Wavelet estimation for operator fractional Brownian motion

PATRICE ABRY[1] and GUSTAVO DIDIER[2]

[1]*Physics Lab, CNRS and École Normale Supérieure de Lyon, 46 allée d'Italie, F-69364, Lyon cedex 7, France. E-mail: patrice.abry@ens-lyon.fr*
[2]*Mathematics Department, Tulane University, 6823 St. Charles Avenue, New Orleans, LA 70118, USA. E-mail: gdidier@tulane.edu*

Operator fractional Brownian motion (OFBM) is the natural vector-valued extension of the univariate fractional Brownian motion. Instead of a scalar parameter, the law of an OFBM scales according to a Hurst matrix that affects every component of the process. In this paper, we develop the wavelet analysis of OFBM, as well as a new estimator for the Hurst matrix of bivariate OFBM. For OFBM, the univariate-inspired approach of analyzing the entry-wise behavior of the wavelet spectrum as a function of the (wavelet) scales is fraught with difficulties stemming from mixtures of power laws. Instead we consider the evolution along scales of the eigenstructure of the wavelet spectrum. This is shown to yield consistent and asymptotically normal estimators of the Hurst eigenvalues, and also of the eigenvectors under assumptions. A simulation study is included to demonstrate the good performance of the estimators under finite sample sizes.

Keywords: operator fractional Brownian motion; operator self-similarity; wavelets

1. Introduction

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

$$\{X(ct)\}_{t \in \mathbb{R}} \overset{\mathcal{L}}{=} \{c^H X(t)\}_{t \in \mathbb{R}}, \qquad c > 0, \tag{1.1}$$

where $c^H = \sum_{k=0}^{\infty} \log^k(c) H^k / k!$ and $\overset{\mathcal{L}}{=}$ denotes the equality of finite-dimensional distributions. No specific assumption on the eigenstructure of H is imposed, for example, canonical vectors are not necessarily eigenvectors. The notion of operator self-similarity underpins the natural multivariate generalization of the univariate fractional Brownian motion (FBM): an operator fractional Brownian motion (OFBM) $B_H = \{B_H(t)\}_{t \in \mathbb{R}}$ is a proper Gaussian, o.s.s., stationary increment stochastic process. In this paper, we propose using the wavelet eigenstructure of OFBM to estimate H. The main motivation behind this methodology is to avoid the difficulties stemming from the extrapolation of univariate techniques to a multivariate context, where operator scaling laws such as (1.1) may arise.

Inferential theory for univariate self-similar processes now comprises a voluminous and well-established literature. A non-exhaustive list includes Fox and Taqqu [23] and Robinson [43,44] on Fourier domain methods, and Wornell and Oppenheim [52], Flandrin [22], and Veitch and Abry [48] on wavelet domain methods, among many others. The multivariate framework evokes

several applications where matrix-based scaling laws are expected to appear, such as in long range dependent time series (Marinucci and Robinson [36], Davidson and de Jong [14], Chung [9], Dolado and Marmol [21], Davidson and Hashimzade [15], Kechagias and Pipiras [29]) and queueing systems (Konstantopoulos and Lin [30], Majewski [33,34], Delgado [17]). Like FBM in the univariate setting, OFBM is a natural starting point in the construction of estimators for operator self-similar processes due to its tight connection to stationary fractional processes and its being Gaussian (on the general theory of o.s.s. processes, see Laha and Rohatgi [31], Hudson and Mason [27], Maejima and Mason [32], Cohen *et al.* [11]).

A characterization of the covariance structure of OFBM can be derived from stochastic integral representations. Under a mild condition on the eigenvalues of the exponent H (see (2.9)), Didier and Pipiras [19] showed that any OFBM B_H admits a harmonizable representation

$$\{B_H(t)\}_{t\in\mathbb{R}} \stackrel{\mathcal{L}}{=} \left\{\int_{\mathbb{R}} \frac{e^{itx}-1}{ix}\left(x_+^{-D}A + x_-^{-D}\overline{A}\right)\widetilde{B}(dx)\right\}_{t\in\mathbb{R}} \tag{1.2}$$

for some complex-valued matrix A. In (1.2), $x_\pm = \max\{\pm x, 0\}$,

$$D = H - \tfrac{1}{2}I, \tag{1.3}$$

and $\widetilde{B}(dx)$ is a complex-valued random measure such that $\widetilde{B}(-dx) = \overline{\widetilde{B}(dx)}$, $E\widetilde{B}(dx)\widetilde{B}(dx)^* = dx$, where $*$ represents Hermitian transposition. Expression (1.2) shows that the law of an OFBM can be fully described based on the scaling matrix H and the spectral parameter A (see Remark 2.1 on the parametrization).

Let $H = PJ_HP^{-1}$, $P \in GL(n, \mathbb{C})$, be the Jordan form of the Hurst parameter in (1.2) (see Section 2 for matrix notation). If H is diagonal, then we can assume that P takes the form of a scalar matrix $P = pI$, where $p \in \mathbb{R}$ and I is the $n \times n$ identity matrix, and that $J_H = \mathrm{diag}(h_1, \ldots, h_n)$. In this case, (1.1) breaks down into simultaneous entry-wise expressions

$$\{X(ct)\}_{t\in\mathbb{R}} \stackrel{\mathcal{L}}{=} \left\{\left(c^{h_1}X_1(t), \ldots, c^{h_n}X_n(t)\right)^*\right\}_{t\in\mathbb{R}}, \qquad c > 0. \tag{1.4}$$

Relation (1.4) is henceforth called *entry-wise scaling*. In particular, under (1.4) an OFBM is a vector of correlated FBM entries (Amblard *et al.* [4], Coeurjolly *et al.* [10]). Several estimators have been developed by building upon the univariate, entry-wise scaling laws, e.g., the Fourier-based multivariate local Whittle (e.g., Shimotsu [46], Nielsen [42]) and the multivariate wavelet regression (Wendt *et al.* [51], Amblard and Coeurjolly [5], Achard and Gannaz [3]). However, if H is non-diagonal, that is, if P is not a scalar matrix, then the relation (1.1) mixes together the several entries of X. The estimation problem under a non-scalar P turns out to be rather intricate and calls for the construction of methods that are multivariate from their inception.

Although the emergence of o.s.s. processes in applications is rightly expected – for example, as functional weak limits of multivariate time series – there is no specific reason to believe a priori that scaling laws occur predominantly entry-wise and *exactly* along the canonical axes. Indeed, this is palpably not true in several applications such as fractional blind source separation (see Didier *et al.* [18]) and fractional cointegration (see Robinson [45] for a bivariate local Whittle estimator). The framework of multivariate mixed fractional time series subsumes both cases.

Let $W = \{W_t\}_{t \in \mathbb{Z}}$ be an unobserved signal. To fix ideas, suppose that the signal is an operator fractional Gaussian noise (OFGN), namely, it has the form $W_t = B_{H_W}(t) - B_{H_W}(t-1)$ with spectral parameter A_W (see (1.2)). Further assume that $H_W = \mathrm{diag}(h_1, \ldots, h_n)$. The observed signal has the form $Y = \{Y_t\}_{t \in \mathbb{Z}} = \{PW_t\}_{t \in \mathbb{Z}}$, for a non-scalar, mixing matrix $P \in GL(n, \mathbb{R})$. Then, the mixed process Y is another OFGN with parameters $H_Y = P \, \mathrm{diag}(h_1, \ldots, h_n) P^{-1}$ and $A_Y = PA_W$. In blind source separation, the entries of W are uncorrelated, whereas in cointegration they are typically correlated.

From a mathematical standpoint, the mixing of scaling laws can be illustrated by means of the expression for the spectral density f_X of an OFBM with Hurst parameter $H = P \, \mathrm{diag}(h_1, h_2) P^{-1}$, $0 < h_1 < h_2 < 1$, $P \in GL(2, \mathbb{R})$ (see condition (2.12)). For $M := P^{-1} A A^* (P^*)^{-1} \in M(n)$ (cf. condition (2.11)) and $x > 0$, the spectral density takes the form $(f_X(x)_{ij}) = x^{-D} A A^* x^{-D^*}$, where

$$f_X(x)_{11} = p_{11}^2 m_{11} x^{-2d_1} + 2 p_{11} p_{12} m_{12} x^{-(d_1+d_2)} + p_{12}^2 m_{22} x^{-2d_2},$$

$$f_X(x)_{21} = p_{11} p_{21} m_{11} x^{-2d_1} + (p_{11} p_{22} + p_{12} p_{21}) m_{12} x^{-(d_1+d_2)} + p_{12} p_{22} m_{22} x^{-2d_2},$$

$$f_X(x)_{22} = p_{21}^2 m_{11} x^{-2d_1} + 2 p_{21} p_{22} m_{12} x^{-(d_1+d_2)} + p_{22}^2 m_{22} x^{-2d_2},$$

and $d_1 = h_1 - 1/2$, $d_2 = h_2 - 1/2$ (see (4.18) for the analogous expression in the wavelet domain). The univariate-inspired approach of setting up a Fourier-domain log-regression – for example, Whittle-type estimators – has to cope with the double-sided challenge of mixed power laws. On one hand, under mild assumptions on the amplitude coefficients, the dominant power law x^{-2d_2} always prevails around the origin of the spectrum. On the other hand, and paradoxically, even if the estimation of d_2 is the target, the magnitude of the amplitude coefficients themselves can arbitrarily bias the estimate over finite samples by masking the dominant power law.

In this work, we build upon the wavelet analysis of (n-dimensional) OFBM to propose a novel wavelet-based estimation method for bivariate, and potentially multivariate, OFBM. The method yields the Hurst eigenvalues of H and, under mild assumptions, also its eigenvectors when $P \in O(2)$. Its essential ingredient, and the main theme of this paper, is a change of perspective: instead of considering the entry-wise behavior of the wavelet spectrum as a function of wavelet scales, it draws upon the evolution along scales of the eigenstructure of the wavelet spectrum. This way, it avoids much of the difficulty associated with inference in the presence of mixed power laws, as we now explain.

For a wavelet function $\psi \in L^2(\mathbb{R})$ with a number N_ψ of vanishing moments (see (2.6)), the (normalized) vector wavelet transforms of OFBM is naturally defined as

$$\mathbb{R}^n \ni D(2^j, k) = 2^{-j/2} \int_{\mathbb{R}} 2^{-j/2} \psi(2^{-j}t - k) B_H(t) \, dt, \qquad j \in \mathbb{N} \cup \{0\}, k \in \mathbb{Z}, \qquad (1.5)$$

provided the integral in (1.5) exists in an appropriate sense. The wavelet-domain process $\{D(2^j, k)\}_{k \in \mathbb{Z}}$ is stationary in k and o.s.s. in j (Proposition 3.1). Moreover, whereas the original stochastic process $B_H(t)$ displayed fractional memory, the covariance between (multivariate) wavelet coefficients decays as a function of $|2^j k - 2^{j'} k'|$ according to an inverse fractional power

controlled by N_ψ (Proposition 3.2). The wavelet spectrum (variance) at scale j is the positive definite matrix

$$ED(2^j, k)D(2^j, k)^* = ED(2^j, 0)D(2^j, 0)^* =: EW(2^j),$$

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

$$W(2^j) = \frac{1}{K_j}\sum_{k=1}^{K_j} D(2^j, k)D(2^j, k)^*, \qquad K_j = \frac{\nu}{2^j}, j = j_1, \ldots, j_m, \tag{1.6}$$

for a dyadic total of ν (wavelet) data points. Within the bivariate framework, the univariate-like entry-wise scaling approach would consist of exploiting the behavior of each component $W(2^j)_{i_1,i_2}$, $i_1, i_2 = 1, 2$, of the sample wavelet transform $W(2^j)$ as a function of the scales 2^j. Apart from an amplitude effect, the entries are then controlled by the dominant Hurst eigenvalue h_2 (see expression (4.16)). Figure 1, top panels, illustrates the fact that this precludes the estimation of h_1.

The proposed estimators of the Hurst eigenvalues h_1 and h_2 are

$$\hat{h}_1(2^j) = \frac{\log \lambda_1(2^j)}{2\log(2^j)}, \qquad \hat{h}_2(2^j) = \frac{\log \lambda_2(2^j)}{2\log(2^j)}, \tag{1.7}$$

where $\lambda_1(2^j) \leq \lambda_2(2^j)$ are the eigenvalues of the positive definite symmetric matrix $W(2^j)$ (see Definition 4.1 for the precise assumptions). However, as usual with operator self-similarity, the finite sample expressions for $\lambda_1(2^j)$ and $\lambda_2(2^j)$ themselves involve a mixture of distinct power laws 2^{j2h_1}, $2^{j(h_1+h_2)}$, 2^{j2h_2} ($h_1 < h_2$). For this reason, one must take the limit at coarse scales, namely, the scale itself must go to infinity. It is a remarkable fact that the power law 2^{j2h_1} ends up prevailing in the expression for $\lambda_1(2^j)$ (see Figure 1, bottom panels, and the striking contrast with the top panels; see Remark 4.1 for a mathematically motivated, intuitive discussion). The convergence of (1.7) in turn allows for the convergence of associated sequences of eigenvectors when P is orthogonal. Moreover, simulation studies show that the estimation procedure is accurate and computationally fast. The asymptotics are mathematically developed in two stages. In the first, the wavelet scales (octaves) are held fixed and the asymptotic distribution of the sample wavelet transform is obtained (Proposition 3.3 and Theorem 3.1). In the second, one takes the limit with respect to the scales themselves. However, the latter must go to infinity slower than the sample size, a feature that our estimators share with Fourier or wavelet-based semiparametric estimators in general (e.g., Robinson [43], Moulines *et al.* [39–41]).

Our results are related to the literature on the estimation of operator stable laws via eigenvalues and eigenvectors of sample quadratic forms (see Meerschaert and Scheffler [37,38]). In this context, one encounters the same problem with the prevalence of some dominant power law (i.e., the tail exponent) in most directions. In Becker-Kern and Pap [8], a similar philosophy is applied in the time domain to produce one of the very few available estimators for authentic, mixed scaling o.s.s. processes of dimension up to 4. However, the asymptotics provided are restricted to consistency. In our work, the wavelet transform is the main tool for ensuring the consistency and asymptotic normality of the proposed estimators.
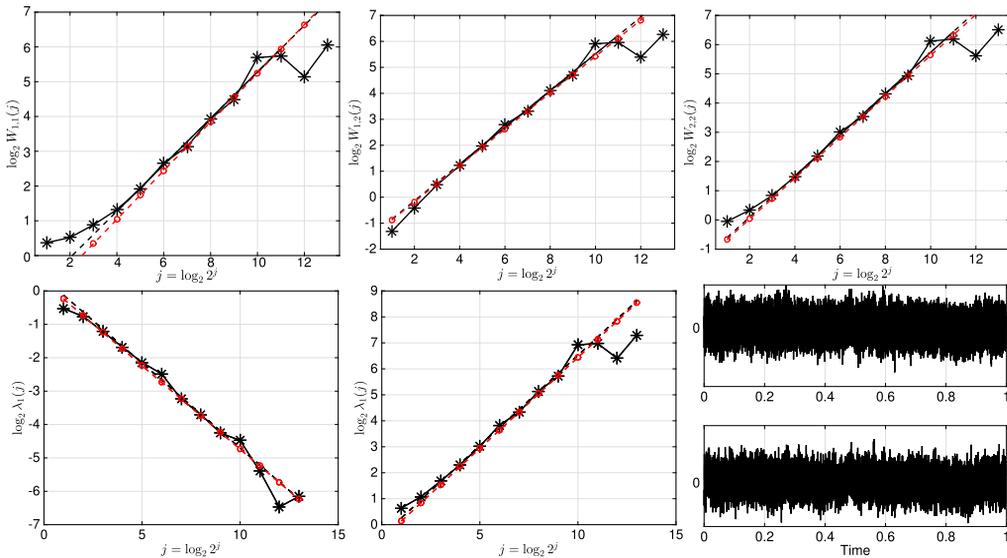
Figure 1. *Entry-wise vs eigenvalue-based estimation.* From one synthetic realization of OFBM, the top row displays (black solid lines with ∗), from left to right, $\log_2 W(2^j)_{1,1}$ vs. j, $\log_2 W(2^j)_{1,2}$ vs j and $\log_2 W(2^j)_{2,2}$ vs. j, on which the asymptotic behavior $j2h_2$ is superimposed (red dashed line with "o"). All auto- and cross-components are then driven by the dominant Hurst eigenvalue h_2, which precludes the estimation of the eigenvalue h_1. The first two bottom row plots display (black solid lines with ∗), from left to right, $\log_2 \lambda_1(2^j)$ vs j and $\log_2 \lambda_2(2^j)$ vs j, with their respective asymptotic trends $j2h_1$ and $j2h_2$ superimposed (red dashed line with "o"). This demonstrates that both Hurst eigenvalues h_1 and h_2 can be estimated (see Section 5.1 for the simulation details). The bottom right plot displays the increments of the generated OFBM.

The paper is organized as follows. Section 2 contains the notation, assumptions and basic concepts. Section 3 is dedicated to the wavelet analysis of n-dimensional OFBM, as well as the asymptotics of the wavelet transform for fixed scales (most of the proofs can be found in the supplementary material file Abry and Didier [2], Section B). In Section 4, the estimation method for the Hurst exponent of bivariate OFBM is laid out in full detail and its asymptotics are established at coarse scales. Section 5 displays finite sample computational studies, including one of the performance of the estimators under blind source separation and cointegrated instances, with the purpose of illustrating the robustness of the estimators' performance with respect to different parametric scenarios. The research contained in this paper leads to a number of interesting open questions, which are mentioned in Section 6. Among these is the extension of the consistency and asymptotic normality to any dimension $n \geq 3$, which will generally require dispensing with explicit formulas for eigenvalues and eigenvectors. The proofs of some basic results can be found in the appendix. In addition, the supplementary file Abry and Didier [2], Section B, contains several additional auxiliary results and proofs. In Section C of the supplementary file, the performance of the estimators is established under the assumption that only discrete observations are available, instead of a full sample path as in (1.5).

2. Notation and assumptions

All through the paper, the dimension of OFBM is denoted by $n \geq 2$.

We shall use throughout the paper the following notation for finite-dimensional operators (matrices). All with respect to the field $\mathbb{R}$, $M(n)$ or $M(n, \mathbb{R})$ is the vector space of all $n \times n$ matrices (endomorphisms), $GL(n)$ or $GL(n, \mathbb{R})$ is the general linear group (invertible matrices, or automorphisms), $O(n)$ is the (orthogonal) group of matrices O such that $OO^* = I = O^*O$ (i.e., the adjoint operator is the inverse), and $\mathcal{S}(n, \mathbb{R})$ is the space of symmetric matrices. We also write I_n to indicate the dimension of the identity matrix I. 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

$$\text{diag}(\mathcal{P}_1, \ldots, \mathcal{P}_n), \qquad \text{diag}_m(\mathcal{P}), \tag{2.1}$$

respectively. The symbol $\| \cdot \|$ represents a generic matrix or vector norm. For $q \in \mathbb{N}$, the l^q entry-wise norm of an $m \times n$ real-valued matrix M is denoted by

$$\|M\|_{l^q} = \left(\sum_{i_1=1}^{m} \sum_{i_2=1}^{n} |M_{i_1, i_2}|^q \right)^{1/q} \tag{2.2}$$

and $\|M\|_{l^\infty} = \sup_{i_1, i_2} |M_{i_1, i_2}|$. A generic matrix $M \in M(n, \mathbb{C})$ has real and imaginary parts $\Re(M)$ and $\Im(M)$, respectively. The functions

$$\pi_i(v), \pi_{i_1, i_2}(M), \qquad \mathbf{v} \in \mathbb{R}^n, M \in M(n, \mathbb{R}), \tag{2.3}$$

denote, respectively, the ith projection (entry) of the vector $\mathbf{v}$ and the (i_1, i_2)th projection (entry) of the matrix M, $i_1, i_2 = 1, \ldots, n$. For $S = (s_{i_1, i_2})_{i_1, i_2 = 1, \ldots, n} \in \mathcal{S}(n, \mathbb{R})$, we define the operator

$$\text{vec}_{\mathcal{S}}(S) = (s_{11}, s_{12}, \ldots, s_{1n}; s_{22}, \ldots, s_{2n}; \ldots; s_{n-1,n-1}, s_{n-1,n}; s_{n,n})^*. \tag{2.4}$$

In other words, $\text{vec}_{\mathcal{S}}(\cdot)$ vectorizes the upper triangular entries of the symmetric matrix S.

When establishing bounds, C stands for a positive constant whose value can change from one line to the next. For a sequence of random vectors $\{X_l, Y_l\}_{l \in \mathbb{N}}$, $P(Y_l = 0) = 0$, we write

$$X_l \overset{P}{\sim} Y_l \tag{2.5}$$

to mean that $X_l / Y_l \overset{P}{\to} 1$, $l \to \infty$. Note that this does not imply that $\{X_l, Y_l\}_{l \in \mathbb{N}}$ converges in probability. Relations of the type (2.5) will often appear in the proofs of the results in Section 4. We write $\mathbf{X} \overset{d}{=} \mathbf{Y}$ when the random vectors $\mathbf{X}$ and $\mathbf{Y}$ have the same distribution.

All through the paper, we will make the following assumptions on the underlying wavelet basis. For this reason, such assumptions will be omitted in the statements.

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

$$\int_{\mathbb{R}} \psi^2(t) \, dt = 1, \qquad \int_{\mathbb{R}} t^q \psi(t) \, dt = 0, \qquad q = 0, 1, \ldots, N_\psi - 1, N_\psi \geq 2. \tag{2.6}$$

Assumption ($W2$).

$$\text{supp}(\psi) \text{ is a compact interval.} \tag{2.7}$$

Assumption ($W3$). There is $\alpha > 1$ such that

$$\sup_{x \in \mathbb{R}} \left| \hat{\psi}(x) \right| \left(1 + |x| \right)^{\alpha} < \infty. \tag{2.8}$$

Under (2.6), (2.7) and (2.8), ψ is continuous, $\hat{\psi}(x)$ is everywhere differentiable and its first $N_\psi - 1$ derivatives are zero at $x = 0$ (see Mallat [35], Theorem 6.1 and the proof of Theorem 7.4).

Example 2.1. If ψ is a Daubechies wavelet with N_ψ vanishing moments, $\text{supp}(\psi) = [0, 2N_\psi - 1]$ (see Mallat [35], Proposition 7.4).

Starting from the harmonizable representation (1.2), throughout the paper we will make the following assumptions on the OFBM $B_H = \{B_H(t)\}_{t \in \mathbb{R}}$.

Assumption (OFBM1). The eigenvalues (characteristic roots) h_k of the matrix exponent H satisfy

$$0 < \Re(h_k) < 1, \qquad k = 1, \ldots, n. \tag{2.9}$$

Assumption (OFBM2).

$$\Re(AA^*) \text{ has full rank.} \tag{2.10}$$

The condition (2.9) generalizes the familiar constraint $0 < H < 1$ on the Hurst parameter of a FBM. As shown in Didier and Pipiras [19], it ensures the existence of the harmonizable representation (1.2). Also, (2.9) implies that the OFBM under consideration has mean zero, which follows by the same reasoning as in Taqqu [47], page 7, property (ii). In turn, recall that a stochastic process is called proper when the distribution of $X(t)$ is full dimensional for $t \neq 0$. The condition (2.10) is sufficient (though not necessary) for the integral on the right-hand side of (1.2) to be a proper stochastic process and hence to define an OFBM.

The next two assumptions will appear in some of the results.

Assumption (OFBM3).

$$\Im(AA^*) = 0. \tag{2.11}$$

Assumption (OFBM4). $B_H = \{B_H(t)\}_{t \in \mathbb{R}}$ is a *bivariate* OFBM with scaling matrix

$$H = PJ_H P^{-1} = P \,\text{diag}(h_1, h_2) P^{-1}, \qquad 0 < h_1 < h_2 < 1, P \in GL(2, \mathbb{R}), \tag{2.12}$$

where the columns of P are unit vectors.

The condition (2.11) is equivalent to time reversibility, namely, $\{B_H(-t)\} \overset{\mathcal{L}}{=} \{B_H(t)\}$. In turn, the latter is equivalent to the existence of a closed form expression for the covariance function, i.e.,

$$E B_H(s) B_H(t)^* = \tfrac{1}{2} \left\{ |t|^H \Sigma |t|^{H^*} + |s|^H \Sigma |s|^{H^*} - |t-s|^H \Sigma |t-s|^{H^*} \right\}, \tag{2.13}$$

where $\Sigma := E B_H(1) B_H(1)^*$ (Didier and Pipiras [19], Proposition 5.2). Time reversibility is used in some of the results in Section 3 and in all Section 4. The bivariate framework (2.12) is only used in limits at coarse scales (Section 4), due to the availability of convenient formulas for eigenvalues and eigenvectors.

Remark 2.1. In regard to the parametrization, the connection between Σ and (H, AA^*) can be obtained implicitly based on the second moment of the harmonizable representation (1.2) at $t = 1$, and it can be worked out explicitly under (2.13) and stronger additional conditions, for example, assuming H is diagonalizable with real eigenvalues and eigenvectors.

The condition (2.12) renders OFBM identifiable, namely, the mapping from the parametrization H into the space of OFBM laws is injective for a fixed spectral parameter AA^*. This is a consequence of only considering Hurst matrices H with real eigenvalues. For a general discussion of the (non)identifiability of OFBM, see Didier and Pipiras [20].

3. Wavelet analysis

In this section, we carry out the wavelet analysis of n-dimensional OFBM. The proof of Proposition 3.1 can be found in the Appendix, whereas Abry and Didier [2], Section B, contains the proofs of Propositions 3.2, 3.3 and Theorem 3.1.

3.1. Basic properties

The normalized wavelet transform (1.5) is itself a vector-valued random field in the scale and shift parameters j and k, respectively. It will be convenient to make the change of variables $z = 2^{-j} t - k$, and reexpress

$$D\big(2^j, k\big) = \int_{\mathbb{R}} \psi(z) B_H\big(2^j z + 2^j k\big) \, dz. \tag{3.1}$$

As in the univariate case, the wavelet coefficients of OFBM exhibit a number of nice properties. The next proposition describes such properties as well as the general form of the wavelet spectrum (variance).

Proposition 3.1. *Under the assumptions* (OFBM1)–(2), *let* $\{D(2^j, k)\}_{j \in \mathbb{N}, k \in \mathbb{Z}}$ *be as in* (3.1). *Then,*

(P1) *the wavelet transform* (1.5) *is well-defined in the mean square sense, and* $E D(2^j, k) = 0$;

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

(P3) *(operator self-similarity over different scales)* $\{D(2^j, k)\}_{k \in \mathbb{Z}} \overset{\mathcal{L}}{=} \{2^{jH} D(1, k)\}_{k \in \mathbb{Z}}$;

(P4) *for some* $C > 0$, *the wavelet spectrum* $E W(2^j, k) \equiv E W(2^j)$ *is given by*

$$E W\big(2^j\big) = C \int_{\mathbb{R}} \big(x_+^{-D} AA^* x_+^{-D^*} + x_-^{-D} \overline{AA^*} x_-^{-D^*}\big) \frac{|\hat{\psi}(2^j x)|^2}{x^2} \, dx; \tag{3.2}$$

(P5) *the wavelet spectrum satisfies the operator scaling relation*

$$E W(2^j) = 2^{jH} E W(1) 2^{jH^*} = 2^{jH} \left\{ \int_\mathbb{R} \int_\mathbb{R} \psi(t) \psi(t') E B_H(t) B_H(t')^* \, dt \, dt' \right\} 2^{jH^*},$$

$j \in \mathbb{N}$. *In particular, under* (2.11),

$$E W(2^j) = -\frac{1}{2} 2^{jH} \left\{ \int_\mathbb{R} \int_\mathbb{R} \psi(t) \psi(t') |t - t'|^H \Sigma |t - t'|^{H^*} \, dt \, dt' \right\} 2^{jH^*}; \tag{3.3}$$

(P6) *in analogy to* (P5), $W(2^j) \overset{d}{=} 2^{jH} (\frac{1}{K_j} \sum_{k=1}^{K_j} D(1, k) D(1, k)^*) 2^{jH^*}$, *where* K_j *is given in* (1.6);

(P7) *the wavelet spectrum has full rank, namely,* $\det E W(2^j) \neq 0$, $j \in \mathbb{N}$.

Fix some $0 < \delta < 1/2$, and consider the range of wavelet parameters j, j', k, k' such that

$$\left| \frac{\max\{2^j, 2^{j'}\}}{2^j k - 2^{j'} k'} \right| \leq \frac{\delta}{\max\{1, \text{length}(\text{supp}(\psi))\}}. \tag{3.4}$$

If the parameters (j, k) and (j', k') of two wavelet coefficients satisfy (3.4), then we can interpret that the latter are "far apart" in the parameter space. The next result provides a notion of decay of the covariance between wavelet coefficients under (3.4). The proof is similar to that for the univariate case, but we provide it in Abry and Didier [2] for the reader's convenience.

Proposition 3.2. *Under the assumptions* (OFBM1)–(3) *and* (3.4), *the covariance between wavelet coefficients* (3.1) *satisfies the relation*

$$E D(2^j, k) D(2^{j'}, k')^*$$
$$= -\frac{1}{2} \frac{2^{(j+j')N_\psi}}{|2^j k - 2^{j'} k'|^{2N_\psi}} \left\{ |2^j k - 2^{j'} k'|^H (O_{i_1, i_2}^{N_\psi}(1))_{i_1, i_2 = 1, \dots, n} |2^j k - 2^{j'} k'|^{H^*} \right\}, \tag{3.5}$$

where $(O_{i_1, i_2}^{N_\psi}(1))_{i_1, i_2 = 1, \dots, n}$ *is an entry-wise bounded symmetric-matrix-valued function that depends only on* N_ψ. *As a consequence,*

$$\left\| E D(2^j, k) D(2^{j'}, k')^* \right\| \leq C(N_\psi, j, j') \frac{|\log^\kappa |2^j k - 2^{j'} k'||}{|2^j k - 2^{j'} k'|^{2N_\psi - 2\max_{h \in \text{eig}(H)} \Re(h)}}, \tag{3.6}$$

where κ *is the dimension of the largest Jordan block in the spectrum of* H, *and* $C(N_\psi, j, j') = C(N_\psi) 2^{(j+j')N_\psi}$ *for some constant* $C(N_\psi) > 0$.

3.2. Asymptotics for sample wavelet transforms: Fixed scales

As typical in the asymptotic study of averages, we begin by investigating the asymptotic covariance of the sample wavelet transforms $W(2^j)$.

For FBM, the asymptotic covariance between wavelet transforms $W(2^j)$, $W(2^{j'}) \in \mathbb{R}$ is not available in closed form since it depends on the wavelet function, which is itself not available in closed form (cf. Bardet [7], Proposition II.3). Operator self-similarity adds a layer of intricacy, since in general exact entry-wise scaling relations are not present.

For notational simplicity, let

$$
\begin{aligned}
\mathbf{X}^* &= (X_1, \dots, X_n) = \big(d_1(2^j, k), \dots, d_n(2^j, k)\big), \\
\mathbf{Y}^* &= (Y_1, \dots, Y_n) = \big(d_1(2^{j'}, k'), \dots, d_n(2^{j'}, k')\big),
\end{aligned}
\tag{3.7}
$$

where $d_i(2^j, k)$, $i = 1, \dots, n$, is the ith entry of the wavelet transform vector $D(2^j, k)$. The bivariate case serves to illustrate the computation of covariances.

Example 3.1. For a zero mean, Gaussian random vector $\mathbf{Z} \in \mathbb{R}^m$, the Isserlis theorem (e.g., Vignat [50]) yields

$$
E(Z_1 \cdots Z_{2k}) = \sum \prod E(Z_i Z_j), \qquad E(Z_1 \cdots Z_{2k+1}) = 0, \qquad k = 1, \dots, \lfloor m/2 \rfloor. \tag{3.8}
$$

The notation $\sum \prod$ stands for adding over all possible k-fold products of pairs $E(Z_i Z_j)$, where the indices partition the set $1, \dots, 2k$. So, let $\mathbf{X}$ and $\mathbf{Y}$ be as in (3.7) with $n = 2$. Then,

$$
\begin{aligned}
&\mathrm{Cov}\left(\begin{pmatrix} X_1^2 \\ X_2^2 \\ X_1 X_2 \end{pmatrix}, \begin{pmatrix} Y_1^2 \\ Y_2^2 \\ Y_1 Y_2 \end{pmatrix} \right) \\
&= \begin{pmatrix} \big(E(X_1 Y_1)\big)^2 & 2\big(E(X_1 Y_2)\big)^2 & 2E(X_1 Y_1)E(X_1 Y_2) \\ 2\big(E(X_2 Y_1)\big)^2 & 2\big(E(X_2 Y_2)\big)^2 & 2E(X_2 Y_1)E(X_2 Y_2) \\ 2E(X_1 Y_1)E(X_2 Y_1) & 2E(X_1 Y_2)E(X_2 Y_2) & c_{33} \end{pmatrix},
\end{aligned}
\tag{3.9}
$$

where $c_{33} = E(X_1 Y_1)E(X_2 Y_2) + E(X_1 Y_2)E(X_2 Y_1)$.

Expression (3.9) shows that the asymptotic behavior of the second moments of $W(2^j)$ involves several cross products. A notationally parsimonious way of tackling this difficulty is by resorting to Kronecker products. For instance, in the bivariate case,

$$
M(4, \mathbb{R}) \ni EXY^* \otimes EXY^* = \begin{pmatrix} E(X_1 Y_1) & E(X_1 Y_2) \\ E(X_2 Y_1) & E(X_2 Y_2) \end{pmatrix} \otimes \begin{pmatrix} E(X_1 Y_1) & E(X_1 Y_2) \\ E(X_2 Y_1) & E(X_2 Y_2) \end{pmatrix}
$$

contains all the 9 terms (two-fold products of cross moments), as well as a few repeated ones, needed to express (3.9). In view of (3.8), this fact extends to general dimension n by means of the relations

$$
\mathrm{Cov}(X_{i_1} X_{i_2}, Y_{j_1} Y_{j_2}) = E(X_{i_1} Y_{j_1})E(X_{i_2} Y_{j_2}) + E(X_{i_1} Y_{j_2})E(X_{i_2} Y_{j_1}), \tag{3.10}
$$

for $i_1, i_2, j_1, j_2 = 1, \dots, n$. The next proposition provides an expression that encompasses the asymptotic fourth moments of the wavelet coefficients.

Proposition 3.3. *Let* $B_H = \{B_H(t)\}_{t \in \mathbb{R}}$ *be an OFBM under the assumptions* (OFBM1)–(3) *and let* K_j *be as in* (1.6). *As* $\nu \to \infty$, *for every pair of octaves* j, j',

(i)

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

$$\to 2^{-(j+j')/2} \gcd(2^j, 2^{j'}) \sum_{z=-\infty}^{\infty} \Phi_{z \gcd(2^j, 2^{j'})} \otimes \Phi_{z \gcd(2^j, 2^{j'})}, \tag{3.11}$$

where

$$\Phi_q := -\frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(t) \psi(t') \left| (2^j t - 2^{j'} t') + q \right|^H \Sigma \left| (2^j t - 2^{j'} t') + q \right|^{H^*} dt\, dt';$$

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

$$\sqrt{K_j}\sqrt{K_{j'}} \operatorname{Cov}\left(\operatorname{vec}_{\mathcal{S}} W(2^j), \operatorname{vec}_{\mathcal{S}} W(2^{j'})\right) \to G_{jj'},$$

where the entries of $G_{jj'}$ *can be retrieved from* (3.11) *by means of* (3.10) (*see* (2.4) *on the notation* $\operatorname{vec}_{\mathcal{S}}$).

Remark 3.1. The definition of wavelet only requires $N_\psi \geq 1$, but $N_\psi \geq 2$ (see (2.6)) is needed for the convergence in Proposition 3.3 (see (B.32)).

The next theorem establishes the asymptotics for the vectorized sample wavelet transforms $(\operatorname{vec}_{\mathcal{S}} W(2^j))_{j=j_1,\dots,j_m}$ at a fixed set of octaves.

Theorem 3.1. *Let* $B_H = \{B_H\}_{t \in \mathbb{R}}$ *be an OFBM under the assumptions* (OFBM1)–(3). *Let* $j_1 < \cdots < j_m$ *be a fixed set of octaves. Then,*

$$\left(\sqrt{K_j}\left(\operatorname{vec}_{\mathcal{S}} W(2^j) - \operatorname{vec}_{\mathcal{S}} E W(2^j)\right)\right)_{j=j_1,\dots,j_m} \xrightarrow{d} \mathcal{N}_{n(n+1)/2 \times m}(0, F), \tag{3.12}$$

as $\nu \to \infty$ (*see* (2.4) *on the notation* $\operatorname{vec}_{\mathcal{S}}$). *In* (3.12), *the matrix* $F \in \mathcal{S}(\frac{n(n+1)}{2}m, \mathbb{R})$ *has the form* $F = (G_{jj'})_{j,j'=1,\dots,m}$, *where each block* $G_{jj'} \in M(n(n+1)/2, \mathbb{R})$ *is described in Proposition* 3.3.

4. A wavelet-based estimator for bivariate OFBM

In this section, we switch to the bivariate framework (2.12), that is, $n = 2$. We draw upon explicit expressions for eigenvalues to establish the consistency and asymptotic normality of the estimators (1.7) as the wavelet scale grows according to a factor $a(\nu) \to \infty$, as $\nu \to \infty$. We also show

the consistency and asymptotic normality, in a sense to be defined, of a sequence of eigenvectors associated with the smallest eigenvalue of the sample wavelet variance matrix.

The proposed estimators make use of the behavior at coarse scales of the sample wavelet variance

$$W\big(a(v)2^j\big) = \frac{1}{K_{a,j}} \sum_{k=1}^{K_{a,j}} D\big(a(v)2^j, k\big) D\big(a(v)2^j, k\big)^*, \qquad K_{a,j} = \frac{v}{a(v)2^j}. \tag{4.1}$$

Recall that $h_1 < h_2$ under (2.12). In (4.1), $\{a(v)\}_{v \in \mathbb{N}}$ is assumed to be a dyadic sequence such that

$$a(v) \le \frac{v}{2^j}, \qquad \frac{a(v)}{v} + \frac{v}{a(v)^{1+2h_1}} \to 0, \qquad v \to \infty \tag{4.2}$$

(see Remark 4.2 on the choice of $a(v)$ in practice).

We will make use of some basic relations for bivariate symmetric positive semidefinite matrices. Recall that for a matrix

$$S = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \tag{4.3}$$

the eigenvalues can be expressed in closed form as

$$\lambda_1 = \frac{(a+c) - \sqrt{\Delta}}{2} \le \lambda_2 = \frac{(a+c) + \sqrt{\Delta}}{2}, \qquad \Delta = (a+c)^2 - 4\big(ac - b^2\big). \tag{4.4}$$

By Sylvester's criterion, positive semidefiniteness implies that $a + c \ge 0$. As a consequence, if $\det(S) > 0$,

$$\lambda_1 = \frac{1}{2}(a+c)\left(1 - \sqrt{1 - \frac{4(ac - b^2)}{(a+c)^2}}\right) = \frac{2(ac - b^2)}{a+c} \frac{(1 - \sqrt{1 - \frac{4(ac-b^2)}{(a+c)^2}})}{4(ac - b^2)/(a+c)^2}, \tag{4.5}$$

and

$$\lambda_2 = \frac{1}{2}(a+c)\left(1 + \sqrt{1 - \frac{4(ac - b^2)}{(a+c)^2}}\right). \tag{4.6}$$

Let $\mathbf{v} = (v_1, v_2)^* \in \mathbb{R}^2$ be an eigenvector associated with an eigenvalue λ. By further assuming $b \ne 0$, the relation $S\mathbf{v} = \lambda\mathbf{v}$ yields

$$v_2 = \frac{\lambda - a}{b} v_1. \tag{4.7}$$

4.1. The weak limit of eigenvalues

The next definition describes the proposed estimators for the Hurst eigenvalues $h_1 < h_2$.

Definition 4.1. *Let $B_H = \{B_H(t)\}_{t \in \mathbb{R}}$ be an OFBM under the assumptions* (OFBM1)–(4). *For a dyadic number $a(v)$, let $W(a(v)2^j)$ be the associated (symmetric) sample wavelet spectrum at scale $a(v)2^j$, and let*

$$\lambda_1\big(a(v)2^j\big) \leq \lambda_2\big(a(v)2^j\big) \tag{4.8}$$

be its eigenvalues. The wavelet estimators at scale $a(v)2^j$ of the eigenvalues $h_1 < h_2$ are defined, respectively, as in expression (1.7) *with $a(v)2^j$ in place of 2^j.*

By analogy to (4.8) and (1.7), we denote the eigenvalues and normalized log-eigenvalues of $EW(a(v)2^j)$, respectively, by

$$
\begin{aligned}
&\lambda_1^E\big(a(v)2^j\big) \leq \lambda_2^E\big(a(v)2^j\big), \\
&h_1^E\big(a(v)2^j\big) = \frac{\log \lambda_1^E(a(v)2^j)}{2\log(a(v)2^j)} \leq h_2^E\big(a(v)2^j\big) = \frac{\log \lambda_2^E(a(v)2^j)}{2\log(a(v)2^j)}.
\end{aligned}
\tag{4.9}
$$

When developing asymptotics for the estimators (4.8), in view of the operator self-similarity property (P6) we will consider the matrix statistics

$$\hat{B}_a\big(2^j\big) := P^{-1}W_a\big(2^j\big)\big(P^*\big)^{-1}, \qquad j = j_1 < \cdots < j_m, \tag{4.10}$$

where $W_a(2^j) = \frac{1}{K_{a,j}}\sum_{k=1}^{K_{a,j}} D(2^j,k)D(2^j,k)^*$ (cf. (1.6) and (4.1)). Each $\hat{B}_a(2^j)$ is only a pseudo-estimator of

$$B\big(2^j\big) := P^{-1}EW\big(2^j\big)\big(P^*\big)^{-1}, \qquad P \in GL(2,\mathbb{R}), \tag{4.11}$$

because its expression involves the unknown matrix parameter P. It will be convenient to describe the matrices (4.10), (4.11) entry-wise as

$$\hat{B}_a\big(2^j\big) = \big(\hat{b}_{i_1,i_2}\big(2^j\big)\big)_{i_1,i_2=1,2}, \qquad B\big(2^j\big) = \big(b_{i_1,i_2}\big(2^j\big)\big)_{i_1,i_2=1,2}. \tag{4.12}$$

The asymptotic distribution of the matrix statistics (4.10) is given in the following lemma.

Lemma 4.1. *For $m \in \mathbb{N}$, let $j_1 < \cdots < j_m$ be a set of fixed octaves j. Let $\Pi = (\pi_{i_1 i_2})_{i_1,i_2=1,2} = P^{-1}$, and let $\hat{B}_a(2^j)$, $B(2^j)$ be as in* (4.10), (4.11), *respectively. Under the assumptions* (OFBM1)–(4) *and* (4.2),

$$\big(\sqrt{K_{a,j}}\big(\mathrm{vec}_{\mathcal{S}}\,\hat{B}_a\big(2^j\big) - \mathrm{vec}_{\mathcal{S}}\,B\big(2^j\big)\big)\big)_{j=j_1,\ldots,j_m} \xrightarrow{d} N\big(0, \Sigma_B(j_1,\ldots,j_m)\big), \qquad v \to \infty. \tag{4.13}$$

In (4.13),

$$\Sigma_B(j_1,\ldots,j_m) = \mathrm{diag}_m(\mathcal{P})\,F\,\mathrm{diag}_m(\mathcal{P})^*, \tag{4.14}$$

where F is the asymptotic covariance matrix in (3.12), *$\mathrm{diag}_m(\mathcal{P})$ is as defined in* (2.1), *and*

$$
\mathcal{P} = \begin{pmatrix}
\pi_{11}^2 & 2\pi_{11}\pi_{12} & \pi_{12}^2 \\
\pi_{11}\pi_{21} & \pi_{11}\pi_{22} + \pi_{12}\pi_{21} & \pi_{12}\pi_{22} \\
\pi_{21}^2 & 2\pi_{21}\pi_{22} & \pi_{22}^2
\end{pmatrix}.
$$

Proof. For any j, a brief calculation shows that $\mathrm{vec}_\mathcal{S}(\Pi W_a(2^j)\Pi^*) = \mathcal{P}\mathrm{vec}_\mathcal{S} W_a(2^j)$. Likewise, $\mathrm{vec}_\mathcal{S}(\Pi E W(2^j)\Pi^*) = \mathcal{P}\mathrm{vec}_\mathcal{S} E W(2^j)$. Therefore, we can recast the left-hand side of (4.13) as

$$\mathrm{diag}_m(\mathcal{P})\big(\sqrt{K_{a,j}}\big(\mathrm{vec}_\mathcal{S} W_a\big(2^j\big) - \mathrm{vec}_\mathcal{S} E W\big(2^j\big)\big)\big)_{j=j_1,\ldots,j_m}.$$

The weak limit in (4.13) is now a consequence of (4.2) and Theorem 3.1. $\qquad\square$

The next lemma contains expressions for the wavelet spectrum and its sample counterpart. The notation $(\cdot)_{j=j_1,\ldots,j_m}$ designates a vector of 2×2 matrices.

Lemma 4.2. *For* $m \in \mathbb{N}$, *let* $j_1 < \cdots < j_m$ *be a set of fixed scales* j. *Let* $\hat{B}_a(2^j)$, $B(2^j)$ *be as in* (4.10), (4.11), *respectively, and let* $\hat{b}_{i_1,i_2}(2^j)$, $b_{i_1,i_2}(2^j)$ *be as in* (4.12). *For a dyadic number* $a(v)$, *under the assumptions* (OFBM1)–(4), *we can express*

$$\big(W\big(a(v)2^j\big)\big)_{j=j_1,\ldots,j_m} \stackrel{d}{=} \begin{pmatrix} \hat{a}_{j,a(v)} & \hat{b}_{j,a(v)} \\ \hat{b}_{j,a(v)} & \hat{c}_{j,a(v)} \end{pmatrix}_{j=j_1,\ldots,j_m}, \tag{4.15}$$

where

$$\hat{a}_{j,a(v)} = p_{11}^2 \hat{b}_{11}\big(2^j\big)a(v)^{2h_1} + 2p_{11}p_{12}\hat{b}_{12}\big(2^j\big)a(v)^{h_1+h_2} + p_{12}^2 \hat{b}_{22}\big(2^j\big)a(v)^{2h_2},$$

$$\begin{aligned}\hat{b}_{j,a(v)} &= p_{11}p_{21}\hat{b}_{11}\big(2^j\big)a(v)^{2h_1} + (p_{11}p_{22} + p_{12}p_{21})\hat{b}_{12}\big(2^j\big)a(v)^{h_1+h_2} \\ &\quad + p_{12}p_{22}\hat{b}_{22}\big(2^j\big)a(v)^{2h_2},\end{aligned} \tag{4.16}$$

$$\hat{c}_{j,a(v)} = p_{21}^2 \hat{b}_{11}\big(2^j\big)a(v)^{2h_1} + 2p_{21}p_{22}\hat{b}_{12}\big(2^j\big)a(v)^{h_1+h_2} + p_{22}^2 \hat{b}_{22}\big(2^j\big)a(v)^{2h_2}.$$

Likewise,

$$E W\big(a(v)2^j\big) = \begin{pmatrix} a_{j,a(v)} & b_{j,a(v)} \\ b_{j,a(v)} & c_{j,a(v)} \end{pmatrix}, \tag{4.17}$$

where

$$a_{j,a(v)} = p_{11}^2 b_{11}\big(2^j\big)a(v)^{2h_1} + 2p_{11}p_{12}b_{12}\big(2^j\big)a(v)^{h_1+h_2} + p_{12}^2 b_{22}\big(2^j\big)a(v)^{2h_2},$$

$$\begin{aligned}b_{j,a(v)} &= p_{11}p_{21}b_{11}\big(2^j\big)a(v)^{2h_1} + (p_{11}p_{22} + p_{12}p_{21})b_{12}\big(2^j\big)a(v)^{h_1+h_2} \\ &\quad + p_{12}p_{22}b_{22}\big(2^j\big)a(v)^{2h_2},\end{aligned} \tag{4.18}$$

$$c_{j,a(v)} = p_{21}^2 b_{11}\big(2^j\big)a(v)^{2h_1} + 2p_{21}p_{22}b_{12}\big(2^j\big)a(v)^{h_1+h_2} + p_{22}^2 b_{22}\big(2^j\big)a(v)^{2h_2}.$$

Proof. The operator scaling properties (P6) and (P5), respectively, yield

$$\begin{aligned}&\big(W\big(a(v)2^j\big)\big)_{j=j_1,\ldots,j_m} \\ &\stackrel{d}{=} \big(P\,\mathrm{diag}\big(a(v)^{h_1}, a(v)^{h_2}\big)\hat{B}_a\big(2^j\big)\,\mathrm{diag}\big(a(v)^{h_1}, a(v)^{h_2}\big)P^*\big)_{j=j_1,\ldots,j_m},\end{aligned} \tag{4.19}$$

$$E W\big(a(v)2^j\big) = P\,\mathrm{diag}\big(a(v)^{h_1}, a(v)^{h_2}\big)B\big(2^j\big)\,\mathrm{diag}\big(a(v)^{h_1}, a(v)^{h_2}\big)P^*, \tag{4.20}$$

from which (4.16) and (4.18) follow. $\qquad\square$

The next theorem establishes the consistency and asymptotic normality of the estimators described in Definition 4.1. Intuitively, the theorem states that

$$\hat{h}_1\big(a(v)2^j\big) \approx h_1, \qquad \hat{h}_2\big(a(v)2^j\big) \approx h_2 \qquad \text{with convergence rate } 2\log\big(a(v)2^j\big)\sqrt{K_{a,j}}.$$

Theorem 4.1. *For $m \in \mathbb{N}$, let $j_1 < \cdots < j_m$ be a set of fixed octaves j. Let $\hat{h}_1$, $\hat{h}_2$, h_1^E, h_2^E be as in* (1.7) *and* (4.9)*. Under the assumptions* (OFBM1)–(4) *and* (4.2)*, as $v \to \infty$,*

(i)

$$\big(\hat{h}_1\big(a(v)2^j\big), \hat{h}_2\big(a(v)2^j\big)\big) \xrightarrow{P} (h_1, h_2),$$
$$\big(h_1^E\big(a(v)2^j\big), h_2^E\big(a(v)2^j\big)\big) \to (h_1, h_2),$$
$$\tag{4.21}$$

for $j = j_1, \ldots, j_m$;

(ii)

$$\begin{pmatrix} 2\log\big(a(v)2^j\big)\sqrt{K_{a,j}}\big[\hat{h}_1\big(a(v)2^j\big) - h_1^E\big(a(v)2^j\big)\big] \\ 2\log\big(a(v)2^j\big)\sqrt{K_{a,j}}\big[\hat{h}_2\big(a(v)2^j\big) - h_2^E\big(a(v)2^j\big)\big] \end{pmatrix}_{j=j_1,\ldots,j_m}$$
$$\xrightarrow{d} N\big(0, \Sigma_{h_1,h_2}(j_1, \ldots, j_m)\big).$$
$$\tag{4.22}$$

In (4.22)*,*

$$\Sigma_{h_1,h_2}(j_1, \ldots, j_m) = \mathcal{Q}\Sigma_B(j_1, \ldots, j_m)\mathcal{Q}^*, \tag{4.23}$$

where $\Sigma_B(j_1, \ldots, j_m)$ is the covariance matrix in (4.14)*, and $\mathcal{Q} = (\mathcal{Q}_{jj'})_{j,j'=j_1,\ldots,j_m}$ is a block matrix whose blocks have dimension 2×3 and satisfy*

$$\mathcal{Q}_{jj} = \begin{pmatrix} \dfrac{b_{22}(2^j)}{\det B(2^j)} & -\dfrac{2b_{12}(2^j)}{\det B(2^j)} & \left(\dfrac{b_{11}(2^j)b_{22}(2^j)}{\det B(2^j)} - 1\right)\dfrac{1}{b_{22}(2^j)} \\ 0 & 0 & \dfrac{1}{b_{22}(2^j)} \end{pmatrix} \quad \text{and}$$

$$\mathcal{Q}_{jj'} = \mathbf{0}, \text{ if } j \neq j'.$$

Proof. Fix an arbitrary j. For notational simplicity we write $B = B(2^j) = (b_{i_1,i_2})_{i_1,i_2=1,2}$. We also drop the subscripts j, v in the expressions (4.16) and (4.18). Recall that the asymptotic distribution for $(\hat{B}_a(2^j))_{j=j_1,\ldots,j_m}$ is given by (4.13).

The statement (i) is a consequence of Lemma A.1 and Theorem 3.1. In regard to statement (ii), the asymptotics will be written out for just one general term indexed j, but the conclusions apply to the whole vector comprising the terms associated with $j = j_1, \ldots, j_m$. Since all meaningful limits will boil down to a function of the variables that appear on the right-hand side of (4.13), they depend on the omitted, fixed octave j.

Note that $2\log(a(v)2^j)[\hat{h}_i(a(v)2^j) - h_i^E(a(v)2^j)] = \log\lambda_i(a(v)2^j) - \log\lambda_i^E(a(v)2^j)$, $i = 1, 2$. With probability going to 1, we can reexpress the latter for $i = 1$ as

$$\left\{\log\frac{\hat{a}\hat{c} - \hat{b}^2}{\hat{a} + \hat{c}} - \log\frac{ac - b^2}{a + c}\right\}$$

$$+ \left\{\log\frac{\sqrt{1 + (-4)(\hat{a}\hat{c} - \hat{b}^2)/(\hat{a} + \hat{c})^2} - 1}{(-4)(\hat{a}\hat{c} - \hat{b}^2)/(\hat{a} + \hat{c})^2} - \log\frac{\sqrt{1 + (-4)(ac - b^2)/(a + c)^2} - 1}{(-4)(ac - b^2)/(a + c)^2}\right\}. \tag{4.24}$$

The mean value theorem will allow us to show that the second term in the sum (4.24) does not contribute to the asymptotic limit. Let

$$f_1(x) = (\sqrt{1 - x} - 1)(-x)^{-1}, \qquad f_2(x) = \log f_1(x). \tag{4.25}$$

By a second order Taylor expansion of the square root function around 1,

$$f_1'(x) = \frac{1}{\sqrt{1 - x}}\frac{-x/2 - (\sqrt{1 - x} - 1)}{x^2} = \frac{1}{\sqrt{1 - x}}\left\{\frac{1}{8} + o(1)\right\} \to \frac{1}{8}, \qquad x \to 0. \tag{4.26}$$

Therefore, $\lim_{x\to 0} f_2'(x) = \frac{1}{4}$. Thus, the second term in the sum (4.24) can be reexpressed as

$$f_2\left(\frac{4(\hat{a}\hat{c} - \hat{b}^2)}{(\hat{a} + \hat{c})^2}\right) - f_2\left(\frac{4(ac - b^2)}{(a + c)^2}\right) = f_2'(\xi_{abc})\left\{\frac{4(\hat{a}\hat{c} - \hat{b}^2)}{(\hat{a} + \hat{c})^2} - \frac{4(ac - b^2)}{(a + c)^2}\right\},$$

where the random variable ξ_{abc} lies between $\frac{4(\hat{a}\hat{c} - \hat{b}^2)}{(\hat{a} + \hat{c})^2}$ and $\frac{4(ac - b^2)}{(a + c)^2}$. As a consequence of Lemma A.1, it satisfies the limit $\xi_{abc} \xrightarrow{P} 0$; hence, $f_2'(\xi_{abc}) \xrightarrow{P} \frac{1}{4}$. By expressions (4.19) and (4.20),

$$\hat{a}\hat{c} - \hat{b}^2 = a(v)^{2(h_1+h_2)}\det W_a(2^j), \qquad ac - b^2 = a(v)^{2(h_1+h_2)}\det EW(2^j). \tag{4.27}$$

Thus,

$$\begin{aligned}
&\frac{(\hat{a}\hat{c} - \hat{b}^2)}{(\hat{a} + \hat{c})^2} - \frac{(ac - b^2)}{(a + c)^2} \\
&\quad = \frac{(\hat{a}\hat{c} - \hat{b}^2) - (ac - b^2)}{(a + c)^2} - (\hat{a}\hat{c} - \hat{b}^2)\left\{\frac{1}{(a + c)^2} - \frac{1}{(\hat{a} + \hat{c})^2}\right\} \\
&\quad = \frac{1}{a(v)^{2(h_2-h_1)}}\left\{\frac{\det W_a(2^j) - \det EW(2^j)}{[O(a(v)^{h_1-h_2}) + (p_{12}^2 + p_{22}^2)b_{22}]^2}\right. \\
&\qquad - \det W_a(2^j)\left[\frac{1}{[O(a(v)^{h_1-h_2}) + (p_{12}^2 + p_{22}^2)b_{22}]^2}\right. \\
&\qquad - \left.\left.\frac{1}{[O_P(a(v)^{h_1-h_2}) + (p_{12}^2 + p_{22}^2)\hat{b}_{22}]^2}\right]\right\}.
\end{aligned} \tag{4.28}$$

After premultiplication by the rate factor $\sqrt{K_{a,j}}$, by (4.2) the expression (4.28) becomes

$$\frac{1}{a(v)^{2(h_2-h_1)}}\left\{\frac{O_P(1)}{[O(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)b_{22}]^2}-\det W_a(2^j)\right.$$

$$\left.\times\left[\frac{O_P(1)}{[O(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)b_{22}]^2[O_P(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)\hat{b}_{22}]^2}\right]\right\} \quad (4.29)$$

$$\xrightarrow{P} 0.$$

The numerators of the first and second terms appearing in the sum on the left-hand side of (4.29) are bounded in probability as a consequence of Theorem 3.1 and of applying the mean value theorem to the function $f_3(x)=x^2$.

We now look at the first term in the sum (4.24). In view of the relations (4.27) and by applying the mean value theorem twice, it can be rewritten as

$$\log\frac{\det W_a(2^j)}{O_P(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)\hat{b}_{22}}-\log\frac{\det EW(2^j)}{O(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)b_{22}}$$

$$\quad (4.30)$$

$$=\frac{1}{\xi_W}\left\{\det W_a(2^j)-\det EW(2^j)\right\}-\frac{1}{\xi_{b_{22}}}\left\{O_P\left(\frac{a(v)^{h_1-h_2}}{\sqrt{K_{a,j}}}\right)+(p_{12}^2+p_{22}^2)(\hat{b}_{22}-b_{22})\right\}.$$

The random variable ξ_W lies between $\det W_a(2^j)$ and $\det EW(2^j)$, whereas the random variable $\xi_{b_{22}}$ lies between $O_P(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)\hat{b}_{22}$ and $O(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)b_{22}$. Moreover, these random variables display asymptotic behavior $\xi_W\xrightarrow{P}\det EW(2^j)$, $\xi_{b_{22}}\xrightarrow{P}(p_{12}^2+p_{22}^2)b_{22}$, $v\to\infty$, where the latter limits stem from Theorem 3.1 and Lemma 4.1, respectively. Therefore, after premultiplication by the rate factor $\sqrt{K_{a,j}}$, the expression (4.30) is asymptotically equivalent in probability to

$$\sqrt{K_{a,j}}\frac{\det\hat{B}_a(2^j)-\det B(2^j)}{\det B(2^j)}-\sqrt{K_{a,j}}\frac{\hat{b}_{22}-b_{22}}{b_{22}}. \quad (4.31)$$

Let $f_4(x,y,z)=xz-y^2$. By the mean value theorem,

$$\det\hat{B}_a(2^j)-\det B(2^j)=\nabla f_4(\xi_{b_{11}},\xi_{b_{12}},\xi_{b_{22}})(\text{vec}_{\mathcal{S}}\hat{B}_a(2^j)-\text{vec}_{\mathcal{S}}B(2^j)), \quad (4.32)$$

where $\nabla f_4(\xi_{b_{11}},\xi_{b_{12}},\xi_{b_{22}})\xrightarrow{P}\nabla f_4(b_{11},b_{12},b_{22})$ as a consequence of Lemma 4.1. Reintroducing j, from (4.29), (4.31) and (4.32), we conclude that

$$\sqrt{K_{a,j}}\left(\log\lambda_1(a(v)2^j)-\log\lambda_1^E(a(v)2^j)\right)$$

$$\overset{P}{\sim}\frac{b_{22}(2^j)}{\det B(2^j)}\sqrt{K_{a,j}}\left(\hat{b}_{11}(2^j)-b_{11}(2^j)\right) \quad (4.33)$$

$$- 2\frac{b_{12}(2^j)}{\det B(2^j)}\sqrt{K_{a,j}}\big(\hat{b}_{12}\big(2^j\big) - b_{12}\big(2^j\big)\big)$$

$$+ \left(\frac{b_{11}(2^j)b_{22}(2^j)}{\det B(2^j)} - 1\right)\sqrt{K_{a,j}}\left(\frac{\hat{b}_{22}(2^j) - b_{22}(2^j)}{b_{22}(2^j)}\right).$$

A similar calculation can be developed for $\lambda_2(a(\nu)2^j) - \lambda_2^E(a(\nu)2^j)$. The latter can be recast as

$$\big\{\log(\hat{a} + \hat{c}) - \log(a + c)\big\} \tag{4.34}$$

$$+ \left\{\log\left(1 + \sqrt{1 - \frac{4(\hat{a}\hat{c} - \hat{b}^2)}{(\hat{a} + \hat{c})^2}}\right) - \log\left(1 + \sqrt{1 - \frac{4(ac - b^2)}{(a + c)^2}}\right)\right\}.$$

By the mean value theorem based on the function $f_5(x) = \log(1 + \sqrt{1 - x})$, after premultiplication by the rate factor $\sqrt{K_{a,j}}$ the second term in the sum (4.34) can be reexpressed as

$$f_5'(\xi_{abc})\sqrt{K_{a,j}}\left\{\frac{4(\hat{a}\hat{c} - \hat{b}^2)}{(\hat{a} + \hat{c})^2} - \frac{4(ac - b^2)}{(a + c)^2}\right\} \xrightarrow{P} 0,$$

since $f_5'(\xi_{abc}) \xrightarrow{P} -\frac{1}{4}$ and by (4.29). As for the first term in the sum (4.34), by the mean value theorem we can rewrite it with probability going to 1 as

$$\log\big\{O_P\big(a(\nu)^{h_1-h_2}\big) + \big(p_{12}^2 + p_{22}^2\big)\hat{b}_{22}\big\} - \log\big\{O\big(a(\nu)^{h_1-h_2}\big) + \big(p_{12}^2 + p_{22}^2\big)b_{22}\big\}$$

$$= \frac{1}{\xi_{b_{22}}}\left\{O_P\left(\frac{a(\nu)^{h_1-h_2}}{\sqrt{K_{a,j}}}\right) + \big(p_{12}^2 + p_{22}^2\big)(\hat{b}_{22} - b_{22})\right\},$$

where $\xi_{b_{22}} \xrightarrow{P} (p_{12}^2 + p_{22}^2)b_{22}$. As a consequence, and reintroducing j,

$$\sqrt{K_{a,j}}\big\{\log\big(\lambda_2\big(a(\nu)2^j\big)\big) - \log\big(\lambda_2^E\big(a(\nu)2^j\big)\big)\big\} \overset{P}{\sim} \sqrt{K_{a,j}}\left(\frac{\hat{b}_{22}(2^j) - b_{22}(2^j)}{b_{22}(2^j)}\right). \tag{4.35}$$

The expressions (4.33) and (4.35) for $j = j_1 < \cdots < j_m$ imply the weak limit (4.22) with asymptotic covariance matrix (4.23). $\square$

Remark 4.1. At this point, we can shed more light on the apparent paradox anticipated in the Introduction: the asymptotic behavior of each entry of the matrices $W(a(\nu)2^j)$ and $EW(a(\nu)2^j)$ (see (4.15) and (4.17)) is generally governed by the power law $a(\nu)^{2h_2}$, whereas that of the eigenvalues $\lambda_1(a(\nu)2^j)$, $\lambda_1^E(a(\nu)2^j)$ is dominated by $a(\nu)^{2h_1}$. For mathematical simplicity, consider only the matrix $EW(a(\nu)2^j)$ under the instance $A = P$ in (1.2), where the columns of P are unit vectors. The operator scaling property (P5) of the wavelet transform yields

$$EW\big(a(\nu)2^j\big) = P\,\mathrm{diag}\big(w_1a(\nu)^{2h_1}, w_2a(\nu)^{2h_2}\big)P^*, \tag{4.36}$$

where $EW(2^j) = P \operatorname{diag}(w_1, w_2) P^*$, for $w_1 = w_1(2^j) > 0$, $w_2 = w_2(2^j) > 0$. Then,

$$\lambda_1^E(a(v)2^j) = \inf_{\mathbf{u} \in S^1} \mathbf{u}^* EW(a(v)2^j)\mathbf{u} \le \mathbf{u}_0^* EW(a(v)2^j)\mathbf{u}_0 = \frac{w_1 a(v)^{2h_1}}{\|\mathbf{u}_0\|^2}, \tag{4.37}$$

where $\mathbf{u}_0 = (P^*)^{-1} e_1 / \|(P^*)^{-1} e_1\|$ and e_1 is the first Euclidean vector. In other words, $\lambda_1^E(a(v)2^j)$ cannot go to infinity faster than $a(v)^{2h_1}$.

Notwithstanding (4.37), it should be stressed that $\lambda_1^E(a(v)2^j)$ does not generally follow an exact power law. This can be seen based on the explicit expression for $\lambda_1^E(a(v)2^j)$ (see (4.4)), that is,

$$\lambda_1^E(a(v)2^j) = \tfrac{1}{2}\{EW(a(v)2^j)_{11} + EW(a(v)2^j)_{22} - \sqrt{\Delta(a(v)2^j)}\} \tag{4.38}$$

where $\Delta(a(v)2^j) = (EW(a(v)2^j)_{11} - EW(a(v)2^j)_{22})^2 + 4(EW(a(v)2^j)_{12})^2$. Under (4.36), the matrix $EW(a(v)2^j) = (EW(a(v)2^j)_{i_1,i_2})_{i_1,i_2=1,2}$ can be expressed entry-wise as in (4.18) with $b_{11}(2^j) = w_1$, $b_{12}(2^j) = 0$ and $b_{22}(2^j) = w_2$. It is of interest to note that the quality of the approximation $\lambda_1^E(a(v)2^j) \sim Ca(v)^{2h_1}$ increases with the difference $h_2 - h_1$. Indeed, by (4.5) and a second order Taylor expansion of the function $f(x) = \sqrt{x}$ around 1,

$$\lambda_1^E(a(v)2^j) = \frac{2 \det EW(2^j)}{w_1 a(v)^{2(h_1-h_2)} + w_2} a(v)^{2h_1}$$

$$\times \left\{ \frac{1}{2} + \frac{1}{2} \frac{\det EW(2^j)}{[w_1 a(v)^{2(h_1-h_2)} + w_2]^2} \frac{1}{a(v)^{2(h_2-h_1)}} + o\left(\frac{1}{a(v)^{2(h_2-h_1)}}\right) \right\}. \tag{4.39}$$

The discussion and conclusions presented in this remark can be easily generalized beyond the instance (4.36).

Remark 4.2. In practice, the choice of $a(v)$ involves a statistical compromise. A large value of $a(v)$ with respect to v implies that the estimator variance is relatively large and the estimator distribution is not very close to Gaussian, but it also yields a relatively small bias. Computational experiments suggest that the ratio $v/a(v)2^j$ should be no less than 2^3.

4.2. The weak limit of unit eigenvectors

For given j, $\log_2 a(v) \in \mathbb{N}$, consider the relation (4.7) for the eigenvector entries

$$\hat{\mathbf{v}}(a(v)2^j) = (\hat{v}_1(a(v)2^j), \hat{v}_2(a(v)2^j))^* \in \mathbb{R}^2 \tag{4.40}$$

associated with the smallest eigenvalue $\lambda_1 := \lambda_1(a(v)2^j)$ of the (symmetric) sample wavelet variance matrix $\hat{S} := W(a(v)2^j)$. Further consider their wavelet variance counterparts $(v_1(a(v)2^j), v_2(a(v)2^j))^*$, $\lambda_1^E := \lambda_1^E(a(v)2^j)$ and $S := EW(a(v)2^j)$. The ratios $(\hat{v}_2/\hat{v}_1)(a(v)2^j)$, $(v_2/v_1)(a(v)2^j)$ represent the tangents of the angles that determine the associated eigenspaces. In the next definition, we propose to use $(\hat{v}_2/\hat{v}_1)(a(v)2^j)$ as an estimator of

$b_{22} \neq 0$, whence $b = b_{j,a(\nu)} \neq 0$ for large ν in expression (4.18). The same applies to $a = a_{j,a(\nu)}$. Therefore, $\nu_1 \neq 0$ in (4.7), and

$$\frac{\nu_2}{\nu_1} = \frac{a}{b}\left(\frac{\lambda_1^E}{a} - 1\right) \sim -\frac{p_{12}}{p_{22}}, \qquad \nu \to \infty, \tag{4.47}$$

by Lemma A.1.

Now assume that $p_{12} = 0$. Then, $p_{11} \neq 0$, and $a = a_{j,a(\nu)} \neq 0$ for large ν. Since $b_{12} \neq 0$ by assumption, then $b = b_{j,a(\nu)} \neq 0$. Again, the limit (4.47) follows.

In regard to $\hat{\theta}(a(\nu)2^j)$, in either case in (4.43) the relations above can be simply rewritten for $\hat{S} = W(a(\nu)2^j)$ with eigenvector $\hat{\mathbf{v}}$. Because of (4.13), $\hat{B}_a(2^j) \xrightarrow{P} B(2^j)$ and thus the expression $\hat{\nu}_2/\hat{\nu}_1$ is well-defined with probability going to 1, as $\nu \to \infty$. Again by Lemma A.1, $\hat{\nu}_2/\hat{\nu}_1 \xrightarrow{P} -\frac{p_{12}}{p_{22}}$.

For the ensuing developments, recall that b_{22}^{-1} is well defined. So, to show (ii), reexpress

$$a(\nu)^{h_2-h_1}\sqrt{K_{a,j}}\left[\left(\frac{\lambda_1 - \hat{a}}{\hat{b}}\right) - \left(\frac{\lambda_1^E - a}{b}\right)\right]$$
$$= a(\nu)^{h_2-h_1}\sqrt{K_{a,j}}\left\{\frac{\lambda_1}{\hat{b}} - \frac{\lambda_1^E}{b}\right\} + a(\nu)^{h_2-h_1}\sqrt{K_{a,j}}\left\{\frac{a}{b} - \frac{\hat{a}}{\hat{b}}\right\}. \tag{4.48}$$

The first term on the right-hand side of (4.48) can be further developed into

$$a(\nu)^{h_2-h_1}\frac{1}{\hat{b}}\sqrt{K_{a,j}}\{\lambda_1 - \lambda_1^E\} + a(\nu)^{h_2-h_1}(-1)\frac{\lambda_1^E}{b}\sqrt{K_{a,j}}\left\{\frac{\hat{b} - b}{\hat{b}}\right\}. \tag{4.49}$$

In view of Lemmas A.1 and 4.1,

$$\frac{\lambda_1^E}{b} \sim \frac{1}{a(\nu)^{2(h_2-h_1)}}\frac{\det EW(2^j)}{p_{12}p_{22}b_{22}^2(p_{12}^2 + p_{22}^2)},$$
$$\sqrt{K_{a,j}}\left\{\frac{\hat{b} - b}{\hat{b}}\right\} \overset{P}{\sim} \sqrt{K_{a,j}}\frac{\{\hat{b}_{22} - b_{22}\}}{\hat{b}_{22}} \xrightarrow{d} \frac{N(0, \sigma^2(b_{22}(2^j)))}{b_{22}(2^j)},$$

where $\sigma^2(b_{22}(2^j))$ comes from the matrix $\Sigma_B(j_1, \ldots, j_m)$ in (4.13). Consequently, in regard to the second term in the sum (4.49),

$$a(\nu)^{h_2-h_1}(-1)\frac{\lambda_1^E}{b}\sqrt{K_{a,j}}\left\{\frac{\hat{b} - b}{\hat{b}}\right\} \xrightarrow{P} 0. \tag{4.50}$$

We now look at the first term in the sum (4.49). Let $\hat{r} = 4\frac{\hat{a}\hat{c} - \hat{b}^2}{(\hat{a}+\hat{c})^2}$ and $r = 4\frac{ac - b^2}{(a+c)^2}$. Then,

$$\sqrt{K_{a,j}}\{\lambda_1 - \lambda_1^E\} = \sqrt{K_{a,j}}\left\{2\left(\frac{\hat{a}\hat{c} - \hat{b}^2}{\hat{a} + \hat{c}}\right) - 2\left(\frac{ac - b^2}{a + c}\right)\right\}\left(\frac{\sqrt{1-\hat{r}} - 1}{-\hat{r}}\right)$$
$$+ 2\left(\frac{ac - b^2}{a + c}\right)\sqrt{K_{a,j}}\left\{\left(\frac{\sqrt{1-\hat{r}} - 1}{-\hat{r}}\right) - \left(\frac{\sqrt{1-r} - 1}{-r}\right)\right\}. \tag{4.51}$$

By the same reasoning leading to (4.29),

$$
\sqrt{K_{a,j}}\left\{2\left(\frac{\hat{a}\hat{c}-\hat{b}^2}{\hat{a}+\hat{c}}\right)-2\left(\frac{ac-b^2}{a+c}\right)\right\}
$$

$$
= a(v)^{2h_1}\left\{\frac{O_P(1)}{O(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)b_{22}}\right. \tag{4.52}
$$

$$
\left.+\det W(2^j)\frac{O_P(1)}{[O_P(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)b_{22}][O(a(v)^{h_1-h_2})+(p_{12}^2+p_{22}^2)\hat{b}_{22}]}\right\}.
$$

Moreover, for $f_1(x)$ as in (4.25), with probability going to 1 the mean value theorem gives $\sqrt{K_{a,j}}\{f_1(\hat{r})-f_1(r)\}=f_1'(\xi_r)\sqrt{K_{a,j}}(\hat{r}-r)$, where the random variable ξ_r lies between r and $\hat{r}$. Thus, by (A.3), (A.4), (4.26) and (4.29),

$$
f_1'(\xi_r)\overset{P}{\to}\frac{1}{8}, \qquad \sqrt{K_{a,j}}(\hat{r}-r)\overset{P}{\sim}\frac{O_P(1)}{a(v)^{2(h_2-h_1)}}. \tag{4.53}
$$

Therefore, by (4.52), (4.53), (A.3) and (A.4), the first term in (4.49) is asymptotically equivalent in probability to

$$
a(v)^{h_2-h_1}\left(\frac{1}{[O_P(a(v)^{h_1-h_2})+p_{12}p_{22}\hat{b}_{22}]}\frac{1}{a(v)^{2h_2}}\right)
$$

$$
\times a(v)^{2h_1}\left\{O_P(1)+\frac{O_P(1)}{a(v)^{2(h_2-h_1)}}\right\}\overset{P}{\to}0. \tag{4.54}
$$

As a consequence, in regard to the first term on the right-hand side of (4.48),

$$
a(v)^{h_2-h_1}\sqrt{K_{a,j}}\left\{\frac{\lambda_1}{\hat{b}}-\frac{\lambda_1^E}{b}\right\}\overset{P}{\to}0. \tag{4.55}
$$

We now turn to the second term on the right-hand side of (4.48). Note that

$$
\frac{a}{b}-\frac{\hat{a}}{\hat{b}}=\frac{a(\hat{b}-b)-b(\hat{a}-a)}{b\hat{b}}\overset{P}{\sim}a(v)^{-4h_2}\frac{a(\hat{b}-b)-b(\hat{a}-a)}{p_{12}^2p_{22}^2b_{22}^2(2^j)}. \tag{4.56}
$$

Note that the ratio on the right-hand side of (4.56) is well defined, since $p_{12}, p_{22}\neq 0$, by assumption. Recall the expressions (4.16) and (4.18) and denote $a(v)^{-2h_2}a$, $a(v)^{-2h_2}b$, $a(v)^{-2h_2}\hat{a}$ and $a(v)^{-2h_2}\hat{b}$, respectively, by

$$
O(a(v)^{2(h_1-h_2)})+\alpha_{12}a(v)^{h_1-h_2}+\alpha_{22}, \qquad O(a(v)^{2(h_1-h_2)})+\beta_{12}a(v)^{h_1-h_2}+\beta_{22},
$$
$$
O_P(a(v)^{2(h_1-h_2)})+\hat{\alpha}_{12}a(v)^{h_1-h_2}+\hat{\alpha}_{22} \quad\text{and}\quad O_P(a(v)^{2(h_1-h_2)})+\hat{\beta}_{12}a(v)^{h_1-h_2}+\hat{\beta}_{22}. \tag{4.57}
$$

For the sake of mathematical clarity, we will henceforth omit the terms of order $a(v)^{2(h_1-h_2)}$ in the expressions (4.57), since the ensuing argument can be easily adapted to account for them. So,

based on (4.57), we can rewrite the numerator in (4.56) as

$$\left(\alpha_{12} a(\nu)^{h_1 - h_2} + \alpha_{22}\right)\left\{(\hat{\beta}_{12} - \beta_{12})a(\nu)^{h_1 - h_2} + (\hat{\beta}_{22} - \beta_{22})\right\}$$

$$- \left(\beta_{12} a(\nu)^{h_1 - h_2} + \beta_{22}\right)\left\{(\hat{\alpha}_{12} - \alpha_{12})a(\nu)^{h_1 - h_2} + (\hat{\alpha}_{22} - \alpha_{22})\right\}$$

$$= a(\nu)^{h_1 - h_2}\left\{\left(\alpha_{12} a(\nu)^{h_1 - h_2} + \alpha_{22}\right)(\hat{\beta}_{12} - \beta_{12}) - \left(\beta_{12} a(\nu)^{h_1 - h_2} + \beta_{22}\right)(\hat{\alpha}_{12} - \alpha_{12})\right\}$$

$$+ a(\nu)^{h_1 - h_2}\left\{\alpha_{12}(\hat{\beta}_{22} - \beta_{22}) - \beta_{12}(\hat{\alpha}_{22} - \alpha_{22})\right\}, \tag{4.58}$$

where the equality is a consequence of the fact that $\alpha_{22}(\hat{\beta}_{22} - \beta_{22}) - \beta_{22}(\hat{\alpha}_{22} - \alpha_{22}) = 0$. When premultiplied by $a(\nu)^{h_2 - h_1}\sqrt{K_{a,j}}$, the first term in the sum (4.58) is asymptotically equivalent in probability to

$$\sqrt{K_{a,j}}\left\{\alpha_{22}(\hat{\beta}_{12} - \beta_{12}) - \beta_{22}(\hat{\alpha}_{12} - \alpha_{12})\right\} \xrightarrow{d} -p_{12}^2 b_{22}(2^j) \det(P) N\left(0, \sigma^2\left(b_{12}(2^j)\right)\right), \tag{4.59}$$

where $\sigma^2(b_{12}(2^j))$ is taken from the matrix $\Sigma_B(j_1, \ldots, j_m)$ in (4.13). In turn, when premultiplied by $a(\nu)^{h_2 - h_1}\sqrt{K_{a,j}}$, the second term in the sum (4.58) is asymptotically equivalent in probability to

$$\sqrt{K_{a,j}}\left\{\alpha_{12}(\hat{\beta}_{22} - \beta_{22}) - \beta_{12}(\hat{\alpha}_{22} - \alpha_{22})\right\} \xrightarrow{d} p_{12}^2 b_{12}(2^j) \det(P) N\left(0, \sigma^2\left(b_{22}(2^j)\right)\right). \tag{4.60}$$

In view of (4.55), and assuming $p_{22} > 0$, the relation (4.46) is obtained by dividing the sum of the weak limits (4.59) and (4.60) by the denominator on the right-hand side of (4.56). □

Remark 4.3. From a different but mathematically equivalent perspective, the statement (4.44) implies that $W(a(\nu)2^j)$ and $EW(a(\nu)2^j)$ have sequences of unit eigenvectors associated with $\lambda_1(a(\nu)2^j)$ and $\lambda_1^E(a(\nu)2^j)$, respectively, which converge (in probability and deterministically, respectively) to the limiting unit vector

$$\left(\frac{|p_{22}|}{\sqrt{p_{12}^2 + p_{22}^2}}, \frac{-p_{12}\operatorname{sign}(p_{22})}{\sqrt{p_{12}^2 + p_{22}^2}}\right). \tag{4.61}$$

If $P \in O(2)$, then eigenvectors of H in two orthogonal directions can be consistently estimated by the eigenvectors of $W(a(\nu)2^j)$. In view of (4.61), there is also a unit eigenvector associated with $\lambda_2(a(\nu)2^j)$ that converges to $(p_{21}, -p_{11})^*$, since the latter is orthogonal to $(p_{22}, -p_{12})^*$. It is easy to see that the latter vectors generate the same (eigen)spaces as $(p_{12}, p_{22})^*$, $(p_{11}, p_{21})^*$, respectively.

Remark 4.4. In regard to the assumptions of Theorem 4.2, when $p_{22} = 0$, it is clear that (4.42) cannot be consistent. By a similar proof to that of Theorem 4.2, asymptotic normality (with different variances) for $\{\hat{\theta}(a(\nu)2^j) - \theta(a(\nu)2^j)\}$ can also be established when $p_{12} = 0$ or $b_{12} = 0$. In the former case, for instance, the convergence rate is $a(\nu)^{h_2 - h_1}\sqrt{K_{a,j}}$ or $\sqrt{K_{a,j}}$, respectively, depending on whether $b_{12} \neq 0$ or both $b_{12} = 0$ and $p_{21} \neq 0$.

Remark 4.5. In Theorem 4.2, the convergence rate is the non-standard $a(\nu)^{h_2-h_1}\sqrt{K_{a,j}} \sim a(\nu)^{h_2-h_1-1/2}\sqrt{\nu/2^j}$, which depends on the parameters to be estimated h_1, h_2. In practice, since $a(\nu)$ is much slower than ν (see condition (4.2)), its effect may not be noticeable (see Section 5 below).

Remark 4.6. Substituting $\lambda_2(a(\nu)2^j)$ for $\lambda_1(a(\nu)2^j)$ in (4.42) does not provide information on p_{11} or p_{21}. Under mild assumptions on the parameters, Lemma A.1 yields

$$\frac{a}{b}\left(\frac{\lambda_2(a(\nu)2^j)}{a} - 1\right) \xrightarrow{P} \frac{p_{22}}{p_{12}}.$$

5. Simulation studies

In this section, we lay out and discuss a broad computational study of the performance of the proposed estimators. OFBM was simulated by means of the Hermir Toolbox, devised in Helgason *et al.* [25,26] and available at www.hermir.org. When describing the results, we drop the asymptotic scaling factor $a(\nu)$ used in Theorems 4.1 and 4.2 and only speak of shifting scales $2^j \in \mathbb{N}$.

The chosen sample path size was $\nu = 2^{16}$. The purpose of picking a large ν was to provide a compelling illustration of the estimators' ability to capture both Hurst eigenvalues. In biomedical applications, typical recordings can be much shorter (of the order $\nu = 2^{10}$; see, for instance, Ivanov [28]). Nevertheless, sample paths of size $\nu = 2^{16}$ are, indeed, encountered in Internet traffic analysis (Abry *et al.* [1]) and hydrodynamic turbulence (Frisch [24]).

In the wavelet analysis, we used least asymmetric orthogonal Daubechies wavelets with $N_\psi = 2$ (Daubechies [13], Chapters 6–8). It was verified that similar conclusions can be obtained when using $N_\psi > 2$ or some other wavelet with a large enough number of vanishing moments.

Remark 5.1. In practice, a continuous time OFBM 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 (Mallat [35], Chapter 7). The algorithm's main input is a discrete time sequence, which, following Veitch *et al.* [49], can be a discrete OFBM sample $\{B_H(k)\}_{k\in T}$, $T \subseteq \mathbb{Z}$. In Section C of Abry and Didier [2], we lay out in detail the mathematical framework for estimation based on discretized wavelet coefficients. In the main result of the section, Theorem C.1, we establish that, under mild conditions, the estimated Hurst eigenvalues and eigenvector stemming from the pyramidal algorithm also satisfy the weak limits (4.22) and (4.46), respectively.

5.1. Entry-wise vs eigenvalue-based estimation

A matrix $W(2^j)$ was computed from a single sample path (of size $\nu = 2^{16}$) of a synthetic OFBM with matrix parameters $P = \begin{pmatrix} 0.98 & 0.57 \\ 0.20 & 0.82 \end{pmatrix}$, $J_H = \mathrm{diag}(0.25, 0.85)$. In the spirit of the entry-wise

approach, the wavelet-based estimation of the Hurst eigenvalues relies on performing a linear regression on a $\log_2 W(2^j)$ vs $j = \log_2 2^j$ diagram, motivated by the log-transformed scalar version of $EW(2^j)$ as in (P5), i.e., $EW(2^j) = C(H)2^{j2H}$ for some $C(H) > 0$. This is shown for each auto- and cross-wavelet components of the bivariate OFBM in Figure 1. The top row displays, in order, plots for $\log_2 W(2^j)_{1,1}$ vs $j = \log_2 2^j$, $\log_2 W(2^j)_{1,2}$ vs $j = \log_2 2^j$ and $\log_2 W(2^j)_{2,2}$ vs $j = \log_2 2^j$. The asymptotic behavior $j2h_2$ is superimposed on each of these plots. In view of the expression (4.16), it is unsurprising that at coarse scales all auto- and cross-components end up driven by the dominant Hurst eigenvalue h_2. In other words, the conspicuous prevalence of the latter precludes the estimation of the Hurst eigenvalue h_1. By contrast, Figure 1, bottom row, displays, in order, plots for $\log_2 \lambda_1(2^j)$ vs $j = \log_2 2^j$ and $\log_2 \lambda_2(2^j)$ vs $j = \log_2 2^j$, as well as the superimposed asymptotic behaviors $j2h_1$ and $j2h_2$, respectively. The eigenvalue-based procedure leads to two scaling laws individually driven by each of the Hurst eigenvalues h_1 and h_2.

5.2. Estimation performance and asymptotic normality

We also numerically synthesized 10,000 bivariate OFBM sample paths to study the finite-sample effectiveness of the normal approximation described in Theorem 4.1. For each path, the estimates $W(2^j)$, $\lambda_1(2^j)$, $\lambda_2(2^j)$ and $\hat{v}_2(2^j)/\hat{v}_1(2^j)$ were computed. Averaging over realizations yields Monte Carlo estimates $\hat{E}\lambda_1(2^j)$, $\hat{E}\lambda_2(2^j)$ and $\hat{E}p_{12}(2^j)/p_{22}(2^j)$ of the ensemble averages $E\lambda_1(2^j)$, $E\lambda_2(2^j)$ and $-E\hat{\theta}(2^j)$, together with estimates of the variances $\hat{\text{Var}}\lambda_1(2^j)$, $\hat{\text{Var}}\lambda_2(2^j)$, $\hat{\text{Var}}\theta(2^j)$.

In Figures 2, 3 and 4, the simulated OFBM has parameter $J_H = \text{diag}(0.25, 0.85)$ and path size $\nu = 2^{16}$ for the three cases, but different mixing matrices of the general form $P = \begin{pmatrix} 1/\sqrt{1+\gamma^2} & \beta/\sqrt{1+\beta^2} \\ \gamma/\sqrt{1+\gamma^2} & 1/\sqrt{1+\beta^2} \end{pmatrix}$. In this parametrization, $\beta = -\theta$ (see (4.41)). The simulated instances are representative of different parametric settings. The choice $\beta = 0.7$, $\gamma = 0.2$ (Figure 2) illustrates the situation of a general mixing matrix P. The choice $\beta = -\gamma$ and $\beta/\sqrt{1+\beta^2} = \sin \pi/6$ (Figure 3) represents the case $P \in O(2)$, whereas $\gamma = 0$ and $\beta = 0.2$ (Figure 4) portrays the case often referred to as fractional cointegration. A comparison of the estimates (black solid lines with 'o') with the theoretical values (red dashed lines) in Figures 2, 3, and 4 reveals the robustness of the proposed estimators (top row). The qq-plots (bottom row) also show that no deviation from a $\mathcal{N}(0, 1)$ distribution can be observed within ± 2 standard deviations for $\log_2 \lambda_1(2^j)$ and $\log_2 \lambda_1(2^j)$, and within ± 2 standard deviations for p_{12}/p_{22}.

The simulations indicate that, beyond asymptotics, Theorem 4.1 for $\hat{h}_1(2^j)$ and $\hat{h}_2(2^j)$, and Theorem 4.2 for $\hat{\theta}(2^j) = -\hat{\beta}(2^j)$ provide effective normal approximations to the finite-sample estimator distributions. Whether or not $P \in O(2)$ does not impact the performance of $\log_2 \lambda_1(2^j)/2j$, $\log_2 \lambda_2(2^j)/2j$ and $\hat{\beta}(2^j)$. However, assuming $P \in O(2)$ additionally allows for the full estimation of P, and thus of H, as discussed in Remark 4.3 and illustrated in Figure 5.
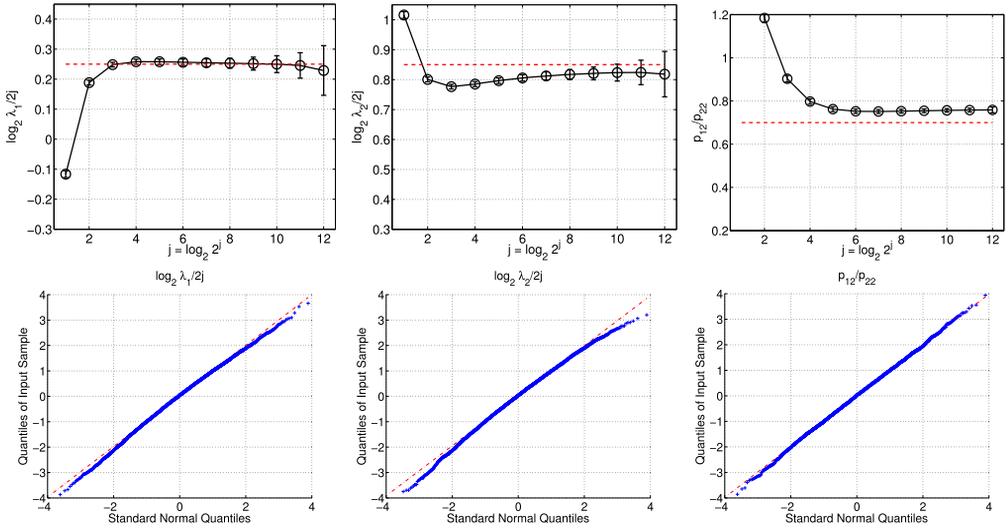
**Figure 2.** *Estimation performance and asymptotic normality*: OFBM with $\gamma = 0.2$ and $\beta = 0.7$. Top row: black solid lines with "o" show $\hat{E}\hat{\vartheta} \pm \sqrt{\hat{\text{Var}}\hat{\vartheta}/n}$ for $\hat{\vartheta} = \log_2 \lambda_1(2^j)/2j$, $\log_2 \lambda_2(2^j)/2j$ and $p_{12}(2^j)/p_{22}(2^j)$ (target parameters: h_1 (left plots), h_2 (center plots), p_{12}/p_{22} (right plots), respectively), red dashed lines corresponding to theoretical values. Bottom row, the corresponding qq-plots (against $\mathcal{N}(0,1)$ distributions) for $j = 10$.

5.3. Beyond the bivariate setting

It is natural to ask whether in higher dimension the eigenvalues of the sample wavelet variance are good estimators of the Hurst eigenvalues. Figure 6 shows eigenvalue-based estimation at work for $n = 4$, with $\log_2 \lambda_p(2^j)/2j$, for $p = 1, 2, 3, 4$, averaged over 2000 realizations of an OFBM with parameters

$$P = \begin{pmatrix} 0.90 & -0.22 & -0.30 & -0.22 \\ 0.43 & 0.45 & 0.63 & 0.46 \\ 0 & -0.85 & 0.40 & 0.30 \\ 0 & 0 & -0.59 & 0.81 \end{pmatrix}$$

and $J_H = \text{diag}(0.20, 0.40, 0.70, 0.90)$, $N = 2^{16}$. The computational results indicate that the smallest and largest eigenvalues of $W(2^j)$ are still good estimators of the smallest and largest entries of J_H. Furthermore, there is evidence that the method produces reasonable estimates of all the intermediate-valued entries of J_H based on the corresponding intermediate eigenvalues of $W(2^j)$. In regard to eigenvector estimation under the assumption $P \in O(4)$, unreported numerical studies suggest that the sample wavelet variance eigenvectors are also good estimators of P.
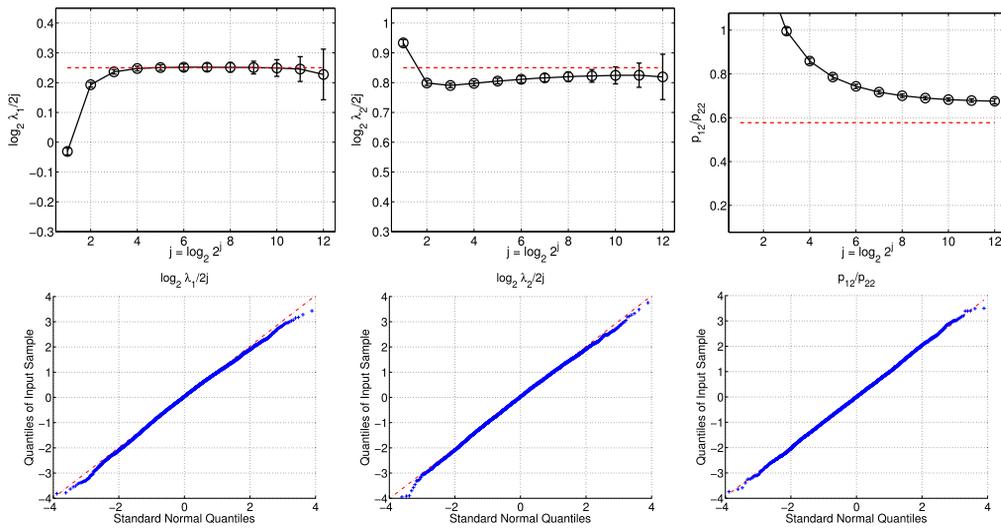
Figure 3. *Estimation performance and asymptotic normality*: OFBM with $\gamma = -\beta$ and $\beta/\sqrt{1+\beta^2} = \sin\pi/6$. Top row: black solid lines with "o" show $\hat{E}\hat{\vartheta} \pm \sqrt{\widehat{\mathrm{Var}}\hat{\vartheta}/n}$ for $\hat{\vartheta} = \log_2\lambda_1(2^j)/2j$, $\log_2\lambda_2(2^j)/2j$ and $p_{12}(2^j)/p_{22}(2^j)$ (target parameters: h_1 (left plots), h_2 (center plots), p_{12}/p_{22} (right plots), respectively), red dashed lines correspond to theoretical values. Bottom row, corresponding qq-plots (against $\mathcal{N}(0,1)$ distributions) for $j = 10$.

6. Perspectives and open issues

OFBM constitutes the natural multivariate extension of the univariate FBM, allowing a different Hurst eigenvalue in each coordinate, in an arbitrary coordinate system. When the mixing matrix P is non-diagonal, the problem of estimating H becomes distinctively multivariate and is not easily amenable to approaches inspired in the univariate context. In this work, we propose a change of perspective from univariate-like, entry-wise scaling relations to the eigenstructure of the sample wavelet variance $W(2^j)$ across scales. In the bivariate setting, this methodology is mathematically shown to yield consistent and asymptotically normal estimates of the Hurst eigenvalues as well as of the eigenspace angle parameter $-p_{12}/p_{22}$. Consequently, the matrix Hurst parameter H can be fully estimated when $P \in O(2)$. Large sample size simulation was used to illustrate and shed further light on the weak limits obtained. The research contained in this paper has lead to five open issues, currently under investigation: (i) the quantitative assessment of the performance of the estimators as a function of sample size; how much data is demanded by the difficult problem of estimating operator-scaling systems, especially in high dimension?; (ii) is there an advantage to using multiple scales in a regression, as in univariate wavelet-based estimation?; (iii) mathematical extensions to OFBM in higher dimension; (iv) the estimation of non-orthogonal coordinate systems; (v) applications in real data. In the near future, a MATLAB toolbox for the estimators proposed in this paper will be made publicly available.
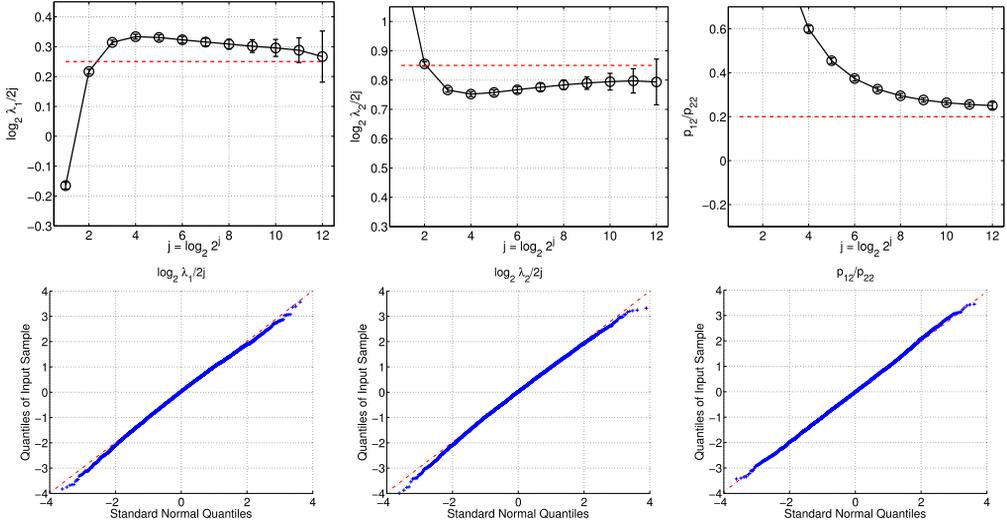
Figure 4. *Estimation performance and asymptotic normality*: OFBM with $\gamma = 0$ and $\beta = 0.2$. Top row: black "o" in solid lines show $\hat{E}\hat{\vartheta} \pm \sqrt{\hat{Var}\hat{\vartheta}/n}$ for $\hat{\vartheta} = \log_2 \lambda_1(2^j)/2j$, $\log_2 \lambda_2(2^j)/2j$ and $p_{12}(2^j)/p_{22}(2^j)$ (target parameters: h_1 (left plots), h_2 (center plots), p_{12}/p_{22} (right plots), respectively), red dashed lines correspond to theoretical values. Bottom row, corresponding qq-plots (against $\mathcal{N}(0, 1)$ distributions) for $j = 10$.

Appendix: Additional results

Proof of Proposition 3.1. We first show (P1). Since the covariance function $EB_H(s)B_H(t)^*$ is continuous, by Cramér and Leadbetter [12], page 86, it suffices to show that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left\| EB_H(s)B_H(t)^* \right\|_{l^1} \left| \psi(s) \right| \left| \psi(t) \right| ds\, dt < \infty. \tag{A.1}$$

In fact,

$$\left\| EB_H(s)B_H(t)^* \right\|_{l^1}$$
$$= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \left| EB_H(s)_{i_1} B_H(t)_{i_2} \right| \tag{A.2}$$
$$\leq \left(\sum_{i_1=1}^{n} \sqrt{EB_H(s)_{i_1}^2} \right) \left(\sum_{i_2=1}^{n} \sqrt{EB_H(t)_{i_2}^2} \right).$$

However, by operator self-similarity, $\|EB_H(t)B_H(t)^*\|_{l^\infty} = \|t^H \Sigma t^{H^*}\|_{l^\infty} \leq C|t|^{2\max \Re \operatorname{eig}(H)}$ for $t \in \mathbb{R}$. Therefore, (A.2) is bounded by $C|t|^{\max \Re \operatorname{eig}(H)}|s|^{\max \Re \operatorname{eig}(H)}$. By conditions (2.6), (2.7)
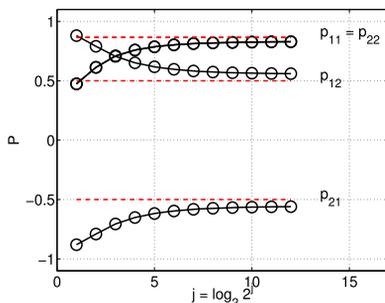
Figure 5. *Estimation performance when $P \in O(2)$: OFBM with $\gamma = -\beta$ and $\beta/\sqrt{1+\beta^2} = \sin \pi/6$. Estimates $\hat{p}_{k,l}(2^j)$, $k, l \in \{1, 2\}^2$ (black solid lines with 'o') of the entries of P, $p_{k,l}$ (red dashed line), from the eigenvectors of the $W(2^j)$. When $P \in O(2)$, both P and J_H can be estimated, and the matrix exponent H is fully estimated.*

and (2.8), (A.1) holds. The former three conditions and the fact that $E\|B_H(t)\|_{l^1} < \infty$ yield

$$E \int_{\mathbb{R}} |\psi(t)| \|B_H(2^j t + 2^j k)\|_{l^1} dt$$

$$\leq C \int_{\mathbb{R}} |\psi(t)| E\|B_H(t+k)\|_{l^1} dt$$

$$\leq CE\|B_H(1)\| \int_{\mathbb{R}} |\psi(s-k)| \|s^H\|_{l^1} ds < \infty,$$

where we used the change of variables $s = t + k$. By Fubini, this yields $ED(2 \cdot k) = 0$, as claimed.

In view of (P1), the properties (P2), (P3) and (P5) and can be established by arguments similar to those for the univariate case (see, for instance, the argument in Delbeke and Abry [16], Theorem 3 for the former two, and Bardet [6], Proposition II.1, for the latter). The property (P6) is a consequence of (P3). Note that $\int_{\mathbb{R}} e^{i2^j xt} \psi(t) dt = C\hat{\psi}(2^j x)$ for some constant $C > 0$. Therefore, the property (P4) is a consequence of the harmonizable representation (1.2), and also of the conditions (2.6), (2.7) and (2.8); the latter ensure that the integrand in (3.2) is well-defined in $\mathbb{R}$. We now show (P7). Since $\psi \in \mathbb{R}$, then $\hat{\psi}(-2^j x) = \overline{\hat{\psi}(2^j x)}$. Thus, by a change of variables $y = -x$ over the integration domain $x < 0$ we can rewrite (3.2) as

$$EW(2^j) = C2^j \int_0^\infty x^{-D} 2\Re(AA^*) x^{-D^*} \frac{|\hat{\psi}(2^j x)|^2}{x^2} dx$$

for some $C > 0$. Since $\text{supp } \hat{\psi}(x)$ has positive Lebesgue measure, then (2.10) yields $v^* EW(2^j)v > 0$, $v \in \mathbb{C}^n \setminus \{0\}$. □
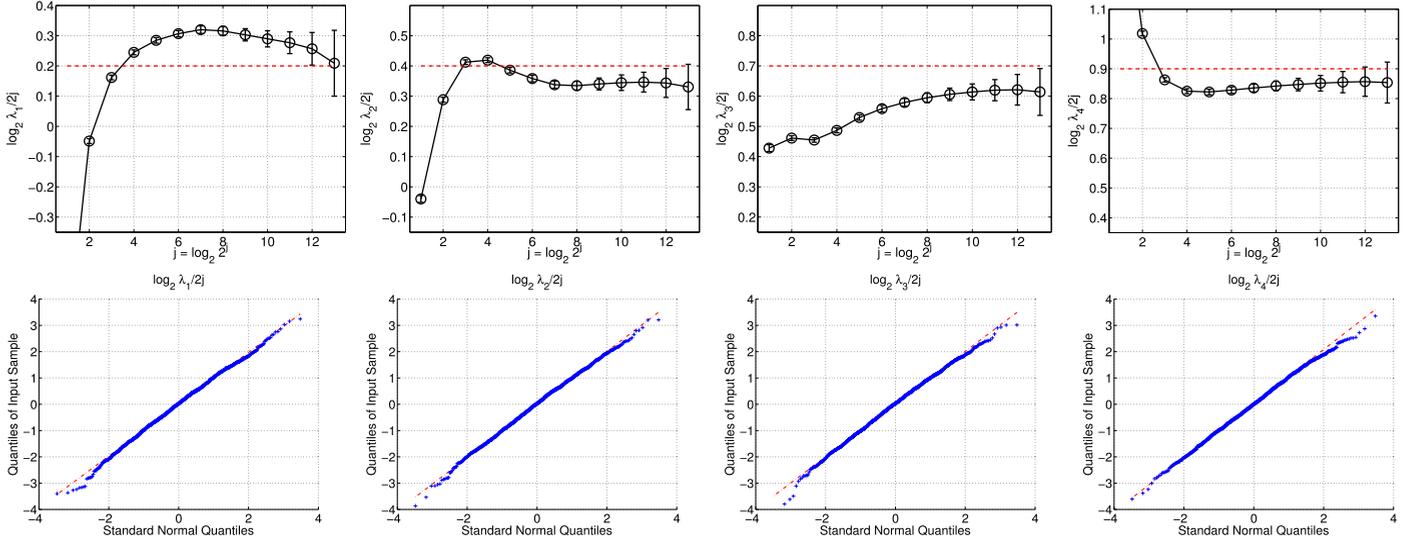
Figure 6. *Estimation performance and asymptotic normality*: OFBM in dimension 4. Top row: black solid lines with "o" show $\hat{E}\hat{\vartheta} \pm \sqrt{\hat{\text{Var}}\hat{\vartheta}/n}$ for $\hat{\vartheta} = \log_2 \lambda_1(2^j)/2j$, $\log_2 \lambda_2(2^j)/2j$, $\log_2 \lambda_3(2^j)/2j$ and $\log_2 \lambda_4(2^j)/2j$ (target parameters: h_1, h_2, h_3, h_4, respectively, with the plots in the same order), red dashed lines correspond to theoretical values. Bottom row, corresponding qq-plots (against $\mathcal{N}(0, 1)$ distributions) for $j = 10$.

Lemma A.1. *Fix $j \in \mathbb{N}$. Under* (OFBM1)–(4) *and* (4.2), *let $EW(a(\nu)2^j)$ and $W(a(\nu)2^j)$ be as in Lemma* 4.2. *Then, the following limits hold, as $\nu \to \infty$:*

$$\frac{4(ac-b^2)}{a+c} \sim \frac{4\det EW(2^j)}{(p_{12}^2+p_{22}^2)b_{22}(2^j)}a(\nu)^{2h_1},$$

$$\frac{4(ac-b^2)}{(a+c)^2} \sim \frac{4\det EW(2^j)}{((p_{12}^2+p_{22}^2)b_{22}(2^j))^2}\frac{1}{a(\nu)^{2(h_2-h_1)}}, \tag{A.3}$$

$$\frac{4(\hat{a}\hat{c}-\hat{b}^2)}{\hat{a}+\hat{c}} \overset{P}{\sim} \frac{4\det W(2^j)}{(p_{12}^2+p_{22}^2)\hat{b}_{22}(2^j)}a(\nu)^{2h_1},$$

$$\frac{4(\hat{a}\hat{c}-\hat{b}^2)}{(\hat{a}+\hat{c})^2} \overset{P}{\sim} \frac{4\det W(2^j)}{((p_{12}^2+p_{22}^2)\hat{b}_{22}(2^j))^2}\frac{1}{a(\nu)^{2(h_2-h_1)}}, \tag{A.4}$$

$$\frac{(1-\sqrt{1-4(ac-b^2)/(a+c)^2})}{4(ac-b^2)/(a+c)^2} \to \frac{1}{2}, \qquad \frac{(1-\sqrt{1-4(\hat{a}\hat{c}-\hat{b}^2)/(\hat{a}+\hat{c})^2})}{4(\hat{a}\hat{c}-\hat{b}^2)/(\hat{a}+\hat{c})^2} \overset{P}{\to} \frac{1}{2}, \quad \text{(A.5)}$$

$$\lambda_1^E \sim \frac{\det EW(2^j)}{(p_{12}^2+p_{22}^2)b_{22}(2^j)}a(\nu)^{2h_1}, \qquad \lambda_1 \overset{P}{\sim} \frac{\det W(2^j)}{(p_{12}^2+p_{22}^2)\hat{b}_{22}(2^j)}a(\nu)^{2h_1}, \tag{A.6}$$

$$\lambda_2^E \sim (p_{12}^2+p_{22}^2)b_{22}(2^j)a(\nu)^{2h_2}, \qquad \lambda_2 \overset{P}{\sim} (p_{12}^2+p_{22}^2)\hat{b}_{22}(2^j)a(\nu)^{2h_2}. \tag{A.7}$$

Proof. In regard to λ_1 and λ_1^E, consider the relation (4.5). By Lemma 4.2,

$$\frac{4(ac-b^2)}{a+c}$$

$$= \frac{4a(\nu)^{2(h_1+h_2)}\det EW(2^j)}{(p_{11}^2+p_{21}^2)b_{11}a(\nu)^{2h_1}+2(p_{11}p_{12}+p_{21}p_{22})b_{12}a(\nu)^{h_1+h_2}+(p_{12}^2+p_{22}^2)b_{22}a(\nu)^{2h_2}}, \tag{A.8}$$

where the determinant is non-trivial due to the property (P7). Again from Lemma 4.2 and by applying Theorem 3.1, an analogous expression holds for $4(\hat{a}\hat{c}-\hat{b}^2)/(\hat{a}+\hat{c})$. Then, the expressions (A.3), (A.4), (A.5), (A.6) follow.

An analogous reasoning applied to λ_2 and λ_2^E leads to (A.7), since the relation (4.6), Lemma 4.2 and Theorem 3.1 show that $\lambda_2^E \sim a+c$, $\lambda_2 \overset{P}{\sim} \hat{a}+\hat{c}$. $\qquad\square$

Acknowledgements

[14] Davidson, J. and De Jong, R. (2000). The Functional Central Limit Theorem and weak convergence to stochastic integrals. *Econometric Theory* **16** 643–666.

[15] Davidson, J. and Hashimzade, N. (2008). Alternative frequency and time domain versions of fractional Brownian motion. *Econometric Theory* **24** 256–293. MR2397267

[16] Delbeke, L. and Abry, P. (2000). Stochastic integral representation and properties of the wavelet coefficients of linear fractional stable motion. *Stochastic Process. Appl.* **86** (2) 177–182. MR1741803

[17] Delgado, R. (2007). A reflected fBm limit for fluid models with ON/OFF sources under heavy traffic. *Stochastic Process. Appl.* **117** 188–201. MR2290192

[18] Didier, G., Helgason, H. and Abry, P. (2015). Demixing multivariate-operator self-similar processes. In *IEEE International Conference on Acoustics*, *Speech and Signal Processing* (*ICASSP*) 1–5. Brisbane, Australia.

[19] Didier, G. and Pipiras, V. (2011). Integral representations and properties of operator fractional Brownian motions. *Bernoulli* **17** 1–33. MR2797980

[20] Didier, G. and Pipiras, V. (2012). Exponents, symmetry groups and classification of operator fractional Brownian motions. *J. Theoret. Probab.* **25** 353–395. MR2914433

[21] Dolado, J.J. and Marmol, F. (2004). Asymptotic inference results for multivariate long-memory processes. *Econom. J.* **7** 168–190. MR2076631

[22] Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory* **38** 910–917. MR1162229

[23] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** 517–532. MR0840512

[24] Frisch, U. (1995). *Turbulence*. Cambridge: Cambridge Univ. Press. The legacy of A.N. Kolmogorov. MR1428905

[25] Helgason, H., Pipiras, V. and Abry, P. (2011). Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding. *Signal Process.* **91** 1123–1133.

[26] Helgason, H., Pipiras, V. and Abry, P. (2011). Synthesis of multivariate stationary series with prescribed marginal distributions and covariance using circulant embedding. *Signal Process.* **91** 1741–1758.

[27] Hudson, W.N. and Mason, J.D. (1982). Operator-self-similar processes in a finite-dimensional space. *Trans. Amer. Math. Soc.* **273** 281–297. MR0664042

[28] Ivanov, P., Amaral, L.N., Goldberger, A., Havlin, S., Rosenblum, M., Struzik, Z. and Stanley, H. (1999). Multifractality in human heartbeat dynamics. *Nature* **399** 461–465.

[29] Kechagias, S. and Pipiras, V. (2015). Definitions and representations of multivariate long-range dependent time series. *J. Time Series Anal.* **36** 1–25. MR3300202

[30] Konstantopoulos, T. and Lin, S.-J. (1996). Fractional Brownian approximations of queueing networks. In *Stochastic Networks* (*New York*, 1995). *Lecture Notes in Statist.* **117** 257–273. New York: Springer. MR1466791

[31] Laha, R.G. and Rohatgi, V.K. (1982). Operator self-similar stochastic processes in $\mathbf{R}_d$. *Stochastic Process. Appl.* **12** 73–84. MR0632393

[32] Maejima, M. and Mason, J.D. (1994). Operator-self-similar stable processes. *Stochastic Process. Appl.* **54** 139–163. MR1302699

[33] Majewski, K. (2003). Large deviations for multi-dimensional reflected fractional Brownian motion. *Stoch. Stoch. Rep.* **75** 233–257. MR1994908

[34] Majewski, K. (2005). Fractional Brownian heavy traffic approximations of multiclass feedforward queueing networks. *Queueing Syst.* **50** 199–230. MR2143121

[35] Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press. MR1614527

[36] Marinucci, D. and Robinson, P.M. (2000). Weak convergence of multivariate fractional processes. *Stochastic Process. Appl.* **86** 103–120. MR1741198

[37] Meerschaert, M. and Scheffler, H.-P. (2003). Portfolio modeling with heavy-tailed random vectors. In *Handbook of Heavy-Tailed Distributions in Finance* (S.T. Rachev, ed.) 595–640. Amsterdam: Elsevier Science B.V.

[38] Meerschaert, M.M. and Scheffler, H.-P. (1999). Moment estimator for random vectors with heavy tails. *J. Multivariate Anal.* **71** 145–159. MR1721965

[39] Moulines, E., Roueff, F. and Taqqu, M.S. (2007). Central limit theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context. *Fractals* **15** 301–313. MR2396718

[40] Moulines, E., Roueff, F. and Taqqu, M.S. (2007). On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter. *J. Time Series Anal.* **28** 155–187. MR2345656

[41] Moulines, E., Roueff, F. and Taqqu, M.S. (2008). A wavelet Whittle estimator of the memory parameter of a nonstationary Gaussian time series. *Ann. Statist.* **36** 1925–1956. MR2435460

[42] Nielsen, F.S. (2011). Local Whittle estimation of multi-variate fractionally integrated processes. *J. Time Series Anal.* **32** (3) 317–335. MR2808259

[43] Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23** 1630–1661. MR1370301

[44] Robinson, P.M. (1995). Log-periodogram regression of time series with long range dependence. *Ann. Statist.* **23** 1048–1072. MR1345214

[45] Robinson, P.M. (2008). Multiple local Whittle estimation in stationary systems. *Ann. Statist.* **36** (5) 2508–2530. MR2458196

[46] Shimotsu, K. (2007). Gaussian semiparametric estimation of multivariate fractionally integrated processes. *J. Econometrics* **137** 277–310. MR2354946

[47] Taqqu, M.S. (2003). Fractional Brownian motion and long-range dependence. In *Theory and Applications of Long-Range Dependence* (P. Doukhan, G. Oppenheim and M.S. Taqqu, eds.) 5–38. Boston, MA: Birkhäuser. MR1956042

[48] Veitch, D. and Abry, P. (1999). A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Trans. Inform. Theory* **45** 878–897. MR1682517

[49] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the automatic selection of the onset of scaling. *Fractals* **11** (4) 377–390. MR2028959

[50] Vignat, C. (2012). A generalized Isserlis theorem for location mixtures of Gaussian random vectors. *Statist. Probab. Lett.* **82** 67–71. MR2863025

[51] Wendt, H., Scherrer, A., Abry, P. and Achard, S. (2009). Testing fractal connectivity in multivariate long memory processes. In *IEEE International Conference on Acoustics, Speech and Signal Processing* 2913–2916. Taipei, Taiwan.

[52] Wornell, G. and Oppenheim, A. (1992). Estimation of fractal signals from noisy measurements using wavelets. *IEEE Trans. Signal Process.* **40** 611–623.