{(2,h+1)}$, we define the multivariate extension of Kendall's τ by*

$$\tau_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) := \operatorname{Cor}\left(\mathbf{1}_{\left\{X^{(1,h+1)}\leq\tilde{X}^{(1,h+1)}\right\}}, \mathbf{1}_{\left\{X^{(2,h+1)}\leq\tilde{X}^{(2,h+1)}\right\}}\right),$$

where $\left(\tilde{X}^{(1,h+1)}, \tilde{X}^{(2,h+1)}\right)^t$ is an independent copy of $\left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t$.

We see a significant difference between this definition to the one of ordinal pattern dependence. We notice that this definition needs an independent copy of the considered two random vectors.

In contrast, ordinal pattern dependence is regarded in a bivariate time series setting in which independence is usually not guaranteed. However, to close this gap and to get a comparable result of ordinal pattern dependence for this classical definition of a multivariate dependence measure, in Chapter 3 in [11], a limit theorem for the case of indepent copies as given in Definition 5.4 in terms of ordinal pattern dependence is provided.
An estimator of multivariate Kendall's τ is derived using the approach given in [29], p.100. Given an independent copy $\left(\tilde{X}^{(1,h+1)}, \tilde{X}^{(2,h+1)}\right)^t$ of the vector $\left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t$, we have

$$\begin{aligned}\tau_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) =& \frac{p_{\left(X^{(1,h+1)},X^{(2,h+1)}\right)} - p_{X^{(1,h+1)}}p_{X^{(2,h+1)}}}{\sqrt{p_{X^{(1,h+1)}}\left(1-p_{X^{(1,h+1)}}\right)p_{X^{(2,h+1)}}\left(1-p_{X^{(1,h+1)}}\right)}}\\ =&\psi\left(p_{\left(X^{(1,h+1)},X^{(2,h+1)}\right)}, p_{X^{(1,h+1)}}, p_{X^{(2,h+1)}}\right),\end{aligned}$$

where

$$\begin{aligned}p_{\left(X^{(1,h+1)},X^{(2,h+1)}\right)} &:= \mathbb{P}\left(X^{(1,h+1)} \leq \tilde{X}^{(1,h+1)}, X^{(2,h+1)} \leq \tilde{X}^{(2,h+1)}\right),\\ p_{X^{(1,h+1)}} &:= \mathbb{P}\left(X^{(1,h+1)} \leq \tilde{X}(1,h+1)\right)\\ p_{X^{(2,h+1)}} &:= \mathbb{P}\left(X^{(2,h+1)} \leq \tilde{X}(2,h+1)\right),\end{aligned}$$

and where $\psi : \mathbb{R}^3 \to \mathbb{R}$ is defined by

$$\psi(x,y,z) := \frac{z - x\,y}{x(1-x)y(1-y)}.$$

The probabilities $p_{\left(X^{(1,h+1)},X^{(2,h+1)}\right)}$, $p_{X^{(1,h+1)}}$, and $p_{X^{(2,h+1)}}$ can be estimated by their sample analogues, defined by

$$\begin{aligned}\hat{p}_{X^{(1,h+1)},n} &:= \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}\mathbf{1}_{\left\{X_i^{(1,h+1)}\leq X_j^{(1,h+1)}\right\}},\\ \hat{p}_{X^{(2,h+1)},n} &:= \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}\mathbf{1}_{\left\{X_i^{(2,h+1)}\leq X_j^{(2,h+1)}\right\}},\\ \hat{p}_{X^{(1,h+1)},X^{(2,h+1)},n} &:= \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}\mathbf{1}_{\left\{X_i^{(1,h+1)}\leq X_j^{(1,h+1)},X_i^{(2,h+1)}\leq X_j^{(2,h+1)}\right\}},\end{aligned}$$

where $X_j^{(1,h+1)} = \left(X_j^{(1)}, X_{j+1}^{(1)}, \ldots, X_{j+h}^{(1)}\right)$ and $X_j^{(2,h+1)} = \left(X_j^{(2)}, X_{j+1}^{(2)}, \ldots, X_{j+h}^{(2)}\right)$.
Then, the plug-in estimator for Kendall's τ is given by

$$\hat{\tau}_n^{(h)}(X^{(1,h+1)}, X^{(2,h+1)}) := \psi(\hat{p}_{X^{(1,h+1)},n}, \hat{p}_{X^{(2,h+1)},n}, \hat{p}_{(X^{(1,h+1)},X^{(2,h+1)}),n}). \tag{5.1}$$

Section 4.2 of [11] deals with the asymptotic properties of this estimator and a limit theorem is provided in case of short-range dependence.

Remark 5.5 *We briefly introduce the blockwise estimation of multivariate Kendall's τ as an analog approach as given in blockwise estimation of ordinal patterns. This is used in the data analysis later on. Hence, we define*

$$\hat{p}_{X^{(1,h+1)},n,bw} := \frac{2}{\tilde{n}(\tilde{n}-1)}\sum_{1\leq i<j\leq n}\mathbf{1}_{\left\{X_{i(h-1)+1}^{(1,h+1)}\leq X_{j(h-1)+1}^{(1,h+1)}\right\}},$$

$$\hat{p}_{X^{(2,h+1)},n,bw} := \frac{2}{\tilde{n}(\tilde{n}-1)} \sum_{1\leq i<j\leq n} \mathbf{1}_{\left\{X^{(2,h+1)}_{i(h-1)+1} \leq X^{(2,h+1)}_{j(h-1)+1}\right\}},$$

$$\hat{p}_{X^{(1,h+1)},X^{(2,h+1)},n,bw} := \frac{2}{\tilde{n}(\tilde{n}-1)} \sum_{1\leq i<j\leq n} \mathbf{1}_{\left\{X^{(1,h+1)}_{i(h-1)+1} \leq X^{(1,h+1)}_{j(h-1)+1}, X^{(2,h+1)}_{i(h-1)+1} \leq X^{(2,h+1)}_{j(h-1)+1}\right\}},$$

with $\tilde{n} := \left\lfloor \frac{n-1}{h-1} \right\rfloor (h-1) + 1$.
Therefore, we use as blockwise estimator for multivariate Kendall's τ*:*

$$\hat{\tau}^{(h)}_{n,bw}(X^{(1,h+1)}, X^{(2,h+1)}) := \psi(\hat{p}_{X^{(1,h+1)},n,bw}, \hat{p}_{X^{(2,h+1)},n,bw}, \hat{p}_{(X^{(1,h+1)},X^{(2,h+1)}),n,bw}). \tag{5.2}$$

In the following, we focus on the comparison of multivariate Kendall's τ and ordinal pattern dependence. Therefore, we need to create a relation between the definition of multivariate Kendall's τ in Definition 5.4 and the definition of ordinal pattern dependence in the time series context in (4.13) first. In the next lemma we constitute a first expression for a stationary bivariate Gaussian process.

Lemma 5.6 *Let* $\left(X^{(1)}_j, X^{(2)}_j\right)_{j\geq 0}$ *denote a centered stationary Gaussian process, and let* $X^{(1,h+1)} = \left(X^{(1)}_0, \ldots, X^{(1)}_h\right)^t$ *and* $X^{(2,h+1)} = \left(X^{(2)}_0, \ldots, X^{(2)}_h\right)^t$*. Then, we have*

$$\tau_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) = \frac{\mathbb{P}\left(X^{(1)}_0 \leq 0, \ldots, X^{(1)}_h \leq 0, X^{(2)}_0 \leq 0, \ldots, X^{(2)}_h \leq 0\right) - p_{X^{(1,h+1)}} p_{X^{(2,h+1)}}}{\sqrt{p_{X^{(1,h+1)}}\left(1-p_{X^{(1,h+1)}}\right) p_{X^{(2,h+1)}}\left(1-p_{X^{(2,h+1)}}\right)}},$$

where $p_{X^{(1,h+1)}} := \mathbb{P}\left(X^{(1)}_0 \leq 0, \ldots, X^{(1)}_h \leq 0\right)$ *and* $p_{(X^{(2,h+1)}} := \mathbb{P}\left(X^{(2)}_0 \leq 0, \ldots, X^{(2)}_h \leq 0\right)$.

Proof. We denote by $\left(\tilde{X}^{(1,h+1)}, \tilde{X}^{(2,h+1)}\right)^t$ an independent copy of $\left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t$ as introduced in Definition 5.4. We obtain

$$\begin{aligned}
&\tau_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) \\
&= \operatorname{Cor}\left(\mathbf{1}_{\left\{X^{(1,h+1)} \leq \tilde{X}^{(1,h+1)},\right\}}, \mathbf{1}_{\left\{X^{(2,h+1)} \leq \tilde{X}^{(2,h+1)},\right\}}\right) \\
&= \operatorname{Cor}\left(\mathbf{1}_{\left\{X^{(1,h+1)} - \tilde{X}^{(1,h+1)} \leq 0\right\}}, \mathbf{1}_{\left\{X^{(2,h+1)} - \tilde{X}^{(2,h+1)} \leq 0\right\}}\right) \\
&= \frac{\mathbb{P}\left(X^{(1)}_0 - \tilde{X}^{(1)}_0 \leq 0, \ldots, X^{(1)}_h - \tilde{X}^{(1)}_h \leq 0, X^{(2)}_0 - \tilde{X}^{(2)}_0 \leq 0, \ldots, X^{(2)}_h - \tilde{X}^{(2)}_h \leq 0\right)}{\sqrt{p_{X^{(1,h+1)}}\left(1-p_{X^{(1,h+1)}}\right) p_{X^{(2,h+1)}}\left(1-p_{X^{(2,h+1)}}\right)}} \\
&\quad - \frac{p_{X^{(1,h+1)}} p_{X^{(2,h+1)}}}{\sqrt{p_{X^{(1,h+1)}}\left(1-p_{X^{(1,h+1)}}\right) p_{X^{(2,h+1)}}\left(1-p_{X^{(2,h+1)}}\right)}}
\end{aligned}$$

with

$$p_{X^{(1,h+1)}} = \mathbb{P}\left(X^{(1)}_0 - \tilde{X}^{(1)}_0 \leq 0, \ldots, X^{(1)}_h - \tilde{X}^{(1)}_h \leq 0\right)$$

and

$$p_{X^{(2,h+1)}} = \mathbb{P}\left(X^{(2)}_0 - \tilde{X}^{(2)}_0 \leq 0, \ldots, X^{(2)}_h - \tilde{X}^{(2)}_h \leq 0\right).$$

Note that since we are considering a centered Gaussian process, we have

$$\left(X^{(1,h+1)} - \tilde{X}^{(1,h+1)}, X^{(2,h+1)} - \tilde{X}^{(2,h+1)}\right)^t \overset{\mathcal{D}}{=} \sqrt{2}\left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t$$

which explicitly implies that the cross-correlations within the vector

$$\left(X^{(1,h+1)} - \tilde{X}^{(1,h+1)}, X^{(2,h+1)} - \tilde{X}^{(2,h+1)}\right)^t$$

equal those within $\left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t$. We conclude that

$$\begin{aligned}
&\frac{\mathbb{P}\left(X_0^{(1)} - \tilde{X}_0^{(1)} \leq 0, \ldots, X_h^{(1)} - \tilde{X}_h^{(1)} \leq 0, X_0^{(2)} - \tilde{X}_0^{(2)} \leq 0, \ldots, X_h^{(2)} - \tilde{X}_h^{(2)} \leq 0\right)}{\sqrt{p_{X^{(1,h+1)}}\left(1 - p_{X^{(1,h+1)}}\right) p_{X^{(2,h+1)}}\left(1 - p_{X^{(2,h+1)}}\right)}} \\
&\qquad - \frac{p_{X^{(1,h+1)}} p_{X^{(2,h+1)}}}{\sqrt{p_{X^{(1,h+1)}}\left(1 - p_{X^{(1,h+1)}}\right) p_{X^{(2,h+1)}}\left(1 - p_{X^{(2,h+1)}}\right)}} \\
=&\frac{\mathbb{P}\left(X_0^{(1)} \leq 0, \ldots, X_h^{(1)} \leq 0, X_0^{(2)} \leq 0, \ldots, X_h^{(2)} \leq 0\right) - p_{X^{(1,h+1)}} p_{X^{(2,h+1)}}}{\sqrt{p_{X^{(1,h+1)}}\left(1 - p_{X^{(1,h+1)}}\right) p_{X^{(2,h+1)}}\left(1 - p_{X^{(2,h+1)}}\right)}},
\end{aligned}$$

with $p_{X^{(1,h+1)}} = \mathbb{P}\left(X_0^{(1)} \leq 0, \ldots, X_h^{(1)} \leq 0\right)$ and $p_{X^{(2,h+1)}} = \mathbb{P}\left(X_0^{(2)} \leq 0, \ldots, X_h^{(2)} \leq 0\right)$. □

Although we do not know an analytical expression, neither for $p_{X^{(1,h+1)}}$, $p_{X^{(2,h+1)}}$ nor for the first probability in the enumerator, for $h \geq 2$, we know that orthant probabilities of a multivariate Gaussian distribution are determined by the entries of the correlation matrix, which includes here the cross-correlation matrix of $X^{(1,h+1)}$ and $X^{(2,h+1)}$. In the case studied here, that the multivariate vectors consist of $h+1$ consecutive points of a bivariate *stationary* Gaussian process, we know that the values having an impact are given by $\mathbb{E}\left(X_0^{(1)} X_k^{(2)}\right)$ for $k = -h, \ldots, h$, beside the autocorrelations $\mathbb{E}\left(X_0^{(1)} X_k^{(1)}\right)$ for $k = 0, \ldots, h$ and $\mathbb{E}\left(X_0^{(2)} X_k^{(2)}\right)$ for $k = 0, \ldots, h$.
Therefore, we have found a multivariate dependence measure that considers exactly the cross-dependencies mentioned above and hence takes the desired dynamical dependence into account, in contrast to multivariate Pearson's Rho. A direct comparison to ordinal pattern dependence is nevertheless not constructive here, since

$$\begin{aligned}
&OPD_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) \\
=&\frac{\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right) = \Pi\left(X^{(2,h+1)}\right)\right) - \sum_{\pi \in S_h} \mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right) = \pi\right) \mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right) = \pi\right)}{1 - \sum_{\pi \in S_h} \mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right) = \pi\right) \mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right) = \pi\right)},
\end{aligned}$$

and concentrating on the first summand this yields

$$\begin{aligned}
&\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right) = \Pi\left(X^{(2,h+1)}\right)\right) \\
=&\mathbb{P}\left(X_0^{(1)} \leq X_1^{(1)} \leq \ldots \leq X_h^{(1)}, X_0^{(2)} \leq X_1^{(2)} \leq \ldots \leq X_h^{(2)}\right) \\
&\quad + \mathbb{P}\left(X_0^{(1)} \geq X_1^{(1)} \geq \ldots \geq X_h^{(1)}, X_0^{(2)} \geq X_1^{(2)} \geq \ldots \geq X_h^{(2)}\right) \\
&\quad + \sum_{\pi \in S_h \setminus \{(h, h-1, \ldots, 0), (0, 1, \ldots, h)\}} \mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right) = \Pi\left(X^{(2,h+1)}\right) = \pi\right).
\end{aligned}$$

We observe that the direct ordinal comparison between the $h+1$ consecutive random variables of each random vector, as it is used in the ordinal pattern dependence approach, is not echoed in multivariate Kendall's τ. However, we can prove a relation between ordinal pattern dependence and multivariate Kendall's τ of the increments of the Gaussian process considered, denoted by $Y_j^{(1)} = X_j^{(1)} - X_{j-1}^{(1)}$ and $Y_j^{(2)} = X_j^{(2)} - X_{j-1}^{(2)}$. We explicitly use the existence of a function $\tilde{\Pi}$, such that $\Pi\left(X_0^{(p)}, \ldots, X_h^{(p)}\right) = \tilde{\Pi}\left(Y_1^{(p)}, \ldots, Y_h^{(p)}\right)$, $p = 1, 2$. This implicates that OPD_h is uniquely determined by the increments.

Theorem 5.7 *Let* $\left(X_j^{(1)}, X_j^{(2)}\right)_{j\in\mathbb{Z}}$ *be a bivariate stationary centered Gaussian process. It holds that*

(i)

$$\begin{aligned}
&OPD_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right)\\
&= \frac{2\tau\left(Y_1^{(1)}, \ldots, Y_h^{(1)}, Y_1^{(2)}, \ldots, Y_h^{(2)}\right)\sqrt{\tilde{p}_{Y^{(1)}}\left(1-\tilde{p}_{Y^{(1)}}\right)\tilde{p}_{Y^{(2)}}\left(1-\tilde{p}_{Y^{(2)}}\right)}}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}\\
&\quad+\frac{\sum_{\pi\in S_h\setminus\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}\\
&\quad-\frac{\sum_{\pi\in S_h\setminus\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)},
\end{aligned}$$

with

$$\begin{aligned}
\tilde{p}_{Y^{(1,h)}} &= \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = (0,1,\ldots,h)\right) = \mathbb{P}\left(Y_1^{(1)} \leq 0, \ldots, Y_h^{(1)} \leq 0\right),\\
\tilde{p}_{Y^{(2,h)}} &= \mathbb{P}\left(\Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right) = (0,1,\ldots,h)\right) = \mathbb{P}\left(Y_1^{(2)} \leq 0, \ldots, Y_h^{(2)} \leq 0\right),
\end{aligned}$$

(ii)

$$\begin{aligned}
&OPD_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right)\\
&=\frac{\sum_{\pi\in S_h}\tau\left(X_{\pi_2}^{(1)}-X_{\pi_1}^{(1)}, \ldots, X_{\pi_{h+1}}^{(2)}-X_{\pi_h}^{(2)}\right)\sqrt{p_{X^{(1,h+1)},\pi}\left(1-p_{X^{(1,h+1)},\pi}\right)p_{X^{(2,h+1)},\pi}\left(p_{X^{(2,h+1)},\pi}\right)}}{1-\sum_{\pi\in S_h}p_{X^{(1,h+1)},\pi}, p_{X^{(2,h+1)},\pi}},
\end{aligned}$$

with $\pi = (\pi_0, \ldots, \pi_h)$, *and*

$$\begin{aligned}
p_{X^{(1,h+1)},\pi} &= \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \pi\right),\\
p_{X^{(2,h+1)},\pi} &= \mathbb{P}\left(\Pi\left(X_0^{(2)} \ldots, X_h^{(2)}\right) = \pi\right).
\end{aligned}$$

Proof. First note that

$$\begin{aligned}
&OPD_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right)\\
&= \frac{\sum_{\pi\in\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}\\
&\quad-\frac{\sum_{\pi\in\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}\\
&\quad+\frac{\sum_{\pi\in S_h\setminus\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}
\end{aligned}$$

$$-\frac{\sum_{\pi\in S_h\setminus\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)},$$

and by focusing on the pattern $\pi = (0, 1, \ldots, h)$ in the first summand, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=(0,1,\ldots,h)\right)\\
=&\mathbb{P}\left(X_0^{(1)}\geq X_1^{(1)}\geq\ldots\geq X_h^{(1)},X_0^{(2)}\geq X_1^{(2)}\geq\ldots\geq X_h^{(2)}\right)\\
=&\mathbb{P}\left(Y_1^{(1)}\leq 0,\ldots,Y_h^{(1)}\leq 0,\ldots,Y_1^{(2)}\leq 0,\ldots,Y_h^{(2)}\leq 0\right)\\
=&\tau\left(Y_1^{(1)},\ldots,Y_h^{(1)},Y_1^{(2)},\ldots,Y_h^{(2)}\right)\sqrt{\tilde{p}_{Y^{(1)}}\left(1-\tilde{p}_{Y^{(1)}}\right)\tilde{p}_{Y^{(2)}}\left(1-\tilde{p}_{Y^{(2)}}\right)}+\tilde{p}_{Y^{(1)}}\tilde{p}_{Y^{(2)}},
\end{aligned}$$

following Lemma 5.6.
Note that the different normalizing constants in the denumerators of the two dependence measures originate from the fact that both dependence measures are normed to 1 in the case $X^{(1,h+1)} = X^{(2,h+1)}$. However, multivariate Kendall's τ does only consider one pattern in the ordinal view, but ordinal pattern dependence cannot be normed to 1 for each summand but must be norned for the entire sum over all possible patterns.
Due to the symmetry property of the multivariate normal distribution, we use

$$\left(X^{(1,h+1)},X^{(2,h+1)}\right)^t\overset{\mathcal{D}}{=}\left(-X^{(1,h+1)},-X^{(2,h+1)}\right)^t$$

and conclude that

$$\begin{aligned}
&\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=(0,1,\ldots,h)\right)\\
=&\,\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=(h,h-1,\ldots,0)\right).
\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
&OPD_h\left(X^{(1,h+1)},X^{(2,h+1)}\right)\\
=&\frac{2\tau\left(Y_1^{(1)},\ldots,Y_h^{(1)},Y_1^{(2)},\ldots,Y_h^{(2)}\right)\sqrt{\tilde{p}_{Y^{(1)}}\left(1-\tilde{p}_{Y^{(1)}}\right)\tilde{p}_{Y^{(2)}}\left(1-\tilde{p}_{Y^{(2)}}\right)}}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}\\
&+\frac{\sum_{\pi\in S_h\setminus\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}\\
&-\frac{\sum_{\pi\in S_h\setminus\{(h,h-1,\ldots,0),(0,1,\ldots,h)\}}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}\mathbb{P}\left(\Pi\left(X^{(1,h+1)}\right)=\pi\right)\mathbb{P}\left(\Pi\left(X^{(2,h+1)}\right)=\pi\right)}.
\end{aligned}$$

We proved (i) and found a possibility to express OPD_h in terms of multivariate Kendall's τ of the increment vectors $\left(Y_1^{(1)},\ldots,Y_h^{(1)},Y_1^{(2)},\ldots,Y_h^{(2)}\right)$ of the considered processes. However, the result might be unsatisfactory, since the last two summands are quite unhandy. Fortunately, it is possible to extend this concept: let $\pi = (\pi_0, \ldots, \pi_h)$ be a permutation in S_h. If

$$\Pi\left(X_0^{(1)},\ldots,X_h^{(1)}\right)=\pi$$

it holds that

$$\left\{X_{\pi_0}^{(1)}\geq X_{\pi_1}^{(1)}\geq\ldots\geq X_{\pi_h}^{(1)}\right\}.$$

Then, we can express ordinal pattern dependence by the following formula:

$$
\begin{aligned}
&OPD_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) \\
=&\frac{\sum_{\pi \in S_h} \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right) = \pi\right) - p_{X^{(1,h+1)},\pi} p_{X^{(2,h+1)},\pi}}{1 - \sum_{\pi \in S_h} \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \pi\right) \mathbb{P}\left(\Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right) = \pi\right)} \\
=&\frac{\sum_{\pi \in S_h} \mathbb{P}\left(X_{\pi_0}^{(1)} \geq X_{\pi_1}^{(1)} \geq \ldots \geq X_{\pi_h}^{(1)}, X_{\pi_0}^{(2)} \geq X_{\pi_1}^{(2)} \geq \ldots \geq X_{\pi_h}^{(2)}\right) - p_{X^{(1,h+1)},\pi} p_{X^{(2,h+1)},\pi}}{1 - \sum_{\pi \in S_h} p_{X^{(1,h+1)},\pi}, p_{X^{(2,h+1)},\pi}} \\
=&\frac{\sum_{\pi \in S_h} \mathbb{P}\left(X_{\pi_1}^{(1)} - X_{\pi_0}^{(1)} \leq 0, \ldots, X_{\pi_h}^{(1)} - X_{\pi_{h-1}}^{(1)} \leq 0, \ldots, X_{\pi_h}^{(2)} - X_{\pi_{h-1}}^{(2)} \leq 0\right) - p_{X^{(1,h+1)},\pi} p_{X^{(2,h+1)},\pi}}{1 - \sum_{\pi \in S_h} p_{X^{(1,h+1)},\pi}, p_{X^{(2,h+1)},\pi}} \\
=&\frac{\sum_{\pi \in S_h} \tau\left(X_{\pi_1}^{(1)} - X_{\pi_0}^{(1)}, \ldots, X_{\pi_h}^{(2)} - X_{\pi_{h-1}}^{(2)}\right) \sqrt{p_{X^{(1,h+1)},\pi}\left(1 - p_{X^{(1,h+1)},\pi}\right) p_{X^{(2,h+1)},\pi}\left(p_{X^{(2,h+1)},\pi}\right)}}{1 - \sum_{\pi \in S_h} p_{X^{(1,h+1)},\pi}, p_{X^{(2,h+1)},\pi}}
\end{aligned}
$$

with

$$
\begin{aligned}
p_{X^{(1,h+1)},\pi} &= \mathbb{P}\left(\Pi\left(X_0^{(1)}, \ldots, X_h^{(1)}\right) = \pi\right) \\
&= \mathbb{P}\left(X_{\pi_0}^{(1)} \geq X_{\pi_1}^{(1)} \geq \ldots \geq X_{\pi_h}^{(1)}\right) \\
&= \mathbb{P}\left(X_{\pi_1}^{(1)} - X_{\pi_0}^{(1)} \leq 0, \ldots, X_{\pi_h}^{(1)} - X_{\pi_{h-1}}^{(1)} \leq 0\right)
\end{aligned}
$$

and

$$
\begin{aligned}
p_{X^{(2,h+1)},\pi} &= \mathbb{P}\left(\Pi\left(X_0^{(2)}, \ldots, X_h^{(2)}\right) = \pi\right) \\
&= \mathbb{P}\left(X_{\pi_0}^{(2)} \geq X_{\pi_1}^{(2)} \geq \ldots \geq X_{\pi_h}^{(2)}\right) \\
&= \mathbb{P}\left(X_{\pi_1}^{(2)} - X_{\pi_0}^{(2)} \leq 0, \ldots, X_{\pi_h}^{(2)} - X_{\pi_{h-1}}^{(2)} \leq 0\right).
\end{aligned}
$$

□

We have found a way to describe OPD_h of $X^{(1,h+1)}$ and $X^{(2,h+1)}$ partwise by multivariate Kendall's τ of the *original* increments in (i) and fully by the *modified* increments, where the differences considered are determined by the regarded ordinal pattern π.

Note that it is a special property of the multivariate Gaussian distribution that the distribution of

$$\left(X_{\pi_1}^{(1)} - X_{\pi_0}^{(1)}, \ldots, X_{\pi_h}^{(1)} - X_{\pi_{h-1}}^{(1)}, X_{\pi_1}^{(2)} - X_{\pi_0}^{(2)}, \ldots, X_{\pi_h}^{(2)} - X_{\pi_{h-1}}^{(2)}\right)^t$$

is uniquely determined by the cross-correlations $\mathbb{E}\left(X_0^{(1)} X_k^{(2)}\right)$ for $k = -h, \ldots, h$ and the autocorrelations $\mathbb{E}\left(X_0^{(1)} X_k^{(1)}\right)$ for $k = 0, \ldots, h$, and $\mathbb{E}\left(X_0^{(2)} X_k^{(2)}\right)$ for $k = 0, \ldots, h$. Therefore, it is possible to express all of the dependencies in the vector above by the two-dimensional marginal distributions. However, since we do not have a closed expression of orthant probabilities of a multivariate Gaussian vector with more than three elements, it is not possible to constitute a closed form for ordinal pattern dependence in terms of Kendall's τ neither of the marginals nor of the joint vector. Therefore, we can only show this relation illustratively in the case $h = 1$ as the following example shows.

Example 5.8 *Recall that for Gaussian random vectors* $X^{(1,2)} = \left(X_0^{(1)}, X_1^{(1)}\right)^t$ *and* $X^{(2,2)} = \left(X_0^{(2)}, X_1^{(2)}\right)^t$ *satisfying the assumptions in Theorem 5.7, it holds that*

$$\begin{aligned} OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) &= \tau_1\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right) \\ &= \frac{2}{\pi}\arcsin\left(\mathrm{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right)\right) \end{aligned}$$

according to Theorem 5.7, (i).
Moreover, if $X^{(1,2)}$ *and* $X^{(2,2)}$ *exhibit the same distribution and have standard normal marginals, it holds that*

$$\mathrm{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right) = \frac{2\mathbb{E}\left(X_0^{(1)}X_0^{(2)}\right) - \mathbb{E}\left(X_0^{(1)}X_1^{(2)}\right) - \mathbb{E}\left(X_0^{(2)}X_1^{(1)}\right)}{2 - 2\mathbb{E}\left(X_0^{(1)}X_1^{(1)}\right)}.$$

In general, we know that $\tau_1(X_0^{(1)}, X_0^{(2)}) = \frac{2}{\pi}\arcsin\left(\mathrm{Cor}\left(X_0^{(1)}, X_0^{(2)}\right)\right)$ *for a Gaussian random vector* $\left(X_0^{(1)}, X_0^{(2)}\right)^t$*, and hence* $\mathrm{Cor}\left(X_0^{(1)}, X_0^{(2)}\right) = \sin\left(\frac{\pi}{2}\tau_1\left(X_0^{(1)}, X_0^{(2)}\right)\right)$.
As a result, we obtain

$$\begin{aligned} &OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) \\ &= \frac{2}{\pi}\arcsin\left(\frac{2\sin\left(\frac{\pi}{2}\tau_1\left(X_0^{(1)}, X_0^{(2)}\right)\right) - \sin\left(\frac{\pi}{2}\tau_1\left(X_0^{(1)}, X_1^{(2)}\right)\right) - \sin\left(\frac{\pi}{2}\tau_1\left(X_1^{(1)}, X_0^{(2)}\right)\right)}{2 - 2\sin\left(\frac{\pi}{2}\tau_1\left(X_0^{(1)}, X_1^{(1)}\right)\right)}\right). \end{aligned}$$

Therefore, the ordinal pattern dependence of order 1 *is determined by* $\tau_1\left(X_0^{(1)}, X_0^{(2)}\right)$, $\tau_1\left(X_0^{(1)}, X_1^{(2)}\right)$, $\tau_1\left(X_1^{(1)}, X_0^{(2)}\right)$, *and* $\tau_1\left(X_1^{(1)}, X_1^{(2)}\right)$.

We have derived theoretical results that relate ordinal pattern dependence to multivariate Kendall's τ. To illustrate these formulas, and to get a better intuitive understanding of the differences between the dependence measures, including those considered in the pilot study in Section 5.1, we restrict ourselves to the cases $h = 1$ and $h = 2$.
To this end, we use a classical bivariate Gaussian time series, namely a bivariate AR(1)-process for $h = 1$ and a bivariate AR(2)-process for $h = 2$ and want to study under which assumptions it is possible that OPD detects a dependence while classical measures like correlation or Kendall's τ do not.

5.2.1 Example of an AR(1)-process for $h = 1$

In the case $h = 1$, first remark the following corollary as a conclusion of Theorem 5.7, since $S_2 \setminus \{(1,0),(0,1)\} = \emptyset$ and $\tilde{p}_{Y^{(1,2)}} = \mathbb{P}\left(Y_1^{(1)} \leq 0\right) = \frac{1}{2}$.

Corollary 5.9 *Under the assumptions of Theorem 5.7 for* $h = 1$ *we have a direct connection between ordinal pattern dependence and the univariate Kendall's* τ*:*

$$\begin{aligned} OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) &= \tau_1\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right) \\ &= \tau_1\left(Y_1^{(1)}, Y_1^{(2)}\right) \\ &= \frac{2}{\pi}\arcsin\left(\mathrm{Cor}\left(Y_1^{(1)}, Y_1^{(2)}\right)\right). \end{aligned}$$

This corollary leads to the conjecture that it is possible to construct a bivariate process, which is uncorrelated, since correlation only compares two random variables at equal points in time, but which has a positive ordinal pattern dependence. In the following, we give a way to create such a process in terms of a bivariate AR(1)-time series.

Lemma 5.10 *Let $(X_j)_{j\geq 0}$, be a multivariate AR(1)-process defined by*

$$X_j := \begin{pmatrix} X_j^{(1)} \\ X_j^{(2)} \end{pmatrix}, j \geq 0,$$

with $X_j = AX_{j-1} + \xi_j$,

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_j := \begin{pmatrix} \varepsilon_j \\ \eta_j \end{pmatrix}.$$

We assume $a^2 + b^2 < 1$ to ensure stationarity and that ξ_j, $j \geq 0$ are bivariate Gaussian random vectors with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix). Then, it holds that $\mathrm{Cov}(X_0^{(1)}, X_0^{(2)}) = 0$ and for $X^{(1,2)} := \left(X_0^{(1)}, X_1^{(1)}\right)$ and $X^{(2,2)} := (X_0^{(2)}, X_1^{(2)})$, we obtain

$$OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = \frac{2}{\pi}\arcsin\left(-\frac{b}{\sqrt{1-a^2}}\right).$$

Proof. The eigenvalues of A are $\lambda_{1,2} = \pm\sqrt{a^2+b^2}$, and thus $(X_j)_{j\geq 0}$ is stationary. Since the AR(1)-equation defines a Markov chain with state space $\mathbb{R}^2$, the joint distribution of $X_j = \left(X_j^{(1)}, X_j^{(2)}\right)^t$ is uniquely characterized by the distributional fixed point equation

$$X \stackrel{\mathcal{D}}{=} A\,X + \xi,$$

where $\xi = (\varepsilon, \eta)$ has a bivariate normal distribution with mean zero and covariance matrix I_2, and where ξ is independent of X. We will now show that $X \sim \mathcal{N}\left(0, \sigma^2 I_2\right)$ satisfies this equation with

$$\sigma^2 = \frac{1}{1-a^2-b^2}.$$

In order to prove this, we need to calculate the distribution of $A\,X + \xi$. Since the distribution is Gaussian, it suffices to calculate the variances and the covariance. We obtain

$$\begin{aligned}
&\mathrm{Cov}(aX^{(1)} + bX^{(2)} + \epsilon, bX - aY + \eta) = ab\sigma^2 - ab\sigma^2 = 0, \\
&\mathrm{Var}(aX^{(1)} + bX^{(2)} + \epsilon) = a^2\sigma^2 + b^2\sigma^2 + 1 = \frac{a^2+b^2}{1-a^2-b^2} + 1 = \frac{1}{1-a^2-b^2} = \sigma^2, \\
&\mathrm{Var}(bX^{(1)} - aX^{(2)} + \epsilon) = b^2\sigma^2 + a^2\sigma^2 + 1 = \frac{a^2+b^2}{1-a^2-b^2} + 1 = \frac{1}{1-a^2-b^2} = \sigma^2,
\end{aligned}$$

which shows that $A\,X + \xi$ has indeed the same distribution as X.

In order to determine the OPD_1 of the two processes, we need to calculcate the correlation of the differences. The covariance of the increments is given by

$$\mathrm{Cov}(X_2^{(1)} - X_1^{(1)}, X_2^{(2)} - X_1^{(2)}) = ((a-1)X_1^{(1)} + bX_1^{(2)} + \epsilon_2, bX_1^{(1)} + (-a-1)X_1^{(2)} + \eta_2)$$

$$= b(a-1)\sigma^2 - b(a+1)\sigma^2$$
$$= -2b\sigma^2 = \frac{-2b}{1-a^2-b^2}$$

and the variances of the increments are given by

$$\mathrm{Var}(X_2^{(1)} - X_1^{(1)}) = (a-1)^2\sigma^2 + b^2\sigma^2 + 1 = \frac{(a-1)^2+b^2}{1-a^2-b^2} + 1 = \frac{2(1-a)}{1-a^2-b^2},$$
$$\mathrm{Var}(X_2^{(2)} - X_1^{(2)}) = b^2\sigma^2 + (a+1)^2\sigma^2 + 1 = \frac{b^2+(a+1)^2}{1-a^2-b^2} + 1 = \frac{2(a+1)}{1-a^2-b^2}.$$

Thus, we obtain the following formula for the correlation of the increments:

$$\mathrm{Cor}(X_2^{(1)} - X_1^{(1)}, X_2^{(2)} - X_1^{(2)}) = \frac{-2b}{\sqrt{4(1-a)(a+1)}} = -\frac{b}{\sqrt{1-a^2}}$$

Using the identity $OPD_1((X_1^{(1)}, X_2^{(1)}), (X_1^{(2)}, X_2^{(2)})) = \frac{2}{\pi} \arcsin \mathrm{Cor}(X_2^{(1)} - X_1^{(1)}, X_2^{(2)} - X_1^{(2)})$, we proved the result. □

Remark 5.11 *(i) The special choice of the matrix A in Lemma 5.10 assures that the two processes $\left(X_j^{(1)}\right)_{j\geq 0}$ and $\left(X_j^{(2)}\right)_{j\geq 0}$ have identical marginals, and that $X_j^{(1)}$ and $X_j^{(2)}$ are independent for each fixed i. In fact, one can show that the latter two properties only hold if A is either of the form in Lemma 5.10 or of the form*

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \tag{5.3}$$

In this case, using similar calculations as above, one obtains $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0$.

(ii) Lemma 5.10 provides an example of a Gaussian process for which Pearson's correlation of $X_j^{(1)}$ and $X_j^{(2)}$ equals 0, i.e., the one-dimensional marginals are independent. However, the processes $\left(X_j^{(1)}\right)_{j\geq 0}$ and $\left(X_j^{(2)}\right)_{j\geq 0}$ are not independent, as can be seen from the identity for $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right)$.

5.2.2 Simulations

We emphasize our results by simulating a bivariate AR(1)-process

$$X_j := \begin{pmatrix} X_j^{(1)} \\ X_j^{(2)} \end{pmatrix}, \ j = 1, \ldots, 10\,000,$$

with $X_j = AX_{j-1} + \xi_i$, where

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_i := \begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix},$$

ξ_i being a multivariate Gaussian random vector with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix). We choose $a^2+b^2 < 1$ but close to 1 in order to obtain $\mathrm{Cov}\left(X_j^{(1)}, X_j^{(2)}\right) = 0$, but OPD_1 close to 1. For our simulation $a = 0.7$ and $b = -0.7$ yield illustrating results.

In Figure 5.8, left plot, we simulate one sample path of the processes $\left(X_j^{(1)}\right)$ and $\left(X_j^{(2)}\right)$, both for $j = 1, \ldots, 500$ and in Figure 5.8, right plot, we do the same for the increment processes. It is easy to see that the two processes displayed in the right plot in Figure 5.8 are positively correlated. The corresponding values of the estimators of the different dependence measures are displayed in Table 5.1.

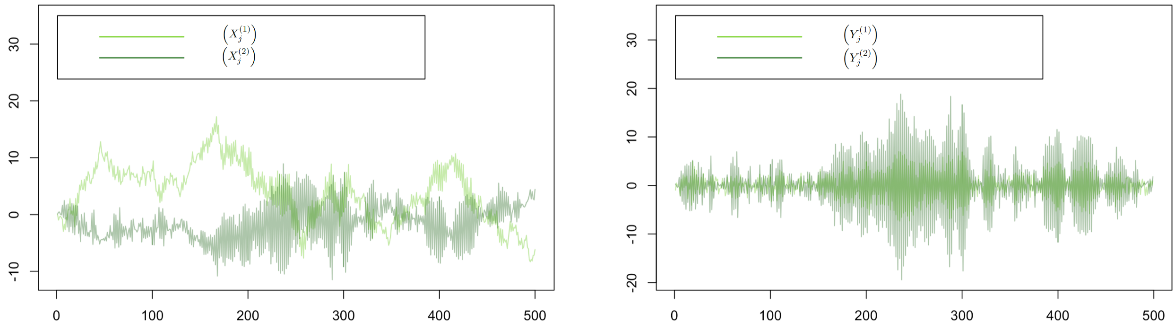

Figure 5.8: One sample path of the original process (X_j) (left) and of the increment process (Y_j) (right), $j = 1, \ldots, 500$.

$\text{Cor}\left(X_0^{(1)}, X_0^{(2)}\right)$	-0.049
$\tau\left(X_0^{(1)}, X_0^{(2)}\right)$	-0.035
$\rho\left(X_0^{(1)}, X_0^{(2)}\right)$	-0.054

$OPD_1\left(X^{(1,2)}, X^{(2,2)}\right)$	0.859
$\text{Cor}\left(Y_1^{(1)}, Y_1^{(2)}\right)$	0.98
$\tau\left(Y_1^{(1)}, Y_1^{(2)}\right)$	0.866
$\rho\left(Y_1^{(1)}, Y_1^{(2)}\right)$	0.975

Table 5.1: Values of the different empirical dependence measures for the simulated $AR(1)$ process.

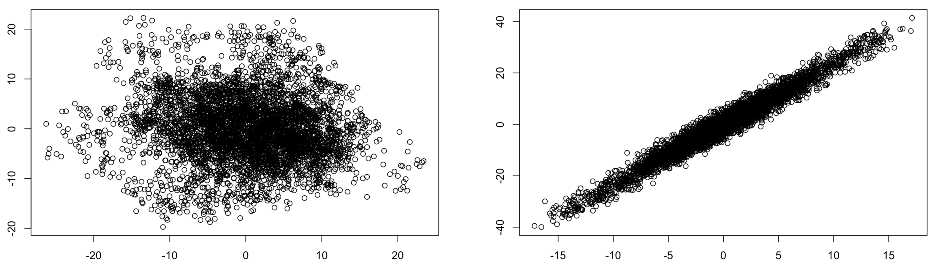

Figure 5.9: Scatterplot of $\left(X_j^{(1)}\right)$ and $\left(X_j^{(2)}\right)$ (left) and of $\left(Y_j^{(1)}\right)$ and $\left(Y_j^{(2)}\right)$ (right), $j = 1, \ldots, 5000$.

The high dependence between the two increment processes is reflected in the right plot of Figure 5.9. In the left plot of Figure 5.9 the uncorrelated processes are shown. To get an impression of the empirical convergence properties of OPD_1 and Pearson's correlation, see [11], Figure 1.

5.2.3 Example of an AR(2)-process for $h = 2$

For the computation of the ordinal pattern dependence of order $h = 1$, the crucial quantity is $\operatorname{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right)$ since, according to Thorem 5.7, $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right)$ is just a monotone transformation of this correlation. Therefore, it is natural to wonder whether it is possible to construct a stationary, bivariate process $\left(X_j^{(1)}, X_j^{(2)}\right)_{j\geq 0}$ with $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0$, but $OPD_2\left(X^{(1,2)}, X^{(2,2)}\right) \neq 0$.
The AR(1)-process in Lemma 5.10 does not fulfill these conditions, since the restriction

$$\operatorname{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right) = -\frac{b}{\sqrt{1-a^2}} = 0$$

implies $b = 0$. As a result, we would consider a process $(X_j)_{j\geq 0}$, given by

$$X_j = \left(X_j^{(1)}, X_j^{(2)}\right)^t = \left(aX_{j-1}^{(1)} + \xi_j, -aX_{j-1}^{(2)} + \eta_j\right)^t,$$

that does not incorporate any *dynamical* dependence between the processes $\left(X_j^{(1)}\right)_{j\geq 0}$ and $\left(X_j^{(2)}\right)_{j\geq 0}$. The only dependence in this model exists within each marginal process. Yet, this does not have any impact on ordinal pattern dependence.
Following Remark 5.11, the choice of the matrix A in (5.3) yields $\operatorname{Cor}\left(X_j^{(1)}, X_j^{(2)}\right) = 0$ for $j = 1, 2$ and $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0$. This leads to the question whether this special construction of an AR(1)-process fulfills $OPD_2\left(X^{(1,3)}, X^{(2,3)}\right) \neq 0$.

Lemma 5.12 *Consider the stationary bivariate Gaussian AR*(1)*-process* $X_j = \left(X_j^{(1)}, Y_j^{(2)}\right)_{j\geq 0}$, *satisfying*

$$X_j = AX_{j-1} + \xi_j,$$

with (ξ_j) *as given in Lemma 5.10 and with matrix A given by* (5.3), *such that* $a^2 + b^2 < 1$. *Then, it holds that*

$$\operatorname{Cov}\left(X_1^{(1)}, X_1^{(2)}\right) = 0,\ \operatorname{Var}\left(X_1^{(1)}\right) = \operatorname{Var}\left(X_1^{(2)}\right) = \sigma^2 = \frac{1}{1-a^2-b^2},$$
$$OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0,\ \operatorname{Cor}\left(X_2^{(1)} - X_1^{(1)}, X_3^{(2)} - X_2^{(2)}\right) = -b,$$
$$\operatorname{Cor}\left(X_3^{(1)} - X_2^{(1)}, X_2^{(2)} - X_1^{(2)}\right) = b.$$

Proof. The first three identities can be shown as in Lemma 5.10. Thus, it remains to show the latter two. It holds that

$$\operatorname{Var}\left(X_2^{(1)} - X_1^{(1)}\right) = (a-1)^2\sigma^2 + b^2\sigma^2 + 1 = \frac{(a-1)^2 + b^2 + 1 - a^2 - b^2}{1-a^2-b^2} = 2(1-a)\sigma^2.$$

Analogously, we obtain

$$\operatorname{Var}\left(X_3^{(2)} - X_2^{(2)}\right) = 2(1-a)\sigma^2.$$

Furthermore, it holds that

$$\begin{aligned}\mathrm{Cov}(X_3^{(2)} - X_2^{(2)}, X_2^{(1)} - X_1^{(1)}) &= \mathbb{E}\left(X_3^{(2)}X_2^{(1)}\right) - \mathbb{E}\left(X_2^{(2)}X_2^{(1)}\right) - \mathbb{E}\left(X_3^{(2)}X_1^{(1)}\right) + \mathbb{E}\left(X_2^{(2)}X_1^{(1)}\right)\\ &= 2b(a-1)\sigma^2,\end{aligned}$$

since

$$\mathbb{E}(X_3^{(2)}X_2^{(1)}) = -b\sigma^2,\ \mathbb{E}(X_3^{(2)}X_1^{(1)}) = -2ab\sigma^2.$$

Altogether, we arrive at

$$\mathrm{Cor}(X_3^{(2)} - X_2^{(2)}, X_2^{(1)} - X_1^{(1)}) = \frac{2ab-2b}{2(1-a)} = -b.$$

$\mathrm{Cor}(X_2^{(2)} - X_1^{(2)}, X_3^{(1)} - X_2^{(1)}) = b$ is derived by similar calculations. □

Lemma 5.12 provides an example of a bivariate process $\left(X_j^{(1)}, X_j^{(2)}\right)_{j\geq 0}$ for which $\mathrm{Cor}\left(X_j^{(1)}, X_j^{(2)}\right) = 0$ and $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0$, but where the processes $\left(X_j^{(1)}\right)_{j\geq 0}$ and $\left(X_j^{(2)}\right)_{j\geq 0}$ are nevertheless dependent. The fact that the increments $X_2^{(1)} - X_1^{(1)}$ and $X_3^{(2)} - X_2^{(2)}$ are dependent, leads us to conjecture that $OPD_2\left(X^{(1,3)}, X^{(2,3)}\right) \neq 0$, but we do not have a proof. The analytic difficulties arising in order to calculate $OPD_2\left(X^{(1,3)}, X^{(2,3)}\right)$ are presented in the subsequent example of a bivariate AR(2)-process $(X_j^{(1)}, X_j^{(2)})_{j\geq 0}$ for which $\mathrm{Cor}(X_j^{(1)}, X_j^{(2)}) = 0$ and $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0$, but where the processes $(X_j^{(1)})_{j\geq 0}$ and $(X_j^{(2)})_{j\geq 0}$ are dependent. For this example, we provide simulation results that confirm $OPD_2\left(X^{(1,3)}, X^{(2,3)}\right) \neq 0$.

Example 5.13 *Let $(X_j)_{j\geq 0}$, be a multivariate AR(2)-process defined by*

$$X_j := \begin{pmatrix} X_j^{(1)} \\ X_j^{(2)} \end{pmatrix},\ j \geq 0,$$

where $X_j = AX_{j-2} + \xi_j$ with

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_j := \begin{pmatrix} \varepsilon_j \\ \eta_j \end{pmatrix}, \tag{5.4}$$

ξ_j, $j \geq 0$, bivariate Gaussian random vectors with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix) and $X_0^{(1)} := \xi_0$, $X_0^{(2)} := \eta_0$, $X_1^{(1)} := \xi_1$, $X_1^{(2)} = \eta_1$. Moreover, we assume that $\sigma^2 := \mathrm{Var}\left(X_0^{(1)}\right) = \mathrm{Var}\left(X_0^{(2)}\right)$.

By definition it holds that $\mathrm{Cov}\left(X_0^{(1)}, X_0^{(2)}\right) = \mathrm{Cov}\left(X_1^{(1)}, X_1^{(2)}\right) = 0$. Moreover, we have $OPD_1\left(X^{(1,2)}, X^{(2,2)}\right) = 0$, since

$$\begin{aligned}\mathrm{Cov}\left(X_2^{(1)} - X_1^{(1)}, X_2^{(2)} - X_1^{(2)}\right) &= \mathbb{E}\left[\left(aX_0^{(1)} + bX_0^{(2)} + \xi_2 - X_1^{(1)}\right)\left(bX_0^{(1)} - aX_0^{(2)} + \eta_2 - X_1^{(2)}\right)\right]\\ &= ab\sigma^2 - ba\sigma^2\\ &= 0.\end{aligned}$$

In order to compute $OPD_2\left(X^{(1,3}, X^{(2,3)}\right)$*, we have to calculate*

$$\mathbb{P}\left(\Pi\left(X_0^{(1)}, X_1^{(1)}, X_2^{(1)}\right) = \Pi\left(X_0^{(2)}, X_1^{(2)}, X_2^{(2)}\right) = \pi\right)$$

for every $\pi \in S_2$*.*
With $\pi = (0,1,2)$ *it follows that*

$$\begin{aligned}
&\mathbb{P}\left(\Pi\left(X_0^{(1)}, X_1^{(1)}, X_2^{(1)}\right) = \Pi\left(X_0^{(2)}, X_1^{(2)}, X_2^{(2)}\right) = \pi\right)\\
=&\mathbb{P}\left(X_0^{(1)} \leq X_1^{(1)} \leq X_2^{(1)}, X_0^{(2)} \leq X_1^{(2)} \leq X_2^{(2)}\right)\\
=&\mathbb{P}\left(X_1^{(1)} - X_0^{(1)} \geq 0, X_2^{(1)} - X_1^{(1)} \geq 0, X_1^{(2)} - X_0^{(2)} \geq 0, X_2^{(2)} - X_1^{(2)} \geq 0\right).
\end{aligned}$$

As a result, computing $OPD_2\left(X^{(1,3)}, X^{(2,3)}\right)$ *boils down to determine the orthant probabilities of a four-dimensional Gaussian vector. To our knowledge a closed expression for these probabilities is not at hand; see [1].*
Nonetheless, we claim that it is no longer sufficient to consider $\operatorname{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_1^{(2)} - X_0^{(2)}\right)$*, which vanishes here, but we have two more quantities determining this probability, namely* $\operatorname{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_2^{(2)} - X_1^{(2)}\right)$ *and* $\operatorname{Cor}\left(X_2^{(1)} - X_1^{(1)}, X_1^{(2)} - X_0^{(2)}\right)$*. These correspond to the cross-correlation function* $r^{(1,2)}(k) = \operatorname{Cor}\left(X_1^{(1)} - X_0^{(1)}, X_{k+1}^{(2)} - X_k^{(2)}\right)$ *at lags* $k = 1$ *and* $k = -1$*. Therefore, these two quantities do not necessarily coincide since the cross-correlation function is in general not symmetric.*
Straightforward calculations yield

$$\begin{aligned}
&\operatorname{Cov}\left(X_1^{(1)} - X_0^{(1)}, X_2^{(2)} - X_1^{(2)}\right)\\
=&\operatorname{Cov}\left(X_2^{(1)} - X_1^{(1)}, X_3^{(2)} - X_2^{(2)}\right)\\
=&\mathbb{E}\left(\left(aX_0^{(1)} + bX_0^{(2)} + \xi_2 - X_1^{(1)}\right)\left(bX_1^{(1)} - aY_1^{(2)} + \eta_3 - bX_0^{(1)} + aX_0^{(2)} - \eta_2\right)\right)\\
=&-ab\sigma^2 + ab\sigma^2 - b\sigma^2\\
=&-b\sigma^2
\end{aligned}$$

and

$$\begin{aligned}
&\operatorname{Cov}\left(X_2^{(1)} - X_1^{(1)}, X_1^{(2)} - X_0^{(2)}\right)\\
=&\operatorname{Cov}\left(X_3^{(1)} - X_2^{(1)}, X_2^{(2)} - X_1^{(2)}\right)\\
=&\mathbb{E}\left(\left(aX_1^{(1)} + bX_1^{(2)} + \xi_3 - aX_0^{(1)} - bX_0^{(2)} - \xi_2\right)\left(bX_0^{(1)} - aX_0^{(2)} + \eta_2 - X_1^{(2)}\right)\right)\\
=&-b\sigma^2 - ab\sigma^2 + ab\sigma^2\\
=&-b\sigma^2.
\end{aligned}$$

With

$$\operatorname{Var}\left(X_j - X_{j-1}\right) = \operatorname{Var}\left(X_j\right) + \operatorname{Var}\left(X_{j-1}\right) = 2\sigma^2,$$

we arrive at

$$r^{(1,2)}(1) = r^{(1,2)}(-1) = \frac{-b\sigma^2}{\sqrt{4\sigma^2}} = -\frac{b}{2}.$$

Note that in this special construction the cross-covariance function at lag 1 *is symmetric.*

This construction of AR(2)-for $h = 2$ *can be extended to AR(h) for* $h \in \mathbb{N}$*, if one wants to obtain* $OPD_h\left(X^{(1,h+1)}, X^{(2,h+1)}\right) \neq 0$ *but* $OPD_i\left(X^{(1,i+1)}, X^{(2,i+1)}\right) = 0$, $i = 1, \ldots, h-1$ *and* $\operatorname{Cor}\left(X_0^{(1)} X_0^{(2)}\right) = 0$ *by using h independent AR(1)-processes and couple them via*

$$X_j = A\left(X_{j-h}^{(1)}, X_{j-h}^{(2)}\right)^t.$$

5.2.4 Simulations

We simulate a bivariate AR(2)-process

$$X_j := \begin{pmatrix} X_j^{(1)} \\ X_j^{(2)} \end{pmatrix}, \; j = 1, \ldots 10\,000,$$

with $X_j = AX_{j-2} + \xi_j$, where

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_j := \begin{pmatrix} \varepsilon_j \\ \eta_j \end{pmatrix},$$

ξ_j being a multivariate Gaussian random vector with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix). Since we need to maximize $r^{(1,2)}(1) = -\frac{b}{2}$ we have chosen $b = -0.98$ and $a = 0.01$ for the following simulations. The values of the dependence measures are displayed in Table 5.2. One sample path of $\left(X_j^{(1)}\right)$ and $\left(X_j^{(2)}\right)$ for $i = 1, \ldots, 500$ as well as one sample path

$\operatorname{Cor}\left(X_0^{(1)}, X_0^{(2)}\right)$	0.009	$OPD_1\left(X^{(1,2)}, X^{(2,2)}\right)$	0.039	$OPD_2\left(X^{(1,3)}, X^{(2,3)}\right)$	0.201
$\tau\left(X_0^{(1)}, X_0^{(2)}\right)$	-0.035	$\operatorname{Cor}\left(Y_1^{(1)}, Y_1^{(2)}\right)$	0.026	$\operatorname{Cor}\left(Y_1^{(1)}, Y_2^{(2)}\right)$	0.478
$\rho\left(X_0^{(1)}, X_0^{(2)}\right)$	-0.054	$\tau\left(Y_1^{(1)}, Y_1^{(2)}\right)$	0.029	$\tau_2\left(Y_1^{(1,h)}, Y_1^{(2,h)}\right)$	0.397
		$\rho\left(Y_1^{(1)}, Y_1^{(2)}\right)$	0.044		

Table 5.2: Values of the different empirical dependence measures for the simulated AR(2)-process.

of the corresponding increment processes are shown in Figure 5.10. Compared to Figure 5.8 one can see that we have no longer a positive correlation. We finally provide three scatterplots in Figure 5.11, in which it is possible to see that there are no dependencies between the observations at a fixed point of time and between the increments at a fixed point of time. However, the third scatterplot illustrates the positive correlation of 0.478 between the two increment processes at lag 1. In [11], p.18, Figure 3, the empirical convergence properties of OPD_1 and OPD_2 illustrate the argumentation in Example 5.13. Further simulations that deal with the comparison of the estimators of OPD_2 and τ_2 can be found in [11], Section 5.4.

We close this chapter with a summary of the results and a brief outlook. Concerning the questions that arise from Section 5.1, we have seen that the univariate dependence measures only take data from equal points in time into account and, therefore, ignore the dynamical dependence that appears in a time series approach. Similar concerns hold for multivariate Pearson's Rho as

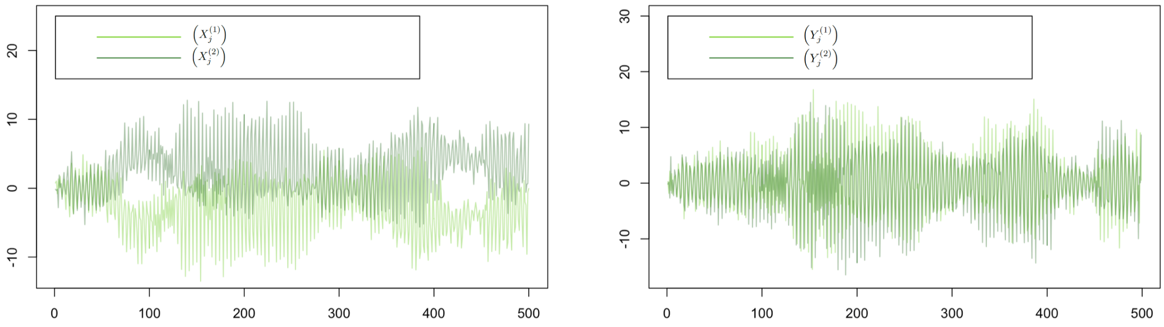

Figure 5.10: One sample path of $\left(X_j^{(1)}\right)$ and $\left(X_j^{(2)}\right)$ (left), $\left(Y_j^{(1)}\right)$ and $\left(Y_j^{(2)}\right)$ (right), $j = 1, \ldots, 500$.

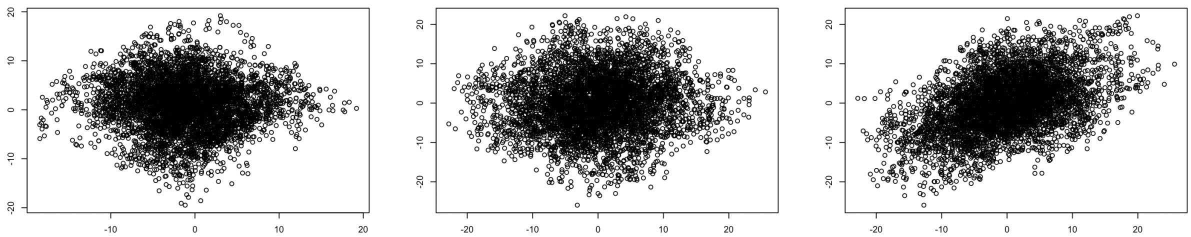

Figure 5.11: Scatterplots of $\left(X_j^{(1)}\right)$ and $\left(X_j^{(2)}\right)$ (left), $\left(Y_j^{(1)}\right)$ and $\left(Y_j^{(2)}\right)$ (middle) and $\left(Y_j^{(1)}\right)$ and $\left(Y_{j+1}^{(2)}\right)$ (right) , $j = 1, \ldots, 5000$.

proposed in [55], p. 2, as the dynamical dependence *within* two time series is not regarded either. Finally, in multivariate Kendall's τ as introduced in [29], p. 100, we have found a measure that reflects the time series approach when defined in terms of two $h+1$-dimensional random vectors. In the centered stationary Gaussian setting, we derived formulas to relate OPD_h to τ_h. In this context, it is to mention that the relation happens to only take the increments into account, as seen in Theorem 5.7. This observation suits to the property of ordinal patterns to be uniquely determined by their increments. Examples to emphasize the differences between the dependence measures are given in Section 5.2.1 by an AR(1)-process for $h = 1$ and in Section 5.2.3 by an AR(2)-process for $h = 2$ and are accompanied by simulations of the corresponding empirical estimators. In the case $h = 2$ we see exemplarily that besides OPD_2 and τ_2 all other dependence measures studied fail to detect the dependecies within the considered two AR(2)-processes.
Summarizing, conditioned especially by the robustness of ordinal pattern dependence as empirically shown in the pilot study in Section 5.1, ordinal pattern dependence yields a wholistic approach in measuring multivariate dependence in a time series context with many advantages in applications.

The field of multivariate dependence measures is of strong interest for practice and applications and still has interesting questions to investigate. A possibility of further research is to use the limit theorems obtained in Chapter 3 in this work to derive a limit theorem for multivariate Kendall's τ under long-range dependence. Another idea is the integration and comparison of multivariate Spearman's ρ, as proposed in [29], Section 3.1, to the set of multivarite dependence measures investigated in this work. Finally, the theoretical analysis of time series that obtain non-linear dependence seems a very interesting field, since linear multivariate dependence measures as Kendall's τ fail here, but ordinal pattern dependence detects the dependence. However, for the construction of the theory of this class of time series, we have to leave the Gaussian setting that only exhibits linear dependence. We motivate this idea with the following example:

Example 5.14 *Let $\left(X_j^{(1)}\right)_{j\geq 0}$ be a univariate standard i.i.d. Gaussian process. Define $X_j^{(2)} := \left(X_j^{(1)}\right)^2 - 1$. Note that these two processes are not independent, as the following calculation shows:*

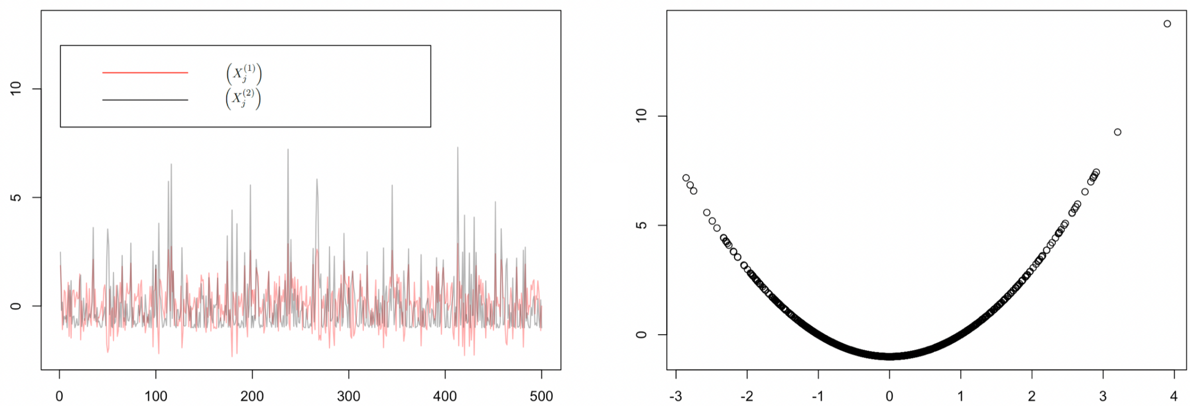

Figure 5.12: One sample path of $\left(X_j^{(1)}\right)$ and $\left(X_j^{(2)}\right)$, $j = 1, \ldots, 500$ (left) and the corresponding scatterplot (right).

$$\mathbb{P}\left(X_j^{(1)} \leq 5, \left(X_j^{(1)}\right)^2 - 1 \leq 5\right) = \mathbb{P}\left(X_j^{(1)} \leq 5, \left|X_j^{(1)}\right| \leq 2\right) = \mathbb{P}\left(\left|X_j^{(1)}\right| \leq 2\right)$$
$$\neq \mathbb{P}\left(X_j^{(1)} \leq 5\right) \mathbb{P}\left(\left|X_j^{(1)}\right| \leq 2\right).$$

The dependence of the two processes is also displayed in Figure 5.12.
However, the processes as well as their increments are uncorrelated and have no cross-correlations at any lag, since $\mathbb{E}\left(\left(X_j^{(1)}\right)^k\right) = 0$, if k is an odd positive integer. Therefore, no linear dependence is detectable. Hence, multivariate Kendall's τ also fails to detect the dependence for $h = 1$ and $h = 2$. Ordinal pattern dependence is nevertheless able to trace the quadratic dependence in this example, in a simulation study with $n = 10\,000$ data points for $h = 2$, we obtain $OPD_2 = 0.1331$.

Having this example in mind, ordinal pattern dependence seems to be a user friendly approach to capture those kinds of dependencies. An interesting field of application is given in the random variables that arises in the context of Hermite processes, see [3], Proposition 3.5 and [23], Theorem 4.1, since they are known to be uncorrelated but not independent.

6 Real-world data analysis

In this chapter we investigate data sets in the field of hydrology, which are provided by "The Global Runoff Data Centre, 56068 Koblenz, Germany". Those data sets are known to be often stationary and long-range dependent which would yield a great possibility to apply the theory we developed in this work to real world data. We consider mean daily discharges of the rivers Elbe

Figure 6.1: Location of the measuring stations Neu-Darchau, Wittenberge, Hofkirchen and Achleiten, created with d-maps.com, see `https://d-maps.com/m/europa/germany/allemagne_de/allemagne_de13.svg`.

and Danube, two data sets for each river, each measured at different locations, which are shown in Figure 6.1. Hence, Neu-Darchau and Wittenberge are cities located at the Elbe and Hofkirchen and Achleiten are located at the Danube. The distance between relevant locations for our data analysis is given in Table 6.1 The data sets of the Elbe contain data from 11/1899-12/2013 which

Location	Location	Distance
Neu-Darchau	Wittenberge	approx. 80km
Hofkirchen	Achleiten	approx. 38km
Neu-Darchau	Hofkirchen	approx. 715km

Table 6.1: Distance between the locations of the measuring stations.

yields a sample size of $n = 41699$. The data sets of the mean daily discharges of the Danbue consist of $n = 41334$ and are derived from measurements between 11/1900 and 12/2013. Since we analyze both rivers in one setting in the following, note that the first year of measurements of the corresponding data set concerning the Elbe is be cut off to make the time series comparable. To avoid ties, we add a small white noise process to each data set, with standard deviation $\sigma = 0.01$, so this white noise does not change the relevant data. The amount of ties in the four considered data sets is from 4% to 8%.

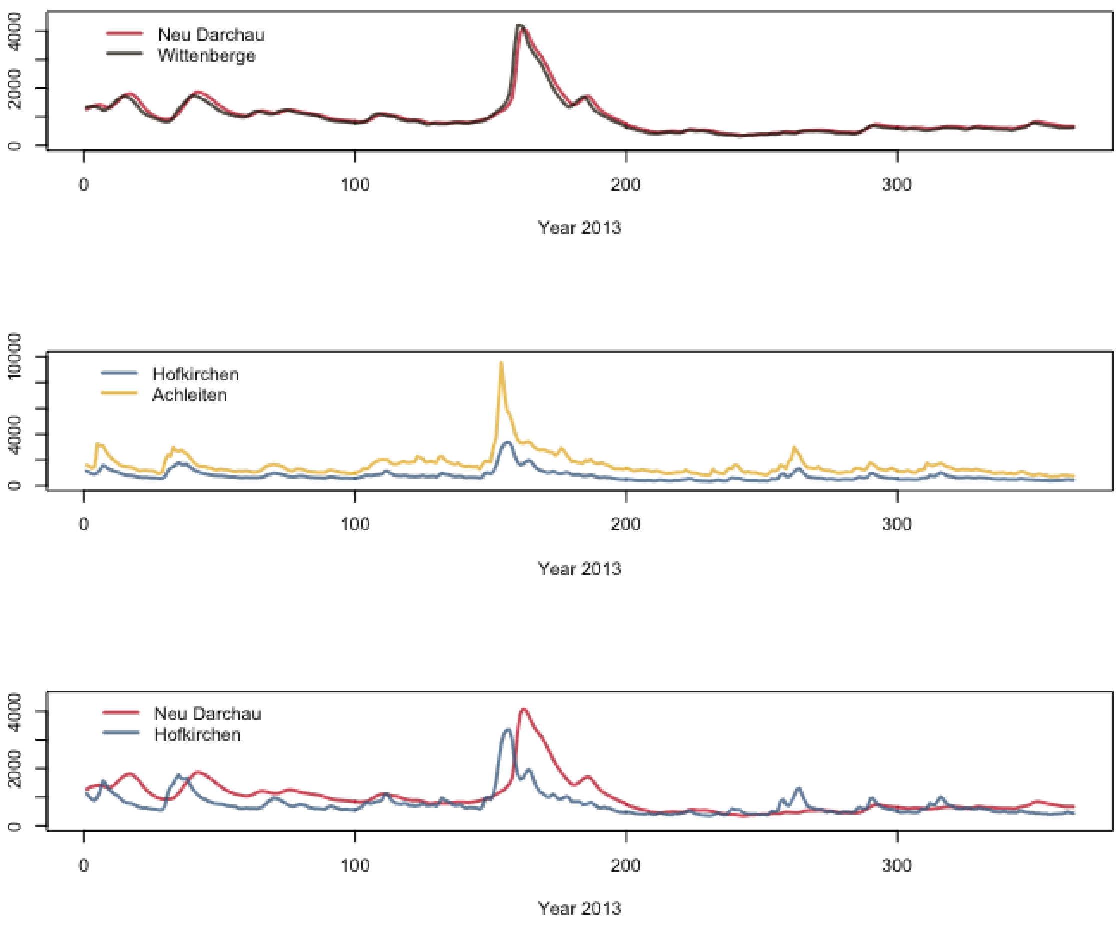

Figure 6.2: Daily mean discharges of corresponding locations in 2013.

First, we have to do a careful pre-processing of the data. We test stationarity using the *Augmented Dickey-Fuller Test* as implemented in the R package "tseries", for theoretical details see [20]. The Hurst parameter of each time series is estimated by using the R function *hurstexp* in the R package "pracma". As this question might arise, note that the estimation of the Hurst parameter based on ordinal pattern analysis is not suitable here. This estimator highly relies on fractional Gaussian noise as underlying process, since the special covariance structure is explicitly used there.

Location	Stationarity	Estimated Hurst parameter
Neu-Darchau	✓ (p-value=0.01)	0.722
Wittenberge	✓ (p-value=0.01)	0.779
Hofkirchen	✓ (p-value=0.01)	0.779
Achleiten	✓ (p-value=0.01)	0.726

Table 6.2: Properties of the data sets.

In the following, we study three settings:

- **Setting 1: Neu-Darchau vs. Wittenberge**
- **Setting 2: Hofkirchen vs. Achleiten**
- **Setting 3: Neu-Darchau vs. Hofkirchen**

The mean daily discharges for each setting in the year 2013 is displayed in Figure 6.2. Concerning the cross-correlation functions of these three settings, it is possible to recognize the similar behaviour of either the auto-correlation functions of the univariate time series as well as the cross-correlation function of the two time series within one setting, as one can observe in Figure 6.3.

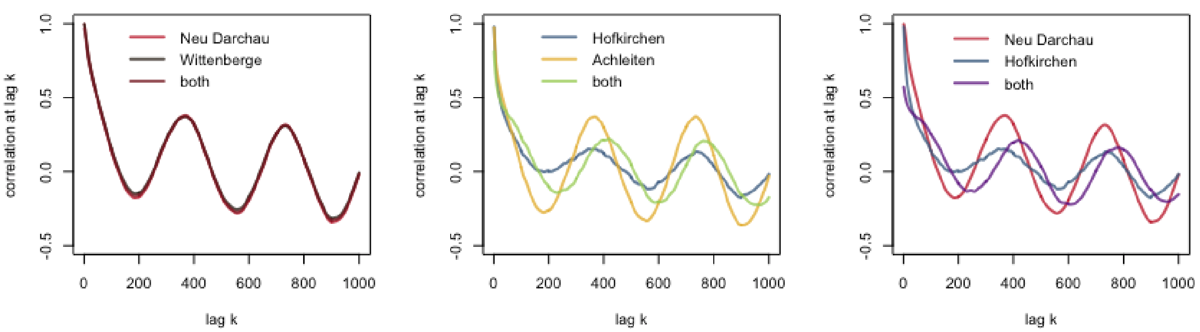

Figure 6.3: (Cross-)correlation functions for each setting.

Following these empirical results, we are hence dealing with long-range dependent stationary time series. We want to apply the methods introduced in Section 4.3, especially the various concepts in Chapter 4.4 and those in Chapter 5 to these data sets. To make them better fit into our theoretical setting, it is possible to transform the data sets to become Gaussian and since this is a monotone transformation, the ordinal patterns within these data sets are not affected. Hence, we know that the estimators we use in the following, have the statistical properties discussed in Section 4.4.3 and we have a real world example of one class of processes we studied in this work. The heatmap in Figure 6.4 show the result of the data analysis. Figure 6.5 provides a comparison ordered by the different settings.

	Correlation	Kendall's τ	Spearman's ρ
Setting 1	0,9738	0,8913	0,9751
Setting 2	0,8373	0,6373	0,8286
Setting 3	0,5859	0,4074	0,5745

	OPD_1	OPD_2	$OPD_{2,bw}$	$AWOPD_5$
Setting 1	0,4626	0,3542	0,2150	0,3238
Setting 2	0,3802	0,2702	0,1780	0,2529
Setting 3	0,0070	0,0070	0,0840	0,0130

	τ_1	$\tau_{1,bw}$	τ_2	$\tau_{2,bw}$
Setting 1	0,8922	0,8922	0,8929	0,8929
Setting 2	0,6378	0,6378	0,6369	0,6372
Setting 3	0,4073	0,4075	0,4080	0,4080

Figure 6.4: Heatmap of values of different dependence measures.

The first table in Figure 6.4 shows the values of the corresponding **univariate** dependence measures, while the second table deals with the different approaches of measuring **ordinal pattern dependence**. The third table provides the results of **multivariate Kendall's** τ measured for $h = 1$ and $h = 2$ estimated stepwise and blockwise, respectively.
We observe that all approaches of OPD measures significantly lower values as the other dependence measures. Especially for the third setting no dependence is registered. As already indicated in the pilot study in Section 5.1, for example larger jumps as they are known to happen often in hydrological time series, f.e. in times with a huge amout of rain falling, let the standard dependence measures increase very fast since they react very sensitive to those “shocks”. The univariate measures almost obtaining 1 as highest value possible in Setting 1 as well as the high value for Setting 3 that deals with two locations 715 km remote confirm this. Therefore, since ordinal pattern dependence is known to be articulately more robust we propose ordinal pattern dependence to be the less sensitive and more appropriate measure here.

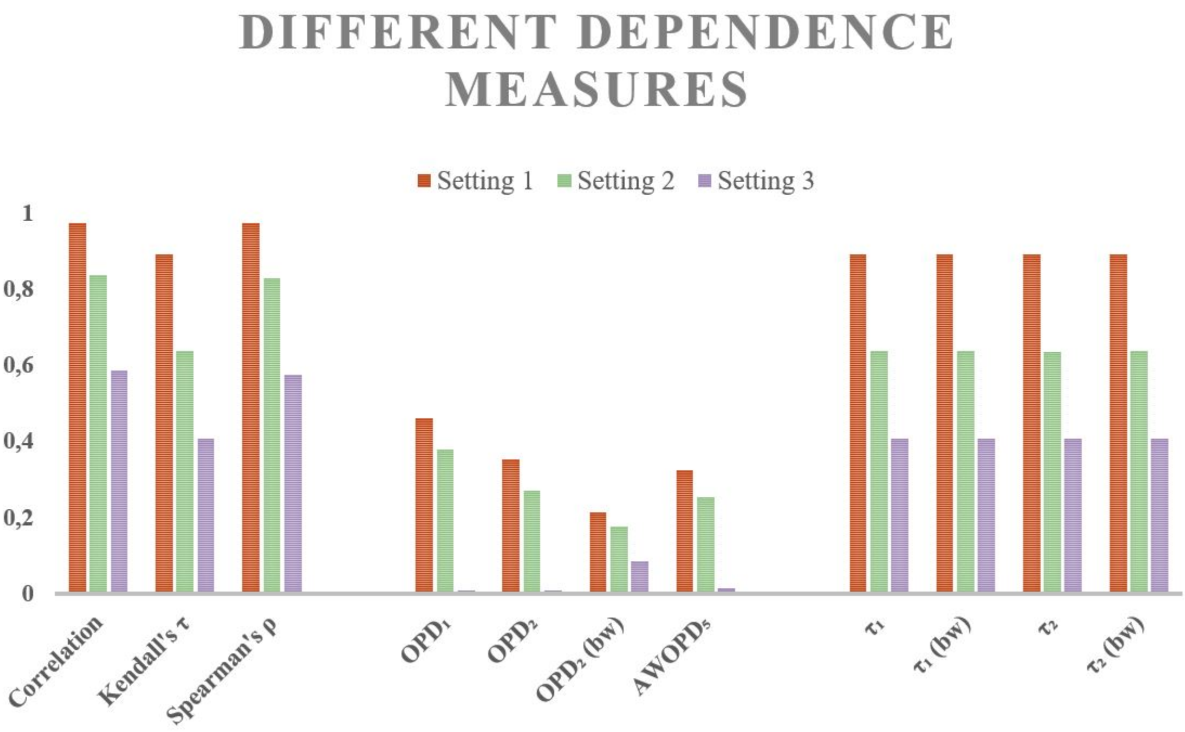

Figure 6.5: Dependence measures for different settings.

The other multivariate measure considered here is given by multivariate Kendall's τ_h for $h = 1$ and $h = 2$, where we used both, the stepwise as well as the blockwise estimator. We do not register an improvement compared to univariate Kendall's τ (that corresponds to $h = 0$ in terms of the multivariate approach). Therefore, this dependence measure provides no further amendment in the data sets investigated and especially due to the large computational quadratic effort has no practical impact here. Hence, we get more into the details of the different estimators of ordinal pattern dependence.

Regarding Figure 6.5, we observe that the values obtained using the stepwise estimators for $h = 1$ and $h = 2$ as well as the estimator of average-weighted ordinal pattern dependence for $h = 5$ have the same structure, although $h = 1$ yields higher values due to the theoretical definition. The similar values of OPD_2 and $AWOPD_5$ show that we obtain a balance for these choices of pattern lengths between absolutely coincident patterns and certain deviations. The values of blockwise estimation for OPD_2 fall out of line and provide values with smaller deviation between the settings. Having the background parameters as the distance between the locations in mind we, therefore, propose the stepwise approaches here.

We also studied time-shifted ordinal pattern dependence here. Remark that for each shift of length 1 the sample size decreases by 1 and therefore, we made the restriction that the sample size for each setting regarded has to be larger than 20000 to get comparable results. In practice this is a negligible condition, since we intuitively might expect a larger dependence if we shift the data sets by a few days due to the distance, but not by more than 50 years.

An interesting finding here is that we obtain a different result using time-shifted ordinal pattern

Setting 1:

	Standard OPD_h	Time-shifted OPD_h with $a = 1$
$h = 1$	0.4626	0.6388
$h = 2$	0.3542	0.5167

Table 6.3: Standard OPD vs. time-shifted OPD.

dependence in Setting 1, in the other two cases the differences are neglectable (due to the closeness in setting 2 and the large distance and the different rivers in setting 3). We obtain a maximum of ordinal pattern dependence for a time shift of $a = 1$, concerning the data of Wittenberge from 01/11/1899 to 30/12/2013 and the data of Neu Darchau from 02/11/1899-31/12/2013. Note that this shift of one day coincides with the direction of the Elbe from east to west. The different values are displayed in Table 6.3.

We turn to the computational cost of the respective dependence measures. In Table 6.4 we observe that the algorithms we use to calculate the values of the estimators in the context of ordinal pattern dependence of the data sets are outstandingly fast compared to the computational effort of multivariate Kendall's τ. The computational effort for estimators in the context of multivariate Kendall's τ is quadratic, see (5.1) and (5.2), whereas the computational cost for estimators in the context of ordinal pattern dependence is linear. The last detail we want to take a closer

OPD_1, OPD_2	0.22 sec	$\tau_{2,\mathrm{bw}}$	5.96 min
$OPD_{2,\mathrm{bw}}$	0.07 sec	$\tau_{1,\mathrm{bw}}$	13.67 min
$AWOPD_5$	34.46 sec	τ_2	51.69 min
Time-shifted OPD_2	2.38 min	τ_1	59.77 min

Table 6.4: Computational cost of the different empirical dependence measures for Setting 1.

look at is the distribution of the single possible patterns compared to all patterns found. The estimators for this approach are given in Section 4.2 and Remark 4.22 for the one-dimensional time series and in Section 4.4.1, stationary case, for both time series. For each setting, we look at the patterns found in each single time series, respectively and finally at the pattern found in both time series at same points in time. A table with the absolute values of pattern found is given in the Appendix in Table B.1 and Table B.2. The percentaged results are shown in Figure 6.6. We observe that within each but also between the different settings, we have close distributions of the single patterns in the cases where we consider the single data sets, where the largest percentage of almost 50% is obtained by the pattern $(0, 1, 2)$, followed by approx. 25% of $(2, 1, 0)$. Interestingly coming to the last column in Figure 6.6, the share of $(0, 1, 2)$ significantly increases in all three settings, while the one of $(2, 1, 0)$ only changes slightly. The pattern dealing with a changing sign of the increments almost have a negligible share in the last column.

Relative frequency of stepwise counted ordinal patterns for $h = 2$:

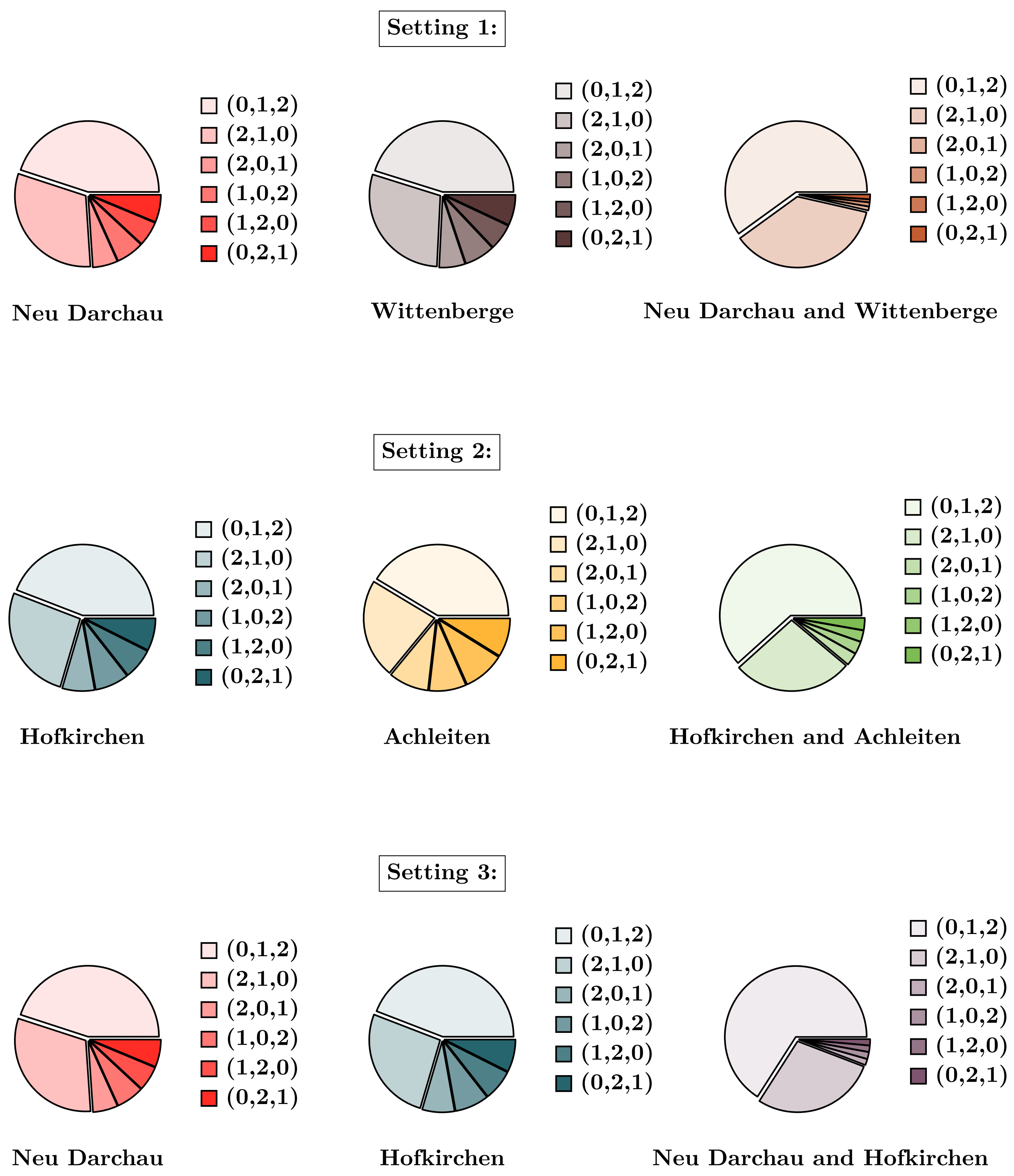

Figure 6.6: Relative frequency of stepwise counted ordinal patterns for $h = 2$.

For $h = 1$ the phenomenon as described above can be observed in the Appendix in Figure B.1 However, since only two patterns appear in this setting, it is not as distinctive as for $h = 2$. A remarkable finding is that with blockwise estimation of ordinal pattern dependence, which does not yield satisfying results concerning the concrete values of ordinal pattern dependence here, nontheless a very similar distribution of the single patterns is derived, as we observe comparing Figure 6.6 and Figure B.2 in the Appendix. Due to the smaller computational cost, if one is interested in the properties of the distributions, we propose the blockwise estimation here.
Finally we see in the Appendix in Figure B.3 the ratio of coincident patterns to the theoretical number of coincident patterns that could have been obtained. This yields an impression of the dependence, however, note that this is **not** a dependence measure and must not be mixed up with ordinal pattern dependence, since for example independence would yield a value of zero here. However, the same properties for stepwise and blockwise estimation for the distribution of the ordinal patterns are confirmed here, too.

We close this data analysis with the conclusion that ordinal pattern dependence proved itself true to be a dependence measure with a high impact on real world data due to its outstanding property of robustness. It confirms to be a suitable measure for this kind of data that naturally often contain external disturbances since ordinal pattern dependence does not overrate the value of single data points as the considered univariate dependence measures happen to do. However, even the multivariate extensions of these dependence measures do not yield an improvement here. The possibility to apply ordinal pattern dependence to stationary and long-range dependent data, with having the theoretical properties derived in the previous chapters of this work in mind, yield a practical improvement for this class of data sets.

7 Conclusion and outlook

In the last years an extensive framework for dependence analysis using ordinal patterns has been developed. Ordinal pattern probabilities in short- as well as in long-range dependence time series as a multi-dimensional problem that is reducable to a univariate time series were discussed in [60], [10] and [58]. The latter one also dealt with ordinal pattern dependence in the short-range dependence setting.

This thesis complemented these results by providing the extension of ordinal pattern dependence to the long-range dependent case as well as to mixed cases that allow for both short-range and for long-range dependent components within the multivariate Gaussian time series under consideration. The necessary limit theorems were developed in Chapter 3. We no longer had the reduction possibility to the asymptotics of the sample covariances of a univariate time series but developed asymptotic results for the multivariate sample cross correlations. Therefore, we successfully reached our aim to make ordinal pattern dependence applicable and interesting for long-range dependent real-world data. We provided asymptotic information on the estimators of ordinal pattern dependence. Furthermore, we illustrated the theoretical findings with a simulation study that shed light on asymptotic distributions in the context of multivariate Hermite polynomials. The practical utility value of the results derived for estimators of ordinal pattern dependence can be increased by using them to develop tests for structural breaks in the dependence structure between the two time series. Furthermore, asymptotic confidence intervals in the context of hypothesis testing can be constructed. We classified ordinal pattern dependence into the framework of multivariate dependence measures. From the user's perspective this a further benefit, since the theoretical results derived allow for a better interpretation of the respective values of the different dependence measures.

One possible extension for future research is to soften the restriction in Definition 2.13 to general slowly varying functions, as it is given in the univariate case in Definition 2.6. It might be an idea to generalize the normalizing constants for the univariate case as given in [9], Sec. 4.2.2, to the multivariate setting. One has to keep the framework developed in [37] with regard to the implications of different definitions of multivariate long-range dependence in mind. A further extension is to take a closer look at the case $d^* = \frac{1}{4}$ for functionals with Hermite rank 2. The univariate case is discussed in [32], Theorem 4, which yields a Gaussian limit distribution, however, using a different normalization. For details, see [18], p. 58.

A very interesting generalization of the limit theorems in Chapter 3, as well as their applications to ordinal pattern analysis, is given by softening the Gaussian assumption and considering multivariate linear processes in general, that were extensively discussed in [37] and limit theorems were provided in [23]. Several challenges arise in the proofs of in this work: the main problem is, that there is no generalization of the reduction theorem at hand. This highly relies on the Gaussian assumption and the corresponding property of the Gaussian distribution to be uniquely determined by the first and second moment. Considering the application to ordinal pattern analysis, we observe that the Rao-Blackwellization, that yields the improved estimators $\hat{q}_{imp,n,\pi}$ and $\hat{p}_{imp,n}$ cannot be generalized anymore due to the strong need of the symmetry properties of the multivariate Gaussian distribution.

Concerning the second part of this chapter dealing with ordinal pattern dependence in contrast to other univariate as well as multivariate dependence measures, we have seen that one interesting further research opportunity arises in the investigation of non-linear dependencies, as they appeared for example in the limit processes in [3], Proposition 3.5 and [23], Theorem 4.1. Another possibility to describe those kinds of dependence is given by copulas, see [51]. In this approach information concerning the marginal distributions is ignored and the copula only provides the dependence structure of a multivariate random vector. Notice that the same challenge as mentioned in the beginning of Chapter 5 arises here, namely that one has to distinguish carefully between a multivariate dependence measure which was invented for measuring dependence within one single random vector, see in this context [56] and the moving-window time series approach, forming the basis of ordinal pattern dependence. Therefore, a first step is to make both approaches comparable before comparing their outcomes.
The last ongoing research we want to mention is ordinal pattern dependence applied to extreme value theory. The investigation of the distribution of ordinal patterns that occur over a certain threshold was empirically done in [50]. In [52], the distribution of clusters of extremes and ordinal patterns as well as asymptotic results for empirical estimators were provided, completed with the application to data of Rhine river discharge at Cologne.

A Details of some limit distributions

In this chapter we provide further information on the limit distributions arising in Chapter 4.

A.1 Hermite coefficients of $\hat{q}_n(\pi)$ for $h = 2$ for the pattern $\pi = (2, 1, 0)$ using the Cholesky decomposition

We recall Example 4.16. Since we look at $h = 2$, the covariance matrix of $Y_{1,2} = (Y_1, Y_2)^t$ is given by

$$\Sigma_2 = \begin{pmatrix} 1 & r^{(1,1)}(1) \\ r^{(1,1)}(1) & 1 \end{pmatrix}.$$

The Cholesky decomposition $\Sigma = AA^t$ has the following form:

$$A = \begin{pmatrix} 1 & 0 \\ r^{(1,1)}(1) & \sqrt{1 - (r^{(1,1)}(1))^2} \end{pmatrix}$$

Note that $Y_{1,2} = AU_{1,2}$, where $U_{1,2} = (U_1, U_2)^t$ has a bivariate standard normal distribution. Following Theorem 4.13, we need to calculate $\alpha = \left(A^{-1}\right)^t b$, where $b = \mathbb{E}\left(Y_{1,2} 1_{\{\tilde{\Pi}(Y_1, Y_2) = \pi\}}\right)$. Since

$$\left(A^{-1}\right)^t = \begin{pmatrix} 1 & -\frac{r^{(1,1)}(1)}{\sqrt{1-(r^{(1,1)}(1))^2}} \\ 0 & \frac{1}{\sqrt{1-(r^{(1,1)}(1))^2}} \end{pmatrix}$$

we need to determine b to calculate the variance in the limit distribution. We consider $\pi = (2, 1, 0)$. From the Cholesky decomposition it follows that $Y_1 = U_1$ and $Y_2 = r^{(1,1)}(1)U_1 + \sqrt{1 - (r^{(1,1)}(1))^2} U_2$ and, therefore, $c_1 = \mathbb{E}\left(Y_1 1_{\{\tilde{\Pi}(Y_1, Y_2) = \pi\}}\right) = b_1$ and

$$c_2 = r^{(1,1)}(1) b_1 + \sqrt{1 - (r^{(1,1)}(1))^2} b_2.$$

For this choice of π we also know by (4.6) and (4.7) that $c_1 = c_2$ and hence we arrive at

$$b_1 = \frac{\sqrt{1 - (r^{(1,1)}(1))^2}}{1 - r^{(1,1)}(1)} b_2.$$

Therefore, it is sufficient to only determine b_2. For this, we rewrite

$$\begin{aligned} \{\tilde{\Pi}(Y_1, Y_2) = (2, 1, 0)\} &= \{Y_1 \geq 0, Y_2 \geq 0\} = \{U_1 \geq 0, r^{(1,1)}(1)U_1 + \sqrt{1 - (r^{(1,1)}(1))^2} U_2 \geq 0\} \\ &= \{U_1 \geq 0, U_2 \geq -\frac{r^{(1,1)}(1)}{\sqrt{1 - (r^{(1,1)}(1))^2}} U_1\}. \end{aligned}$$

Hence, we need to determine

$$
\begin{aligned}
b_2 &= \mathbb{E}\left(U_2 1_{\{\tilde{\Pi}(Y_1,Y_2)=\pi\}}\right) \\
&= \int_0^\infty \int_{-\frac{r^{(1,1)}(1)}{\sqrt{1-(r^{(1,1)}(1))^2}}U_1}^\infty u_2\varphi(u_2)\varphi(u_1)\mathrm{d}u_2\mathrm{d}u_1 \\
&= \int_0^\infty \varphi\left(\frac{r^{(1,1)}(1)}{\sqrt{1-(r^{(1,1)}(1))^2}}u_1\right)\varphi(u_1)\mathrm{d}u_1 \\
&= \frac{1}{2\pi}\int_0^\infty \exp\left(-\frac{\left(1+\frac{(r^{(1,1)}(1))^2}{1-(r^{(1,1)}(1))^2}\right)u_1^2}{2}\right)\mathrm{d}u_1 \\
&= \frac{1}{2\pi}\int_0^\infty \exp\left(-\frac{\left(\frac{1}{1-(r^{(1,1)}(1))^2}\right)u_1^2}{2}\right)\mathrm{d}u_1 \\
&= \frac{1}{2\sqrt{2\pi}}\sqrt{1-(r^{(1,1)}(1))^2}.
\end{aligned}
$$

Finally, we obtain

$$
\begin{aligned}
\sum_{j=1}^2 \alpha_j &= b_1 + \frac{1-r^{(1,1)}(1)}{\sqrt{1-(r^{(1,1)}(1))^2}}b_2 \\
&= \frac{\sqrt{1-(r^{(1,1)}(1))^2}}{1-r^{(1,1)}(1)}b_2 + \frac{1-r^{(1,1)}(1)}{\sqrt{1-(r^{(1,1)}(1))^2}}b_2 \\
&= \left(\sqrt{\frac{1+r^{(1,1)}(1)}{1-r^{(1,1)}(1)}} + \sqrt{\frac{1-r^{(1,1)}(1)}{1+r^{(1,1)}(1)}}\right)b_2 \\
&= \frac{1}{2\sqrt{2\pi}}\frac{2}{\sqrt{1-(r^{(1,1)}(1))^2}}\sqrt{1-(r^{(1,1)}(1))^2} \\
&= \frac{1}{\sqrt{2\pi}}.
\end{aligned}
$$

As a result, we confirm the result from Example 4.16 for the pattern $\pi = (2,1,0)$. For $\pi = (2,0,1)$, the analytical calculations work analogously.

A.2 Table of Hermite coefficients for $\hat{p}_n(\pi)$

π	$\mathcal{S}(\pi)$	$\mathcal{T}(\pi)$ $(\mathcal{S}(\pi))$	$\mathcal{T}\circ\mathcal{S}(\pi)$ (π)
$c_1^{\pi,2}$	$c_1^{\mathcal{S}(\pi),2}$	$c_1^{\mathcal{T}(\pi),2}$ $(c_1^{\mathcal{S}(\pi),2})$	$c_1^{\mathcal{T}\circ\mathcal{S}(\pi),2}$ $(c_1^{\pi,2})$
$c_2^{\pi,2}$	$c_2^{\mathcal{S}(\pi),2}$	$c_2^{\mathcal{T}(\pi),2}$ $(c_2^{\mathcal{S}(\pi),2})$	$c_2^{\mathcal{T}\circ\mathcal{S}(\pi),2}$ $(c_2^{\pi,2})$
$c_3^{\pi,2}$	$c_3^{\mathcal{S}(\pi),2}$	$c_3^{\mathcal{T}(\pi),2}$ $(c_3^{\mathcal{S}(\pi),2})$	$c_3^{\mathcal{T}\circ\mathcal{S}(\pi),2}$ $(c_3^{\pi,2})$
$c_4^{\pi,2}$	$c_4^{\mathcal{S}(\pi),2}$	$c_4^{\mathcal{T}(\pi),2}$ $(c_4^{\mathcal{S}(\pi),2})$	$c_4^{\mathcal{T}\circ\mathcal{S}(\pi),2}$ $(c_4^{\pi,2})$

Table A.1: Hermite coefficients for Hermite rank 1 for each pattern of the group $\bar{\pi}$.

We refer to (4.43). In Table A.1 we consider the case $h = 2$ and emphasize the importance of the symmetry properties. Note that each entry of the respective vector of Hermite coefficients is a 4-dimensional integral, keeping the computational cost in mind. Therefore, without considering any of the properties listed above, we would have to compute 16 of these four-dimensional integrals, only in the case $d = 2$ and $h = 2$. In Table A.2 it is shown under which assumptions it is possible to significantly reduce the number of integrals

With (I) using $c_i^{\pi,2} = -c_i^{\mathcal{S}(\pi),2}$, $i = 1, \ldots, 4$

π	$\mathcal{S}(\pi)$	$\mathcal{T}(\pi)$ $(\mathcal{S}(\pi))$	$\mathcal{T} \circ \mathcal{S}(\pi)$ (π)
$c_1^{\pi,2}$	$-c_1^{\pi,2}$	$c_1^{\mathcal{T}(\pi),2}$ $(-c_1^{\pi,2})$	$-c_1^{\mathcal{T}(\pi),2}$ $(c_1^{\pi,2})$
$c_2^{\pi,2}$	$-c_2^{\pi,2}$	$c_2^{\mathcal{T}(\pi),2}$ $(-c_2^{\pi,2})$	$-c_2^{\mathcal{T}(\pi),2}$ $(c_2^{\pi,2}$
$c_3^{\pi,2}$	$-c_3^{\pi,2}$	$c_3^{\mathcal{T}(\pi),2}$ $(-c_3^{\pi,2})$	$-c_3^{\mathcal{T}(\pi),2}$ $(c_3^{\pi,2})$
$c_4^{\pi,2}$	$-c_4^{\pi,2}$	$c_4^{\mathcal{T}(\pi),2}$ $(-c_4^{\pi,2})$	$-c_4^{\mathcal{T}(\pi),2}$ $(c_4^{\pi,2})$

With (I), (II) using $c_i^{\pi,2} = c_{5-i}^{\mathcal{T}\circ\mathcal{S}(\pi),2}$ $(c_i^{\pi,2} = c_{5-i}^{\pi,2})$, $i = 1, \ldots 4$:

π	$\mathcal{S}(\pi)$	$\mathcal{T}(\pi)$ $(\mathcal{S}(\pi))$	$\mathcal{T} \circ \mathcal{S}(\pi)$ (π)
$c_1^{\pi,2}$	$-c_1^{\pi,2}$	$-c_4^{\pi,2}$ $(-c_1^{\pi,2})$	$c_4^{\pi,2}$ $(c_1^{\pi,2})$
$c_2^{\pi,2}$	$-c_2^{\pi,2}$	$-c_3^{\pi,2}$ $(-c_2^{\pi,2})$	$c_3^{\pi,2}$ $(c_2^{\pi,2})$
$c_3^{\pi,2}$ $(c_2^{\pi,2})$	$-c_3^{\pi,2}$ $(-c_2^{\pi,2})$	$-c_2^{\pi,2}$	$c_2^{\pi,2}$
$c_4^{\pi,2}$ $(c_1^{\pi,2})$	$-c_4^{\pi,2}$ $(-c_1^{\pi,2})$	$-c_1^{\pi,2}$	$c_1^{\pi,2}$

With (I),(III) using $c_k^{\pi,2} = c_{3-k}^{\mathcal{T}\circ\mathcal{S}(\pi),2}$ $(c_k^{\pi,2} = c_{3-k}^{\pi,2})$, $k = 1, 2$:

π	$\mathcal{S}(\pi)$	$\mathcal{T}(\pi)$ $(\mathcal{S}(\pi))$	$\mathcal{T} \circ \mathcal{S}(\pi)$ (π)
$c_1^{\pi,2}$	$-c_1^{\pi,2}$	$-c_2^{\pi,2}$ $(-c_1^{\pi,2})$	$c_2^{\pi,2}$ $(c_1^{\pi,2})$
$c_2^{\pi,2}$ $(c_1^{\pi,2})$	$-c_2^{\pi,2}$ $(-c_1^{\pi,2})$	$-c_1^{\pi,2}$	$c_1^{\pi,2}$
$c_3^{\pi,2}$	$-c_3^{\pi,2}$	$-c_4^{\pi,2}$ $(-c_3^{\pi,2})$	$c_4^{\pi,2}$ $(c_3^{\pi,2})$
$c_4^{\pi,2}$ $(c_3^{\pi,2})$	$-c_4^{\pi,2}$ $(-c_3^{\pi,2})$	$-c_3^{\pi,2}$	$c_3^{\pi,2}$

With (I),(II),(III) using additionally $c_k^{\pi,2} = c_{2+k}^{\pi,2}$, $k = 1, 2$:

π	$\mathcal{S}(\pi)$	$\mathcal{T}(\pi)$ $(\mathcal{S}(\pi))$	$\mathcal{T} \circ \mathcal{S}(\pi)$ (π)
$c_1^{\pi,2}$	$-c_1^{\pi,2}$	$-c_2^{\pi,2}$ $(-c_1^{\pi,2})$	$c_2^{\pi,2}$ $(c_1^{\pi,2})$
$c_2^{\pi,2}$ $(c_1^{\pi,2})$	$-c_2^{\pi,2}$ $(-c_1^{\pi,2})$	$-c_1^{\pi,2}$	$c_1^{\pi,2}$
$c_1^{\pi,2}$	$-c_1^{\pi,2}$	$-c_2^{\pi,2}$ $(-c_1^{\pi,2})$	$c_2^{\pi,2}$ $(c_1^{\pi,2})$
$c_2^{\pi,2}$ $(c_1^{\pi,2})$	$-c_2^{\pi,2}$ $(-c_1^{\pi,2})$	$-c_1^{\pi,2}$	$c_1^{\pi,2}$

Table A.2: Hermite coefficients for Hermite rank 1 for each pattern of the group $\bar{\pi}$ using different assumptions on the dependence structure on the underlying Gaussian vector.

A.3 Description of the Matlab algorithms

To calculate the limit variance in Theorem 4.26 and Theorem 4.27 for $h = 2$ two Matlab functions were implemented, following the efficiency criterions we derived in Section 4.3.1 to keep the computational cost low.

- *cHR2*:
  This Matlab function provides the matrix $\tilde{\alpha}_{\pi,2}^{(p,q)}$, $p, q = 1, 2$, given in Theorem 4.26 for a pattern $\pi \in S_2$. The respective pattern is passed as input parameter as well as the entries of the correlation matrix $\Sigma_{2,2}$. According to these entries, the conditions (II) and/or (III) are verified so that the matrix is computed in an efficient way.

- *HR2LimVar*:
  This function depends on the entries of $\Sigma_{2,2}$, to get a satisfactory result concerning the computational cost. Another input parameter is the matrix $(L_{p,q})_{p,q=1,2}$, as limit of the slowly varying functions as given in (3.2). The Matlab function computes the matrix $\tilde{\alpha}^{(p,q)}$ as defined in (4.24) to (4.26). A further output is the limit variance in Theorems 4.26 and 4.27.

To get further information on the limit variance in Theorem 4.52 for $h = 2$ we present two more Matlab algorithms based on the considerations in Section 4.4.1.

- *cHR1*:
  The input parameters of the Matlab function are given in the covariance matrix $\Sigma_{2,2}$ to use all possibilities to reduce the computational cost. Furthermore, the pattern $\pi \in S_2$ of interest and the matrix of the limits of the slowly-varying functions $(L_{p,q})_{p,q=1,2}$ are passed. The possible output parameters are the vector of Hermite coefficients, the vector $\left(\tilde{\alpha}_{\pi,2}^{(1)}, \tilde{\alpha}_{\pi,2}^{(2)}\right)^t$ and the total limit variance of Theorem 4.52

- *HR1allpattern*:
  This function can be used if one is interested in the limit variances for all permutations in S_2. It is possible to reduce the computational cost compared to using *cHR1* for all $\#S_2 = 6$ pattern significantly.

A.4 Simulation study

We extend the simulation study in Section 4.3.5. Since the asymptotic distributions of $\hat{r}_n^{(1,1)}(0) - r_H^{(1,1)}(0)$ and $\hat{r}_n^{(1,2)}(0) - r_H^{(1,2)}(0)$ do not coincide for $H > \frac{3}{4}$ we turn to the extreme case that $H = 1$ and we consider two independent standard normal random variables $U_1^{(1)}$ and $U_2^{(1)}$ and use $N = 50000$ repetitions. Following [66], p. 983, the Rosenblatt distribution approaches the χ^2-distribution with mean 0 and variance 1. This is shown in Figure A.1. We observe the χ^2-distribution for $H_2\left(U_1^{(1)}\right)$. For $H_1\left(U_1^{(1)}\right) H_1\left(U_1^{(2)}\right)$ a distribution arises that differs significantly from the χ^2-distribution.

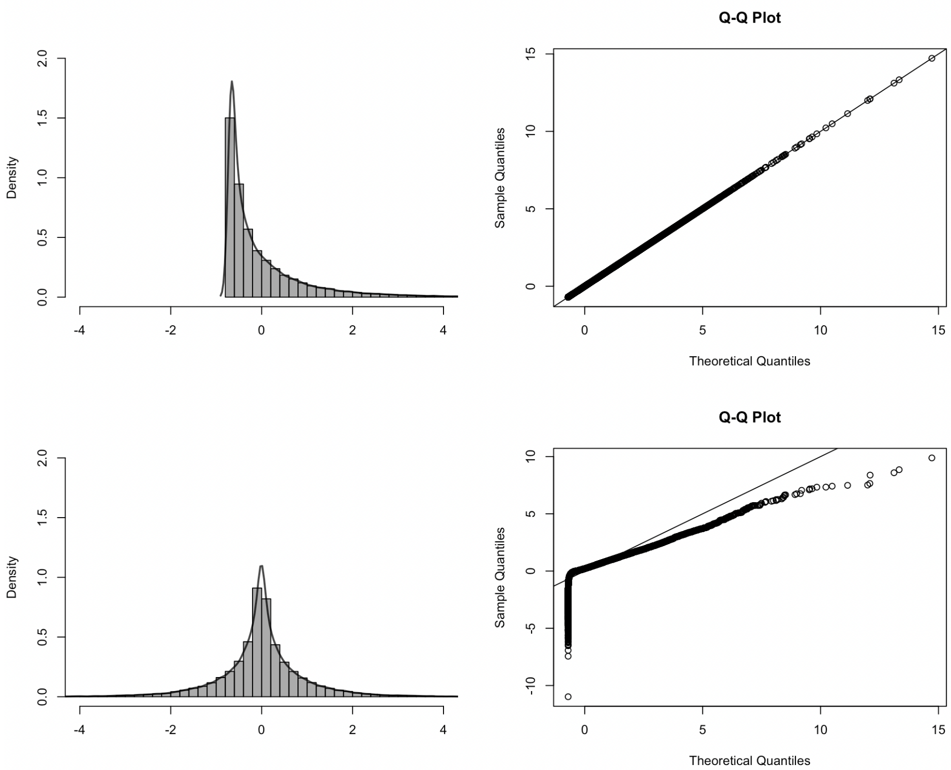

Figure A.1: Histogram, kernel density estimation and Q-Q plot with respect to the standardized χ^2-distribution of appropriately normalized $H_2\left(U_1^{(1)}\right)$ (top) and $H_1\left(U_1^{(1)}\right) H_1\left(U_1^{(2)}\right)$ (bottom).

B Further information on the real-world data analysis

B.1 Absolute number of pattern in the real-world data analysis for different measuring stations

Stepwise:

Pattern	Neu-Darchau	Wittenberge	both	Hofkirchen	Achleiten	both	Neu-Darchau	Hofkirchen	both
$(0,1)$	23747	24207	18500	24349	24542	18522	23539	24349	13937
$(1,0)$	17950	17490	12243	16983	16790	10696	17793	16983	7381
$(0,1,2)$	18752	18833	13847	18244	17109	11887	18597	18244	8266
$(2,1,0)$	12956	12117	8420	10877	9356	5258	12852	10877	3545
$(2,0,1)$	2375	2466	200	3043	3781	560	2351	3043	174
$(1,0,2)$	2638	2891	210	3208	3489	542	2614	3208	196
$(1,2,0)$	2356	2482	154	2898	3945	494	2327	2898	180
$(0,2,1)$	2620	2908	256	3062	3652	534	2591	3062	177

Table B.1: Absolute number of ordinal patterns occured in each setting for stepwise ordinal pattern dependence.

Blockwise:

Pattern	Neu-Darchau	Wittenberge	both	Hofkirchen	Achleiten	both	Neu-Darchau	Hofkirchen	both
$(0, 1, 2)$	9379	9412	6900	9138	8575	5966	9296	9138	4141
$(2, 1, 0)$	6480	6054	4195	5455	4699	2616	6423	5455	1778
$(2, 0, 1)$	1197	1220	96	1525	1895	276	1166	1525	91
$(1, 0, 2)$	1342	1456	108	1617	1663	259	1284	1617	96
$(1, 2, 0)$	1152	1262	81	1428	1995	261	1189	1428	91
$(0, 2, 1)$	1298	1444	112	1502	1838	259	1307	1502	92

Table B.2: Absolute number of ordinal patterns occured in each setting for blockwise ordinal pattern dependence.

B.2 Frequency of ordinal patterns for $h = 1$

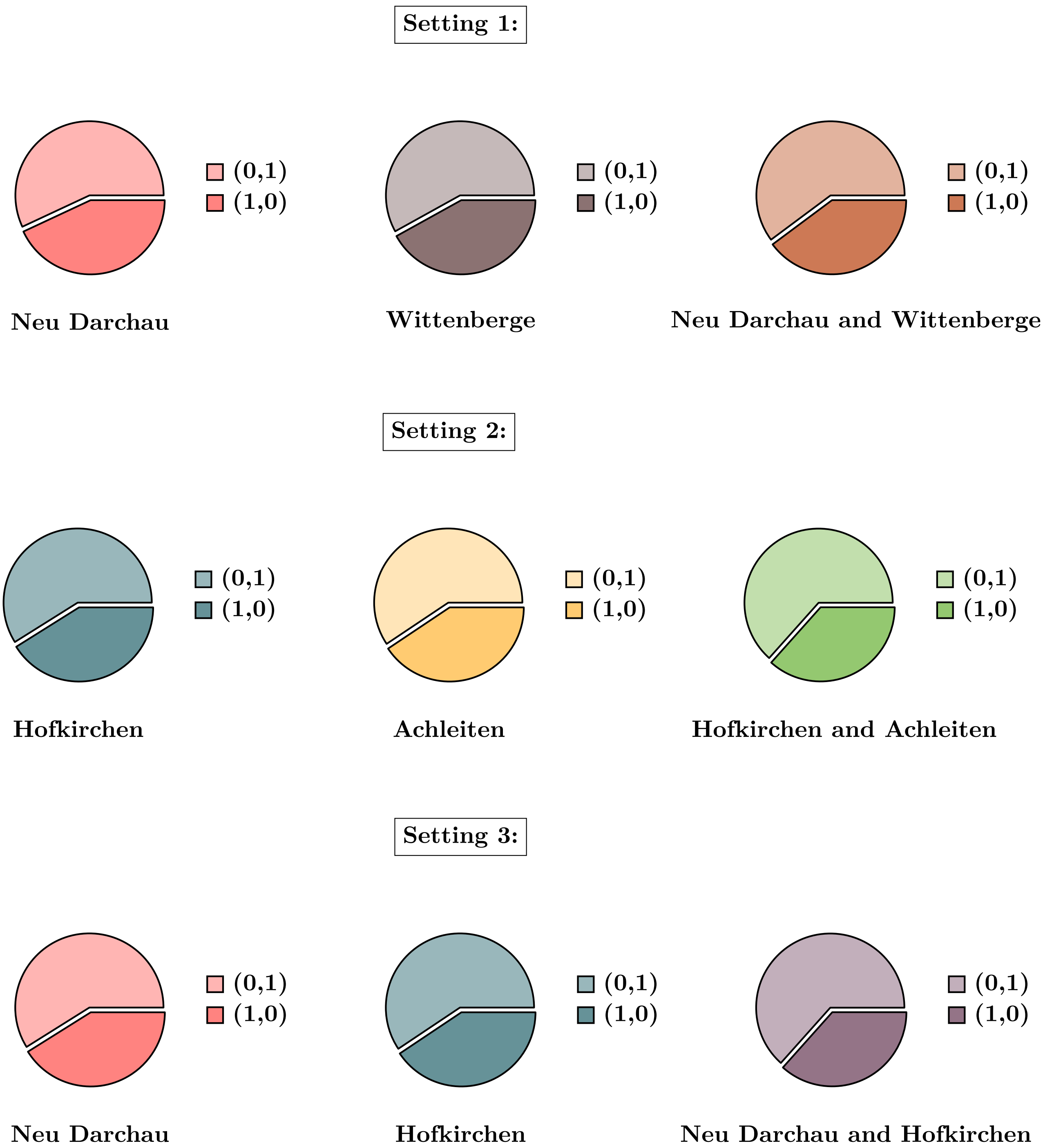

Figure B.1: Frequency of ordinal patterns for $h = 1$.

B.3 Relative frequency of blockwise counted ordinal patterns for $h = 2$

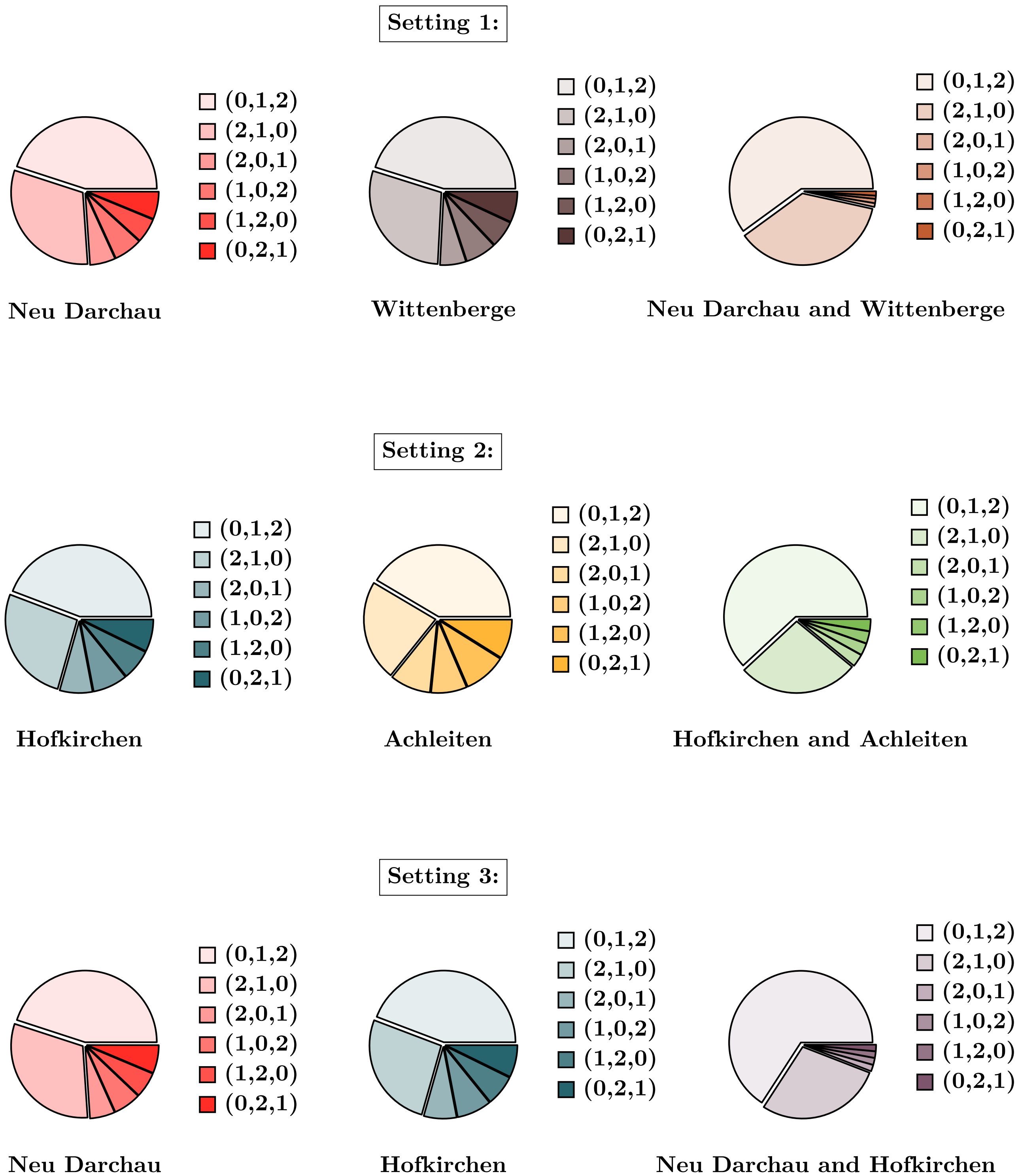

Figure B.2: Relative frequency of blockwise counted ordinal patterns for $h = 2$.

B.4 Ratio of coincident patterns to sample size

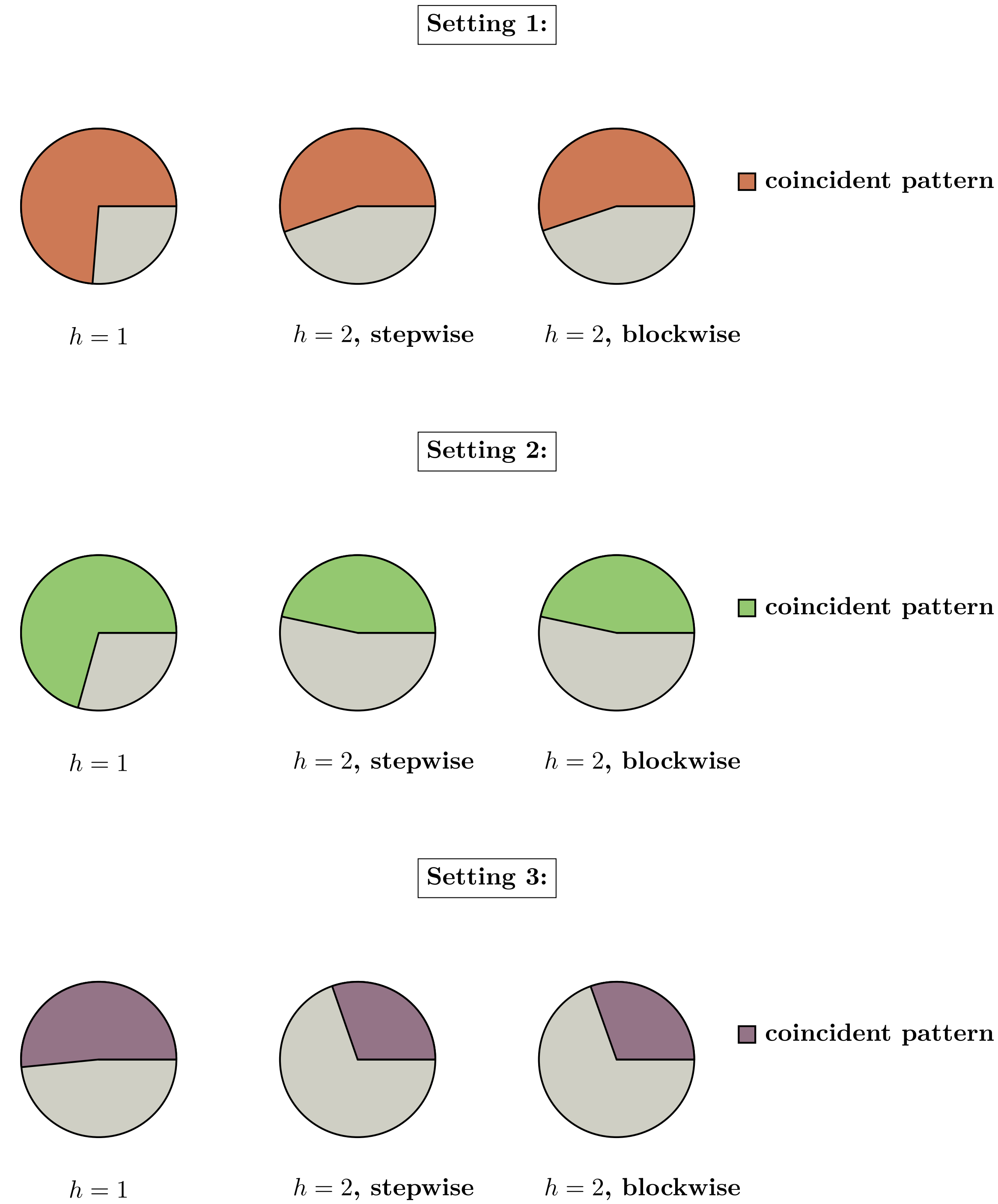

Figure B.3: Ratio of coincident patterns to sample size..

Bibliography

[1] I.G. Abrahamson. “Orthant Probabilities for the Quadrivariate Normal Distribution”. In: *The Annals of Mathematical Statistics* 35.4 (1964), pp. 1685–1703.

[2] Miguel A. Arcones. “Limit Theorems for Nonlinear Functionals of a Stationary Gaussian Sequence of Vectors”. In: *The Annals of Probability* (1994), pp. 2242–2274.

[3] Shuyang Bai and Murad S. Taqqu. “Multivariate Limit Theorems In The Context Of Long-Range Dependence”. In: *Journal of Time Series Analysis* 34.6 (2013), pp. 717–743.

[4] Christoph Bandt. “Ordinal time series analysis”. In: *Ecological Modelling* 182.3-4 (2005), pp. 229–238.

[5] Christoph Bandt and Bernd Pompe. “Permutation Entropy: A Natural Complexity Measure for Time Series”. In: *Physical Review Letters* 88.17 (2002), p. 174102.

[6] Christoph Bandt and Faten Shiha. “Order Patterns in Time Series”. In: *Journal of Time Series Analysis* 28.5 (2007), pp. 646–665.

[7] Heinz Bauer. *Measure and Integration Theory.* Vol. 26. Walter de Gruyter, 2011.

[8] Francois Belletti, Minmin Chen, and Ed H. Chi. “Quantifying Long Range Dependence in Language and User Behavior to improve RNNs”. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.* 2019, pp. 1317–1327.

[9] Jan Beran, Yuanhua Feng, Sucharita Ghosh, and Rafał Kulik. *Long-Memory Processes.* Springer-Verlag Berlin Heidelberg, 2013.

[10] Annika Betken, Jannis Buchsteiner, Herold Dehling, Ines Münker, Alexander Schnurr, and Jeannette H.C. Woerner. “Ordinal Patterns in Long-Range Dependent Time Series”. In: *arXiv preprint arXiv:1905.11033* (2019).

[11] Annika Betken, Herold Dehling, Ines Nüßgen, and Alexander Schnurr. “Ordinal pattern dependence as a multivariate dependence measure”. In: *arXiv preprint arXiv:2012.02445* (2021).

[12] Albrecht Beutelspacher. *Lineare Algebra.* Springer, 2003.

[13] Martin Bilodeau and David Brenner. *Theory of Multivariate Statistics.* Springer Science & Business Media, 2008.

[14] David Blackwell. “Conditional Expectation and Unbiased Sequential Estimation”. In: *The Annals of Mathematical Statistics* (1947), pp. 105–110.

[15] Peter Breuer and Péter Major. “Central Limit Theorems for Non-Linear Functionals of Gaussian Fields”. In: *Journal of Multivariate Analysis* 13.3 (1983), pp. 425–441.

[16] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods.* Springer-Verlag New York, Inc., 1991.

[17] Krzysztof Burnecki. “FARIMA processes with application to biophysical data”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2012.05 (2012), P05015.

[18] Ching-Fan Chung. “Sample Means, Sample Autocovariances, and Linear Regression of Stationary Multivariate Long Memory Processes”. In: *Econometric Theory* 18.1 (2002), pp. 51–78.

[19] Hongshuai Dai, Guangjun Shen, and Lingtao Kong. “Limit theorems for functionals of Gaussian vectors”. In: *Frontiers of Mathematics in China* 12.4 (2017), pp. 821–842.

[20] David A. Dickey and Wayne A. Fuller. “Distribution of the Estimators for Autoregressive Time Series With a Unit Root”. In: *Journal of the American Statistical Association* 74.366a (1979), pp. 427–431.

[21] Gustavo Didier and Vladas Pipiras. “Integral representations and properties of operator fractional Brownian motions”. In: *Bernoulli* 17.1 (2011), pp. 1–33.

[22] Marie-Christine Düker. “Limit Theorems for Multivariate Long-Range Dependent Processes”. In: *arXiv preprint arXiv:1704.08609v3* (2017).

[23] Marie-Christine Düker. “Limit theorems in the context of multivariate long-range dependence”. In: *Stochastic Processes and their Applications* 130.9 (2020), pp. 5394–5425.

[24] Svenja Fischer, Andreas Schumann, and Alexander Schnurr. “Ordinal pattern dependence between hydrological time series”. In: *Journal of Hydrology* 548 (2017), pp. 536–551.

[25] Robert Fox and Murad S. Taqqu. “Multiple stochastic integrals with dependent integrators”. In: *Journal of Multivariate Analysis* 21.1 (1987), pp. 105–127.

[26] Eduardo Garcia Portugues. *Nonparametric Statistics.* URL: `https://bookdown.org/egarpor/NP-UC3M/kde-ii-mult.html`.

[27] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* Vol. 3. JHU press, 2013.

[28] G. Graff, B. Graff, A. Kaczkowska, D. Makowiec, J.M. Amigó, J. Piskorski, K. Narkiewicz, and P. Guzik. “Ordinal pattern statistics for the assessment of heart rate variability”. In: *The European Physical Journal Special Topics* 222.2 (2013), pp. 525–534.

[29] Oliver Grothe, Julius Schnieders, and Johan Segers. “Measuring association and dependence between random vectors”. In: *Journal of Multivariate Analysis* 123 (2014), pp. 96–110.

[30] Allan Gut. *Probability: A Graduate Course.* Vol. 75. Springer Science & Business Media, 2013.

[31] Hwai-Chung Ho and Tze-Chien Sun. “Limiting Distributions of Nonlinear Vector Functions of Stationary Gaussian Processes”. In: *The Annals of Probability* 18.3 (1990), pp. 1159–1173.

[32] Jonathan R.M. Hosking. “Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series”. In: *Journal of Econometrics* 73.1 (1996), pp. 261–284.

[33] Harold E. Hurst. "Long-Term Storage Capacity of Reservoirs". In: *Transactions of the American Society of Civil Engineers* 116 (1951), pp. 770–808.

[34] Scripts originally by Jan Beran; Datasets via Brandon Whitcher. Toplevel R functions and much more by Martin Maechler. *longmemo: Statistics for Long-Memory Processes (Book Jan Beran), and Related Functionality.* R package version 1.1-2. 2020. URL: `https://CRAN.R-project.org/package=longmemo`.

[35] Olav Kallenberg. *Random Measures, Theory and Applications.* Vol. 1. Springer, 2017.

[36] Mandar Karlekar and Anubha Gupta. "Stochastic modeling of EEG rhythms with fractional Gaussian noise". In: *2014 22nd European Signal Processing Conference (EUSIPCO).* IEEE. 2014, pp. 2520–2524.

[37] Stefanos Kechagias and Vladas Pipiras. "Definitions and representations of multivariate long-range dependent time series". In: *Journal of Time Series Analysis* 36.1 (2015), pp. 1–25.

[38] Karsten Keller, Heinz Lauffer, and Mathieu Sinn. "Ordinal analysis of EEG time series". In: *Chaos and Complexity Letters* 2 (2007), pp. 247–258.

[39] Karsten Keller and Mathieu Sinn. "Ordinal analysis of time series". In: *Physica A: Statistical Mechanics and its Applications* 356.1 (2005), pp. 114–120.

[40] Gebhard Kirchgässner, Jürgen Wolters, and Uwe Hassler. *Introduction to Modern Time Series Analysis.* Springer Science & Business Media, 2012.

[41] Frédéric Lavancier, Anne Philippe, and Donatas Surgailis. "Covariance function of vector self-similar processes". In: *Statistics & Probability Letters* 79.23 (2009), pp. 2415–2421.

[42] Xian Liu. *Survival Analysis: Models and Applications.* John Wiley & Sons, 2012.

[43] Helmut Lütkepohl. *New Introduction to Multiple Time Series Analysis.* Springer Science & Business Media, 2005.

[44] Jan R. Magnus and Heinz Neudecker. "The Commutation Matrix: Some Properties and Applications". In: *The Annals of Statistics* (1979), pp. 381–394.

[45] Péter Major. "Non-central limit theorem for non-linear functionals of vector valued Gaussian stationary random fields". In: *arXiv preprint arXiv:1901.04086* (2019).

[46] Voichiţa Maxim, Levent Şendur, Jalal Fadili, John Suckling, Rebecca Gould, Rob Howard, and Ed Bullmore. "Fractional Gaussian noise, functional MRI and Alzheimer's disease". In: *Neuroimage* 25.1 (2005), pp. 141–158.

[47] Henry P. McKean. "Geometry of Differential Space". In: *The Annals of Probability* 1.2 (1973), pp. 197–206.

[48] Ines Münker. "Grenzwertsätze für langzeitabhängige stationäre Zeitreihen mit Anwendung auf Wahrscheinlichkeiten ordinaler Muster". MA thesis. University of Siegen, 2017.

[49] Ines Münker and Alexander Schnurr. "Ordinal Pattern Dependence in Contrast to Other Concepts of Dependence". In: *Proceedings of the ISI World Congress 2017* (2017).

[50] Ines Münker and Alexander Schnurr. “Space Dependent Ordinal Pattern Probabilities in Time Series”. In: *Proceedings of the ISI World Congress 2019* (2019).

[51] Roger B. Nelsen. *An Introduction to Copulas*. Springer-Verlag, 2006.

[52] Marco Oesting and Alexander Schnurr. “Ordinal patterns in clusters of subsequent extremes of regularly varying time series”. In: *Extremes* 23.4 (2020), pp. 521–545.

[53] Gaoxiang Ouyang, Chuangyin Dang, Douglas A. Richards, and Xiaoli Li. “Ordinal pattern based similarity analysis for EEG recordings”. In: *Clinical Neurophysiology* 121.5 (2010), pp. 694–703.

[54] Vladas Pipiras and Murad S. Taqqu. *Long-Range Dependence and Self-Similarity*. Vol. 45. Cambridge University Press, 2017.

[55] Giovanni Puccetti. “Measuring Linear Correlation Between Random Vectors”. In: *Available at SSRN 3116066* (2019).

[56] Friedrich Schmid, Rafael Schmidt, Thomas Blumentritt, Sandra Gaißer, and Martin Ruppert. “Copula-Based Measures of Multivariate Association”. In: *Copula Theory and Its Applications*. Springer, 2010, pp. 209–236.

[57] Alexander Schnurr. “An Ordinal Pattern Approach to Detect and to Model Leverage Effects and Dependence Structures Between Financial Time Series”. In: *Statistical Papers* 55.4 (2014), pp. 919–931.

[58] Alexander Schnurr and Herold Dehling. “Testing for Structural Breaks via Ordinal Pattern Dependence”. In: *Journal of the American Statistical Association* 112.518 (2017), pp. 706–720.

[59] Yantai Shu, Zhigang Jin, Lianfang Zhang, Lei Wang, and Oliver W.W. Yang. “Traffic prediction using FARIMA models”. In: *1999 IEEE International Conference on Communications (Cat. No. 99CH36311)*. Vol. 2. IEEE. 1999, pp. 891–895.

[60] Mathieu Sinn and Karsten Keller. “Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments”. In: *Computational Statistics & Data Analysis* 55.4 (2011), pp. 1781–1790.

[61] Robert D. Stuart. *An Introduction to Fourier Analysis*. Methuen, 1961.

[62] Murad S. Taqqu. “Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 40.3 (1977), pp. 203–238.

[63] Murad S. Taqqu. “The Rosenblatt Process”. In: *Selected Works of Murray Rosenblatt*. Ed. by Richard A. Davis, Keh-Shin Lii, and Dimitris N. Politis. Springer, 2011, pp. 29–45.

[64] Murad S. Taqqu. “Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31 (1975), pp. 287–302.

[65] Aad W. van der Vaart. *Asymptotic Statistics*. Vol. 3. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2000.

[66] Mark S. Veillette and Murad S. Taqqu. “Properties and numerical evaluation of the Rosenblatt distribution”. In: *Bernoulli* 19.3 (2013), pp. 982–1005.

[67] Robert E. Whaley. “Understanding the VIX”. In: *The Journal of Portfolio Management* 35.3 (2009), pp. 98–105.

[68] Akira M. Yaglom. *Correlation Theory of Stationary and Related Random Functions*. Springer-Verlag New York, 1987.

[69] Xuechao Yan, Dongbing Tong, Qiaoyu Chen, Wuneng Zhou, and Yuhua Xu. “Adaptive State Estimation of Stochastic Delayed Neural Networks with Fractional Brownian Motion”. In: *Neural Processing Letters* 50 (2019), pp. 2007–2020.

[70] Daoqiang Zhang, Jiashuang Huang, Biao Jie, Junqiang Du, Liyang Tu, and Mingxia Liu. “Ordinal pattern: A New Descriptor for Brain Connectivity Networks”. In: *IEEE Transactions on Medical Imaging* 37.7 (2018), pp. 1711–1722.

Notation

Basics

$n!$	n factorial
$\mathbb{N}$	natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$
$\mathbb{N}_0$	natural numbers with zero; $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$
$\mathbb{R}_+$	positive half axis $(0, \infty)$
$\Vert A\Vert_F$	Frobenius norm; $\Vert A\Vert_F = \sqrt{\sum_{p,q=1}^{d} \vert a_{pq}\vert^2}$ for $A = (a_{pq})_{p,q=1,\ldots,d} \in \mathbb{R}^{d\times d}$.
x_+, x_-	$\max\{x, 0\}$, $\max\{-x, 0\}$, respectively
$\operatorname{vec}(A)$	the vec operator transforms a matrix A into a vector by stacking the columns of the matrix underneath each other
$A \otimes A$	Kronecker product of a matrix A
I_d	d-dimensional identity matrix
K_d	commutation matrix that transforms $\operatorname{vec}(A)$ into $\operatorname{vec}\left(A^t\right)$ for a matrix $A \in \mathbb{R}^{d\times d}$
$A^{1/2}$	principal square root of a matrix A
$\operatorname{tr}(A)$	trace of a matrix A
$\Re$, $\Im$	real and imaginary parts, respectively
$\bar{z}$	complex conjugate of $z \in \mathbb{C}$
A^*	Hermitian transpose of a matrix A
$\lfloor x \rfloor$	floor integer part of x
$\hat{f}$	Fourier transform of a function f

$L^2(E, m(\mathrm{d}x))$	space of functions $f: E \to \mathbb{R}$ such that $\int_E \|f(x)\|^2 m(\mathrm{d}x) < \infty$
$L^2\left(\mathbb{R}^d\right)$	space of functions $f: \mathbb{R}^d \to \mathbb{R}$ such that $\int_{\mathbb{R}^d} \|f(x)\|^2 \mathrm{d}x < \infty$
$\\|f\\|_{L^2}$	L_2-norm; $\\|f\\|_{L^2} = \left(\int_{\mathbb{R}^d} \|f(x)\|^2 \mathrm{d}x\right)^{1/2}$
$\int'_{\mathbb{R}^k}$	integration excludes the diagonals $x_i = x_j$, $i \neq j$, $i, j = 1, \dots, k$
$\int''_{\mathbb{R}^k}$	integration excludes the diagonals $\|x_i\| = \|x_j\|$, $i \neq j$, $i, j = 1, \dots, k$

Framework of Hermite polynomials

H_k	Hermite polynomial of order k
$m(f, \Sigma_d)$	Hermite rank m with respect to the covariance matrix Σ_d
c_k	Hermite coefficients

Probability and stochastic processes

$(\Omega, \mathcal{A}, \mathbb{P})$	probability space, with Ω sample space, $\mathcal{A}$ σ-algebra, $\mathbb{P}$ probability measure
$\mathcal{B}\left(\mathbb{R}^d\right)$	Borel σ-algebra of $\mathbb{R}^d$
$\mathcal{N}(\mu, \sigma^2)$	univariate Normal distribution with mean μ and variance σ^2
$\mathcal{N}(\mu, \Sigma_X)$	multivariate Normal distribution with mean vector μ and covariance matrix Σ_X
$\overset{\mathcal{D}}{=}$	equality in distribution
$\xrightarrow{\mathcal{D}}$, $\xrightarrow{\mathbb{P}}$, $\xrightarrow{L_2}$	convergence in distribution, in probability, in the L_2 sense, respectively
$(X_k)_{k \in T}$	discrete d-variate stochastic process
$(X(t))_{t \in T}$	time-continuous d-variate stochastic process
$\Gamma_X(l)$	cross-covariance matrix of the d-variate stochastic process $(X_k)_{k \in \mathbb{Z}}$ at lag l
$\gamma^{(p,q)}(\cdot)$	cross-covariance function, $p, q = 1, \dots, d$

$r^{(p,q)}(\cdot)$	cross-correlation function, $p, q = 1, \ldots, d$
$\hat{\Gamma}_{n,X}(l)$	sample cross-covariance matrix of the d-variate stochastic process $(X_k)_{k\in\mathbb{Z}}$ at lag l
$\hat{r}_n^{(p,q)}(l)$	sample cross-covariances at lag $l = 0, 1, 2 \ldots$
$(B(t))_{t\in\mathbb{R}}$	standard (d-variate) Brownian motion
$(B_H(t))_{t\in\mathbb{R}}$	fractional Brownian motion, operator fractional Brownian motion
$(W_{k,H})_{k\in\mathbb{Z}}$	(vector-) fractional Gaussian noise
$(Z_{m,H}(t))_{t\in\mathbb{R}}$	Hermite-Rosenblatt process of order m
$(Y_j)_{j\in\mathbb{Z}}$	d-variate Gaussian process, stationary, componentwise long- or short-range dependent
$(Y_{j,h})_{j\in\mathbb{Z}}$	corresponding dh-dimensional process, $Y_{j,h} = \left(Y_j^{(1)}, \ldots, Y_{j+h-1}^{(1)}, Y_j^{(d)}, \ldots, Y_{j+h-1}^{(d)}\right)^t$
$\Sigma_{d,h}$	covariance matrix of $Y_{j,h}$; $\Sigma_d := \Sigma_{d,1}$
$I_{d,h}$	dh-dimensional identity matrix; $I_d := I_{d,1}$
$d_1, \ldots, d_d$	long-range dependence parameters
d^*	strongest long-range dependence parameter; $d^* = \max\{d_1, \ldots, d_d\}$
P^*	set of indices corresponding to d^*; $P^* = \{p \in \{1, \ldots, d : d_p = d^*\}\}$
C_1, C_2	normalizing constants for limit theorems for functionals with Hermite rank 1 or 2, respectively
c	vector of first-order Hermite coefficients; $c = \mathbb{E}\left(f\left(Y_{1,h}\right) Y_{1,h}\right)$
α, $\alpha_i^{(p)}$	factors determining the limit variance for functionals of Hermite rank 1; $\alpha = \Sigma_{d,h}^{-1} c$, $\alpha_i^{(p)} = \alpha_{i+(p-1)h}$, $p = 1, \ldots, d$, $i = 1, \ldots, h$

C	matrix of second-order Hermite coefficients; $C = \mathbb{E}\left(Y_{1,h}\left(f\left(Y_{1,h}\right) - \mathbb{E}\left(f\left(Y_{1,h}\right)\right)\right) Y_{1,h}^t\right)$
α, $\tilde{\alpha}^{(p,q)}$	factors determining the limit variance for functionals of Hermite rank 1; $\alpha = \Sigma_{d,h}^{-1} C \Sigma_{d,h}^{-1}$, $\tilde{\alpha}^{(p,q)} = \sum_{i,k=1}^{h} \alpha_{i+(p-q)h,k+(q-1)h}$, $p, q = 1, \ldots, d$

Ordinal pattern analysis

S_h	set of $(h+1)$-dimensional permutations, $h \in \mathbb{N}$
$\bar{\pi}$	reversion group of $\pi \in S_h$, $\bar{\pi} = \{\pi, \mathcal{S}(\pi), \mathcal{T}(\pi), \mathcal{T} \circ \mathcal{S}(\pi)\}$
S_h^*	subset of S_h, whose elements π are choosen from S_h such that their spatial reverse $\mathcal{S}(\pi)$ must not be contained in S_h^*
$S_{[2],h}$, $S_{[4],h}$	groups of generating pattern of reversion groups with two or four elements, respectively

$(X_j)_{j \in \mathbb{Z}}$	d-dimensional Gaussian process, in which we study the pattern, $d = 1$ in Section 4.2 and $d = 2$ in Section 4.3, in general **not** stationary in Chapter 4, **only** in Remark 4.22 and Section 4.4.3
$(X_{j,h+1})_{j \in \mathbb{Z}}$	corresponding $d(h+1)$-dimensional process
$(Y_j)_{j \in \mathbb{Z}}$	increment process of $(X_j)_{j \in \mathbb{Z}}$, **always** stationary, componentwise long- or short-range dependent
$(Y_{j,h})_{j \in \mathbb{Z}}$	increment process of $(X_{j,h+1})_{j \in \mathbb{Z}}$

$q_{Y^{(p)}}(\pi)$	ordinal pattern probability of $\pi \in S_h$ in $\left(Y_j^{(p)}\right)_{j \in \mathbb{Z}}$, $p = 1, 2$; $q(\pi) := q_{Y^{(1)}}(\pi)$ in Section 4.2
p	probability of coincident ordinal patterns in the two time series $\left(X_j^{(1)}\right)_{j \in \mathbb{Z}}$ and $\left(X_j^{(2)}\right)_{j \in \mathbb{Z}}$
q	probability of coincident ordinal patterns in the hypothetical case of independence

OPD	value of ordinal pattern dependence
$\hat{q}_{n,Y^{(p)}}(\pi)$	estimator of $q_{Y^{(p)}}(\pi)$; $\hat{q}_n(\pi) := \hat{q}_{Y^{(1)},n}(\pi)$ in Section 4.2
$\hat{q}_{n,Y^{(p)},\mathrm{imp}}(\pi)$	improved estimator of $q_{Y^{(p)}}(\pi)$ based on Rao-Blackwellization; $\hat{q}_{n,\mathrm{imp}}(\pi) := \hat{q}_{Y^{(1)},n,\mathrm{imp}}(\pi)$ in Section 4.2
$\hat{p}_n$	estimator for p
$\hat{q}_n$	estimator for q based on $\hat{q}_{n,Y^{(p)}}(\pi)$, $p = 1, 2$
$\hat{q}_{n,\mathrm{alt}}$	alternative estimator for q, based on $\hat{q}_{n,Y^{(p)}}(\pi)$, $p = 1, 2$
$\hat{q}_{n,\mathrm{imp}}$	improved estimator for q, based on $\hat{q}_{n,Y^{(p)},\mathrm{imp}}(\pi)$, $p = 1, 2$
$\widehat{OPD}_n$	estimator of OPD based on $\hat{p}_n$ and $\hat{q}_n$
$\widehat{OPD}_{n,\mathrm{alt}}$	alternative estimator of OPD based on $\hat{p}_n$ and $\hat{q}_{n,\mathrm{alt}}$
$\widehat{OPD}_{n,\mathrm{imp}}$	improved estimator of OPD based on $\hat{p}_n$ and $\hat{q}_{n,\mathrm{imp}}$
$p(\pi)$	probability of a certain coincident ordinal pattern $\pi \in S_h$ in both time series $\left(X_j^{(1)}\right)_{j\in\mathbb{Z}}$ and $\left(X_j^{(2)}\right)_{j\in\mathbb{Z}}$
$\hat{p}_n(\pi)$	estimator for $p(\pi)$
$\hat{p}_{n,\mathrm{imp}}(\pi)$	improved estimator for $p(\pi)$
$\hat{H}_n^{(p,q)}$	Hurst parameter estimator for vector fractional Gaussian noise, $p, q = 1, 2$
p^a	largest possible probability of coincident ordinal patterns by additionally allowing for a time-shift a
$\hat{p}_n^a$	estimator for p^a
$\hat{p}_n^{block}$	blockwise estimation of p
$AWOPD$	value of average-weighted ordinal pattern dependence
$\hat{p}_n^{\mathrm{awopd}}$	estimator for "almost" coincident patterns in the sense of AWOPD

$\hat{q}_n^{\text{awopd}}$	estimator for "almost" coincident patterns in the hypothetical case of independence
C^{π}	$C^{\pi} = \left(c_{i,j}^{\pi}\right)_{i,j=1,\dots,h}$ matrix of second order Hermite coefficients $\hat{q}_n(\pi)$
$C^{\pi,2}$	$C^{\pi,2} = \left(c_{i,j}^{\pi,2}\right)_{i,j=1,\dots,h}$ matrix of second order Hermite coefficients of $\hat{p}_n(\pi)$
$\alpha^{\pi,2}, \tilde{\alpha}_{\pi,2}^{(p,q)}$	factors in the limit variance; $\alpha^{\pi,2} = \Sigma_{2,h}^{-1} C^{\pi,2} \Sigma_{2,h}$, $\tilde{\alpha}_{\pi,2}^{(p,q)} = \sum_{i,k=1}^{h} \alpha_{(p-1)h+i,(q-1)h+k}^{\pi,2}$, $p,q = 1,2$
$c^{\pi,2}$	vector of first order Hermite coefficients of $\hat{p}_n(\pi)$
$\tilde{\alpha}_{\pi,2}^{(p)}$, $\alpha^{\pi,2,\star}$	factors in the limit variance; $\alpha^{\pi,2,\star} = \Sigma_{2,h}^{-1} c^{\pi,2}$; $\tilde{\alpha}_{\pi,2}^{(p)} = \sum_{i=1}^{h} \alpha_{(p-1)h+i}^{\pi,2,\star}$, $p = 1,2$

Ordinal pattern dependence in contrast to other measures of dependence

$X^{(1,h+1)}$, $X^{(2,h+1)}$	$X^{(p,h+1)} = \left(X_0^{(p)}, \dots, X_h^{(p)}\right)^t$, $p = 1,2$, $\left(X^{(1,h+1)}, X^{(2,h+1)}\right)^t = X_{0,h+1}$
$\Sigma_{1,h+1}$, $\tilde{\Sigma}_{1,h+1}$	corresponding covariance matrices, if $X_{0,h+1}$ is Gaussian
$\Sigma_{X^{(1,h+1)},X^{(2,h+1)}}$	cross-covariance matrix of Gaussian $X^{(1,h+1)}$, $X^{(2,h+1)}$
OPD_h	value of ordinal pattern dependence with emphasis on h increments under consideration
ρ_h	multivariate Pearson's correlation coefficient
τ_h	multivariate Kendall's τ
$\hat{\tau}_n^{(h)}$	estimator of τ_h
$\hat{\tau}_{n,bw}^{(h)}$	blockwise estimator of τ_h

Miscellaneous

A^t, x^t	transpose of matrix $A \in \mathbb{R}^{d \times d}$ or vector $x \in \mathbb{R}^d$

$c \cdot A$	componentwise multiplication of a scalar $c \in \mathbb{R}$ and a matrix $A \in \mathbb{R}^{d\times d}$
$c \cdot x$	componentwise multiplication of a scalar $c \in \mathbb{R}$ and a vector $x \in \mathbb{R}^d$
c^D	$c^D := \operatorname{diag}\left(c^{d_1}, \dots, c^{d_d}\right)$ for $c \in \mathbb{R}$ and a diagonal matrix $D = \operatorname{diag}(d_1, \dots, d_d) \in \mathbb{R}^d$
$\#M$	cardinality of a set M
$X \sim \mathcal{N}(0,1)$	X is standard normally distributed
$\simeq$	asymptotic equality; $a_n \simeq b_n \Leftrightarrow \lim_{n\to\infty} \frac{a_n}{b_n} = 1$
$o_{\mathbb{P}}(n)$	small stochastic o; convergence to zero in probability is faster than n^{-1}; $R_n = o_{\mathbb{P}}(n)$ has the meaning that $\lim_{n\to\infty} \frac{R_n}{n} \xrightarrow{\mathbb{P}} 0$
$\mathbf{1}_A$	indicator function of set A
δ_0	Dirac measure in 0
$\delta_{x,y}$	Kronecker delta

List of Figures

List of Tables