Wavelet-based Estimation for Gaussian and Non-Gaussian Mixed Fractional Processes

AN ABSTRACT
SUBMITTED ON THE FIFTH DAY OF AUGUST, 2017
TO THE DEPARTMENT OF MATHEMATICS
OF THE SCHOOL OF SCIENCE AND ENGINEERING OF
TULANE UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
BY

Hui Li

HUI LI

APPROVED: Gustavo Didier

GUSTAVO DIDIER, PH.D.
CHAIRMAN

Ricardo Cortez

RICARDO CORTEZ, PH.D.

LISA J. FAUCI, PH.D.

MICHELLE LACEY, PH.D.

KUN ZHAO, PH.D.

Abstract

In this thesis, we tackle the statistical problem of demixing a multivariate stochastic process made up of independent, fractional process entries. We consider both Gaussian and non-Gaussian frameworks. The observable, mixed process is then a multivariate fractional stochastic process. In particular, when the components of the unmixed process are self-similar, the mixed process is operator self-similar. Multivariate mixed fractional processes are parameterized by a vector of Hurst parameters and a mixing matrix. We propose a 2-step wavelet-based estimation method to produce estimators of both the demixing matrix and the Hurst parameters. In the first step, an estimator of the demixing matrix is obtained by applying a classical joint diagonalization algorithm to two wavelet variance matrices of the mixed process. In the second step, a univariate-like wavelet regression method is applied to each entry of the demixed process to provide estimators of each individual Hurst parameter. The limiting distribution of the estimators is established for both Gaussian and non-Gaussian (Rosenblatt-like) instances. Monte Carlo experiments show that the finite sample estimation performance is very satisfactory. As an application, we model bivariate series of annual tree ring measurements from bristlecone pine trees in White Mountains, California.

Wavelet-based Estimation for Gaussian and Non-Gaussian Mixed Fractional Processes

A DISSERTATION
SUBMITTED ON THE FIFTH DAY OF AUGUST, 2017
TO THE DEPARTMENT OF MATHEMATICS
OF THE SCHOOL OF SCIENCE AND ENGINEERING OF
TULANE UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
BY

Hui Li

APPROVED: Gustavo Didier, Ph.D.
CHAIRMAN

Ricardo Cortez, Ph.D.

Lisa J. Fauci, Ph.D.

Michelle Lacey, Ph.D.

Kun Zhao, Ph.D.

Acknowledgement

I would like to dedicate this thesis to all those who have offered me tremendous assistance and encouragement during these 5 years in the Math department of Tulane University, without whom, I cannot finish my Ph.D study.

Firstly, My deepest gratitude goes to my advisor, Professor Gustavo Didier, for his constant guidance, professional advice and selfless support. He lent me many useful books and provided me summer support for the past two years. He has spent much time reading through each draft and offered me with inspiring advice, and walked me through all the stages of the writing of this thesis. I have learned a lot from him, not only the professional knowledge, but also meticulous attitude towards research.

Secondly, I would like to express my sincere gratitude to other professors in our department: Professor Lisa Fauci, Professor Alexander Kurganov, Professor Michelle Lacey, Professor Victor Hugo Moll and Professor Morris Kalka, for their wonderful courses. I also thank my other committee members Professor Kun Zhao and Professor Ricardo Cortez.

I thank my Ph.D peers in Tulane: Li Guan, Shanshan Jiang, Xiao Guan, Zhuolin Qu, Yuanzhen Cheng, etc, for the great time we have spent together.

Finally, I especially thank my beloved parents and husband Kui Zhang, for their loving considerations and great confidence in me all through these years.

List of Tables

List of Figures

Contents

Chapter 1

Introduction

Fractional processes are widely found in science, technology and engineering systems. Numerous signals and systems from a wide range of applications have been analyzed by means of fractional models. Examples include natural systems (hydrodynamic turbulence [1], geophysics [2], heart rate variability [3], infraslow (i.e., below 1Hz) brain activity [4], or man-made systems (e.g., Internet traffic [5]).

This thesis is dedicated to a subclass of multivariate processes, i.e., those of the form

$$\{Y(t)\}_{t\in\mathbb{R}} = \{PX(t)\}_{t\in\mathbb{R}}, \tag{1.1}$$

where P is a non-singular matrix and $\{X(t)\}_{t\in\mathbb{R}}$ is a vector of independent fractional processes, either Gaussian or non-Gaussian. It was recently shown ([6]) that such processes naturally emerge as a consequence of the aggregation of measurements from relevant classes of fractional SDE models (see Section 2.4). In fact, many real data sets are aggregate. For example, sales of products, industrial production, tree ring widths (see Section 4.2), river flows, and rainfall are obtained through aggregation over a certain time interval. Multivariate fractional processes of the form (1.1) are also a relevant subcase of the so-named operator self-similar processes, which have been attracting much attention recently ([7–9]). In addition, they also provide an extension to the framework of fractional processes of the so-named mixed processes from the blind source separation literature in signal processing (e.g., [10, 11]).

This thesis is organized as follows. In Chapter 2, we introduce Gaussian multivariate processes with integer integration orders. In Chapter 3, we construct a two-step wavelet-based estimation method for both P and the Hurst parameters. In addition, assuming Y in (1.1) is a fractional Gaussian process observed in continuous time, we prove the asymptotic normality of the estimators. The performance of the method on finite samples is studied by means of Monte Carlo experiments. We also establish the joint asymptotic distribution of the eigenvalues and eigenvectors of wavelet variance matrices, which is of independent interest. In Chapter 4, we make the more realistic assumption that Y in (1.1) is observed in discrete time instead and prove that the asymptotic properties of the proposed estimators do not change. In addition, we apply the proposed method on a bivariate tree-ring data set, and show that the latter can be modeled by the mixed form (1.1). Chapter 5 illustrates that the proposed methodology also works for non-Gaussian case (mixed Rosenblatt processes). However, as expected, the limiting distribution of the estimator is no longer Gaussian, and hence calls for particular mathematical techniques. All proofs can be found in the appendix. Moreover, also in Appendix, we analyze the bivariate COS2-NO diffusion data sets provided by the David B. Hill Lab (UNC-Chapel Hill), and find that we cannot reject the null hypothesis that the Hurst exponents are equal.

Chapter 2

Gaussian Fractional Mixed Multivariate Processes

Consider the models of the form

$$\{Y(t)\}_{t\in\mathbb{R}} = \{PX(t)\}_{t\in\mathbb{R}}, \tag{2.1}$$

where P is a $n \times n$ non-singular matrix, and $X(t)$ is a vector of independent fractional processes with so-called Hurst parameters h_k, $k = 1, \cdots, n$. The goal of this thesis is to develop an efficient estimation methodology both for P and $h_1, \cdots, h_n$. In this chapter, we lay out the mathematical framework underlying (2.1) in Gaussian case, i.e., stationary increments, Gaussian factional stochastic processes.

2.1 Processes with N-th order stationary increments

A stochastic process is (strictly) stationary when its finite dimensional distributions do not change when shifted in time. Stationary processes can be used to describe the time variation of some trait of a steady state phenomenon, for which no choice of time has any advantage over any other choice. We call wide sense stationary a finite variance process Y whose covariance function $\mathrm{Cov}(Y(t), Y(s))$ only depends on $t - s$.

Let $X(t)$ be a real-valued random process on $\mathbb{R}$. For and $\tau > 0$, define the shift operator B_τ by

$$B_\tau X(t) = X(t - \tau)$$

and the difference operator $\triangle_\tau$ by

$$\triangle_\tau = I - B_\tau.$$

The N-th order difference of $X(t)$ with lag τ is defined to be

$$\triangle_\tau^N X(t) = \sum_{k=0}^{N} (-1)^k \binom{N}{k} X(t - k\tau).$$

Definition 2.1.1. A finite variance process X is said to have N-th order stationary increments if $\triangle_\tau^N X$ is wide sense stationary for all $\tau > 0$.

Yaglom and Pinsker established a harmonizable representation for this class of processes in the 1950s [12–14]. A process with N-th order stationary increments can be expressed as

$$X(t) \stackrel{d}{=} \int_{\mathbb{R}} \left(\frac{e^{ixt} - 1 - ixt - \cdots - \frac{(ixt)^{N-1}}{(N-1)!}}{(ix)^N} \right) \widetilde{Y}(dx) + \Upsilon_0 + \Upsilon_1 t + \cdots + \Upsilon_N t^N, \quad t \in \mathbb{R}, \tag{2.2}$$

where Υ_0, Υ_1, $\cdots$, Υ_N are some random variables and $\widetilde{Y}(dx)$ is a $\mathbb{C}$-valued orthogonal-increment random measure.

2.2 Self-similar processes

The paradigm of scale invariance is applied in the analysis of dynamic signals or systems where no characteristic scale is present. Under scale invariance, a continuum of time scales contribute to the observed dynamics, and the analyst's focus is on identifying mechanisms that relate the scales, often in the form of the so-named scaling exponents [15–17]. An important form of scale invariance is self-similarity.

Definition 2.2.1. A real-valued stochastic process $X = \{X(t)\}_{t\in\mathbb{R}}$ is self-similar with Hurst parameter $h > 0$ if, for any $a > 0$,

$$\{X(at)\}_{t\in\mathbb{R}} \stackrel{d}{=} \{a^h X(t)\}_{t\in\mathbb{R}},$$

where $\stackrel{d}{=}$ denotes the equality of the finite dimensional distributions. In other words, $(X(at_1), \cdots, X(at_n))$ has the same joint distribution as $a^h(X(t_1), \cdots, X(t_n))$, for any $a > 0$, $(t_1, \cdots, t_n) \in \mathbb{R}^n$, $n \in \mathbb{N}$.

Example 2.2.1. The celebrated fractional Brownian motion (fBm) is the only Gaussian, self-similar process with stationary increments [18]. The covariance function of a fBm $B_h(t)$ with Hurst parameter h is given by

$$\mathrm{Cov}(B_h(t), B_h(s)) = \frac{\sigma^2}{2}(|t|^{2h} + |s|^{2h} - |t-s|^{2h}),$$

where $\sigma^2 = \mathrm{Var}(B_h(1))$. It is called standard if $\sigma^2 = 1$.

Proposition 2.2.1. *Let $X = \{X(t)\}_{t\in\mathbb{R}}$ be a zero mean, Gaussian, self-similar process with Hurst parameter*

$$h \in (N-1, N). \tag{2.3}$$

Suppose, in addition, that X has N-th order stationary increments. Then, $\{X(t)\}_{t\in\mathbb{R}}$

admits the harmonizable representation

$$\{X(t)\}_{t\in\mathbb{R}} \overset{d}{=} C\int_{\mathbb{R}} \frac{e^{itx} - \sum_{k=0}^{N-1}\frac{1}{k!}(itx)^k}{(ix)^N}|x|^{-(h-(N-1/2))}\widetilde{B}(dx), \tag{2.4}$$

for some constant $C \in \mathbb{R}$, where $\widetilde{B}(dx)$ is a $\mathbb{C}$-valued Gaussian random measure satisfying $\widetilde{B}(-dx) = \overline{\widetilde{B}(dx)}$.

Proof. We first show that in (2.2),

$$\Upsilon_0 + \Upsilon_1 t + \cdots + \Upsilon_N t^N \overset{\text{a.s}}{=} 0, \quad t \in \mathbb{R}. \tag{2.5}$$

Let $F(dx)$ be the control measure associated with $\widetilde{Y}(dx)$, i.e., $F(x) = \mathbb{E}\widetilde{Y}(dx)\overline{\widetilde{Y}(dx)}$. Then, by self-similarity,

$$c^h X(t) \overset{d}{=} X(ct) \overset{d}{=} \int_{\mathbb{R}} \frac{e^{ixt} - \sum_{k=0}^{N-1}\frac{1}{k!}(itx)^k}{(ix)^N} c^N Y(c^{-1}dx) + \Upsilon_0 + \Upsilon_1 t + \cdots + \Upsilon_N t^N,$$

for $t \in \mathbb{R}$. However, the measure F is uniquely determined by $X(t)$, therefore,

$$c^{2h}F(dx) = c^{2N}F(c^{-1}dx), \tag{2.6}$$

$$c^h(\Upsilon_0 + \Upsilon_1 t + \cdots + \Upsilon_N t^N) \overset{d}{=} \Upsilon_0 + \Upsilon_1 ct + \cdots + \Upsilon_N c^N t^N. \tag{2.7}$$

Take variance on both sides of (2.7) and let $t = 1/c$,

$$\infty > \mathrm{Var}(\Upsilon_0 + \cdots + \Upsilon_N) = c^{2h}\mathrm{Var}(\Upsilon_0 + \Upsilon_1/c + \cdots + \Upsilon_N/c^N). \tag{2.8}$$

Since $N-1 < h < N$, the leading term in the righthand side of (2.8) is $c^{2h-2N}\mathrm{Var}(\Upsilon_N)$ as $c \to 0$. This implies $\mathrm{Var}(\Upsilon_N) = 0$.

On the otherhand, let $c \to \infty$, the leading term in the righthand side of (2.8) is $c^{2h}\mathrm{Var}(\Upsilon_0)$, this implies $\mathrm{Var}(\Upsilon_0) = 0$. By using $\mathrm{Var}(\Upsilon_0) = 0$, the leading term in the righthand side of (2.8) is $c^{2h-2}\mathrm{Var}(\Upsilon_1)$, this implies $\mathrm{Var}(\Upsilon_1) = 0$. By applying the same procedure, we conclude $\mathrm{Var}(\Upsilon_k) = 0$, $k = 0, 1, \cdots, N$. Thus, Υ_k's are a.s constants, $k = 0, 1, \cdots, N$, and by (2.7),

$$\Upsilon_k \overset{\text{a.s}}{=} 0, \quad k = 0, 1, \cdots, N.$$

This proves (2.5). Moreover, by using (2.6), apply the same argument as in the proof of Theorem 3.1 in [7], it can be shown that $F(dx)$ is absolute continuous. Thus, there exists a density function $f(x) = F(dx)/dx$, and X can be represented as

$$X(t) = \int_{\mathbb{R}} \frac{e^{itx} - \sum_{k=0}^{N-1} \frac{1}{k!}(itx)^k}{(ix)^N} f(x)\widetilde{B}(dx).$$

By self-similarity,

$$X(ct) \overset{d}{=} c^h \int_{\mathbb{R}} \frac{e^{itx} - \sum_{k=0}^{l-1} \frac{1}{k!}(itx)^k}{(ix)^N} f(x)\widetilde{B}(dx),$$

for $c > 0$. On the other hand, through a change of variable $x = c^{-1}\xi$,

$$X(ct) \overset{d}{=} \int_{\mathbb{R}} \frac{e^{it\xi} - \sum_{k=0}^{l-1} \frac{1}{k!}(it\xi)^k}{(i\xi)^N} c^{N-1/2} f(\xi/c)\widetilde{B}(dx),$$

then $f(x) = Cx^{-(h-N+1/2)}$, for some constant C. Let g_t be the integrand of (2.4), then $g_t(x)$ behaves like $x^{N-1/2-h}$ at the origin and like $x^{N-3/2-h}$ at infinity, we obtain that $\int_{\mathbb{R}} |g_t(x)|^2 dx < \infty$ by using (2.3).

□

Example 2.2.2. The spectral representation of a standard fBm $B_h(t)$ with Hurst parameter $h \in (0, 1)$ is given by

$$B_h(t) \overset{d}{=} \frac{1}{C(h)} \int_{\mathbb{R}} \frac{e^{itx} - 1}{ix} |x|^{-(h-1/2)} \widetilde{B}(dx),$$

where

$$C^2(h) = \frac{\pi}{h\Gamma(2h)\sin h\pi}.$$

2.3 Gaussian fractional mixed multivariate processes

In (2.1), we consider the case where the source signal X satisfies

$$\{X(t)\}_{t\in\mathbb{R}} = \{(X_{h_1}(t), \cdots, X_{h_n}(t))^T\}_{t\in\mathbb{R}}, \tag{2.9}$$

where, for $i = 1, \cdots, n$, the entries $X_{h_i}(t)$ are independent and have the form

$$X_{h_i}(t) = \int_{\mathbb{R}} \frac{e^{itx} - \sum_{k=0}^{N_i-1} \frac{1}{k!}(itx)^k}{(ix)^{N_i}} |x|^{-(h_i-(N_i-1/2))} g_i(x)\widetilde{B}(dx), \tag{2.10}$$

with $N_i - 1 < h_i < N_i$, i.e., X_{h_i} has N_i-th order stationary increments. The so-named high frequency functions $g_i(x)$ are bounded with $|g_i(0)| > 0$, $i = 1, \cdots, n$. Then, each individual $X_{h_i}(t)$ is a Gaussian fractional process, but is not necessarily self-similar.

In (2.10), if $g_i(x)$ is a constant, $i = 1 \cdots, n$, then the source signal X satisfies the so-named entry-wise scaling property

$$\{X(ct)\}_{t\in\mathbb{R}} \overset{d}{=} \{c^{\operatorname{diag}(h_1,\cdots,h_n)} X(t)\}_{t\in\mathbb{R}}, \quad c > 0. \tag{2.11}$$

However, for most mixing matrices P, the observed signal $Y = PX$ doesn't generally satisfy (2.11). Instead, it satisfies the operator self-similarity property, defined next.

Definition 2.3.1. An $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t\in\mathbb{R}}$ is operator self-similar when its law scales according to a matrix (Hurst) exponent H, i.e.,

$$\{X(ct)\}_{t\in\mathbb{R}} \overset{d}{=} \{c^H X(t)\}_{t\in\mathbb{R}}, \quad c > 0,$$

where $c^H = \sum_{k=0}^{\infty} \log^k(c)H^k/k!$.

Thus, if $g_i(x)$ is a constant, $i = 1\cdots,n$, the signal $Y = PX$ of the form (2.1) is operator self-similar with Hurst matrix

$$H = P\text{diag}(h_1, \cdots, h_n)P^{-1}.$$

Example 2.3.1. Operator fractional Brownian motion is a proper Gaussian, operator self-similar, stationary increment stochastic process. Under mild conditions, any operator fractional Brownian motion B_H admits a harmonizable representation

$$\{B_H(t)\}_{t\in\mathbb{R}} \stackrel{d}{=} \left\{ \int_{\mathbb{R}} \frac{e^{itx}-1}{ix}(x_+^{-(H-1/2I)}A + x_-^{-(H-1/2I)}\overline{A})\widetilde{B}(dx)\right\}_{t\in\mathbb{R}}$$

for some complex-valued matrix A, and $x_\pm = \max\{\pm x, 0\}$ ([7]).

2.4 Mixed processes and aggregation

Consider the following multi-dimensional fractional Brownian motion driven Langevin equation for $t > 0$

$$dY(t) = \Phi Y(t) + \Sigma dB^{\mathbf{h}}(t). \tag{2.12}$$

In (2.12), where $B^{\mathbf{h}}(t) = (B_{h_1}(t), \cdots, B_{h_n}(t))^T$, is a vector of independent fractional Brownian motions $\{B_{h_k}(t)\}$ with Hurst parameters h_k, $k = 1, \cdots, n$. The solution of (2.12) can be a.s. written as

$$Y(t) = e^{\Phi t}Y(0) + \int_0^t e^{\Phi(t-u)}\Sigma dB^{\mathbf{h}}(u). \tag{2.13}$$

Suppose all the eigenvalues of Φ have strictly negative real parts. Consider the case that the continuous-time process $\{Y(t)\}$ defined by (2.13) is digitalized by aggregation over interval $\triangle$, i.e.,

$$Y_i^{\triangle} = \int_{(i-1)\triangle}^{i\triangle} Y(u)du, \quad i \in \mathbb{Z}^+.$$

Then, it can be shown that, as $\triangle \to \infty$,

$$-\begin{pmatrix} \triangle^{-h_1} & & \\ & \ddots & \\ & & \triangle^{-h_n} \end{pmatrix} \Sigma^{-1}\Phi Y_i^{\triangle} \xrightarrow{d} \begin{pmatrix} B_i^{h_1} - B_{i-1}^{h_1} \\ \vdots \\ B_i^{h_n} - B_{i-1}^{h_n} \end{pmatrix}.$$

This is proceed in [6] for bivariate case, but the argument extends to any dimension n. Therefore, for large $\triangle$, $Y_i^{\triangle}$ can be approximated by

$$\widetilde{Y}_i = PX_i, \quad i \in \mathbb{Z}. \tag{2.14}$$

In (2.14), X_i is a vector of n independent fractional Gaussian noises with Hurst parameters h_k, $k = 1, \cdots, n$, and $P = -\Phi^{-1}\Sigma \begin{pmatrix} \triangle^{h_1} & & \\ & \ddots & \\ & & \triangle^{h_n} \end{pmatrix}$. Note that (2.14) is a particular case of (2.1), when the latter is measured in discrete time.

Chapter 3

Two-Step Wavelet-Based Estimation for Mixed Fractional Gaussian Processes

In this chapter, we consider the problem of demixing the Gaussian fractional process

$$Y(t) = PX(t), \tag{3.1}$$

where $X(t)$ is defined in (2.9). In order to estimate the demixing matrix P^{-1} and recover the original source signal X by estimating the Hurst parameters $h_1, \cdots, h_n$, we propose a two-step wavelet-based procedure, namely,

(1) demix the process;

(2) estimate $h_1, \cdots, h_n$ by wavelet log-regression.

In step (1), the demixing matrix estimator $\widehat{P^{-1}}$ is obtained by applying a classical joint diagonalization algorithm to two wavelet variances of the mixed process Y (Section 3.5), and the demixed process is then given by $\widehat{X} = \widehat{P^{-1}}Y$. Once the demixing is done, univariate-like estimation methodology can be applied (e.g., ([17, 19–21]). In step (2), we use the wavelet-regression method (Section 3.6), which works very well (see simulation results in Section 3.7.1).

This chapter is organized as follows. Section 3.1 contains the notation, assumptions and theoretical background of the chapter. In Section 3.2, we establish some properties of wavelet transform for the mixed process Y. Section 3.3 is dedicated to asymptotic properties of the wavelet transform at fixed scales. In Section 3.4, the asymptotic normality of the eigenstructure of the sample wavelet variance matrix is developed. Sections 3.5 and 3.6 contain the main mathematical results of the chapter. In Section 3.5, the demixing method for the matrix P^{-1} is laid out in full detail and its asymptotic properties are established. The post-demixing Hurst parameter estimation method is put forward in Section 3.6, and the asymptotic normality of the Hurst estimator is established. The performance of the estimation for both demixing matrix and the hurst parameters is further investigated by means of Monte Carlo experiments in Section 3.7. All the proofs can be found in Appendix A.3.

3.1 Preliminaries and assumptions

All through this chapter, the dimension of the mixed process Y is denoted by $n \geq 2$.

We shall use throughout the paper the following matrix notation. $M(m, n, \mathbb{R})$ is the vector space of all $m \times n$ real-valued matrices, whereas $M(n, \mathbb{R})$ is a shorthand for $M(n, n, \mathbb{R})$. $GL(n, \mathbb{R})$ is the general linear group (invertible matrices), $O(n)$ is the orthogonal group of matrices O such that $OO^* = I = O^*O$, where * represents the matrix adjoint and ^{T} is reserved for vector transpose. $\mathcal{S}(n, \mathbb{R})$ and $\mathcal{S}_+(n, \mathbb{R})$ are, respectively, the space of symmetric and cone of symmetric positive semidefinite matrices. The symbol $\mathbf{0}$ represents a vector or matrix of zeroes. A block-diagonal matrix with main diagonal blocks $\mathcal{P}_1, \ldots, \mathcal{P}_n$ or m times repeated diagonal block $\mathcal{P}$ is represented by

$$\mathrm{diag}(\mathcal{P}_1, \ldots, \mathcal{P}_n), \quad \mathrm{diag}_m(\mathcal{P}), \tag{3.2}$$

respectively. The symbol $\|\cdot\|$ represents a generic matrix or vector norm. The l_p entry-wise norm of the matrix A is denoted by

$$\|A\|_{l_p} = \|(a_{i_1,i_2})_{\substack{i_1=1,\dots,m \\ i_2=1,\dots,n}}\|_{l_p} = \Big(\sum_{i_1=1}^{m}\sum_{i_2=1}^{n} |a_{i_1,i_2}|^p\Big)^{1/p}. \tag{3.3}$$

For $S = (s_{i_1,i_2})_{i_1,i_2=1,\dots,n} \in M(n,\mathbb{R})$, let

$$\text{vec}_{\mathcal{S}}(S) = (s_{11}, s_{21}, \dots, s_{n1}, s_{22}, s_{32}, \dots, s_{n2}, \dots, s_{nn}),$$

$$\text{vec}_{\mathcal{D}}(S) = (s_{11}, s_{22}, \dots, s_{nn}), \quad \text{vec}(S) = (s_{11}, \dots, s_{n1}, s_{12}, \dots, s_{n2}, \dots, s_{nn}). \tag{3.4}$$

In other words, the operator $\text{vec}_{\mathcal{S}}(\cdot)$ vectorizes the lower triangular entries of S, $\text{vec}_{\mathcal{D}}(\cdot)$ vectorizes the diagonal entries of S, and $\text{vec}(\cdot)$ vectorizes all the entries of S. Note that the expressions in (3.4) are defined as row vectors; this will make the notation in several statements simpler.

Throughout this chapter, we will make the following assumptions on Y.

Assumption (A1): the observed signal has the mixed form $Y = PX$, where P is nonsingular, X is defined in (2.9) and satisfy

$$0 < h_1 < h_2 < \ldots < h_n, \quad N_i - 1 < h_i < N_i, \quad i = 1 \cdots, n. \tag{3.5}$$

Assumption (A2):

$$P \in GL(n,\mathbb{R}), \quad \|\mathbf{p}_{\cdot l}\| = 1, \quad p_{ll} \geq 0, \quad l = 1, \dots, n. \tag{3.6}$$

Assumption (A3): the functions $g_i(x) \in C^2(\mathbb{R})$ in (2.10) are bounded and satisfy

$$||g_i(x)|^2 - |g_i(0)|^2| < L|x|^{\beta}, \quad L > 0, \quad i = 1, \cdots, n, \tag{3.7}$$

for any $x \in (-\pi, \pi)$. In (3.7), β satisfies

$$1/2 < \beta < 2h_1 + 2\alpha. \tag{3.8}$$

for some

$$\alpha > 1. \tag{3.9}$$

Remark 3.1.1. In (3.5), one incurs no loss of generality by assuming that the Hurst eigenvalues (individual Hurst exponents) are disposed in ascending order. This fact can be easily illustrated in dimension $n = 2$. Suppose that the mixed signal has the form $Y(t) = P(X_{h_2}(t), X_{h_1}(t))^T$, where $X_{h_i}(t)$, $i = 1, 2$, are independent fractional processes defined in (2.10) with parameters $h_1 < h_2$. Let

$$R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then, $Y(t) = PR(X_{h_1}(t), X_{h_2}(t))^T$, whence PR can be treated as the mixing matrix with unit vector columns.

Remark 3.1.2. Assumption (A3) is typical in the semi-parametric estimation setting (e.g., [20, 22, 23]). The larger the value of β, the smoother the function at the origin.

When establishing bounds, C denotes a positive constant whose value can change from one inequality to another. All through the chapter, we will make the following assumptions on the underlying wavelet basis, which are then omitted from the statements.

ASSUMPTION $(W1)$: $\psi \in L^1(\mathbb{R})$ is a wavelet function, namely,

$$\int_{\mathbb{R}} \psi^2(t)dt = 1, \quad \int_{\mathbb{R}} t^q \psi(t)dt = 0, \quad q = 0, 1, \ldots, N_\psi - 1, \quad N_\psi \geq N_n + 1. \tag{3.10}$$

Assumption $(W2)$:

$$\text{supp}(\psi) \text{ is a compact interval.} \tag{3.11}$$

Assumption $(W3)$: for α as in (3.9),

$$\sup_{x\in\mathbb{R}} |\widehat{\psi}(x)|(1+|x|)^{\alpha} < \infty. \tag{3.12}$$

Under (3.10), (3.11) and (3.12), ψ is continuous, $\widehat{\psi}(x)$ is everywhere differentiable and its first $N_\psi - 1$ derivatives are zero at $x = 0$ (see [24], Theorem 6.1 and the proof of Theorem 7.4). The condition $(W1)$ is equivalent to asserting that the first $N_\psi - 1$ derivatives of $\widehat{\psi}$ vanish at the origin. This implies, using a Taylor expansion, that

$$|\widehat{\psi}^{(l)}(x)| = O(|x|^{N_\psi - l}), \quad l = 0, 1 \cdots, N_\psi, \quad x \to 0. \tag{3.13}$$

Example 3.1.1. If ψ is a Daubechies wavelet with N_ψ vanishing moments, $\text{supp}(\psi) = [0, 2N_\psi - 1]$ (see [24], Proposition 7.4).

Remark 3.1.3. Assumption (A1) requires using a number of vanishing moments N_ψ larger than the unknown integration order N_n. In practice, though, the latter parameter is rarely greater than 2, so the requirement is easily met even for low values of N_ψ.

3.2 Wavelet transform of the mixed Gaussian fractional process

In this section, we will define and carry out the basic properties of the wavelet transform for the mixed fractional process Y.

For a wavelet function $\psi \in L^2(\mathbb{R})$ with a number N_ψ of vanishing moments, the (normalized) vector wavelet transforms of Y as in (3.1) is naturally defined as

$$\mathbb{R}^n \ni D(2^j,k) = 2^{-j/2}\int_{\mathbb{R}} 2^{-j/2}\psi(2^{-j}t-k)Y(t)dt, \quad j \in \mathbb{N}\cup\{0\}, \quad k \in \mathbb{Z}, \quad (3.14)$$

provided the integral in (3.14) exists in an appropriate sense. It will be convenient to make the change of variable $z = 2^{-j}t - k$, and reexpress

$$D(2^j,k) = \int_{\mathbb{R}} \psi(z)Y(2^j z + 2^j k)dz.$$

The wavelet domain process $\{D(2^j,k)\}_{k\in\mathbb{Z}}$ is stationary in k (Proposition 3.2.1). The wavelet spectrum (variance) at scale j is the positive definite matrix

$$\mathbb{E}D(2^j,k)D(2^j,k)^* = \mathbb{E}D(2^j,0)D(2^j,0)^* =: \mathbb{E}W(2^j), \quad (3.15)$$

and its natural estimator, the sample wavelet variance, is the random matrix

$$W(2^j) = \frac{1}{K_j}\sum_{k=1}^{K_j} D(2^j,k)D(2^j,k)^*, \quad K_j = \frac{\nu}{2^j}, \quad j = j_1,\ldots,j_m, \quad (3.16)$$

for a total of ν available (wavelet) data points.

The next proposition describes some nice properties of the wavelet coefficients (3.14) as well as the general form of the wavelet spectrum (3.15).

Proposition 3.2.1. *Under the assumptions ($A1-2$), let $D(2^j,k)$ and $\mathbb{E}W(2^j,k)$ be as in (3.14) and (3.15), respectively. Then,*

($P1$) the wavelet transform (3.14) is well-defined in the mean square sense, and $\mathbb{E}D(2^j,k) = 0$;

($P2$) (stationarity for a fixed scale) $\{D(2^j,k+h)\}_{k\in\mathbb{Z}} \stackrel{d}{=} \{D(2^j,k)\}_{k\in\mathbb{Z}}$, $h \in \mathbb{Z}$;

(P3) the wavelet transform (3.14) satisfies $\{D(2^j,k)\}_{k\in\mathbb{Z}} \overset{d}{=} \{2^{jH}\widetilde{D}_j(1,k)\}_{k\in\mathbb{Z}}$*, where*

$$\widetilde{D}_j(1,k) = P\begin{pmatrix} \int_{\mathbb{R}} \overline{\widehat{\psi}(t)} \frac{1}{(ix)^{N_1}} |x|^{-(h_1-N_1+1/2)} g_1(\frac{x}{2^j}) dt \widetilde{B}(dx) \\ \vdots \\ \int_{\mathbb{R}} \overline{\widehat{\psi}(t)} \frac{1}{(ix)^{N_n}} |x|^{-(h_n-N_n+1/2)} g_n(\frac{x}{2^j}) dt \widetilde{B}(dx) \end{pmatrix}. \tag{3.17}$$

In (3.17), $\widehat{\psi}(t) := \int_{\mathbb{R}} \psi(t) e^{-itx} dt$*;*

(P4) the wavelet spectrum (3.15) can be expressed as

$$\mathbb{E}W(2^j) = 2^{jH}\left\{ \int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-(H+I/2)} G\left(\frac{x}{2^j}\right) |x|^{-(H+I/2)^*} dx \right\} 2^{jH^*}, \tag{3.18}$$

where

$$G(x) = P\mathrm{diag}(|g_1(x)|^2, \cdots, |g_n(x)|^2) P^*, \tag{3.19}$$

and

$$H = P\mathrm{diag}(h_1, \cdots, h_n) P^{-1}; \tag{3.20}$$

(P5) the wavelet spectrum has full rank, namely, $\det\mathbb{E}W(2^j) \neq 0$*,* $j \in \mathbb{N}$*.*

Remark 3.2.1. In regard to (P3) and (P4), we have

$$2^{-jH} D(2^j,k) \overset{P}{\to} P\begin{pmatrix} \int_{\mathbb{R}} \overline{\widehat{\psi}(t)} \frac{1}{(ix)^{N_1}} |x|^{-(h_1-N_1+1/2)} g_1(0) dt \widetilde{B}(dx) \\ \vdots \\ \int_{\mathbb{R}} \overline{\widehat{\psi}(t)} \frac{1}{(ix)^{N_n}} |x|^{-(h_n-N_n+1/2)} g_n(0) dt \widetilde{B}(dx) \end{pmatrix},$$

and

$$2^{-jH}\mathbb{E}W(2^j)2^{-jH^*} \to \int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-(H+I/2)} G(0) |x|^{-(H+I/2)^*} dx,$$

as $j \to \infty$. In this sense, $\{D(2^j,k)\}_{k\in\mathbb{Z}}$ and $\{\mathbb{E}W(2^j)\}_{k\in\mathbb{Z}}$ satisfy asymptotic scaling laws.

Moreover, by a standard calculation, the wavelet variance (3.18) can be recast as

$$\mathbb{E}W(2^j) = P\mathcal{E}(2^j)^{1/2}\mathrm{diag}(2^{2jh_1}, \cdots, 2^{2jh_n})\mathcal{E}(2^j)^{1/2}P^*, \tag{3.21}$$

where

$$\mathcal{E}(2^j) = \mathrm{diag}\Bigg(\int_{\mathbb{R}} |\widehat{\psi}(y)|^2|y|^{-2(h_1+1/2)}\Bigg|g_1\Big(\frac{y}{2^j}\Big)\Bigg|^2 dy, \cdots, \int_{\mathbb{R}} |\widehat{\psi}(y)|^2|y|^{-2(h_n+1/2)}\Bigg|g_n\Big(\frac{y}{2^j}\Big)\Bigg|^2 dy\Bigg). \tag{3.22}$$

3.3 Asymptotic theory for sample wavelet transforms: fixed scales

In this section, we derive the asymptotic behavior of the sample wavelet variance at fixed scales. As typical in the asymptotic study of averages, we begin by investigating the asymptotic covariance of the sample wavelet transforms $W(2^j)$.

Recall that for a zero mean, Gaussian random vector $\mathbf{Z} \in \mathbb{R}^m$, the Isserlis theorem (e.g., [25]) yields

$$\mathbb{E}(Z_1 \dots Z_{2k}) = \sum\prod \mathbb{E}(Z_i Z_j), \quad \mathbb{E}(Z_1 \dots Z_{2k+1}) = 0, \quad k = 1, \dots, \lfloor m/2 \rfloor. \tag{3.23}$$

The notation $\sum\prod$ stands for adding over all possible k-fold products of pairs $\mathbb{E}(Z_i Z_j)$, where the indices partition the set $1, \dots, 2k$. Proposition 3.3.1 below describes the asymptotic covariance matrix for the wavelet transform of the mixed fractional process Y at fixed octaves.

Proposition 3.3.1. *Let $Y = \{Y(t)\}_{t\in\mathbb{R}}$ satisfies the assumptions (A1 – 3). As $\nu \to \infty$, for every pair of octaves j, j',*

(*i*)

$$\sqrt{K_j}\sqrt{K_{j'}}\frac{1}{K_j}\frac{1}{K_{j'}}\sum_{k=1}^{K_j}\sum_{k'=1}^{K_{j'}}\mathbb{E}D(2^j,k)D(2^{j'},k')^*\otimes\mathbb{E}D(2^j,k)D(2^{j'},k')^*$$

$$\to 2^{-(j+j')/2}\,\mathrm{gcd}(2^j,2^{j'})\sum_{z=-\infty}^{\infty}\Phi_{z\mathrm{gcd}(2^j,2^{j'})}\otimes\Phi_{z\mathrm{gcd}(2^j,2^{j'})}, \tag{3.24}$$

where

$$\Phi_z := \int_{\mathbb{R}}\overline{\widehat{\psi}(2^j x)}\widehat{\psi}(2^{j'}x)e^{-izx}|x|^{-(H+I/2)}G(x)|x|^{-(H+I/2)^*}dx; \tag{3.25}$$

(*ii*) *there is a matrix* $G_{jj'}\in M(n(n+1)/2,\mathbb{R})$*, not necessarily symmetric, such that*

$$\sqrt{K_j}\sqrt{K_{j'}}\,\mathrm{Cov}(\mathrm{vec}_{\mathcal{S}}W(2^j),\mathrm{vec}_{\mathcal{S}}W(2^{j'}))\to G_{jj'}, \tag{3.26}$$

where the entries of $G_{jj'}$ *can be retrieved from* (3.24) *by means of* (3.23) *(see* (3.4) *on the notation* $\mathrm{vec}_{\mathcal{S}}$*).*

The following theorem establishes the asymptotic distribution of the vectorized sample wavelet spectrum at a fixed set of octaves.

Theorem 3.3.1. *Let* $Y=\{Y(t)\}_{t\in\mathbb{R}}$ *satisfies the assumptions (A1 – 3). Let* $j_1<\cdots<j_m$ *be a fixed set of octaves. Then*

$$\Big(\sqrt{K_j}(\mathrm{vec}_{\mathcal{S}}(W(2^j)-\mathbb{E}W(2^j))\Big)^T_{j=j_1,\ldots,j_m}\stackrel{d}{\to}\mathcal{N}_{\frac{n(n+1)}{2}\times m}(\mathbf{0},F), \tag{3.27}$$

as $\nu\to\infty$ *(see* (3.4) *on the notation* $\mathrm{vec}_{\mathcal{S}}$*). In* (3.27)*, the matrix* $F\in\mathcal{S}(\frac{n(n+1)}{2}m,\mathbb{R})$ *has the form* $F=(G_{jj'})_{j,j'=1,\ldots,m}$*, where each block* $G_{jj'}\in M(n(n+1)/2,\mathbb{R})$ *is described in Proposition 3.3.1.*

3.4 The asymptotic behavior of the eigenstructure of sample wavelet variance matrices

Asymptotic results on the eigenstructure of sample wavelet variance matrices are by themselves of interest because the latter do not generally follow a Wishart distribution. This results from the presence of residual correlation after the application of the wavelet transform. In order to state Theorem 3.4.1 below, consider the matrix spectral decompositions

$$W(2^j) = \widehat{O}_j L_j \widehat{O}_j^*, \quad \mathbb{E}W(2^j) = O_j \Lambda_j O_j^*, \quad \widehat{O}_j, O_j \in O(n), \tag{3.28}$$

where $L_j := \text{diag}(l_{j,1}, \ldots, l_{j,n})$, $\Lambda_j := \text{diag}(\lambda_{j,1}, \ldots, \lambda_{j,n})$, $\widehat{O}_j$, O_j have columns $\widehat{\mathbf{o}}_{j,\cdot i}$, $\mathbf{o}_{j,\cdot i}$, respectively, for $i = 1, \ldots, n$, and

$$l_{j,1} \leq \ldots \leq l_{j,n}, \quad \lambda_{j,1} \leq \ldots \leq \lambda_{j,n}, \quad \widehat{\mathbf{o}}_{j,1i} \geq 0, \quad \mathbf{o}_{j,1i} \geq 0, \quad i = 1, \ldots, n, \quad j = j_1, \ldots, j_m. \tag{3.29}$$

In other words, the eigenvalues appearing on the main diagonal entries of L_j and Λ_j are ordered from smallest to largest, and the entries on the first row of O_j and $\widehat{O}_j$ are all nonnegative, which makes these orthogonal matrices identifiable. Following [26], p. 427, we recall the definition of the so-named duplication matrix $D \in M(n^2, \frac{1}{2}n(n+1), \mathbb{R})$. It consists of the (unique) operator D that performs the transformation

$$D(\text{vec}_{\mathcal{S}}(A))^T = (\text{vec}(A + A^* - \text{dg}(A)))^T, \quad A = (a_{i_1 i_2})_{i_1, i_2 = 1, \ldots, n} \in M(n, \mathbb{R}), \tag{3.30}$$

where $\text{dg}(A) := \text{diag}(a_{11}, \ldots, a_{nn})$. Moreover, for $S \in \mathcal{S}(n, \mathbb{R})$ with ordered eigenvalues $\lambda_1 < \ldots < \lambda_n$ and their respective normalized eigenvectors $\mathbf{o}_{\cdot 1}, \ldots, \mathbf{o}_{\cdot n}$, we further define the operator

$$\mathcal{J}(S) = \begin{pmatrix} (\mathbf{o}_{\cdot 1}^T \otimes \mathbf{o}_{\cdot 1}^T)D \\ \vdots \\ (\mathbf{o}_{\cdot n}^T \otimes \mathbf{o}_{\cdot n}^T)D \\ (\mathbf{o}_{\cdot 1}^T \otimes (\lambda_1 I_n - S)^+)D \\ \vdots \\ (\mathbf{o}_{\cdot n}^T \otimes (\lambda_n I_n - S)^+)D \end{pmatrix}_{(n+n^2)\times n(n+1)/2}, \tag{3.31}$$

where we can apply the relation

$$\mathrm{vec}(A)D = \mathrm{vec}_{\mathcal{S}}(A + A^* - \mathrm{dg}(A)) \tag{3.32}$$

(see Lemma 3.7, (i), in [26]). The proof of Theorem 3.4.1 relies on Theorem 3.3.1, Theorem A.6.1 (on the weak convergence of eigenvalues and eigenvectors) and the Delta method.

Theorem 3.4.1. *Under assumption (A1–2), let* $\{W(2^j)\}_{j=j_1,\dots,j_m}$ *be a set of sample wavelet variance matrices (see (3.16)). Suppose*

$$\mathbb{E}W(2^j) \text{ has pairwise distinct eigenvalues,} \quad j = j_1,\dots,j_m, \tag{3.33}$$

and let F *be as in (3.27). Let the matrices* L_j*,* Λ_j*,* $\widehat{O}_j$*,* O_j *be as in (3.28). Then,*

$$\Big(\sqrt{K_j}\mathrm{vec}_{\mathcal{D}}(L_j - \Lambda_j), \sqrt{K_j}\mathrm{vec}(\widehat{O}_j - O_j)\Big)^T_{j=j_1,\dots,j_m} \xrightarrow{d} \mathcal{N}_{n(n+1)m}(\mathbf{0}, JFJ^*), \quad \nu \to \infty, \tag{3.34}$$

where $J = \mathrm{diag}(\mathcal{J}_1,\dots,\mathcal{J}_m)$ *and* $\mathcal{J}_i$*,* $i = 1,\dots,m$*, is given by* $\mathcal{J}(S)$ *in (3.31) with* $S := \mathbb{E}W(2^{j_i})$*.*

Remark 3.4.2. The conclusions of Theorem 3.4.1 may not hold when the condition

(3.33) is not in place. Proposition A.1.1 in Appendix A.1 illustrates this fact in a particular case.

In spite of Remark 3.4.2, the next proposition shows that the condition (3.33) is always satisfied at coarse scales. This fact is useful in the context of Theorem 3.5.2 below.

Proposition 3.4.1. *Under the assumptions (A1–2), let* $W(2^j)$ *be the sample wavelet variance* (3.16) *at octave* $j \in \mathbb{N}$. *Then, for large enough* j, *the matrix* $\mathbb{E}W(2^j)$ *has pairwise distinct eigenvalues.*

3.5 Wavelet-based estimation for the demixing matrix

In this section, we address the first step of the two-step wavelet-based approach, namely, demixing the process by jointly diagonalizing two wavelet variances.
The joint diagonalization of two matrices is a well-known problem. For the case of symmetric matrices, its description and full characterization can be stated as follows (see Theorem 4.5.17, (b), in [27]). Suppose C_0 and C_1 are symmetric and C_0 is nonsingular. Then, there are a nonsingular $S \in M(n, \mathbb{R})$ and complex diagonal matrices Λ_0 and Λ_1 such that

$$C_0 = S\Lambda_0 S^T, \quad C_1 = S\Lambda_1 S^T, \tag{3.35}$$

if and only if the matrix $C_0^{-1}C_1$ is diagonalizable (in its Jordan form). In light of this, we can cast a joint diagonalization algorithm in the form of pseudo-code.

Pseudo-code for exact joint diagonalization (EJD)
Input: C_0, C_1 are symmetric matrices and the former is positive definite;
Step 1: set $W = C_0^{-1/2}$ so that $C_0^{-1} = W^* W$;
Step 2: compute $Q \in O(n)$ in the spectral decomposition $W C_1 W^* = Q^* D_1 Q$;
Step 3: compute the demixing matrix $B := QW$;
Step 4: stop and exit.

Example 3.5.1. In view of (3.21), it is clear that $C_0 = \mathbb{E}W(2^{J_1})$, $C_1 = \mathbb{E}W(2^{J_2})$, $J_1 < J_2$, can be jointly diagonalized, where the underlying process is defined by (3.1) under the assumptions (A1–2). In addition,

$$C_0^{-1} C_1 = (P^*)^{-1} \Big(\mathrm{diag}(2^{2(J_2 - J_1)\, h_1}, \ldots, 2^{2(J_2 - J_1)\, h_n}) \mathcal{E}(2^{J_1})^{-1} \mathcal{E}(2^{J_2}) \Big) P^*.$$

This expression constitutes a diagonal Jordan decomposition, whence (3.35) holds.

The proposed wavelet-based estimator of a demixing matrix B is defined next.

Definition 3.5.1. Consider two octaves $0 \leq J_1 < J_2$ for which the eigenvalues of $\mathbb{E}W(2^{J_1})$ are pairwise distinct. For $\nu \in \mathbb{N}$, the wavelet-based demixing estimator $\widehat{B}_\nu$ is the output of the EJD algorithm when setting

$$C_0 = W(2^{J_1}) \text{ and } C_1 = W(2^{J_2}). \tag{3.36}$$

In Theorem 3.5.2, stated next, we establish the consistency and asymptotic normality of the estimator put forward in Definition 3.5.1. The result involves characterizing the set of solutions to the EJD algorithm. In light of (3.21), this relies on reexpressing

$$(C_0 =)\ \mathbb{E}W(2^{J_1}) = P\mathcal{E}(2^{J_1})^{1/2}\mathrm{diag}(2^{2J_1\,h_1}, 2^{2J_1\,h_2}, \ldots, 2^{2J_1\,h_n})\mathcal{E}(2^{J_1})^{1/2}P^* =: RR^*,$$

$$(C_1 =)\ \mathbb{E}W(2^{J_2}) = P\mathcal{E}(2^{J_2})^{1/2}\mathrm{diag}(2^{2J_2\,h_1}, 2^{2J_2\,h_2}, \ldots, 2^{2J_2\,h_n})\mathcal{E}(2^{J_2})^{1/2}P^* =: R\Lambda R^*, \tag{3.37}$$

where

$$R := P\mathcal{E}(2^{J_1})^{1/2}\mathrm{diag}(2^{J_1 h_1}, \ldots, 2^{J_1 h_n}),$$

$$\Lambda := \mathrm{diag}(2^{2(J_2-J_1)\,h_1}, \ldots, 2^{2(J_2-J_1)\,h_n})\mathcal{E}(2^{J_1})^{-1}\mathcal{E}(2^{J_2}), \tag{3.38}$$

and then making use of the matrix polar decomposition of R. Then, consistency and asymptotic normality stem from obtaining the behavior of the sample counterparts $W(2^{J_1})$ and $W(2^{J_2})$ vis-à-vis (3.37) by means of Theorem 3.3.1 and Theorem A.6.1, plus the Delta method when developing limits in distribution.

Theorem 3.5.2. *Suppose the assumptions (A1–3) hold, and let* $\mathcal{E}(2^j)$ *be as in* (3.22)*. Also let*

$$\mathcal{I} = \{\Pi \in M(n, \mathbb{R}) : \Pi \text{ has the form } \mathrm{diag}(\pm 1, \ldots, \pm 1)\}. \tag{3.39}$$

(i) Then,

$$\mathcal{M}_{\mathrm{EJD}} = \{\Pi\,\mathrm{diag}(2^{-J_1 h_1}, \ldots, 2^{-J_1 h_n})\mathcal{E}(2^{J_1})^{-1/2}P^{-1}, \Pi \in \mathcal{I}\} \tag{3.40}$$

is the set of matrix solutions produced by the EJD algorithm when setting

$$C_0 = \mathbb{E}W(2^{J_1}) \text{ and } C_1 = \mathbb{E}W(2^{J_2}); \tag{3.41}$$

(ii) for any estimator sequence $\{\widehat{B}_\nu\}_{\nu\in\mathbb{N}}$*,*

$$\widehat{B}_\nu \xrightarrow{P} \Pi\,\mathrm{diag}(2^{-J_1 h_1}, \ldots, 2^{-J_1 h_n})\mathcal{E}(2^{J_1})^{-1/2}P^{-1}, \quad \nu \to \infty, \tag{3.42}$$

for some matrix $\Pi \in \mathcal{I}$*;*

(iii) in addition, assume that condition (3.33) *holds for* $j = J_1$*. Then, an estimator sequence* $\{\widehat{B}_\nu\}_{\nu\in\mathbb{N}}$ *as described in (ii) satisfies*

$$\sqrt{\nu}(\mathrm{vec}(\widehat{B}_\nu - \Pi\,\mathrm{diag}(2^{-J_1 h_1}, \ldots, 2^{-J_1 h_n})\mathcal{E}(2^{J_1})^{-1/2}P^{-1}))^T \stackrel{d}{\to} \mathcal{N}(\mathbf{0}, \Sigma_F(J_1, J_2)) \tag{3.43}$$

for some matrix $\Pi \in \mathcal{I}$*, where the covariance matrix* $\Sigma_F(J_1, J_2)$ *is a function of* F*, and* F *is defined in Theorem 3.3.1, with* $m = 2$*.*

Remark 3.5.3. By (3.42), any sequence $\widehat{B}_\nu^{-1}$ has a limit in probability of the form $B^{-1} := P\mathrm{diag}(\beta_1, \ldots, \beta_n)$, $|\beta_i| \neq 0$, $i = 1, \ldots, n$, i.e., involving a non-identifiability factor post-multiplying the mixing matrix P. However, note that $H = P\mathrm{diag}(h_1, \ldots, h_n)P^{-1} = B^{-1}\mathrm{diag}(h_1, \ldots, h_n)B$, i.e., the columns of B^{-1} consist of (non-unit) eigenvectors of the Hurst matrix H. Consequently, $\widehat{H} := \widehat{B}_\nu^{-1}\mathrm{diag}(\widehat{h}_1, \ldots, \widehat{h}_n)\widehat{B}_\nu$ is a natural estimator of the latter, where $\widehat{h}_1, \ldots, \widehat{h}_n$ are univariate (e.g., wavelet-based) estimators of the individual Hurst exponents obtained from the demixed signal.

Nevertheless, producing a direct estimator of P is straightforward. Just normalize each column of the matrix estimator $\widehat{B}_\nu^{-1}$ and multiply it by -1 if necessary as to arrive at a matrix $\widehat{P}$ with positive diagonal entries (cf. (3.6)). This procedure is used in Section 3.7 below.

Remark 3.5.4. More precisely, the covariance matrix in the limit (3.43) can be written as $\Sigma_F(J_1, J_2) = A_3\Sigma_3 A_3^*$, where Σ_3 and A_3 are given by expressions (A.54) and (A.55), respectively. It is clear that the expression for $\Sigma_F(J_1, J_2)$ is quite intricate, and the construction of theoretical confidence intervals is a matter for future investigation (cf. Combrexelle et al. [28]).

Remark 3.5.5. Let T_P be a matrix such that

$$(\mathrm{vec}(\widehat{I}_\nu - I))^T = T_P(\mathrm{vec}(\widehat{B}_\nu - \Pi\,\mathrm{diag}(2^{-J_1 h_1}, \ldots, 2^{-J_1 h_n})\mathcal{E}(2^{J_1})^{-1/2}P^{-1}))^T,$$

where

$$\widehat{I}_\nu = \widehat{B}_\nu P(\Pi \,\mathrm{diag}(2^{-J_1 h_1}, \ldots, 2^{-J_1 h_n})\mathcal{E}(2^{J_1})^{-1/2})^{-1}. \tag{3.44}$$

Then, by (3.43) and the Delta method

$$\sqrt{\nu}(\mathrm{vec}(\widehat{I}_\nu - I))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma(J_1, J_2)), \tag{3.45}$$

where $\Sigma(J_1, J_1) = T_P \Sigma_F(J_1, J_2) T_P^*$.

We will use (3.45) in next section to establish limiting distribution of the Hurst parameter estimation.

Remark 3.5.6. Removing the condition (3.5) can alter the limits (3.42) and (3.43). For example, if $h_1 = h_2 = \ldots = h_n =: h$, then in **Step 2** of the algorithm, $\widehat{W}\widehat{C}_1\widehat{W}^* \xrightarrow{P} 2^{2h\,(J_2 - J_1)} I$. Thus, the eigenvectors of $\widehat{W}\widehat{C}_1\widehat{W}^*$ do not have a limit in probability.

3.6 Wavelet-based estimation for the Hurst parameters

In this section, we focus on the second step of the two-step wavelet-based approach, namely, developing the estimation of the Hurst parameters by applying univariate wavelet analysis on entries of the demixed process. We establish the asymptotic normality of the Hurst parameters' estimators as the wavelet scale grows according to a factor $a(\nu) \to \infty$, as $\nu \to \infty$. It is necessary to take the coarse scale limit due to the fact that the source signal X is not self-similar (c.f. (2.10)). Throughout this section, the factor $a(\nu)$ is assumed to be a dyadic sequence such that

$$\frac{a(\nu)}{\nu} + \frac{\nu}{a(\nu)^{1+2\beta}} \to 0, \quad \nu \to \infty \tag{3.46}$$

(see Remark 3.6.4 below on the choice of $a(\nu)$ in practice).

Let $\widehat{B}_\nu$ be the demixing matrix described by (3.42), let $\widehat{I}_\nu$ be defined by (3.44), and let $\mathfrak{D} := \Pi\,\mathrm{diag}(2^{-J_1 h_1},\ldots,2^{-J_1 h_n})\mathcal{E}(2^{J_1})^{-1/2}$. Then, the demixed process is defined by

$$\widehat{X} := \widehat{B}_\nu Y = \widehat{B}_\nu P \mathfrak{D}^{-1}\mathfrak{D}X = \widehat{I}_\nu \mathfrak{D} X. \tag{3.47}$$

The following proposition gives the asymptotic distribution of the sample wavelet variance of $\widehat{X}$.

Proposition 3.6.1. *Let $\widehat{X}$ be the demixed process defined by (3.47), let $W_{\widehat{X}}$ be the sample wavelet variance of $\widehat{X}$, and let $\mathbb{E}W_X$ be the wavelet variance of the source signal X. Then,*

$$\Big(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1},\cdots,a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(W_{\widehat{X}}(a(\nu)2^j)-\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}))^T\Big)_{j=j_1,\cdots,j_m}$$

$$\stackrel{d}{\to} \mathcal{N}(0,\mathcal{K}\mathbf{W}\mathcal{K}^*), \tag{3.48}$$

as $\nu\to\infty$ (see (3.4) on the notation $\mathrm{vec}_{\mathcal{D}}$). In (3.48), the matrices $\mathbf{W}$ and $\mathcal{K}$ are defined by (A.59) and (A.58), respectively.

Remark 3.6.1. Intuitively, the demixing estimator $\widehat{B}_\nu$ yields a signal $\widehat{X}$ that is close to the unknown source X up to a non-identifiability factor $\mathfrak{D}$. In fact the limiting distribution of

$$\Big(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1},\cdots,a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(W_{\mathfrak{D}X}(a(\nu)2^j)-\mathbb{E}W_{\mathfrak{D}X}(a(\nu)2^j)))^T\Big)_{j=j_1,\cdots,j_m}$$

is also $\mathcal{N}(0,\mathcal{K}\mathbf{W}\mathcal{K}^*)$. In particular, the diagonal entries of the sample wavelet variance $W_{\widehat{X}}(a(\nu)2^j)$ of the demixed signal $\widehat{X}$ are asymptotically independent.

Let

$$\widehat{\sigma}^2_{\widehat{X}_i} \text{be the } (i,i)\text{-th entry of } W_{\widehat{X}}, \tag{3.49}$$

and let

$$\sigma^2_{X_i} \text{be the } (i,i)\text{-th entry of } \mathbb{E}W_X. \tag{3.50}$$

The wavelet based regression estimator of the Hurst parameters h_i, $i = 1, \cdots, n$, involves regressing the terms $\widehat{\sigma}_{\widehat{X}_i}(a(\nu)2^j)$ on the scale index j, i.e.,

$$\begin{pmatrix} \widehat{h}_1 \\ \vdots \\ \widehat{h}_n \end{pmatrix} := \begin{pmatrix} \sum_{j=j_1}^{j_m} w_j^1 \log(\widehat{\sigma}^2_{\widehat{X}_1}(a(\nu)2^j)) \\ \vdots \\ \sum_{j=j_1}^{j_m} w_j^n \log(\widehat{\sigma}^2_{\widehat{X}_n}(a(\nu)2^j)) \end{pmatrix}, \tag{3.51}$$

where the weight vectors

$$\mathbf{w}^i = (w_1^i, \cdots, w_m^i)^T \tag{3.52}$$

satisfy

$$\sum_{j=1}^{m} w_j^i = 0, \quad 2\log 2 \sum_{j=1}^{m} j w_j^i = 1, \quad i = 1, \cdots, n. \tag{3.53}$$

The next proposition gives the difference between the wavelet variance of X_i and 2^{j2h_i}, $i = 1, \cdots, n$, and will be used to prove Theorem 3.6.2.

Proposition 3.6.2. *For $i = 1, \cdots, n$, let $\sigma^2_{X_i}$ be defined by (3.50). Then,*

$$|\sigma^2_{X_i}(2^j) - 2^{j2h_i}|g_i(0)|^2 K(h_i)| \le C 2^{j(2h_i - \beta)}, \tag{3.54}$$

where $K(h) = \int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-2h-1} dx$.

Heuristic justification for the regression estimator: By Proposition 3.6.2,

$$|\sigma^2_{X_i}(a(\nu)) - a(\nu)^{2h_i}|g_i(0)|^2 K(h_i)| \le C a(\nu)^{2h_i - \beta}. \tag{3.55}$$

By (3.48), we have

$$\sqrt{\nu/a(\nu)}a(\nu)^{-2h_i}(\widehat{\sigma}^2_{\widehat{X}_i}(a(\nu)) - \mathfrak{D}(i,i)^2\sigma^2_{X_i}(a(\nu))) \xrightarrow{d} \mathcal{N}(0,b),$$

for some $b > 0$. Thus, $\widehat{\sigma}^2_{\widehat{X}_i}(a(\nu)) \overset{d}{\sim} \mathfrak{D}(i,i)^2\sigma^2_{X_i}(a(\nu)) + \sqrt{a(\nu)/\nu}a^{2h_i}Z$, $Z \overset{d}{\sim} \mathcal{N}(0,b)$. Now consider

$$\log \widehat{\sigma}^2_{\widehat{X}_i}(a(\nu)) - \log(\mathfrak{D}(i,i)^2K(h_i)|g_i(0)|^2a(\nu)^{2h_i})$$
$$\overset{d}{\sim} \log\left(\mathfrak{D}(i,i)^2\sigma^2_{X_i}(a(\nu)) + \sqrt{a(\nu)/\nu}a^{2h_i}Z\right) - \log(\mathfrak{D}(i,i)^2K(h_i)|g_i(0)|^2a(\nu)^{2h_i})$$
$$\overset{d}{=} \log\mathfrak{D}(i,i)^2\sigma^2_{X_i}(a(\nu)) - \log(\mathfrak{D}(i,i)^2K(h_i)|g_i(0)|^2a(\nu)^{2h_i}) + \log\left(1+\frac{\sqrt{a(\nu)}a^{2h_i}}{\sqrt{\nu}\mathfrak{D}(i,i)^2\sigma^2_{X_i}(a(\nu))}Z\right)$$
$$\overset{d}{=} \log\left(\frac{\sigma^2_{X_i}(a(\nu))}{K(h_i)|g_i(0)|^2a(\nu)^{2h_i}} - 1 + 1\right) + \log\left(1 + \frac{\sqrt{a(\nu)}a^{2h_i}}{\sqrt{\nu}\mathfrak{D}(i,i)^2\sigma^2_{X_i}(a(\nu))}Z\right)$$
$$\overset{d}{\sim} Ca(\nu)^{-\beta} + \sqrt{a(\nu)/\nu}Z/\mathfrak{D}(i,i)^2.$$

As a consequence,

$$\sqrt{\nu/a(\nu)}\left(\log\widehat{\sigma}^2_{\widehat{X}_i}(a(\nu)) - 2h_i\log(a(\nu)) - \log\mathfrak{D}(i,i)^2K(h_i)g_i^2(0)\right)$$
$$\overset{d}{=} Z/\mathfrak{D}(i,i)^2 + Ca(\nu)^{-\beta}\sqrt{\nu/a(\nu)}.$$

Condition (3.46) ensures that $\sqrt{\nu/a(\nu)} \to \infty$, and the bias term $Ca(\nu)^{-\beta}\sqrt{\nu/a(\nu)}$ vanishes. Thus, we have the following central limit theorem for the estimators $\widehat{h}_i$, $i = 1, \cdots, n$.

Theorem 3.6.2. *Suppose the estimator $(\widehat{h}_1, \cdots, \widehat{h}_n)$ be defined by (3.51). Then,*

$$\sqrt{\nu/a(\nu)}\left[\begin{pmatrix}\widehat{h}_1\\ \vdots\\ \widehat{h}_n\end{pmatrix} - \begin{pmatrix}h_1\\ \vdots\\ h_n\end{pmatrix}\right] \xrightarrow{d} \mathcal{N}(0,\mathcal{W}). \tag{3.56}$$

In (3.56),

$$\mathcal{W} = \operatorname{diag}((\mathbf{w}^1)^T V(h_1)\mathbf{w}^1, \cdots, (\mathbf{w}^n)^T V(h_n)\mathbf{w}^n),$$

the weight vectors $\mathbf{w}^i$, $i = 1, \cdots, n$ *satisfy* (3.53), *and the matrix* $V(h) = \{V_{k_1,k_2}(h)\}_{k_1,k_2=1,\cdots,m}$ *is defined entrywise by*

$$V_{k_1,k_2}(h) = \frac{4\pi b_{j_{k_1},j_{k_2}}^{4h+1}}{2^{2(j_{k_1}+j_{k_2})h} K^2(h)} \int_{\mathbb{R}} x^{-(4h+2)} |\widehat{\psi}(2^{j_{k_1}} x / b_{j_{k_1},j_{k_2}})|^2 |\widehat{\psi}(2^{j_{k_2}} x / b_{j_{k_1},j_{k_2}}))|^2 dx, \tag{3.57}$$

where $K(h) = \int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-2h-1} dx$ *and* $b_{j_{k_1},j_{k_2}} = \gcd(2^{j_{k_1}}, 2^{j_{k_2}})$.

Remark 3.6.3. Theorem 3.6.2 shows that the Hurst estimators $\widehat{h}_1, \cdots, \widehat{h}_n$ are asymptotically independent. The asymptotic distribution of the Hurst estimator by estimating from $\widehat{X}$ is equal to the asymptotic distribution of directly estimate the Hurst parameters from the source X.

Remark 3.6.4. In practice, the choice of $a(\nu)$ involves a statistical compromise. A large value of $a(\nu)$ with respect to ν implies a relatively small bias, but also cause a relatively large variance of the estimator. Simulation results suggest the ratio $\nu/a(\nu)2^j$ should be no less than 2^3.

3.7 Simulation studies

3.7.1 Performance of demixing procedure for operator fractional Brownian motion

To study the performance of the wavelet demixing method over finite samples, we simulated $R = 500$ sample paths of sizes $n = 2^{20}$ and 2^{10} of 4-dimensional operator fractional Brownian motion with parameter $H = P\mathrm{diag}(0.2, 0.4, 0.6, 0.8)P^{-1}$ and

$$P = \begin{pmatrix} 0.6834 & -0.7142 & 0.6960 & -0.1165 \\ -0.0096 & 0.4539 & -0.0908 & 0.7740 \\ 0.4771 & -0.2345 & 0.3359 & -0.4243 \\ 0.5525 & -0.4784 & -0.6281 & 0.4553 \end{pmatrix} \tag{3.58}$$

(see also Remark 3.7.2 on the choice of P). For notational simplicity, denote $X := B_{\mathbf{h}}$, $Y := B_H$. To investigate the effect of mixing and demixing, we applied the univariate wavelet regression method of Veitch and Abry [17] and estimated the entry-wise Hurst exponents $h_{X,i}$, $h_{Y,i}$, $i = 1, \dots, n$, starting from simulated sample paths of the unmixed and mixed processes X and Y. We further used the same procedure to estimate the entry-wise Hurst exponents $h_{Z,i}$, $i = 1, \dots, n$, of the demixed sequence $Z = \widehat{P}^{-1} B_H$ for demixing matrix estimates $\widehat{P}^{-1}$.

The results consist of comparisons of the Monte Carlo log-averages $\log_2 \widehat{\mathbb{E}} W_{X,i}(2^j)$, $\log_2 \widehat{\mathbb{E}} W_{Y,i}(2^j)$ and $\log_2 \widehat{\mathbb{E}} W_{Z,i}(2^j)$ for each of the $n = 4$ components for the sample sizes 2^{20} and 2^{10} (Figures 3.1 and 3.4); boxplots for $\widehat{h}_{X,i} - h_i$, $\widehat{h}_{Y,i} - h_i$ and $\widehat{h}_{Z,i} - h_i$, $i = 1, 2, 3, 4$ (Figures 3.2 and 3.5); and boxplots for the 16 entries of $\widehat{P^{-1}}P - I$ (Figures 3.3 and 3.6). Following the procedure described in Remark 3.5.3, the columns of $\widehat{P}$ were adjusted as to eliminate the non-identifiability factor. In all cases, the sample wavelet variance matrices were computed based on Daubechies wavelet filters with $N_\psi = 2$ vanishing moments. Using a different wavelet with $N_\psi \geq 2$ yields similar conclusions.

In Figures 3.1 and 3.4, as expected for the mixed data Y all components of $\log \widehat{\mathbb{E}} W_{Y,i}(2^j)$ display patent departures from the original data $\log \widehat{\mathbb{E}} W_{X,i}(2^j)$. After demixing, all components of $\log \widehat{\mathbb{E}} W_{Z,i}(2^j)$ remarkably superimpose those of $\log \widehat{\mathbb{E}} W_{X,i}(2^j)$, with the possible exception of a few coarse scales for $h = 0.2$ and 0.4. In addition, the boxplots in Figures 3.2 and 3.5 show that the Monte Carlo distributions for $\widehat{h}_{Z,i} - h_i$

resemble those of $\widehat{h}_{X,i} - h_i$, which illustrates the successful demixing of Y. Figures 3.3 and 3.6 further indicate that $\widehat{P^{-1}}$ is very well estimated with negligible biases. In all comparisons, as expected the observed estimator properties improve significantly when passing from the relatively small sample size 2^{10} to the large sample size 2^{20}, hence reflecting the asymptotic statement of Theorem 3.5.2, (iii). In addition, simulation results not displayed also show that the standard deviation of the estimates decreases with the sample size according to the scaling ratio $C/\sqrt{\nu}$ for some $C > 0$, as anticipated.

Remark 3.7.1. Theorem 3.5.2 leaves open the question of how to optimally choose the octaves $j_1 < j_2$. For multiple choices of wavelet octaves, namely, $j_1 = 1$ (which involves the largest number of sum terms in (3.16)) and $j_2 = 2, \dots, 6$, Table 3.1 shows the performance of the individual Hurst eigenvalues' estimators in terms of Monte Carlo bias, standard deviation and (square root) mean squared error. For sample sizes 2^{20} and 2^{10}, the results indicate that for low values of the Hurst eigenvalues, the use of two widely separated wavelet octaves produces better results in terms of mean squared error, whereas for large values of the Hurst eigenvalues the choice of octaves has little impact on the estimation.

Remark 3.7.2. Simulation studies not included show that the choice of the mixing matrix (3.58) does not substantially affect the finite sample results. Moreover, the demixing estimator is very robust with respect to the condition number of the mixing matrix P. The distributions of the estimated scalar Hurst eigenvalues after demixing are barely affected for condition numbers of the order of at least 10^5.

Remark 3.7.3. In practice, a continuous time sample path is not available and thus the theoretical wavelet coefficient $D(2^j, k)$ cannot be computed. Instead, one approximates the latter by means of the classical recursive (or pyramidal) discrete filter bank algorithm (see chapter 7 in [24]). In Section 4.1, we lay out the mathematical frame work for estimation based on discrete time observations. We prove that, under mild

h	j_1, j_2	$\widehat{h}$ (2^{20})	bias	sd	$\sqrt{\text{MSE}}$	$\widehat{h}$ (2^{10})	bias	sd	$\sqrt{\text{MSE}}$
0.20	1,2	0.25	0.05	0.04	0.06	0.31	0.11	0.10	0.14
	1,3	0.22	0.02	0.03	0.04	0.25	0.05	0.08	0.10
	1,4	0.22	0.02	0.03	0.03	0.24	0.04	0.08	0.09
	1,5	0.21	0.01	0.02	0.03	0.23	0.03	0.08	0.09
	1,6	0.21	0.01	0.02	0.03	0.22	0.02	0.08	0.08
0.40	1,2	0.40	-0.00	0.02	0.02	0.45	0.05	0.08	0.10
	1,3	0.40	-0.00	0.01	0.02	0.41	0.01	0.07	0.07
	1,4	0.39	-0.01	0.01	0.02	0.40	0.00	0.07	0.07
	1,5	0.40	-0.00	0.01	0.01	0.40	0.00	0.07	0.07
	1,6	0.39	-0.01	0.01	0.01	0.40	-0.00	0.07	0.07
0.60	1,2	0.59	-0.01	0.01	0.02	0.60	-0.00	0.07	0.07
	1,3	0.59	-0.01	0.01	0.02	0.58	-0.02	0.07	0.07
	1,4	0.59	-0.01	0.01	0.02	0.58	-0.02	0.07	0.07
	1,5	0.59	-0.01	0.01	0.02	0.58	-0.02	0.07	0.07
	1,6	0.59	-0.01	0.01	0.02	0.58	-0.02	0.07	0.07
0.80	1,2	0.79	-0.01	0.01	0.02	0.76	-0.04	0.07	0.08
	1,3	0.79	-0.01	0.01	0.02	0.77	-0.03	0.07	0.07
	1,4	0.79	-0.01	0.01	0.02	0.77	-0.03	0.07	0.07
	1,5	0.79	-0.01	0.01	0.02	0.77	-0.03	0.07	0.07
	1,6	0.79	-0.01	0.01	0.02	0.77	-0.03	0.07	0.07

Table 3.1: **Choice of scales** 1,000 Monte Carlo runs, sample sizes 2^{20} and 2^{10}, $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$.

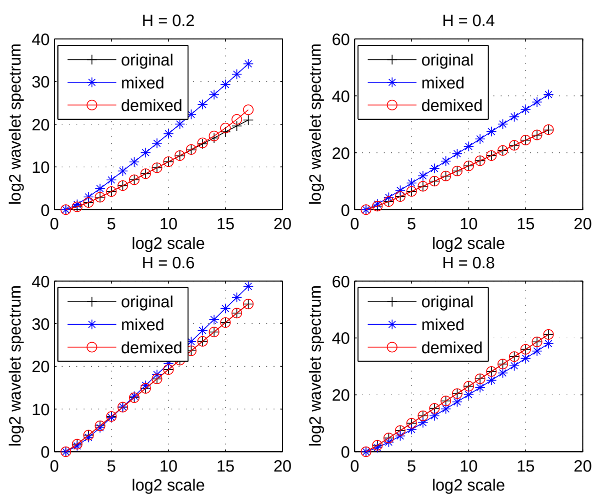

Figure 3.1: **Scaling** $\log W_{\cdot,\cdot}(2^j)$ vs. j for each of the $n = 4$ components based on the wavelet variance scales 2^1 and 2^2. The plots were produced by means of 500 Monte Carlo runs of sample size 2^{20}, with parameter values $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$ and $N_\psi = 2$.

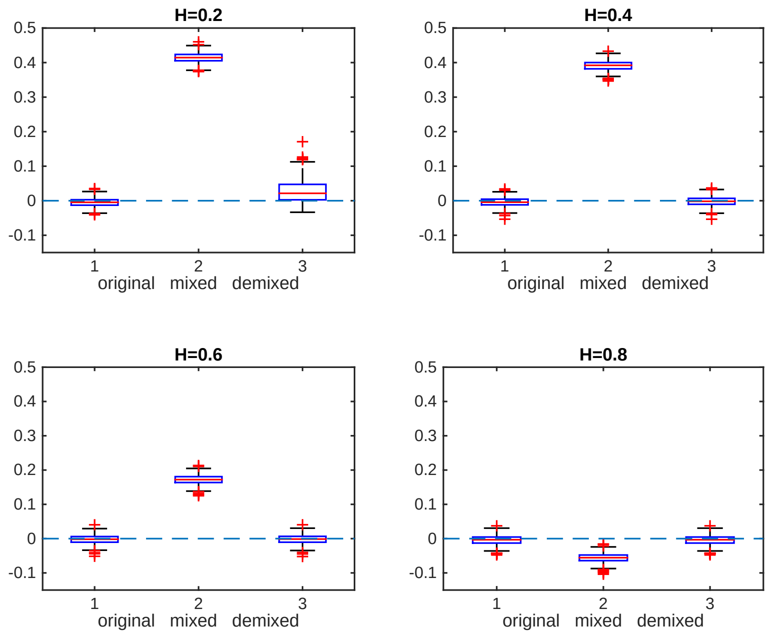

Figure 3.2: **Boxplots** based on the wavelet variance scales 2^1 and 2^2 for $i = 1, 2, 3, 4$, $\widehat{h}_{X,i} - h_i$ (original, left), $\widehat{h}_{Y,i} - h_i$ (mixed, middle) and $\widehat{h}_{Z,i} - h_i$ (demixed, right), for each of the $n = 4$ components, sorted by ascending order in terms of h. The plots were produced by means of 500 Monte Carlo runs of sample size 2^{20}, with parameter values $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$ and $N_\psi = 2$.

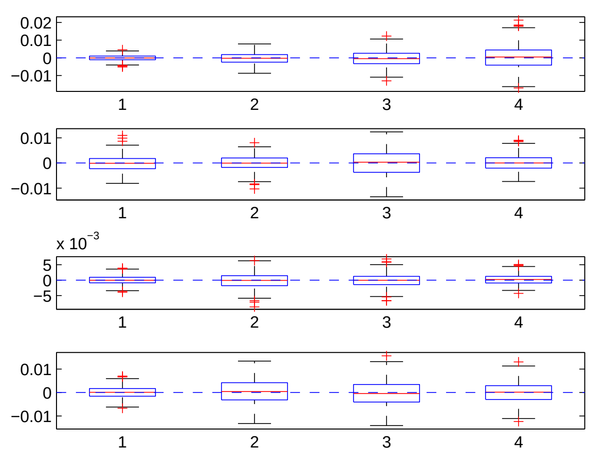

Figure 3.3: **Boxplots** based on the wavelet variance scales 2^1 and 2^2 for the 16 entries of $\widehat{P^{-1}}P - I$. The (i_1, i_2)-th boxplot denotes the (i_1, i_2)-th entry of $\widehat{P^{-1}}P - I$. The plots were produced by means of 500 Monte Carlo runs of sample size 2^{20}, with parameter values $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$ and $N_\psi = 2$.

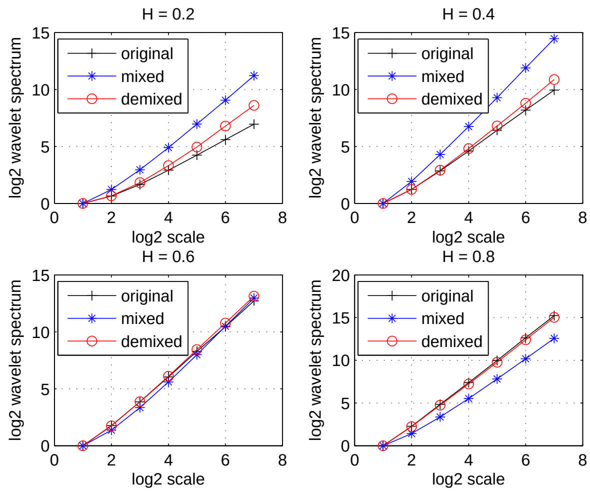

Figure 3.4: $\log W_{\cdot,\cdot}(2^j)$ vs. j for each of the $n = 4$ components based on the wavelet variance scales 2^1 and 2^2. The plots were produced by means of 500 Monte Carlo runs of sample size 2^{10}, with parameter values $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$ and $N_\psi = 2$.

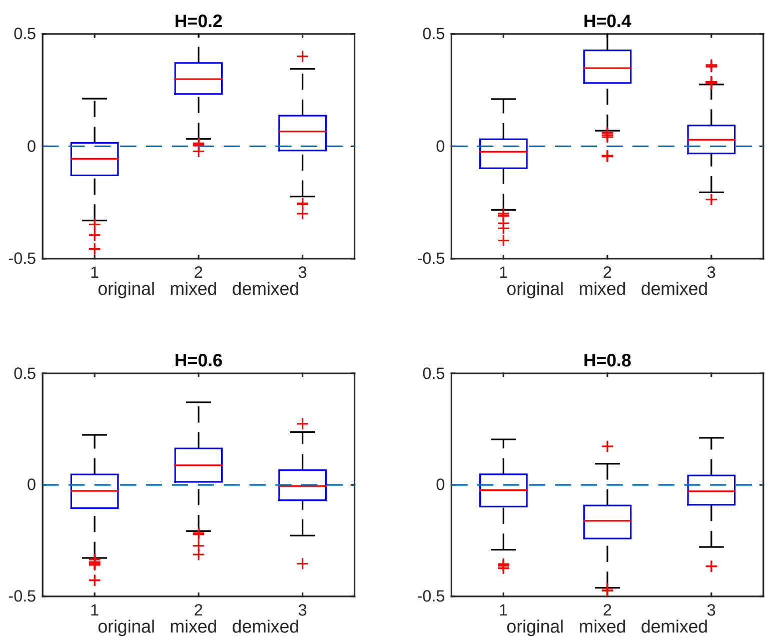

Figure 3.5: **Boxplots** based on the wavelet variance scales 2^1 and 2^2 for $i = 1, 2, 3, 4$, $\widehat{h}_{X,i} - h_i$ (original, left), $\widehat{h}_{Y,i} - h_i$ (mixed, middle) and $\widehat{h}_{Z,i} - h_i$ (demixed, right), for each of the $n = 4$ components, sorted by ascending order in terms of h. The plots were produced by means of 500 Monte Carlo runs of sample size 2^{10}, with parameter values $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$ and $N_\psi = 2$.

conditions, the estimated Hurst parameters stemming from the pyramidal algorithm also satisfy the weak limits (3.56).

3.7.2 Wavelet-based estimation v.s. spectral maximum likelihood estimation

Spectral maximum likelihood estimation is an alternative way to estimate mixed operator fractional Gaussian noise (see detail in [29]). In this section, we compare the two-step wavelet-based estimation with spectral maximum likelihood estimation by means of Monte Carlo simulation. We briefly introduce spectral maximum likelihood method for the read's convenience.

Let $\{Y(t)\}_{t\in\mathbb{R}} = \{PX(t)\}_{t\in\mathbb{R}}$, where P is the mixing matrix, X is a vector of two independent fractional Gaussian noises with Hurst parameters h_i, $i = 1, 2$. Then, the matrix-valued spectral density of Y is

$$f(x; h_1, h_2, A) = 2(1 - \cos x)AG(x; h_1, h_2)A^*,$$

where $A = P\text{diag}(e(h_1), e(h_2))$, $e(h_i) = \{\Gamma(2h_i + 1)\sin(\pi h_i)/2\pi\}^{1/2}$, $G(x; h_1, h_2) = \text{diag}(R(x; h_1), R(x; h_2))$, $R(x; h_i) = \sum_{\nu=-\infty}^{\infty} |x + 2\nu\pi|^{-2h_i-1}$, for $i = 1, 2$. Use the method in [30], f can be approximated by

$$\widetilde{f}(x; h_1, h_2, A) = 2(1 - \cos x)A\widetilde{G}(x; h_1, h_2)A^*,$$

where $\widetilde{G}(x; h_1, h_2) = \text{diag}(\widetilde{R}(x, h_1), \widetilde{R}(x, h_2))$, and $\widetilde{R}(x, h_i)) = \frac{1}{4\pi h_i}\{(2\pi M - x)^{-2h_i} + (2\pi M + x)^{-2h_i}\} + \sum_{k=-M}^{M} |x + 2k\pi|^{-2h_i-1}$ for some large integer M. In turn, the (negative) Whittle log-likelihood function of Y can be approximated by

$$l(h_1,h_2,A) = \sum_{i=1}^{T}\left[\log\det\widetilde{f}(x_i;h_1,h_2,A) + \mathrm{tr}\{\widetilde{f}(x_i;h_1,h_2,A)^{-1}I_Y(x_i)\}\right]$$

$$= 2T\log|\det A| + \sum_{i=1}^{T}\left\{\log|2(1-\cos x_i)\det\left(\widetilde{G}(x_i;h_1,h_2)\right)|\right\}$$

$$+\sum_{i=1}^{T}\mathrm{tr}\left[(A^*)^{-1}\{2(1-\cos x_i)\widetilde{G}(x_i;h_1,h_2)\}^{-1}A^{-1}I_Y(x_i)\right], \tag{3.59}$$

where $T = [(\nu-1)/2]$, ν be the sample size, $I_Y(x) = J_Y(x)J_Y(x)^*/(2\pi\nu)$, $J_Y(x) = \sum_{t=1}^{\nu} Y_t \exp(itx)$, $x_j = 2\pi j/\nu$ ([31]). And the spectral maximum likelihood estimator is defined by

$$\widehat{\theta} := \mathrm{argmin}_{\theta} l(\theta). \tag{3.60}$$

In (3.60), $l(\cdot)$ is defined by (3.59), $\widehat{\theta} = (\widehat{h}_1, \widehat{h}_2, \widehat{P})$. When implementing the estimator (3.60) in Matlab, we use the function fminsearch.m to minimize $l(h_1,h_2,A)$ with respect to the unknown parameters h_1, h_2 and A.

In the following simulation study, set

$$\nu = 1024, \quad M = \nu^{0.7}, \quad (h_1,h_2) = (0.4, 0.8), \quad P = \begin{pmatrix} 0.4472 & 0.3162 \\ 0.8944 & 0.9487 \end{pmatrix}.$$

Averages of 100 simulations of the spectral maximum likelihood estimators and wavelet estimators for the parameters h_1, h_2 and P are shown in the following table.

From the simulation study, although the spectral maximum likelihood estimation is a bit more accurate, the coupled demixing-wavelet method is far more computationally efficient. In particular, for the former estimator, we need to minimize (3.59) with respect to $n + n^2$ unknown parameters, which can be numerically very difficult for large values of n. Our computational studies also indicate that the minimization procedure is somewhat sensitive to the initial guess.

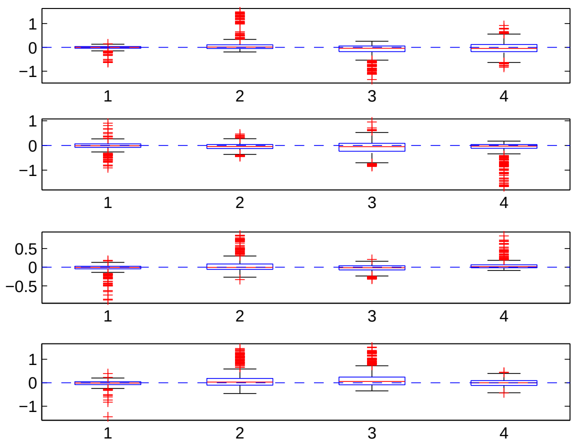

Figure 3.6: **Boxplots** based on the wavelet variance scales 2^1 and 2^2 for the 16 entries of $\widehat{P^{-1}}P - I$. The (i_1, i_2)-th boxplot denotes the (i_1, i_2)-th entry of $\widehat{P^{-1}}P - I$. The plots were produced by means of 500 Monte Carlo runs of sample size 2^{10}, with parameter values $\mathbf{h} = (0.2, 0.4, 0.6, 0.8)$ and $N_\psi = 2$.

method and time	parameter	bias	sd	$\sqrt{MSE}$
SML	h_1	0.0134	0.0632	0.0643
99.6548min	h_2	-0.0094	0.0707	0.0710
	$p_{1,1}$	-0.0040	0.0220	0.0223
	$p_{1,2}$	-0.0234	0.1314	0.1329
	$p_{2,1}$	0.0017	0.0092	0.0093
	$p_{2,2}$	-0.0551	0.3158	0.3190
wavelet	h_1	0.0346	0.1043	0.1094
0.0555min	h_2	-0.0193	0.1096	0.1108
	$p_{1,1}$	0.0001	0.0139	0.0138
	$p_{1,2}$	-0.0096	0.0348	0.0359
	$p_{2,1}$	-0.0002	0.0070	0.0070
	$p_{2,2}$	0.0025	0.0105	0.0107

Table 3.2: Biases, standard deviations and mean squared errors (in terms of square root) of 100 replications of the two underlined estimation methods for the parameters h_1, h_2 and $P = (p_{ij})_{i,j=1,2}$. Wavelet estimation: Number of vanishing moments: 1, $(j_1, j_2) = (3, 5)$, $[J_1, J_2] = [1, 2]$. SML estimation: M=$\nu^{0.7}$. Sample size: $\nu = 2^{10}$.

Chapter 4

Extensions and Applications

Since in practice, only observations in discrete time are available, the theoretical wavelet coefficients $D(2^j, k)$ cannot be computed, it is necessary to study the performance of the estimators under the assumption that only a discrete sample

$$\{Y(k)\}_{k\in\mathbb{Z}} \tag{4.1}$$

of (3.1) is available. In Section 4.1, we develop the asymptotic distribution of the estimators, and Section 4.2 provides an application in tree ring data. All the proofs can be found in Appendix A.4.

4.1 Estimation based on the discretized wavelet transform

4.1.1 Notation and assumptions

Throughout this chapter, we suppose the wavelet approximation coefficients stem from Mallat's pyramidal algorithm, under a multiresolution analysis of $L^2(\mathbb{R})$ (MRA; see [24], chapter 7). Accordingly, we need to replace $(W2)$ with the following more restrictive condition.

Assumption $(W2')$: φ and ψ are compactly supported, integrable and

$$\varphi(0) = 1, \quad \text{and} \quad \int_{\mathbb{R}} \psi^2(t)dt = 1.$$

Moreover, we also need the following additional condition.

Assumption $(W4)$: the function

$$\sum_{k\in\mathbb{Z}} k^m \varphi(\cdot - k)$$

is a polynomial of degree m for all $m = 0, \cdots, N_\psi - 1$.

All through this chapter, we assume the $(W1)$,$(W2')$ and $(W3-4)$ hold. Assumption $(W1)$ and $(W4)$ imply that

$$\int_{\mathbb{R}} \psi(2^{-j}t) \sum_{l\in Z} \varphi(t+l) l^m dt = 0, \quad j \geq 0, \quad m = 0, \cdots, N_\psi - 1. \tag{4.2}$$

4.1.2 Main results

Given (4.1), we initialize the algorithm with the vector-valued sequence

$$\mathbb{R}^n \ni \widetilde{a}_{0,k} := Y(k), \quad k \in \mathbb{Z},$$

also called the approximation coefficients at scale $2^0 = 1$. At coarser scales 2^j, Mallat's algorithm is characterized by the iterative procedure

$$\widetilde{a}_{j+1,k} = \sum_{k'\in\mathbb{Z}} h_{k'-2k} \widetilde{a}_{j,k'}, \quad \widetilde{d}_{j+1,k} = \sum_{k'\in\mathbb{Z}} g_{k'-2k} \widetilde{a}_{j,k'}, \quad j \in \mathbb{N}, \quad k \in \mathbb{Z},$$

where the filter sequences $\{h_k\}_{k\in\mathbb{Z}}$, $\{g_k\}_{k\in\mathbb{Z}}$ are called low- and high-pass MRA filters, respectively. Due to $(W2')$, only a finite number of filter terms is non-zero, which is convenient for computational purposes [32]. The normalized discretized the wavelet coefficients are defined by

$$\mathbb{R}^n \ni \widetilde{D}(2^j, k) := 2^{-j}\widetilde{d}_{j,k}.$$

By applying (4.2) and direct computation, the integral representation of wavelet covariance is given in the following proposition.

Proposition 4.1.1. *Let $\{Y(k)\}_{k\in\mathbb{Z}}$ be the sequence (4.1). For all $j, j' \geq 0$ and $k, k' \in \mathbb{Z}$,*

$$\mathrm{Cov}(\widetilde{D}(2^j, k), \widetilde{D}(2^{j'}, k')) = \int_{\mathbb{R}} H_j(x)\overline{H_{j'}(x)}e^{ix(2^j k - 2^{j'} k')}|x|^{-(H+I/2)}G(x)|x|^{-(H+I/2)^*}dx,$$

where

$$H_j(x) = 2^{-j}\int_{\mathbb{R}}\sum_{l\in\mathbb{Z}}\psi(2^{-j}s)\varphi(s+l)e^{-ixl}ds. \tag{4.3}$$

Let

$$\mathbb{E}\widetilde{W}(2^j) = \mathbb{E}\widetilde{D}(2^j, 0)\widetilde{D}(2^j, 0)^*, \quad \widetilde{W}(2^j) = \frac{1}{K_j}\sum_{k=1}^{K_j}\widetilde{D}(2^j, k)\widetilde{D}(2^j, k)^* \tag{4.4}$$

be the wavelet variance matrix and its sample counterpart, respectively. The next result is analogous to Proposition 3.3.1.

Proposition 4.1.2. *Let $\{Y(k)\}_{k\in\mathbb{Z}}$ be the sequence (4.1). For every pair of octaves $j, j' \geq 0$,*

(i)

$$\sqrt{K_j}\sqrt{K_{j'}}\frac{1}{K_j}\frac{1}{K_{j'}}\sum_{k=1}^{K_j}\sum_{k'=1}^{K_{j'}}\mathbb{E}\widetilde{D}(2^j, k)\widetilde{D}(2^{j'}, k')^* \otimes \mathbb{E}\widetilde{D}(2^j, k)\widetilde{D}(2^{j'}, k')^*$$

$$\to 2^{-(j+j')/2}\,\mathrm{gcd}(2^j, 2^{j'})\sum_{z=-\infty}^{\infty}\widetilde{\Phi}_{z\mathrm{gcd}(2^j,2^{j'})} \otimes \widetilde{\Phi}_{z\mathrm{gcd}(2^j,2^{j'})}, \quad \nu \to \infty, \tag{4.5}$$

where

$$\widetilde{\Phi}_z := \int_{\mathbb{R}} \overline{H_{j'}(x)} H_j(x) e^{-izx} |x|^{-(H+I/2)} G(x) |x|^{-(H+I/2)^*} dx; \tag{4.6}$$

(*ii*) *there is a matrix* $\widetilde{G}_{jj'} \in M(n(n+1)/2, \mathbb{R})$*, not necessarily symmetric, such that*

$$\sqrt{K_j}\sqrt{K_{j'}}\, \mathrm{Cov}(\mathrm{vec}_{\mathcal{S}} \widetilde{W}(2^j), \mathrm{vec}_{\mathcal{S}} \widetilde{W}(2^{j'})) \to \widetilde{G}_{jj'}, \quad \nu \to \infty, \tag{4.7}$$

where the entries of $\widetilde{G}_{jj'}$ *can be retrieved from* (4.5) *by means of* (3.23) *(see* (3.4) *on the notation* $\mathrm{vec}_{\mathcal{S}}$*).*

The following theorem establishes the asymptotic distribution of the wavelet variance matrices at fixed octaves.

Theorem 4.1.1. *Let* $\{Y(k)\}_{k\in\mathbb{Z}}$ *be the sequence* (4.1)*. Let* $j_1 < \cdots < j_m$ *be a fixed set of octaves. Then*

$$\Big(\sqrt{K_j}(\mathrm{vec}_{\mathcal{S}}(\widetilde{W}(2^j) - \mathbb{E}\widetilde{W}(2^j))\Big)^T_{j=j_1,\ldots,j_m} \stackrel{d}{\to} \mathcal{N}_{\frac{n(n+1)}{2}\times m}(\mathbf{0}, \widetilde{F}), \tag{4.8}$$

as $\nu \to \infty$ *(see* (3.4) *on the notation* $\mathrm{vec}_{\mathcal{S}}$*). In* (4.8)*, the matrix* $\widetilde{F} \in \mathcal{S}(\frac{n(n+1)}{2}m, \mathbb{R})$ *has the form* $\widetilde{F} = (\widetilde{G}_{jj'})_{j,j'=1,\ldots,m}$*, where each block* $\widetilde{G}_{jj'} \in M(n(n+1)/2, \mathbb{R})$ *is described in* (4.7)*.*

By Proposition 4.1.1, $\mathbb{E}\widetilde{W}(2^j)$ can be recast as

$$\mathbb{E}\widetilde{W}(2^j) = P\widetilde{\Lambda}_j P^*, \tag{4.9}$$

where

$$\widetilde{\Lambda}_j = \mathrm{diag}\Big(\int_{\mathbb{R}} |H_j(x)|^2 x^{-(2h_1+1)} |g_1|^2(x) dx, \cdots, \int_{\mathbb{R}} |H_j(x)|^2 x^{-(2h_n+1)} |g_n|^2(x) dx\Big). \tag{4.10}$$

Expression (4.9) indicates that an estimator $\widetilde{B}_\nu$ of P^{-1} can be estimated by jointly diagonalizing $\widetilde{W}(2^{J_1})$ and $\widetilde{W}(2^{J_2})$, for $J_1 \neq J_2$. Note that, the conclusion in Theorem 3.4.1 also holds when replacing $W(2^j)$ by $\widetilde{W}(2^j)$. Thus, as a consequence of Theorem 4.1.1 and by following the same argument as in the proof of Theorem 3.5.2, we obtain the limiting distribution of $\widetilde{B}_\nu$.

Theorem 4.1.2. *Consider two octaves $0 < J_1 < J_2$ for which the eigenvalues of $\mathbb{E}\widetilde{W}(2^{J_1})$ are pairwise distinct. Let $\widetilde{B}_\nu$ be the output of the EJD algorithm (see Section 3.5) when setting*

$$C_0 = \widetilde{W}(2^{J_1}) \quad \text{and} \quad C_1 = \widetilde{W}(2^{J_2}).$$

Then,

$$\sqrt{\nu}(\text{vec}(\widetilde{B}_\nu - \Pi\widetilde{\Lambda}_{J_1}^{-1/2}P^{-1}))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma_{\widetilde{F}}(J_1, J_2)), \tag{4.11}$$

where $\widetilde{\Lambda}_{J_1}$ is defined by (4.10), for some matrix

$$\Pi \in \{\Pi \in M(n, \mathbb{R}) : \Pi \text{ has the form diag}(\pm 1, \dots, \pm 1)\}.$$

In (4.11), the covariance matrix $\Sigma_{\widetilde{F}}(J_1, J_2)$ is a function of $\widetilde{F}$, and $\widetilde{F}$ is defined in Theorem 4.1.1 with $m = 2$.

Let the demixed process be

$$\widetilde{X} := \widetilde{B}_\nu Y = \widetilde{I}_\nu \widetilde{\mathfrak{D}} X. \tag{4.12}$$

In (4.12), $\widetilde{I}_\nu = \widetilde{B}_\nu P \widetilde{\Lambda}_{J_1}^{1/2}\Pi$, $\widetilde{\mathfrak{D}} = \Pi\widetilde{\Lambda}_{J_1}^{-1/2}$.

The following proposition gives the asymptotic distribution of the sample wavelet variance of $\widetilde{X}$.

Proposition 4.1.3. *Let the demixed process $\widetilde{X}$ is defined by (4.12), let $\widetilde{W}_{\widetilde{X}}$ be the sample wavelet variance for $\widetilde{X}$, $\mathbb{E}\widetilde{W}_X$ be the wavelet variance of the source signal X.*

Then,

$$\left(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1},\cdots,a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(\widetilde{W}_{\widetilde{X}}(a(\nu)2^j)-\widetilde{\mathfrak{D}}\mathbb{E}\widetilde{W}_X(a(\nu)2^j)\widetilde{\mathfrak{D}}))^T\right)_{j=j_1,\cdots,j_m}$$

$$\xrightarrow{d}\mathcal{N}(0,\widetilde{\mathcal{K}}\widetilde{\mathbf{W}}\widetilde{\mathcal{K}}^*), \tag{4.13}$$

as $\nu\to\infty$ *(see* (3.4) *on the notation* $\mathrm{vec}_{\mathcal{D}}$*). In* (4.13)*,* $\widetilde{\mathcal{K}}=\mathrm{diag}(\underbrace{\widetilde{\mathfrak{D}}^2,\cdots,\widetilde{\mathfrak{D}}^2}_{m})$*. The* (k_1,k_2)*-th entry of the limiting covariance matrix is*

$$\widetilde{\mathbf{W}}(k_1,k_2)=\begin{cases}\widetilde{w}_{l,v,i}, & k_1=ln+i,k_2=vn+i;\\ 0, & \text{otherwise,}\end{cases}$$

where $\widetilde{w}_{l,v,i}=4\pi(f_i^*(0))^2 2^{4d_i\max(j_l,j_v)+\min(j_l,j_v)}\int_{-\pi}^{\pi}|D_{|j_l-j_v|}(x;d_i)|^2dx$*, for* $l,v=0,\cdots,m-1$*,* $i=1,\cdots,n$*,* f_i^* *is defined in* (A.70)*,* $d_i=h_i+1/2$*,*

$$D_u(x;d)=\sum_{k\in\mathbb{Z}}|x+2k\pi|^{-2d}\mathbf{e}_u(x+2k\pi)\overline{\widehat{\psi}(x+2k\pi)}\widehat{\psi}(2^{-u}(x+2k\pi)), \tag{4.14}$$

for all $u\geq 0$*,*

$$\mathbf{e}_u(x)=2^{-u/2}(1,e^{i2^{-u}x},\cdots,e^{-i(2^u-1)2^{-u}x})^T,\quad x\in\mathbb{R}.$$

Define the Hurst parameters' estimators

$$\begin{pmatrix}\widetilde{h}_1\\ \vdots\\ \widetilde{h}_n\end{pmatrix}:=\begin{pmatrix}\sum_{j=j_1}^{j_m}w_j^1\log(\widetilde{\sigma}_1^2(a(\nu)2^j))\\ \vdots\\ \sum_{j=j_1}^{j_m}w_j^n\log(\widetilde{\sigma}_n^2(a(\nu)2^j))\end{pmatrix}, \tag{4.15}$$

where $\widetilde{\sigma}_i$ is the ith diagonal entry of the sample wavelet variance of $\widetilde{X}$, and the weight vectors $\mathbf{w}^i=(w_1^i,\cdots,w_m^i)^T$, $i=1,\cdots,n$, satisfy (3.53). We have the following CLT.

Theorem 4.1.3. *Let* $(\widetilde{h}_1,\cdots,\widetilde{h}_n)^T$ *be the estimator for* $(h_1,\cdots,h_n)^T$ *defined by*

(4.15). *Suppose* $a(\nu)$ *satisfies*

$$\frac{a(\nu)}{\nu} + \frac{\nu}{a(\nu)^{1+2\beta_*}} \to 0, \quad \nu \to \infty,$$

where β_* *is defined in* (A.71). *Then,*

$$\sqrt{\nu/a(\nu)}\left[\begin{pmatrix} \widetilde{h}_1 \\ \vdots \\ \widetilde{h}_n \end{pmatrix} - \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}\right] \xrightarrow{d} \mathcal{N}(0, \widetilde{\mathcal{W}}), \tag{4.16}$$

where

$$\widetilde{\mathcal{W}} = \mathrm{diag}((\mathbf{w}^1)^T \widetilde{V}(h_1)\mathbf{w}^1, \cdots, (\mathbf{w}^n)^T \widetilde{V}(h_n)\mathbf{w}^n),$$

the weight vectors $\mathbf{w}^i$, $i = 1, \cdots, n$ *satisfy* (3.53), $\widetilde{V}(h)$ *is a* $m \times m$ *matrix whose* (i,k)*-th entry is*

$$\widetilde{V}_{i,k}(h) = \frac{4\pi 2^{2(h+1/2)|j_i - j_k|} 2^{\min(j_i, j_k)}}{K(h)} \int_{-\pi}^{\pi} |D_{|j_i - j_k|}(x; h+1/2)|^2 dx, \quad i, k = j_1, \cdots, j_m,$$

$D_{|j_i - j_k|}(\lambda, h+1/2)$ *is defined in* (4.14), *and* $K(h) = \int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-2h-1} dx$.

4.2 Empirical application

In this section, we illustrate the two-step method with a bivariate series of annual tree ring measurements from bristlecone pine trees in White Mountains, California. The period covered ranges from 5142 BC to 1962 AD, yielding a total of 7104 data points. The data is available on the website https://datamarket. com/data/list/?q=cat:ecw%20 provider:tsdl.

Many tree-ring data sets exhibit long-range dependence properties [33], and annual tree-ring measurements can be modeled as aggregates of the underlying continuous-time growth rate process over time intervals between two consecutive sampling time points, which in turn, can be approximated by a mixed fractional process. Figure 4.1 shows the time series plot of the tree ring measurements. The comparative analysis over different time periods of wavelet scaling plots (Figure 4.2) indicate the presence of disturbance in the first 1000 data points. For this reason, we only model the data from 4141 BC to 1962 AD.

As expected, the sample ACFs (Figure 4.3) suggest that the time series of tree-ring measurements have long memory. However, it is known that spurious cross-correlation may occur as a result of the presence of fractional memory in each sequence, so it is pivotal to study the cross-correlation after pre-whitening (e.g., [34]). Figure 4.4 shows the sample cross-correlation for the pre-whitened data. It reveals that the sequences are contemporaneously strongly correlated but not cross-correlated at any non-zero lag values.

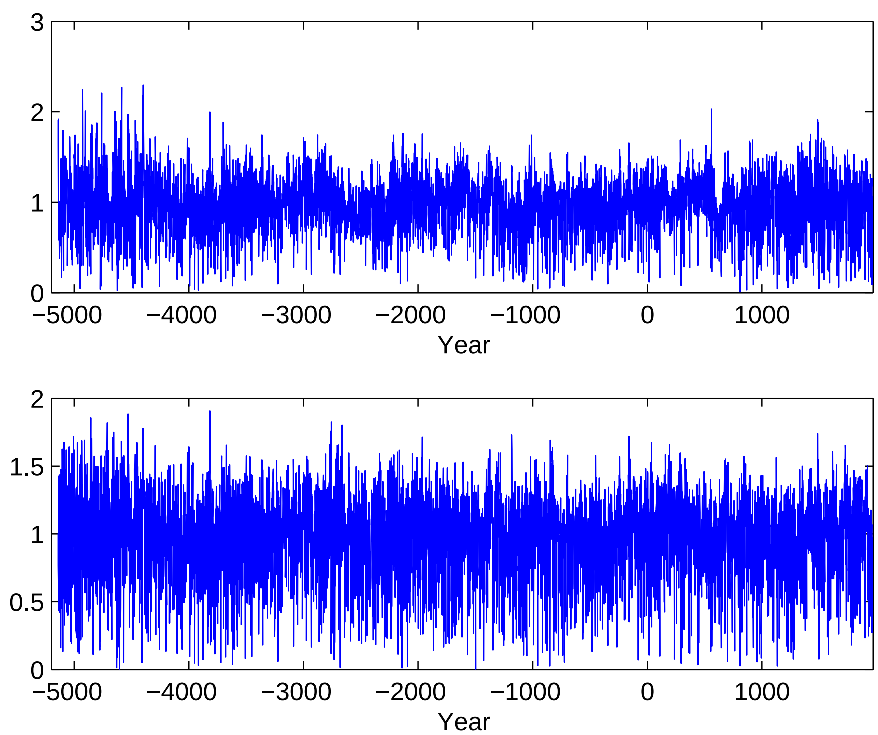

Figure 4.1: Time series plots of the tree-ring measurements.

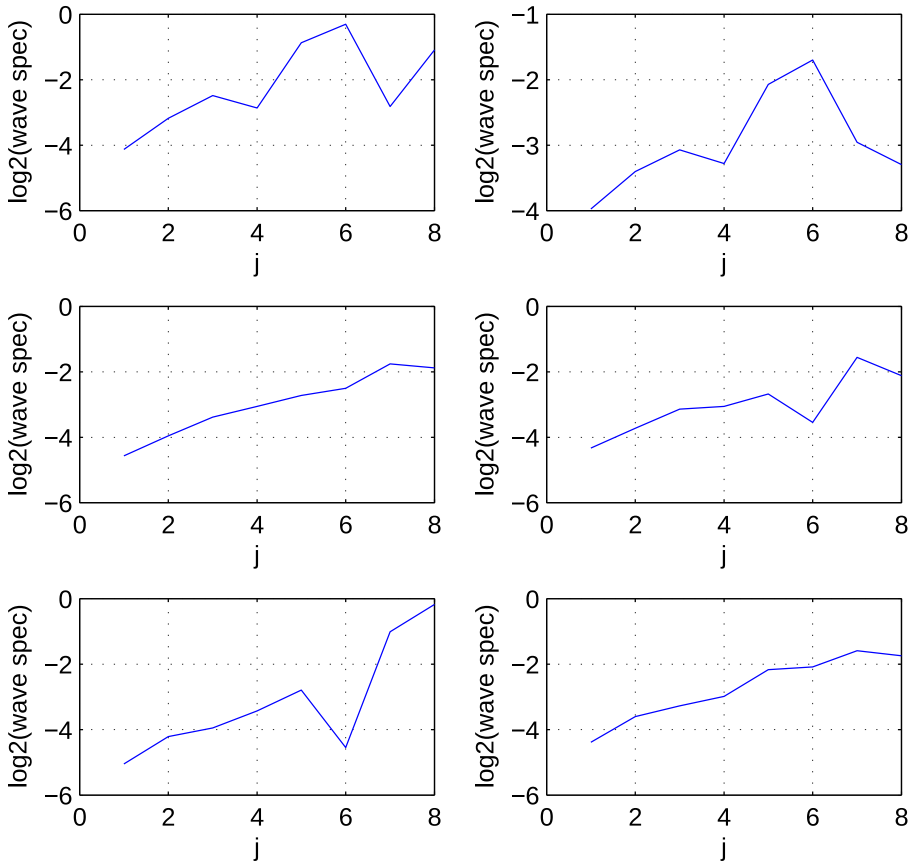

Figure 4.2: $\log_2(W(2^j))$ vs j. First row: scaling plots for the 1-1000 tree-ring measurements of the bivariate data set; Second row: scaling plots for the 1000-2000 tree-ring measurements of the bivariate data set; Third row: scaling plots for the 2000-3000 tree-ring measurements of the bivariate data set.

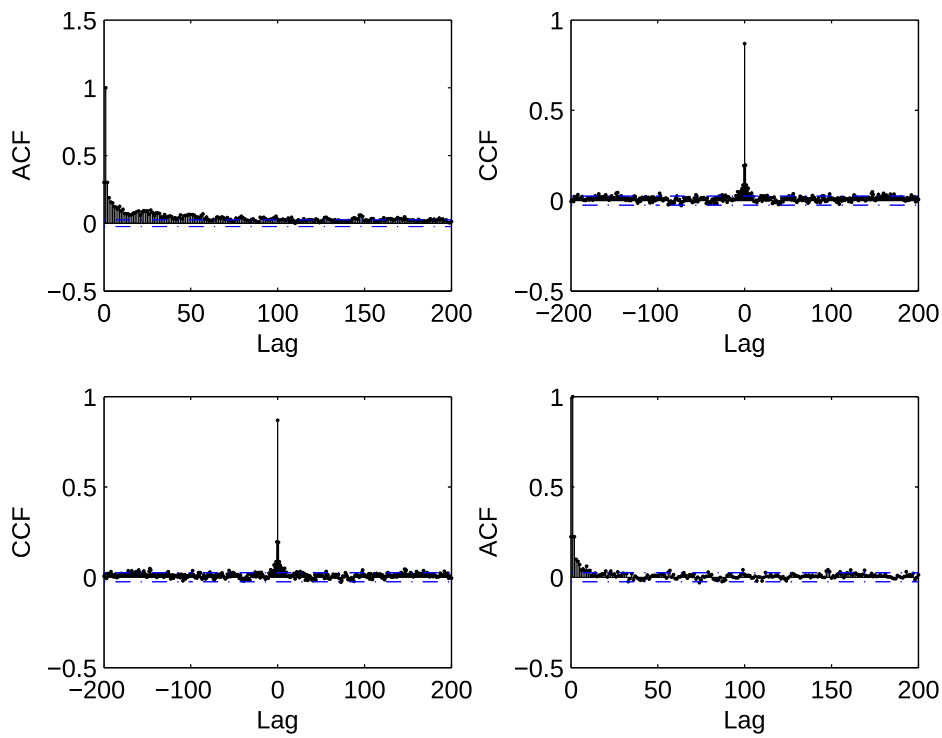

Figure 4.3: Sample auto-correlations and cross-correlations of the tree-ring measurements.

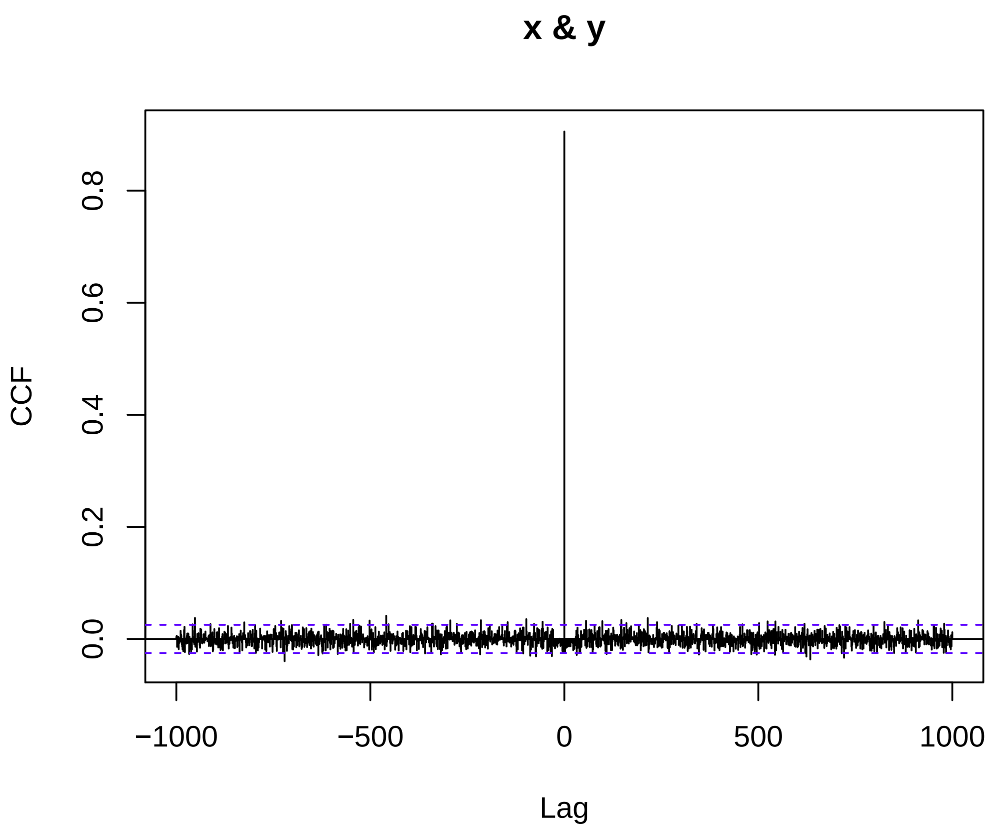

Figure 4.4: Cross-correlation between tree-ring measurements, after pre-whitening. Dash line is the threshold $\pm 1.96/\sqrt{N}$ of 5% significant level.

The wavelet based demixing method gives the estimated demixing matrix

$$\widehat{P^{-1}} = \begin{pmatrix} 10.4048 & -9.3864 \\ 1.4360 & 2.6275 \end{pmatrix}.$$

After demixing, wavelet-based estimates of the Hurst parameters The estimated are $\widehat{h}_1 = 0.6470$, $\widehat{h}_2 = 0.9295$ (using scales (j_1, j_2)=(3,7)), and $\widehat{h}_1 = 0.6517$, $\widehat{h}_2 = 0.9564$ (using scales (j_1, j_2)=(3,9)).

Table A.3 shows a Monte Carlo study of the sample mean and sample standard deviation of $\widehat{h_2 - h_1}$ for the case when $h_1 = h_2 = h$. The difference between the estimated Hurst parameters for the tree ring data is $\widehat{h}_2 - \widehat{h}_1 = 0.9564 - 0.6517 = 0.3047 > 1.645 * \text{sd}(\widehat{h_2 - h_1})$. Therefore, we conclude that there is evidence that $h_2 > h_1$. The cross-correlation of the demixed tree-ring data after pre-whitening is showed in Figure 4.5. The rare occurrence of significant spikes in the cross-correlation plot suggest that the wavelet demixing method demixed the data.

true h	parameter	mean	sd
h=0.7	$\widehat{h}_1$	0.6985	0.0183
	$\widehat{h}_2$	0.7229	0.0176
	$\widehat{h_2 - h_1}$	0.0244	0.0185
h=0.8	$\widehat{h}_1$	0.7957	0.0191
	$\widehat{h}_2$	0.8229	0.0195
	$\widehat{h_2 - h_1}$	0.0272	0.0202

Table 4.1: **wavelet estimation**: (j_1, j_2)=(3,9), sample size=6000, number of MC runs=1000.

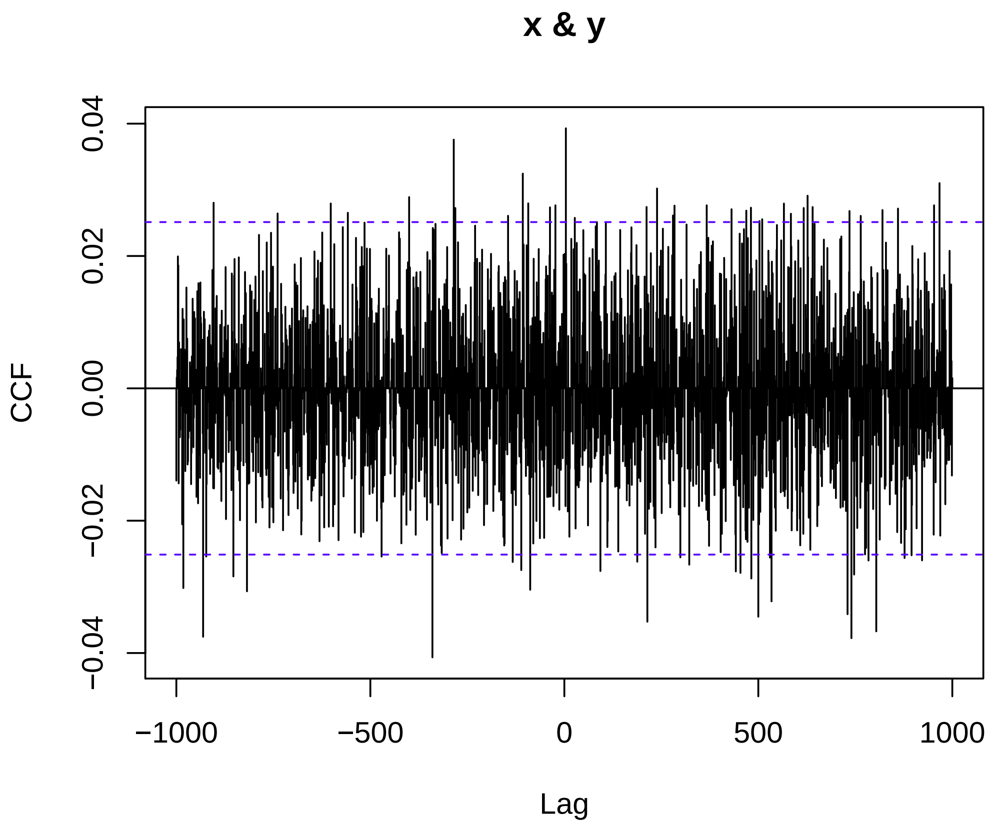

Figure 4.5: Sample cross-correlation of demixed data, after prewhitening.

Chapter 5

Two-Step Wavelet-Based Estimation for Mixed Fractional Non-Gaussian Processes

5.1 Introduction

In modern applications, it often happens that data are multivariate and characterized jointly by scale-free dynamics (self-similarity) and non-Gaussianity. The modeling of non-Gaussian, multivariate fractional signals, while of great importance in applications, is an essentially unexplored research topic, with the exception of related work in the econometric literature (e.g., [35]).

Hermite processes are typically non-Gaussian, self-similar, stationary increment processes. They appear as a consequence of non-central limit theorems, i.e., by means of nonlinear transformations of stationary long-range dependent sequences [36–38]. Fractional Brownian motion (fBm) corresponds to the Hermite process of rank 1, and is the only Gaussian such process, whereas the Hermite process of rank 2 is called Rosenblatt process (or fractional Rosenblatt motion, fRm). The integral representation for fRm is given by the following definition.

Definition 5.1.1. The standard Rosenblatt process of index $d \in (1/4, 1/2)$ is the continuous time process

$$R_{2d}(t) = a_h \int_{\mathbb{R}^2}'' \frac{e^{i(u_1+u_2)t} - 1}{i(u_1 + u_2)} |u_1|^{-d} |u_2|^{-d} \widetilde{B}(du_1)\widetilde{B}(du_2), \tag{5.1}$$

where

$$a_h = \sqrt{\frac{h(2h-1)}{2(2\Gamma(1-h)\sin(h\pi/2))^2}}, \tag{5.2}$$

$\widetilde{B}(du)$ is a $\mathbb{C}$-valued Gaussian random measure, and the symbol $\int_{\mathbb{R}^2}''$ indicates that the diagonal $u_1 = u_2$ is excluded from the integration domain.

The integral in (5.1) has finite $L^2(P)$ norm, $\mathbb{E}R_{2d}^2(1) = 1$, and the process is self-similar with Hurst parameter $h = 2d \in (1/2, 1)$. The random variable $R_{2d}(1)$ of (5.1) evaluated at $t = 1$ is called a standard Rosenblatt variable with Hurst index $2d$.
Statistical inference for Hermite-type processes is challenging: in the univariate context, it was recently shown that wavelet-based estimators may display nonstandard convergence rates and asymptotic distributions [39, 40]. Mathematically, arguments involve chaos expansions and Malliavin calculus.
In this chapter, we take the first step in the construction of statistical inference for Rosenblatt-type operator fractional stochastic processes. We consider the case where multiple independent subparts of a system are univariate Rosenblatt processes which get mixed by a number of recording sensors. In other words, the underlying process $Y(t)$ is defined by

$$\{Y(t)\}_{t\in\mathbb{R}} = \{PR(t)\}_{t\in\mathbb{R}}, \tag{5.3}$$

where P is a $n \times n$ non-singular matrix, and $R = (R_{h_1}, \cdots, R_{h_n})^T$ is a vector of independent fRms with Hurst parameters

$$1/2 < h_1 < \cdots < h_n < 1. \tag{5.4}$$

The two-step wavelet-based approach (see Chapter 3) will be applied to $Y(t)$ defined in (5.3) to estimate both the mixing matrix P and the Hurst parameters $h_1, \cdots, h_n$. We stress that, to the best of our knowledge, this is the first estimation methodology ever proposed for multivariate Hermite-type processes.

5.2 Wavelet transform of the mixed non-Gaussian fractional process

Throughout this chapter, let $\psi : \mathbb{R} \to \mathbb{R}$ be a wavelet function with number of vanishing moments $N \geq 1$, and has compact support. The vector wavelet transform of Y as in (5.3) is defined as

$$\mathbb{R}^n \ni D(2^j, k) = 2^{-j/2} \int_{\mathbb{R}} 2^{-j/2} \psi(2^{-j}t - k) Y(t) dt, \quad j \in \mathbb{N} \cup \{0\}, \quad k \in \mathbb{Z}. \tag{5.5}$$

The wavelet spectrum (variance) at scale j is the positive definite matrix

$$\mathbb{E}W(2^j) := \mathbb{E}D(2^j, k) D(2^j, k)^* = \mathbb{E}D(2^j, 0) D(2^j, 0)^*, \tag{5.6}$$

and its natural estimator, the sample wavelet variance, is the random matrix

$$W(2^j) = \frac{1}{K_j} \sum_{k=1}^{K_j} D(2^j, k) D(2^j, k)^*, \quad K_j = \frac{\nu}{2^j}, \quad j = j_1, \ldots, j_m, \tag{5.7}$$

for a total of ν available (wavelet) data points.

By direct computation, we can show the following properties of the wavelet coefficients and the general form of the wavelet variance.

Proposition 5.2.1. *Let $D(2^j, k)$ and $\mathbb{E}W(2^j)$ be defined by (5.5) and (5.6), respectively. Then,*

(P1) the wavelet transform (5.5) is well-defined in the mean square sense, and $\mathbb{E}D(2^j, k) =$

0;

(P2) for a fixed scale j, $\{D(2^j,k)\}_{k\in\mathbb{Z}}$ is stationary in k;

(P3)

$$\mathbb{E}W(2^j) = P\mathrm{diag}(2^{2jh_1}C_\psi(h_1),\cdots,2^{2jh_n}C_\psi(h_n))P^*, \tag{5.8}$$

where

$$C_\psi(h) = a_h\int_{\mathbb{R}^2}\frac{|\widehat{\psi}(u_1+u_2)|^2}{(u_1+u_2)^2}|u_1u_2|^{-h}du_1du_2, \tag{5.9}$$

and a_h is defined in (5.2).

The following theorem establishes the asymptotic distribution of the vectorized sample wavelet variance at a fixed set of octaves.

Theorem 5.2.1. *Let $Y = \{Y(t)\}_{t\in\mathbb{R}}$ be defined by (5.3). Let $j_1 < \cdots < j_m$ be a fixed set of octaves. Then,*

$$\nu^{1-h_n}\Big(\mathrm{vec}_{\mathcal{S}}(W(2^j)-\mathbb{E}W(2^j))\Big)^T_{j=j_1,\cdots,j_m} \xrightarrow{d} \mathcal{R}. \tag{5.10}$$

as $\nu\to\infty$. In (5.10),

$$\mathcal{R} \overset{d}{=} \mathrm{vec}_{\mathcal{S}}\left(P\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2^{2jh_n}C_\psi(h_n)S(h_n)R_j^{h_n}\end{pmatrix}P^*\right)^T_{j=j_1,\cdots,j_m}. \tag{5.11}$$

In (5.11),

$$S(h) = \frac{4}{h+1}\left(\frac{h}{2(2h-1)}\right)^{-1/2}\frac{\int_{\mathbb{R}^2}\psi(x)\psi(x')|x-x'|^{h+1}dxdx'}{\int_{\mathbb{R}^2}\psi(x)\psi(x')|x-x'|^{2h}dxdx'}, \tag{5.12}$$

$R_j^{h_n}$ are normalized Rosenblatt random variables with self-similarity index h_n for all $j = j_1,\cdots,j_m$, and for all $\lambda_1,\cdots,\lambda_m\in\mathbb{R}$, the k-th cumulant of the random variable

$\sum_{k=1}^{m} \lambda_k R_{j_m}^{h_n}$ is

$$\sum_{l_1,\cdots,l_k=1}^{m} \lambda_{l_1} \cdots \lambda_{l_k} c_k(R^{h_n}),$$

where $c_k(R^{h_n})$ denotes the k-th cumulant of the normalized Rosenblatt random variable with self-similarity index h_n.

5.3 Main results

Note that expression (5.8) suggests estimating P by the joint diagonalization algorithm.

Definition 5.3.1. Consider two octaves $0 \leq J_1 < J_2$ for which the eigenvalues of $\mathbb{E}W(2^{J_1})$ are pairwise distinct. For $\nu \in \mathbb{N}$, the wavelet-based demixing estimator $\widehat{B}_\nu$ is the output of the EJD algorithm when setting

$$C_0 = W(2^{J_1}) \text{ and } C_1 = W(2^{J_2}). \tag{5.13}$$

By using Theorem 5.2.1, Theorem A.6.1 and the Delta method, and then by following the same argument as in the proof of Theorem 3.5.2, we can establish the limiting distribution of $\widehat{B}_\nu$.

Theorem 5.3.2. *Let $\widehat{B}_\nu$ be as in Definition 5.3.1, then*

$$\nu^{1-h_n}(\widehat{B}_\nu - \mathfrak{D}P^{-1}) \xrightarrow{d} \mathfrak{R}, \quad \nu \to \infty. \tag{5.14}$$

In (5.14), the finite variance random matrix $\mathfrak{R}$ is a function of $\mathcal{R}$ in (5.11), and

$$\mathfrak{D} = \Pi(\text{diag}(2^{2J_1h_1}C_\psi(h_1), \cdots, 2^{2J_1h_n}C_\psi(h_n)))^{-1/2}, \tag{5.15}$$

for some matrix $\Pi \in \mathcal{I}$, *where*

$$\mathcal{I} = \{\Pi \in M(n, \mathbb{R}) : \Pi \text{ has the form } \mathrm{diag}(\pm 1, \ldots, \pm 1)\}. \tag{5.16}$$

In (5.15), $C_\psi(h)$ *is defined in* (5.9).

Define $\widehat{I}_\nu := \widehat{B}_\nu P \mathfrak{D}^{-1}$, then $\widehat{I}_\nu - I = O_p(\nu^{h_n - 1})$. The demixed process is defined by

$$\widehat{R} := \widehat{B}_\nu Y = \widehat{B}_\nu P \mathfrak{D}^{-1} \mathfrak{D} R = \widehat{I}_\nu \mathfrak{D} R. \tag{5.17}$$

The following proposition gives the asymptotic distribution of the sample wavelet variance of $\widehat{R}$.

Proposition 5.3.1. *Let* $\widehat{R}$ *be the demixed process defined by* (5.17), *let*

$$W_{\widehat{R}} \text{ be the sample wavelet variance of } \widehat{R}, \tag{5.18}$$

and let

$$\mathbb{E} W_R \text{ be the wavelet variance of the source signal } R. \tag{5.19}$$

Suppose the scaling factor $a(\nu)$ *satisfies*

$$\lim_{\nu \to \infty} \max_{l=1,\cdots,n} \left(\frac{\nu^{h_n - h_l}}{a(\nu)^{1-h_l}} \right) = 0. \tag{5.20}$$

Then,

$$\Big(\mathrm{diag}((\nu/a(\nu))^{1-h_1} a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n} a(\nu)^{-2h_n})$$
$$(\mathrm{vec}_{\mathcal{D}}(W_{\widehat{R}}(a(\nu)2^j) - \mathfrak{D} \mathbb{E} W_R(a(\nu)2^j) \mathfrak{D})^T \Big)_{j=j_1,\cdots,j_m} \xrightarrow{d} \mathcal{K}\mathcal{G}, \tag{5.21}$$

as $\nu \to \infty$ *(see* (3.4) *on the notation* $\mathrm{vec}_{\mathcal{D}}$*). In* (5.21),

$$\mathcal{K} = \mathrm{diag}(\underbrace{\mathfrak{D}^2, \cdots, \mathfrak{D}^2}_{m}), \tag{5.22}$$

$$\mathcal{G} = (2^{2j_1 h_1} C_\psi(h_1) S(h_1) R_{j_1}^{h_1}, \cdots, 2^{2j_1 h_n} C_\psi(h_n) S(h_n) R_{j_1}^{h_n},$$

$$\cdots, 2^{2j_m h_1} C_\psi(h_1) S(h_1) R_{j_m}^{h_1}, \cdots, 2^{2j_m h_n} C_\psi(h_n) S(h_n) R_{j_m}^{h_n})^T, \tag{5.23}$$

In (5.23), $R_{j_{k_1}}^{h_i}$ is independent of $R_{j_{k_2}}^{h_l}$, for $i \neq l$, $k_1, k_2 = 1, \cdots, m$. Moreover, for any fixed $l = 1, \cdots, n$, $R_j^{h_l}$ are normalized Rosenblatt random variables with self-similarity index h_l for all $j = j_1, \cdots, j_m$, and for all $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$, the k-th cumulant of the random variable $\sum_{k=1}^m \lambda_k R_{j_m}^{h_l}$ is

$$\sum_{l_1, \cdots, l_k = 1}^{m} \lambda_{l_1} \cdots \lambda_{l_k} c_k(R^{h_l}),$$

where $c_k(R^{h_l})$ denotes the k-th cumulant of the Rosenblatt random variable with self-similarity index h_l, $l = 1, \cdots, n$.

Let $W_{\widehat{R}}$ and $\mathbb{E}W_R$ be as in (5.18) and (5.19), respectively. Let

$$\widehat{\sigma}^2_{\widehat{R}_i}(2^j) \text{be the } (i,i)\text{-th entry of } W_{\widehat{R}}(2^j), \tag{5.24}$$

and let

$$\sigma^2_{R_i}(2^j) \text{be the } (i,i)\text{-th entry of } \mathbb{E}W_R(2^j). \tag{5.25}$$

Then, by direct computation, $\sigma^2_{R_i}(2^j) = 2^{2jh_i} C_\psi(h_i)$, $i = 1, \cdots, n$, where $C_\psi(h)$ is defined in (5.9). The wavelet regression estimator of the Hurst parameters h_i, $i = 1, \cdots, n$, involves regressing the terms $\widehat{\sigma}_{\widehat{R}_i}(a(\nu)2^j)$ on the scale index j, i.e.,

$$\begin{pmatrix} \widehat{h}_1 \\ \vdots \\ \widehat{h}_n \end{pmatrix} := \begin{pmatrix} \sum_{j=j_1}^{j_m} w_j^1 \log \widehat{\sigma}^2_{\widehat{R}_1}(a(\nu)2^j) \\ \vdots \\ \sum_{j=j_1}^{j_m} w_j^n \log \widehat{\sigma}^2_{\widehat{R}_n}(a(\nu)2^j) \end{pmatrix}, \tag{5.26}$$

where the weight vectors

$$\mathbf{w}^i = (w_1^i, \cdots, w_m^i)^T \tag{5.27}$$

satisfy

$$\sum_{j=1}^{m} w_j^i = 0, \quad 2\log 2 \sum_{j=1}^{m} j w_j^i = 1, \quad i = 1, \cdots, n. \tag{5.28}$$

Then, the following theorem develops the limiting distribution of the estimators $\widehat{h}_1, \cdots, \widehat{h}_n$.

Theorem 5.3.3. *Under the assumptions of Proposition 5.3.1, let estimator* $(\widehat{h}_1, \cdots, \widehat{h}_n)$ *be defined by* (5.26)*. Then,*

$$\mathrm{diag}((\nu/a(\nu))^{1-h_1}, \cdots, (\nu/a(\nu))^{1-h_n}) \left[\begin{pmatrix} \widehat{h}_1 \\ \vdots \\ \widehat{h}_n \end{pmatrix} - \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \right] \xrightarrow{d} \begin{pmatrix} L_{h_1} \\ \vdots \\ L_{h_n} \end{pmatrix}, \tag{5.29}$$

In (5.29)*,* $L_{h_i} = S(h_i) \sum_{k=1}^{m} w_k^i R_{jk}^{h_i}$ *are independent random variables depends on* h_i*,* $i = 1, \cdots, n$*, and* $S(h)$ *is defined in* (5.12)*. The random variables* $R_{jk}^{h_i}$ *are defined in Proposition 5.3.1,* $k = 1, \cdots, m$*,* $i = 1 \cdots, n$*.*

Remark 5.3.4. Even though the underlying process is exactly operator self-similar, we still need to set the scale $a((\nu))$ goes to infinity when estimating the Hurst parameters. By doing so, the bias introduced by demixing in the first step converges to zero when multiplying the convergence rate in (5.29).

5.4 Simulation studies

To study the performance of the wavelet demixing method over finite samples, we conducted Monte Carlo experiments as follows. For $n = 2$ and sample (path) size $\nu = 2^{14}$, we simulated 500 realizations by mixing 2 independent fRms with Hurst parameters 0.6 and 0.9, and the mixed process was given by $Y = PR$. For notational simplicity, let R, Y and $\widehat{R}$ denote the original, mixed and demixed processes. For the analysis, Daubechies least asymmetric wavelet with $N_\psi = 2$ vanishing moment were used. We set the octaves $(J_1, J_2) = (1, 2)$ for the demixing, and $(j_1, j_2) = (3, \log_2 \nu - N_\psi)$ for the Hurst parameter estimation, respectively.

The univariate log-wavelet variance $W_{R_i}(2^j)$, $W_{Y_i}(2^j)$ and $W_{\widehat{R}_i}(2^j)$, as well as the univariate wavelet regression estimator $\widehat{h}_{R_i}$, $\widehat{h}_{Y_i}$ and $\widehat{h}_{\widehat{R}_i}$ were computed for each component of R, Y and $\widehat{R}$, respectively. The results are reported for one arbitrarily chosen matrix P, since the performance for all instances of P was comparable.

We plot the log-averages $\log_2\langle W_{R_i}(2^j)\rangle$, $\log_2\langle W_{Y_i}(2^j)\rangle$ and $\log_2\langle W_{\widehat{R}_i}(2^j)\rangle$ for each of the $n = 2$ respective components, where $\langle\cdot\rangle$ denotes the Monte Carlo average (Figure 5.1), as well as the distributions (boxplots) of $\widehat{h}_{R_i} - h_i$, $\widehat{h}_{Y_i} - h_i$ $\widehat{h}_{\widehat{R}_i} - h_i$ (Figure 5.2). In Figure 5.1, as expected all components of $\log_2\langle W_{Y_i}(2^j)\rangle$ diverge from their unmixed original data counterparts $\log_2\langle W_{R_i}(2^j)\rangle$. After demixing, all components of $\log_2\langle W_{\widehat{R}_i}(2^j)\rangle$ remarkably superimpose those of $\log_2\langle W_{R_i}(2^j)\rangle$. The boxplots in Figure 5.2 show that the Monte Carlo distributions for $\widehat{h}_{\widehat{R}_i} - h_i$ resemble those of $\widehat{h}_{R_i} - h_i$, which illustrates the successful dimixing of Y.

To show the versatility of the 2-step demixing method, the boxplot in Figure 5.3 shows the successful estimation of Y in the situation where the 4-dimensional process Y is obtained by mixing 4 independent self-similar processes (2 fBms with Hurst parameters 0.2 and 0.4 and 2 fRms with Hurst parameters 0.6 and 0.8). Figure 5.4 shows that the variances of the Hurst parameter estimators decrease at the expected rate. Indeed, for fBm entries, the decrease is proportional to $1/\nu$, whereas for fRm entries, the decrease is h_i-dependent and proportional to $1/\nu^{2-2h_i}$.

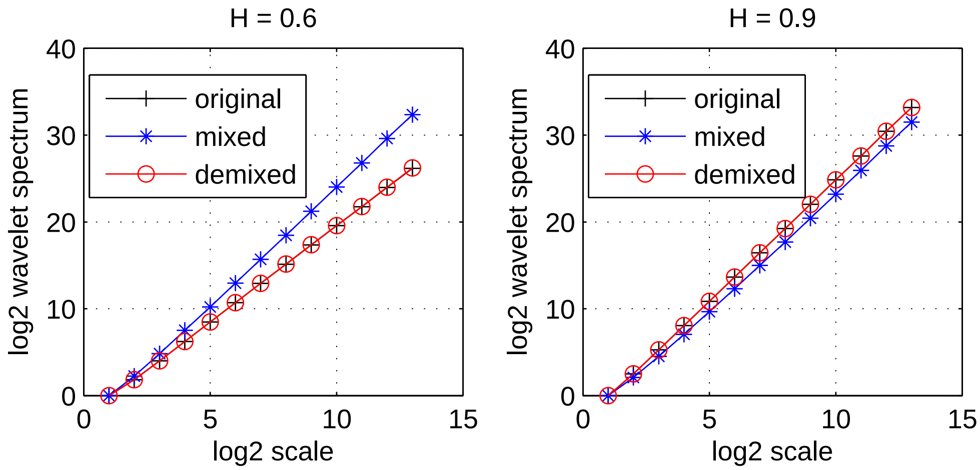

Figure 5.1: $\log W_{\cdot}(2^j)$ **vs.** j superimposed for processes R (original, black '+'), Y (mixed, blue '$*$') and $\widehat{R}$ (demixed, red 'o'), for each of the $n = 2$ components, sorted by ascending order of h_i.

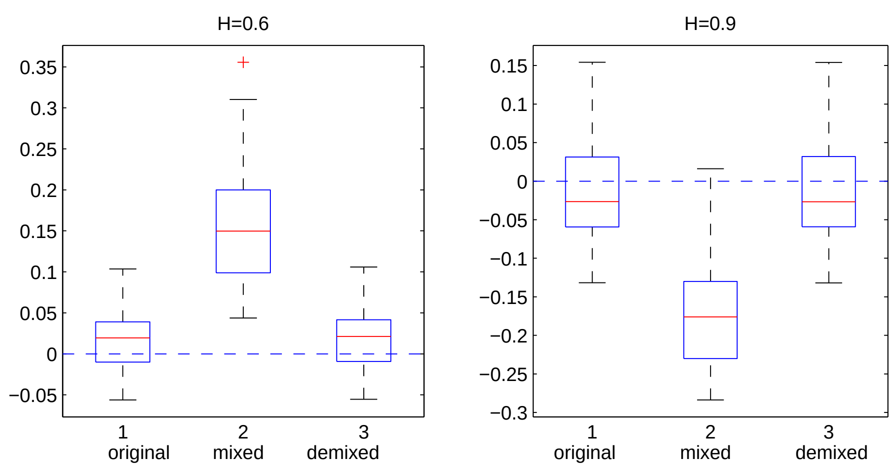

Figure 5.2: **Boxplots** $\hat{h}_i - h_i$ obtained from R (original, left), Y (mixed, middle) and $\widehat{R}$ (demixed, right), for each of the $n = 2$ components, sorted by ascending order of h_i.

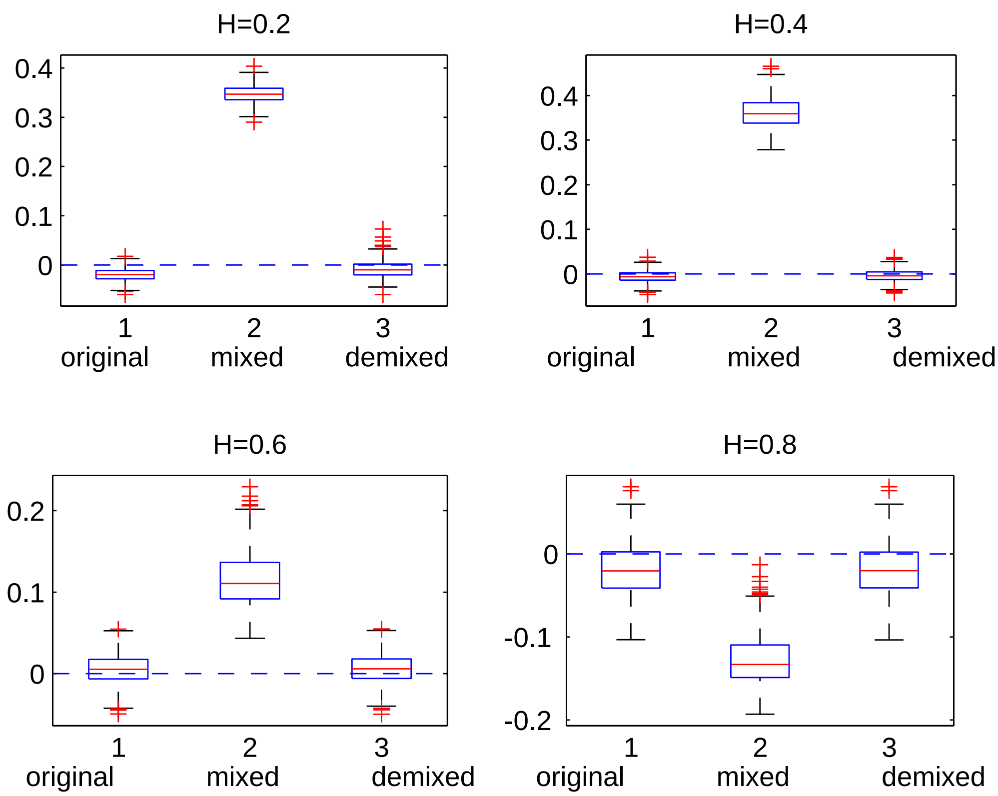

Figure 5.3: **Boxplots** $\hat{h}_i - h_i$ obtained from R (original, left), Y (mixed, middle) and $\widehat{R}$ (demixed, right), for each of the $n = 4$ components, sorted by ascending order of h_i.

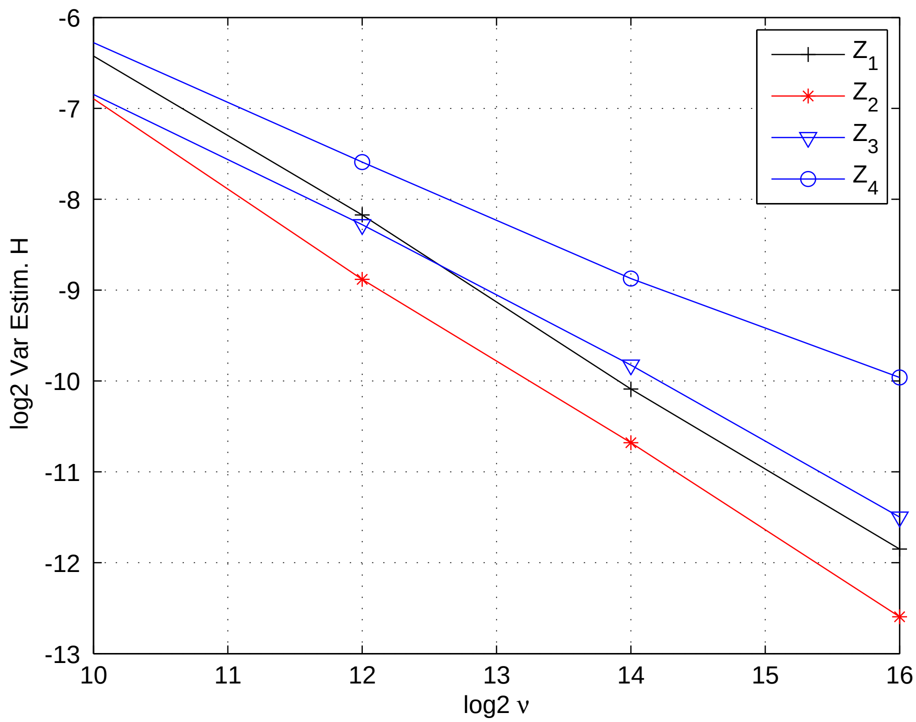

Figure 5.4: **Decrease of the estimator variance $\widehat{\text{Var}\hat{h}_l}$ as a function of the sample size ν (fBm with $h_1 = 0.2$ and $h_2 = 0.4$; fRm with $h_3 = 0.6$ and $h_4 = 0.8$.** The slopes for the three segments of $\widehat{R}_1$ (black, +) and $\widehat{R}_2$ (red, $*$) are, respectively, -0.8744, -0.9568, -0.8808, and -0.9925, -0.8987, -0.9577, reflecting the theoretical value -1. The slopes for the three segments of $\widehat{R}_3$ (blue, $\triangledown$) and $\widehat{R}_4$ (blue, o), respectively, are -0.7161, -0.7720, -0.8367, and -0.6572, -0.6410, -0.5432, indicating convergence to the theoretical values -0.8 and -0.4, respectively.

Appendix A

Appendix

A.1 Repeated eigenvalues

Following up on the discussion in Remark 3.4.2, the next proposition describes the limiting distribution for the eigenvalues of $W(2^j)$ for a special case where $\mathbb{E}W(2^j)$ has one repeated eigenvalue. In its statement, we use the multivariate gamma function $\Gamma_q(\cdot)$, which is defined by

$$\Gamma_q(t) = \pi^{q(q-1)/4} \prod_{i=1}^{q} \Big(t - \frac{1}{2}(i-1)\Big).$$

Moreover, we replace (A1) with the following assumption.

Assumption (A1′): the observed signal $Y = \{Y(t)\}_{t \in \mathbb{R}}$ has the form (3.1) with

$$h_1 = h_2 = \ldots = h_n =: h, \quad N-1 < h < N, \tag{A.1}$$

and the high frequency functions $g_i(x)$, $i = 1, \cdots, n$ are constants, i.e.,

$$g_1(x) = g_1, \cdots, g_n(x) = g_n.$$

Proposition A.1.1. *Suppose the assumptions (A1′–A2) hold. Let*

$$\mathbb{E}W(2^j) = O\Lambda O^*, \quad W(2^j) = \widehat{O}L\widehat{O}^*, \tag{A.2}$$

be the matrix spectral decompositions of the wavelet and sample wavelet variance matrices, respectively. Assume the diagonal matrix Λ *has the form*

$$\Lambda = \begin{pmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \lambda_* I_q \end{pmatrix} \tag{A.3}$$

for some $1 < q < n$*, where the main diagonal entries of the matrix* Λ_1 *are pairwise distinct and less than* λ_**. Let*

$$L = \text{diag}(l_1, \dots, l_n), \quad \Lambda_1 = \text{diag}(\lambda_1, \dots, \lambda_{n-q}). \tag{A.4}$$

Then, as $\nu \to \infty$*,*

$$\sqrt{K_j}\big((l_1 - \lambda_1, \dots, l_{n-q} - \lambda_{n-q}), (l_{n-q+1} - \lambda_*, \dots, l_n - \lambda_*)\big)^T \stackrel{d}{\to} (\mathcal{L}_1^T, \mathcal{L}_2^T)^T, \tag{A.5}$$

where $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *are independent random vectors. Moreover,*

$$\mathcal{L}_1 \sim \mathcal{N}(0, 2b\,\text{diag}(\lambda_1^2, \dots, \lambda_{n-q}^2)), \tag{A.6}$$

where

$$b := 2^{-j} \sum_{z=-\infty}^{\infty} \Big\{ \int_{\mathbb{R}} |\widehat{\psi}(2^j x)|^2 e^{-izx} |x|^{-1-h} dx \Big/ \int_{\mathbb{R}} |\widehat{\psi}(2^j x)|^2 |x|^{-1-h} dx \Big\}^2, \tag{A.7}$$

and $\mathcal{L}_2$ has density

$$2^{-\frac{1}{2}q}(\sqrt{b}\lambda_*\pi)^{q(q-1)/4}\Gamma_q^{-\frac{1}{2}}\Big(\frac{q}{2}\Big)\exp\Big\{-\frac{1}{2\sqrt{b}\lambda_*}\sum_{i=n-q+1}^{n} d_i^2\Big\}\prod_{l<i}(d_i-d_l), \tag{A.8}$$

where

$$d_i = l_i - \lambda_*, \quad i = n-q+1,\dots,n. \tag{A.9}$$

Proof. Let O and $\widehat{O}$ be as in expression (A.2), and define

$$T = O^*W(2^j)O, \quad U = \sqrt{K_j}(T-\Lambda), \tag{A.10}$$

where O is the orthogonal matrix in the expression (A.2). Then, we can write

$$T = YLY^*, \quad Y = O^*\widehat{O} \in O(n), \tag{A.11}$$

and thus

$$U = O^*\sqrt{K_j}(W(2^j) - \mathbb{E}W(2^j))O. \tag{A.12}$$

Let h be as in (A.1). From (3.21), we obtain

$$\Lambda = O^*P\text{diag}(g_1^2,\dots,g_n^2)P^*O2^{2jh}\int_{\mathbb{R}}|\widehat{\psi}(2^jx)|^2|x|^{-1-h}dx.$$

For $z \in \mathbb{Z}$, let Φ_z be as in (3.25) (for $j = j'$). Under the condition (A.1),

$$O^*\Phi_zO = O^*P\text{diag}(g_1^2,\cdots,g_n^2)P^*O\frac{1}{2}\int_{\mathbb{R}}|\widehat{\psi}(2^jx)|^2|x|^{-1-h}dx$$

$$= \Lambda\Big\{\int_{\mathbb{R}}|\widehat{\psi}(2^jx)|^2e^{-izx}|x|^{-1-h}dx\Big/\int_{\mathbb{R}}|\widehat{\psi}(2^jx)|^2|x|^{-1-h}dx\Big\}.$$

By (3.24) (which also holds under (A.1)),

$$\sqrt{K_j}\sqrt{K_j}\frac{1}{K_j}\frac{1}{K_j}\sum_{k=1}^{K_j}\sum_{k'=1}^{K_j} O^*\mathbb{E}D(2^j,k)D(2^j,k')^*O \otimes O^*\mathbb{E}D(2^j,k)D(2^j,k')^*O$$

$$\to 2^{-j}\sum_{z=-\infty}^{\infty} O^*\Phi_{z2^j}O \otimes O^*\Phi_{z2^j}O = b(\Lambda\otimes\Lambda), \quad \nu\to\infty, \tag{A.13}$$

where the scalar b is given by (A.7). Thus, from (A.12),

$$U \overset{d}{\to} \mathcal{U} = \{u_{i_1 i_2}\}_{i_1,i_2=1,\dots,n}, \tag{A.14}$$

where $(\mathrm{vec}_{\mathcal{S}}(\mathcal{U}))^T \sim \mathcal{N}_{\frac{n(n+1)}{2}}(\mathbf{0},\Omega)$ and Ω can be retrieved from (A.13) by means of (3.23). In particular, all entries of $(\mathrm{vec}_{\mathcal{S}}(\mathcal{U}))^T$ are independent. Moreover, for $\lambda_\bullet$ as in (A.4),

$$\mathrm{Var}(u_{i_1 i_1}) = 2b\,\lambda_{i_1}^2, \quad \mathrm{Var}(u_{i_1 i_2}) = b\,\lambda_{i_1}\lambda_{i_2}, \quad 1\le i_1,i_2\le n-q, \tag{A.15}$$

$$\mathrm{Var}(u_{i_1 i_1}) = 2b\,\lambda_*^2, \quad \mathrm{Var}(u_{i_1 i_2}) = b\,\lambda_*^2, \quad n-q+1\le i_1,i_2\le n \tag{A.16}$$

(the remaining entries of $\mathcal{U}$ will not play a role in the ensuing development). It now suffices to follow the same arguments as in Sections 13.5.1 and 13.5.2 of Anderson [41]. For the reader's convenience, we lay out the main steps. Recast the random matrices T, Y, U and L in (A.10) and (A.11) as

$$T = \begin{pmatrix} T_{11} & T_{12}\\ T_{21} & T_{22}\end{pmatrix}, \quad Y = \begin{pmatrix} Y_{11} & Y_{12}\\ Y_{21} & Y_{22}\end{pmatrix}, \quad U = \begin{pmatrix} U_{11} & U_{12}\\ U_{21} & U_{22}\end{pmatrix}, \quad L = \mathrm{diag}(L_1,L_2), \tag{A.17}$$

where $T_{11}, Y_{11}, U_{11}, L_1 \in M(n-q,\mathbb{R})$, and let

$$D = \sqrt{K_j}(L-\Lambda) = \mathrm{diag}(D_1,D_2).$$

Define

$$Y_{22} = EJF, \quad C_2 = EF \in O(q), \tag{A.18}$$

where the first relation is a singular value decomposition, J is diagonal and $E, F \in O(q)$ are orthogonal. Also let

$$W_{11} = \sqrt{K_j}(Y_{11} - I), \quad W_{12} = \sqrt{K_j}Y_{12}, \quad W_{21} = \sqrt{K_j}Y_{21}, \quad W_{22} = \sqrt{K_j}(Y_{22} - C_2). \tag{A.19}$$

Based on (A.17) and (A.19), we can reexpress the system of equalities $T = \Lambda + \frac{1}{\sqrt{K_j}}U = YLY^*$ as

$$T = \begin{pmatrix} \Lambda_1 & \\ & \lambda_* I_q \end{pmatrix} + \frac{1}{\sqrt{K_j}} \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} = \left[\begin{pmatrix} I_{n-q} & \\ & C_2 \end{pmatrix} + \frac{1}{\sqrt{K_j}} \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \right]$$

$$\cdot \left[\begin{pmatrix} \Lambda_1 & \\ & \lambda_* I_q \end{pmatrix} + \frac{1}{\sqrt{K_j}} \begin{pmatrix} D_1 & \\ & D_2 \end{pmatrix} \right] \cdot \left[\begin{pmatrix} I_{n-q} & \\ & C_2^* \end{pmatrix} + \frac{1}{\sqrt{K_j}} \begin{pmatrix} W_{11}^* & W_{21}^* \\ W_{12}^* & W_{22}^* \end{pmatrix} \right]$$

$$= \begin{pmatrix} \Lambda_1 & \\ & \lambda_* I_q \end{pmatrix} + \frac{1}{\sqrt{K_j}} \left[\begin{pmatrix} D_1 & \\ & C_2 D_2 C_2^* \end{pmatrix} + \begin{pmatrix} W_{11}\Lambda_1 & \lambda_* W_{12} C_2^* \\ W_{21}\Lambda_1 & \lambda_* W_{22} C_2^* \end{pmatrix} \right.$$

$$\left. + \begin{pmatrix} \Lambda_1 W_{11}^* & \Lambda_1 W_{21}^* \\ \lambda_* C_2 W_{12}^* & \lambda_* C_2 W_{22}^* \end{pmatrix} \right] + O_P\left(\frac{1}{K_j}\right). \tag{A.20}$$

On the other hand, $I = YY^*$ and the relations (A.19) yield

$$I_n = \begin{pmatrix} I_{n-q} & \\ & I_q \end{pmatrix} + \frac{1}{\sqrt{K_j}} \left[\begin{pmatrix} W_{11} & W_{12}C_2^* \\ W_{21} & W_{22}C_2^* \end{pmatrix} + \begin{pmatrix} W_{11}^* & W_{21}^* \\ C_2 W_{12}^* & C_2 W_{22}^* \end{pmatrix} \right] + O_P\left(\frac{1}{K_j}\right). \tag{A.21}$$

From (A.20) and (A.21), we obtain the system of equations

$$U_{11} = W_{11}\Lambda_1 + D_1 + \Lambda_1 W_{11}^* + O_P\left(\frac{1}{\sqrt{K_j}}\right), \quad \mathbf{0} = W_{11} + W_{11}^* + O_P\left(\frac{1}{\sqrt{K_j}}\right), \tag{A.22}$$

$$U_{22} = C_2 D_2 C_2^* + O_P\left(\frac{1}{\sqrt{K_j}}\right). \tag{A.23}$$

Recall that the limiting joint distribution of (U_{11}, U_{22}) is given by $\mathcal{U}_{11} := \{u_{i_1 i_2}\}_{i_1,i_2=1,\dots,n-q}$ and $\mathcal{U}_{22} := \{u_{i_1 i_2}\}_{i_1,i_2=n-q+1,\dots,n}$ from expression (A.14), where

$$\mathcal{U}_{11} \text{ and } \mathcal{U}_{22} \text{ are independent.} \tag{A.24}$$

By following the same argument as on pp. 546 and 547 in Anderson [41], expressions (A.22) can be used to show that the limiting distribution of the diagonal entries of D_1 is (A.6). Next note that D_2 and Y_{22} are functions of U depending on ν (see (A.10) and (A.11)), and C_2, in turn, is a function of Y_{22} depending on ν (see (A.18)). Therefore, by the same argument as in Anderson [41], p. 549, the limiting distribution of D_2 and C_2 is the distribution of $\mathcal{D}_2$ and $\mathcal{Y}_{22}$ defined by the expression

$$\mathcal{U}_{22} = \mathcal{Y}_{22}\mathcal{D}_2\mathcal{Y}_{22}^*.$$

In particular, the limiting distribution of the diagonal entries of D_2 is (A.8). In view of (A.24), the established limiting distributions for the diagonal entries of D_1 and D_2 yield (A.5). □

Example A.1.1. For $n = 3$, consider the OFBM for which $h_1 = h_2 = h_3 =: h$, $P \in O(3)$, and $0 < g_1 < g_2 = g_3$. Then, by (3.21), the eigenvalues of $\mathbb{E}(2^j)$ are $\lambda_1 = 2^{2jh} g_1^2 \int_{\mathbb{R}} |\widehat{\psi}(y)|^2 |y|^{-2(h+1/2)} dy < \lambda_* = 2^{2jh} g_2^2 \int_{\mathbb{R}} |\widehat{\psi}(y)|^2 |y|^{-2(h+1/2)} dy$, where the latter has multiplicity 2. Now let $l_1 \le l_2 \le l_3$ be the ordered eigenvalues of the sample wavelet variance $W(2^j)$ (cf. (3.29)). Then, by Proposition A.1.1,

$$\sqrt{K_j}\big(l_1-\lambda_1, l_2-\lambda_*, l_3-\lambda_*\big)^T \xrightarrow{d} (\mathcal{L}_1^T, \mathcal{L}_2^T)^T, \quad \nu\to\infty. \tag{A.25}$$

In (A.25), $\mathcal{L}_1$ is independent of $\mathcal{L}_2$, $\mathcal{L}_1 \sim \mathcal{N}(0, 2b\lambda_1^2)$ and $\mathcal{L}_2$ has density

$$\frac{1}{2}(\sqrt{b}\lambda_*\pi)^{1/2}\Gamma_2^{-\frac{1}{2}}(1)\exp\Big\{-\frac{1}{2\sqrt{b}\lambda_*}(d_2^2+d_3^2)\Big\}(d_3-d_2),$$

where $d_i = l_i - \lambda_*$, $i = 2, 3$, and b is given by (A.7).

A.2 Central limit theorem for the wavelet variance of univariate Gaussian fractional processes

In this section, we will establish the asymptotic normality of the wavelet variance of univariate Gaussian fractional processes. All through the appendix, we assume the wavelet basis $\psi \in L^2(\mathbb{R})$ satisfies the condition (W1-3). Suppose a real-valued process $\{X(t)\}_{t\in\mathbb{R}}$ has the form

$$\{X(t)\}_{t\in\mathbb{R}} \stackrel{d}{=} \int_{\mathbb{R}} \frac{e^{itx} - \sum_{k=0}^{N-1}\frac{1}{k!}(itx)^k}{(ix)^N}|x|^{-(h-(N-1/2))}g(x)\widetilde{B}(dx), \tag{A.26}$$

where $\widetilde{B}(dx)$ is a $\mathbb{C}$-valued Gaussian random measure satisfying $\widetilde{B}(-dx) = \overline{\widetilde{B}(dx)}$. In (A.26), $N-1 < h < N$, and the high frequency function g satisfies condition (A3).

The wavelet transform of X in (A.26) is defined as

$$d(2^j,k) = 2^{-j/2}\int_{\mathbb{R}} 2^{-j/2}\psi(2^{-j}t-k)X(t)dt, \quad j\in\mathbb{N}\cup\{0\}, \quad k\in\mathbb{Z}.$$

The wavelet variance at scale j is

$$\sigma^2(2^j) =: \mathbb{E}d(2^j,0)^2, \tag{A.27}$$

and its natural estimator, the sample wavelet variance, is

$$\widehat{\sigma}^2(2^j) =: \frac{1}{K_j}\sum_{k=0}^{K_j} d^2(2^j,k), \quad K_j = \nu/2^j, \quad j = j_1,\cdots,j_m, \tag{A.28}$$

for a total number of ν available data points.

Throughout this section, we assume a sequence $\{a(\nu)\}$ satisfying (3.46).

The following lemma will be used in the subsequent proposition.

Lemma A.2.1. *For any fix two octaves $j, j' > 0$,*

$$\lim_{\nu\to\infty} a(\nu)^{-4h}\int_{\mathbb{R}} x^{-(4h+2)}|\widehat{\psi}(a(\nu)2^{j'}x)|^2|\widehat{\psi}(a(\nu)2^j x)|^2||g(x)|^2-|g(0)|^2|^2 dx = 0.$$

Proof.

$$a(\nu)^{-4h}\int_{\mathbb{R}} x^{-(4h+2)}|\widehat{\psi}(a(\nu)2^{j}x)|^2|\widehat{\psi}(a(\nu)2^{j'} x)|^2||g(x)|^2-|g(0)|^2|^2 dx$$

$$= a^{-4h}\int_{|x|\le\pi} x^{-(4h+2)}|\widehat{\psi}(a(\nu)2^{j}x)|^2|\widehat{\psi}(a(\nu)2^{j'} x)|^2||g(x)|^2-|g(0)|^2|^2 dx$$

$$+ a^{-4h}\int_{|x|>\pi} x^{-(4h+2)}|\widehat{\psi}(a(\nu)2^{j}x)|^2|\widehat{\psi}(a(\nu)2^{j'} x)|^2||g(x)|^2-|g(0)|^2|^2 dx.$$

By (3.7),

$$a(\nu)^{-4h}\int_{|x|\le\pi} x^{-(4h+2)}|\widehat{\psi}(a(\nu)2^{j}x)|^2|\widehat{\psi}(a(\nu)2^{j'} x)|^2||g(x)|^2-|g(0)|^2|^2 dx$$

$$\le a(\nu)^{-4h}\int_{|x|\le\pi} x^{-(4h+2)}|\widehat{\psi}(a(\nu)2^{j}x)|^2|\widehat{\psi}(a(\nu)2^{j'} x)|^2 x^{2\beta} dx$$

$$= a(\nu)^{1-2\beta} \int_{|x|\le a(\nu)\pi} x^{-(4h+2)+2\beta} |\widehat{\psi}(a(\nu)2^j x)|^2 |\widehat{\psi}(a(\nu)2^{j'} x)|^2 dx$$

$$\le a(\nu)^{1-2\beta} \int_{\mathbb{R}} x^{-(4h+2)+2\beta} |\widehat{\psi}(a(\nu)2^j x)|^2 |\widehat{\psi}(a(\nu)2^{j'} x)|^2 dx,$$

by (3.12), (3.13) and (3.7), $\int_{\mathbb{R}} x^{-(4h+2)+2\beta} |\widehat{\psi}(a(\nu)2^j x)|^2 |\widehat{\psi}(a(\nu)2^{j'} x)|^2 dx < \infty$, so

$$a(\nu)^{-4h} \int_{|x|\le\pi} x^{-(4h+2)} |\widehat{\psi}(a(\nu)2^j x)|^2 |\widehat{\psi}(a(\nu)2^{j'} x)|^2 ||g(x)|^2 - |g(0)|^2|^2 dx \to 0,$$

as $\nu \to \infty$. On the other hand, by (3.12),

$$a(\nu)^{-4h} \int_{|x|>\pi} x^{-(4h+2)} |\widehat{\psi}(a(\nu)2^j x)|^2 |\widehat{\psi}(a(\nu)2^{j'} x)|^2 ||g(x)|^2 - |g(0)|^2|^2 dx$$

$$\le C a(\nu)^{-4h} \int_{|x|>\pi} x^{-(4h+2)} (a(\nu)x)^{-4\alpha} dx$$

$$+ C a(\nu)^{-4h-4\alpha} \int_{|x|>\pi} x^{-(4h+2)-4\alpha} dx \to 0,$$

as $\nu \to \infty$. □

Proposition A.2.1. *Let $\widehat{\sigma}^2$ be defined by (A.28). Then,*

$$a(\nu)^{-4h} \frac{\nu}{a} \mathrm{Cov}(\widehat{\sigma}^2(a(\nu)2^j), \widehat{\sigma}^2(a(\nu)2^{j'})) \to 4\pi b^{4h+1} |g(0)|^4 \int_{\mathbb{R}} x^{-(4h+2)} |\widehat{\psi}(2^j x/b)|^2 |\widehat{\psi}(2^{j'} x/b)|^2 dx, \tag{A.29}$$

as $\nu \to \infty$, where $b = \gcd(2^j, 2^{j'})$.

Proof. The main proof is similar to the proof of Proposition 3.1 in [42], so we outline the main steps for the reader's convenience.

It suffices to consider the subsequence $\nu = a(\nu)2^{j+j'}\nu_*$. Then

$$a(\nu)^{-4h} \frac{\nu}{a(\nu)} \mathrm{Cov}(\widehat{\sigma}^2(a(\nu)2^j), \widehat{\sigma}^2(a(\nu)2^{j'}))$$

$$= a(\nu)^{-4h} \frac{1}{\nu_*} \sum_{k=1}^{2^{j'}\nu_*} \sum_{k'=1}^{2^j\nu_*} \mathrm{Cov}(d^2(a(\nu)2^j, k), d^2(a(\nu)2^{j'}, k'))$$

$$= 2a(\nu)^{-4h} \frac{1}{\nu_*} \sum_{k=1}^{2^{j'}\nu_*} \sum_{k'=1}^{2^j\nu_*} \Big(\mathbb{E} d(a(\nu)2^j, k) d(a(\nu)2^{j'}, k') \Big)^2$$

$$= 2a(\nu)^{-4h} \frac{1}{\nu_*} \sum_{k=1}^{2^{j'}\nu_*} \sum_{k'=1}^{2^j\nu_*} \Big(\int_{\mathbb{R}} e^{ia(\nu)(2^jk - 2^{j'}k)x} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2$$

$$= 2a(\nu)^{-4h} \frac{1}{\nu_*} \sum_{k=1}^{2^{j'}\nu_*} \sum_{k'=1}^{2^j\nu_*} \Big\{ \Big(\int_{\mathbb{R}} e^{ia(\nu)(2^jk - 2^{j'}k)x} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2$$
$$- \Big(\int_{\mathbb{R}} e^{ia(\nu)(2^jk - 2^{j'}k)x} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2 \Big\}$$
$$+ 2a(\nu)^{-4h} \frac{1}{\nu_*} \sum_{k=1}^{2^{j'}\nu_*} \sum_{k'=1}^{2^j\nu_*} \Big(\int_{\mathbb{R}} e^{ia(\nu)(2^jk - 2^{j'}k)x} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2,$$

where the second equality is a consequence of the Isserlis theorem. We now show that

$$a(\nu)^{-4h} \frac{1}{\nu_*} \sum_{k=1}^{2^{j'}\nu_*} \sum_{k'=1}^{2^j\nu_*} \Big\{ \Big(\int_{\mathbb{R}} e^{ia(\nu)(2^jk - 2^{j'}k)x} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2$$
$$- \Big(\int_{\mathbb{R}} e^{ia(\nu)(2^jk - 2^{j'}k)x} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2 \Big\} \to 0, \quad \nu \to \infty. \tag{A.30}$$

The summation in (A.30) can be reexpressed as (for the details, see the proof of Proposition 3.1, (iv) in [42])

$$a(\nu)^{-4h} \sum_{r \in \Pi(\nu_*)} \frac{\xi_r(\nu_*)}{\nu_*} \Big\{ \Big(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2$$
$$- \Big(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx \Big)^2 \Big\} =: \Theta_1. \tag{A.31}$$

In (A.31), $\Pi(\nu_*) = \gcd(a(\nu)2^j, a(\nu)2^{j'})\mathbb{Z} \cap B_{jj'}(\nu_*)$, $B_{jj'}(\nu_*)$ is the range for r such that the pairs (k, k') satisfying $2^jk - 2^{j'}k' = \gcd(2^j, 2^{j'})w$ for some $w \in \mathbb{Z}$ lie in the region

$$1 \le k \le 2^{j'}\nu_*, \quad 1 \le k' \le 2^{j}\nu_*,$$

and

$$\frac{\xi_r(\nu_*)}{\nu_*} \to \gcd(2^j, 2^{j'}), \quad \nu \to \infty. \tag{A.32}$$

By Parseval's Theorem, the sequences

$$\left\{\left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right\}_{r\in\mathbb{Z}}$$

and

$$\left\{\left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right\}_{r\in\mathbb{Z}}$$

are summable. Moreover, by (A.32), for large enough ν,

$$\Theta_1 < (\gcd(2^j, 2^{j'})+1) a(\nu)^{-4h} \Bigg| \sum_{r\in\Pi(\nu_*)} \left\{\left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right.$$

$$\left. - \left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right\}\Bigg|.$$

By Minkowski's inequality,

$$\Bigg| a(\nu)^{-4h} \left\{\sum_{r\in\Pi(\nu_*)} \left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(x)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right\}^{1/2}$$

$$- \left\{\sum_{r\in\Pi(\nu_*)} \left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} |g(0)|^2 \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right\}^{1/2} \Bigg|$$

$$\le \left\{\sum_{r\in\Pi(\nu_*)} a(\nu)^{-4h} \left(\int_{\mathbb{R}} e^{ia(\nu)rx} x^{-(2h+1)} (|g(x)|^2 - |g(0)|^2) \overline{\widehat{\psi}(a(\nu)2^j x)} \widehat{\psi}(a(\nu)2^{j'}x) dx\right)^2\right\}^{1/2}$$

$$\le \left(2\pi a(\nu)^{-4h} \int_{\mathbb{R}} x^{-(4h+2)} |\widehat{\psi}(a(\nu)2^j x)|^2 |\widehat{\psi}(a(\nu)2^{j'}x)|^2 ||g(x)|^2 - |g(0)|^2|^2 dx\right)^{1/2} \to 0,$$

as $\nu \to \infty$. The last inequality is a consequence of Parseval's theorem, and the limit follows from Lemma A.2.1. This proves (A.30), as desired. By an analogous procedure, we obtain

$$2a(\nu)^{-4h}\frac{1}{\nu_*}\sum_{k=1}^{2^{j'}\nu_*}\sum_{k'=1}^{2^{j}\nu_*}\Big(\int_{\mathbb{R}} e^{ia(\nu)(2^j k-2^{j'}k)x}x^{-(2h+1)}|g(0)|^2\overline{\widehat{\psi}(a(\nu)2^j x)}\widehat{\psi}(a(\nu)2^{j'}x)dx\Big)^2$$

$$= 2\frac{1}{\nu_*}\sum_{k=1}^{2^{j'}\nu_*}\sum_{k'=1}^{2^{j}\nu_*}\Big(\int_{\mathbb{R}} e^{i(2^j k-2^{j'}k)x}x^{-(2h+1)}|g(0)|^2\overline{\widehat{\psi}(2^j x)}\widehat{\psi}(2^{j'}x)dx\Big)^2$$

$$\to 2\gcd(2^j,2^{j'})\sum_{z=-\infty}^{\infty}\Big(\int_{\mathbb{R}} e^{i\gcd(2^j,2^{j'})zx}x^{-(2h+1)}|g(0)|^2\overline{\widehat{\psi}(2^j x)}\widehat{\psi}(2^{j'}x)dx\Big)^2$$

$$= 2b^{4h+1}\sum_{z=-\infty}^{\infty}\Big(\int_{\mathbb{R}} e^{izx}x^{-(2h+1)}|g(0)|^2\overline{\widehat{\psi}(2^j x/b)}\widehat{\psi}(2^{j'}x/b)dx\Big)^2$$

$$= 4\pi b^{4h+1}|g(0)|^4\int_{\mathbb{R}} x^{-(4h+2)}|\widehat{\psi}(2^j x/b)|^2|\widehat{\psi}(2^{j'}x/b)|^2dx, \tag{A.33}$$

where the last equality is a consequence of Parseval's theorem. By (A.30) and(A.33), (A.29) holds. □

Theorem A.2.1. *Let σ^2 and $\widehat{\sigma}^2$ be defined as (A.27) and (A.28), and fix the octaves $j_1 < \cdots < j_m$. Then,*

$$\sqrt{\nu/a}\left(\begin{pmatrix}\widehat{\sigma}^2(a2^{j_1})/a^{2h}\\ \vdots \\ \widehat{\sigma}^2(a2^{j_m})/a^{2h}\end{pmatrix} - \begin{pmatrix}\sigma^2(a2^{j_1})/a^{2h}\\ \vdots \\ \sigma^2(a2^{j_m})/a^{2h}\end{pmatrix}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0},W),$$

as $\nu\to\infty$, where

$$W_{ik} = 4\pi b_{j_ij_k}^{4h+1}|g(0)|^4\int_{\mathbb{R}} x^{-(4h+2)}|\widehat{\psi}(2^j x/b_{j_ij_k})|^2|\widehat{\psi}(2^i x/b_{j_ij_k}))|^2dx,$$

and $b_{j_ij_k} = \gcd(2^{j_i},2^{j_k})$, $i,k = 1,\cdots,m$.

Proof. The proof can be written as a simple adaption of the proof of Theorem 3.3.1. □

A.3 Proofs for Chapter 3

PROOF OF PROPOSITION 3.3.1: The statement (ii) is a direct consequence of (i), so we only prove the latter. We proceed as in the proof of Proposition 3.3 (i) in [43]. It suffices to consider the subsequence $\nu = 2^{j+j'}\nu_*$, $\nu_* \to \infty$. Then, $K_j = 2^{j'}\nu_*$, $K_{j'} = 2^j\nu_*$, and $\sqrt{K_j}\sqrt{K_{j'}}K_j^{-1}K_{j'}^{-1} = 2^{-(j+j')/2}/\nu_*$. The covariance between wavelet coefficients can be expressed as

$$\mathbb{E}D(2^j,k)D(2^{j'},k')^* = \mathbb{E}\int_{\mathbb{R}}\int_{\mathbb{R}}\psi(t)\psi(t')Y(2^jt+2^jk)Y(2^{j'}t'+2^{j'}k')dtdt'$$

$$= \int_{\mathbb{R}}dx\int_{\mathbb{R}}\int_{\mathbb{R}}\psi(t)\psi(t')e^{i(2^jt+2^jk)x}|x|^{-(H+I/2)}G(x)|x|^{-(H+I/2)^*}\overline{e^{i(2^{j'}t'+2^{j'}k')x}}dtdt'$$

$$= \int_{\mathbb{R}}\overline{\widehat{\psi}(2^jx)}\widehat{\psi}(2^{j'}x)e^{i(2^jk-2^{j'}k')x}|x|^{-(H+I/2)}G(x)|x|^{-(H+I/2)^*}dx,$$

$$=: \Phi_{2^jk-2^{j'}k'}.$$

Let $\Xi_{2^jk-2^{j'}k'} = \Phi_{2^jk-2^{j'}k'} \otimes \Phi_{2^jk-2^{j'}k'}$. By Theorem 1.8 in [44], p.10, the range of indices spanned by $2^jk - 2^{j'}k'$ is $\mathbb{Z}\mathrm{gcd}(2^j, 2^{j'})$. Thus, we would like to show that

$$\sum_{z=-\infty}^{\infty}\|\Xi_{z\mathrm{gcd}(2^j,2^{j'})}\| < \infty. \tag{A.34}$$

Note that $\|\Xi_{2^jk-2^{j'}k'}\|_{l_1} = \|\mathrm{vec}(\Phi_{z\mathrm{gcd}(2^j,2^{j'})})\mathrm{vec}(\Phi_{z\mathrm{gcd}(2^j,2^{j'})})^*\|_{l_1} \leq \|\Phi_{z\mathrm{gcd}(2^j,2^{j'})}\|_{l_1}^2$. Thus, if $\sum_{z=-\infty}^{\infty}\|\Phi_z\|^2 < \infty$, the expression (3.24) is now a consequence of Lemma A.6.4 below. In fact,

$$\|\Phi_z\|^2 = \left\|P\int_{\mathbb{R}}e^{izx}\overline{\widehat{\psi}(2^jx)}\widehat{\psi}(2^{j'}x)\mathrm{diag}(x^{-(2h_1+1)}|g_1(x)|^2,\cdots,x^{-(2h_n+1)}|g_n(x)|^2)dxP^*\right\|^2$$

$$\leq C\|P\|^4 \max_{1\leq k\leq n} \left| \int_{\mathbb{R}} e^{izx} \overline{\widehat{\psi}(2^j x)} \widehat{\psi}(2^{j'} x) x^{-(2h_k+1)} |g_k(x)|^2 dx \right|^2 .$$

For any $1 \leq k \leq n$, $\overline{\widehat{\psi}(2^j x)} \widehat{\psi}(2^{j'} x) x^{-(2h_k+1)} |g_k(x)|^2 \in L^2(\mathbb{R})$. Thus, by Parseval's Theorem,

$$\sum_{z=-\infty}^{\infty} \left| \int_{\mathbb{R}} e^{izx} \overline{\widehat{\psi}(2^j x)} \widehat{\psi}(2^{j'} x) x^{-(2h_k+1)} |g_k(x)|^2 dx \right|^2$$

$$= 2\pi \int_{\mathbb{R}} \left| \overline{\widehat{\psi}(2^j x)} \widehat{\psi}(2^{j'} x) x^{-(2h_k+1)} |g_k(x)|^2 \right|^2 dx < \infty,$$

this proves $\sum_{z=-\infty}^{\infty} \|\Phi_z\|^2 < \infty$, as claimed. □

Proof of Theorem 3.3.1: For notation simplicity, we will restrict ourselves to the bivariate context ($n = 2$). the argument for general n can be worked out by a simple adaptation.

The proof is by means of Cramér-Wold device. Form the vector of wavelet coefficients

$$V_\nu = (d_1(2^{j_1}, 1), d_2(2^{j_1}, 1), \cdots, d_1(2^{j_1}, K_{j_1}), d_2(2^{j_1}, K_{j_1}); \cdots;$$

$$d_1(2^{j_m}, 1), d_2(2^{j_m}, 1), \cdots, d_1(2^{j_m}, K_{j_m}), d_2(2^{j_m}, K_{j_m}))^T \in \mathbb{R}^{\Upsilon(\nu)},$$

where $\Upsilon(\nu) = 2\sum_{j=j_1}^{j_m} K_j$. Notice that $m, j_1, \cdots, j_m$ are fixed, but each K_j goes to infinity with ν. let

$$\alpha = (\alpha_{j_1} \cdots, \alpha_{j_m})^T \in \mathbb{R}^{3m}$$

where

$$\alpha_j = (\alpha_{j,1}, \alpha_{j,12}, \alpha_{j,3})^T \in \mathbb{R}^3, \quad j = j_1, \cdots, j_m.$$

Now form the block-diagonal matrix

$$D_\nu = \text{diag}\Big(\underbrace{\frac{1}{K_{j_1}} \sqrt{\frac{1}{2^{j_1}}} \Omega_{j_1}, \cdots, \frac{1}{K_{j_1}} \sqrt{\frac{1}{2^{j_1}}} \Omega_{j_1}}_{K_{j_1}}; \cdots; \underbrace{\frac{1}{K_{j_m}} \sqrt{\frac{1}{2^{j_m}}} \Omega_{j_m}, \cdots, \frac{1}{K_{j_m}} \sqrt{\frac{1}{2^{j_m}}} \Omega_{j_m}}_{K_{j_m}} \Big),$$

where

$$\Omega_j = \begin{pmatrix} \alpha_{j,1} & \alpha_{j,12}/2 \\ \alpha_{j,12}/2 & \alpha_{j,2} \end{pmatrix}, \quad j = j_1, \cdots, j_m.$$

Let $\Gamma(\nu)$ be the covariance matrix of V_ν.

We would like to show $\sqrt{\nu}(V_\nu^* D_\nu V_\nu - \mathbb{E}V_\nu^* D_\nu V_\nu) \xrightarrow{d} N(0, \sigma^2)$ for some $\sigma^2 < \infty$. By Lemma A.6.1, we only need to prove that

(1) $\sigma^2 := \lim_{\nu\to\infty} \mathrm{Var}(\sqrt{\nu} V_\nu^* D_\nu V_\nu) < \infty$;

(2) $\lim_{\nu\to\infty} \rho(\sqrt{\nu} D_\nu)\rho(\Gamma(\nu)) = 0$,

where $\rho(\cdot)$ is the spectral radius of the matrix.

Statement (1) is a consequence of Proposition 3.3.1, i.e.,

$$\mathrm{Var}(\sqrt{\nu} V^* DV) = \sum_{j=j_1}^{j_m} \sum_{j'=j_1}^{j_m} \alpha_j^T \Big\{ \sqrt{\frac{\nu}{2^j}} \sqrt{\frac{\nu}{2^{j'}}} \mathrm{Cov}(\mathrm{vec}_{\mathcal{S}} W(2^j), \mathrm{vec}_{\mathcal{S}} W(2^{j'})) \Big\} \alpha_{j'} \to$$

$$\sum_{j=j_1}^{j_m} \sum_{j'=j_1}^{j_m} \alpha_j^T G_{jj'} \alpha_{j'} < \infty, \quad \nu \to \infty.$$

To show statement (2), note that, by Lemma A.6.2,

$$\rho(\Gamma(\nu)) \le \rho(\Gamma_1) + \cdots + \rho(\Gamma_m),$$

where Γ_i is the covariance matrix of $V_i := (d_1(2^{j_i}, 1), d_2(2^{j_i}, 1), \cdots, d_1(2^{j_i}, K_{j_i}), d_2(2^{j_i}, K_{j_i}))^T$. Let T_i be the permutation matrix such that

$$T_i V_i = (d_1(2^{j_i}, 1), d_1(2^{j_i}, 2), \cdots, d_1(2^{j_i}, K_{j_i}); d_2(2^{j_i}, 1), d_2(2^{j_i}, 2), \cdots, d_2(2^{j_i}, K_{j_i}))^T =: \widetilde{V}_i,$$

and let $\widetilde{\Gamma}_i$ be the covariance matrix of $\widetilde{V}_i$. Then,

$$\Gamma_i = \mathbb{E} V_i V_i^T = \mathbb{E}(T_i^{-1} \widetilde{V}_i \widetilde{V}_i^T T_i) =: T_i^{-1} \widetilde{\Gamma}_i T_i.$$

Since a similarity transformation of a matrix does not change the eigenvalues, we have $\rho(\Gamma_i) = \rho(\widetilde{\Gamma}_i)$. By Lemma A.6.2 again,

$$\rho(\widetilde{\Gamma}_i) \le \rho(\Gamma_{i1}) + \rho(\Gamma_{i2}),$$

where Γ_{is} is the covariance matrix of

$$V_{iv} := (d_v(2^{j_i},1), d_v(2^{j_i},2), \cdots, d_v(2^{j_i},K_{j_i}))^T, \quad i = 1, \cdots, m, \quad v = 1, 2.$$

Let Y_v be the vth entry of Y. Then,

$$Y_v(t) = \sum_{l=1}^{2} p_{vl} \int_{\mathbb{R}} \frac{e^{itx} - \sum_{k=1}^{N_i-1} \frac{1}{k!}(itx)^k}{(ix)^{N_i}} |x|^{-(h_i - N_i + 1/2)} g_l(x) \widetilde{B}(dx).$$

Let $\{d_v(2^j,k)\}_{k\in\mathbb{Z}}$ be the wavelet transform of Y_v at octave j and shift k. Then, the covariance function of $d_v(2^j,k)$ and $d_v(2^j,k')$ is given by

$$\mathbb{E} d_v(2^j,k) d_v(2^j,k') = \sum_{l=1}^{2} p_{vl}^2 2^{2jh_l} \int_{\mathbb{R}} e^{i(k-k')y} |\widehat{\psi}(y)|^2 y^{-2(h_l+1/2)} \left| g_l\left(\frac{y}{2^j}\right)\right|^2 dy.$$

Thus, $\{d_v(2^j,k)\}_{k\in\mathbb{Z}}$ is a stationary sequence for a fixed octave j and its the spectral density is

$$f_{j,s}(y) = \sum_{l=1}^{2} p_{sl}^2 2^{2jh_l} \sum_{k=-\infty}^{\infty} |\widehat{\psi}(y+2k\pi)|^2 |y+2k\pi|^{-2(h_l+1/2)} \left| g_l\left(\frac{y+2k\pi}{2^j}\right)\right|^2$$

$$\le C_j \sum_{l=1}^{2} \sum_{k=-\infty}^{\infty} |\widehat{\psi}(y+2k\pi)|^2 |y+2k\pi|^{-2(h_l+1/2)}, \quad -\pi < y < \pi. \tag{A.35}$$

Fix $l = 1, 2$, the summation in (A.35) is bounded on $(-\pi, \pi)$ by using (3.13) for $k = 0$, and the decay of $\widehat{\psi}$ given by (W3) for bounding the remaining terms $\sum_{k\neq 0}$. By Lemma A.6.3 below, $\rho(\Gamma_{iv}) < \infty$, $i = 1, \cdots, m$, $v = 1, 2$. Thus, $\rho(\Gamma_\nu) < \infty$. Since $\rho(\sqrt{\nu} D_\nu) = O(\frac{1}{\sqrt{\nu}})$, then $\lim_{\nu\to\infty} \rho(\sqrt{\nu} D_\nu)\rho(\Gamma_\nu) = 0$. □

Proof of Theorem 3.4.1: For any matrix $S \in \mathcal{S}_+(n,\mathbb{R})$, define the vector-valued function

$$f : \mathrm{vec}_{\mathcal{S}}(S) \to (\xi_1,\ldots,\xi_n,\mathrm{vec}(\mathcal{O})) \tag{A.36}$$

such that $S = \mathcal{O}\mathrm{diag}(\xi_1,\ldots,\xi_n)\mathcal{O}^*$, $\mathcal{O} \in O(n)$, $\xi_1 < \ldots < \xi_n$, is the spectral decomposition of S, and $\mathcal{O} = (o_{i_1 i_2})_{i_1,i_2=1,\ldots,n}$ satisfies $o_{1i} \geq 0$, $i = 1,\ldots,n$ (cf. (3.29)). Since $\mathbb{E}W(2^j)$ has pairwise distinct eigenvalues, Theorem A.6.1 implies that f is infinitely differentiable on a neighborhood of $\mathbb{E}W(2^j)$. Moreover, the Jacobian matrix $\mathcal{J}_j$ of f at the point $\mathbb{E}W(2^j)$ is given by (3.31) with $S = \mathbb{E}W(2^j)$. So, let $J = \mathrm{diag}(\mathcal{J}_1,\ldots,\mathcal{J}_m)$. Recall the notation (3.2) for block-diagonal matrices. The Delta method and Theorem 3.3.1, imply that

$$(\sqrt{K_j}(\mathrm{vec}_{\mathcal{D}}(L_j - \Lambda_j)), \sqrt{K_j}(\mathrm{vec}(\widehat{O}_j - O_j)))^T_{j=j_1,\ldots,j_m}$$

$$= (\sqrt{K_j}(f(\mathrm{vec}_{\mathcal{S}}(W(2^j)) - f(\mathrm{vec}_{\mathcal{S}}(\mathbb{E}W(2^j))))^T_{j=j_1,\ldots,j_m} \stackrel{d}{\to} \mathcal{N}_{mn(n+1)}(\mathbf{0}, JFJ^*), \tag{A.37}$$

as claimed. $\Box$

Proof of Propostion 3.4.1: From (3.21), write out the spectral decomposition

$$P\mathcal{E}(2^0)P^* = \mathbb{E}W(2^0) = O\mathrm{diag}(a_1,\ldots,a_n)O^*, \quad O \in O(n), \quad 0 < a_1 \leq \ldots \leq a_n.$$

Then, there is $Q \in O(n)$ such that $\mathcal{E}(2^0)^{-1/2}P^{-1}O\mathrm{diag}(\sqrt{a_1},\ldots,\sqrt{a_n}) = Q$. Solving for P yields

$$P = O\mathrm{diag}(\sqrt{a_1},\ldots,\sqrt{a_n})Q^*\mathcal{E}(2^0)^{-1/2}. \tag{A.38}$$

By plugging (A.38) into (3.21),

$$\mathbb{E}W(2^j) = O\mathrm{diag}(\sqrt{a_1},\ldots,\sqrt{a_n})\{Q^*\mathcal{E}(2^0)^{-1/2}\mathcal{E}(2^j)^{1/2}$$

$$\mathrm{diag}(2^{2jh_1},\ldots,2^{2jh_n})\mathcal{E}(2^j)^{1/2}\mathcal{E}(2^0)^{-1/2}Q\}\mathrm{diag}(\sqrt{a_1},\ldots,\sqrt{a_n})O^*.$$

$$=: OMNMO^*,$$

where

$$M = \text{diag}(\sqrt{a_1}, \dots, \sqrt{a_n}),$$

and

$$N = Q^*\mathcal{E}(2^0)^{-1/2}\mathcal{E}(2^j)^{1/2}\text{diag}(2^{2jh_1}, \dots, 2^{2jh_n})\mathcal{E}(2^j)^{1/2}\mathcal{E}(2^0)^{-1/2}Q.$$

The eigenvalues of $\mathbb{E}W(2^j)$ are equal to the eigenvalues of MNM. Write the ordered eigenvalues of $\mathbb{E}W(2^j)$ as $\lambda_1(j) \leq \dots \leq \lambda_n(j)$. We will show that

$$\lambda_1(j) < \dots < \lambda_n(j) \tag{A.39}$$

for large enough j. Let $\mathcal{U}_k$ be a k-dimensional subspace of $\mathbb{R}^n$ and let $\langle x, y\rangle = x^T y$, $x, y \in \mathbb{R}^n$. By the min-max theorem (e.g., Teschl [45], section 4.4),

$$\lambda_k = \min_{\mathcal{U}_k} \max_{x\in\mathcal{U}_k, x\neq\mathbf{0}} \frac{\langle MNMx, x\rangle}{\langle x, x\rangle} = \min_{\mathcal{U}_k} \max_{x\in\mathcal{U}_k, x\neq\mathbf{0}} \frac{\langle NMx, Mx\rangle}{\langle Mx, Mx\rangle}\frac{\langle Mx, Mx\rangle}{\langle x, x\rangle}.$$

Moreover,

$$a_1 \leq \frac{\langle M^2 x, x\rangle}{\langle x, x\rangle} \leq a_n,$$

$$\min_{\mathcal{U}_k} \max_{x\in\mathcal{U}_k, x\neq\mathbf{0}} \frac{\langle NMx, Mx\rangle}{\langle Mx, Mx\rangle} = \min_{\mathcal{U}_k} \max_{x\in\mathcal{U}_k, x\neq\mathbf{0}} \frac{\langle Nx, x\rangle}{\langle x, x\rangle} = \min_{\mathcal{U}_k} \max_{x\in\mathcal{U}_k, x\neq\mathbf{0}} \frac{\langle \widetilde{N}x, x\rangle}{\langle x, x\rangle},$$

where the first equality is a consequence of the fact that M is a full rank matrix, and

$$\widetilde{N} = \mathcal{E}(2^0)^{-1/2}\mathcal{E}(2^j)^{1/2}\text{diag}(2^{2jh_1}, \dots, 2^{2jh_n})\mathcal{E}(2^j)^{1/2}\mathcal{E}(2^0)^{-1/2}.$$

However $\mathcal{E}(2^0)^{-1/2}\mathcal{E}(2^j)^{1/2} \to \text{diag}(b_1, \cdots, b_n)$, as $j \to \infty$, where

$$b_i = \int_{\mathbb{R}} |\widehat{\psi}(y)|^2 |y|^{-h_i-1} |g_i(0)|^2 dy \Big/ \int_{\mathbb{R}} |\widehat{\psi}(y)|^2 |y|^{-h_i-1} |g_i(y)|^2 dy > 0, \quad i = 1, \cdots, n.$$

Thus, for large enough j,

$$b_k 2^{2jh_k} - 1 < \min_{\mathcal{U}_k} \max_{x \in \mathcal{U}_k, x \neq \mathbf{0}} \frac{\langle \widetilde{N}x, x\rangle}{\langle x, x\rangle} < b_k 2^{2jh_k} + 1,$$

and $a_1(b_k 2^{2jh_k} - 1) \leq \lambda_k \leq a_n(b_k 2^{2jh_k} + 1)$. So, (A.39) holds for any j such that $a_n(b_k 2^{2jh_k} + 1) < a_1(b_{k+1} 2^{2jh_{k+1}} - 1)$, $k = 1, \dots, n-1$. $\square$

Proof of Theorem 3.5.2: We first show (i). Let C_0, C_1, R and Λ be as in (3.37) and (3.38). We now show that, under (3.41), any solution B produced by the EJD algorithm is in the set $\mathcal{M}_{\mathrm{EJD}}$. In view of (3.21), consider the polar decomposition

$$R = \mathcal{P}O, \quad \mathcal{P} \text{ is positive definite}, \quad O \in O(n). \tag{A.40}$$

The decomposition (A.40) always exists for nonsingular, real matrices, and is unique. Thus,

$$C_0 = RR^* = \mathcal{P}OO^*\mathcal{P}^* = \mathcal{P}^2,$$

Since square roots are unique, **Step 1** yields

$$W = \mathcal{P}^{-1}. \tag{A.41}$$

Step 2 and (3.37) imply that

$$WC_1W^* = W(\mathcal{P}O\Lambda O^*\mathcal{P}^*)W^* = O\Lambda O^*. \tag{A.42}$$

By (3.5), we can assume that the eigenvalues of Λ (see (3.38)) are ordered from smallest to largest, in which case the column vector $\mathbf{o}_{\cdot i}$ in O is associated with the eigenvalue $2^{2(J_2-J_1)\,h_i}$, $i = 1, \dots, n$. However, in the spectral decomposition in **Step 2**, each orthogonal eigenvector is determined up to multiplication by -1. Thus, for $\mathcal{I}$ as in (5.16), $Q^* \in O\mathcal{I}$, and any demixing matrix B produced by the EJD algorithm has the form

$$B = QW \in \mathcal{I}(\mathcal{P}O)^{-1} = \mathcal{I}R^{-1} = \mathcal{I}(P\mathcal{E}(2^{J_1})^{1/2}\text{diag}(2^{J_1 h_1}, \dots, 2^{J_1 h_n}))^{-1}. \tag{A.43}$$

In other words, $B \in \mathcal{M}_{\text{EJD}}$. Conversely, it is clear that any matrix in $\mathcal{M}_{\text{EJD}}$ can be attained as a solution to the EJD algorithm under (3.41). This establishes (i).

To show (ii), consider the EJD algorithm with input matrices $\widehat{C}_0 = W(2^{J_1})$, $\widehat{C}_1 = W(2^{J_1})$ (we write $\widehat{C}_k$ to avoid confusion with their deterministic counterparts $C_k = \mathbb{E}W(2^{J_k})$, $k = 1, 2$). By replacing all matrices in the proof of (i) with their sample counterparts and following the same argument, the set of solutions to the EJD algorithm is made up of matrices of the form

$$\widehat{B}_\nu = \Pi \widehat{O}^* \widehat{C}_0^{-1/2}, \quad \text{where } \Pi \in \mathcal{I}, \quad \widehat{C}_0^{-1/2} \widehat{C}_1 \widehat{C}_0^{-1/2} = \widehat{O}\widehat{\Lambda}\widehat{O}^*,$$

for some spectral decomposition with orthogonal $\widehat{O}$ and diagonal $\widehat{\Lambda}$. Note that $\widehat{C}_0 \xrightarrow{P} C_0$, by Theorem 3.3.1. Since the square root is unique and C_0 is invertible, then Theorem A.6.1 implies that, with probability going to 1, the inverse square root $\widehat{C}_0^{-1/2}$ exists. Thus, by Theorem 3.3.1, and Slutsky's theorem, $\widehat{C}_0^{-1/2}\widehat{C}_1\widehat{C}_0^{-1/2} \xrightarrow{P} \mathcal{P}^{-1}C_1\mathcal{P}^{-1}$. However, $\mathcal{P}^{-1}C_1\mathcal{P}^{-1}$ is a symmetric positive definite matrix that admits the spectral decomposition $O\Lambda O^*$ with pairwise distinct eigenvalues (see (A.41) and (A.42)). Then, by Theorem A.6.1, so is $\widehat{C}_0^{-1/2}\widehat{C}_1\widehat{C}_0^{-1/2}$ with probability going to 1. Therefore, Theorem A.6.1 implies that there is a spectral decomposition of $\widehat{C}_0^{-1/2}\widehat{C}_1\widehat{C}_0^{-1/2}$ whose eigenvector and eigenvalue matrices $\widehat{O}$ and $\widehat{\Lambda}$, respectively, satisfy $\widehat{O} \xrightarrow{P} O$, $\widehat{\Lambda} \xrightarrow{P} \Lambda$. So, $\widehat{B}_\nu = \Pi\widehat{O}^*\widehat{C}_0^{-1/2} \xrightarrow{P} \Pi O^* C_0^{-1/2} = \Pi(\mathcal{P}O)^{-1}$, i.e., the sequence $\widehat{B}_\nu$ satisfies (3.42).

We now show (iii). From Theorem 3.3.1, $\sqrt{\nu}(\text{vec}_{\mathcal{S}}(\widehat{C}_0 - C_0), \text{vec}_{\mathcal{S}}(\widehat{C}_1 - C_1))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, F)$, where $F \in \mathcal{S}_+(n(n+1), \mathbb{R})$. Consider the spectral decompositions $C_0 =$

$O_0\Lambda_0 O_0^T$ and $\widehat{C}_0 = \widehat{O}_0\widehat{\Lambda}_0\widehat{O}_0^T$. Then, by (3.33) at $j = J_1$ and the Delta method,

$$\sqrt{\nu}(\mathrm{vec}_{\mathcal{D}}(\widehat{\Lambda}_0 - \Lambda_0), \mathrm{vec}(\widehat{O}_0 - O_0), \mathrm{vec}_{\mathcal{S}}(\widehat{C}_1 - C_1))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, JFJ^*). \tag{A.44}$$

In (A.44), J is the block-diagonal matrix

$$J = \mathrm{diag}\Big(\mathcal{J}_{J_1}, I_{\frac{n(n+1)}{2}}\Big) \in M\Big(n + n^2 + \frac{n(n+1)}{2}, n(n+1), \mathbb{R}\Big),$$

and $\mathcal{J}_{J_1}$ is given by (3.31) with $S = C_0$. Starting from (A.44), again the Delta method yields

$$\sqrt{\nu}(\mathrm{vec}_{\mathcal{D}}(\widehat{\Lambda}_0^{-\frac{1}{2}} - \Lambda_0^{-\frac{1}{2}}), \mathrm{vec}(\widehat{O}_0 - O_0), \mathrm{vec}_{\mathcal{S}}(\widehat{C}_1 - C_1))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma),$$

where

$$\Sigma = \begin{pmatrix} -\frac{1}{2}\Lambda_0^{-3/2} & \\ & I_{n^2+\frac{n(n+1)}{2}} \end{pmatrix} JFJ^T \begin{pmatrix} -\frac{1}{2}\Lambda_0^{-3/2} & \\ & I_{n^2+\frac{n(n+1)}{2}} \end{pmatrix} \in M\Big(n+n^2+\frac{n(n+1)}{2}, \mathbb{R}\Big).$$

Thus, we arrive at the matrix system of equations

$$\widehat{\Lambda}_0^{-\frac{1}{2}} = \Lambda_0^{-\frac{1}{2}} + \frac{1}{\sqrt{\nu}}Z_{1,\nu}, \quad \widehat{O}_0 = O_0 + \frac{1}{\sqrt{\nu}}Z_{2,\nu}, \quad \widehat{C}_1 = C_1 + \frac{1}{\sqrt{\nu}}Z_{3,\nu}, \tag{A.45}$$

where

$$(\mathrm{vec}_{\mathcal{D}}(Z_{1,\nu}), \mathrm{vec}(Z_{2,\nu}), \mathrm{vec}_{\mathcal{S}}(Z_{3,\nu}))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma). \tag{A.46}$$

Since $C_0^{-1} = O_0\Lambda_0^{-1}O_0^*$, $\widehat{C}_0^{-1} = \widehat{O}_0\Lambda_0^{-1}\widehat{O}_0^*$, consider the EJD algorithm with $W = O_0\Lambda_0^{-\frac{1}{2}}O_0^*$ in **Step 1**, and also with its estimated counterpart $\widehat{W} = \widehat{O}_0\widehat{\Lambda}_0^{-\frac{1}{2}}\widehat{O}_0^*$. By (A.45),

$$\sqrt{\nu}(\widehat{W} - W) = O_0\Lambda_0^{-\frac{1}{2}}Z_{2,\nu}^* + O_0 Z_{1,\nu}O_0^* + Z_{2,\nu}\Lambda_0^{-\frac{1}{2}}O_0^* + O_P\Big(\frac{1}{\sqrt{\nu}}\Big), \tag{A.47}$$

and

$$\sqrt{\nu}(\widehat{W}\widehat{C}_1\widehat{W}^* - WC_1W^*) = O_0\Lambda_0^{-\frac{1}{2}}O_0^*C_1O_0\Lambda_0^{-\frac{1}{2}}Z_{2,\nu}^* + O_0\Lambda_0^{-\frac{1}{2}}O_0^*C_1O_0Z_{1,\nu}O_0^*$$

$$+O_0\Lambda_0^{-\frac{1}{2}}O_0^*C_1Z_{2,\nu}\Lambda_0^{-\frac{1}{2}}O_0^* + O_0\Lambda_0^{-\frac{1}{2}}O_0^*Z_{3,\nu}O_0\Lambda_0^{-\frac{1}{2}}O_0^* + O_0\Lambda_0^{-\frac{1}{2}}Z_{2,\nu}^*C_1O_0\Lambda_0^{-\frac{1}{2}}O_0^*$$

$$+ O_0Z_{1,\nu}O_0^*C_1O_0\Lambda_0^{-\frac{1}{2}}O_0^* + Z_{2\nu}\Lambda_0^{-\frac{1}{2}}O_0^*C_1O_0\Lambda_0^{-\frac{1}{2}}O_0^* + O_P\Big(\frac{1}{\sqrt{\nu}}\Big). \tag{A.48}$$

As a consequence, there are matrices

$$A_1 = A_1(O_0, \Lambda_0) \in M\Big(n^2, n + n^2 + \frac{n(n+1)}{2}, \mathbb{R}\Big)$$

and

$$A_2 = A_2(O_0, \Lambda_0, C_1) \in M\Big(\frac{n(n+1)}{2}, n + n^2 + \frac{n(n+1)}{2}, \mathbb{R}\Big)$$

such that

$$(\mathrm{vec}(O_0\Lambda_0^{-\frac{1}{2}}Z_{2,\nu}^* + O_0Z_{1,\nu}O_0^* + Z_{2,\nu}\Lambda_0^{-\frac{1}{2}}O_0^*))^T$$

$$= A_1(\mathrm{vec}_{\mathcal{D}}(Z_{1,\nu}), \mathrm{vec}(Z_{2,\nu}), \mathrm{vec}_{\mathcal{S}}(Z_{3,\nu}))^T, \tag{A.49}$$

$$(\mathrm{vec}_{\mathcal{S}}(O_0\Lambda^{-\frac{1}{2}}O_0^*C_1O_0\Lambda^{-\frac{1}{2}}Z_{2,\nu}^* + O_0\Lambda^{-\frac{1}{2}}O_0^*C_1O_0Z_{1,\nu}O_0^* + O_0\Lambda^{-\frac{1}{2}}O_0^*C_1Z_{2,\nu}\Lambda^{-\frac{1}{2}}O_0^*$$

$$+O_0\Lambda^{-\frac{1}{2}}O_0^*Z_{3,\nu}O_0\Lambda^{-\frac{1}{2}}O_0^* + O_0\Lambda^{-\frac{1}{2}}Z_{2,\nu}^*C_1O_0\Lambda^{-\frac{1}{2}}O_0^*$$

$$+O_0Z_{1,\nu}O_0^*C_1O_0\Lambda^{-\frac{1}{2}}O_0^* + Z_{2\nu}\Lambda^{-\frac{1}{2}}O_0^*C_1O_0\Lambda^{-\frac{1}{2}}O_0^*))^T$$

$$= A_2(\mathrm{vec}_{\mathcal{D}}(Z_{1,\nu}), \mathrm{vec}(Z_{2,\nu}), \mathrm{vec}_{\mathcal{S}}(Z_{3,\nu}))^T. \tag{A.50}$$

By (A.46)–(A.50),

$$\sqrt{\nu}(\mathrm{vec}(\widehat{W} - W), \mathrm{vec}_{\mathcal{S}}(\widehat{W}\widehat{C}_1\widehat{W}^* - WC_1W^*))^T \xrightarrow{d} N(0, \Sigma_2), \tag{A.51}$$

where

$$\Sigma_2 = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \Sigma \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}^* \in M(n^2 + n(n+1)/2, \mathbb{R}). \tag{A.52}$$

In **Step 2** of the EJD algorithm, write out the spectral decomposition $WC_1W^* = Q^*D_1Q$ and also its estimated counterpart $\widehat{W}\widehat{C}_1\widehat{W}^* = \widehat{Q}^*\widehat{D}_1\widehat{Q}$. From the ordering of eigenvalues in (A.42) and expression (3.38), WC_1W^* has pairwise distinct eigenvalues $2^{2(J_2-J_1)h_1} < \ldots < 2^{2(J_2-J_1)h_n}$. So, by the Delta method,

$$\sqrt{\nu}(\mathrm{vec}(\widehat{W} - W), \mathrm{vec}(\widehat{Q}^* - Q^*))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, J_Q\Sigma_2 J_Q^*),$$

where

$$J_Q = \mathrm{diag}(I_{n^2}, \mathcal{J}_q) \in M(2n^2, n^2 + n(n+1)/2), \tag{A.53}$$

and J_q is given by

$$\mathcal{J}_q = \begin{pmatrix} \big(\mathbf{q}_{1\cdot} \otimes (2^{2h_1\,(J_2-J_1)} I_n - WC_1W^*)^+\big) D \\ \vdots \\ \big(\mathbf{q}_{n\cdot} \otimes (2^{2h_n\,(J_2-J_1)} I_n - WC_1W^*)^+\big) D \end{pmatrix} \in M(n^2, n(n+1)/2, \mathbb{R})$$

(cf. expression (3.31)), where the vector $\mathbf{q}_{i\cdot}$ denotes the i-th row of $Q \in O(n)$. Let $T = (t_{i_1 i_2})_{i_1, i_2 = 1,\ldots,n^2}$ be the permutation operator defined by the transformation $T(\mathrm{vec}(R^*))^T = (\mathrm{vec}(R))^T$, $R \in M(n, \mathbb{R})$, i.e.,

$$t_{i_1 i_2} = \begin{cases} 1, & i_1 = (k-1)n + p,\, i_2 = (p-1)n + k, \quad k, p = 1, \ldots, n; \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$$\sqrt{\nu}(\mathrm{vec}(\widehat{W} - W), \mathrm{vec}(\widehat{Q} - Q))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma_3),$$

where

$$\Sigma_3 = \mathrm{diag}(I_{n^2}, T) J_Q \Sigma_2 J_Q^* \mathrm{diag}(I_{n^2}, T)^*. \tag{A.54}$$

Then, we arrive at

$$\widehat{W} = W + \frac{1}{\sqrt{\nu}} Z_{4,\nu}, \quad \widehat{Q} = Q + \frac{1}{\sqrt{\nu}} Z_{5,\nu},$$

where $(\mathrm{vec}(Z_{4,\nu}), \mathrm{vec}(Z_{5,\nu}))^T \xrightarrow{d} N(\mathbf{0}, \Sigma_3)$. Therefore,

$$\sqrt{\nu}(\widehat{Q}\widehat{W} - QW) = QZ_{4,\nu} + Z_{5,\nu}W + O_P\left(\frac{1}{\sqrt{\nu}}\right).$$

Therefore, for some matrix

$$A_3 = A_3(O_0, \Lambda_0, C_1) \in M(n^2, 2n^2),$$

we can write

$$(\mathrm{vec}(QZ_{4,\nu} + Z_{5,\nu}W))^T = A_3(\mathrm{vec}(Z_{4,\nu}), \mathrm{vec}(Z_{5,\nu}))^T. \tag{A.55}$$

Hence,

$$\sqrt{\nu}(\mathrm{vec}(\widehat{Q}\widehat{W}) - \mathrm{vec}(QW))^T \xrightarrow{d} \mathcal{N}(\mathbf{0}, A_3\Sigma_3 A_3^*),$$

as claimed. □

Proof of Proposition 3.6.1: Since $\widehat{X} = \widehat{I}_\nu \mathfrak{D} X$, then,

$$W_{\widehat{X}}(a(\nu)2^j) = (\widehat{I}_\nu)\mathfrak{D} W_X(a(\nu)2^j)\mathfrak{D}(\widehat{I}_\nu)^*.$$

Thus,

$$W_{\widehat{X}}(a(\nu)2^j) - \mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D} = (\widehat{I}_\nu)\mathfrak{D}W_X(a(\nu)2^j)\mathfrak{D}(\widehat{I}_\nu)^* - \mathfrak{D}\mathbb{E}W_X(a(\nu)2^j))\mathfrak{D}$$

$$= (\widehat{I}_\nu)\Big\{\mathfrak{D}W_X(a(\nu)2^j)\mathfrak{D} - (\widehat{I}_\nu)^{-1}\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}((\widehat{I}_\nu)^{-1})^*\Big\}(\widehat{I}_\nu)^*$$

$$= (\widehat{I}_\nu)\Big\{\Big[\mathfrak{D}(W_X(a(\nu)2^j) - \mathbb{E}W_X(a(\nu)2^j))\mathfrak{D}\Big] - \Big[((\widehat{I}_\nu)^{-1} - I)\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}\Big]$$

$$-\Big[\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}\Big(((\widehat{I}_\nu)^{-1})^* - I\Big)\Big]$$

$$- \Big[((\widehat{I}_\nu)^{-1} - I)\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}\Big(((\widehat{I}_\nu)^{-1})^* - I\Big)\Big]\Big\}(\widehat{I}_\nu)^*. \tag{A.56}$$

Recall that the operator $\text{vec}_{\mathcal{D}}(W_X(a(\nu)2^j))$ picks out the diagonal entries of the matrix $W_X(a(\nu)2^j)$, which are independent. Therefore, by Theorem A.2.1 for univariate processes,

$$\Big(\sqrt{\nu/a(\nu)}\text{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})(\text{vec}_{\mathcal{D}}(\mathfrak{D}(W_X(a(\nu)2^j) - \mathbb{E}W_X(a(\nu)2^j))\mathfrak{D}))^T\Big)_{j=j_1,\cdots,j_m}$$

$$= \mathcal{K}\Big(\sqrt{\nu/a(\nu)}\text{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})(\text{vec}_{\mathcal{D}}(W_X(a(\nu)2^j) - \mathbb{E}W_X(a(\nu)2^j)))^T\Big)_{j=j_1,\cdots,j_m}$$

$$\xrightarrow{d} \mathcal{N}(0, \mathcal{K}\mathbf{W}\mathcal{K}^*), \quad \nu \to \infty. \tag{A.57}$$

In (A.57),

$$\mathcal{K} = \text{diag}(\underbrace{\mathfrak{D}^2, \cdots, \mathfrak{D}^2}_{m}), \tag{A.58}$$

and

$$\mathbf{W}(k_1, k_2) = \begin{cases} w_{l,v,i}, & k_1 = ln + i, k_2 = vn + i; \\ 0, & \text{otherwise.} \end{cases} \tag{A.59}$$

where

$$w_{l,v,i} = 4\pi b_{lv}^{4h_i+1} g_i^4(0) \int_{\mathbb{R}} x^{-(4h_i+2)} |\widehat{\psi}(2^l x/b_{lv})|^2 |\widehat{\psi}(2^v x/b_{lv}))|^2 dx,$$

and $b_{lv} = \gcd(2^l, 2^v)$, for $l, v = 0, \cdots, m-1$, $i = 1, \cdots, n$. By (3.45) and the Delta method,

$$(\sqrt{\nu}\text{vec}((\widehat{I}_\nu)^{-1} - I))^T \xrightarrow{d} \mathcal{N}(0, \Sigma(J_1, J_2)).$$

Since $\mathbb{E}W_X(a(\nu)2^j) = \mathrm{diag}(\mathbb{E}W_{X_1}(a(\nu)2^j), \cdots, \mathbb{E}W_{X_n}(a(\nu)2^j))$, then

$$(\mathrm{vec}_{\mathcal{D}}(((\widehat{I}_\nu)^{-1} - I)\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j))\mathfrak{D})^T =$$

$$\mathfrak{D}^2 \mathrm{diag}(\mathbb{E}W_{X_1}(a(\nu)2^j), \cdots, \mathbb{E}W_{X_n}(a(\nu)2^j))(\mathrm{vec}_{\mathcal{D}}((\widehat{I}_\nu)^{-1} - I))^T.$$

Therefore,

$$\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(((\widehat{I}_\nu)^{-1} - I)\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}))^T$$

$$= \mathfrak{D}^2\Big(\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})\mathrm{diag}(\mathbb{E}W_{X_1}(a(\nu)2^j), \cdots, \mathbb{E}W_{X_n}(a(\nu)2^j))\Big)$$

$$\cdot \Big((\sqrt{\nu}\mathrm{vec}_{\mathcal{D}}((\widehat{I}_\nu)^{-1} - I))^T\Big) \cdot \frac{1}{\sqrt{a(\nu)}} = O_P\Big(\frac{1}{\sqrt{a(\nu)}}\Big). \qquad \text{(A.60)}$$

Similarly,

$$\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}((\widehat{I}_\nu)^{-1} - I)))^T$$

$$= O_P\Big(\frac{1}{\sqrt{a(\nu)}}\Big), \qquad \text{(A.61)}$$

and

$$\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})$$

$$\cdot\, (\mathrm{vec}_{\mathcal{D}}((\widehat{I}_\nu)^{-1} - I)\mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}(((\widehat{I}_\nu)^{-1})^* - I))^T = O_P\Big(\frac{1}{\sqrt{\nu}}\Big). \qquad \text{(A.62)}$$

Consequently, by (A.78)-(A.62) and Slutsky's Theorem, the limiting distribution of

$$\Big(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(W_{\widehat{X}}(a(\nu)2^j) - \mathfrak{D}\mathbb{E}W_X(a(\nu)2^j)\mathfrak{D}))^T\Big)_{j=j_1,\cdots,j_m}$$

is equal to the limiting distribution of

$$\mathcal{K}\Big(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(W_X(a(\nu)2^j) - \mathbb{E}W_X(a(\nu)2^j)))^T\Big)_{j=j_1,\cdots,j_m},$$

as claimed. $\square$

Proof of Proposition 3.6.2: In fact, for $k = 1, \cdots .n$,

$$|\sigma_k^2(2^j) - 2^{j2h_k} g_k^2(0) K(h)| = \left| \int_{\mathbb{R}} |\widehat{\psi}(2^j x)|^2 (|g_k(x)|^2 - |g_k(0)|^2) |x|^{-(2h_k+1)} dx \right|$$

$$\leq \int_{|x|\leq\pi} |\widehat{\psi}(2^j x)|^2 ||g_k(x)|^2 - |g_k(0)|^2| |x|^{-(2h_k+1)} dx + \int_{|x|>\pi} |\widehat{\psi}(2^j x)|^2 ||g_k(x)|^2 - |g_k(0)|^2| |x|^{-(2h_k+1)} dx.$$

By (3.7), we have

$$\int_{|x|\leq\pi} |\widehat{\psi}(2^j x)|^2 ||g_k(x)|^2 - |g_k(0)|^2| |x|^{-(2h_k+1)} dx \leq C \int_{|x|\leq\pi} |\widehat{\psi}(2^j x)|^2 |x|^{\beta} |x|^{-(2h_k+1)} dx$$

$$= C 2^{j(2h_k-\beta)} \int_{|x|\leq 2^j\pi} |\widehat{\psi}(x)|^2 |x|^{-(2h_k+1)+\beta} dx$$

$$\leq C 2^{j(2h_k-\beta)} \int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-(2h_k+1)+\beta} dx. \tag{A.63}$$

By (3.13), the integrand in (A.63) behaves like $|x|^{2N_\psi-(2h_k+1)+\beta}$ around the origin. By (3.12), the integrand is bounded by $|x|^{\beta-2\alpha-2h_k-1}$ as $|x| \to \infty$, where $\beta-2\alpha-2h_k-1 < -1$ as a consequence of (3.8). Thus, $\int_{\mathbb{R}} |\widehat{\psi}(x)|^2 |x|^{-(2h_k+1)+\beta} dx < \infty$ and

$$\int_{|x|\leq\pi} |\widehat{\psi}(2^j x)|^2 ||g_k(x)|^2 - |g_k(0)|^2| |x|^{-(2h_k+1)} dx \leq C 2^{j(2h_k-\beta)}.$$

Moreover, since $g_k(x)$ is bounded and by (3.12),

$$\int_{|x|>\pi} |\widehat{\psi}(2^j x)|^2 ||g_k(x)|^2 - |g_k(0)|^2| |x|^{-(2h_k+1)} dx \leq C 2^{-2j\alpha} \int_{|x|>\pi} |x|^{-(2h_k+1+2\alpha)} dx$$

$$\leq C 2^{j(2h_k-\beta)}.$$

The last inequality holds because $\int_{|x|>\pi} |x|^{-(2h_k+1+2\alpha)} dx < \infty$ and $-2\alpha < 2h_k - \beta$. Consequently,

$$|\sigma_k^2(2^j) - 2^{j2h_k} |g_k(0)|^2 K(h)| < C 2^{j(2h_k-\beta)},$$

as claimed. □

Proof of Theorem 3.6.2: (This proof is follow the same arguments in the proof of Proposition 3 in [46]). By rewriting (3.48), we have

$$\sqrt{\frac{\nu}{a(\nu)}}\left(\begin{pmatrix} a(\nu)^{-2h_1}\widehat{\sigma}^2_{\widehat{X}_1}(a(\nu)2^{j_1}) \\ \vdots \\ a(\nu)^{-2h_1}\widehat{\sigma}^2_{\widehat{X}_1}(a(\nu)2^{j_m}) \\ \vdots \\ a(\nu)^{-2h_n}\widehat{\sigma}^2_{\widehat{X}_n}(a(\nu)2^{j_1}) \\ \vdots \\ a(\nu)^{-2h_n}\widehat{\sigma}^2_{\widehat{X}_n}(a(\nu)2^{j_m}) \end{pmatrix} - \begin{pmatrix} a(\nu)^{-2h_1}\sigma^2_{X_1}(a(\nu)2^{j_1})\mathfrak{D}(1,1)^2 \\ \vdots \\ a(\nu)^{-2h_1}\sigma^2_{X_1}(a(\nu)2^{j_m})\mathfrak{D}(1,1)^2 \\ \vdots \\ a(\nu)^{-2h_n}\sigma^2_{X_n}(a(\nu)2^{j_1})\mathfrak{D}(n,n)^2 \\ \vdots \\ a(\nu)^{-2h_n}\sigma^2_{X_n}(a(\nu)2^{j_m})\mathfrak{D}(n,n)^2 \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}(0,\mathcal{G}), \tag{A.64}$$

where $\widehat{\sigma}^2_{\widehat{X}_i}$ and $\sigma^2_{X_i}$ are defined by (3.49) and (3.50), respectively. The limiting covariance matrix is block diagonal and can be written as $\mathcal{G} = \mathrm{diag}(\mathcal{G}_1, \cdots, \mathcal{G}_n)$. For $i = 1, \cdots, n$, $\mathcal{G}_i$ is a $m \times m$ matrix whose (k_1, k_2)-th entry is given by

$$\mathcal{G}_i(k_1,k_2) = 4\pi b^{4h_i+1}_{j_{k_1},j_{k_2}} |g_i(0)|^2 \mathfrak{D}(i,i)^2 \int_{\mathbb{R}} x^{-(4h_i+2)} |\widehat{\psi}(2^{j_{k_1}}x/b_{j_{k_1},j_{k_2}})|^2 |\widehat{\psi}(2^{j_{k_2}}x/b_{j_{k_1},j_{k_2}}))|^2 dx,$$

where $b_{j_{k_1},j_{k_2}} = \gcd(2^{j_{k_1}}, 2^{j_{k_2}})$ for $i = 1\cdots,n,\ k_1,k_2 = 1\cdots,m$. However, under condition (3.46), relation (3.54) implies that

$$\sqrt{\nu/a(\nu)}a^{-2h_i}|\sigma^2_{X_k}(a(\nu)2^j) - |g_i(0)|^2 K(h_i)(a(\nu)2^j)^{2h_i}|$$

$$\le C\sqrt{\nu/a(\nu)}a^{-2h_i}a^{2h_i-\beta}2^{2jh_i-\beta} \le C\sqrt{\nu/a(\nu)}a^{-\beta} \to 0, \quad \nu \to \infty, \tag{A.65}$$

for $i = 1, \cdots, n$. As a consequence of (A.64) and (A.65),

$$\sqrt{\frac{\nu}{a(\nu)}}\left(\begin{pmatrix} a(\nu)^{-2h_1}\widehat{\sigma}^2_{\widehat{X}_1}(a(\nu)2^{j_1}) \\ \vdots \\ a(\nu)^{-2h_1}\widehat{\sigma}^2_{\widehat{X}_1}(a(\nu)2^{j_m}) \\ \vdots \\ a(\nu)^{-2h_n}\widehat{\sigma}^2_{\widehat{X}_n}(a(\nu)2^{j_1}) \\ \vdots \\ a(\nu)^{-2h_n}\widehat{\sigma}^2_{\widehat{X}_n}(a(\nu)2^{j_m}) \end{pmatrix} - \begin{pmatrix} |g_1(0)|^2K(h_1)2^{2j_1h_1}\mathfrak{D}(1,1)^2 \\ \vdots \\ |g_1(0)|^2K(h_1)2^{2j_mh_1}\mathfrak{D}(1,1)^2 \\ \vdots \\ |g_n(0)|^2K(h_n)2^{2j_1h_n}\mathfrak{D}(n,n)^2 \\ \vdots \\ |g_n(0)|^2K(h_n)2^{2j_mh_n}\mathfrak{D}(n,n)^2 \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}(0,\mathcal{G}). \tag{A.66}$$

Define

$$f(\mathbf{x}) = \left(\sum_{k=1}^{m} w_k^1 \log(x_{1k}), \cdots, \sum_{k=1}^{m} w_k^n \log(x_{nk})\right)^T,$$

for $\mathbf{x} = (x_{11}, \cdots, x_{1m}; \cdots; x_{n1}, \cdots, x_{nm})^T \in \mathbb{R}_+^{nm}$ and $\mathbf{w}^i$ as in (3.52), $i = 1, \cdots, n$. Let $\mathbf{y}_\nu$ and $\mathbf{y}_0$ be the left and right vectors in the difference between parentheses on the left-hand side of (A.66). Then, $f(\mathbf{y}_\nu) = (\widehat{h}_1, \cdots, \widehat{h}_n)$ and $f(\mathbf{y}_0) = (h_1, \cdots, h_n)$. By (A.66) and the Delta method,

$$\sqrt{\nu/a(\nu)}\left[\begin{pmatrix} \widehat{h}_1 \\ \vdots \\ \widehat{h}_n \end{pmatrix} - \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}\right] \xrightarrow{d} \mathcal{N}(\mathbf{0}, \nabla f(\mathbf{y}_0)\mathcal{G}\nabla f(\mathbf{y}_0)^T),$$

where

$$\nabla f(\mathbf{y}_0) = \text{diag}(\mathcal{A}_1, \cdots, \mathcal{A}_n),$$

and

$$\mathcal{A}_i = \left(\frac{w_1^i}{|g_i(0)|^2K(h_i)2^{2j_1h_i}\mathfrak{D}(i,i)^2}, \cdots, \frac{w_m^i}{|g_i(0)|^2K(h_i)2^{2j_mh_i}\mathfrak{D}(i,i)^2}\right), \quad i = 1, \cdots, n.$$

□

A.4 Proofs for Chapter 4

PROOF OF PROPOSITION 4.1.1: Let $\widetilde{Y}_t = \sum_{l=-\infty}^{\infty} Y_l \varphi(t-l)$. Then, $\widetilde{D}(2^j, k) = 2^{-j/2}\int_{\mathbb{R}} \widetilde{Y}_t \psi(2^{-j}t - k)dt$. Therefore,

$$\mathrm{Cov}(\widetilde{D}(2^j,k), \widetilde{D}(2^{j'},k'))$$

$$= 2^{-j}\mathbb{E}\int_{\mathbb{R}}\int_{\mathbb{R}}\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\psi(2^{-j}t-k)\psi(2^{-j'}t'-k')\varphi(t-l)\varphi(t'-l')Y_lY_{l'}^* dtdt'$$

$$= 2^{-j}\mathbb{E}\int_{\mathbb{R}}\int_{\mathbb{R}}\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\psi(2^{-j}t)\psi(2^{-j'}t')\varphi(t+l)\varphi(t'+l')Y_{2^jk-l}Y_{2^{j'}k'-l'}^* dtdt'.$$

By (4.2)

$$\mathrm{Cov}(\widetilde{D}(j,k), \widetilde{D}(j',k'))$$

$$= 2^{-j}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\psi(2^{-j}t)\psi(2^{-j'}t')\varphi(t+l)\varphi(t'+l')$$

$$e^{i(2^jk-l)x}e^{-i(2^{j'}k'-l')}|x|^{-(H+I/2)}G(x)|x|^{-(H+I/2)^*}dtdt'dx,$$

$$= \int_{\mathbb{R}} H_j(x)\overline{H_{j'}(x)}e^{ix(2^jk-2^{j'}k')}|x|^{-(H+I/2)}G(x)|x|^{-(H+I/2)^*}dx, \tag{A.67}$$

where $G(x)$ and $H_j(x)$ are defined in (3.19) and (4.3), respectively. Note that, by Proposition 3 in [22],

$$|H_j(x)| = O(|x|^{N_\psi}), \quad x \to 0, \tag{A.68}$$

$$|H_j(x)| = O(|x|^{\alpha}), \quad x \to \infty, \tag{A.69}$$

so the integral on the right-hand side of (A.67) is finite. □

PROOF OF PROPOSITION 4.1.2: Following the same argument in the proof of Proposition 3.3.1, we only need to show $\|\widetilde{\Phi}_z\|^2$ is summable, where

$$\widetilde{\Phi}_z := \int_{\mathbb{R}} H_j(x)\overline{H_{j'}(x)}e^{ixz}|x|^{-(H-I/2)}G(x)|x|^{-(H-I/2)^*}dx.$$

Since

$$\|\widetilde{\Phi}_z\| = \left\| P\int_{\mathbb{R}} e^{izx}\overline{H_{j'}(x)}H_j(x)\mathrm{diag}(x^{-(2h_1+1)}g_1^2(x),\cdots,x^{-(2h_n+1)}g_n^2(x))dxP^*\right\|^2$$

$$\leq C\|P\|^4 \max_{1\leq k\leq n}\left|\int_{\mathbb{R}} e^{izx}\overline{H_{j'}(x)}H_j(x)(2^{j'}x)x^{-(2h_k+1)}dx\right|^2.$$

Moreover, for any $1 \leq k \leq n$, by (A.68) and (A.69), $\overline{H_{j'}(x)}H_j(x)x^{-(2h_k+1)}g_k^2(x) \in L^2(\mathbb{R})$. Thus, by Parseval's Theorem,

$$\sum_{z=-\infty}^{\infty}\left|\int_{\mathbb{R}} e^{izx}\overline{H_{j'}(x)}H_j(x)x^{-(2h_k+1)}g_k^2(x)dx\right|^2$$

$$= \int_{\mathbb{R}}\left|\overline{H_{j'}(x)}H_j(x)x^{-(2h_k+1)}g_k^2(x)\right|^2 dx < \infty,$$

so $\|\widetilde{\Phi}_z\|^2$ is summable. Therefore, (4.5) and (4.7) hold. □

Proof of Proposition 4.1.3: Adapting the argument for the proof of Proposition 3.6.1 by replacing $\widehat{X}$ by $\widetilde{X}$, the limiting distribution of

$$\left(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1},\cdots,a(\nu)^{-2h_n})(\mathrm{vec}_{\mathcal{D}}(\widetilde{W}_{\widetilde{X}}(a(\nu)2^j)-\widetilde{\mathfrak{D}}\mathbb{E}\widetilde{W}_X(a(\nu)2^j)\widetilde{\mathfrak{D}}))^T\right)_{j=j_1,\cdots,j_m}$$

is equal to the limiting distribution of

$$\widetilde{\mathcal{K}}\left(\sqrt{\nu/a(\nu)}\mathrm{diag}(a(\nu)^{-2h_1},\cdots,a(\nu)^{-2h_n})\mathrm{vec}_{\mathcal{D}}(\widetilde{W}_X(a(\nu)2^j)-\mathbb{E}\widetilde{W}_X(a(\nu)2^j))\right)_{j=j_1,\cdots,j_m}.$$

By (2.10), the spectral density of the l-th component of X is

$$f_l(x) = \sum_{k=-\infty}^{\infty} |x + 2k\pi|^{-2d_l} g_l^2(x + 2k\pi), \quad x \in (-\pi, \pi),$$

where $d_l = h_l + 1/2$, $l = 1, \cdots, n$. Reexpress f_l as

$$f_l(x) = |1 - e^{-ix}|^{-2d_l} f_l^*(x),$$

where

$$f_l^*(x) = \left|\frac{2\sin(x/2)}{\lambda}\right|^{2d_l} g_l^2(x) + |2\sin(x/2)|^{2d_l} \sum_{k\neq 0} |x + 2k\pi|^{-2d_l} g_l^2(x + 2k\pi). \quad \text{(A.70)}$$

Then, $f_l^*(0) = g_l^2(0)$, and

$$|f_l^*(x) - f_l^*(0)|$$

$$\leq |g_l^2(x)| \left| \left|\frac{2\sin(x/2)}{x}\right|^{2d_l} - 1 \right| + |g_l^2(x) - g_l^2(0)| + \left| 2\sin(x/2)^{2d_l} \sum_{k\neq 0} |x+2k\pi|^{-2d_l} g_l^2(x+2k\pi) \right|$$

$$= O(|x|^2) + O(|x|^\beta) + O(|x|^{2d_l}), \quad x \to 0.$$

So, $|f_l^*(x) - f_l^*(0)| < C|x|^{\beta_*}$, where

$$\beta_* = \min\{2, \beta, 2h_1 + 1\}. \quad \text{(A.71)}$$

Thus, by Theorem 2 in [46],

$$\left(\sqrt{\nu/a(\nu)}a(\nu)^{2h_l}(\widetilde{W}_{X_l}(a(\nu)2^j) - \mathbb{E}\widetilde{W}_{X_l}(a(\nu)2^j))\right)_{j=j_1,\cdots,j_m} \xrightarrow{d} \mathcal{N}(\mathbf{0}, W(d_l)).$$

The (i,k)-th entry of the limiting covariance matrix is given by

$$W_{i,k}(d_l) = 4\pi(f_l^*(0))^2 2^{4d_l \max(j_i,j_k)+\min(j_i,j_k)} \int_{-\pi}^{\pi} |D_{|j_i-j_k|}(\lambda; d_l)|^2 d\lambda, \quad i,k = 1, \cdots, m,$$

where $D_{|j_i - j_k|}(\lambda; d_l)$ is defined in (4.14). Moreover, the entries of X are independent, thus (4.13) holds. □

Proof of Theorem 4.1.3: The proof can be written as a direct adaption of Theorem 3.6.2 by using Theorem 1 in [22] as the counterpart of Proposition 3.6.2. □

A.5 Proof for Chapter 5

Proof of Theorem 5.2.1: Let $d_l(2^j, k)$ be the wavelet coefficient of each R_{h_l} in (5.3), $l = 1, \cdots, n$. Then, the sample wavelet variance of Y can be written as

$$W(2^j) = P\bigg(\frac{1}{K_j}\sum_{k=1}^{K_j} d_i(2^j,k)d_l(2^j,k)\bigg)_{i,l=1,\cdots,n} P^*, \quad K_j = \frac{\nu}{2^j}.$$

By Theorem 3 in [39],

$$\mathrm{Var}\bigg(\frac{1}{K_j}\sum_{k=1}^{K_j} d_l^2(2^j,k)\bigg) = O(\nu^{2h_l-2}), \quad l = 1, \cdots, n.$$

Then, by the Cauchy-Schwarz inequality,

$$\mathrm{Var}\bigg(\frac{1}{K_j}\sum_{k=1}^{K_j} d_i(2^j,k)d_l(2^j,k)\bigg)$$

$$\leq \sqrt{\mathrm{Var}\bigg(\frac{1}{K_j}\sum_{k=1}^{K_j} d_l^2(2^j,k)\bigg)\mathrm{Var}\bigg(\frac{1}{K_j}\sum_{k=1}^{K_j} d_i^2(2^j,k)\bigg)}$$

$$= O(\nu^{(h_i+h_l)-2}).$$

Thus,

$$\nu^{1-h_n}\bigg((W(2^j) - \mathbb{E}W(2^j))\bigg)_{j=j_1,\cdots,j_m}$$

$$= \left(P \begin{pmatrix} O_p(\nu^{h_1-h_n}) & \cdots & O_p(\nu^{(h_1-h_n)/2}) \\ \vdots & \ddots & \vdots \\ O_p(\nu^{(h_1-h_n)/2}) & \cdots & \nu^{1-h_n}((1/K_j)\sum_{k=1}^{K_j} d_n^2(2^j,k) - \sigma^2_{R_n}(2^j)) \end{pmatrix} P^* \right)_{j=j_1,\cdots,j_m}, \tag{A.72}$$

where $\sigma^2_{R_n}(2^j)$ is the n-th diagonal term of $\mathbb{E}W(2^j)$, i.e., the univariate wavelet variance of the process R_n. By (A.72) and Theorem 4 in [39], (5.10) holds. $\square$

PROOF OF PROPOSITION 5.3.1: Since $\widehat{R} = \widehat{I}_\nu \mathfrak{D} R$, then,

$$W_{\widehat{R}}(a(\nu)2^j) = \widehat{I}_\nu \mathfrak{D} W_R(a(\nu)2^j)\mathfrak{D}\widehat{I}^*_\nu.$$

Thus,

$$W_{\widehat{R}}(a(\nu)2^j) - \mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D} = \widehat{I}_\nu \mathfrak{D} W_R(a(\nu)2^j)\mathfrak{D}\widehat{I}^*_\nu - \mathfrak{D}\mathbb{E}W_R(a(\nu)2^j))\mathfrak{D}$$

$$= \widehat{I}_\nu \Bigg\{ \Big[\mathfrak{D}(W_R(a(\nu)2^j) - \mathbb{E}W_R(a(\nu)2^j))\mathfrak{D}\Big] - \Big[(\widehat{I}_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}\Big]$$

$$-\Big[\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}\Big((\widehat{I}_\nu^{-1})^* - I\Big)\Big]$$

$$- \Big[(\widehat{I}_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}\Big((\widehat{I}_\nu^{-1})^* - I\Big)\Big]\Bigg\}\widehat{I}^*_\nu. \tag{A.73}$$

Recall that the operator $\text{vec}_{\mathcal{D}} W_R(a(\nu)2^j)$ picks out the diagonal entries of the matrix $W_R(a(\nu)2^j)$, which are independent. Therefore, by Theorem 4 in [39] for univariate processes,

$$\Big(\text{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\text{vec}_{\mathcal{D}}(\mathfrak{D}(W_R(a(\nu)2^j) - \mathbb{E}W_R(a(\nu)2^j))\mathfrak{D})^T\Big)_{j=j_1,\cdots,j_m}$$

$$= \mathcal{K}\Big(\text{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot\left(\mathrm{vec}_{\mathcal{D}}(W_R(a(\nu)2^j) - \mathbb{E}W_R(a(\nu)2^j))^T\right)_{j=j_1,\cdots,j_m} \xrightarrow{d} \mathcal{K}\mathcal{G}, \tag{A.74}$$

as $\nu \to \infty$. In (A.79), $\mathcal{K}$ and $\mathcal{G}$ are defined in (5.22) and (5.23), respectively. Since $\mathrm{Var}(\widehat{I}_\nu^{-1} - I) = O(\nu^{2h_n-2})$, then

$$\Big(\mathrm{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\mathrm{vec}_{\mathcal{D}}((\widehat{I}_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D})\Big)^T$$

$$= \mathfrak{D}^2\Big(\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})\mathrm{diag}(\mathbb{E}W_{R_{h_1}}(a(\nu)2^j), \cdots, \mathbb{E}W_{R_{h_n}}(a(\nu)2^j))\Big)$$

$$\cdot\mathrm{diag}\left(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\right) \cdot \left(\nu^{1-h_n}\mathrm{vec}_{\mathcal{D}}(\widehat{I}_\nu^{-1} - I))^T\right)$$

$$= O_P\left(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\right). \tag{A.75}$$

Similarly,

$$\Big(\mathrm{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\mathrm{vec}_{\mathcal{D}}(\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}((\widehat{I}_\nu^{-1} - I)^*))\Big)^T$$

$$= O_P\left(\left(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\right)^T\right), \tag{A.76}$$

and

$$\Big(\mathrm{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\mathrm{vec}_{\mathcal{D}}((I_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}(\widehat{I}_\nu^{-1} - I)^*)\Big)^T$$

$$= O_p\left(a(\nu)^{h_n-1}\left(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\right)^T\right). \tag{A.77}$$

Consequently, by condition (5.20), (A.79)-(A.82), and Slutsky's Theorem, (5.21) holds. $\square$

Proof of Proposition 5.3.1: Proof of Proposition 5.3.1: Since $\widehat{R} = \widehat{I}_\nu \mathfrak{D} R$, then,

$$W_{\widehat{R}}(a(\nu)2^j) = \widehat{I}_\nu \mathfrak{D} W_R(a(\nu)2^j)\mathfrak{D}\widehat{I}_\nu^*.$$

Thus,

$$W_{\widehat{R}}(a(\nu)2^j) - \mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D} = \widehat{I}_\nu \mathfrak{D} W_R(a(\nu)2^j)\mathfrak{D}\widehat{I}_\nu^* - \mathfrak{D}\mathbb{E}W_R(a(\nu)2^j))\mathfrak{D}$$

$$= \widehat{I}_\nu \Big\{ \Big[\mathfrak{D}(W_R(a(\nu)2^j) - \mathbb{E}W_R(a(\nu)2^j))\mathfrak{D}\Big] - \Big[(\widehat{I}_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}\Big]$$

$$-\Big[\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}\Big((\widehat{I}_\nu^{-1})^* - I\Big)\Big]$$

$$- \Big[(\widehat{I}_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}\Big((\widehat{I}_\nu^{-1})^* - I\Big)\Big]\Big\}\widehat{I}_\nu^*. \tag{A.78}$$

Recall that the operator $\text{vec}_{\mathcal{D}} W_R(a(\nu)2^j)$ picks out the diagonal entries of the matrix $W_R(a(\nu)2^j)$, which are independent. Therefore, by Theorem 4 in [39] for univariate processes,

$$\Big(\text{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\text{vec}_{\mathcal{D}}(\mathfrak{D}(W_R(a(\nu)2^j) - \mathbb{E}W_R(a(\nu)2^j))\mathfrak{D})^T\Big)_{j=j_1,\cdots,j_m}$$

$$= \mathcal{K}\Big(\text{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot (\text{vec}_{\mathcal{D}}(W_R(a(\nu)2^j) - \mathbb{E}W_R(a(\nu)2^j))^T\Big)_{j=j_1,\cdots,j_m} \xrightarrow{d} \mathcal{K}\mathcal{G}, \tag{A.79}$$

as $\nu \to \infty$. In (A.79), $\mathcal{K}$ and $\mathcal{G}$ are defined in (5.22) and (5.23), respectively. Since $\text{Var}(\widehat{I}_\nu^{-1} - I) = O(\nu^{2h_n - 2})$, then

$$\Big(\text{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\mathrm{vec}_{\mathcal{D}}((\widehat{I}_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D})\Big)^T$$

$$= \mathfrak{D}^2\Big(\mathrm{diag}(a(\nu)^{-2h_1}, \cdots, a(\nu)^{-2h_n})\mathrm{diag}(\mathbb{E}W_{R_{h_1}}(a(\nu)2^j), \cdots, \mathbb{E}W_{R_{h_n}}(a(\nu)2^j))\Big)$$

$$\cdot\mathrm{diag}\Big(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\Big) \cdot \Big(\nu^{1-h_n}\mathrm{vec}_{\mathcal{D}}(\widehat{I}_\nu^{-1} - I))^T\Big)$$

$$= O_P\Big(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\Big). \tag{A.80}$$

Similarly,

$$\Big(\mathrm{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\mathrm{vec}_{\mathcal{D}}(\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}((\widehat{I}_\nu^{-1} - I)^*))\Big)^T$$

$$= O_P\Big(\Big(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\Big)^T\Big), \tag{A.81}$$

and

$$\Big(\mathrm{diag}((\nu/a(\nu))^{1-h_1}a(\nu)^{-2h_1}, \cdots, (\nu/a(\nu))^{1-h_n}a(\nu)^{-2h_n})$$

$$\cdot(\mathrm{vec}_{\mathcal{D}}((I_\nu^{-1} - I)\mathfrak{D}\mathbb{E}W_R(a(\nu)2^j)\mathfrak{D}(\widehat{I}_\nu^{-1} - I)^*)\Big)^T$$

$$= O_p\Big(a(\nu)^{h_n-1}\Big(\frac{\nu^{h_n-h_1}}{a(\nu)^{1-h_1}}, \cdots, \frac{\nu^{h_n-h_n}}{a(\nu)^{1-h_n}}\Big)^T\Big). \tag{A.82}$$

Consequently, by condition (5.20), (A.79)-(A.82), and Slutsky's Theorem, (5.21) holds. □

PROOF OF THEOREM 5.3.3: By rewriting (5.21), we have

$$\Big(\frac{\nu}{a(\nu)}\Big)^{1-h_i}\left(\begin{pmatrix} a(\nu)^{-2h_i}\widehat{\sigma}^2_{\widehat{R}_i}(a(\nu)2^{j_1}) \\ \vdots \\ a(\nu)^{-2h_i}\widehat{\sigma}^2_{\widehat{R}_i}(a(\nu)2^{j_m}) \end{pmatrix} - \begin{pmatrix} \mathfrak{D}(i,i)^2 C_\psi(h_i)2^{2j_1h_1} \\ \vdots \\ \mathfrak{D}(i,i)^2 C_\psi(h_i)2^{2j_1h_1} \end{pmatrix}\right)$$

$$\xrightarrow{d} \begin{pmatrix} \mathfrak{D}(i,i)^2 C_\psi(h_i) 2^{2j_1 h_1} S(h_i) R_{j_1}^{h_i} \\ \vdots \\ \mathfrak{D}(i,i)^2 C_\psi(h_i) 2^{2j_m h_1} S(h_i) R_{j_m}^{h_i} \end{pmatrix}, \tag{A.83}$$

for $i = 1, \cdots, n$. Define

$$f_i(\mathbf{x}) = \sum_{k=1}^{m} w_k^i \log(x_k),$$

for $\mathbf{x} = (x_1, \cdots, x_m)^T \in \mathbb{R}_+^m$ and $\mathbf{w}^i$ as in (5.27), $i = 1, \cdots, n$. Let $\mathbf{y}_\nu^i$ and $\mathbf{y}_0^i$ be the left and right vectors in the difference between parentheses on the left-hand side of (A.83). Then, by (5.28), $f_i(\mathbf{y}_\nu^i) = \widehat{h}_i$ and $f_i(\mathbf{y}_0^i) = h_i$, $i = 1 \cdots, n$. By (A.83) and the Delta method,

$$\left(\frac{\nu}{a(\nu)}\right)^{1-h_i} (\widehat{h}_i - h_i) \xrightarrow{d} \nabla f_i(\mathbf{y}_0^i) \cdot \begin{pmatrix} \mathfrak{D}(i,i)^2 C_\psi(h_i) 2^{2j_1 h_1} S(h_i) R_{j_1}^{h_i} \\ \vdots \\ \mathfrak{D}(i,i)^2 C_\psi(h_i) 2^{2j_m h_1} S(h_i) R_{j_m}^{h_i} \end{pmatrix}$$

$$\stackrel{d}{=} S(h_i) \sum_{k=1}^{m} w_k^i R_{j_k}^{h_i},$$

for $i = 1, \cdots, n$. This establishes (5.29). $\square$

A.6 Useful lemmas and theorems

Lemma A.6.1. *([23]) Let $\{\xi_n, n \geq 1\}$ be a sequence of centered Gaussian vectors and let Γ_n be the covariance matrix of ξ_n. Let $(A_n)_{n\geq 1}$ be a sequence of deterministic matrices with adapted dimensions such that*

$$\lim_{n\to\infty} \mathrm{Var}(\xi_n^T A_n \xi_n) = \sigma^2 \in [0, \infty].$$

Assume that

$$\lim_{n\to\infty} \rho(A_n)\rho(\Gamma_n) = 0,$$

where $\rho(\cdot)$ denotes the spectral radius. Then

$$\xi_n^T A_n \xi_n - E(\xi_n^T A_n \xi_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

Lemma A.6.2. *([46]) Let $m \geq 2$ be an integer and Γ be a $m \times m$ covariance matrix. Let p be an integer between 1 and $m-1$. let Γ_1 be the top left submatrix with size $p \times p$ and Γ_2 the bottom right submatrix with size $(m-p) \times (m-p)$. Then*

$$\rho(\Gamma) \leq \rho(\Gamma_1) + \rho(\Gamma_2).$$

Lemma A.6.3. *([46]) Let $\{\xi_k, k \in \mathbb{Z}\}$ be a stationary process with spectral density function f and let Γ_n be the covariance matrix of $(\xi_1, \cdots, \xi_n)$. Then, $\rho(\Gamma_n) \leq 2\pi \parallel f \parallel_\infty$.*

The following theorem provides the partial derivatives of the eigenvalues and eigenvectors of a symmetric matrix with respect to the latter.

Theorem A.6.1. *([47], Theorem 1) Let $S_0 \in \mathcal{S}(n, \mathbb{R})$, and let u_0 be a normalized eigenvector associated with a simple eigenvalue λ_0 of S_0. Then, we can define a real-valued and a vector function λ and u, respectively, for all symmetric matrix S in some neighborhood $N(S_0) \in \mathcal{S}(n, \mathbb{R})$ of S_0, where*

$$\lambda(S_0) = \lambda_0, \qquad u(S_0) = u_0,$$

and

$$Su = \lambda u, \qquad u^T u = 1, \qquad S \in \mathcal{S}(n, \mathbb{R}).$$

Moreover, the functions λ and u are infinitely differentiable on $N(S_0)$, and their *differentials at S_0 are given by*

$$\frac{\partial \lambda}{\partial [\mathrm{vec}_{\mathcal{S}}(S)]} = (u_0^T \otimes u_0^T)D, \quad \frac{\partial u}{\partial [\mathrm{vec}_{\mathcal{S}}(S)]} = [u_0^T \otimes (\lambda_0 I_n - S_0)^+]D. \tag{A.84}$$

In (A.84), the symbol $\otimes$ and the superscript $+$ denote the Kronecker product and the Moore-Penrose inverse, respectively, and D is the duplication matrix defined by (3.30).

Lemma A.6.4. *([43], Lemma B.3) Let $\{\phi_\cdot\} \in \mathbb{R}$ be a sequence such that $\sum_{z=-\infty}^{\infty} |\phi_{z\mathrm{gcd}(a_j,a_{j'})}| < \infty$. Then,*

$$\frac{1}{\nu} \sum_{k=1}^{a_{j'}\nu} \sum_{k'=1}^{a_j \nu} \phi_{a_j k - a_{j'} k'} \to \mathrm{gcd}(a_j, a_{j'}) \sum_{z=-\infty}^{\infty} \phi_{z\mathrm{gcd}(a_j,a_{j'})}, \quad \nu \to \infty.$$

A.7 Empirical study of diffusion data

This section consists a study of bivariate diffusion data provided by the David B. Hill Lab (UNC-Chapel Hill). The main goal is to verify whether the diffusion data have different diffusion exponents along different directions. To achieve this goal, the tools we use are the two-step wavelet-based method and the wavelet eigen-structure method developed in Chapter 3 and [43], respectively. The statistical analysis covers COS2-NO data set. The data set alternates between the abscissa and ordinate coordinates that make up the two-dimensional sample path of each microparticle under the microscope.

There are mainly two kinds of data paths: diffusive and quasi-stationary (Figure A.1). The range of quasi-stationary paths are much narrower than the diffusive ones, which suggests that the former particles are trapped by some physical barrier. In this section, we focus on the diffusive sample paths, which can be models as stationary increment processes. 50 sample paths were randomly selected for each of the 4 concentration levels: 2,4,8,16 mg ml^{-1}, Table A.1 and Table A.2 show the estimates of

h_1, h_2 and $h_2 - h_1$ by means of the two-step wavelet-based method and the wavelet eigen structure method, respectively. Based on the Monte Carlo simulation studies shown in Table A.3-A.4, we cannot reject that $h_1 = h_2 = h$ for each concentration level of COS2-NO diffusion data.

COS2-NO (mg ml^{-1})	$\overline{\widehat{h}_1}$	$\overline{\widehat{h}_2}$	$\overline{\widehat{h}_2 - \widehat{h}_1}$
2	0.27988	0.33198	0.052097
4	0.17661	0.22743	0.050821
8	0.22266	0.31085	0.088192
16	0.4513	0.51841	0.067112

Table A.1: Two-step wavelet-based estimation for Hurst parameters of bivariate diffusion data. 50 independent paths of length 1800 are randomly selected from each group.

COS2-NO (mg ml^{-1})	$\overline{\widehat{h}_1}$	$\overline{\widehat{h}_2}$	$\overline{\widehat{h}_2 - \widehat{h}_1}$
2	0.22493	0.33439	0.10946
4	0.1066	0.17659	0.069991
8	0.20289	0.28833	0.085443
16	0.40256	0.51271	0.11014

Table A.2: Wavelet eigen structure estimation for Hurst parameters of bivariate diffusion data. 50 independent paths of length 1800 are randomly selected from each group.

true h	$\overline{\widehat{h}_1}$	$\overline{\widehat{h}_2}$	$\overline{\widehat{h}_2 - \widehat{h}_1}$	sd($\widehat{h}_2 - \widehat{h}_1$)
0.2	0.12818	0.23181	0.10364	0.081683
0.5	0.4458	0.5532	0.1075	0.0837

Table A.3: Two-step wavelet-based estimation for the Hurst parameters of simulated bivariate operator fractional Brownian motion with $h_1 = h_2 = h$. Sample size=1800, number of MC runs=1000.

true h	$\overline{\widehat{h}_1}$	$\overline{\widehat{h}_2}$	$\overline{\widehat{h}_2 - \widehat{h}_1}$	$\widehat{\mathrm{sd}(\widehat{h}_2 - \widehat{h}_1)}$
0.2	0.067089	0.19416	0.12707	0.093029
0.5	0.37157	0.5060	0.13443	0.11003

Table A.4: Wavelet eigen structure estimation for the Hurst parameters of simulated bivariate operator fractional Brownian motion with $h_1 = h_2 = h$. Sample size=1800, number of MC runs=1000.

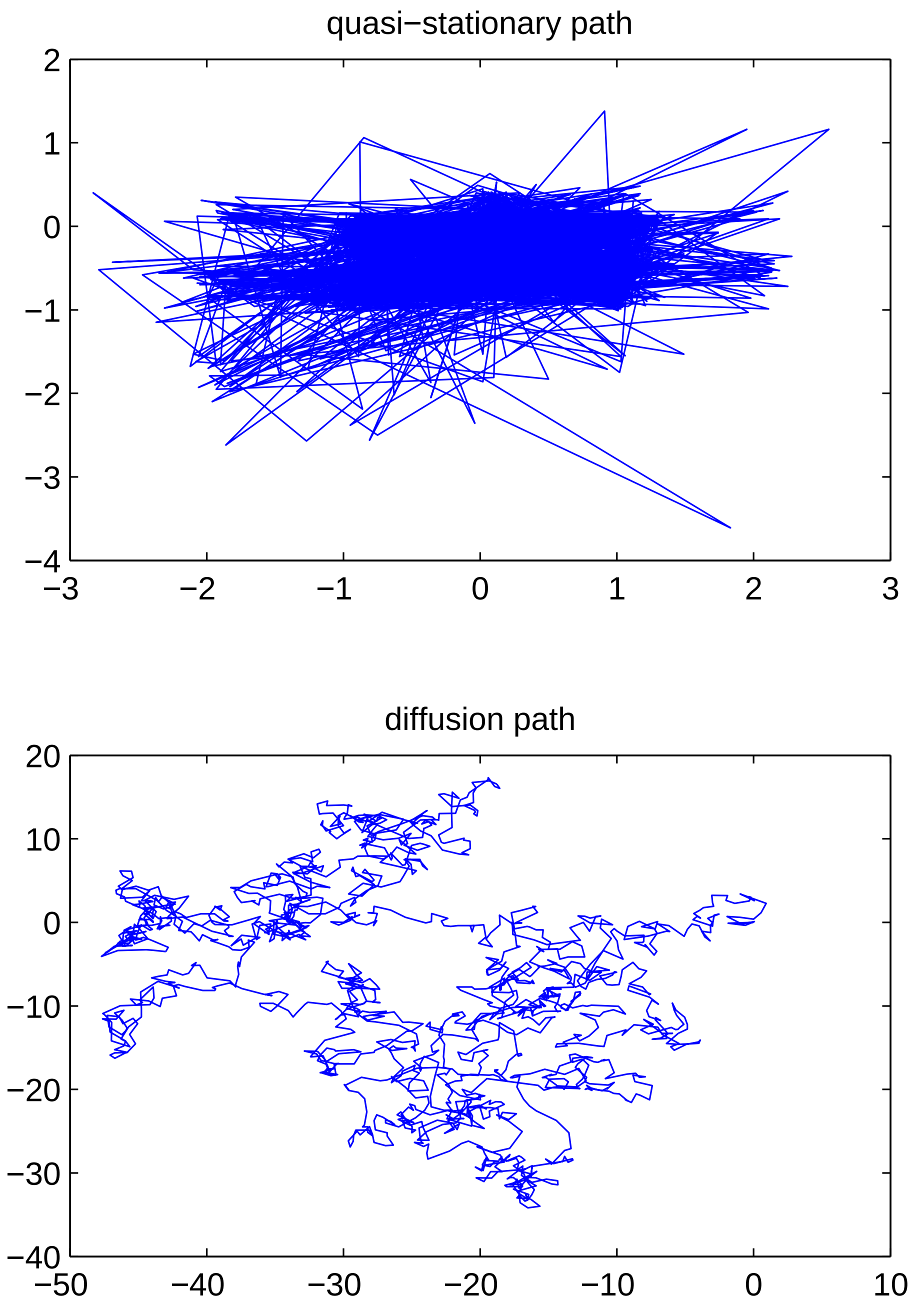

Figure A.1: Two kinds of diffusion sample paths.

Bibliography

[1] B. Mandelbrot. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, 62:331–358, 1974.

[2] E. Foufoula-Georgiou and P. Kumar. *Wavelets in Geophysics.* Academic Press, San Diego, Calif., 1994.

[3] P. Ivanov, L. Nunes Amaral, A. Goldberger, S. Havlin, M. Rosenblum, Z. Struzik, and H. Stanley. Multifractality in human hearbeat dynamics. *Narure*, 399:461–465, 1999.

[4] P. Ciuciu, P. Abry, and B. He. Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks. *NeuroImage*, 95(186):246–263, 2014.

[5] M. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *SIGCOMM Comput. Commun. Rev.*, 27(2):5–23, 1997.

[6] T. H., R. H., and C. K.S. Inference of bivariate long-memory aggregate time series. *Statistica Sinica*, 2016.

[7] G. Didier and V. Pipiras. Integral representations and properties of operator fractional Brownian motions. *Bernoulli*, 17(1):1–33, 2011.

[8] W. Hudson and J. Mason. Operator-self-similar processes in a finite-dimensional space. *Transactions of the American Mathematical Society*, 273(1):281–297, 1982.

[9] Z. Jurek and J. Mason. *Operator-Limit Distributions in Probability Theory.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, NY, 1993.

[10] J. Stone. *Independent Component Analysis: a Tutorial Introduction.* Cambridge, MA; London: MIT Press, 2004.

[11] P. Comon and C. Jutten. *Handbook of Blind Source Separation: Independent Component Analysis and Applications.* Academic Press, 2010.

[12] M. S. Pinsker. The theory of curves in Hilbert space with stationary nth increments. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 19:319–344, 1955.

[13] M. S. Pinsker and A. M. Yaglom. Random processes with stationary increments of the nth order. *Dokl. Akad. Nauk. SSSR*, 94:385–388, 1954.

[14] A. M. Yaglom. Correlation theory of processes with random nth increments. *Amer. Math. Soc. Transl.*, 8(2):87–141, 1958.

[15] G. Wornell and A. Oppenheim. Estimation of fractal signals from noisy measurements using wavelets. *IEEE Transactions on Signal Processing*, 40(3):611–623, 1992.

[16] P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*, 38:910 – 917, March 1992.

[17] D. Veitch and P. Abry. A wavelet-based joint estimator of the parameters of long–range dependence. *IEEE Transactions on Information Theory*, 45(3):878–897, April 1999.

[18] M. S. Taqqu. Fractional Brownian motion and long range dependence. In *Theory and Applications of Long-Range Dependence (P. Doukhan, G. Oppenheim and M. S. Taqqu, eds.)*, pages 5–38. Birkhäuser, Boston, 2003.

[19] G. Didier and K. Zhang. The asymptotic distribution of the pathwise mean squared displacement in single particle tracking experiments. *To appear in Journal of Time Series Analysis*, 2016.

[20] P. Robinson. Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23(5):1630–1661, 1995.

[21] P. Robinson. Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, pages 1048–1072, 1995.

[22] E. Moulines, F. Roueff, and M. Taqqu. On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter. *Journal of Time Series Analysis*, 28(2):155–187, March 2007.

[23] E. Moulines, F. Roueff, and M. Taqqu. A wavelet whittle estimator of the memory parameter of a nonstationary Gaussian time series. *Annals of Statistics*, pages 1925–1956, 2008.

[24] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[25] C. Vignat. A generalized Isserlis theorem for location mixtures of Gaussian random vectors. *Statistics and Probability Letters*, 82(1):67–71, 2012.

[26] J. Magnus and H. Neudecker. The elimination matrix: some lemmas and applications. *SIAM Journal on Algebraic Discrete Methods*, 1(4):422–449, 1980.

[27] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[28] S. Combrexelle, G. Didier, H. Wendt, J.-Y. Tourneret, and P. Abry. Multivariate Hadamard self-similarity: testing fractal connectivity. *Preprint*, pages 1–40, 2016.

[29] K. Chan and K. Tsai. Inference of bivariate long-memory aggregate time series. *available at* ***http://public.econ.duke.edu/ brossi/NBERNSF/Tsai.pdf***, pages 1–17, 2010.

[30] M. Chambers. The estimation of continuous parameter long-memory time series models. *Econometric Theory*, 12:374–390, 1996.

[31] Y. Hosoya. The quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence. *Journal of Econometrics*, 73:217–236, 1996.

[32] I. Daubechies. *Ten Lectures on Wavelets*, volume 61. Society for Industrial and Applied Mathematics, Philadelphia-PA, 1992.

[33] H. Tsai and C. K.S. Quasi-maximum likelihood estimation for a class of continuous-time long-memory processes. *Journal of Time Series Analysis*, 26(5):691–713, 2005.

[34] J. D. Cryer and K. S. Chan. *Time Series Analysis With Application in R*. New York: Springer, 2008.

[35] P. Robinson. Multivariate Local Whittle estimation in stationary systems. *Annals of Statistics*, 36(5):2508–2530, 2008.

[36] M. Taqqu. Weak convergence to the fractional brownian motion and to the rosenblatt process. *Z. Wahrsch. Verw. Gebiete*, 31:287–302, 1975.

[37] M. Taqqu. Convergence of integrated processes of arbitrary hermite rank. *Z. Wahrsch. Verw. Gebiete*, 50:53–83, 1979.

[38] P. Breuer and P. Major. Central limit theorems for nonlinear functionals of gaussian fields. *Journal of Multivariate Analysis*, pages 425–441, 1983.

[39] B. J.-M. and T. C.A. A wavelet analysis of the rosenblatt process: chaos expansion and estimation of the self-similarity parameter. *Stochastic Processes and their Applications*, 120(12):2331–2362, 2010.

[40] C. M., R. F., T. M.S., and T. C. Wavelet estimation of the long memory parameter for hermite polynomial of gaussian processes. *ESAIM: Probability and Statistics*, 18:42–76, 2014.

[41] T. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley, 3 edition, 2003.

[42] H. Wendt, G. Didier, S. Combrexelle, and P. Abry. Multivariate Hadamard self-similarity: testing fractal connectivity. *available at https://arxiv.org/pdf/1701.04366v1.pdf*, pages 1–51, 2016.

[43] P. Abry and G. Didier. Wavelet estimation for operator fractional Brownian motion. To appear in *Bernoulli*, pages 1–30, 2016.

[44] G. A. Jones and J. M. Jones. *Elementary Number Theory*. Berlin: Springer-Verlag, 1998.

[45] G. Teschl. *Mathematical Methods in Quantum Mechanics*, volume 99. American Mathematical Society, 2009.

[46] E. Moulines, F. Roueff, and M. Taqqu. Central limit theorem for the log-regression wavelet estimation of the memory parameter in the gaussian semi-parametric context. *Fractals*, 15(4):301–313, 2007.

[47] J. Magnus. On differentiating eigenvalues and eigenvectors. *Econometric Theory*, 1(2):179–191, Aug. 1985.

Biography

Hui Li got her Bachelor's degree in Chongqing University. She started PhD program at Tulane University mathematics department in August 2012, eventually completing the program in August 2017 under the supervision of Professor Gustavo Didier.