# Maximum Likelihood Estimation of the Parameters of Fractional Brownian Motions.

Brett Ninness *
Centre for Industrial Control Science at Department of Electrical Engineering
University of Newcastle
Callaghan 2308, Australia.

## Abstract

This paper provides convergence analysis for maximum likelihood estimation of the parameters describing a particular multi-scale process known as fractional Brownian motion. We concentrate on schemes that 'pre-whiten' the available noise corrupted data via the Fast Wavelet Transform and show that such schemes are strongly consistent and asymptotically efficient. We also analyse the rate of convergence of the maximum likelihood estimates and show that this rate depends on the memory of the fractional process.

## 1 Introduction

Recently there has been significant interest in so-called 'multi-scale' stochastic processes [3, 2, 17, 16]. In this paper we concentrate on the modelling of particular examples of multi-scale processes which are more colloquially known as '1/f processes'. For a long time these processes have been important to physicists and the signal processing community [14].

A difficulty with these 1/f processes is that, as their name implies, they have hyperbolic non-integrable spectra. This means that they cannot be modelled as a stationary stochastic process. However, due to their unusual distribution of energy with frequency, realisations of their sample paths are fractal in nature. One could therefore model them via a deterministic but non-linear system with fractal attractor. Benoit Mandelbrot and John Van Ness [21] have put forward an alternative stochastic embedding description which is a fractional integral of the increments of a classical Wiener process.

This latter description is called fractional Brownian motion. In [31, 30] Wornell and Oppenheim have proposed estimating the parameterisation of a fractional Brownian motion by maximising the likelihood of the wavelet transform of the available data. However, apart from a Monte-Carlo simulation study, Wornell and Oppenheim did not study the convergence properties of the estimators they proposed.

The purpose of this paper is to prove that subject to regularity conditions, a normalisation of the estimator proposed by Wornell and Oppenheim is strongly consistent, asymptotic Normal and asymptotically efficient. Using these results and an evaluation of the Cramér-Rao lower bound we gauge the convergence rate of the maximum likelihood estimator and show that it slows as the hyperbolicity of the spectra increases. This is in accordance with the simulation results presented by Wornell and Oppenheim [31, 30] and others [18].

## 2 Fractional Brownian Motions

Noise referred to as '1/f' noise is a stochastic process $\{x_k\}$ whose sample spectral density, or periodogram[1], $|\widehat{x}_N(\omega)|^2$ is of the form $|\widehat{x}_N(\omega)|^2 \approx \sigma_x^2|\omega|^{-\gamma}$ for some finite non-zero $\sigma_x$ and $\gamma$. Such noise has been found to occur in a wide variety of physical processes such as [14] Voltages and currents in semiconductors, Resistances in electronic components, Frequencies of quartz crystal oscillators, and Geophysical records.

This 1/f noise is therefore of great interest in communications systems; see for example [1] for a discussion of the signal detection problem in 1/f noise. These noises also occur in situations where accurate modelling of them is important for controlling a process. Such examples include [14] Rate of insulin uptake by diabetics, Economic data, Rate of Traffic flow, and more recently is has also proved important in the control of critically ill patients via 1/f modelling of heart rate [10].

There are some key features of the 1/f noise representation which make it very interesting, but also difficult to model. For example, for $\gamma \in (0, 1)$, no stationary process can be associated with a $f^{-\gamma}$ spectral density. Intuitively, apart from the non-integrability of the spectrum, this is so since the spectral density diverges to $\infty$ at low frequencies which implies that a covariance function, if it existed, would have to be very smooth. But $f^{-\gamma}$ for $\gamma \in (0, 1)$ decays quite slowly with increasing f. This implies a fast decay of the associated covariance function, which is incompatible with a smoothness requirement.

This leads to a consideration of non-stationary processes, for which a spectral density is not readily defined.

*This author gratefully acknowledges support from the Australian Research Council. The author can be contacted at FAX +61 49 21 69 93 or email:brett@tesla.newcastle.edu.au.

[1] The absolute value of the N point DFT $\widehat{x}_N(\omega)$ of $\{x_k\}$.

suggested considering a definition of spectra conditional upon the observation time. Solo [24] has suggested using the expected value of the sample periodogram. This, and other interpretations have been formalised by calling on formulations arising from fractional integrals with respect to the increments of ordinary Brownian motion. This latter formulation is called fractional Brownian motion (fBm).

To be more specific, for the case of $\gamma = 0$ the attendant difficulties of defining a process associated with a constant spectrum are traditionally handled by defining the classical non-stationary Brownian motion process $B(t)$ and then considering its increments $dB$ which lead to the required spectra. That is

$$B(t) = \sigma_B^2 \int_0^t dB(\sigma) \tag{1}$$

where the increments are a stationary process with a white spectrum. Therefore, speaking very loosely for the purposes of motivation, we consider the (formal) derivative of $B(t)$ to have a spectral representation

$$\frac{dB}{dt} = \int_{-\infty}^{\infty} e^{j\omega t} d\mu(\omega) \tag{2}$$

where the measure $d\mu$ satisfies $E\{|d\mu(\omega)|^2\} = \sigma_B^2 d\omega^2, \sigma_B^2 < \infty$ in which case $B(t)$ should, at least intuitively since it is the integral of $\dot{B}(t)$, have a spectrum like $\sigma_B^2 \omega^{-2}$. Following this heuristic line of reasoning, if we integrate again

$$B_2(t) = \sigma_B^2 \int_0^t \int_0^{\xi} dB(\sigma)\, d\xi \tag{3}$$

then $B_2(t)$, being a double integral of $\dot{B}$, should have a spectrum like $\sigma_B^2 \omega^{-4}$ and so on so that using Liouville's formula [13]

$$B_n(t) = \sigma_B^2 \int_0^t \int_0^{t_n} \cdots \int_0^{t_2} dB(\sigma)\, dt_2 \cdots dt_n = \frac{\sigma_B^2}{\Gamma(n)} \int_0^t (t-\xi)^n \, dB(\xi) \tag{4}$$

should have a spectrum like $\sigma_B^2 \omega^{-2n}$. However, we are interested in fractional $n$ since we are interested in spectra like $\sigma_B^2 |\omega|^{-\gamma}$ with $\gamma \in (0,1)$. This suggests, at least heuristically, that we should consider a fractional Brownian motion $B_H(t)$ and associated fractional Gaussian noise $\dot{B}_H(t)$ derived from the right hand side of (4) as

$$B_H(t) = \frac{\sigma_B^2}{\Gamma(H+1/2)} \int_0^t (t-\xi)^{H-1/2} dB(\xi) \tag{5}$$

where $H \in (1/2, 1)$. The reason for the offsetting of the index by 1/2 is for compatibility with accepted notation [21, 1]. The definition in (5) is in terms of what has become known as a 'fractional integral' and around which a complete theory of calculus has arisen [23].

Unfortunately, this heuristic derivation falls short in that the increments of $B_H$ are not stationary [21] making rectified, as Mandelbrot and Van Ness discuss [21] by modifying (5) to

$$B_H(t) = \frac{\sigma_B^2}{\Gamma(H+1/2)} \left\{ \int_0^t |t-\xi|^{H-1/2} dB(\xi) + \int_{-\infty}^0 \left( |t-\xi|^{H-1/2} - |\xi|^{H-1/2} \right) dB(\xi) \right\} \tag{6}$$

and this latter equation is what is now known as 'fractional Brownian motion'. A property of these fractional Brownian motions is that they are self similar in the sense that [8]

$$B_H(\lambda t) \overset{\mathcal{D}}{=} \lambda^H B_H(t)$$

where $\overset{\mathcal{D}}{=}$ denotes equality in distribution. This property of scale invariance makes them a particular example of the multi-scale stochastic processes recently studied in [3, 2, 17, 16].

To continue with a frequency domain interpretation of these processes, the increments of $B_H(t)$ as defined by (6) have the spectral representation [21]

$$B_H(t_2) - B_H(t_1) = \sigma_B^2 \int_0^{\infty} \frac{e^{j\omega t_2} - e^{j\omega t_1}}{\omega^{H+1/2}} dB(\omega) \tag{7}$$

and therefore, indeed, we can think of $\dot{B}_H(t)$ having spectral density like $\sigma_B^2 |\omega|^{2H-1}$ in the same way that we think of the white spectrum of the derivative of ordinary Brownian motion $(H = 1/2)$.

This again is intuitively reasonable but not completely clear since, like ordinary Brownian motion, the sample paths of $B_H(t)$ are almost surely non-differentiable. Flandrin [8, 7] has attempted to address the problem by using tools targeted at spectral analysis of non-stationary signals. He considers the Wigner-Ville and the Wavelet transform of $B_H(t)$ in order to infer the spectrum of $\dot{B}_H(t)$. We will concentrate in this paper on the Wavelet interpretations.

In [22] Masry showed that if one takes the Wavelet transform $(\mathcal{W}B_H)(a,t)$ of a fBm,

$$(\mathcal{W}B_H)(a,t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} B_H(\sigma) \psi\left(\frac{\sigma - t}{a}\right) d\sigma \tag{8}$$

then for $a$ fixed $\{(\mathcal{W}B_H)(a,t)\}$ is a stationary process with spectrum $a|\widehat{\psi}(a\omega)|^2 |\omega|^{2H+1}$ and since from (8) the transform is in fact $B_H$ 'filtered' by $\psi(t/a)$ this implies that over the support of $\widehat{\psi}(a\omega)$ we can interpret $B_H(t)$ as having a 'local spectral density' of $1/|\omega|^{2H+1}$ and its increments (fGn) as having a spectral density of $1/|\omega|^{2H-1}$. Now, $a$ can be made arbitrarily small, so that the 'local' interpretation can apply over very large frequency ranges, but never down to zero since $\widehat{\psi}(0) = 0$.

Apart from providing this useful spectral interpretation, the Wavelet transform is of independent interest since (as Mallat [19] points out) it suggests a means for estimating $H$ (and $\gamma$) from observed realisations of fBm sity, the variance of $(\mathcal{W}B_H)(a,t)$ obeys[2]

$$\sigma_a^2 \triangleq \mathrm{E}\left\{[(\mathcal{W}B_H)(a,t)]^2\right\} = a\int_{-\infty}^{\infty} \frac{|\widehat{\psi}(a\omega)|^2}{|\omega|^{2H+1}}\,d\omega \quad (9)$$

so that for $a' \neq a$

$$\sigma_{a'}^2 = a'\int_{-\infty}^{\infty} \frac{|\widehat{\psi}(a'\omega)|^2}{|\omega|^{2H+1}}\,d\omega = \left(\frac{a'}{a}\right)^{2H+1}\sigma_a^2 \quad (10)$$

then since $\{(\mathcal{W}B_H)(a,t)\}$ is stationary, we can take an estimate ($\Delta$ is a sampling interval)

$$\widehat{\sigma}_a^2 = \frac{1}{N}\sum_{k=0}^{N-1}[(\mathcal{W}B_H)(a,t+k\Delta)]^2 \quad (11)$$

for a range of $a = \{a_1, a_2, \cdots, a_M\}$ such that for some $\beta$ we have $a_{k+1} = \beta a_k$ and by the relationship (10) an intuitive estimate of H then is

$$\widehat{H} = \frac{\widehat{\gamma}+1}{-2} \quad (12)$$

where $\widehat{\gamma}$ is the slope of a graph of $\widehat{\sigma}_{a_k}$ versus $a_k$ with respect to $\log-\log_\beta$ axes. Of course, one could equivalently take the sample periodogram and take $\widehat{\gamma}$ in (12) to be its slope with respect to $\log-\log$ axes. This is illustrated in figure 1 for the case of no measurement noise corrupting the observations of the fBm, and in figure 2 for the case of noise corruption present. This estimation method, without the corrupting influence of measurement noise, is analyzed in [15, 24, 9]. Figures 1 and 2 illustrate the sensitivity to measurement noise of this simple linear regression approach.

To combat this, it is necessary to note a key feature of the Wavelet transform of a fBm. It approximately 'whitens' the non-stationary fractional process in the same way the the discrete Fourier transform whitens a stationary process. The whitening of fBm depends upon the regularity of the analysing wavelet $\psi(t)$ and has been analysed partially by Flandrin [7] for the Haar Wavelet, by Masry [22] in terms of spectral representations, and more completely by Dijkerman and Mazumder [6] and Tewfik and Kim [25]. These latter authors consider the dyadic situation which has become common in Wavelet analysis and corresponds to setting $a = 2$ in (8) in which case the following simpler and standard [5] notation can be adopted

$$c_k^n \triangleq (\mathcal{W}B_H)(2^{-n}, 2^{-n}k) \quad (13)$$

so that from (9) with $a_0 = 1$ we have

$$\mathrm{E}\{c_k^n c_\ell^n\} = \left[2^{-(2H+1)}\right]^n \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(\omega)|^2}{|\omega|^{2H+1}} e^{j(k-\ell)\omega}\,d\omega. \quad (14)$$

Masry [22] has pointed out that this formula illustrates that the $\{c_k^n\}$ cannot be perfectly uncorrelated as Wornell and Oppenheim [29, 31] assume[3] $\psi(t)$ is compactly supported on $[0,\xi]$ then one may achieve arbitrary uncorrelatedness by making $\psi(t)$ regular enough. More specifically, the smoothness of a wavelet $\psi(t)$ is generally measured by the number P of vanishing moments that exist

$$\int_{-\infty}^{\infty} t^p\psi(t)\,dt = 0 \quad p \in [0, P-1] \quad (15)$$

In which case, by Theorem 1 in [25] for some $C < \infty$

$$\mathrm{E}\{c_k^n c_\ell^m\} \leq \frac{C 2^{2(m+n)(P-H)}}{(2^n k - 2^m \ell)^{2(P-H)}} \quad (16)$$

provided that $|2^{-n}k - 2^{-m}\ell| > \max(2^{-n}\xi, 2^{-m}\xi)$. Therefore, for $P \gg 1$ the correlation decay along scales is rapid as the figures in [25] illustrate. These latter figures also show rapid decay across scales, with the rapidity again proportional to the number P of vanishing moments. Moreover, we can find compactly supported $\psi(t)$ with P as large as we wish by using (for example) the Wavelet construction techniques of Daubechies [5]. Recently Dijkerman and Mazumdar [6] have supported these 'whiteness' results by showing that the correlation of the Wavelet co-efficients of fBm decays exponentially fast across scales and hyperbollicaly fast along time.

This leads us back to the issue raised in (11)-(12) of estimating H. As well as the Wavelet approach proposed by Mallat [19], a direct maximum likelihood approach was suggested by Lundahl et. al. [18]. They found that slight measurement noise on top of a fGn process (as illustrated in figure 2) profoundly affected their ability to estimate H, especially for H near 1 (corresponding via $\gamma = 2H-1$ to $\gamma$ near 1). They suggested, therefore, that the measurement noise be included in the model and its variance be estimated. This has been investigated for the wavelet case by Oppenheim and Wornell who assume measurements are available of the form

$$r_n^m = c_n^m + \nu_n \quad (17)$$

where $\{\nu_k\}$ is i.i.d. zero mean and has variance $\mathrm{E}\{\nu_k^2\} = \sigma_\nu^2$. In view of the Gaussian nature of fBm the $\{c_n^m\}$ are zero mean Gaussian distributed and so by (14)

$$r_n^m \sim \mathcal{N}(0, \alpha\lambda^m + \sigma_\nu^2) = \mathcal{N}(0, \sigma_m^2(\theta)) \quad (18)$$

where $\theta^T = [\alpha, \lambda, \sigma_\nu^2]$ and $\lambda \triangleq 2^{-(2H+1)}$ for fractional Brownian motion and $\lambda \triangleq 2^{-(2H-1)}$ for fractional Gaussian noise. Also

$$\alpha \triangleq \int_{\infty}^{\infty} |\widehat{\psi}(\omega)|^2 |\omega|^{2H\pm 1}\,d\omega$$

and if we are modelling $f^{-\gamma}$ processes with $\gamma \in (0,1)$ then this implies $\lambda \in (1/2, 1)$. Finally we note that $\gamma$ and $\lambda$ are related by $\gamma = |\log_2\lambda|$.

# 3 Convergence Analysis for Embedding Parameter Estimates

We now wish to make this rigorous. Specifically, we examine the Maximum likelihood estimate that Wornell

[2]Modulo a scaling depending on the $L_2$ norm of $\psi(t)$, and also modulo $\sigma_B^2$ which affects the variance of the increments of $B_H(t)$.

[3]They, however, address $f^{-\gamma}$ noise in general citing empirical evidence for the uncorrelatedness; they do not seek motivation from fractional Brownian motion.

that $\{c_n^m\}$ to be uncorrelated and to be calculated via the fast wavelet transform. In this case, we need assume a data record of length $N_0 2^M$ for some $M$ in which case the negative log-likelihood function $\ell(R_M \mid \theta)$ for the data given the parameter vector $\theta$ is (modulo an additive constant):

$$\begin{aligned}\ell(R_M \mid \theta) &= -\ln\left(\prod_{m=1}^{M}\prod_{n=1}^{N_0 2^{m-1}} \frac{e^{-\frac{1}{2}(r_n^m/\sigma_m(\theta))^2}}{\sqrt{2\pi\sigma_m^2(\theta)}}\right) \\ &= \sum_{m=1}^{M}\sum_{n=1}^{N_0 2^{m-1}} \ln\sigma_m^2(\theta) + \left(\frac{r_n^m}{\sigma_m(\theta)}\right)^2\end{aligned}$$

Equivalently, we can use a formulation suggested by Wornell and Oppenheim of

$$\ell(Z_M \mid \theta) \triangleq \sum_{m=1}^{M} N_0 2^{m-1} \ln\sigma_m^2(\theta) + \frac{z_m^2}{\sigma_m^2(\theta)} \tag{19}$$

where

$$Z_M \triangleq \{z_1, \cdots, z_M\} \tag{20}$$

$$z_m^2 = \sum_{n=1}^{N_0 2^{m-1}} (r_n^m)^2 \tag{21}$$

Now, define the maximum likelihood estimate $\widehat{\theta}_M$ as

$$\widehat{\theta}_M = \arg\min_{\theta\in\Theta}\{Q_M(Z_M \mid \theta)\} \tag{22}$$

where

$$Q_M(Z_M \mid \theta) \triangleq \frac{1}{2^M}\ell(Z_M \mid \theta). \tag{23}$$

Wornell and Oppenheim [31] do not use the normalising factor of $2^{-M}$ in (23) and claim that '*It is well known that ML estimators are generally asymptotically efficient and consistent. This, specifically, turns out to be the case here*'. Presumably this comment is based on empirical evidence since in [31] Wornell and Oppenheim offer Monte-Carlo simulations, but no theoretical analysis.

If we review available work on convergence of maximum likelihood estimators in a general setting then we must begin with the seminal analysis of Wald [26] in 1947 which applies only to the stationary white noise case. Some years after Wald's work extensions to the stationary coloured noise case were made by Whittle [27] in 1953 and Hannan [11] in 1973. About the same time Hoadley [12] treated the non-stationary but independently distributed case. Very recently Caines [4] has considered the case where exogenous inputs satisfy Wiener's conditions [28] for generalised harmonic analysis.

In all these cases the consistency and efficiency properties mentioned by Wornell and Oppenheim are shown to hold. Unfortunately, none of these results apply to our problem since it is non-stationary, and also 'segmented' in an unusual way due to the multi-scale wavelet decomposition employed.

mators enjoy pleasant properties. As a simple example, if we consider the problem where we have observations $\{y_k\}$ of a scalar $\theta$ that is corrupted by additive uniformly distributed white noise $\{\nu_k\}$

$$y_k = \theta + \nu_k, \qquad \nu \sim U(-\delta, \delta)$$

then the probability density function for data $Y_N \triangleq \{y_0, \cdots, y_{N-1}\}$ conditional upon $\theta$ is of the form

$$p(Y_N \mid \theta) = \begin{cases} k & ;\theta : |y_k - \theta| \le \delta \ \forall k \in [0,1] \\ 0 & ;\theta : \exists k : |y_k - \theta| > \delta \end{cases}$$

where $k$ is some constant so that the maximum likelihood estimate of $\theta$ cannot even be uniquely defined, much less be consistent or distributed in some fashion. Because of examples such as these, the statistics literature is studded with articles examining the properties of ML estimators for specific probability distributions of the noise process.

Following these comments, the convergence of the maximum likelihood estimate (22) needs to be rigorously examined. We have done so, and have found the normalising factor of $2^M$ in the definition (23) and a regularity condition on $\Theta$ to be vital in order to obtain the asymptotic[4] consistency and efficiency that was empirically observed by Wornell and Oppenheim. To begin with, we have the following strong consistency result.

***Theorem 1.*** *Define $\theta_0$ to be a vector containing the true values of $[\alpha, \lambda, \sigma_\nu^2]$, then provided $\theta_0 \in \mathrm{int}\{\Theta\}$, $0 \notin \Theta$ and $\gamma \in (0,1)$*

$$\widehat{\theta}_M \xrightarrow{a.s} \theta_0 \quad \textit{as } M \to \infty. \tag{24}$$

In [31] Wornell and Oppenheim calculate the Cramér-Rao lower bound for the problem. Unfortunately, this only applies to unbiased estimates and Wornell and Oppenheim comment that (again presumably based on empirical evidence) this maximum likelihood scheme is biased. Nevertheless, the Cramér-Rao bound is still of interest since many maximum likelihood schemes have the property that their (possibly biased) estimates converge in distribution to a Gaussian with covariance equal to the Cramér-Rao bound. To be rigorous then, one needs to prove this convergence for the Cramér-Rao bound to be of relevance. This is done in the following theorem.

***Theorem 2.*** *Under the conditions of Theorem 1 and provided $\Theta$ is convex*

$$\sqrt{2^M} P_M^{1/2}\left(\widehat{\theta}_M - \theta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I) \quad \textit{as } M \to \infty \tag{25}$$

*where*

$$P_M = E\left\{\frac{d^2 Q_M(Z \mid \theta_0)}{d\theta d\theta^T}\right\}. \tag{26}$$

Therefore, Theorem 2 allows us to conclude that the maximum likelihood estimator (22) is indeed asymptotically efficient as Wornell and Oppenheim hoped. Furthermore, we can calculate $\mathcal{I}_M(\theta_0)$ as

[4] The normalising clearly has no effect on any finite data estimate.

$$\mathcal{I}_M(\theta_0) = \frac{N_0}{4}\sum_{m=1}^{M}\frac{2^m}{(\alpha\lambda^m+\sigma_v^2)^2}\times$$

$$\begin{bmatrix} \lambda^{2m} & (\alpha/\lambda)m\lambda^{2m} & \lambda^m \\ (\alpha/\lambda)m\lambda^{2m} & (\alpha/\lambda)^2m^2\lambda^{2m} & (\alpha/\lambda)m\lambda^m \\ \lambda^m & (\alpha/\lambda)m\lambda^m & 1 \end{bmatrix}$$

This leads to the following corollary to Theorem 2

***Corollary 1.*** *Define*

$$\widehat{\beta}_M^T \triangleq \sqrt{2^M}\left[\lambda^M\widehat{\alpha}, M\lambda^M\left(\frac{\widehat{\lambda}}{\alpha}\right), \sigma_v^2\right] \tag{27}$$

*then*

$$\widehat{\beta}_M - \beta_0 \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathcal{I}^{-1}(\beta_0)) \quad \textit{as } M\to\infty \tag{28}$$

*where* $0 < \mathcal{I}(\beta_0) < \infty$ *and*

$$\mathcal{I}(\beta_0) = \lim_{M\to\infty}\mathcal{I}_M(\beta_0). \tag{29}$$

This tells us the rate of convergence of the maximum likelihood estimator. Since use of the fast wavelet transform to obtain the data $\{r_n^m\}$ implies that $M = \log_2(N/N_0)$ and remembering that $\gamma = |\log_2\lambda|$ we have that for $f^{-\gamma}$ noise

$$\lambda^{-M} = \left(\frac{N}{N_0}\right)^{\gamma} \tag{30}$$

and so corollary 1 allows us to conclude that

$$\mathrm{var}\{\widehat{\alpha}\} = \mathcal{O}\left(\frac{1}{N^{1-\gamma}}\right) \quad \text{as } n\to\infty \tag{31}$$

$$\mathrm{var}\left\{\widehat{\lambda}\right\} = \mathcal{O}\left(\frac{1}{N^{1-\gamma}\log^2 N}\right) \quad \text{as } n\to\infty \tag{32}$$

$$\mathrm{var}\left\{\widehat{\sigma}_v^2\right\} = \mathcal{O}\left(\frac{1}{N}\right) \quad \text{as } n\to\infty \tag{33}$$

Therefore, we should expect estimates of $\alpha$ and $\lambda$ to be more accurate for smaller $\gamma$ which corresponds to $\lambda$ near 1. This is in agreement with the simulation results presented by Lundahl et.al. [18] and Wornell and Oppenheim [31] and is also in accordance with intuition. For $\gamma$ small near 1/2 the spectrum of the process is nearly flat so its size and slope at high frequencies is highly indicative of it's overall behavior. For $\gamma$ large near 1, the spectrum of the process is much more hyperbolic and hence we need to observe it at low frequencies in order to characterise it; observing the spectrum at low frequencies requires observing more data than does observing the spectrum at high frequencies.

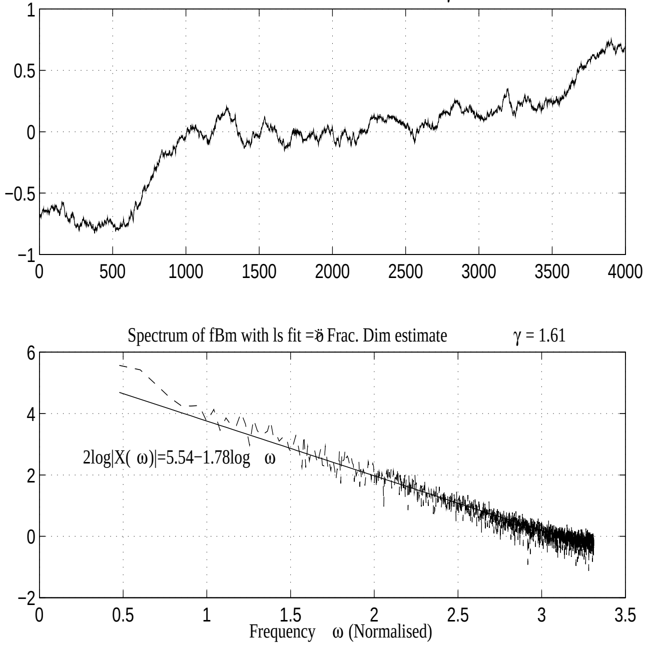


Figure 1: *A sample realisation of fBm with Hurst exponent* $H = 0.57$ *together with an illustration of the least squares estimate found by linear regression to observed periodogram.*

## REFERENCES


[1] R. J. Barton and H. V. Poor, *Signal detection in fractional Gaussian noise*, IEEE Transactions on Information Theory, IT-34 (1988), pp. 943–959.

[2] M. Basseville, A. Benveniste, K. Chou, S. Golden, R. Nikoukah, and A. Willsky, *Modeling and estimation of multiresolution stochastic processes*, IEEE Transactions on Information Theory, 38 (1992), pp. 766–784.

[3] A. Benveniste, R. Nikoukah, and A. Willsky, *Multiscale system theory*, IEEE Transactions on Circuits and Systems I, 41 (1994), pp. 2–15.

[4] P. Caines, *R.E.Kalman Festschrift*, Springer-Verlag, 1991, ch. Finite Dimensional Linear Stochastic System Identification.

[5] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[6] R. Dijkerman and R. Mazumdar, *On the correlation structure of the wavelet coefficients of fractional brownian motion*, IEEE Transactions on Information Theory, IT-40 (1994), pp. 1609–1612.

[7] P. Flandrin, *On the spectrum of fractional Brownian motion*, IEEE Transactions on Informaton Theory, IT-35 (1989), pp. 197–199.

[8] ——, *Wavelet analysis and synthesis of fractional Brownian motions*, IEEE Transactions on Informaton Theory, IT-38 (1992), pp. 910–917.

[9] J. Geweke and S. Porter-Hudak, *The estimation and application of long memory time series*

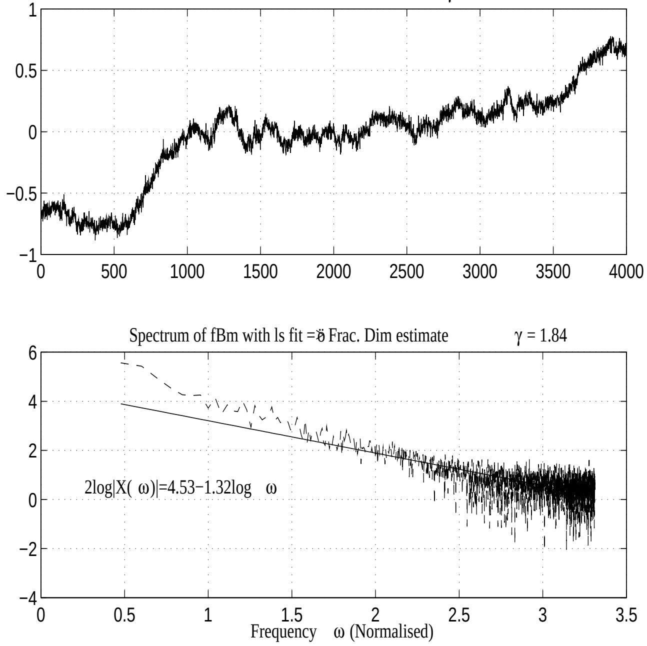


Figure 2: *Same as the previous figure, but now white measurement noise obscures the observation of the fBm.*

*models*, Journal of Time Series Analysis, 4 (1983), pp. 221–238.

[10] A. L. Goldberger and B. West, *Fractals in physiology and medicine*, Yale Journal of Medicine and Biology, 60 (1987), pp. 421–435.

[11] E. Hannan, *The asymptotic theory of linear time series*, Journal of Applied Probability, 10 (1973), pp. 130–145.

[12] B. Hoadley, *Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case*, The Annals of Mathematical Statistics, 42 (1971), pp. 1977–1991.

[13] G. Jumarie, *Stochastic differential equations with fractional Brownian motion input*, International Journal of Systems Science, 24 (1993), pp. 1113–1131.

[14] M. Keshner, 1/f *noise*, Proceedings of the IEEE, 70 (1982), pp. 212–218.

[15] J.-S. Leu and A. Papamarcou, *On estimating the spectral exponent of fractional brownian motion*, IEEE Transactions on Information Theory, 41 (1995), pp. 233–244.

[16] M. Luettgen and A. Willsky, *Likelihood calculation for a class of multiscale stochastic models*, in Proceedings of the 32nd IEEE Conference on Decision and Control, 1993, pp. 1241–1246.

[17] M. Luettgen and A. Willsky, *Multiscale smoothing error models*, IEEE Transactions on Automatic Control, 40 (1995), pp. 173–175.

fert, *Fractional Brownian motion:A maximum likelihood estimator and its application to image texture*, IEEE Transactions on Medical Imaging, MI-5 (1986), pp. 152–161.

[19] S. G. Mallat, *A theory for multiresolution signal decomposition:The wavelet representation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 11 (1989), pp. 674–693.

[20] B. B. Mandelbrot, *Some noises with* 1/f *spectrum, a bridge between direct current and white noise*, IEEE Transactions on Information Theory, IT-13 (1967), pp. 289–298.

[21] B. B. Mandelbrot and J. W. V. Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Review, 10 (1968), pp. 422–437.

[22] E. Masry, *The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion*, IEEE Transactions on Information Theory, IT-39 (1993), pp. 260–264.

[23] A. McBride, *Fractional calculus and Integral transforms of generalised functions*, vol. 31 of Research Notes in Mathematics, Pitman, San Fransisco, 1979.

[24] V. Solo, *Intrinsic random functions and the paradox of* 1/f *noise*, Siam Journal of Applied Mathematics, 52 (1992), pp. 270–291.

[25] A. Tewfik and M. Kim, *Correlation structure of the discrete wavelet co-efficients of fractional Brownian motion*, IEEE Transactions on Information Theory, 38 (1992), pp. 904–909.

[26] A. Wald, *Note on the consistency of the maximum likelihood estimate*, Annals of Mathematical Statistics, 20 (1949), pp. 595–601.

[27] P. Whittle, *Estimation and information in stationary time series*, Arkiv För Matematik, 2 (1953), pp. 423–434.

[28] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, M.I.T. Press, 1949.

[29] G. Wornell, *A Karhunen-Loève-like expansion for* 1/f *processes*, IEEE Transactions on Information Theory, IT-36 (1990), pp. 859–861.

[30] G. Wornell, *Wavelet-based representations for the* 1/f *family of fractal processes*, Proceedings of the IEEE, 81 (1993), pp. 1428–1450.

[31] G. Wornell and A. Oppenheim, *Estimation of fractal signals from noise measurements using wavelets*, IEEE Transactions on Signal Processing, 40 (1992), pp. 611–623.