Minimax Estimation via Wavelets for Indirect Long-Memory Data

Yazhen Wang*

University of Missouri-Columbia

To appear in Journal of Statistical Planning and Inference

Abstract

In this paper we model linear inverse problems with long-range dependence by a fractional Gaussian noise model and study function estimation based on observations from the model. By using two wavelet-vaguelette decompositions, one for the inverse problem which simultaneously quasi-diagonalizes both the operator and the prior information and one for long-range dependence which decorrelates fractional Gaussian noise, we establish asymptotics for minimax risks, and show that the wavelet shrinkage estimate can be tuned to achieve the minimax convergence rate and significantly outperform linear estimates.

Running head: Inverse Problem with Long-Memory Noise

Key words: Fractional Brownian motion, fractional Gaussian noise, linear inverse problem, long-range dependence, minimax risk, wavelets, wavelet-vaguelette decomposition.

AMS 1980 Subject Classification. Primary: 62G07; 62C20; 42C15.

*Yazhen Wang is an assistant professor at Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211, U.S.A, email: yzwang@stat.missouri.edu

1 Introduction

Suppose we wish to recover a function $f(x)$ from observations about a transformation $(Kf)(x)$ of the function $f(x)$, where K is a linear transformation such as Radon transformation, Abel transformation and convolution. Such problems arise in many scientific areas and are often called linear inverse problems. In the seminal work of Donoho (1995), wavelet-vaguelette decomposition (WVD) is proposed to solve linear inverse problems in the presence of white noise. WVD can achieve the goal of simultaneous quasi-diagonalization of the operator K and a prior regularity on f. Unlike singular value decomposition, WVD not only extracts information of the underlying function f from indirect observations but also enjoys the localized property. Wang (1995) used these two properties to study change-points for indirect data and Wang (1996) employed WVD to decorrelate noise with long-range dependence and investigate function estimation for data with long-range dependence. Slow decay in correlation between far apart observations is referred to as long-range dependence or long-memory [Beran (1992 1994)]. Long-range dependence often occurs in fields such as agronomy, astronomy, chemistry, economics, engineering, environmental sciences, geosciences, hydrology, physics, and signal and image processing [Beran (1992), Keshner (1982) and Wornell and Oppenheim (1992)]. In this paper we use WVD to study minimax estimation of $f(x)$ for linear inverse problems in the presence of long-range dependent noise.

As fractional Gaussian noise provides a useful model for phenomenon exhibiting long-range dependence, we model linear inverse problems with long-range dependence by a fractional Gaussian noise model and then investigate function estimation by data observed from the model. We establish asymptotics for minimax risks. The minimax risk and the minimax linear risk converge to zero at rates that depend on both transformation and long-range dependence, and thus differ from those for problems with either direct observations or indirect observations with independence or short-range dependence. We prove that the wavelet shrinkage estimates can be tuned to achieve the minimax rate over a wide range of spaces and significantly outperform any linear estimates.

The approach is to employ two WVDs, one for the inverse problem which simultaneously quasi-diagonalizes both the operator and the prior information and one for long-range dependence which decorrelates fractional Gaussian noise, to convert the linear inverse problem into a new inverse problem with white noise. The transformation operator of the new inverse problem is the composition of the operator K and a fractional differential operator, and has a WVD which can be determined by the fractional differential operator and the WVD for K. Therefore, we can solve the new inverse problem and establish asymptotics for minimax risks by the method developed in Donoho (1995).

The rest of this paper is organized as follows. Section 2 introduces a fractional Gaussian noise model and Section 3 reviews WVD. Section 4 presents the wavelet shrinkage estimate of f and derives asymptotics of minimax risks. Proofs are given in Section 5.

2 Fractional Gaussian Noise Model

Suppose we have observations from fractional Gaussian noise model

$$Y(dx) = (Kf)(x)dx + \epsilon^{2-2H}\, B_H(dx), \qquad x \in I\!R, \tag{1}$$

where f is an unknown function with support contained in $[0,1]$, K is a linear transformation, not necessarily bounded, from $L^2(I\!R)$ to $L^2(I\!R)$ [see examples in Section 3], ϵ is noise level, and $B_H(dx)$ is fractional Gaussian noise. [Below we describe fractional Gaussian noise and discuss approximation of discrete model by model (1).] Our goal is to estimate f on the basis of the data $Y(x)$ and a priori information that $f \in \mathcal{F}$, a class of functions. We use squared-error loss and are interested in the minimax risk

$$R(\epsilon; \mathcal{F}) = \inf_{\hat{f}} \sup_{f \in \mathcal{F}} E\|\hat{f} - f\|^2_{L^2[0,1]}.$$

Fractional Gaussian noise is the generalized derivative of standard fractional Brownian motion $B_H(x)$ with parameter H $(1/2 \le H < 1)$, where B_H is a Gaussian process with covariance function $cov(B_H(s), B_H(t)) = V_H \{|s|^{2H} + |t|^{2H} - |t-s|^{2H}\}/2$, $V_H = var\{B_H(1)\} = cos(\pi H)\,\Gamma(1-2H)/(\pi H)$ [see Benassi and Jaffard (1993), Mandelbrot (1982), and Mandelbrot and Van Ness (1968)].

Long-range dependence (or long-memory) is referred to as that correlations between observations that are far apart decay to zero at a slow rate, and fractional Gaussian noise and fractional Brownian motion are often used to model phenomenon exhibiting long-range dependence [Beran (1992 1994) and Mandelbrot and Van Ness (1968)]. As white noise approximates models with i.i.d errors [Brown and Low (1996), Donoho (1995) and Donoho and Johnstone (1996)], we can show that model (1) is an approximation of the following discrete inverse problem with long-range dependence

$$y_i = (Kf)(x_i) + \varepsilon_i, \qquad i = 1, \cdots, n, \tag{2}$$

where $x_i = i/n$, and random errors ε_i are zero mean stationary normal random variables with long-range dependence. In deed, according to definition 2.1 in Beran (1994, p. 42), we have $corr(\varepsilon_i, \varepsilon_j) \asymp |i-j|^{-\alpha}$, for some $0 < \alpha < 1$. Let $L_n^2 = \sum_{i=1}^n \sum_{j=1}^n E(\varepsilon_i \varepsilon_j)$. Then $L_n \asymp n^H$, as $n \to \infty$, with $H = 1 - \alpha/2$. By theorem 5.1 in Taqqu (1975) we obtain that $\sum_{i=1}^{[n\,x]} \varepsilon_i / L_n$ can be approximated by $B_H(x)$. Define the process $Y_n(x)$ at the point $x_i = i/n$ by $Y_n(x_i) = \sum_{j=1}^i y_j/n$, and interpolate between the points $x_i = i/n$ using B_H. Then on $[0,1]$ Y_n satisfies (1) with $\epsilon_n \asymp \{var(\varepsilon)/n\}^{1/2}$ and $(K_n f)(x) = \sum_{i=1}^n (Kf)(x_i) 1_{\{x_{i-1} \le x < x_i\}}$, and thus minimax results derived in Section 4 for model (1) can formally be applied to model (2) [see also Donoho (1995, section 6.3)].

3 Wavelets and WVD

Wavelets. Let ψ be a Daubechies mother wavelet of compact support, having M vanishing moments and M continuous derivatives and unit L^2 norm. Wavelets are derived from ψ by a process of dyadic dilations and translations of ψ. Let $\lambda = (j,k)$ with j and k integers,

$$\psi_\lambda(x) = 2^{j/2} \psi(2^j x - k), \qquad j, k \in \mathbb{Z}.$$

Daubechies' construction guarantees that $(\psi_\lambda)_\lambda$ furnishes the space $L^2(I\!R)$ with a complete orthonormal basis. Hence, for a function $f \in L^2(I\!R)$, we have the wavelet decomposition

$$f(x) = \sum_\lambda \theta_\lambda \psi_\lambda(x), \tag{3}$$

with wavelet coefficients of f

$$\theta_\lambda = \int f(x)\psi_\lambda(x)dx. \tag{4}$$

[See Daubechies (1992).]

For f supported in $[0,1]$, we select those ψ_λ, $\lambda \in \Lambda$, whose supports either included in $[0,1]$ or intersect with $[0,1]$ and have

$$f(x) = \sum_{\lambda\in\Lambda} \theta_\lambda \psi_\lambda(x).$$

[See Donoho (1995, section 5.3)].

Vaguelettes. Vaguelettes u_λ are formally a wavelet like function system which is obtained by a process of dyadic dilations and translations of a mother function u. In general neither the mother function u has all the special properties of mother wavelet ψ nor vaguelettes u_λ are orthogonal. [See Meyer (1990) and Donoho (1995).]

WVD. For a class of linear inverse problems, Donoho (1995) developed a WVD to simultaneously quasi-diagonalize both an operator K and a priori information $f \in \mathcal{F}$ such that f can be recovered through a reproducing formula by use of information about Kf. A WVD is defined as follows.

(1) Three sets of functions $(\psi_\lambda)_\lambda$ – an orthogonal wavelet basis – and $(u_\lambda)_\lambda$ and $(v_\lambda)_\lambda$ – near-orthogonal vaguelettes.

(2). Quasi-singular value relations $K\psi_\lambda = \kappa_j v_\lambda$ and $K^* u_\lambda = \kappa_j \psi_\lambda$ for $\lambda = (j,k)$ with quasi-singular values (κ_j), depending on resolution index j but not spatial index k.

(3). Biorthogonality relations $[u_\lambda, v_\mu] = \delta_{\lambda,\mu}$.

(4). Near-Orthogonality relations

$$\|\sum_\lambda a_\lambda u_\lambda\|_2 \asymp ||(a_\lambda)||_{\ell^2}, \qquad \|\sum_\lambda a_\lambda v_\lambda\|_2 \asymp ||(a_\lambda)||_{\ell^2}.$$

From the decomposition we can easily derive the following reproducing formula

$$f = \sum_\lambda [Kf, u_\lambda]\, \kappa_j^{-1}\, \psi_\lambda. \tag{5}$$

Examples. Donoho (1995) explicitly constructed WVD for commonly used transformations such as integration, fractional integration, deconvolution operator and Radon Transform.

Example 1. Integration. Let $(Kf)(u) = \int_{-\infty}^{u} f(t)dt$. K has domain $\{f : \int |\omega|^{-2}\, |\tilde{f}(\omega)|^2\, d\omega < \infty\}$ and range $\{f : \int |\omega|^2\, |\tilde{f}(\omega)|^2\, d\omega < \infty\}$, where $\tilde{f}$ is Fourier transformation of f. It has the WVD with $\kappa_j = 2^{-j}$, $u_\lambda = \kappa_j \psi'_\lambda = 2^{j/2}\, \psi'(2^j x - k)$ and $v_\lambda = \kappa_j^{-1} K \psi_\lambda$.

Example 2. Fractional Integration. Let $\Omega(\cdot)$ be a not identically vanishing function, homogeneous of degree 0, let $\alpha \in (0,1)$, and set

$$(Kf)(u) = \int f(t) \frac{\Omega(t-u)}{|t-u|^{1-\alpha}} dt.$$

If $\Omega =$ the Heaviside function and $\alpha = 1/2$, then K is Abel transform.

K has the WVD with $\kappa_j = 2^{-j\alpha}$, $u_\lambda = \kappa_j \gamma_\lambda$ and $v_\lambda = \kappa_j^{-1} \xi_\lambda$, where $\tilde{\gamma}_\lambda(\omega) = \tilde{\psi}_\lambda(\omega)\, |\omega|^\alpha / \tilde{\Omega}(\omega)$ and $\xi_\lambda = \tilde{\psi}_\lambda(\omega)\, |\omega|^{-\alpha}\, \tilde{\Omega}(\omega)$.

Example 3. Convolution. $(Kf)(u) = \int k(u-t) f(t) dt$ with the kernel k obeying $|\tilde{k}(\omega)| \asymp |\omega|^{-\alpha}$ as $|\omega| \to \infty$. For example, if $k(x) = e^x 1_{\{x<0\}}$, then $\kappa_j = \min(1, 2^{-j})$, $u_\lambda = \kappa_j(\psi_\lambda - \psi'_\lambda)$ and $v_\lambda = \kappa_j^{-1} K \psi_\lambda$; if $k(x) = \frac{1}{2} e^{-|x|}$, then $\kappa_j = \min(1, 2^{-2j})$, $u_\lambda = \kappa_j(\psi_\lambda - \psi''_\lambda)$ and $v_\lambda = \kappa_j^{-1} K \psi_\lambda$.

4 Main Results

For observations $Y(x)$ from model (1), the reproducing formula (5) of WVD leads to a simple wavelet shrinkage estimate of f

$$\hat{f}_{(t_j)}(x) = \sum_\lambda \delta_{t_j}\left([Y, u_\lambda] \kappa_j^{-1}\right)\, \psi_\lambda(x),$$

where t_j are level-dependent thresholds, $\delta_t(y) = sgn(y)\,(|y| - t)_+$ for soft threshold rule and $\delta_t(y) = y\,1(|y| \geq t)$ for hard threshold rule.

Denote by $B^\sigma_{p,q}$ the Besov space with index (σ, p, q), which is the set of all functions $f : I\!R \to I\!R$ with $f \in L^p(I\!R)$ and $\|f\|_{B^\sigma_{p,q}} < \infty$, where $\|f\|_{B^\sigma_{p,q}}$ is Besov seminorm of index (σ, p, q) and defined for $r > \sigma$ by

$$\|f\|_{B^\sigma_{p,q}} = \left(\int_0^1 \left(\frac{w_{r,p}(f;h)}{h^\sigma} \right)^q \frac{dh}{h} \right)^{1/q} \quad \text{if } q < \infty, \quad \text{and} \quad \|f\|_{B^\sigma_{p,\infty}} = \sup_{0<h<1} \frac{w_{r,p}(f;h)}{h^\sigma},$$

and

$$w_{r,p}(f;h) = \|\Delta_h^{(r)} f\|_{L^p}, \qquad \Delta_h^{(r)} f = \sum_{k=0}^{r} \binom{r}{k} (-1)^k f(t + kh).$$

$\|f\|_{B^\sigma_{p,q}}$ is equivalent to

$$\left(\sum_j 2^{jsq} \left(\sum_k |\theta_{j,k}|^p \right)^{q/p} \right)^{1/q},$$

where $\theta_{j,k}$ are wavelet coefficients of f and $s = \sigma + 1/2 - 1/p$, and thus Besov space can be characterized through wavelet expansions [see Donoho and Johnstone (1996, section 2), Meyer (1990, chapter 6) and Triebel (1992, chapter 2)].

The Besov measure of smoothness includes commonly used Hölder and Soblev measures. For example, a function f satisfying Hölder condition $|f^{(m)}(s) - f^{(m)}(t)| \leq c\,|s - t|^\delta$ for some $c > 0$, and $0 < \delta < 1$ is equivalent to $\|f\|_{B^{m+\delta}_{\infty,\infty}} < \infty$; f with $\int |f^{(m)}(x)|^2 dx < \infty$, if and only if $\|f\|_{B^m_{2,2}} < \infty$ [see Donoho and Johnstone (1996, section 2) and Meyer (1990, chapter 6)].

Let $\mathcal{F} = \{f : \|f\|_{B^\sigma_{p,q}} \leq C\}$, a ball of Besov space $B^\sigma_{p,q}$. We define the minimax-wavelet threshold risk of the wavelet shrinkage estimate

$$R_W(\epsilon; \mathcal{F}) = \inf_{(t_j)} \sup_{\mathcal{F}} E\|\hat{f}_{(t_j)} - f\|^2_{L^2[0,1]},$$

and the minimax linear risk

$$R_L(\epsilon; \mathcal{F}) = \inf_{\hat{f}\ \text{linear}} \sup_{\mathcal{F}} E\|\hat{f} - f\|^2_{L^2[0,1]}.$$

Theorem 1 *Let $(\psi_\lambda, u_\lambda, v_\lambda)_\lambda$ be the WVD for K. Suppose that the quasi-singular-values $\kappa_j = 2^{-j\,\alpha}$ and $(u_\lambda)_\lambda$ satisfies the condition that $\{(-\Delta)^{1/4-H/2} u_\lambda\}_\lambda$ (after normalized) is near orthogonal, and assume $\sigma > \frac{2-p}{p}(\alpha+1-H) + 1/p - 1/2$. Then*

$$R(\epsilon; \mathcal{F}) \asymp \epsilon^{2r}, \qquad \text{as } \epsilon \to 0,$$

with rate exponent

$$r = \frac{2\sigma(1-H)}{\sigma+1+\alpha-H};$$

and

$$R_W(\epsilon; \mathcal{F}) \leq \text{constant} \cdot R(\epsilon; \mathcal{F}), \qquad \text{as } \epsilon \to 0.$$

However,

$$R_L(\epsilon; \mathcal{F}) \asymp \epsilon^{2r'}, \qquad \text{as } \epsilon \to 0,$$

with the rate

$$r' = \frac{(2\sigma+1-2/p_-)(1-H)}{\sigma+3/2-1/p_-+\alpha-H},$$

where $p_- = \min(2, p)$.

Remark 1. From Theorem 1 we see that the wavelet shrinkage estimate can be tuned to achieve the minimax convergence rate. For $p < 2$, $r > r'$, so linear estimates can not achieve even the minimax rate, and thus the wavelet shrinkage estimate can significantly outperform any linear estimates.

Remark 2. As exponents r and r' decrease in both H and α, the convergence rates depend on both transformation and long-range dependence. The $H = 1/2$ case corresponds to linear inverse problems with white noise considered in Donoho (1995) while nonparametric regression with long-range dependence considered in Wang (1996) matches the case that $K = I$ and $H > 1/2$. Also the results can recover convergence rates derived in Donoho (1995) and Wang (1996).

Remark 3. The near orthogonality condition for $\{(-\Delta)^{1/4-H/2} u_\lambda\}_\lambda$ can be heuristically explained as follows: fractional Gaussian noise can be decorrelated through a WVD for $(-\Delta)^{H/2-1/4}$ [see (13) in Section 5]. By decorrelating fractional Gaussian noise we can convert the linear inverse problem into an inverse problem with white noise. The transformation for the new inverse problem is the composition of K and $(-\Delta)^{H/2-1/4}$. The near orthogonality condition is to insure that the composition operator has a WVD.

Remark 4. The near orthogonality condition is quite reasonable. Transformations considered in Section 3 meet the condition. For example, for the integration transformation, $u_\lambda(x) = 2^{j/2}\,\psi'(2^j x - k)$. The Fourier transformation of $(-\Delta)^{1/4-H/2}u_{0,0}$ is $\hat{\psi}(\omega)\,|\omega|^{3/2-H}$. Therefore, by definition 1 and lemma 3 (or lemma 4) in Donoho (1995, section 4), we have that $\{(-\Delta)^{1/4-H/2}u_\lambda\}_\lambda$ (after normalized with constants of orders $2^{j(1/2-H)}$) form vaguelettes, and hence the near orthogonality is just a consequence of theorem 2 in Donoho (1995, section 4).

5 Proofs

Define

$$z_\lambda \;=\; \left(Var\left\{\int u_\lambda(s)dB_H(s)\right\}\right)^{-1/2}\int u_\lambda(s)dB_H(s). \tag{6}$$

Note that

$$\begin{aligned}
Var\left\{\int u_\lambda(s)dB_H(s)\right\} &= E\left\{\int u_\lambda(s)dB_H(s)\int u_\lambda(t)dB_H(t)\right\}\\
&= E\left\{\int\int u_\lambda(s)u_\lambda(t)dB_H(s)dB_H(t)\right\}\\
&= V_H\,H\,(2H-1)\int\int u_\lambda(s)u_\lambda(t)|s-t|^{2H-2}dsdt\\
&= \left\{V_H\,H\,(2H-1)\int\int u(x)u(y)|x-y|^{2H-2}dxdy\right\}\;2^{j(1-2H)}.
\end{aligned}$$

and hence

$$z_\lambda = \tau^{-1}\,2^{j\,(H-1/2)}\int u_\lambda(s)dB_H(s) \tag{7}$$

with

$$\tau = \left\{V_H\,H\,(2H-1)\int\int u(s)u(t)|s-t|^{2H-2}dsdt\right\}^{1/2}. \tag{8}$$

Let $y_\lambda = [Y, u_\lambda]/\kappa_j$, $\lambda \in \Lambda$. From (1), (3), (4), (5), (6), (7) and (8) we obtain

$$y_\lambda = \theta_\lambda + \epsilon^{2-2H}\, \tau_j\, \sigma_j\, z_\lambda, \qquad \lambda \in \Lambda, \tag{9}$$

where θ_λ is the wavelet coefficient of f defined in (4), $\sigma_j = \kappa_j^{-1}$, $\tau_j = \tau\, 2^{j\,(1/2-H)}$, and the noise process (z_λ) is defined in (6). We will show that (z_λ) is nearly-independent noise, that is,

$$\gamma_0^2 \le Var(z_\lambda | (z_\lambda)_{\lambda \ne \mu}) \le Var(z_\lambda) \le \gamma_1^2, \tag{10}$$

for each $\lambda \in \Lambda$, where $0 < \gamma_0 < \gamma_1 < \infty$ are positive generic constants.

Model (1) is equivalent to model (9) and the problem of estimating f is equivalent to recovering θ_λ, because wavelets provide a complete orthonormal system and $\|\hat{f} - f\|^2_{L^2[0,1]}$ is essentially the same as $\|\hat{\theta} - \theta\|^2_{\ell^2(\Lambda)}$. Let $\Theta^\sigma_{p,q}$ be the set of all wavelet coefficients of f in $\mathcal{F}$. For sequence observations from model (9), we define the minimax risk

$$R_z(\epsilon) = \inf_{\hat{\theta}} \sup_{\Theta^\sigma_{p,q}} E\|\hat{\theta}(y) - \theta\|^2_{\ell^2(\Lambda)}.$$

For observations from model (9) with near-independent noise, Donoho (1995, section 7.3) showed that the minimax risk $R_z(\epsilon)$ is equivalent to the minimax risk $R_w(\epsilon)$ for observations from the discrete white noise model

$$x_\lambda = \theta_\lambda + \epsilon^{2-2H}\, \tau_j\, \sigma_j\, w_\lambda, \qquad \lambda \in \Lambda \tag{11}$$

where (w_λ) is a Gaussian white noise, and

$$R_w(\epsilon) = \inf_{\hat{\theta}} \sup_{\Theta^\sigma_{p,q}} E\|\hat{\theta}(x) - \theta\|^2_{\ell^2(\Lambda)}.$$

For the white noise model (11), $R_w(\epsilon)$ can be explicitly given by the renormalization arguments in Section 7.4 - 7.6 of Donoho (1995). Also the wavelet minimax risk and the minimax linear risk can be derived by the methods developed in Section 8 and Section 9 in Donoho (1995).

Now we prove (10) using the following WVD decorrelation of fractional fractional Gaussian noise. Fractional Brownian motion B_H is linked to the operator $(-\Delta)^{H+1/2}$.

Let ψ_λ be a usual wavelet orthonormal basis of $L^2(I\!R)$, and

$$\bar{u}_\lambda = (-\Delta)^{H/2-1/4}\,\psi_\lambda, \qquad \bar{v}_\lambda = (-\Delta)^{1/4-H/2}\,\psi_\lambda. \tag{12}$$

By theorem 2 in Wang (1996, section 5), we have that $(\psi_\lambda)_\lambda$, $\left(2^{j_\lambda(1/2-H)}\,\bar{u}_\lambda\right)_\lambda$ and $\left(2^{j_\lambda(H-1/2)}\,\bar{v}_\lambda\right)_\lambda$ form a WVD of $(-\Delta)^{H/2-1/4}$, and hence the WVD decorrelates fractional Gaussian noise $B_H(dx)$, that is, $B_H(dx)$ has a representation

$$B_H(dx) = \sum_\lambda \xi_\lambda\,\tilde{v}_\lambda(x)\,dx, \tag{13}$$

where (ξ_λ) is a white noise.

The upper bound γ_1 in (10) can be taken as $\gamma_1 = Var(z_\lambda) = 1$. To derive the lower bound, we let $\hat{z}_\lambda = E(z_\lambda|(z_\mu)_{\lambda\neq\mu})$. Then $Var(z_\lambda|(z_\mu)_{\lambda\neq\mu}) = Var(z_\lambda - \hat{z}_\lambda)$. As $(z_\lambda)_{\lambda\in\Lambda}$ is a zero-mean Gaussian random field, $\hat{z}_\lambda$ and $z_\lambda - \hat{z}_\lambda$ are independent, and there exists a sequence $(a_\mu)_{\mu\neq\lambda}$ such that

$$\hat{z}_\lambda = \sum_{\mu\neq\lambda} a_\mu\,z_\mu. \tag{14}$$

For convenience, let $a_\lambda = -1$ and for $\lambda = (j_\lambda, k_\lambda)$, let $\tau_\lambda = \tau 2^{j_\lambda\,(H-1/2)}$ with τ defined in (8).

On the other hand, by (13) we have for any μ

$$z_\mu = \tau_\mu\,[B_H(dx), u_\mu] = \tau_\mu\,\sum_{\lambda'} [\bar{v}_{\lambda'}, u_\mu]\,\xi_{\lambda'}. \tag{15}$$

By (14) and (15), we obtain

$$\begin{aligned} z_\lambda - \hat{z}_\lambda &= z_\lambda - \sum_{\mu\neq\lambda} a_\mu\,z_\mu \\ &= \tau_\lambda\,\sum_{\lambda'} [\bar{v}_{\lambda'}, u_\lambda]\,\xi_{\lambda'} - \sum_{\mu\neq\lambda} \tau_\mu\,a_\mu\,\sum_{\lambda'} [\bar{v}_{\lambda'}, u_\mu]\,\xi_{\lambda'} \\ &= \sum_{\lambda'} [\bar{v}_{\lambda'}, \tau_\lambda\,u_\lambda - \sum_{\mu\neq\lambda} \tau_\mu\,a_\mu u_\mu]\,\xi_{\lambda'} \\ &= -\sum_{\lambda'} [\bar{v}_{\lambda'}, \sum_\mu \tau_\mu\,a_\mu u_\mu]\,\xi_{\lambda'}. \end{aligned} \tag{16}$$

However,

$$[\bar{v}_{\lambda'}, u_\mu] = [(-\Delta)^{1/4-H/2}\psi_{\lambda'}, u_\mu] = [\psi_{\lambda'}, (-\Delta)^{1/4-H/2}u_\mu]. \tag{17}$$

Let

$$u_\mu^* = (-\Delta)^{1/4-H/2} u_\mu. \tag{18}$$

Note that $u_\lambda(x) = 2^{j/2} u_{0,0}(2^j x - k)$ and $u_{0,0}$ is bounded in L^2-norm. Then the L^2-norm of u_μ^* has the order τ_μ. Hence, τ_μ normalizes u_μ^* and by the assumption $(\tau_\mu\, u_\mu^*)_\mu$ is near orthogonal.

By (16), (17) and (18), we get

$$z_\lambda - \hat{z}_\lambda \;=\; -\sum_{\lambda'} [\psi_{\lambda'}, \sum_\mu \tau_\mu\, a_\mu u_\mu^*]\, \xi_{\lambda'}$$

and hence by complete orthonormal bases of wavelets, the near orthogonality of $(\tau_\mu\, u_\mu^*)$, and $\|(\alpha_\mu)\|_{\ell^2} \geq 1$, we obtain

$$\begin{aligned} Var(z_\lambda - \hat{z}_\lambda) &= \sum_{\lambda'} [\psi_{\lambda'}, \sum_\mu \tau_\mu\, a_\mu\, u_\mu^*]^2 \\ &= \|\sum_\mu \tau_\mu\, a_\mu\, u_\mu^*\|^2 \\ &= \|\sum_\mu (a_\mu)\, (\tau_\mu\, u_\mu^*)\|^2 \\ &\geq C_0\, \|(\alpha_\mu)\|_{\ell^2} \\ &\geq C_0 = \gamma_0 > 0. \end{aligned}$$

This proves the lower bound in (10).

Acknowledgements. The author would like to thank Iain Johnstone and two anonymous referees for suggestions. The research is partially supported by NSF Grant DMS 94-04142.

REFERENCES

Benassi, A. and Jaffard, S. (1993). Wavelet decomposition of one and several dimensional Gaussian processes. In Recent Advances in Wavelet Analysis (L. L. Schumaker and G. Webb, eds.) 119-154. Academic Press, New York.

Beran, J. (1992). Long-range dependence (with discussion). Statistical Science **7**, 404-427.

Beran, J. (1994). Statistics for Long Memory Processes. Chapman and Hall, New York.

Brown, L. and Low, M. (1996). Asymptotic equivalence of nonparametric regression and white noise. Ann. Statist. To appear.

Daubechies, I. (1992). Ten Lecture on Wavelets. Society for Industrial and Applied Mathematics, Philadelphia.

Donoho, D. L. (1995). Nonlinear solutions of linear inverse problems by wavelet-vaguelette decomposition. Appl. Comput. Harm. Anal. **2**, 101-126.

Donoho, D. L. and Johnstone, I. M. (1996). Minimax estimation via wavelet shrinkage. Ann. Statist. To appear.

Keshner, M. S. (1982). $1/f$ noise. Proc. IEEE **70**, 212-218.

Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. SIAM Review **10**, 422-437.

Mandelbrot, B. B. (1982). Fractional Brownian motions and fractional Gaussian noises. Encyclopedia of Statistical Sciences (vol. 3, S. Kotz and N. L. Johnson, eds.) 186-189. Wiley, New York.

Meyer, Y. (1990). Ondelettes et Opérateurs. I and II. Paris: Hermann (in French). (Wavelets and Operators. University Press, Cambridge, English translation, 1992.)

Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. Z. W. **31**, 287-302.

Triebel, H. (1992). Theory of Function Spaces. II. Boston: Birkhäuser Verlag.

Wang, Y. (1996). Change point analysis via wavelets for indirect data. Technical report, Dept. Statist., Univ. of Missouri-Columbia.

Wang, Y. (1996). Function estimation via wavelet shrinkage for long-memory data. Ann. Statist. **24**, 466-484.

Wornell, G. W and Oppenheim, A. V. (1992). Estimation of fractal signals from noisy measurements using wavelets. IEEE Trans. Signal Proc. **40**, 611-623.