Anisotropic functional deconvolution with long-memory noise: the case of a multi-parameter fractional Wiener sheet

Rida Benhaddou [*] and Qing Liu

Department of Mathematics, Ohio University, Athens, OH 45701

Abstract

We look into the minimax results for the anisotropic two-dimensional functional deconvolution model with the two-parameter fractional Gaussian noise. We derive the lower bounds for the L^p-risk, $1 \leq p < \infty$, and taking advantage of the Riesz poly-potential, we apply a wavelet-vaguelette expansion to de-correlate the anisotropic fractional Gaussian noise. We construct an adaptive wavelet hard-thresholding estimator that attains asymptotically quasi-optimal convergence rates in a wide range of Besov balls. Such convergence rates depend on a delicate balance between the parameters of the Besov balls, the degree of ill-posedness of the convolution operator and the parameters of the fractional Gaussian noise. A limited simulations study confirms theoretical claims of the paper. The proposed approach is extended to the general r-dimensional case, with $r > 2$, and the corresponding convergence rates do not suffer from the curse of dimensionality.

Keywords and phrases: Anisotropic functional deconvolution, Besov space, anisotropic fractional Brownian sheet, minimax convergence rate

AMS (2000) Subject Classification: 62G05, 62G20, 62G08

[*]E-mail address: Benhaddo@ohio.edu

1

1 Introduction.

Consider the problem of estimating a periodic two-dimensional function, $f(\cdot, \cdot) \in L^2(U)$, based on noisy convolutions that are described by the equations

$$dY(t, x) = q(t, x)dtdx + \varepsilon^{\overline{\alpha}}dB^{\alpha}(t, x), \ q(t, x) = \int_0^1 f(s, x)g(t - s, x)ds. \tag{1}$$

Here, $U = [0, 1]^2$, $g(\cdot, \cdot)$ is the convolution kernel and it is assumed to be known, $B^{\alpha}(t, x)$ is an anisotropic two-dimensional fractional Brownian sheet (fBs), $\overline{\alpha} = (\alpha_1 + \alpha_2)/2$, and $\alpha_i = 2 - 2H_i \in (0, 1]$, $i = 1, 2$, are the parameters of the long-memory in the direction of t and x, respectively. A two-dimensional fractional Brownian sheet is defined by the formula

$$B^{\alpha}(t, x) = \frac{1}{C_H} \int_0^t \int_0^x Q(s, y)dW(s, y), \tag{2}$$

where $W = \{W(t, x), (t, x) \in U\}$ is a two-dimensional standard Brownian sheet, C_H is some explicit constant and

$$Q(s, y) = [(t - s)_+]^{H_1 - 1/2} [(x - y)_+]^{H_2 - 1/2}, \tag{3}$$

with $Y_+ = \max(0, Y)$. In addition, the two-parameter fBs is characterized by a covariance function of the form

$$\text{Cov}\left(B^{\alpha}(t_1, s_1), B^{\alpha}(t_2, s_2)\right) = c\left[|t_1|^{2H_1} + |t_2|^{2H_1} - |t_1 - t_2|^{2H_1}\right]\left[|s_1|^{2H_2} + |s_2|^{2H_2} - |s_1 - s_2|^{2H_2}\right], \tag{4}$$

for some $c \in R$, $(t_i, s_i) \in U$, $i = 1, 2$, where $H_i = 1 - \frac{\alpha_i}{2}$ are the Hurst parameters in the direction of t and s, respectively (see Kamont (1996)).

The discrete, and the more realistic, version of model (1) is given by

$$Y(t_i, x_l) = q(t_i, x_l) + \sigma\xi_{il}, \ (i, l) = 1, 2, \cdots, N, \tag{5}$$

where $t_i = \frac{i}{N}$, $x_l = \frac{l}{N}$, σ is a positive variance constant, and ξ_{il} are zero-mean second-order stationary Gaussian random variables satisfying the auto-covariance structure

$$\gamma(h_1, 0) \asymp h_1^{-\alpha_1}, \quad \gamma(0, h_2) \asymp h_2^{-\alpha_2} \ and \ \gamma(h_1, h_2) \asymp h_1^{-\alpha_1} h_2^{-\alpha_2}. \tag{6}$$

Assumption A.1. The error structure $\{\xi_{il} : i, l \in Z\}$ in model (5) is a zero-mean second-order stationary process satisfying

$$X_N \xrightarrow{D} B^{\alpha}, \tag{7}$$

where

$$X_N(t, x) = \frac{1}{N^{2-\alpha_1/2-\alpha_2/2}} \sum_{i=1}^{\lceil Nt \rceil} \sum_{l=1}^{\lceil Nx \rceil} \xi_{il}, \ (t, x) \in U^2, \tag{8}$$

and B^{α} is a two-parameter fractional Brownian sheet.

Assumption **A.1** is valid under auto-covariance structure (6) and it appeared in Adu and Richardson (2018). It guarantees that, with the calibration $\varepsilon^{\alpha_1+\alpha_2} \asymp n^{\alpha_1/2+\alpha_2/2}$, models (1) and (5) are asymptotically equivalent. Therefore, for sufficiently large sample size $n = N^2$, model (1) can be used to approximate model (5).

Deconvolution model has been the subject of a great deal of papers since late 1980s, but the most significant contribution was that of Donoho (1995) who was the first to devise a wavelet solution to the problem. Other attempts include, Abramovich and Silverman (1998), Walter and Shen (1999), Johnstone et al. (2004), Donoho and Raimondo (2004), among others. In the case of functional deconvolution model with $f(t, x) \equiv f(t)$, Pensky and sapatinas (2009, 2010, 2011) pioneered into the formulation and further development of the problem.

Functional deconvolution problem of type (1) with $\alpha_1 = \alpha_2 = 1$ corresponds to the white noise case and it was investigated under the L^2-risk in Benhaddou et al. (2013), and under L^p-risk, $1 \leq p < \infty$, in Benhaddou (2017), where they constructed an adaptive hard-thresholding wavelet estimator, and showed that it is asymptotically quasi-optimal within a logarithmic factor of ε over a wide range of Besov balls of mixed regularity. This model is motivated by experiments in which one needs to recover a two-dimensional function using observations of its convolutions along profiles $x = x_l$. This situation occurs, for example, in seismic inversions (see

Robinson (1999)). In these articles, it is assumed that the error terms are white noise processes or i.i.d noise. However, empirical evidence has shown that, even at large lags, the correlation structure in the errors takes the power-like form (6). This phenomenon is referred to as long-memory (LM) or long-range dependence (LRD). The presence of long memory in oceanic seismic data was pointed out in Wood et al. (2014) for instance, and it may be due for example to sea floor temperature variations.

Long-memory has been investigated quite considerably in the standard deconvolution model. One can list; Wang (1996, 1997), Wishart (2013), Benhaddou et al. (2014) and Kulik et al. (2015). In a few other relevant contexts, LM was also investigated in density deconvolution in Comte et al. (2008), and in the Laplace deconvolution in Benhaddou (2018) where the unknown response function $f(\cdot)$ is non-periodic and defined on the entire positive real half-line.

The objective of the paper is to extend the work of Benhaddou et al. (2013) to the case when the noise is an anisotropic multi-parameter fractional Brownian sheet. Following a standard procedure, we derive minimax lower bounds for the L^p-risk, with $1 \leq p < \infty$, when $f(t,x)$ belongs to an anisotropic Besov ball and the blurring function $g(t,x)$ is regular smooth. Taking advantage of the Riesz poly-potential operator, and inspired by the work of Wang (1996, 1997), we apply the wavelet-vaguelette expansion (WVD) to de-correlate the anisotropic fBs in two directions. In addition, we take advantage of the flexibility of the Meyer wavelet basis in the Fourier domain to construct a wavelet hard-thresholding estimator that is adaptive and asymptotically quasi-optimal within a logarithmic factor of ε over a large array of Besov balls. Furthermore, the proposed estimator attains convergence rates that depend on a delicate balance between the parameters of the Besov ball, the degree of ill-posedness of the convolution, and the parameters α_1 and α_2 associated with the fBs. Similar to the white noise case studied in Benhaddou et al. (2013), the proposed approach is easily extended to recovering an r-dimensional function, $r > 2$, and the corresponding convergence rates turn out to be dimension-free. Finally, a simulation study is carried out to further confirm the results of our theoretical findings.

The rest of the paper is organized as follows. Section 2 introduces some notation as well as the estimation algorithm. Section 3 describes the derivation of the lower bounds for the L^p-risk

of estimators of f as well as the upper bounds and establishes the asymptotic optimality of the estimator. Section 4 presents a limited simulations study to complement the theoretical results from Section 3. Section 5 extends the results in Sections 2 and 3 to the general r-dimensional case. Finally, Section 6 contains the proofs of the theoretical results.

2   Estimation Algorithm.

In what follows, denote the complex conjugate of a by $\bar{a}$. Let $\tilde{Y}(m_1, m_2)$, $\tilde{B}^\alpha(m_1, m_2)$, $\tilde{g}(m_1, m_2)$ and $\tilde{f}(m_1, m_2)$ be the two-dimensional Fourier coefficients of functions $dY(t,x)$, $dB^\alpha(t,x)$, $g(t,x)$, and $f(t,x)$, respectively.

Consider a bandlimited wavelet basis $\psi_{j,k}(y)$ (e.g., Meyer-type) on $[0,1]$, and form the product wavelet basis

$$\Psi_\omega(t,x) = \psi_{j_1,k_1}(t)\psi_{j_2,k_2}(x), \ (t,x) \in U, \tag{9}$$

where $\omega \in \Omega$, and

$$\Omega = \left\{ \omega = (j_1, k_1; j_2, k_2) : j_i = m_0 - 1, \cdots, \infty; k_i = 0, 1, \cdots, 2^{j_i} - 1, i = 1, 2 \right\}. \tag{10}$$

Let m_0 be the lowest resolution level for the Meyer basis and denote the scaling function for the wavelet by ψ_{m_0-1,k_i}, $i = 1, 2$. Since the functions (9) form an orthonormal basis of the $L^2(U)$-space, the function $f(t,x)$ can be expanded over these bases with coefficients β_ω into a wavelet series

$$f(t,x) = \sum_{\omega \in \Omega} \beta_\omega \Psi_\omega(t,x), \tag{11}$$

where $\beta_\omega = \int_0^1 \int_0^1 f(t,x)\Psi_\omega(t,x)dtdx$. Denote the two-dimensional Fourier transform of $\Psi_\omega(t,x)$ by $\tilde{\Psi}_\omega(m_1, m_2) = \tilde{\psi}_{j_1,k_1}(m_1)\tilde{\psi}_{j_2,k_2}(m_2)$. It is well-known (see, e.g, Johnstone et al (2004), section 3.1) that under the Fourier domain and for any $j_i \geq m_0$, $i = 1, 2$, one has

$$W_{j_i} = \left\{ m_i : \tilde{\psi}_{j_i,k_i}(m_i) \neq 0 \right\} \subseteq 2\pi/3 \left[-2^{j_i+2}, -2^{j_i} \right] \cup \left[2^{j_i}, 2^{j_i+2} \right]. \tag{12}$$

Apply the two-dimensional Fourier transform to equation (1) to obtain

$$\tilde{Y}(m_1, m_2) = \overline{\tilde{f}(m_1, m_2)} \tilde{g}(m_1, m_2) + \varepsilon^{\overline{\alpha}} \tilde{B}^{\alpha}(m_1, m_2). \tag{13}$$

Now, applying Plancherel formula in both directions, the wavelet coefficients β_{ω} in (11) can be rewritten as

$$\beta_{\omega} = \sum_{m_1 \in W_{j_1}} \sum_{m_2 \in W_{j_2}} \overline{\tilde{\psi}_{j_1, k_1}(m_1)} \tilde{\psi}_{j_2, k_2}(m_2) \tilde{f}(m_1, m_2). \tag{14}$$

Therefore, an unbiased estimator for β_{ω} is given by

$$\tilde{\beta}_{\omega} = \sum_{m_1 \in W_{j_1}} \sum_{m_2 \in W_{j_2}} \overline{\tilde{\psi}_{j_1, k_1}(m_1)} \tilde{\psi}_{j_2, k_2}(m_2) \frac{\tilde{Y}(m_1, m_2)}{\tilde{g}(m_1, m_2)}. \tag{15}$$

Finally, consider the hard thresholding estimator for $f(t, x)$

$$\widehat{f}_{\varepsilon}(t, x) = \sum_{\omega \in \Omega(J_1, J_2)} \tilde{\beta}_{\omega} \mathbb{I}\left(|\tilde{\beta}_{\omega}| > \lambda_{j;\varepsilon}^{\alpha}\right) \Psi_{\omega}(t, x), \tag{16}$$

where

$$\Omega(J_1, J_2) = \left\{\omega = (j_1, k_1; j_2, k_2) : j_i = m_0 - 1, \cdots, J_i; k_i = 0, 1, \cdots, 2^{j_i} - 1, i = 1, 2\right\}, \tag{17}$$

and the quantities J_1, J_2 and $\lambda_{j;\varepsilon}^{\alpha}$ are to be specified.

Assumption A.2. In the Fourier domain, the convolution kernel $g(t, x)$, for some positive constants ν, and C_1 and C_2, independent of m_1 and m_2, is such that

$$C_1|m_1|^{-2\nu} \leq |\tilde{g}(m_1, m_2)|^2 \leq C_2|m_1|^{-2\nu}. \tag{18}$$

Let us now evaluate the variance of (15).

Lemma 1 *Let $\tilde{\beta}_{\omega}$ be defined in (15). Then, under the condition (18), for $1 \leq p < \infty$, one has*

$$\mathbb{E}\left|\tilde{\beta}_{\omega} - \beta_{\omega}\right|^{2p} \leq K_1 \varepsilon^{2p\overline{\alpha}} 2^{j_1 p(2\nu + \alpha_1 - 1) + j_2 p(\alpha_2 - 1)}. \tag{19}$$

6

In addition,

$$\mathbb{E}\left|\tilde{\beta}_\omega - \beta_\omega\right|^p \le K_2 \varepsilon^{p\overline{\alpha}} 2^{j_1 \frac{p}{2}(2\nu+\alpha_1-1)+j_2\frac{p}{2}(\alpha_2-1)}. \tag{20}$$

where K_1 and K_2 are some positive constants independent of ε.

According to Lemma 1, choose the thresholds $\lambda_{j;\varepsilon}^\alpha$ of the form

$$\lambda_{j;\varepsilon}^\alpha = \gamma\varepsilon^{\overline{\alpha}}\sqrt{|\ln(\varepsilon)|}2^{\frac{j_1}{2}(2\nu+\alpha_1-1)}2^{\frac{j_2}{2}(\alpha_2-1)}, \tag{21}$$

where γ is some positive constant independent of ε. Furthermore, the finest resolution levels J_1 and J_2 are such that

$$2^{J_2} = \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{-\frac{1}{\alpha_2}}, \quad 2^{J_1} = \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{-\frac{1}{2\nu+\alpha_1}}. \tag{22}$$

3 Convergence rates and asymptotic optimality.

Denote

$$s_i^* = s_i + 1/2 - 1/\pi, \tag{23}$$

$$s_i' = s_i + 1/2 - 1/\pi', \tag{24}$$

$$s_i'' = s_i + 1/p - 1/p', \tag{25}$$

where $\pi' = \min\{2,\pi\}$ and $p' = \min\{p,\pi\}$.

Assumption A.3. The function $f(t,x)$ belongs to an anisotropic two-dimensional Besov space. In particular, its wavelet coefficients β_{j_1,k_1,j_2,k_2} satisfy

$$B_{\pi,q}^{s_1,s_2}(A) = \left\{f \in L^2(U) : \left(\sum_{j_1,j_2} 2^{(j_1 s_1^* + j_2 s_2^*)q}\left(\sum_{k_1,k_2}|\beta_{j_1,k_1,j_2,k_2}|^\pi\right)^{q/\pi}\right)^{1/q} \le A\right\} \tag{26}$$

To construct minimax lower bounds for the L^p-risk, we define the L^p-risk over the set V as

$$R(V) = \inf_{\tilde{f}} \sup_{f \in V} \mathbb{E}\|\tilde{f}_\varepsilon - f\|_p^p, \tag{27}$$

where $\|g\|_p$ is the L^p-norm of a function g and the infimum is taken over all possible estimators $\tilde{f}$ of f.

Theorem 1 *Let* $\min\{s_1, s_2\} \geq \max\{\frac{1}{\pi}, \frac{1}{2}\}$ *with* $1 \leq \pi, q \leq \infty$, *and* $A > 0$. *Then, under conditions (18) and (26), for* $1 \leq p < \infty$, *as* $\varepsilon \to 0$,

$$
R(B^{s_1,s_2}_{\pi,q}(A)) \geq CA^p \begin{cases} \left[\frac{\varepsilon^{2\bar{\alpha}}}{A^2}\right]^{\frac{ps_2}{2s_2+\alpha_2}}, & if \ \ s_1 > \frac{s_2}{\alpha_2}(2\nu + \alpha_1), \ \& \ \frac{s_2}{\alpha_2} \geq \frac{p}{2}(\frac{1}{p'} - \frac{1}{p}) \\[2mm] \left[\frac{\varepsilon^{2\bar{\alpha}}}{A^2}\right]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}, & if \ \ \frac{p}{2}(2\nu + \alpha_1)(\frac{1}{\pi} - \frac{1}{p}) \leq s_1 \leq \frac{s_2}{\alpha_2}(2\nu + \alpha_1), \\[2mm] \left[\frac{\varepsilon^{2\bar{\alpha}}}{A^2}\right]^{\frac{p(s_1+\frac{1}{p}-\frac{1}{\pi})}{2s_1^*+2\nu+\alpha_1-1}}, & if \ s_1 < (2\nu + \alpha_1)\frac{p}{2}(\frac{1}{\pi} - \frac{1}{p}), \ \& \ \frac{s_2}{\alpha_2} \geq \frac{p}{2}(\frac{1}{p'} - \frac{1}{p}). \end{cases}
$$
(28)

Proof of Theorem 1. In order to prove the theorem, first we establish the equivalence between a process that involves the fractional Wiener sheet (eq. (1)) and another that involves the standard Wiener sheet by applying an appropriate integral operator, taking advantage of relation (2). Then, we consider two cases, the case when $f(t,x)$ is dense in both dimensions (the dense-dense case) and the case when $f(t,x)$ is dense in x but sparse in t. Lemma $A.1$ of Bunea et al. (2007) is then applied to find such lower bounds using conditions (18) and (26), combined with some useful properties of Meyer wavelet basis. To complete the proof, we choose the highest of the lower bounds. $\square$

The derivation of upper bounds of the L^p-risk relies on the following two lemmas.

Lemma 2 *Under condition (26), one has*

$$
\sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} |\beta_{j_1,k_1,j_2,k_2}|^p \leq A^p 2^{-p[(j_1s_1+j_2s_2)+(\frac{1}{2}-\frac{1}{p'})(j_1+j_2)]}, \ \forall j_1, j_2 \geq 0.
$$
(29)

Lemma 3 *Let* $\tilde{\beta}_\omega$ *and* $\lambda^\alpha_{j;\varepsilon}$ *be defined by (15) and (21), respectively. Define, for some positive constant* η, *the set*

$$
\Theta_{\omega,\eta} = \left\{\Theta : |\tilde{\beta}_\omega - \beta_\omega| > \eta\lambda^\alpha_{j;\varepsilon}\right\}.
$$
(30)

Then, under condition (18) and as $\varepsilon \to 0$, one has

$$\Pr\left(\Theta_{\omega,\eta}\right) = O\left(\frac{\left[\varepsilon^{2\overline{\alpha}}\right]^{\tau}}{|\ln(\varepsilon)|^{\frac{1}{2}}}\right), \tag{31}$$

where $\tau = \frac{\eta^2 \gamma^2 C_1^2}{8\overline{\alpha}} \left(\frac{3}{8\pi}\right)^{2\nu+\alpha_1-1+\alpha_2-1}$, C_1 and γ appear in (18) and (21), respectively, and $\overline{\alpha} = \frac{\alpha_1+\alpha_2}{2}$.

Theorem 2 *Let* $\widehat{f}(.,.)$ *be the wavelet estimator in (16), with* J_1 *and* J_2 *given by (22). Let* $\min\{s_1, s_2\} \geq \max\{\frac{1}{\pi}, \frac{1}{2}\}$ *with* $1 \leq \pi, q \leq \infty$, *and let conditions (18) and (26) hold. If* γ *in (21) is large enough, then, for* $1 \leq p \leq \infty$ *and as* $\varepsilon \to 0$, *one has*

$$R(B_{\pi,q}^{s_1,s_2}(A)) \leq CA^p \begin{cases} \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{ps_2}{2s_2+\alpha_2}} [|\ln(\varepsilon)|]^{\xi_1+\frac{ps_2}{2s_2+\alpha_2}}, & \text{if } s_1 \geq \frac{s_2}{\alpha_2}(2\nu+\alpha_1) \text{ \& } \frac{s_2}{\alpha_2} \geq \frac{p}{2}(\frac{1}{p'}-\frac{1}{p}), \\[2ex] \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}} [|\ln(\varepsilon)|]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}, & \text{if } \frac{p}{2}(2\nu+\alpha_1)(\frac{1}{\pi}-\frac{1}{p}) < s_1 < \frac{s_2}{\alpha_2}(2\nu+\alpha_1), \\[2ex] \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{p(s_1+\frac{1}{p}-\frac{1}{\pi})}{2s_1^*+2\nu+\alpha_1-1}} [|\ln(\varepsilon)|]^{\xi_2+\frac{p(s_1+\frac{1}{p}-\frac{1}{\pi})}{2s_1^*+2\nu+\alpha_1-1}}, & \text{if } s_1 \leq (2\nu+\alpha_1)\frac{p}{2}(\frac{1}{\pi}-\frac{1}{p}) \text{ \& } \frac{s_2}{\alpha_2} \geq \frac{p}{2}(\frac{1}{p'}-\frac{1}{p}), \end{cases} \tag{32}$$

where ξ_1 *and* ξ_2 *are defined as*

$$\xi_1 = \mathbb{I}\left(s_1 = \frac{s_2}{\alpha_2}(2\nu+\alpha_1)\right), \tag{33}$$

$$\xi_2 = \mathbb{I}\left(s_1 = (2\nu+\alpha_1)\frac{p}{2}\left(\frac{1}{\pi}-\frac{1}{p}\right)\right) + \mathbb{I}\left(\frac{s_2}{\alpha_2} = \frac{p}{2}\left(\frac{1}{p'}-\frac{1}{p}\right)\right). \tag{34}$$

The proof of Theorem 2. The proof is very similar to that of Theorem 2 in Benhaddou. (2017).

□

Remark 1 *Notice that the finest resolution levels* J_1 *and* J_2 *in (22) and the thresholds* $\lambda_{j;\varepsilon}^{\alpha}$ *in (21), are independent of the unknown parameters of the Besov ball (26) and therefore estimator (16) is adaptive with respect to those parameters.*

Remark 2 *Theorems 1 and 2 imply that, for the* L^p-*risk, the estimator in (16) is asymptotically quasi-optimal within a logarithmic factor of* ε, *over a wide range of anisotropic Besov balls* $B_{\pi,q}^{s_1,s_2}(A)$.

9

Remark 3 *The convergence rates depend on a delicate balance between the parameters of the Besov ball, smoothness of the convolution kernel ν and the parameters of the anisotropic fractional Brownian sheet, α_1 and α_2. These rates deteriorate as the values of α_1 and α_2 get closer and closer to zero.*

Remark 4 *To de-correlate the two-dimensional fractional Gaussian noise a wavelet-vaguelette expansion based on Meyer wavelet bases has been applied to the fractional Brownian sheet. The validity of a wavelet-based series representation for the fractional Brownian motion has been established in Wang (1997), Meyer, Sellan and Taqqu (1999) and Ayache and Taqqu (2002). In particular, Ayache and Taqqu (2002) established the optimality of wavelet-based series representation of the fractional Brownian sheets.*

Remark 5 *For $\alpha_1 = \alpha_2 = 1$, the rates of convergence match exactly those in Benhaddou (2017), and, with $p = 2$, those in Benhaddou et al. (2013) in their bivariate white noise case.*

Remark 6 *If we hold the second dimension x fixed, our rates are comparable to those in Wang (1997) and Wishart (2013) in their standard deconvolution with the one-parameter fractional Gaussian noise case.*

4   Simulation Study

In order to investigate the effect of the long-memory on the performance of our estimator, we carried out a limited simulation study. We implemented an estimation algorithm for the model (5) using a modification of WaveD method of Raimondo and Stewart (2007). We evaluated mean integrated square error (MISE) $\mathbb{E}\|\hat{f} - f\|^2$ of the functional deconvolution estimator.

1. We generated the data using equation (5) with convolution kernel $g(t, x) = 0.5\exp(-|t|(1 + (x - 0.5)^2))$, and $f(t, x) = f(t)f(x)$. In particular, we chose $f(t)$ to be a LIDAR or Doppler signal over the grid $t_i = \frac{i}{N}$ with $i = 1, 2, ..., N$, and we chose $f(x) = \exp(-|x - 0.5|x^3)$ over the grid $x_i = \frac{i}{N}$ with $i = 1, 2, ..., N$, and $N = 2^{10} = 1024$. The LIDAR and Doppler signals were generated from the `waved` package. All test functions were scaled to have a unit norm.

Bear in mind that, though $f(t, x)$ is a product of two univariate functions, the method does not "recognize" this and, therefore, cannot take advantage of this information. Also, notice that with our choice of $g(t, x)$, the degree of ill-posedness (DIP) of the convolution is DIP=0.5.

2. We simulated the LM error in the direction of both t and x using the `fracdiff` package of R available from CRAN. In particular, the `fracdiff.sim` command was used to simulate two one-dimensional fractionally differenced $\mathrm{ARIMA}(0, d, 0)$ (fARIMA) sequences which behave similar to a fractional Gaussian noise, and then multiply them together to generate a two-dimensional error structure. For the LM parameters $\alpha = (\alpha_1, \alpha_2)$, we used different combinations of the levels 0.8, 0.6, 0.4 and 0.2. To create a dependence structure that is anisotropic, we only used combinations such that $\alpha_1 \neq \alpha_2$. Note that the fractional differencing parameter d is obtained from α by the relation $d = \frac{1-\alpha}{2}$.

3. The choice of σ in (5) was determined by the blurred signal-to-noise ratio (SNR), where $\mathrm{SNR} = 10 \log_{10} \left(\frac{\|q(t,x)\|^2}{\sigma^2} \right)$. We considered three choices, SNR=10dB (high noise), 20dB (medium noise) and 30dB (low noise).

4. To compute the estimator, we used the function `FWaveD` and `IWaveD` from the `waved` package. First, we applied the Meyer wavelet transform to the data in t-direction by `FWaveD`, with no thresholding by setting `thr` to be a zero vector. Then, we applied the second thresholded Meyer wavelet transform in x-direction by `FWaveD`, however, the level dependent threshold were modified according to (21). For the tuning parameter γ, we tried different values ranging from $\gamma = \sqrt{2}$ to $\gamma = \sqrt{6}$, but the performance of the estimator reached its best at the latter value. Therefore, our default choice of γ was $\sqrt{6}$. The finest resolution level J_1 was estimated using the default method in the `waved` package, together with the choice $J_2 = 5$.

Table 1 reports the averages of the errors over 2000 simulation runs using a sample size of $n = N^2 = (2^{10})^2$. Figure (1) illustrates the function $f(t.x)$ and its estimator $\widehat{f}(t, x)$ and shows a relatively good precision that deteriorates as the anisotropic pair of LM parameters (α_1, α_2)

11

decreases from 0.8 to 0.2. Notice here that according to Table 1, the influence of LM is a little more pronounced in the direction of the dimension x than it is in the direction of t. Table 1 complements Figure (1) and confirms our theoretical results in the sense that as the level of LM increases (α gets smaller), the mean integrated squared error increases (the performance deteriorates).

Figure 1: The function $f(t,x)$ (top) and its estimate $\widehat{f}(t,x)$ (bottom) at different combinations of (α_1, α_2).
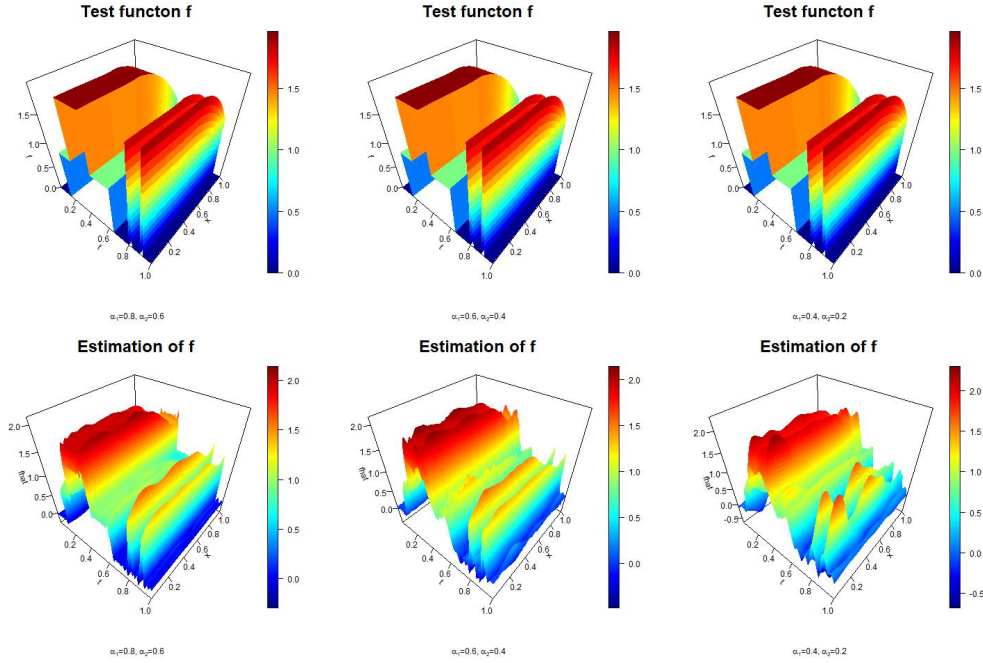

Table 1:

Finest resolution level J_1 are listed in the parentheses.

DIP=0.5		(α_1, α_2)					
signal	σ	(0.8,0.6)	(0.8,0.4)	(0.8,0.2)	(0.6,0.4)	(0.6,0.2)	(0.4,0.2)
Lidar (10dB)	0.20	0.0779 (3)	0.0948 (3)	0.1296 (3)	0.1134 (3)	0.2097 (3)	0.3232 (3)
Lidar (20dB)	0.06	0.0434 (4)	0.0498 (4)	0.0641 (4)	0.0572 (4)	0.0796 (4)	0.0948 (4)
Lidar (30dB)	0.02	0.0132 (5)	0.0183 (5)	0.0292 (5)	0.0219 (5)	0.0358 (5)	0.0444 (5)
Doppler (10dB)	0.15	0.2613 (3)	0.2724 (3)	0.3160 (3)	0.2935 (3)	0.3828 (3)	0.4823 (3)
Doppler (20dB)	0.05	0.1254 (4)	0.1298 (4)	0.1437 (4)	0.1371 (4)	0.1584 (4)	0.1867 (4)
Doppler (30dB)	0.015	0.0516 (5)	0.0530 (5)	0.0568 (5)	0.0544 (5)	0.0610 (5)	0.0667 (5)

DIP=0.5 signal	σ	(α_1,α_2) (0.6,0.8)	(0.4,0.8)	(0.2,0.8)	(0.4,0.6)	(0.2,0.6)	(0.2,0.4)
Lidar (10dB)	0.20	0.0760 (3)	0.0851 (3)	0.0956 (3)	0.1038 (3)	0.1293 (3)	0.2251 (3)
Lidar (20dB)	0.06	0.0423 (4)	0.0453 (4)	0.0486 (4)	0.0536 (4)	0.0612 (4)	0.0791 (4)
Lidar (30dB)	0.02	0.0120 (5)	0.0132 (5)	0.0149 (5)	0.0180 (5)	0.0210 (5)	0.0331 (5)
Doppler (10dB)	0.15	0.2601 (3)	0.2640 (3)	0.2749 (3)	0.2836 (3)	0.3090 (3)	0.3922 (3)
Doppler (20dB)	0.05	0.1247 (4)	0.1264 (4)	0.1292 (4)	0.1331 (4)	0.1376 (4)	0.1578 (4)
Doppler (30dB)	0.015	0.0513 (5)	0.0510 (5)	0.0522 (5)	0.0527 (5)	0.0541 (5)	0.0591 (5)

5 The r-dimensional case: estimation, convergence rates and asymptotic optimality.

Consider the r-dimensional case of model (1)

$$dY(t,\mathbf{x}) = q(t,\mathbf{x})dtd\mathbf{x} + \varepsilon^{\overline{\alpha}}dB^{\alpha}(t,\mathbf{x}), \ q(t,\mathbf{x}) = \int_0^1 f(s,\mathbf{x})g(t-s,\mathbf{x})ds. \tag{35}$$

Here, $U = [0,1]^r$, $\mathbf{x} = (x_1, x_2, \cdots, x_{r-1})$, $g(t,\mathbf{x})$ is the convolution kernel, $B^{\alpha}(t,\mathbf{x})$ is an anisotropic r-dimensional fBs, $\overline{\alpha} = \sum_{k=1}^r \alpha_k/r$, and $\alpha_k = 2 - 2H_k \in (0,1]$, $k = 1, 2, \cdots, r$, is the parameter of the long-memory in the direction of the kth variable. An r-parameter fBs is defined as a zero-mean Gaussian process $\{B^{\alpha}(\mathbf{t}) : \mathbf{t} \in [0,1]^r\}$ whose covariance function is of the form

$$\text{Cov}\ (B^{\alpha}(\mathbf{t_1}), B^{\alpha}(\mathbf{t_2})) = c\Pi_{k=1}^r \left[|t_{k1}|^{2H_k} + |t_{k2}|^{2H_k} - |t_{k1} - t_{k2}|^{2H_k} \right], \tag{36}$$

for some $c \in \mathbb{R}$, $(t_{k1}, t_{k2}) \in [0,1]^2$, $k = 1, 2, \cdots, r$, where $H_k = 1 - \alpha_k/2$ is the Hurst parameter in the direction of the kth dimension.

Consider a bandlimited wavelet basis $\psi_{j,k}(y)$ (e.g., Meyer-type) on $[0,1]$, and form the product wavelet basis

$$\Psi_{\omega_r}(t,\mathbf{x}) = \psi_{j_1,k_1}(t)\Pi_{i=2}^r\psi_{j_i,k_i}(x_{i-1}), \ (t,\mathbf{x}) \in U, \tag{37}$$

where $\omega_r \in \Omega_r$, and

$$\Omega_r = \left\{ \omega_r = (j_1, k_1; j_2, k_2; \cdots; j_r, k_r) : j_i = m_0 - 1, \cdots, \infty; k_i = 0, 1, \cdots, 2^{j_i} - 1, i = 1, 2, \cdots, r \right\}. \tag{38}$$

13

Then, the function $f(t, \mathbf{x})$ can be expanded over these bases with coefficients β_{ω_r} into a wavelet series

$$f(t, \mathbf{x}) = \sum_{\omega_r \in \Omega_r} \beta_{\omega_r} \Psi_{\omega_r}(t, \mathbf{x}), \tag{39}$$

where $\beta_{\omega_r} = \int_U f(t, \mathbf{x}) \Psi_{\omega_r}(t, \mathbf{x}) dt d\mathbf{x}$. Apply the r-dimensional Fourier transform to equation (35) to obtain

$$\tilde{Y}(\mathbf{m_r}) = \tilde{f}(\mathbf{m_r}) \tilde{g}(\mathbf{m_r}) + \varepsilon^{\overline{\alpha}} \tilde{B}^{\alpha}(\mathbf{m_r}), \tag{40}$$

where $\mathbf{m_r} = (m_1, m_2, \cdots, m_r)$. Then, applying Plancherel formula in all directions, the wavelet coefficients β_{ω_r} in (39) can be rewritten as

$$\beta_{\omega_r} = \sum_{m_1 \in W_{j_1}} \sum_{m_2 \in W_{j_2}} \cdots \sum_{m_r \in W_{j_r}} \tilde{\Psi}_{\omega_r}(\mathbf{m_r}) \tilde{f}(\mathbf{m_r}). \tag{41}$$

Therefore, an unbiased estimator for β_{ω_r} is given by

$$\tilde{\beta}_{\omega_r} = \sum_{m_1 \in W_{j_1}} \sum_{m_2 \in W_{j_2}} \cdots \sum_{m_r \in W_{j_r}} \tilde{\Psi}_{\omega_r}(\mathbf{m_r}) \frac{\tilde{Y}(\mathbf{m_r})}{\tilde{g}(\mathbf{m_r})}. \tag{42}$$

Finally, consider the hard thresholding estimator for $f(t, \mathbf{x})$

$$\widehat{f}_{\varepsilon}(t, \mathbf{x}) = \sum_{\omega_r \in \Omega_r(\mathbf{J})} \tilde{\beta}_{\omega_r} \mathbb{I}\left(|\tilde{\beta}_{\omega_r}| > \lambda_{\mathbf{j};\varepsilon}^{\alpha}\right) \Psi_{\omega_r}(t, \mathbf{x}), \tag{43}$$

where

$$\Omega_r(\mathbf{J}) = \left\{ \omega_r = (j_1, k_1; j_2, k_2; \cdots; j_r, k_r) : j_i = m_0 - 1, \cdots, J_i; k_i = 0, 1, \cdots, 2^{j_i} - 1, i = 1, 2, \cdots, r \right\}, \tag{44}$$

with $\mathbf{j} = (j_1, j_2, \cdots, j_r)^T$, and the quantities J_1, J_2, $\cdots$, J_r and $\lambda_{\mathbf{j};\varepsilon}^{\alpha}$ are to be determined.

Assumption A.4. In the Fourier domain, the convolution kernel $g(t, \mathbf{x})$, for some positive constants ν, and C_{r1} and C_{r2}, independent of $\mathbf{m_r} = (m_1, m_2, \cdots, m_r)$, is such that

$$C_{r1} |m_1|^{-2\nu} \le |\tilde{g}(\mathbf{m_r})|^2 \le C_{r2} |m_1|^{-2\nu}. \tag{45}$$

14

Lemma 4 *Let $\tilde{\beta}_{\omega_r}$ be defined in (42). Then, under the condition (45), for $1 \leq p < \infty$, one has*

$$\mathbb{E}\left|\tilde{\beta}_{\omega_r} - \beta_{\omega_r}\right|^{2p} \leq K_3 \varepsilon^{p 2\overline{\alpha}} 2^{j_1 p(2\nu + \alpha_1 - 1) + j_2 p(\alpha_2 - 1) + \cdots + j_r p(\alpha_r - 1)}. \tag{46}$$

In addition,

$$\mathbb{E}\left|\tilde{\beta}_{\omega_r} - \beta_{\omega_r}\right|^{p} \leq K_4 \varepsilon^{p \overline{\alpha}} 2^{j_1 \frac{p}{2}(2\nu + \alpha_1 - 1) + j_2 \frac{p}{2}(\alpha_2 - 1) + \cdots + j_r \frac{p}{2}(\alpha_r - 1)}. \tag{47}$$

where K_3 and K_4 are some positive constants independent of ε.

Consequently, choose the thresholds $\lambda_{\mathbf{j};\varepsilon}^{\alpha}$ of the from

$$\lambda_{\mathbf{j};\varepsilon}^{\alpha} = \rho \varepsilon^{\overline{\alpha}} \sqrt{|\ln(\varepsilon)|} 2^{j_1 \nu} \Pi_{i=1}^{r} 2^{\frac{j_i}{2}(\alpha_i - 1)}, \tag{48}$$

where ρ is some positive constant that is independent of ε, $\overline{\alpha} = \frac{1}{r}\sum_{k=1}^{r} \alpha_k$ and J_1 and J_i, $i = 2, 3, \cdots, r$, are such that

$$2^{J_1} = \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{-\frac{1}{2\nu + \alpha_1}}, \quad 2^{J_i} = \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{-\frac{1}{\alpha_i}}. \tag{49}$$

Lemma 5 *Let $\tilde{\beta}_{\omega_r}$ and $\lambda_{\mathbf{j};\varepsilon}^{\alpha}$ be defined by (42) and (48), respectively. Define, for some positive constant η, the set*

$$\Theta_{\omega_r,\eta}^{r} = \left\{\Theta : |\tilde{\beta}_{\omega_r} - \beta_{\omega_r}| > \eta \lambda_{\mathbf{j};\varepsilon}^{\alpha}\right\}. \tag{50}$$

Then, under condition (45), as $\varepsilon \to 0$, one has

$$\Pr\left(\Theta_{\omega_r,\eta}^{r}\right) = O\left(\frac{\left[\varepsilon^{2\overline{\alpha}}\right]^{\tau_r}}{|\ln(\varepsilon)|^{\frac{1}{2}}}\right), \tag{51}$$

where $\tau_r = \frac{\eta^2 \rho^2 C_{r1}^2}{8\overline{\alpha}}\left(\frac{3}{8\pi}\right)^{2\nu + \sum_{i=1}^{r}(\alpha_i - 1)}$, and C_{r1} and ρ appear in (45) and (48), respectively.

15

Assumption A.5. The r-dimensional function $f(t, \mathbf{x})$ belongs to an anisotropic Besov space. In particular, its wavelet coefficients $\beta_{j_1,k_1;j_2,k_2;\cdots;j_r,k_r}$ satisfy

$$B_{\pi,q}^{s_1,s_2,\cdots,s_r}(A) = \left\{ f \in L^2(U) : \left(\sum_{j_1,j_2,\cdots,j_r} 2^{\mathbf{j}^T \mathbf{s}^* q} \left(\sum_{k_1,k_2,\cdots,k_r} |\beta_{j_1,k_1;j_2,k_2;\cdots;j_r,k_r}|^\pi \right)^{q/\pi} \right)^{1/q} \leq A \right\}$$
(52)

where $\mathbf{j} = (j_1, j_2, \cdots, j_r)^T$, $\mathbf{s}^* = (s_1^*, s_2^*, \cdots, s_r^*)^T$ and s_i^* are defined in (23).

Define the quantity

$$\frac{s_{2o}}{\alpha_{2o}} = \min \left\{ \frac{s_2}{\alpha_2}, \frac{s_3}{\alpha_3}, \cdots, \frac{s_r}{\alpha_r} \right\},$$
(53)

and denote the dimension of f that corresponds to (53) by i_o.

Theorem 3 *Let* $\min\{s_1, s_2\} \geq \max\{\frac{1}{\pi}, \frac{1}{2}\}$ *with* $1 \leq \pi, q \leq \infty$, *and* $A > 0$. *Then, under conditions (45) and (52), for* $1 \leq p < \infty$, *as* $\varepsilon \to 0$,

$$R(B_{\pi,q}^{s_1,s_2,\cdots,s_r}(A)) \geq CA^p \begin{cases} \left[\frac{\varepsilon^{2\alpha}}{A^2} \right]^{\frac{ps_{2o}}{2s_{2o}+\alpha_{2o}}}, & \text{if } s_1 > \frac{s_{2o}}{\alpha_{2o}}(2\nu + \alpha_1), \ \frac{s_{2o}}{\alpha_{2o}} \geq \frac{p}{2}(\frac{1}{p'} - \frac{1}{p}) \\[2ex] \left[\frac{\varepsilon^{2\alpha}}{A^2} \right]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}, & \text{if } \frac{p}{2}(2\nu+\alpha_1)(\frac{1}{\pi} - \frac{1}{p}) \leq s_1 \leq \frac{s_{2o}}{\alpha_{2o}}(2\nu + \alpha_1), \\[2ex] \left[\frac{\varepsilon^{2\alpha}}{A^2} \right]^{\frac{p(s_1 + \frac{1}{p} - \frac{1}{\pi})}{2s_1^* + 2\nu + \alpha_1 - 1}}, & \text{if } s_1 < (2\nu + \alpha_1)\frac{p}{2}(\frac{1}{\pi} - \frac{1}{p}), \ \frac{s_{2o}}{\alpha_{2o}} \geq \frac{p}{2}(\frac{1}{p'} - \frac{1}{p}). \end{cases}$$
(54)

Theorem 4 *Let* $\widehat{f}(.,.)$ *be the wavelet estimator in (43), with* J_1 *and* J_i *with* $i = 2, \cdots, r$ *given by (49). Let* $\min\{s_1, s_2, \cdots, s_r\} \geq \max\{\frac{1}{\pi}, \frac{1}{2}\}$ *with* $1 \leq \pi, q \leq \infty$, *and let conditions (45) and (52) hold. If* ρ *in (48) is large enough, then, for* $1 \leq p \leq \infty$, *as* $\varepsilon \to 0$,

$$R(B_{\pi,q}^{s_1,\cdot,s_r}(A)) \leq CA^p \begin{cases} \left[\frac{\varepsilon^{2\alpha}}{A^2} \right]^{\frac{ps_{2o}}{2s_{2o}+\alpha_{2o}}} [|\ln(\varepsilon)|]^{\xi_1 + \frac{ps_{2o}}{2s_2+\alpha_{2o}}}, & \text{if } s_1 \geq \frac{s_{2o}}{\alpha_{2o}}(2\nu + \alpha_1), \ \frac{s_{2o}}{\alpha_{2o}} \geq \frac{p}{2}(\frac{1}{p'} - \frac{1}{p}), \\[2ex] \left[\frac{\varepsilon^{2\alpha}}{A^2} \right]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}} [|\ln(\varepsilon)|]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}, & \text{if } \frac{p}{2}(2\nu+\alpha_1)(\frac{1}{\pi} - \frac{1}{p}) < s_1 < \frac{s_{2o}}{\alpha_{2o}}(2\nu + \alpha_1), \\[2ex] \left[\frac{\varepsilon^{2\alpha}}{A^2} \right]^{\frac{p(s_1 + \frac{1}{p} - \frac{1}{\pi})}{2s_1^* + 2\nu + \alpha_1 - 1}} [|\ln(\varepsilon)|]^{\xi_2 + \frac{p(s_1 + \frac{1}{p} - \frac{1}{\pi})}{2s_1^* + 2\nu + \alpha_1 - 1}}, & \text{if } s_1 \leq (2\nu + \alpha_1)\frac{p}{2}(\frac{1}{\pi} - \frac{1}{p}), \ \frac{s_{2o}}{\alpha_{2o}} \geq \frac{p}{2}(\frac{1}{p'} - \frac{1}{p}), \end{cases}$$
(55)

where ξ_1 and ξ_2 are defined as

$$\xi_1 = \mathbb{I}\left(s_1 = \frac{s_{2o}}{\alpha_{2o}}(2\nu + \alpha_1)\right) + \sum_{i=2, i\neq i_o}^{r} \mathbb{I}\left(\frac{s_{2o}}{\alpha_{2o}} = \frac{s_i}{\alpha_i}\right), \tag{56}$$

$$\xi_2 = \mathbb{I}\left(s_1 = (2\nu + \alpha_1)\frac{p}{2}\left(\frac{1}{\pi} - \frac{1}{p}\right)\right) + \sum_{i=2, i\neq i_o}^{r} \mathbb{I}\left(\frac{p}{2}\left(\frac{1}{p'} - \frac{1}{p}\right) = \frac{s_i}{\alpha_i}\right). \tag{57}$$

Proof of Theorem 4. The proof is very similar to that of Theorem 5 in Benhaddou et al. (2013) so we skip it. $\square$

Remark 7 *Notice that the convergence rates in Theorems 3 and 4 depend on a delicate balance between s_1, π, ν, α_1, $\min\left\{\frac{s_i}{\alpha_i}, i = 2, 3, \cdots, r\right\}$ and the average of the parameters of the anisotropic fractional Brownian sheet, α_i, $i = 1, 2, \cdots, r$. Such rates are comparable to those in Benhaddou et al. (2013) in their white noise case when $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 1$. In addition, the rates are independent of the dimension r as expected.*

6   Proofs.

In the proofs of Theorems 1 and 2 we will use the properties described in Petsa and Sapatinas (2009) associated with Meyer wavelet basis. Namely, the properties of concentration, unconditionality and Temlyakov.

6.1   Proof of the lower bounds.

In order to prove Theorem 1, we use Lemma A.1 of [Bunea, Tsybakov & Wegkamp (2007)].

Lemma 6 *Let Θ be a set of functions of cardinality $card(\Theta) \geq 2$ such that*
(i) $\|f - g\|_p^p \geq 4\delta^p$, for $f, g \in \Theta$, $f \neq g$,
(ii) the Kullback divergences $K(P_f, P_g)$ between the measures P_f and P_g satisfy the inequality
$K(P_f, P_g) \leq \log(card(\Theta))/16$, for $f, g \in \Theta$.
Then, for some absolute positive constant C_1, one has

$$\inf_{f_n} \sup_{f \in \Theta} \mathbb{E}_f \|f_n - f\|_p^p \geq C_1 \delta^p,$$

where $\inf_{f_n}$ *denotes the infimum over all estimators.*

Proof of Theorem 1. Let P_f be the probability law of the process $\{q(t,x)+\varepsilon^{\overline{\alpha}}dB^{\alpha}(t,x), (t,x) \in U\}$ when f is true. Let us first transform this process into a process that involves the standard Brownian sheet, taking into account relation (2). This can be accomplished using the following functions

$$\Gamma_{\alpha}(t,u) = \frac{C_H}{\Gamma(H+1/2)\Gamma(3/2-H)} [(t-u)_+]^{1/2-H}, \ for \ t \in [0,1], \tag{58}$$

where C_H is some explicit constant, $\Gamma(\cdot)$ is the Gamma function. Consider $0 < u < t < 1$, then (58) reduces to

$$\Gamma_{\alpha}(t,u) = C(t-u)^{1/2-H}. \tag{59}$$

Now define the operator K_Γ such that

$$K_\Gamma f(t,x) = \int_R \int_R \Gamma_{\alpha_1}(t,u)\Gamma_{\alpha_2}(x,v)f(u,v)dudv. \tag{60}$$

Taking into account relation (2), it is well known that operator (60) converts an anisotropic two-dimensional fractional Brownian sheet with Hurst parameters $(H_1, H_2) = (1 - \alpha_1/2, 1 - \alpha_2/2)$ into a standard Brownian sheet. Let us then apply operator (60) to equation (1) to obtain,

$$\widehat{Y}(t,x) = K_\Gamma dY(t,x) = K_\Gamma q(t,x) + \varepsilon^{\overline{\alpha}}B(t,x), \tag{61}$$

where $B(t,x)$ are standard Brownian sheet. Finally, differentiating both sides of (61) with respect to t and x, yields

$$d\widehat{Y}(t,x) = K'_\Gamma q(t,x)dtdx + \varepsilon^{\overline{\alpha}}dB(t,x), \tag{62}$$

where $K'_\Gamma q(t,x)$ is

$$K'_\Gamma q(t,x) = \frac{d^2}{dtdx}K_\Gamma q(t,x) = c \int_0^x (x-v)^{-1/2-H_2} \int_0^t f(s,v) \int_0^{t-s} (t-s-r)^{-1/2-H_1}g(r,v)drdsdv. \tag{63}$$

Notice that in the Fourier domain, by the convolution theorem and condition (18), (63) has the form

$$\tilde{K}'_\Gamma q(m_1, m_2) \asymp \tilde{f}(m_1, m_2)|m_1|^{-\nu}|m_1|^{(1-\alpha_1)/2}|m_2|^{(1-\alpha_2)/2}. \tag{64}$$

The dense-dense case. Let η be the matrix with components $\eta_{k_1 k_2} = \{0, 1\}$, $k_i = 0, 1, \cdots, 2^{j_i} - 1$, $i = 1, 2$, and denote the set of all possible values η by H. Define the functions

$$f_{j_1 j_2}(t, x) = \gamma_{j_1 j_2} \sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} \eta_{k_1 k_2} \psi_{j_1 k_1}(t) \psi_{j_2 k_2}(x). \tag{65}$$

Note that the matrix η has $N = 2^{j_1+j_2}$ components and therefore, card$(H) = 2^N$. To guarantee that functions (65) satisfy (26), choose $\gamma_{j_1 j_2} = A2^{-j_1(s_1+1/2)}2^{-j_2(s_2+1/2)}$. Now, take another function of the form (65), $\tilde{f}_{j_1 j_2}$, with $\tilde{\eta}_{k_1 k_2} \in H$ instead of $\eta_{k_1 k_2}$ and compute the L^p-norm of the difference between $f_{j_1 j_2}$ and $\tilde{f}_{j_1 j_2}$, with the application of the Varshamov-Gilbert Lemma ([Tsybakov (2008)], p 104), to obtain

$$\|\tilde{f}_{j_1 j_2} - f_{j_1 j_2}\|_p^p \geq \gamma_{j_1 j_2}^p 2^{j_1 p/2} 2^{j_2 p/2}/8. \tag{66}$$

In addition, define $\widehat{P}_f$, the probability law of the process $\left\{K'_\Gamma q(t, x) + \varepsilon^{\overline{\alpha}} dW(t, x), (t, x) \in U\right\}$ when f is true, where $W(t, x)$ is a two-parameter standard Wiener sheet. Note that the probability measures $\widehat{P}_f$ and P_f are stochastically equivalent (see, e.g., Huang et al. (2006), Theorem 3.1). Hence, by the multi-parameter Girsanov formula (see, e.g., Dozzi (1989), p.89), the Kullback divergence takes the form

$$K(P_f, P_{\tilde{f}}) = K(\widehat{P}_f, \widehat{P}_{\tilde{f}}) = \mathbb{E}\left[\log\left(\widehat{P}_f/\widehat{P}_{\tilde{f}}\right)\right]_{\widehat{P}_f} = \left(2\varepsilon^{2\overline{\alpha}}\right)^{-1}\left\|K'_\Gamma g * (\tilde{f} - f)(t, x)\right\|^2. \tag{67}$$

Since $|\tilde{\eta}_{k_1 k_2} - \eta_{k_1 k_2}| \leq 1$, plugging $f_{j_1 j_2}$ and $\tilde{f}_{j_1 j_2}$ into (67), applying Plancherel's formula, $|\psi_{j,k,m}| \leq 2^{-j/2}$ and (64), it can be shown that

$$K(\widehat{P}_f, \widehat{P}_{\tilde{f}}) \leq C\left(2\varepsilon^{2\overline{\alpha}}\right)^{-1}\gamma_{j_1 j_2}^2 2^{-j_1(2\nu+\alpha_1-1)}2^{-j_2(\alpha_2-1)}2^{j_1+j_2}. \tag{68}$$

Part (ii) of Lemma 6 gives the constraint

$$2^{-j_1(2s_1+2\nu+\alpha_1)}2^{-j_2(2s_2+\alpha_2)} \leq \frac{C\varepsilon^{\alpha_1+\alpha_2}}{A^2}. \tag{69}$$

Now, define

$$\tau_\varepsilon = \log_2\left(\frac{CA^2}{\varepsilon^{\alpha_1+\alpha_2}}\right). \tag{70}$$

Therefore, we need to find combination $\{j_1, j_2\}$ which is the solution to the following optimization problem

$$2j_1s_1 + 2j_2s_2 \xrightarrow{\min} j_1(2s_1+2\nu+\alpha_1) + j_2(2s_2+\alpha_2) \geq \tau_\varepsilon, \;\; j_1, j_2 \geq 0. \tag{71}$$

It is easy to check that the solution is $\{j_1, j_2\} = \{\frac{\tau_\varepsilon}{2s_1+2\nu+\alpha_1}, 0\}$, if $\alpha_2 s_1 < s_2(2\nu+\alpha_1)$, and $\{j_1, j_2\} = \{0, \frac{\tau_\varepsilon}{2s_2+\alpha_2}\}$, if $\alpha_2 s_1 \geq s_2(2\nu+\alpha_1)$. Hence, the lower bounds are

$$\delta^p = CA^p \begin{cases} \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}, & \text{if } \alpha_2 s_1 \leq s_2(2\nu+\alpha_1), \\ \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{ps_2}{2s_2+\alpha_2}}, & \text{if } \alpha_2 s_1 > s_2(2\nu+\alpha_1). \end{cases} \tag{72}$$

The sparse-dense case. Using the same test functions $\tilde{f}_{j_1j_2}$ and $f_{j_1j_2}$, as in Benhaddou et al. (2013), with a finitely supported basis $\eta_{j_2k_2}$ replaced by a band-limited basis $\psi_{j_2k_2}$, and following the same procedure as in the dense-dense case, it can be shown that the lower bounds are

$$\delta^p = CA^p \begin{cases} \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{ps_2}{2s_2+\alpha_2}}, & \text{if } \alpha_2 s_1^* \geq s_2(2\nu+\alpha_1-1), \\ \left[\frac{\varepsilon^{2\overline{\alpha}}}{A^2}\right]^{\frac{p(s_1+1/p-1/\pi)}{2s_1^*+2\nu+\alpha_1-1}}, & \text{if } \alpha_2 s_1^* < s_2(2\nu+\alpha_1-1). \end{cases} \tag{73}$$

To complete the proof, notice that the highest of the lower bounds corresponds to

$$d = \min\left\{\frac{ps_1}{2s_1+2\nu+\alpha_1}, \frac{ps_2}{2s_2+\alpha_2}, \frac{p(s_1+\frac{1}{p}-\frac{1}{\pi})}{2s_1^*+2\nu+\alpha_1-1}\right\}. \tag{74}$$

□

Proof of Theorem 3.

The dense-dense case. Let η be the matrix with components $\eta_{\mathbf{k}} = \{0, 1\}$, $\mathbf{k} = (k_1, k_2, \cdots, k_r)$, $k_i = 0, 1, \cdots, 2^{j_i}-1$, $i = 1, 2, \cdots, r$, and denote the set of all possible values η by H. Define

the functions

$$f_{\mathbf{j}}(t,x) = \gamma_{\mathbf{j}} \sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} \cdots \sum_{k_r=1}^{2^{j_r}-1} \eta_{\mathbf{k}} \psi_{j_1,k_1}(t) \Pi_{i=1}^{r-1} \psi_{j_i,k_i}(x_i). \tag{75}$$

Note that here $\operatorname{card}(H) = 2^N$, with $N = 2^{j_1+j_2+\cdots+j_r}$. Hence, following the same steps as in the proof of Theorem 1, it can be shown that Note that the lower bounds are

$$\delta^p = CA^p \begin{cases} \left[\frac{\varepsilon^{2\alpha}}{A^2}\right]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}, & \text{if } \alpha_{2o}s_1 \le s_{2o}(2\nu+\alpha_1), \\ \left[\frac{\varepsilon^{2\alpha}}{A^2}\right]^{\frac{ps_i}{2s_i+\alpha_i}}, & \text{if } \alpha_i s_1 > s_i(2\nu+\alpha_1), \ \frac{s_i}{\alpha_i} < \min_{i\neq k}\{\frac{s_k}{\alpha_k}, k=2,3,\cdots,r\}, \end{cases} \tag{76}$$

and $i = 2, 3, \cdots, r$.

The sparse-dense case. We apply similar approach to test functions

$$f_{\mathbf{j}}(t,x) = \gamma_{\mathbf{j}} \sum_{k_2=0}^{2^{j_2}-1} \cdots \sum_{k_r=1}^{2^{j_r}-1} \eta_{\mathbf{k}} \psi_{j_1,k_1}(t) \Pi_{i=1}^{r-1} \psi_{j_i,k_i}(x_i). \tag{77}$$

Here $N = 2^{j_2+\cdots+j_r}$ and $\operatorname{card}(H) = 2^N$ and it can be shown that the lower bounds are

$$\delta^p = CA^p \begin{cases} \left[\frac{\varepsilon^{2\alpha}}{A^2}\right]^{\frac{ps_i}{2s_i+\alpha_i}}, & \text{if } \alpha_i s_1^* \ge s_i(2\nu+\alpha_1-1), \ \frac{s_i}{\alpha_i} < \min_{i\neq k}\{\frac{s_k}{\alpha_k}, k=2,3,\cdots,r\} \\ \left[\frac{\varepsilon^{2\alpha}}{A^2}\right]^{\frac{p(s_1+1/p-1/\pi)}{2s_1^*+2\nu+\alpha_1-1}}, & \text{if } \alpha_i s_1^* < s_i(2\nu+\alpha_1-1), \end{cases} \tag{78}$$

and $i = 2, 3, \cdots, r$. Hence, the lower bounds of the L^p-risk correspond to

$$d' = \min\left\{ \frac{ps_1}{2s_1+2\nu+\alpha_1}, \frac{p(s_1+\frac{1}{p}-\frac{1}{\pi})}{2s_1^*+2\nu+\alpha_1-1}, \frac{ps_{2o}}{2s_{2o}+\alpha_{2o}} \right\}, \tag{79}$$

and the pair (s_{2o}, α_{2o}) is defined in (53). $\square$

6.2   Proof of the upper bounds.

Proof of Lemma 1. Define the quantities

$$\tilde{\beta}_\omega - \beta_\omega = \varepsilon^{\overline{\alpha}} \sum_{m_1 \in W_{j_1}} \sum_{m_2 \in W_{j_2}} \overline{\tilde{\psi}_{j_1 k_1}(m_1)} \tilde{\psi}_{j_2 k_2}(m_2) \frac{\tilde{Z}^\alpha(m_1, m_2)}{\tilde{g}(m_1, m_2)}. \tag{80}$$

Consider the Riesz poly-potential operator

$$I^\alpha f(t,x) = \frac{1}{\gamma_2(\alpha)} \int_{R^2} \frac{f(t_2, x_2)}{|t_1 - t_2|^{1/2 + \alpha_1/2} |x_1 - x_2|^{1/2 + \alpha_2/2}} dt_2 dx_2, \tag{81}$$

where $\gamma_2(\alpha) = 4\Gamma\left(\frac{1}{2} - \frac{\alpha_1}{2}\right)\Gamma\left(\frac{1}{2} - \frac{\alpha_2}{2}\right)cos((1-\alpha_1)\frac{\pi}{4})cos((1-\alpha_2)\frac{\pi}{4})$. Then, the anisotropic two-dimensional fractional Brownian sheet $Z^\alpha(t,x) = dW^\alpha(t,x)$ allows the wavelet-vaguelette representation

$$dW^\alpha(t,x) = \sum_{j_1, k_1} \sum_{j_2, k_2} \zeta_{j_1 k_1; j_2 k_2} I^\alpha \left\{ \psi_{j_1 k_1}(t)\psi_{j_2 k_2}(x) \right\}, \tag{82}$$

where $\zeta_{j_1 k_1; j_2 k_2}$ are white noise processes, and ψ_{jk} is a Meyer-type wavelet basis. Then applying the two dimensional Fourier transform to (82) yields

$$\tilde{W}^\alpha(\mathbf{m}) = |m_1|^{\frac{\alpha_1 - 1}{2}} |m_2|^{\frac{\alpha_2 - 1}{2}} \sum_{j_1, k_1} \sum_{j_2, k_2} \zeta_{j_1 k_1; j_2 k_2} \overline{\tilde{\psi}_{j_1 k_1}(m_1)} \tilde{\psi}_{j_2 k_2}(m_2), \tag{83}$$

where $\mathbf{m} = (m_1, m_2)$. Let us evaluate the covariance of (83). Indeed,

$$\text{Cov}\left[\tilde{Z}^\alpha(\mathbf{m}), \tilde{Z}^\alpha(\mathbf{l})\right] = |m_1 l_1|^{\frac{\alpha_1 - 1}{2}} |m_2 l_2|^{\frac{\alpha_2 - 1}{2}} \sum_{\omega \in \Omega} \overline{\tilde{\psi}_{j_1 k_1}(m_1)} \tilde{\psi}_{j_2 k_2}(m_2) \overline{\tilde{\psi}_{j_1 k_1}(l_1)} \tilde{\psi}_{j_2 k_2}(l_2). \tag{84}$$

Consequently,

$$\text{Cov}\left[\tilde{Z}^\alpha(\mathbf{m}), \tilde{Z}^\alpha(\mathbf{l})\right] = |m_1 l_1|^{\frac{\alpha_1 - 1}{2}} |m_2 l_2|^{\frac{\alpha_2 - 1}{2}} \left(\sum_{j_1; k_1} \tilde{\psi}_{j_1 k_1}(m_1)\tilde{\psi}_{j_1 k_1}(l_1) \right) \left(\sum_{j_2; k_2} \tilde{\psi}_{j_2 k_2}(m_2)\tilde{\psi}_{j_2 k_2}(l_2) \right). \tag{85}$$

Now, evaluating the magnitude of (85), and using Hölder's Inequality along with the fact that $|\tilde{\psi}_{jk}(m)| \leq 2^{-j/2}$, yields

$$
\begin{aligned}
|\text{Cov}\left[\tilde{Z}^\alpha(\mathbf{m}), \tilde{Z}^\alpha(\mathbf{l})\right]|^2 &\leq |m_1 l_1|^{\alpha_1 - 1} |m_2 l_2|^{\alpha_2 - 1} \left(\sum_{j_1; k_1} |\tilde{\psi}_{j_1 k_1}(m_1)|^2 |\tilde{\psi}_{j_1 k_1}(l_1)|^2 \right) \left(\sum_{j_2; k_2} |\tilde{\psi}_{j_2 k_2}(m_2)|^2 |\tilde{\psi}_{j_2 k_2}(l_2)|^2 \right) \\
&\leq 4|m_1 l_1|^{\alpha_1 - 1} |m_2 l_2|^{\alpha_2 - 1}.
\end{aligned}
$$

Hence,

$$|\mathrm{Cov}\,[\tilde{Z}^{\alpha}(\mathbf{m}), \tilde{Z}^{\alpha}(\mathbf{l})]| \le 2|m_1 l_1|^{\frac{\alpha_1-1}{2}}|m_2 l_2|^{\frac{\alpha_2-1}{2}}. \tag{86}$$

Finally, let us evaluate the variance of $\tilde{\beta}_\omega$. Indeed, using (80) and (86), the variance has the form

$$
\begin{aligned}
\mathbb{E}\left|\tilde{\beta}_\omega - \beta_\omega\right|^2 &= \varepsilon^{\alpha_1+\alpha_2} \sum_{m_1,m_2}\sum_{l_1,l_2} \overline{\tilde{\psi}_{j_1 k_1}(m_1)}\tilde{\psi}_{j_2 k_2}(m_2)\overline{\tilde{\psi}_{j_1 k_1}(l_1)}\tilde{\psi}_{j_2 k_2}(l_2) \frac{\mathbb{E}[\overline{\tilde{Z}^{\alpha}(m_1,m_2)}\tilde{Z}^{\alpha}(l_1,l_2)]}{\tilde{g}(m_1,m_2)\tilde{g}(l_1,l_2)} \\
&\le 2\varepsilon^{\alpha_1+\alpha_2}\left(\sum_{m_1,m_2}\tilde{\Psi}_\omega(m_1,m_2)\frac{|m_1|^{\alpha_1-1}|m_2|^{\alpha_2-1}}{\tilde{g}(m_1,m_2)}\right)\left(\sum_{l_1,l_2}\tilde{\Psi}_\omega(l_1,l_2)\frac{|l_1|^{\alpha_1-1}|l_2|^{\alpha_2-1}}{\tilde{g}(l_1,l_2)}\right) \\
&= 2\varepsilon^{\alpha_1+\alpha_2}\left(\sum_{m_1,m_2}\overline{\tilde{\psi}_{j_1 k_1}(m_1)}\tilde{\psi}_{j_2 k_2}(m_2)\frac{|m_1|^{\alpha_1-1}|m_2|^{\alpha_2-1}}{\tilde{g}(m_1,m_2)}\right)^2 \\
&\le 2\varepsilon^{\alpha_1+\alpha_2}C_1^{-2}\left(\frac{8\pi}{3}\right)^{2\nu+\alpha_1+\alpha_2-2}2^{j_1(2\nu+\alpha_1-1)+j_2(\alpha_2-1)}. \tag{87}
\end{aligned}
$$

Results (19) and (20) follow from properties of Gaussian random variables. $\square$

Proof of Lemma 2. First, note that under **Assumption A.3.**, one has

$$\sum_{k_1,k_2}\mid \beta_{j_1,k_1,j_2,k_2}\mid^\pi \le A^\pi 2^{-\pi[(j_1 s_1+j_2 s_2)+(\frac{1}{2}-\frac{1}{\pi})(j_1+j_2)]},\ \forall j_1, j_2 \ge 0. \tag{88}$$

If $p \ge \pi$, one has

$$
\begin{aligned}
\sum_{k_1,k_2}\mid \beta_{j_1,k_1,j_2,k_2}\mid^p &\le \sum_{k_1,k_2}\left(\mid \beta_{j_1,k_1,j_2,k_2}\mid^\pi \{\max|\beta_{j_1,k_1,j_2,k_2}|^\pi\}^{\frac{p-\pi}{\pi}}\right) \\
&= \{\max|\beta_{j_1,k_1,j_2,k_2}|^\pi\}^{\frac{p-\pi}{\pi}}\sum_{k_1,k_2}\mid \beta_{j_1,k_1,j_2,k_2}\mid^\pi \\
&\le A^p 2^{-p[(j_1 s_1+j_2 s_2)+(\frac{1}{2}-\frac{1}{\pi})(j_1+j_2)]}.
\end{aligned}
$$

23

If $p < \pi$, then by Hölder's Inequality, one has

$$
\begin{aligned}
\sum_{k_1,k_2} | \beta_{j_1,k_1,j_2,k_2} |^p &\leq \left(\sum_{k_1,k_2} | \beta_{j_1,k_1,j_2,k_2} |^\pi \right)^{\frac{p}{\pi}} \left(\sum_{k_1,k_2} 1 \right)^{1-\frac{p}{\pi}} \\
&\leq A^p 2^{-p[(j_1 s_1 + j_2 s_2)+(\frac{1}{2}-\frac{1}{\pi})(j_1+j_2)]} 2^{(j_1+j_2)(1-\frac{p}{\pi})} \\
&\leq A^p 2^{-p[(j_1 s_1 + j_2 s_2)+(\frac{1}{2}-\frac{1}{p})(j_1+j_2)]}.
\end{aligned}
$$

Combining the above results completes the proof. $\square$

Proof of Lemma 3. Since the quantities (80) are zero mean Gaussian random variables having variance of form (87), by the Gaussian tail probability inequality, one has

$$
\begin{aligned}
\Pr(\Theta_{\omega,\eta}) &\leq 2 \Pr \left(Z > \eta \gamma \sqrt{|\ln(\varepsilon)|} \left(\frac{C_1^2}{2} \right)^{1/2} \left(\frac{3}{8\pi} \right)^{\nu + \alpha_1/2 + \alpha_2/2 - 1} \right) \\
&= \frac{C}{\sqrt{|\ln(\varepsilon)|}} \exp \left\{ -2\overline{\alpha} |\ln(\varepsilon)| \tau \right\},
\end{aligned}
\tag{89}
$$

where Z is a standard normal, C is some explicit positive constant and τ appears in Lemma 2. $\square$

Proof of Theorem 2. The proof follows a procedure which is similar to that in Benhaddou et al. (2013) with the adaptation to the L^p-norm case. Denote

$$
\chi_\varepsilon = \left[A^{-2} \varepsilon^{2\overline{\alpha}} |\ln(\varepsilon)| \right], \quad 2^{j_{i0}} = [\chi_\varepsilon]^{-\frac{d}{ps''_i}}, \quad i = 1, 2,
\tag{90}
$$

and observe that with the choice of J_1 and J_2 given by (22), the L^p-risk allows the decomposition

$$
\mathbb{E} \| \widehat{f}_\epsilon - f \|_p^p \leq R_1 + R_2 + R_3 + R_4,
\tag{91}
$$

where

$$R_1 = 2^{2(p-1)} \mathbb{E} \left\| \sum_{k_1=0}^{2^{m_0}-1} \sum_{k_2=0}^{2^{m_0}-1} (\tilde{\beta}_{m_0,k_1,m_0,k_2} - \beta_{m_0,k_1,m_0,k_2}) \psi_{m_0,k_1}(t) \psi_{m_0,k_2}(x) \right\|_p^p,$$

$$R_2 = 2^{3(p-1)} \mathbb{E} \int \int \left(\sum_{\boldsymbol{\omega} \in \Omega(J_1,J_2)} \left| \tilde{\beta}_{\boldsymbol{\omega}} - \beta_{\boldsymbol{\omega}} \right|^2 \mathbb{I} \left(\left| \tilde{\beta}_{\boldsymbol{\omega}} \right| > \lambda_{j;\varepsilon}^\alpha \right) \psi_{j_1,k_1}^2(t) \psi_{j_2,k_2}^2(x) \right)^{p/2} dt\, dx,$$

$$R_3 = 2^{3(p-1)} \mathbb{E} \int \int \left(\sum_{\boldsymbol{\omega} \in \Omega(J_1,J_2)} |\beta_{\boldsymbol{\omega}}|^2 \mathbb{I} \left(\left| \tilde{\beta}_{\boldsymbol{\omega}} \right| < \lambda_{j;\varepsilon}^\alpha \right) \psi_{j_1,k_1}^2(t) \psi_{j_2,k_2}^2(x) \right)^{p/2} dt\, dx,$$

$$R_4 = 2^{(p-1)} \int \int \left(\left(\sum_{j_1=J_1}^{\infty} \sum_{j_2=m_0+1}^{\infty} + \sum_{j_1=m_0+1}^{\infty} \sum_{j_2=J_2}^{\infty} \right) \sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} |\beta_{\boldsymbol{\omega}}|^2 \psi_{j_1,k_1}^2(t) \psi_{j_2,k_2}^2(x) \right)^{p/2} dt\, dx.$$

For R_1 and R_4, using (19) in R_1 and (26) in R_4, as $\varepsilon \to 0$, yields

$$R_1 + R_4 = O \left(\varepsilon^{2\overline{\alpha}} + \sum_{j_1=J_1}^{\infty} A^p 2^{-p j_1 s_1''} + \sum_{j_2=J_2}^{\infty} A^p 2^{-p j_2 s_2''} \right) = O \left(A^p [\chi_\varepsilon]^d \right). \qquad (92)$$

Using Minkowski's Inequality for $p \geq 2$ and Jensen's Inequality for $p < 2$, and Temlyakov property, R_2 and R_3 can be partitioned as $R_2 \leq R_{21} + R_{22}$ and $R_3 \leq R_{31} + R_{32}$, where

$$R_{21} = O \left(\sum_{\boldsymbol{\omega} \in \Omega(J_1,J_2)} 2^{(j_1+j_2)(\frac{p}{2}-1)} \mathbb{E} \left| \tilde{\beta}_{\boldsymbol{\omega}} - \beta_{\boldsymbol{\omega}} \right|^p \left[\Pr \left(\left| \tilde{\beta}_{\boldsymbol{\omega}} - \beta_{\boldsymbol{\omega}} \right| > \frac{1}{2} \lambda_{j;\varepsilon}^\alpha \right) \right]^{1/2} \right), \quad (93)$$

$$R_{22} = O \left(\sum_{\boldsymbol{\omega} \in \Omega(J_1,J_2)} 2^{(j_1+j_2)(\frac{p}{2}-1)} \mathbb{E} \left| \tilde{\beta}_{\boldsymbol{\omega}} - \beta_{\boldsymbol{\omega}} \right|^p \mathbb{I} \left(|\beta_{\boldsymbol{\omega}}| > \frac{1}{2} \lambda_{j;\varepsilon}^\alpha \right) \right), \qquad (94)$$

$$R_{31} = O \left(\sum_{\boldsymbol{\omega} \in \Omega(J_1,J_2)} 2^{(j_1+j_2)(\frac{p}{2}-1)} |\beta_{\boldsymbol{\omega}}|^p \Pr \left(\left| \tilde{\beta}_{\boldsymbol{\omega}} - \beta_{\boldsymbol{\omega}} \right| > \frac{1}{2} \lambda_{j;\varepsilon}^\alpha \right) \right), \qquad (95)$$

$$R_{32} = O \left(\sum_{\boldsymbol{\omega} \in \Omega(J_1,J_2)} 2^{(j_1+j_2)(\frac{p}{2}-1)} |\beta_{\boldsymbol{\omega}}|^p \mathbb{I} \left(|\beta_{\boldsymbol{\omega}}| < \frac{3}{2} \lambda_{j;\varepsilon}^\alpha \right) \right). \qquad (96)$$

Combining (93) and (95), similar calculations as in Benhaddou et al. (2013), yield

$$R_{21} + R_{31} = O \left(A^p \left(\varepsilon^{2\overline{\alpha}} \right)^{\frac{\tau}{2} - \frac{p}{2}} \right) = O \left(A^p [\chi_\varepsilon]^d \right). \qquad (97)$$

Now, combining (94) and (96), and using (19) and (21), gives

$$\Delta = R_{22} + R_{32} = O\left(\sum_{\boldsymbol{\omega} \in \Omega(J_1, J_2)} 2^{(j_1+j_2)(\frac{p}{2}-1)} \min\left\{|\beta_{\boldsymbol{\omega}}|^p, 2^{\frac{p}{2}[j_1(2\nu+\alpha_1-1)+j_2(\alpha_2-1)]}\left[A^{-2}\varepsilon^{2\overline{\alpha}}|\ln(\varepsilon)|\right]^{\frac{p}{2}}\right\}\right). \tag{98}$$

Finally, Δ can be decomposed into the following components

$$\Delta_1 = O\left(\left\{\sum_{j_1=j_{10}+1}^{J_1-1}\sum_{j_2=m_0}^{J_2-1} + \sum_{j_1=m_0}^{J_1-1}\sum_{j_2=j_{20}+1}^{J_2-1}\right\} A^p 2^{-j_1 p s_1''} 2^{-j_2 p s_2''}\right), \tag{99}$$

$$\Delta_2 = O\left(\sum_{j_1=m_0}^{j_{10}}\sum_{j_2=m_0}^{j_{20}} A^p 2^{\frac{p}{2}(j_1(2\nu+\alpha_1)+\alpha_2 j_2)}\left[A^{-2}\varepsilon^2|\ln(\varepsilon)|\right]^{\frac{p}{2}} \mathbb{I}\left(\Xi\right)\right), \tag{100}$$

$$\Delta_3 = O\left(\sum_{j_1=m_0}^{j_{10}}\sum_{j_2=m_0}^{j_{20}} A^{p'} 2^{-p'(j_1 s_1'' + j_2 s_2'')}\left[A^p[\chi_\varepsilon]^{\frac{p}{2}} 2^{j_1(\frac{p}{2}(2\nu+\alpha_1)-1)+j_2(\alpha_2\frac{p}{2}-1)}\right]^{1-\frac{p'}{p}} \mathbb{I}\left(\Xi^c\right)\right). \tag{101}$$

where $\Xi = \left\{j_1, j_2 : 2^{\frac{p}{2}(j_1(2\nu+\alpha_1)+\alpha_2 j_2)} \le [\chi_\varepsilon]^{d-\frac{p}{2}}\right\}$.

Case 1: $s_1 \ge \frac{s_2}{\alpha_2}(2\nu+\alpha_1)$ and $\frac{s_2}{\alpha_2} > \frac{p}{2}(\frac{1}{p'}-\frac{1}{p})$. In this case, $d = \frac{ps_2}{2s_2+\alpha_2}$, and

$$\begin{aligned}
\Delta_3 &= O\left(A^p[\chi_\varepsilon]^{\frac{p}{2}(1-\frac{p'}{p})}\sum_{j_1=m_0}^{j_{10}} 2^{-p'j_1[s_1-\frac{p}{2}(\frac{1}{p'}-\frac{1}{p})(2\nu+\alpha_1)]}\sum_{j_2=m_0}^{j_{20}} 2^{-j_2 p'[s_2-\frac{p}{2}\alpha_2(\frac{1}{p'}-\frac{1}{p})]}\mathbb{I}\left(\Xi^c\right)\right)\\
&= O\left(A^p[\chi_\varepsilon]^{\frac{ps_2}{2s_2+\alpha_2}}\sum_{j_1=m_0}^{j_{10}} 2^{-p'j_1[s_1-\frac{s_2}{\alpha_2}(2\nu+\alpha_1)]}\right)\\
&= O\left(A^p[\chi_\varepsilon]^{\frac{ps_2}{2s_2+\alpha_2}}[|\ln(\epsilon)|]^{\mathbb{I}(s_1=\frac{s_2}{\alpha_2}(2\nu+\alpha_1))}\right). \tag{102}
\end{aligned}$$

Case 2: $(2\nu+\alpha_1)\frac{p}{2}(\frac{1}{\pi}-\frac{1}{p}) < s_1 < \frac{s_2}{\alpha_2}(2\nu+\alpha_1)$. In this case, $d = \frac{ps_1}{2s_1+2\nu+\alpha_1}$, and

$$\begin{aligned}
\Delta_3 &= O\left(A^p[\chi_\varepsilon]^{\frac{p}{2}(1-\frac{p'}{p})}\sum_{j_1=m_0}^{j_{10}} 2^{-p'j_1[s_1-(2\nu+\alpha_1)\frac{p}{2}(\frac{1}{p'}-\frac{1}{p})]}\sum_{j_2=m_0}^{j_{20}} 2^{-j_2 p'[s_2-\frac{p}{2}\alpha_2(\frac{1}{p'}-\frac{1}{p})]}\mathbb{I}\left(\Xi^c\right)\right)\\
&= O\left(A^p[\chi_\varepsilon]^{\frac{ps_1}{2s_1+2\nu+\alpha_1}}\sum_{j_2=m_0}^{j_{20}} 2^{-j_2 p'\left(s_2-\frac{\alpha_2 s_1}{2\nu+\alpha_1}\right)}\right)\\
&= O\left(A^p[\chi_\varepsilon]^d\right). \tag{103}
\end{aligned}$$

Case 3: $s_1 \leq (2\nu + \alpha_1)\frac{p}{2}(\frac{1}{\pi} - \frac{1}{p})$ and $\frac{s_2}{\alpha_2} > \frac{p}{2}(\frac{1}{p'} - \frac{1}{p})$. In this case, $d = \frac{p(s_1 + \frac{1}{p} - \frac{1}{\pi})}{2s_1^* + 2\nu + \alpha_1 - 1}$, and

$$
\begin{aligned}
\Delta_3 &= O\left(A^p[\chi_\varepsilon]^{\frac{p}{2}(1 - \frac{p'}{p})} \sum_{j_1 = m_0}^{j_{10}} 2^{-p'j_1[s_1 - \frac{p}{2}(\frac{1}{p'} - \frac{1}{p})(2\nu + \alpha_1)]} \sum_{j_2 = m_0}^{j_{20}} 2^{-p'j_2[s_2 - \frac{p}{2}(\frac{1}{p'} - \frac{1}{p})\alpha_2]}\right) \\
&= O\left(A^p[\chi_\varepsilon]^{\frac{p(s_1 + \frac{1}{p} - \frac{1}{\pi})}{2s_1^* + 2\nu + \alpha_1 - 1}} [|\ln(\epsilon)|]^{\mathbb{I}\left(s_1 = (2\nu + \alpha_1)\frac{p}{2}(\frac{1}{p'} - \frac{1}{p})\right) + \mathbb{I}\left(s_2 = \alpha_2 \frac{p}{2}(\frac{1}{p'} - \frac{1}{p})\right)}\right). \quad (104)
\end{aligned}
$$

Combining the results from (92) to (104) completes the proof. $\square$

Proof of Lemma 4. The proof will be similar to the proof of **Lemma 1**. Define the quantities

$$
\tilde{\beta}_{\omega_r} - \beta_{\omega_r} = \varepsilon^{\overline{\alpha}} \sum_{m_1 \in W_{j_1}} \cdots \sum_{m_r \in W_{j_r}} \tilde{\Psi}_{\omega_r}(\mathbf{m_r}) \frac{\tilde{Z}^\alpha(\mathbf{m_r})}{\tilde{g}(\mathbf{m_r})}. \quad (105)
$$

Consider the Riesz poly-potential operator

$$
I^\alpha f(t, x_1, \ldots, x_{r-1}) = \frac{1}{\gamma_r(\alpha)} \int_{R^r} \frac{f(t', x_1', \ldots, x_{r-1}')}{|t - t'|^{1/2 + \alpha_1/2} \Pi_{i=2}^r |x_{i-1} - x_{i-1}'|^{1/2 + \alpha_i/2}} dt' dx_1' \cdots dx_{r-1}', \quad (106)
$$

where $\gamma_r(\alpha) = 2^r \Pi_{i=1}^r \left[\Gamma\left(\frac{1}{2} - \frac{\alpha_i}{2}\right) \cos((1 - \alpha_i)\frac{\pi}{4})\right]$. Then, the anisotropic r-dimensional fractional Brownian sheet $Z^\alpha(t, x_1, \ldots, x_{r-1}) = dW^\alpha(t, x_1, \ldots, x_{r-1})$ allows the wavelet-vaguelette representation

$$
dW^\alpha(t, x_1, \ldots, x_{r-1}) = \sum_{\omega_r \in \Omega_r} \zeta_{\omega_r} I^\alpha \{\Psi_{\omega_r}(t, \mathbf{x})\}, \quad (107)
$$

Applying the two dimensional Fourier transform to (107), yields

$$
\tilde{W}^\alpha(\mathbf{m_r}) = \Pi_{i=1}^r |m_i|^{\frac{\alpha_i - 1}{2}} \sum_{\omega_r \in \Omega_r} \zeta_{\omega_r} \tilde{\Psi}_{\omega_r}(\mathbf{m_r}), \quad (108)
$$

The covariance of (108) is given by

$$
\text{Cov}\left[\tilde{Z}^\alpha(\mathbf{m_r}), \tilde{Z}^\alpha(\mathbf{l_r})\right] = \Pi_{i=1}^r \left(|m_i l_i|^{\frac{\alpha_i - 1}{2}} \sum_{j_i; k_i} \tilde{\psi}_{j_i k_i}(m_i) \tilde{\psi}_{j_i k_i}(l_i)\right). \quad (109)
$$

Evaluating the magnitude of (109), yields

$$|\text{Cov} \, [\tilde{Z}^\alpha(\mathbf{m_r}), \tilde{Z}^\alpha(\mathbf{l_r})]| \le 2^{\frac{r}{2}} \Pi_{i=1}^r |m_i l_i|^{\frac{\alpha_i - 1}{2}}. \tag{110}$$

Using (105) and (110), similar to (87), the variance has the form

$$\begin{aligned}
\mathbb{E}\left|\tilde{\beta}_{\omega_r} - \beta_{\omega_r}\right|^2 &= \varepsilon^{2\overline{\alpha}} \sum_{m_1,\ldots,m_r} \sum_{l_1,\ldots,l_r} \overline{\tilde{\Psi}_{\omega_r}(\mathbf{m_r})} \tilde{\Psi}_{\omega_r}(\mathbf{l_r}) \frac{\mathbb{E}[\overline{\tilde{Z}^\alpha(\mathbf{m_r})} \tilde{Z}^\alpha(\mathbf{l_r})]}{\tilde{g}(\mathbf{m_r})\tilde{g}(\mathbf{l_r})} \\
&\le K\varepsilon^{2\overline{\alpha}} 2^{j_1(2\nu+\alpha_1-1)+j_2(\alpha_2-1)+\cdots+j_r(\alpha_r-1)}.
\end{aligned} \tag{111}$$

Finally, results (46) and (47) follow from properties of Gaussian random variables. $\square$

Proof of Lemma 5. The proof is very similar to that of Lemma 3, so we skip it. $\square$

References

[Abramovich & Silverman (1998)] Abramovich, F. & Silverman, B.W. (1998), 'Wavelet decomposition approaches to statistical inverse problems', *Biometrika*, **85**, 115–129.

[Adu & Richardson (2018)] Adu, N. & Richardson, G. (2018), 'Unit root test: spatial model with long memory errors', *Statistics and Probability Letters,* **140**, 126-131.

[Ayache & Taqqu (2002)] Ayache, A. & Taqqu, M.S. (2002), 'Rate optimality of wavelet series approximations of fractional Brownian motion', *Journal of Fourier Analysis and Applications,* **9**, 451-471.

[Benhaddou (2018)] Benhaddou, R. (2018), 'Laplace deconvolution with dependent errors: a minimax study' *Journal of Nonparametric Statistics,* **30(4)**, 1032-1048.

[Benhaddou (2017)] Benhaddou, R. (2017), 'On minimax convergence rates under L^p-risk for the anisotropic functional deconvolution model', *Statistics and Probability Letters,* **130**, 120-125.

[Benhaddou, Kulik, Pensky, & Sapatinas (2014)] Benhaddou, R., Kulik, R., Pensky, M., Sapatinas, T. (2014), 'Multichannel deconvolution with long-range dependence: a minimax study', *Journal of Statistical Planning and Inference*, **148**, 1-19.

[Benhaddou, Pensky & Picard (2013)] Benhaddou, R., Pensky, M., Picard, D. (2013), 'Anisotropic denoising in functional deconvolution model with dimension-free convergence rates', *Electronic Journal of Statistics*, **7**, 1686-1715.

[Bunea, Tsybakov & Wegkamp (2007)] Bunea, F., Tsybakov, A. & Wegkamp, M.H. (2007), 'Aggregation for Gaussian regression', *Annals of Statistics,* **35**, 1674–1697.

[Comte, Dedecker & Taupin (2008)] Comte, F., Dedecker, J., Y., Taupin, M.L. (2008), 'Adaptive density deconvolution with dependent inputs', *Mathematical Methods in Statistics*, **17**, 87–112.

[Donoho (1995)] Donoho, D.L. (1995), 'Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition', *Applied and Computational Harmonic Analysis*, **2**, 101–126.

[Donoho & Raimondo (2004)] Donoho, D.L. & Raimondo, M. (2004), 'Translation invariant deconvolution in a periodic setting' *International Journal of Wavelets, Multiresolution and Information Processing*, **14**, 415–432.

[Dozzi (1989)] Dozzi, M. (1989). *Stochastic Processes with a Multidimensional Parameter,* Longman Scientific & Technical, New York.

[Huang, Li, Wan & Wu (2006)] Huang, Z.Y., Li, C.J., Wan, J.P. & Wu, Y. (2006), 'Fractional Brownian motion and sheet as white noise functionals', *Acta Mathematica Sinica, English Series*, **22**, 1183-1188.

[Johnstone, Kerkyacharian, Picard & Raimondo (2004)] Johnstone, I.M., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004), 'Wavelet deconvolution in a periodic setting', *Journal of the Royal Statistical Society*, Series B. **66**, 547–573. (with discussion, 627–657).

[Kamont (1996)] Kamont, A. (1996), 'On the fractional anisotropic Wiener field', *Probability. Math. Statist*, **16**, 85-98.

[Kulik, Sapatinas & Wishart (2015)] Kulik, R., Sapatinas, T. Wishart, J. R. (2015), 'Multi-channel deconvolution with long-range dependence: upper bounds on the L_p-risk', *Applied and Computational Harmonic Analysis*, **38**, 357-384.

[Meyer, Sellan & Taqqu (1999)] Meyer, Y., Sellan, F. & Taqqu, M.S. (1999), 'Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion', *Journal of Fourier Analysis and Applications,* **5**, 465-494.

[Pensky & Sapatinas (2009)] Pensky, M., Sapatinas, T. (2009), 'Functional deconvolution in a periodic setting: uniform case', *Annals of Statistics*, **37**, 73–104.

[Pensky & Sapatinas (2010)] Pensky, M., Sapatinas, T. (2010), 'On convergence rates equivalency and sampling strategies in a functional deconvolution model' *Annals of Statistics.* **38**, 1793–1844.

[Pensky & Sapatinas (2011)] Pensky, M., Sapatinas, T. (2011), 'Multichannel boxcar deconvolution with growing number of channels' *Electronic Journal of Statistics*, **5**, 53-82.

[Petsa & Sapatinas (2009)] Petsa, A., Sapatinas, T. (2009), 'Minimax convergence rates under L^P-risk in functional deconvolution model', *Statistics and Probability Letters*, **79**, 1568-1576.

[Raimondo & Stewart (2007)] Raimondo, M., Stewart, T. (2007), 'The WaveD transform in R: performs fast transaction-invariant wavelet deconvolution', *Journal of Statistical Software*, **21**.

[Tsybakov (2008)] Tsybakov, A.B. (2008). *Introduction to Nonparametric Estimation*, Springer, New York.

[Walter & Shen (1999)] Walter, G., Shen, X. (1999), 'Deconvolution using Meyer wavelets' *Journal of Integral Equations and Applications*, **11**, 515–534.

[Wang (1997)] Wang, Y. (1996), 'Functional estimation via wavelets shrinkage for Long-memory Data', *Annals of Statistics*, **24**, 466-484.

[Wang (1997)] Wang, Y. (1997), 'Minimax estimation via wavelets for indirect long-memory data', *Journal of Statistical Planning and Inference*, **1**, 45-55.

[Wishart (2013)] Wishart, J. M. (2013), 'Wavelet deconvolution in a periodic setting with long-range dependent errors', *Journal of Statistical Planning and Inference*, **5**, 867-881.

[Wood, Martin, Jung & Sample (2014)] Wood, W.T., Martin, K.M., Jung, W., Sample, J. (2014), Seismic reflectivity effects from seasonal seafloor temperature variation', *Geophysical Research Letters*, **41**, 6826-6832.