Wavelet-based Estimator for the Hurst Parameters of Fractional Brownian Sheet

Liang Wu and Yiming Ding

Abstract

It is proposed a class of statistical estimators $\hat{H} = (\hat{H}_1, \ldots, \hat{H}_d)$ for the Hurst parameters $H = (H_1, \ldots, H_d)$ of fractional Brownian field via multi-dimensional wavelet analysis and least squares, which are asymptotically normal. These estimators can be used to detect self-similarity and long-range dependence in multi-dimensional signals, which is important in texture classification and improvement of diffusion tensor imaging (DTI) of nuclear magnetic resonance (NMR). Some fractional Brownian sheets will be simulated and the simulated data are used to validate these estimators. We find that when $H_i \geq 1/2$, the estimators are efficient, and when $H_i < 1/2$, there are some bias.

Index Terms

Detection of long-range dependence, Self-similarity, Hurst parameters, Wavelet analysis, Fractional Brownian sheet

I. INTRODUCTION

In numerous fields such as hydrology, biology, telecommunication and economics, the data available for analysis usually have scaling behavior (or long-range dependence and self-similarity) that need to be detected. The key point for detecting scaling behavior is the estimation of the Hurst parameters $H$, which are used in characterizing self-similarity and long-range dependence. The literature on the topic is vast (see e.g. [1]–[9]). Most of the published research concerns the scaling behavior in 1D data. However,
it is also important to detect the scaling behavior in multi-dimensional signals, which is important in
texture classification and improvement of diffusion tensor imaging (DTI) of nuclear magnetic resonance
(NMR). Moreover, many multidimensional data from various scientific areas also have anisotropic nature
in the sense that they have different geometric and probabilistic along different directions. The study
of long-range dependence and self-similarity in multidimensional case is usually based on the model of
anisotropy fractional Gaussian fields such as fractional Brownian sheet (fBs) [10], [11], operator scaling
Gaussian random field (OSGRF) [12] and extended fractional Brownian field (EFBF) [13].

We focus on fBs, which plays an important role in modeling anisotropy random fields with self-
similarity and long-range dependence (see e.g. [11], [14], [15]) since it was first introduced by Kamont
[10]. Fractional Brownian sheet has arisen in the study of texture analysis for classification [16], [17].
Besides research on the texture analysis, fractional Brownian sheets can also be used to drive stochastic
partial differential equations [18]. Moreover, Diffusion-tensor images (DTI) of nuclear magnetic resonance
(NMR) in brain is now based on three-dimensional Brownian motion [19] \( H_1 = H_2 = H_3 \). Compared
with this model, fractional Brownian sheet may be a better model because it can describe the long-range
dependence, which may exist in brain.

The main purpose of this paper is to estimate the Hurst parameters of fractional Brownian sheets
via wavelet analysis. Wavelet analysis has been an efficient tool for the estimation of Hurst parameter,
but most of the published research on the topic focus on the 1D case. Wavelet-based estimators for the
parameter of single-parameter scaling processes (self-similarity and long-range dependence) have been
proposed and studied by Abry et al. (see [11–8]). Such estimators are based on the wavelet expansion of
the scaling process and a regression on the log-variance of wavelet coefficient. It was usually assumed that
wavelet coefficients are uncorrelated within and across octaves. Bardet [20] and Morales [21] removed this
assumption and proved the asymptotic normality of these estimators in the case of fractional Brownian
motion (fBm), proposed a two-step estimator, which has the lowest variance in all of the least square
estimators. Compared with other estimators, such as the R/S method, the periodogram, the maximum-
likelihood estimation, and so on, the wavelet-based estimator performs better in both the statistical and
computational senses, and is superior in robustness (see [4–6] and the references therein). It not only
has small bias and low variance but also leads to a simple, low-cost, scalable algorithm. Besides, the
wavelet-based method can also eliminate some trends (linear trends, polynomial trend, or more) by the
property of its vanishing moments [4], which makes the estimator robust to some nonstationarities.

The paper is organized as follows. In Section II, some properties of Gaussian random variables and
some limiting theorems used in this paper are introduced. In Section III, we obtain some properties of
wavelet coefficient of fBs (see Proposition [I]). Based on these properties, A class of estimators \( \hat{H} \) for the Hurst parameter of fBs are proposed (see Section [III-B]). We also prove that these estimators \( \hat{H} \) are asymptotically normal following the approach in [21] (see Theorem 4 and Section [III-C]). Using the two-step procedure that proposed by Bardet [22], we realize the generalized least squares (GLS) estimator, which has the lowest variance of \( \hat{H} \) (see Section [III-D]). Finally, this two-step estimator \( \hat{H}_{og} \) is applied to the synthetic fBs generated by the method of circulant embedding in the two-dimensional case and three-dimensional case (see Section [IV]).

II. PRELIMINARIES

Definition 1. Given a vector \( H = (H_1, \ldots, H_d) \in (0, 1)^d \), a fractional Brownian sheet (fBs) \( B^H = \{ B^H(t), t \in \mathbb{R}^d_+ \} \) with Hurst parameter \( H \) is a real-valued mean-zero Gaussian random fields with covariance function given by

\[
E(B^H(s)B^H(t)) = \frac{1}{2d} \prod_{i=1}^d (|s_i|^{2H_i} + |t_i|^{2H_i} - |s_i - t_i|^{2H_i}), \quad s, t \in \mathbb{R}^d_+.
\]

(1)

It follows that \( B^H \) is an anisotropic Gaussian random field.

Note that if \( d = 1 \), then \( B^H \) is a fractional Brownian motion (fBm) with Hurst parameter \( H_1 \in (0, 1) \), which as a moving-average Gaussian process is introduced by Mandelbrot and Van Ness [23]. If \( d > 1 \) and \( H_1 = \cdots = H_d = 1/2 \), then \( B^H \) is the Brownian sheet.

Moreover, \( B^H \) has a continuous version for all \( H = (H_1, \ldots, H_d) \in (0, 1)^d \) and is self-similar in the sense that for any diagonal matrix \( A = \text{diag}(a_1, a_2, \ldots, a_d) \), where \( a_i > 0 \) for all \( 1 \leq i \leq d \), the random field \( \{ B^H(At), t \in \mathbb{R}^d \} \) has the same probability law as \( \{ a^H B^H(t), t \in \mathbb{R}^d \} \), where \( a^H = a_1^{H_1} \cdots a_d^{H_d} \) (see e.g. [24]). \( B^H \) also has stationary increments with respect to each variable (see e.g. [25]).

First we will use some properties of Gaussian random variables and some limiting theorems.

Lemma 1. \((X, Y)\) are two-dimensional mean-zero normal variable, i.e., \((X, Y) \sim N(0, \sigma_1^2; 0, \sigma_2^2; \rho)\), then

\[
\text{cov}(X^2, Y^2) = 2\text{cov}^2(X, Y).
\]

Proof: Assume \( \text{cov}(X, Y) = v \). Let \( Z = aX - Y \) s.t. \( \text{cov}(Z, Y) = 0 \), So we have \( a = \sigma_2^2/v \).

\[
\mathbb{E}Z^2 = a^2\mathbb{E}X^2 + \mathbb{E}Y^2 - 2a\mathbb{E}XY = a^2\sigma_1^2 + \sigma_2^2 - 2av.
\]

\[
\mathbb{E}X^2Y^2 = \frac{1}{a^2}\mathbb{E}Z^2Y^2 + \frac{2}{a}\mathbb{E}ZY^3 + \frac{1}{a^2}\mathbb{E}Y^4
\]

\[
= \frac{1}{a^2}\mathbb{E}Z^2Y^2 + \frac{2}{a}\mathbb{E}Z\mathbb{E}Y^3 + \frac{1}{a^2}\mathbb{E}Y^4.
\]
It’s easy to check $\mathbb{E}Z^3 = 0$. Put $a$ and $\mathbb{E}Z^2$ into the above formula, we have $\mathbb{E}X^2Y^2 = \sigma_1^2\sigma_2^2 + 2\nu^2$. Then $\text{cov}(X^2, Y^2) = 2\text{cov}(X, Y)$.

**Definition 2.** Let $f : \mathbb{R}^d \to \mathbb{R}$. $X = (X^{(1)}, \ldots, X^{(d)})'$ is a Gaussian vector, and $f(X)$ has finite second moment, we define the Hermite rank of $f$ with respect to $X$ as

$$\text{rank}(f) := \inf\{ \tau : \exists \text{ polynomial } P \text{ of degree } \tau \text{ with } \mathbb{E}[(f(X) - \mathbb{E}f(X))P(X^{(1)}, \ldots, X^{(d)})] \neq 0 \}.$$ 

The following two lemmas have been proved in [21] (see Lemma 4 and Lemma 5 of [21]).

**Lemma 2.** If $X \sim N(0, \sigma^2)$, then for $q > -1/2$, the rank of $f_q(X) = |X|^q$ is 2.

**Lemma 3.** Let $h_i(x) = f_q(x) = |x|^q, q > -1/2$ for every $1 \leq i \leq d$. Let $t_1, \ldots, t_d$ be $d$ real numbers, and $X = (X^{(1)}, \ldots, X^{(d)})'$ a zero-mean Gaussian vector. Then

$$h(X) = \sum_{i=1}^d t_i h_i(X^{(i)})$$

is a function $g : \mathbb{R}^d \to \mathbb{R}$ of Hermite rank 2.

The following three limiting theorems have been proved in [26], [27] and [28] (see Theorem 9.0 of [26], Theorem 4 of [27] and Theorem 3.3A of [28]).

**Theorem 1 (Cramér-Wold theorem).** Let $X_n = (X^{(1)}_n, \ldots, X^{(d)}_n)$ and $X = (X^{(1)}, \ldots, X^{(d)})$ be random vectors. As $n \to +\infty$, $X_n \to X$ if and only if:

$$\sum_{i=1}^d t_i X^{(i)}_n \to \sum_{i=1}^d t_i X^{(i)}.$$

for each $(t_1, \ldots, t_d) \in \mathbb{R}^d$.

**Theorem 2 (Arcones, 1994).** Let $\{X_j\}_{j=1}^{\infty}$ be a stationary mean-zero Gaussian sequence of $\mathbb{R}^d$-valued vectors. Set $X_j = (X^{(1)}_j, \ldots, X^{(d)}_j)$. Let $f$ be a function on $\mathbb{R}^d$ with rank $\tau, 1 \leq \tau < \infty$. We define

$$r^{(i_1, i_2)}(k) = \mathbb{E}[X^{(i_1)}_m X^{(i_2)}_{m+k}],$$

for $k \in \mathbb{Z}$, where $m$ is any large enough number such that $m, m + k \geq 1$. Suppose that

$$\sum_{k=-\infty}^{\infty} |r^{(i_1, i_2)}(k)|^\tau < \infty,$$

for each $1 \leq i_1, i_2 \leq d$. Then

$$n^{-1/2} \sum_{j=1}^n (f(X_j) - \mathbb{E}f(X_j)) \overset{d}{\to} N(0, \sigma^2),$$

(3)
where
\[ \sigma^2 = \text{Var}(f(X_1)) + 2 \sum_{k=1}^{\infty} \text{Cov}(f(X_1), f(X_{1+k})). \] (4)

**Theorem 3.** Let \( X_n = (X_n^{(1)}, \ldots, X_n^{(d)}) \). Suppose that as \( n \to \infty \),
\[ \frac{X_n - \mu}{b_n} \overset{d}{\to} N(0, \Sigma), \]
where \( \Sigma \) is covariance matrix and \( b_n \to \infty \). Let \( g(x) = (g_1(x), \ldots, g_m(x)), x = (x_1, \ldots, x_d) \), be a vector-valued function for which each component function \( g_i(x) \) is real-valued and has a nonzero differential \( g_i(\mu; t), t = (t_1, \ldots, t_d) \), at \( x = \mu \). Put
\[ D = \left[ \frac{\partial g_i}{\partial x_j} \bigg|_{x=\mu} \right]_{m \times d}. \]

Then
\[ \frac{g(X_n) - g(\mu)}{b_n} \overset{d}{\to} N(0, D\Sigma D'). \]

### III. WAVELET-BASED ESTIMATOR

**A. Wavelet coefficients of fractional Brownian sheet**

Let \( \psi(t) \in L^2(\mathbb{R}) \) be an orthonormal wavelet with compact support and have \( N \geq 1 \) vanishing moments. A wavelet \( \psi(t) \) is said to have a number \( N \geq 1 \) of vanishing moments if
\[ \int t^k \psi(t) dt \equiv 0, \quad k = 0, 1, 2, \ldots, N - 1, \quad \text{and} \quad \int t^N \psi(t) dt \neq 0. \] (5)

The dilations and translations of \( \psi(t) \)
\[ \psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k), \quad j, k \in \mathbb{Z} \] (6)
construct the wavelet orthonormal basis for \( L^2(\mathbb{R}) \). The factors \( 2^j \) and \( j \) are called the scale and octave respectively. The factor \( k \) is called translation. The wavelet orthonormal basis for \( L^2(\mathbb{R}^d) \) can be obtained by simply taking the tensor product functions generated by \( d \) one-dimensional bases: [29], [30]
\[ \psi_{j_1, \ldots, j_d, k_1, \ldots, k_d}(t_1, t_2, \ldots, t_d) = \psi_{j_1, k_1}(t_1) \psi_{j_2, k_2}(t_2) \cdots \psi_{j_d, k_d}(t_d). \] (7)

For convenience, let \( J = (j_1, \ldots, j_d) \in \mathbb{Z}^d, K = (k_1, \ldots, k_d) \in \mathbb{Z}^d, t = (t_1, \ldots, t_d) \in \mathbb{R}^d. \) Then the wavelet coefficients of fractional Brownian sheet \( B^H \) in \( \mathbb{R}^d \) are:
\[ d_{B^H}(J, K) := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \psi_{j_1, k_1}(t_1) \psi_{j_2, k_2}(t_2) \cdots \psi_{j_d, k_d}(t_d) B^H(t) dt_1 \cdots dt_d. \] (8)

Before introducing our estimator for \( B^H \), we present some properties of the wavelet coefficient of \( B^H \).
Proposition 1. Let $\psi(t) \in L^2(\mathbb{R})$ be an orthonormal with compact support and have $N \geq 1$ vanishing moments. The wavelet coefficients of fractional Brownian sheet $B^H$ defined by Eq. (8) have the following properties:

(i) $\mathbb{E}d_{B^H}(J, K) = 0$, and $d_{B^H}(J, K)$ is Gaussian, for any $J, K \in \mathbb{Z}^d$.

(ii) For fixed $J = (j_1, j_2, \ldots, j_d) \in \mathbb{Z}^d$, and $\forall K_1, \ldots, K_n \in \mathbb{Z}^d$, we have

$$d_{B^H}(J, K_1), \ldots, d_{B^H}(J, K_n) \overset{d}{=} 2^{\sum_{i=1}^d j_i(H_i+1/2)}(d_{B^H}(0, K_1), \ldots, d_{B^H}(0, K_n)), \quad (9)$$

where $0$ denotes $(0, 0, \cdots, 0) \in \mathbb{Z}^d$.

(iii) For fixed $J \in \mathbb{Z}^d$, and $\forall K_1, \ldots, K_n, L \in \mathbb{Z}^d$, we have

$$d_{B^H}(J, K_1 + L), \ldots, d_{B^H}(J, K_n + L) \overset{d}{=} (d_{B^H}(J, K_1), \ldots, d_{B^H}(J, K_n)), \quad (10)$$

(iv) For fixed $J = (j_1, \ldots, j_d) \in \mathbb{Z}^d$, $J' = (j'_1, \ldots, j'_d) \in \mathbb{Z}^d$, $K = (k_1, \ldots, k_d) \in \mathbb{Z}^d$, and $K' = (k'_1, \ldots, k'_d) \in \mathbb{Z}^d$, when \( \min_{1 \leq i \leq d} |2^{j_i} k_i - 2^{j'_i} k'_i| \to +\infty \)

$$\mathbb{E}d_{B^H}(J, K)d_{B^H}(J', K') \approx \prod_{i=1}^d |2^{j_i} k_i - 2^{j'_i} k'_i|^{2H_i - 2N}. \quad (11)$$

In the above, $\overset{d}{=} $ means equality in distribution.

Proof: (i)

$$\mathbb{E}d_{B^H}(J, K) = \int_{\mathbb{R}^d} \int \mathbb{B}^H(t) \psi_{j_1,k_1}(t_1) \psi_{j_2,k_2}(t_2) \cdots \psi_{j_d,k_d}(t_d) dt,$$

$\mathbb{B}^H(t)$ is a constant, and $\int \psi_{j,k}(t) dt = 0$, so $\mathbb{E}d_{B^H}(J, K) = 0$. The Gaussianity of $d_{B^H}(J, K)$ follows from the fact that the wavelet coefficient of Gaussian process is Gaussian (see [31]).

(ii) This property can be obtained by the self-similarity of $B^H$. For $\forall K \in \mathbb{Z}^d$

$$d_{B^H}(J, K) = 2^{-\frac{1}{2} \sum_{i=1}^d j_i} \int_{\mathbb{R}^d} \psi(2^{-j_1} t_1 - k_1) \cdots \psi(2^{-j_d} t_d - k_d) B^H(t) dt$$

$$= 2^{\frac{1}{2} \sum_{i=1}^d j_i} \int_{\mathbb{R}^d} \psi(t_1' - k_1) \cdots \psi(t_d' - k_d) B^H(2^{j_1} t_1', \ldots, 2^{j_d} t_d') dt'$$

$$\overset{d}{=} 2^{\sum_{i=1}^d j_i(H_i+1/2)} d_{B^H}(0, K),$$

where $t' = (t_1', \ldots, t_d')$.

Similarly, one can check that, for $\theta_i \in \mathbb{R}$, $1 \leq i \leq n$,

$$\sum_{i=1}^n \theta_i d_{B^H}(J, K_i) \overset{d}{=} 2^{\sum_{i=1}^d j_i(H_i+1/2)} \sum_{i=1}^n \theta_i d_{B^H}(0, K_i).$$
(iii) Let \( d(J, k_1, \ldots, k_d) = d(J, K), \ L = (l_1, \ldots, l_d) \). Without loss of generality, we can take \( J = 0 \).

First we check the formula below,

\[
d_B^{u}(0, k_1 + l_1, \ldots, k_d + l_d) \overset{d}{=} d_B^{u}(0, k_1, k_2 + l_2, \ldots, k_d + l_d).
\]

In fact,

\[
d_B^{u}(0, k_1 + l_1, \ldots, k_d + l_d)
\]

\[
= \int_{\mathbb{R}^d} \psi(t_1 - k_1 - l_1) \cdots \psi(t_d - k_d - l_d)B^H(t_1, \ldots, t_d)dt_1 \cdots dt_d
\]

\[
= \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} \psi(t_1' - k_1)B^H(t_1' + l_1, t_2, \ldots, t_d)dt_1' \cdot \prod_{i=2}^{d} \psi(t_i - k_i - l_i)dt_2 \cdots dt_d.
\]

Since \( \int_{\mathbb{R}} \psi(t_1' - k_1)dt_1' = 0 \) and \( B^H \) has stationary increments with respect to each variable, we have

\[
\int_{\mathbb{R}} \psi(t_1' - k_1)B^H(t_1' + l_1, t_2, \ldots, t_d)dt_1' = \int_{\mathbb{R}} \psi(t_1' - k_1)B^H(t_1', t_2, \ldots, t_d)dt_1'.
\]

The Eq.(12) is proved.

Applying Eq.(12) repeatedly,

\[
d_B^{u}(0, K + L) \overset{d}{=} d_B^{u}(0, k_1, k_2 + l_2, \ldots, k_d + l_d) \overset{d}{=} d_B^{u}(0, k_1, k_2, k_3 + l_3, \ldots, k_d + l_d)
\]

\[
\overset{d}{=} \ldots \overset{d}{=} d_B^{u}(0, k_1, \ldots, k_d) = d_B^{u}(0, K).
\]

Similarly, one can see that, for \( \theta_i \in \mathbb{R}, 1 \leq i \leq n, \)

\[
\sum_{i=1}^{n} \theta_id_B^{u}(0, K_i + L) \overset{d}{=} \sum_{i=1}^{n} \theta_id_B^{u}(0, K_i).
\]

(iv) The covariance function of \( B^H \) is given by Eq.(1). And

\[
\mathbb{E}d_B^{u}(J, K)d_B^{u}(J', K') = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{E}B^H(t)B^H(s) \prod_{i=1}^{d} \psi_j, k_i(t_i)\psi_j', k_i'(s_i)ds dt.
\]

So we have:

\[
\mathbb{E}d_B^{u}(J, K)d_B^{u}(J', K')
\]

\[
= \prod_{i=1}^{d} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_j, k_i(t_i)\psi_j', k_i'(s_i) \frac{1}{2}(|s_i|^{2H_i} + |t_i|^{2H_i} - |s_i - t_i|^{2H_i})ds dt_i
\]

\[
= \prod_{i=1}^{d} \mathbb{E}d_B^{u}, (j_i, k_i)d_B^{u}, (j'_i, k'_i),
\]

where \( B^H \) is fBm with Hurst exponent \( H_i \).
It has been shown that the correlations between wavelet coefficients of fBm satisfy this asymptotic equation \[6\], \[32\], \[33\]

\[ \mathbb{E}d_{B^H}(j,k)d_{B^H}(j',k') \approx |2^j k - 2^j' k'|^{2H-2N}, |2^j k - 2^j' k'| \to \infty, \]

where \(j, j', k, k' \in \mathbb{Z}\), and \(B^H\) is fBm with Hurst exponent \(H\), and the wavelet used has compact support.

The Eq.\((11)\) is proved. \(\blacksquare\)

**Remark 1.** In view of Eq.\((11)\), to avoid long-range dependence for \(d_{B^H}(J,K)\), i.e., to ensure that \(\sum_{J,K \in \mathbb{Z}^d} \mathbb{E}|d_{B^H}(J,K)d_{B^H}(0,0)| < \infty\), one can choose

\[ N > \max_{1 \leq i \leq d} H_i + 1/2, i.e., \max_{1 \leq i \leq d} (2H_i - 2N) < -1. \quad \text{(14)} \]

Under this condition, the correlation of \(d_{B^H}(J,K)\) tends rapidly to 0 at large lags.

**B. Estimation of the Hurst parameter and its asymptotic behavior**

In this subsection, we estimate the Hurst parameter of fractional Brownian sheet in the \(d\)-dimensional finite interval. The following lemma can be obtained by direct calculation.

**Lemma 4.** \(\{B^H(t), t \in \prod_{i=1}^d [0,T_i], T = (T_1, \ldots, T_d) \in \mathbb{R}^d\}\) is a fractional Brownian sheet in the \(d\)-dimensional finite interval. \(\psi(t)\) is an orthogonal wavelet with compact support \([-M, M]\), \(M > 0\) and has \(N \geq 1\) vanishing moments. We define:

\[ d_{B^H,T}(J,K) := \int_0^{T_d} \cdots \int_0^{T_1} \psi_{j_1,k_1}(t_1)\psi_{j_2,k_2}(t_2) \cdots \psi_{j_d,k_d}(t_d)B^H(t)dt_1 \cdots dt_d. \quad \text{(15)} \]

For ease of writing, let \(M \leq 1\). So when \(1 \leq k_i \leq 2^{-j_i}T_i - M, i = 1, \ldots, d,\)

\[ d_{B^H,T}(J,K) = d_{B^H}(J,K), \quad \text{(16)} \]

where \(d_{B^H}(J,K)\) is defined by Eq.\((8)\).

**Definition 3.** Under the conditions of Lemma\((4)\) we define that the available wavelet coefficients for the finite interval fractional Brownian sheet are the wavelet coefficients \(d_{B^H,T}(J,K)\) that satisfy Eq.\((16)\).

Let \(n_J\) be the number of available wavelet coefficients at octave \(J \in \mathbb{Z}^d\).

\[ n_J = \prod_{i=1}^d n_i, \quad n_i := \lfloor 2^{-j_i}T_i - M \rfloor. \]

If there exists \(i_0 \in \{1, \ldots, d\}\) s.t. \(T_{i_0} \to +\infty\), then \(n_J \to +\infty.\)
Under the conditions of Lemma 4, according to Eq. (9), Eq. (10) and Lemma 4, the wavelet coefficients of $B^H(t)$ satisfy the equation,

$$
\mathbb{E}d^2_{B^H,T}(J, K) = \mathbb{E}d^2_{B^H}(J, K) = C_{B^H} 2^{i_1(2H_1+1)+i_2(2H_2+1)+\cdots+i_d(2H_d+1)},
$$

(17)

where the Hurst parameter $H = (H_1, \ldots, H_d) \in (0, 1)^d$, $C_{B^H} = \mathbb{E}d^2_{B^H}(0, 1)$ only depends on $B^H$, $J = (j_1, \ldots, j_d)$, $K = (k_1, \ldots, k_d) \in \mathbb{Z}^d$.

Take the 2-logarithm of both sides:

$$
\log_2 \mathbb{E}d^2_{B^H,T}(J, K) = j_1(2H_1 + 1) + \cdots + j_d(2H_d + 1) + \log_2 C_{B^H}.
$$

(18)

So the estimation of $H$ can be realized by a linear regression in the left part versus $J$ diagram. The $\mathbb{E}d^2_{B^H,T}(J, K)$ can be estimated by

$$
S(J) := 1/n_J \sum_{k_d=1}^{n_d} \cdots \sum_{k_1=1}^{n_1} d^2_{B^H,T}(J, K).
$$

According to Lemma 4

$$
S(J) = 1/n_J \sum_{k_d=1}^{n_d} \cdots \sum_{k_1=1}^{n_1} d^2_{B^H}(J, K).
$$

(19)

In more detail, the $H$ is estimated by the regression

$$
\begin{pmatrix}
\log_2 S(J_1) \\
\vdots \\
\log_2 S(J_m)
\end{pmatrix} =
\begin{pmatrix}
J_1 & 1 \\
\vdots & \vdots \\
J_m & 1
\end{pmatrix}
\begin{pmatrix}
2H_1 + 1 \\
\vdots \\
2H_d + 1 \\
\log_2 C_{B^H}
\end{pmatrix} +
\begin{pmatrix}
\varepsilon(J_1) \\
\vdots \\
\varepsilon(J_m)
\end{pmatrix},
$$

where $J_l = (j_{l,1}, \ldots, j_{l,d}) \in \mathbb{Z}^d$, $l \in \{1, \ldots, m\}$, and $J_1 < \cdots < J_m$. (J_l < J_k in the sense that $j_{l,i} \leq j_{k,i}$, for $i \in \{1, \ldots, d\}$ and the equality can’t hold for all $i$).

Let

$$
L :=
\begin{pmatrix}
\log_2 S(J_1) \\
\vdots \\
\log_2 S(J_m)
\end{pmatrix},
A :=
\begin{pmatrix}
J_1 & 1 \\
\vdots & \vdots \\
J_m & 1
\end{pmatrix},
\alpha :=
\begin{pmatrix}
2H_1 + 1 \\
\vdots \\
2H_d + 1 \\
\log_2 C_{B^H}
\end{pmatrix},
\varepsilon :=
\begin{pmatrix}
\varepsilon(J_1) \\
\vdots \\
\varepsilon(J_m)
\end{pmatrix},
$$

suppose that $m \geq d + 1$, rank($A$) = $d + 1$.

Then least squares estimation gives the estimator:

$$
\hat{\alpha} = (A'\Sigma^{-1}A)^{-1}A\Sigma^{-1}L,
$$

$$
\hat{H} = \hat{\alpha}/2 - 1/2,
$$

(20)
where $\hat{H} = (\hat{H}_1, \cdots, \hat{H}_d, \log_2 C_B H / 2 - 1/2)'$. $\hat{H}_i$ is the wavelet-based estimator of $H_i$, $i \in \{1, \ldots, d\}$, $\Sigma$ is a $m \times m$ full-rank matrix. $\hat{H}$ is selected to minimize

$$\|L - A\alpha\|_\Sigma^2 = (L - A\alpha)'\Sigma^{-1}(L - A\alpha).$$

If $\Sigma = I_{m \times m}$, $\hat{H}$ is the ordinary least squares (OLS) estimator. If $\Sigma = \text{Var} L = \text{Var} \varepsilon$, $\hat{H}$ is the generalized least squares (GLS) estimator.

The following central limit theorem indicates the asymptotic normality of the estimators $\hat{H}$.

**Theorem 4.** Under the conditions of Lemma 4, $\hat{H}$ is defined by Eq. (20), $N > \max_{1\leq i \leq d} H_i + 1/4$. For any $c \in \{1, \ldots, d\}$, $T_c \rightarrow +\infty$, so $n_{l_c} \rightarrow +\infty$ and $n_{H_l} \rightarrow +\infty$, $l \in \{1, \ldots, m\}$ ($n_{H_l} = \prod_{i=1}^d n_{l_i}$ is the same as that in Definition 3), we have

$$n_{m,c}^{1/2}(\hat{H} - U) \overset{d}{\to} N(0, \Sigma_{\hat{H}}),$$

(21)

where $U = (H', \log_2 C_B H / 2 - 1/2)'$ and

$$\Sigma_{\hat{H}} = \frac{1}{4}(A'\Sigma^{-1}A)^{-1}A'\Sigma^{-1}\Sigma_L^{-1}A(A'\Sigma^{-1}A)^{-1},$$

(22)

with $\Sigma_L = D'\Sigma_S D$, $\Sigma_S = (\sigma_{l,h}^2)_{m \times m}$, $\sigma_{l,h}^2$ is defined by (37), and

$$D = \log_2 e \cdot \text{diag}(\text{Ed}_B H / (J_1, K)^{-1}, \ldots, (\text{Ed}_B H / (J_m, K))^{-1}).$$

**Remark 2.** $\Sigma_L/n_{m,c}$ is the asymptotic covariance of $L$. For GLS estimator, $\Sigma$ is the covariance matrix of $L$. If $T_c \rightarrow +\infty$, $\Sigma \rightarrow \Sigma_L/n_{m,c}$, then

$$\Sigma_{\hat{H}} = \frac{1}{4}(A'\Sigma_L^{-1}A)^{-1}.$$  

(23)

**Remark 3.** According to the equation between $T_c$ and $n_{m,c}$ (see Definition 3), Theorem 4 indicates that the convergence rate of $\hat{H}$ is $O(1/\sqrt{T_c})$ as $T_c \rightarrow +\infty$ for any $c \in \{1, \ldots, d\}$. Similarly, we can also obtain that the convergence rate of $\hat{H}$ is $O(1/\sqrt{\prod_{i=1}^d T_i})$ as $T_i \rightarrow +\infty$ for all $i \in \{1, \ldots, d\}$. That is to say, the estimation accuracy of $\hat{H}$ can be improved by increasing the square root of sample volume $\sqrt{\prod_{i=1}^d T_i}$.

**C. Proof of Theorem 4**

The following limit theorem of $S(J)$ can be obtained by using Theorem 2.
Lemma 5. Under the conditions of Lemma 3, $S(J)$ is defined by Eq. (19), $N > \max_{1 \leq i \leq d} H_i + 1/4$. For any $c \in \{1, \ldots, d\}$, $T_c \to +\infty$, so $n_c \to +\infty$ and $n_J \to +\infty$, we have

$$n_c^{1/2}[S(J) - u(J)] \xrightarrow{d} N(0, \sigma^2(J)),$$

where $n_J$ is the same as that in Definition 3 $u(J) = E d_{Bu}(J, K)$.

$$\sigma^2(J) = \frac{1}{(n_J/n_c)^2} \text{Var}( \sum_{k_c=1,k_i \in \{1, \ldots, n_i\}, i \neq c} d_{BU}^2(J, K) ) + \frac{2}{(n_J/n_c)^2} \sum_{k=1}^{\infty} \text{Cov}( \sum_{k_c=1,k_i \in \{1, \ldots, n_i\}, i \neq c} d_{BU}^2(J, K), \sum_{k_c=1+k,k_i \in \{1, \ldots, n_i\}, i \neq c} d_{BU}^2(J, K) ) \tag{25}$$

**Proof:** We construct a sequence of random vectors $X_v$ from the wavelet coefficients $d_{BU}(J, K)$,

$$X_v = (X_v(1), \ldots, X_v(n_J/n_c))' := [d_{BU}(J, K)|_{k_c=v,k_i \in \{1, \ldots, n_i\}, i \neq c}]_{(n_J/n_c) \times 1}, 1 \leq v \leq n_c.$$  

When $n_c \to +\infty$, using Proposition 1, $\{X_v\}_{v=1}^\infty$ is a stationary mean-zero Gaussian sequence. Define

$$h(X_v) = \frac{1}{n_J/n_c} \sum_{i=1}^{n_J/n_c} |X_v(i)|^2.$$  

According to Lemma 3, the Hermite rank of $h(X_v)$ is 2. So by Proposition 1 when $N > \max_{1 \leq i \leq d} H_i + 1/4$, for each $1 \leq i_1, i_2 \leq n_J/n_c$,

$$\sum_{k=-\infty}^{\infty} |r^{(i_1,i_2)}(k)|^2 < \infty,$$

where $r^{(i_1,i_2)}(k) = E[X_m(i_1)X_{m+k}(i_2)]$.

Hence, by Theorem 2, the lemma is true. \hfill \blacksquare

Now consider the asymptotic behavior of the full vector $S = (S(J_1), \ldots, S(J_m))'$. In order to apply Theorem 1 (Cramér-Wold theorem), we need to study the asymptotic behavior of $aS$, $a = (a_1, \ldots, a_m)$ is a fixed but arbitrary element of $\mathbb{R}^m$.

Lemma 6. Under the conditions of Lemma 4 let $a = (a_1, \ldots, a_m)$ be a fixed but arbitrary element of $\mathbb{R}^m$, $S = (S(J_1), \ldots, S(J_m))'$. $N > \max_{1 \leq i \leq d} H_i + 1/4$. For any $c \in \{1, \ldots, d\}$, $T_c \to +\infty$, $n_J = \prod_{i=1}^d n_i$, $u(J_l) = E d_{Bu}^2(J_l, K)$, $l \in \{1, \ldots, m\}$, we have

$$n_m^{1/2}[aS - \sum_{l=1}^m a_l u(J_l)] \xrightarrow{d} N(0, \sigma_{aS}^2), \tag{26}$$
\[
\sigma_{\delta S}^2 = \text{Var} \left( \sum_{l=1}^{m} a_l \frac{n_l c}{u_l n_{J_l}} \sum_{k_c \in \{1, \ldots, u_l\}, k_c \in \{1, \ldots, n_{t,c}\}, i \neq c} d_B^2(J_l, K) \right) + \\
2 \sum_{k=1}^{\infty} \text{Cov} \left( \sum_{l=1}^{m} a_l \frac{n_l c}{u_l n_{J_l}} \sum_{k_c \in \{1, \ldots, u_l\}, k_c \in \{1, \ldots, n_{t,c}\}, i \neq c} d_B^2(J_l, K), \sum_{l=1}^{m} a_l \frac{n_l c}{u_l n_{J_l}} \sum_{k_c \in \{u_l k+1, \ldots, u_l (k+1)\}, k_c \in \{1, \ldots, n_{t,c}\}, i \neq c} d_B^2(J_l, K) \right),
\]

and \(K = (k_1, \ldots, k_d), u_l = 2^{-l m,c - j_{l,c}}, l \in \{1, \ldots, m\}.

**Proof:** We construct a sequence of random vectors \(X_{t,s}\) from the wavelet coefficients \(d_B(J_l, K)\), \(K = (k_1, \ldots, k_d)\).

\[X_{t,s} = (X_{t,s}(1), \ldots, X_{t,s}(n_{J_l}/n_{t,c})) := [d_B(J_l, K)|_{k_c = s, k_c \in \{1, \ldots, n_{t,c}\}, i \neq c}]_{1 \times (n_{J_l}/n_{t,c})}, 1 \leq s \leq n_{t,c}.\]

Because \(J_1 < \cdots < J_m\), according to Definition 3, \(n_{l,c} \geq \cdots \geq n_{m,c}\). One can check \(u_l \approx \frac{n_{l,c}}{n_{m,c}}\). Without loss of generality, we suppose that \(n_{t,c} = u_l n_{m,c}, l \in \{1, \ldots, m\}\).

Let

\[X_v = (X_{1,u_1(v-1)+1}, \ldots, X_{1,u_1v}, \ldots, X_{m,u_mv(v-1)+1}, \ldots, X_{m,u_mv(v-1)}, 1 \leq v \leq n_{m,c}.\]

When \(T_c \rightarrow +\infty, n_{m,c} \rightarrow +\infty\), using Proposition 1, \(\{X_v\}_{v=1}^{\infty}\) is a stationary mean-zero Gaussian sequence. Define

\[h(X_v) = \sum_{l=1}^{m} a_l \frac{1}{u_l n_{J_l}/n_{t,c}} \sum_{j=1}^{n_{J_l}/n_{t,c}} \sum_{i=1}^{u_l} |X_{l,u_i(v-1)+j}(i)|^2.\]

By Lemma 3 the Hermite rank of \(h(X_v)\) is 2.

We rewrite \(X_v\) as

\[X_v = (X_v(1), \ldots, X_v(\sum_{l=1}^{m} u_l n_{J_l}/n_{t,c}))'.\]

So by Proposition 1 when \(N > \max_{1 \leq l \leq d} H_l + 1/4\), it’s easy to check, for each \(1 \leq i_1, i_2 \leq \sum_{l=1}^{m} u_l n_{J_l}/n_{t,c},\)

\[\sum_{k=-\infty}^{\infty} |r(i_1, i_2)(k)|^2 < \infty,\]

where \(r(i_1, i_2)(k) = \mathbb{E}[X_m(i_1)X_{m+k}(i_2)].\)

Hence, by Theorem 2, Lemma 5 is proved.

So by Theorem 1 (Cramér-Wold theorem), the asymptotic behavior of the vector \(S\) can be obtained.
Lemma 7. Under the conditions of Lemma 4, $S = (S(J_1), \ldots, S(J_m))'$. $N > \max_{1 \leq i \leq d} H_i + 1/4$. For any $c \in \{1, \ldots, d\}$, $T_c \to +\infty$, we have

$$n_{m,c}^{1/2}[S - u] \xrightarrow{d} N(0, \Sigma_S),$$

(28)

where

$$u = (\mathbb{E}d_{BH}^2(J_1, K), \ldots, \mathbb{E}d_{BH}^2(J_m, K)),$$

(29)

and

$$\Sigma_S = (\sigma_{l,h}^2)_{m \times m},$$

(30)

with

$$\sigma_{l,h}^2 = \frac{1}{u_i n_i / n_{l,c}} \frac{1}{u_i n_i / n_{h,c}} \left[ \operatorname{Cov} \left( \sum_{k=1}^{\infty} d_{BH}^2(J_k, K), \sum_{k=1}^{\infty} d_{BH}^2(J_k, K) \right) \right]$$

$$+ \sum_{k=1}^{\infty} \operatorname{Cov} \left( \sum_{k=1}^{\infty} d_{BH}^2(J_k, K), \sum_{k=1}^{\infty} d_{BH}^2(J_k, K) \right)$$

$$+ \sum_{k=1}^{\infty} \operatorname{Cov} \left( \sum_{k=1}^{\infty} d_{BH}^2(J_k, K), \sum_{k=1}^{\infty} d_{BH}^2(J_k, K) \right),$$

(31)

$1 \leq l, h \leq m$, and $u_i = 2^{m-c-J_i, c}$. $n_{l,c} = \prod_{i=1}^{d} n_{l,i}$, $l \in \{1, \ldots, m\}$.

$\sigma_{l,h}^2$ can be calculated by $\sigma_S^2$ in Lemma 6.

Note that $\sigma_{l,l}^2$ is equal to $\sigma^2(J_i)/u_i$ in Lemma 5. In fact,

$$\sigma_{l,l}^2 = \frac{1}{(u_i n_i / n_{l,c})^2} \left[ \operatorname{Var} \left( \sum_{k=1}^{\infty} d_{BH}^2(J_k, K) \right) \right]$$

$$+ 2 \sum_{k=1}^{\infty} \operatorname{Cov} \left( \sum_{k=1}^{\infty} d_{BH}^2(J_k, K), \sum_{k=1}^{\infty} d_{BH}^2(J_k, K) \right)$$

$$= \frac{1}{(u_i n_i / n_{l,c})^2} \left[ \sum_{j=1}^{u_i} \operatorname{Var} \left( \sum_{k=1}^{\infty} d_{BH}^2(J_k, K) \right) \right].$$
+ \sum_{j=1}^{u} 2 \sum_{k=1}^{\infty} \text{Cov}\left( \sum_{k_c=j, k_c \in \{1, \ldots, n, l_i\}, i \neq c} d^2_{B_H}(J_i, K), \sum_{k_c=j+k, k_c \in \{1, \ldots, n, l_i\}, i \neq c} d^2_{B_H}(J_i, K) \right) \right].

Since the sequence of random vectors $X_v$ in Lemma 5 is stationary,

$$\sigma^2_{l,l} = \frac{1}{u_l(n_J/n_l, c)^2} \text{Var}\left( \sum_{k_c=1, k_c \in \{1, \ldots, n, l_i\}, i \neq c} d^2_{B_H}(J_i, K) \right)$$

$$+ 2 \sum_{k=1}^{\infty} \text{Cov}\left( \sum_{k_c=1, k_c \in \{1, \ldots, n, l_i\}, i \neq c} d^2_{B_H}(J_i, K), \sum_{k_c=1+k, k_c \in \{1, \ldots, n, l_i\}, i \neq c} d^2_{B_H}(J_i, K) \right) \right] = \sigma^2(J)/u_l.$$

By Theorem 3 one gets the asymptotic behavior of the logarithm of $S$.

**Lemma 8.** Under the conditions of Lemma 4, $L = (\log_2 S(J_1), \ldots, \log_2 S(J_m))^T$. $N > \max_{1 \leq i \leq d} H_i + 1/4$. For any $c \in \{1, \ldots, d\}$, $T_c \to +\infty$, we have

$$n^{1/2}_{m,c} [L - Lu] \overset{d}{\to} N(0, \Sigma_L), \quad (32)$$

where

$$Lu = (\log_2 E d^2_{B_H}(J_1, K), \ldots, \log_2 E d^2_{B_H}(J_m, K)), \quad (33)$$

and

$$\Sigma_L = D \Sigma_S D^T. \quad (34)$$

with $D = \log_2 e \cdot \text{diag}(\{E d^2_{B_H}(J_1, K)\}^{-1}, \ldots, \{E d^2_{B_H}(J_m, K)\}^{-1})$, and $\Sigma_S$ is defined by (30) and (31), the same as that in Lemma 7.

The estimator $\hat{H}$ denoted by Eq. (20) is linear combinations of the elements of $L$. So the asymptotic normality of $\hat{H}$ follows from the asymptotic normality of $L$ by Theorem 3.

**D. Two-Step estimator**

We adopt the GLS estimator, which has the minimum variance in these estimators defined by (20) (see Gauss-Markov theorem [34]). But the covariance matrix $\text{Var} L$ used by GLS estimator is a function of the unknown parameter $H$ we want to estimate. So a two-step estimator may be feasible, namely, we first obtain the OLS estimator of $H$, and then use the OLS estimator to estimate the covariance matrix. So the GLS estimator can be obtained using the estimate of the covariance matrix. This estimator has
been used to the estimation of Hurst parameter of fBm and is proved to be asymptotically normal (see [20]–[22]).

In this subsection, we apply two-step estimator to the case of fractional Brownian sheet.

First we obtain the covariance matrix of $L$ with known $H$. By the regularity on $R_+$ of logarithm function and Theorem 3 when $n, j_h, n, j_i$ is large, the $(h, l)$ element of the covariance matrix $\text{Var} L$ is given by

$$
\text{Cov}(\log_2 S(J_h), \log_2 S(J_l)) \approx \frac{(\log_2 e)^2 \text{Cov}(S(J_h), S(J_l))}{\mathbb{E}d_{BH}^2(J_h, K)\mathbb{E}d_{BH}^2(J_l, K)}
$$

$$
\approx \frac{\sum_{h_i, h_n} \text{Cov}(d_{BH}^2(J_h, K_h), d_{BH}^2(J_l, K_l))}{\mathbb{E}d_{BH}^2(J_h, K)\mathbb{E}d_{BH}^2(J_l, K)}
$$

$$
\approx \frac{2\sum_{h_i, h_n} \text{Cov}(d_{BH}^2(J_h, K_h), d_{BH}^2(J_l, K_l))}{\mathbb{E}d_{BH}^2(J_h, K)\mathbb{E}d_{BH}^2(J_l, K)}.
$$

(by Lemma 1) (35)

Consider $\text{Cov}(d_{BH}^2(J_h, K_h), d_{BH}^2(J_l, K_l))$, by Eq. (13).

$$
\text{Cov}(d_{BH}^2(J_h, K_h), d_{BH}^2(J_l, K_l)) = \prod_{i=1}^{d} \int \int \psi_{j_h, i, k_h, i}(t)\psi_{j_l, i, k_l, i}(s)(-|t - s|^{2H_i}/2)dt\,ds,
$$

where $J_l = (j_l, 1, \ldots, j_l, d) \in \mathbb{Z}_{+}^{d}$, $J_h = (j_h, 1, \ldots, j_h, d) \in \mathbb{Z}_{+}^{d}$, $K_l = (k_l, 1, \ldots, k_l, d) \in \mathbb{Z}_{+}^{d}$ and $K_h = (k_h, 1, \ldots, k_h, d) \in \mathbb{Z}_{+}^{d}$, $l, h \in \{1, \ldots, m\}$.

Using Eq. (36), $\text{Cov}(d_{BH}^2(J_h, K_h), d_{BH}^2(J_l, K_l))$ can be calculated via two-dimensional wavelet transform of function $g(t, s) = -|t - s|^{2H_i}/2$ for each $i \in \{1, \ldots, d\}$. $\mathbb{E}d_{BH}^2(J_h, K)$ and $\mathbb{E}d_{BH}^2(J_l, K)$ can be calculated similarly.

When $H$ is known, the covariance matrix of $L$ can be calculated approximately. We denote the approximately calculated covariance matrix by $G(H)$, which is a function of $H$.

In view of Eq.(35) and Eq. (36), $G(H)$ is a positive definite symmetric matrix and differentiable in $H$ for $H \in (0, 1)^d$. So $G(H)^{-1}$ exists and is continuous in $H$ for $H \in (0, 1)^d$. Now, we can write the two-step estimator,

$$
\hat{\alpha} = (A'G(\hat{H}_o)^{-1}A)^{-1} AG(\hat{H}_o)^{-1}L,
$$

$$
\hat{H}_{o_g} = \frac{\hat{\alpha}}{2} - 1/2.
$$

(37)

where $\hat{H}_o$ is the OLS estimator.

**Theorem 5.** The two-step estimator $\hat{H}_{o_g}$ denoted by Eq. (37) is asymptotically normal, and has the same asymptotic covariance as the GLS estimator.
**Proof:** The asymptotic normality follow from Theorem 4.

By Eq. (35) and Eq. (36), it’s easy to check that $G(H)$ is a positive symmetric matrix and differentiable in $H$ for $H \in (0, 1)$. And from Theorem 4, for any $c \in \{1, \ldots, d\}$, $\hat{H}_o \overset{P}{\to} T_c \to \infty$ $H$. So we have

$$G(\hat{H}_o)^{-1} \overset{P}{\to} T_c \to \infty G(H)^{-1}. \tag{38}$$

Another, $G(H) = D^\prime \text{Var}(S) D$, where $D = \log_2 e \cdot \text{diag}( (\mathbb{E}d_{BH}(J_1, K))^{-1}, \ldots, (\mathbb{E}d_{BH}(J_m, K))^{-1} )$, $S = (S(J_1), \ldots, S(J_m))'$. By Lemma 7, $\text{Var}(S) \to T_c \to \infty \Sigma_S/n_{m,c}$. So

$$G(H) \to T_c \to \infty \Sigma_L/n_{m,c}. \tag{39}$$

$\Sigma_L/n_{m,c}$ is the asymptotic covariance of $L$.

Combining (38) and (39), the two-step estimator has the same asymptotic covariance as the GLS estimator.

**IV. SIMULATION RESULTS AND DISCUSSION**

In this section, we use the estimator $\hat{H}_{og}$ to estimate the Hurst parameter of generated fBs. The data we test are generated by the method of circulant embedding of the covariance matrix [35], [36].

**a) Selection of parameters:** Before the estimation, the octaves $J$ and the number of vanishing moments $N$ (or wavelet) must be chosen. For octaves $J = (j_1, \cdots, j_d)$, we choose all the octaves $J$ that satisfy $J_1 \leq J \leq J_m$, $J_1 = (j_{1,1}, \cdots, j_{1,d})$ is the lower bound of $J$, $J_m = (j_{m,1}, \cdots, j_{m,d})$ is the lower bound of $J$. Based on prior studies [5], [7], for each $i \in \{1, \ldots, d\}$, $j_{m,i}$ is chosen the largest possible given the data length of dimension $i$. $j_{1,i}$ is chosen 3 by the minimum mean square error (MSE). It is known that there is a bias-variance trade-off for choosing octaves: the selection of small octave increase the bias, but decrease the variance of the estimator. The MSE allows the tradeoff between variance and bias.

We choose the classical Daubechies wavelet, which are orthonormal and have compact support. From Theorem 4, the number of vanishing moments must be chosen $N \geq 2$. An increase in the number of vanishing moments $N$ comes with an enlargement of the compact support $([-M, M])$ [7]. In the case of finite time, this will lead to the decrease of the number of available wavelet coefficients $n_J$ at each octaves. So we choose $N = 3$.

**b) Estimation performance:** The estimator $\hat{H}_{og}$ for each $H$ is applied to 500 independent copies of size $512 \times 512$ for the two-dimensional case and 100 independent copies of size $256 \times 256 \times 256$ for three-dimensional case. All the results are shown in Tab. I.
Tab. I shows that both \( \hat{H}_o \) and \( \hat{H}_{og} \) have small MSE. For all the estimated \( H \), the means of \( \hat{H}_o \) and \( \hat{H}_{og} \) are both less than \( H \), and the two estimators have the similar bias, but the variance and MSE of \( \hat{H}_{og} \) are less than that of \( \hat{H}_o \). Besides, we can also find that the smaller the \( H_i \) is, the larger the bias and MSE are.

**c) Discussion:** From our simulation results, we can conclude that both two-step estimator \( \hat{H}_{og} \) and OLS estimator \( \hat{H}_o \) perform very well in the sense of small MSE. It can be also seen that \( \hat{H}_{og} \) perform better than \( \hat{H}_o \) because \( \hat{H}_{og} \) has smaller variance than \( \hat{H}_o \), which is consistent with the theoretical result we have discussed in Section III-D.

Both estimators are not unbiased because

\[
\mathbb{E} \log_2 S(J) \neq \log_2 \mathbb{E} S(J) = \log_2 \mathbb{E} d_H^2(J, K).
\]

and all the theorems are based on the case of continuous time, but the data we used is discrete, the means of both estimators are less than truth parameters as we can see the simulation results. The bias becomes larger as the value of \( H \) becomes smaller. For \( H_i \geq 0.5 \), the bias of \( \hat{H}_i \) is small and can be ignored. For \( H_i < 0.5 \), the bias of \( \hat{H}_i \) is obvious but still small. The estimator for \( H_i \geq 0.5 \) performs better than that for \( H_i < 0.5 \).

**V. Conclusions**

In this paper, we proposed a class of statistical estimators \( \hat{H} = (\hat{H}_1, \ldots, \hat{H}_d) \) for the Hurst parameters \( H = (H_1, \ldots, H_d) \) of fractional Brownian field via multidimensional wavelet analysis and least squares. The proposed estimators are based on the wavelet expansion of fractional Brownian sheet and a regression on the log-variance of wavelet coefficient. We also prove that our estimators \( \hat{H} \) are asymptotically normal following the approach in [21]. The convergence rate of \( \hat{H} \) indicates that the estimation accuracy of \( \hat{H} \) can be improved by increasing the square root of sample volume \( \sqrt{\prod_{i=1}^dT_i} \). The main difficulty of our proof is the asymmetric normality of linear combination of \( S \) (see Lemma 6), which is achieved by the construction of the sequence of random vectors \( \{X_v\}_{v=1}^\infty \) from wavelet coefficients of fBs and the function \( h(X_v) \) that satisfy the condition of Theorem 2. Using the two-step procedure, we realize the generalized least squares (GLS) estimator, which has the lowest variance of \( \hat{H} \). We simulated some fractional Brownian sheets and used them to validate the two-step estimator. We find that when \( H_i \geq 1/2 \), the estimators are efficient, and when \( H_i < 1/2 \), there are some bias.

The computation code of our method is available upon request.
TABLE I

ESTIMATION PERFORMANCE FOR $\hat{H}_o$ AND $\hat{H}_{og}$

| Dimension | $H$   | $\hat{H}_o$ | $\hat{H}_{og}$ |
|-----------|-------|-------------|----------------|
|           | Mean  | Std         | RMSE           | Mean  | Std         | RMSE           |
| 2D        |       |             |                |       |             |                |
| $H_1 = 0.3$ | 0.2733 | 0.0297     | 0.0399         | 0.2706 | 0.0144     | 0.0327         |
| $H_2 = 0.3$ | 0.2732 | 0.0291     | 0.0396         | 0.2705 | 0.0139     | 0.0326         |
| $H_1 = 0.5$ | 0.4897 | 0.0301     | 0.0318         | 0.4911 | 0.0150     | 0.0174         |
| $H_2 = 0.5$ | 0.4855 | 0.0322     | 0.0353         | 0.4889 | 0.0155     | 0.0191         |
| $H_1 = 0.8$ | 0.7915 | 0.0323     | 0.0334         | 0.7958 | 0.0165     | 0.0170         |
| $H_2 = 0.8$ | 0.7917 | 0.0344     | 0.0354         | 0.7961 | 0.0171     | 0.0176         |
| $H_1 = 0.3$ | 0.2731 | 0.0303     | 0.0405         | 0.2708 | 0.0140     | 0.0324         |
| $H_2 = 0.3$ | 0.2736 | 0.0312     | 0.0340         | 0.4892 | 0.0146     | 0.0181         |
| $H_1 = 0.8$ | 0.7931 | 0.0296     | 0.0304         | 0.7980 | 0.0150     | 0.0151         |
| $H_2 = 0.8$ | 0.7911 | 0.0337     | 0.0348         | 0.7971 | 0.0152     | 0.0155         |
| $H_1 = 0.7$ | 0.6886 | 0.0349     | 0.0367         | 0.6937 | 0.0174     | 0.0185         |
| $H_2 = 0.7$ | 0.7902 | 0.0334     | 0.0347         | 0.7959 | 0.0169     | 0.0174         |
| $H_1 = 0.6$ | 0.5929 | 0.0171     | 0.0184         | 0.5943 | 0.0085     | 0.0102         |
| 3D        |       |             |                |       |             |                |
| $H_2 = 0.7$ | 0.6942 | 0.0172     | 0.0181         | 0.6973 | 0.0080     | 0.0084         |
| $H_3 = 0.8$ | 0.7955 | 0.0152     | 0.0158         | 0.7970 | 0.0087     | 0.0091         |
| $H_1 = 0.7$ | 0.6941 | 0.0186     | 0.0195         | 0.6967 | 0.0095     | 0.0100         |
| $H_2 = 0.8$ | 0.7931 | 0.0181     | 0.0193         | 0.7959 | 0.0089     | 0.0098         |
| $H_3 = 0.9$ | 0.8962 | 0.0167     | 0.0171         | 0.8980 | 0.0105     | 0.0106         |

The estimation performance for OLS estimator is given on the left, and for two-step estimator is given on the right. Std denotes standard deviation, RMSE denotes root of MSE. These values are obtained from 500 realizations of generated fBs of size $512 \times 512$, or 100 realizations of generated fBs of size $256 \times 256 \times 256$. It is shown that the means of $\hat{H}_o$ and $\hat{H}_{og}$ are both less than $H$, and the two estimators have the similar bias, but the Std and RMSE of $\hat{H}_{og}$ are less than that of $\hat{H}_o$ (for more detail see Section IV).

REFERENCES

[1] P. Abry, P. Gonçalvès, and P. Flandrin, *Wavelets, spectrum analysis and 1/f processes*, ser. Lecture Notes in Statistics. Springer New York, 1995, vol. 103, book section 2, pp. 15–29.

[2] P. Flandrin, “Wavelet analysis and synthesis of fractional brownian motion,” *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 910–917, 1992.

[3] L. Delbeke and W. Van Assche, “A wavelet based estimator for the parameter of self-similarity of fractional brownian motion,” in *3rd International Conference on Approximation and Optimization in the Caribbean (Puebla, 1995)*, vol. 24, Conference Proceedings, pp. 65–76.
[4] P. Abry and D. Veitch, “Wavelet analysis of long-range-dependent traffic,” *Information Theory, IEEE Transactions on*, vol. 44, no. 1, pp. 2–15, 1998.

[5] D. Veitch and P. Abry, “A wavelet-based joint estimator of the parameters of long-range dependence,” *Information Theory, IEEE Transactions on*, vol. 45, no. 3, pp. 878–897, 1999.

[6] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch, “Wavelets for the analysis, estimation and synthesis of scaling data,” *Self-similar network traffic and performance evaluation*, pp. 39–88, 2000.

[7] ——, “Self-similarity and long-range dependence through the wavelet lens,” *Theory and applications of long-range dependence*, pp. 527–556, 2003.

[8] P. Abry, H. Helgason, and V. Pipiras, “Wavelet-based analysis of non-gaussian long-range dependent processes and estimation of the hurst parameter,” *Lithuanian Mathematical Journal*, vol. 51, no. 3, pp. 287–302, 2011. [Online]. Available: [http://dx.doi.org/10.1007/s10986-011-9126-4](http://dx.doi.org/10.1007/s10986-011-9126-4)

[9] Y. Ding, Y. Hu, W. Xiao, and L. Yan, “Long-memory processes and applications,” *Abstract and Applied Analysis*, vol. 2014, 2014.

[10] A. Kamont, “On the fractional anisotropic wiener field,” *Probability and Mathematical Statistics-PWN*, vol. 16, no. 1, pp. 85–98, 1996.

[11] B. Pesquet-Popescu and J. L. Vehel, “Stochastic fractal models for image processing,” *Signal Processing Magazine, IEEE*, vol. 19, no. 5, pp. 48–62, 2002.

[12] H. Biermé, M. M. Meerschaert, and H.-P. Scheffler, “Operator scaling stable random fields,” *Stochastic Processes and their Applications*, vol. 117, no. 3, pp. 312–332, 2007. [Online]. Available: [http://www.sciencedirect.com/science/article/pii/S0304414906001104](http://www.sciencedirect.com/science/article/pii/S0304414906001104)

[13] A. Bonami and A. Estrade, “Anisotropic analysis of some gaussian models,” *Journal of Fourier Analysis and Applications*, vol. 9, no. 3, pp. 215–236, 2003. [Online]. Available: [http://dx.doi.org/10.1007/s00041-003-0012-2](http://dx.doi.org/10.1007/s00041-003-0012-2)

[14] P. Doukhan, G. Oppenheim, and M. S. Taqqu, *Theory and applications of long-range dependence*. Springer, 2003.

[15] D. Wu and Y. Xiao, “Geometric properties of fractional brownian sheets,” *Journal of Fourier Analysis and Applications*, vol. 13, no. 1, pp. 1–37, 2007. [Online]. Available: [http://dx.doi.org/10.1007/s00041-005-5078-y](http://dx.doi.org/10.1007/s00041-005-5078-y)

[16] A. M. Atto, Y. Berthoumieu, and P. Bolon, “2-d wavelet packet spectrum for texture analysis,” *Image Processing, IEEE Transactions on*, vol. 22, no. 6, pp. 2495–2500, 2013.

[17] S. G. Roux, M. Clausel, B. Vedel, S. Jaffard, and P. Abry, “Self-similar anisotropic texture analysis: The hyperbolic wavelet transform contribution,” *Image Processing, IEEE Transactions on*, vol. 22, no. 11, pp. 4353–4363, 2013.

[18] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang, *Stochastic calculus for fractional Brownian motion and applications*. Springer London, 2008.

[19] M. Moseley, “Diffusion tensor imaging and aging a review,” *NMR in Biomedicine*, vol. 15, no. 7-8, pp. 553–560, 2002. [Online]. Available: [http://dx.doi.org/10.1002/nbm.785](http://dx.doi.org/10.1002/nbm.785)

[20] J. M. Bardet, “Statistical study of the wavelet analysis of fractional brownian motion,” *Information Theory, IEEE Transactions on*, vol. 48, no. 4, pp. 991–999, 2002.

[21] C. J. Morales, “Wavelet-based multifractal spectra estimation: Statistical aspects and applications,” Ph.D. thesis, Boston University, 2002.

[22] J.-M. Bardet, “Testing for the presence of self-similarity of gaussian time series having stationary increments,” *Journal of Time Series Analysis*, vol. 21, no. 5, pp. 497–515, 2000. [Online]. Available: [http://dx.doi.org/10.1111/1467-9892.00195](http://dx.doi.org/10.1111/1467-9892.00195)
[23] B. Mandelbrot and J. Van Ness, “Fractional brownian motions, fractional noises and applications,” SIAM Review, vol. 10, no. 4, pp. 422–437, 1968. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/1010093
[24] Y. Hu, “Heat equations with fractional white noise potentials,” Applied Mathematics and Optimization, vol. 43, no. 3, pp. 221–243, 2001.
[25] A. Brouste, J. Istas, and S. Lambert-Lacroix, “On fractional gaussian random fields simulations,” Journal of Statistical Software, vol. 23, no. 1, pp. 1–23, 2007.
[26] A. K. Basu, Measure theory and probability. PHI Learning Pvt. Ltd., 2004.
[27] M. A. Arcones, “Limit theorems for nonlinear functionals of a stationary gaussian sequence of vectors,” The Annals of Probability, vol. 22, no. 4, pp. 2242–2274, 1994.
[28] R. J. Serfling, Approximation theorems of mathematical statistics. John Wiley & Sons, 1980, vol. 162.
[29] I. Daubechies, Ten lectures on wavelets. SIAM, 1992, vol. 61.
[30] J. C. Goswami and A. K. Chan, Fundamentals of wavelets: theory, algorithms, and applications. John Wiley & Sons, 2011, vol. 233.
[31] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, S. Stoev, and M. S. Taqqu, “Semi-parametric estimation of the long-range dependence parameter: a survey,” Theory and applications of long-range dependence, pp. 557–577, 2003.
[32] A. H. Tewfik and M. Kim, “Correlation structure of the discrete wavelet coefficients of fractional brownian motion,” IEEE transactions on information theory, vol. 38, no. 2, pp. 904–909, 1992.
[33] R. W. Dijkerman and R. R. Mazumdar, “On the correlation structure of the wavelet coefficients of fractional brownian motion,” Information Theory, IEEE Transactions on, vol. 40, no. 5, pp. 1609–1612, 1994.
[34] A. C. Aitken, “On least squares and linear combinations of observations,” Proceedings of the Royal Society of Edinburgh, vol. 55, pp. 42–48, 1935.
[35] R. B. Davies and D. Harte, “Tests for hurst effect,” Biometrika, vol. 74, no. 1, pp. 95–101, 1987.
[36] D. P. Kroese and Z. I. Botev, “Spatial process generation,” arXiv preprint arXiv:1308.0399, 2013.