Estimation of Hurst Parameter of Mono-fractal signals
EE-678, Wavelets, Prof. V.M. Gadre

Agrim Gupta, Saurabh Pinjani, Asim Ukaye
PhD Supervisor: Shivam Bhardwaj, MTech Mentor: Harsh Solanki

April 23, 2018

Abstract

In the main course assignment, we explored self similar signals and how to estimate them using techniques in wavelet domain. A good starting point was PCA Eigenspectrum paper [1] by Li et al. The proposed method works good for the H parameter in range (0.5,1) and doesn't scale up with sample size as well. Our proposed algorithm using detail coefficients of DWT improves upon these shortcomings. In addition, we studied and implemented paper on "Bayesian Estimation of γ for 1/f signals" and experimented with different wavelet basis for the same, which hadn't been attempted by the authors before

1 Introduction

1.1 Fractals

In mathematics, a self-similar object is exactly or approximately similar to a part of itself i.e. the whole has the same shape as one or more of the parts. Self-similarity is a typical property of fractals.
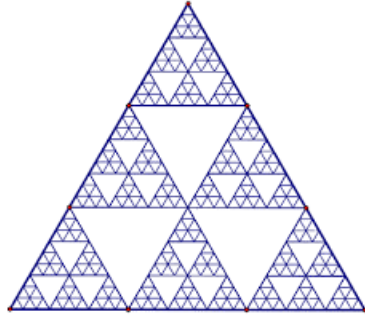
(a) Triangular Fractals   (b) Close-up of a Romanesco broccoli

Figure 1: Examples of fractals are abound.

1.2 Self-Similar Signals

A random process is said to be self-similar if its statistics are invariant to compressions and dilations of the waveform in time i.e. its statistical properties don't change with scaling. In mathematical terms

$$x(t) \equiv a^{-H} x(at)$$

Here the 'H' in the exponent is called the Hurst Parameter. This parameter is used to characterize such self similar random processes. We will be looking at methods of extraction of this parameter going forward in this report.

In this report we will be primarily dealing with two particular classes of self similar signals : (i) Fractional Brownian Motion (ii) $\frac{1}{f}$ processes

2 Self-similar Signals studied

2.1 Fractional Brownian Motion (fBm) signal

fBm, a generalization of Brownian motion, is a continuous-time Gaussian process BH(t) on [0, T], which starts at zero, has expectation zero for all t in [0, T]. Unlike classical Brownian motion, the
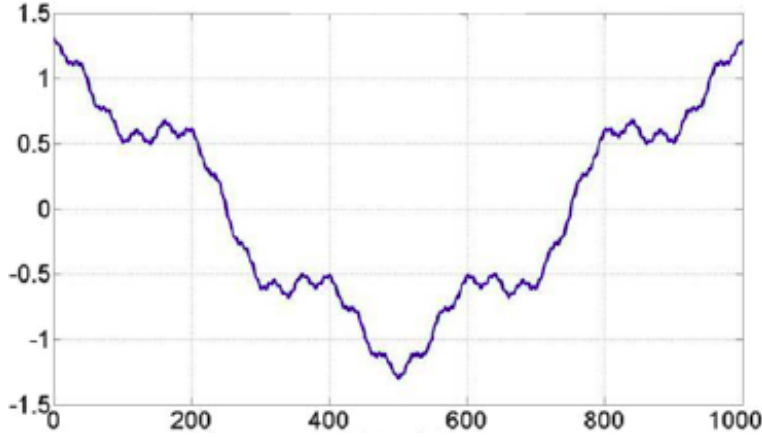
Figure 2: Weierstrass signal is an example of a self-similar signal. If we cut out the part of the signal from t=400 to t=600, the part is similar to the whole.

increments of fBm need not be independent and can be positively or negatively correlated. The degree of correlation is what characterises the process and is captured in the Hurst parameter(H). When H assumes a value lower than 0.5 the increments are negatively correlated and for values above 0.5 they are positively correlated. If H is 0.5 the steps are uncorrelated and hence the fBm becomes a vanilla brownian motion. It has the following covariance function [1]:

$$R_{B_H}(s,t) = E[B_H(s)B_H(t)] = \frac{V_H}{2}(|s|^{2H} + |t|^{2H} + |s-t|^{2H})$$

where $V_H = var[B_H(1)] = \frac{-\Gamma(2-2H)cos(\pi H)}{\pi H(2H-1)}$

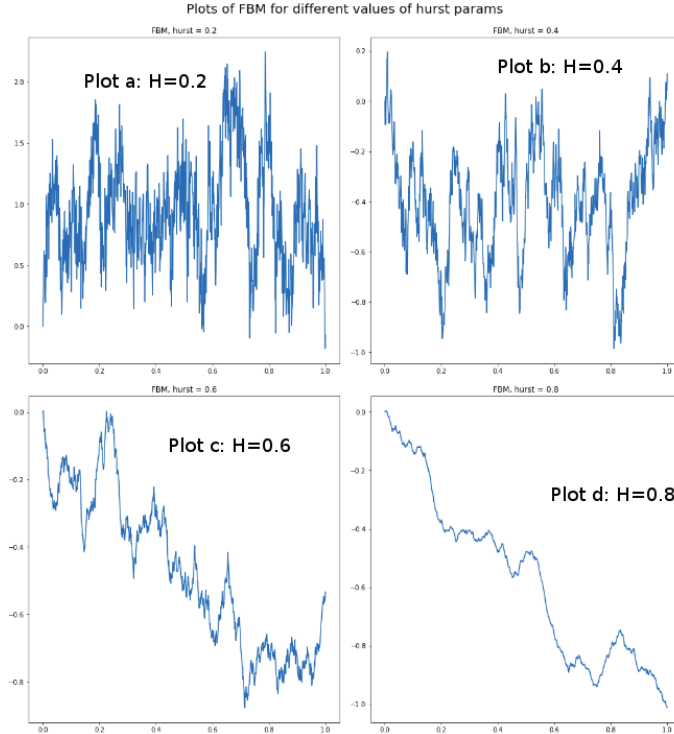We saw various forms of the covariance functions in [1],[2] and [3]. The most general version is presented above.

Figure 3: The Hurst parameter fully characterizes a monofractal signal such as fBm. The graphs above illustrate this point. In graph (a) H=0.2. Hence the incremental steps are negatively correlated and hence the graph is tremulous. However in graph (d) where H=0.8 the graph is quite smooth due to the positive correlation between subsequent steps. So therefore as we move from graph (a) to (d) the graph gets smoother with an increasing Hurst parameter.

2.2   1/f Processes

This is a class of self similar signals that possess a power spectrum obeying the following law:

$$S(\omega) = \frac{\sigma_x^2}{|\omega|^\gamma}$$

2

where $\gamma = 2H + 1$. γ typically lies between 0 and 2. It is generally convenient to extend the notion of l/f processes to include nearly 1/f processes that are defined as having power spectra bounded according to:

$$\frac{k_1}{|\omega|^\gamma} \leq S(\omega) \leq \frac{k_2}{|\omega|^\gamma}$$

where k_1 and k_2 satisfy $0 < k_1 \leq k_2 < \infty$

3 Extraction of Hurst Parameter in fBm

Since the Hurst parameter characterises the fBm, its extraction from a given signal is a problem of great relevance. This has been studied in existing literature. The method discussed in literature is discussed below. We will also be analyzing its shortfalls and domain of applicability.

3.1 H parameter estimation using PCA Spectrum (Li et al. 2009) [2]

A method to estimate the hurst parameter from the PCA eigenspectrum of a FBM signal was proposed. The PCA eigenspectrum of the autocovariance matrix of FBM signal is linear in log log scale, and the slope is indicative of the Hurst parameter. The method was theoretically proved by showing equivalence of eigenvalues arising from discrete K-L expansion and autocovariance matrix. However this method is plagued with certain issues:

- This method works for signals with H in the range (0.5,1). This is a serious shortcoming as many signals in real life have negative correlation.

- As sample size increased the difference between the estimated and actual value of Hurst increases. Ref. Figure below.
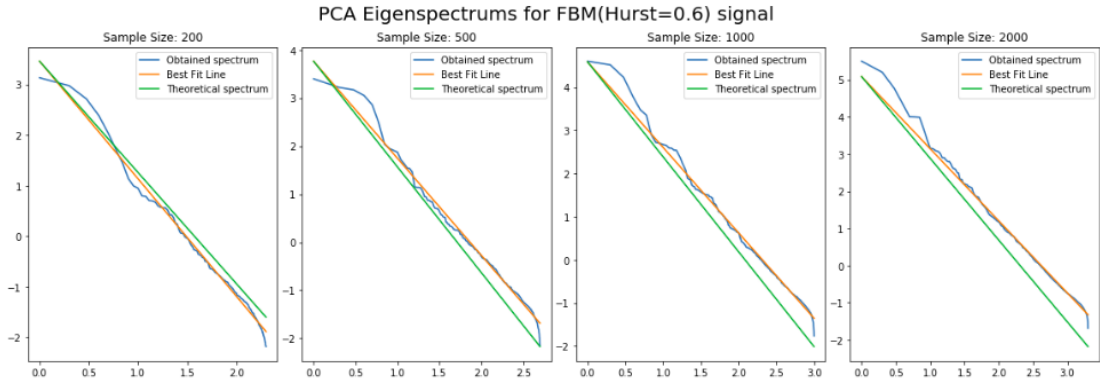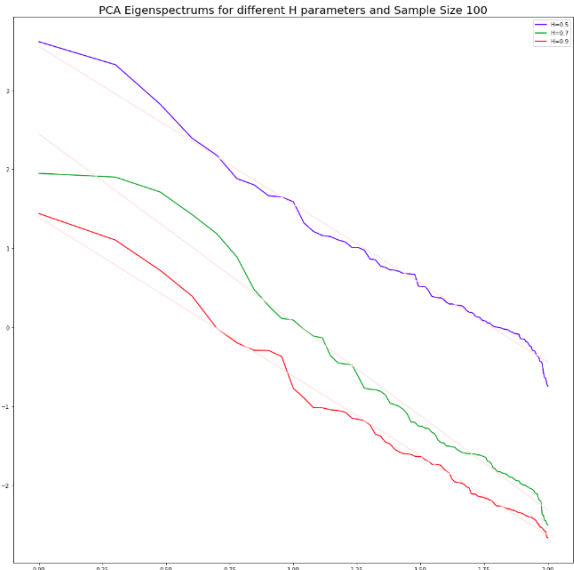

Figure 4: As Sample Size increases, deviation between actual slope and expected slope increases

3.2 H parameter estimation using Eigenspectrum of Detail coefficients

Instead of forming autocovariance matrix of the FBM signal, we first take DWT of the FBM signal. Then we use the detail coefficients (high pass branch) to form the autocovariance. Then we plot the eigenspectrum of the modified autocovariance matrix in log log scale. However, instead of

looking at the slope we look at the y-intercept. We can map the y-intercept (which is nothing but the highest eigenvalue) to the Hurst parameter. Now to judge the effectiveness of the method we look at the graph for the PCA spectrum method reported in literature. One can clearly see that separability of the curves have increased in our method. Hence our method is clearly advantageous.
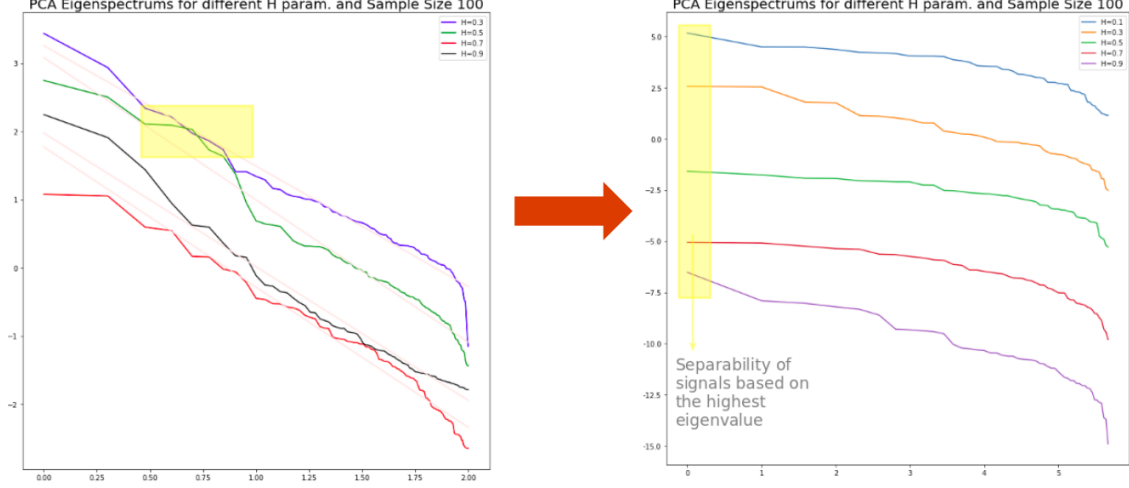

Figure 5: Separability of curves generated by different H increases when eigenspectrum of detail coefficients is plotted

Next step is to fit in the y highest eigenvalues against the Hurst parameter of the signal. A linear fit suffices and gives sufficiently good results, indicating the linear relationship between Hurst parameter and highest eigenvalue. It is clear from the graph that this method gives a fairly accurate estimate of the Hurst parameter, even for hurst parameters in the range (0,0.5) which could not be obtained accurately via the PCA eigenspectrum method proposed by Li et al.
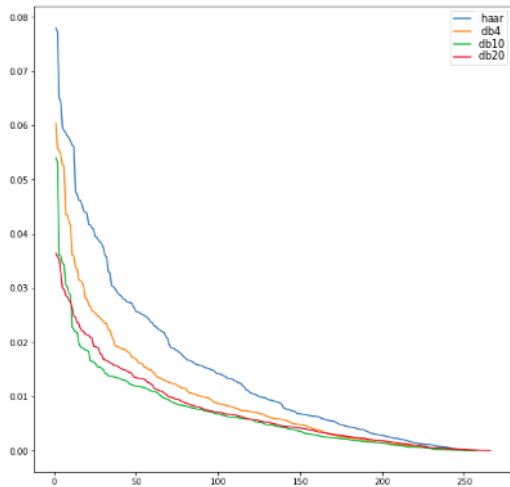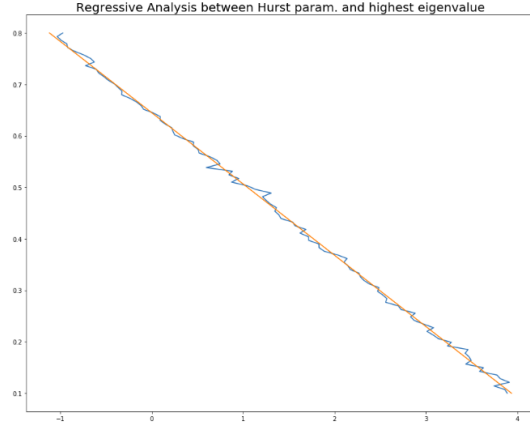

Another Interesting observation was that when eigenspectrum is plotted straightaway and not in log log scale, we see a hyperbolic decrease. The paper by H. Tewfik and M. Kim[1] reported a similar hyperbolic decrease, but for the autocorrelation after taking DWT (but not for the eigenvalues). As vanishing moments increase, the hyperbolic decrease is steeper, which was also observed in [1]. This indicates a deeper connect between eigenspectrum and direct DWT.

Figure 6: Hyperbolic decay of eigenvalues

3.2.1 Mathematical analysis of the proposed method

- Using the fact that the highest eigenvalue($\lambda_1 1$) of the Auto covariance matrix with all positive entries is $1 + (M - 1)\sum_{i=1}^{M}(p * R(p))$ [4], mathematical analysis of the results obtained was attempted. To make all the entries in the Auto covariance matrix positive, $|min_p(R(p))|$ can be added. Hence, the following formula was obtained, from which linearity when plotted in log log scale can be explained.

$$\lambda_1 = 1 + (M - 1)(\frac{1}{\frac{M(M-1)}{2}}\sum_{p=1}^{M-1} pR[p] + |min_p(R[p])|)$$

- However, to obtain the exact values of slope, we need to solve the minimization problem formally and account in for a more accurate expression for R[0], which has been left as a potential future work.

- Explaining this via Discrete KL Expansion was also attempted, but it did not give fruitful result. Another issue faced was with various forms of R[k]. Using approximations to R[k] results in counterfeit results, and as a future work expression used for R[k] should be improved upon to explain the obtained results completely

3.2.2 Future Work Possible

- Completing the analysis by solving the optimization problem directly and calculating the highest eigen value. However, there are 2 issues with this approach, since firstly, in presence of noise, the problem will become probabilistic and may be untrackable, and secondly, the optimization problem may not have a closed form. Moreover, the analysis done will be just for the highest eigenvalue and not for the full spectrum.

- Given the drawbacks of the approaches considered, proving the results using discrete K-L expansion might be the correct way to proceed. The paper by Li et al [2] presented an unusual proof which connects the discrete KL expansion to the PCA method. The proof however isn't very general and can't be readily applied to other basis functions [5]. Li et al, in January 2018, although have replied to the comment and have come up with a stronger proof concerning with Reisez basis and other such tools, and it will be a good exercise to try and generalize it to Haar wavelet [6].

- Since we are using Detail coefficients here, it makes sense to try and generalize it across the scales via wavelet packet transform. [7] by Wang et al also have proposed estimating Hurst parameter using wavelet packet transform instead of DWT.

4 Estimation of Fractality parameters (H & γ) for 1/f processes

In our assignment we also aimed to study and implement [8] presented by Wornell and Oppenheim. This paper shows a robust, computationally efficient and consistent iterative algorithm to estimate Fractality Parameters (H), and variance estimates of 1/f processes using wavelet expansion By implementing the ideas and algorithms proposed in this paper we were able to understand the wavelet domain analysis of fractals through implementation. We also aimed at replicating some of the results proposed in the paper and observe and comment on certain parts which the paper did not discuss, such as the effect of choice of wavelet basis on estimates of H.

The ultimate goal of this activity was to research the possibility of extending the framework (given for 1/f processes) of this paper to f-ARIMA processes.

4.1 Outline of the paper

The paper can be summarized in 4 parts: A background on concepts involved, Problem Formulation & Iterative Algorithm, Verification & Signal Reconstruction,and Simulations & Results.

4.1.1 Background

The paper discusses the lack of research on 1/f processes and their parameterization. $1/f$ processes are described in the first part of this report. The problem is to find the best estimate for γ for a given $1/f$ signal if embedded in a Gaussian White Noise.

The suggested algorithm uses wavelet decomposition to convert the problem in Wavelet Domain. A section of the paper is devoted to the definition of Orthonormal Wavelet Decomposition of Signals. The paper generally uses Daubechies-5 wavelet for most of the simulations. The proposed algorithm uses the the wavelet-based Karhunen-Loéve (KL)-like expansions for 1/f processes. It is shown that one can construct a class of nearly $1/f$ processes using wavelet expansions in terms of uncorrelated transform coefficients having the variance progression:

$$var\ x_n^m = \sigma^2 2^{-\gamma m}$$

where γ is the exponent of the nearly $1/f$ spectrum, and σ^2 is a positive constant proportional to σ_x^2.

4.1.2 Problem Formulation & Algorithm

Let us suppose we have observations $r(t)$ of a zero mean Gaussian $1/f$ process $x(t)$ embedded in zero-mean additive stationary white Gaussian noise $w(t)$ that is statistically independent of $x(t)$.

$$r(t) = x(t) + w(t) \qquad -\infty < t < \infty$$

From this data, the Wavelet coefficients are obtained as such:

$$r_n^m = \int_{-\infty}^{\infty} \psi_n^m(t) r(t) dt.$$

We use a DWT where the finite set representing all the scales is given by:

$$\mathfrak{M} = \{m_1, m_2, \ldots, m_M\}$$

And at each scale m, the of available coefficients is given by:

$$\mathfrak{N}(m) = \{n_1(m), n_2(m), \ldots, n_{N(m)}(m)\}$$

Hence, the available data is:

$$r = \{r_n^m \in \mathfrak{R}\} = \{r_n^m, m \in \mathfrak{M}, n \in \mathfrak{N}(m)\}$$

Exploiting the wavelet decomposition's role as a whitening filter for 1/f processes, and using the fact that the wnm are independent of the xnm and are decorrelated for any wavelet basis, the resulting observation coefficients

$$r_n^m = x_n^m + w_n^m$$

can be modeled as mutually independent zero-mean, Gaussian random variables with variance

$$var\ r_n^m = \sigma_m^2 = \sigma^2 \beta^{-m} + \sigma_w^2$$

Where,

$$\beta = 2^{\gamma}$$

Thus, the parameter set defined below is the one that is to be estimated.

$$\Theta = (\beta, \sigma^2, \sigma_w^2)$$

The likelihood function for the above parameter set is expressed as:

$$\mathfrak{L}(\Theta) = p_r(r;\ \Theta) = \prod_{m,n \in \mathfrak{R}} \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left[-\frac{(r_n^m)^2}{2\sigma_m^2} \right]$$

for which the log-likelihood function is

$$L(\Theta) = -\frac{1}{2} \sum_{m,n \in \mathfrak{R}} \left\{ \frac{1}{\sigma_m^2}(r_n^m)^2 + \ln(2\pi\sigma_m^2) \right\}$$

Equivalently,

$$L(\Theta) = -\frac{1}{2} \sum_{m \in \mathfrak{M}} N(m) \left\{ \frac{\hat{\sigma_m^2}}{\sigma_m^2} + \ln(2\pi\sigma_m^2) \right\}$$

Where the M sample variances given by

$$\hat{\sigma_m^2} = \frac{1}{N(m)} \sum_{n \in \mathfrak{N}(m)} (r_n^m)^2$$

summarize the aspects of the data required in the estimation.

4.1.3   E-M Algorithm

In order to estimate Θ, the log likelihood function $L(\Theta)$ needs to be maximized. The paper solves this problem for 3 different cases: σ_w^2 :unknown, σ_w^2 :known and, $\sigma_w^2 = 0$. Here I am just highlighting the first case and showing the steps of the algorithm proposed which effectively achieves the goal of maximizing the function. The detailed proof of the algorithm is given in the Appendix of the paper. Estimates of parameters $\beta, \sigma^2, \sigma_w^2$ generated at l^{th} step are denoted as $\hat{\beta}^{(l)}, \hat{\sigma}^{2(l)}, \hat{\sigma}_w^{2(l)}$.
E step: This step reduces to estimating the noise and signal portions of the wavelet coefficient variances at each scale m using current estimates of the parameters.

$$S_m^w(\hat{\Theta}^{(l)}) = A_m(\hat{\Theta}^{(l)}) + B_m^w(\hat{\Theta}^{(l)})\hat{\sigma}_m^2$$
$$S_m^x(\hat{\Theta}^{(l)}) = A_m(\hat{\Theta}^{(l)}) + B_m^x(\hat{\Theta}^{(l)})\hat{\sigma}_m^2$$

where,

$$A_m(\hat{\Theta}^{(l)}) = \frac{\hat{\sigma}_w^{2(l)} \cdot \hat{\sigma}^{2(l)}[\hat{\beta}^{(l)}]^{-m}}{\hat{\sigma}_w^{2(l)} + \hat{\sigma}^{2(l)}[\hat{\beta}^{(l)}]^{-m}}$$

$$B_m^w(\hat{\Theta}^{(l)}) = \left(\frac{\hat{\sigma}_w^{2(l)}}{\hat{\sigma}_w^{2(l)} + \hat{\sigma}^{2(l)}[\hat{\beta}^{(l)}]^{-m}} \right)^2$$

$$B_m^x(\hat{\Theta}^{(l)}) = \left(\frac{\hat{\sigma}^{2(l)}[\hat{\beta}^{(l)}]^{-m}}{\hat{\sigma}_w^{2(l)} + \hat{\sigma}^{2(l)}[\hat{\beta}^{(l)}]^{-m}} \right)^2$$

M step: This step reduces to using these signal and noise variance estimates to obtain the new parameter estimates.

$$\hat{\beta}^{(l+1)} \leftarrow \sum_{m \in \mathfrak{M}} C_m N(m) S_m^x(\hat{\Theta}^{(l)})\beta^m = 0$$

$$\hat{\sigma}^{2(l+1)} = \frac{\sum_{m \in \mathfrak{M}} N(m) S_m^x(\hat{\Theta}^{(l)})[\hat{\beta}^{(l+1)}]^m}{\sum_{m \in \mathfrak{M}} N(m)}$$

$$\hat{\sigma}_w^{2(l+1)} = \frac{\sum_{m \in \mathfrak{M}} N(m) S_m^w(\hat{\Theta}^{(l)})}{\sum_{m \in \mathfrak{M}} N(m)}$$

where,

$$C_m \triangleq \frac{m}{\sum_{m \in \mathfrak{M}} m N(m)} - \frac{1}{\sum_{m \in \mathfrak{M}} N(m)}$$

This concludes the problem formulation and the algorithm.

4.1.4   Verification & Signal Reconstruction

The verification part shows the convergence of the algorithm by determining the Cramer-Rao bounds on the obtained estimates. We are omitting this section as it does not contribute towards the implementation of the paper. The second part shows the use of classical estimation theory and the fact that the input signal and the noise are jointly Gaussian and uncorrelated in order to draw an elegant formula to express the most optimal estimate of the 1/f signal as:

$$\hat{x}(t) = \sum_{m,n} \hat{x}_n^m \psi_n^m(t) = \sum_{m,n} \left[\frac{\sigma^2 \beta^{-m}}{\sigma^2 \beta^{-m} + \sigma_w^2} \right] r_n^m \psi_n^m(t)$$

4.2 Simulation and Comparison with Published Results

4.2.1 Implementation in MATLAB :

- Generate a $1/f$ process $x(t)$, with known parameters (γ, N)

- Find signal parameters (β, σ^2) ; $\beta = 2^\gamma$

- Add Gaussian White Noise of known SNR. $(r(t) = x(t) + w(t))$

- Obtain wavelet coefficients for noisy signal (r_n^m)

- Feed the variance of coefficients (σ^2) of noisy signal in the E-M algorithm

- Obtain estimated signal parameters

$$\hat{\Theta}_{ML} = [\hat{\beta}_{ML}, \hat{\sigma}^2_{ML}, \hat{\sigma}^2_{w,ML}]$$

- Establish the effect on estimation error with input parameters:

  1. Resolution
  2. Choice of Wavelet basis
  3. SNR

- Reconstruct the original signal using estimated parameters.

4.2.2 Spectral Density of the Signal.

Comments:

- The power spectral density curve was obtained for the generated signal with $\gamma = 1$ and N = 65536 (2^{16}) samples.

- The red curve shows the theoretical decay of a $1/f$ process on a log-log plot

- From the estimated power spectrum, it can be seen that the slope of the input signal closely matches with that of the theoretical curve
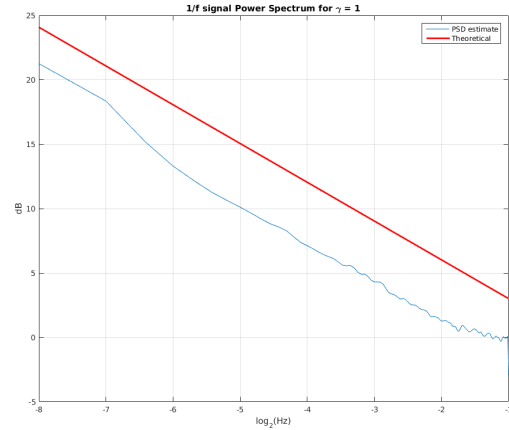

Figure 7: Power Spectral Density

4.2.3 Estimation of γ, σ^2 using linear regression

Comments:

- Fig. **??** is a log-log plot of the wavelet coefficients of the input signal x(t), versus the scale (m) .

- This plot gives us the value of γ, σ^2 using the relation:

$$var \ x_n^m = \sigma^2 2^{-\gamma m}$$

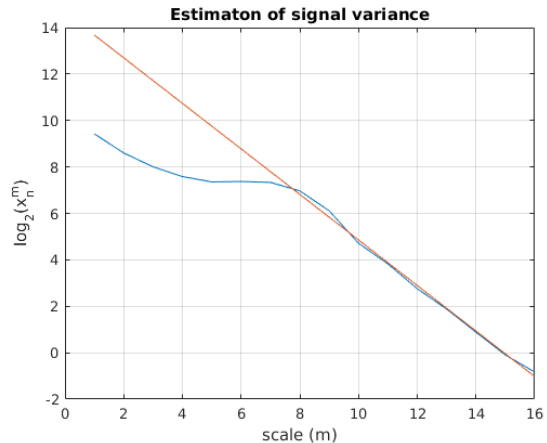- The slope of the linear fit gives $-\gamma$, and the y-intercept gives an estimate of σ^2.

Figure 8: Linear fit of K-L expansion

- An interesting observation while performing simulations shows that the curve is highly non-linear for lower values of m (coarser scales), but is linear for finer scales. As a result of this observation for multiple sample sizes, the linear fit was obtained considering the higher values of m only, i.e. for $m > M/2$. Where M is the highest scale available.

4.2.4 Absolute Error in estimation of γ for varying SNR. (Fig.9)

Sample size (N) = 2048; Wavelet Basis = Daubechies 5th order
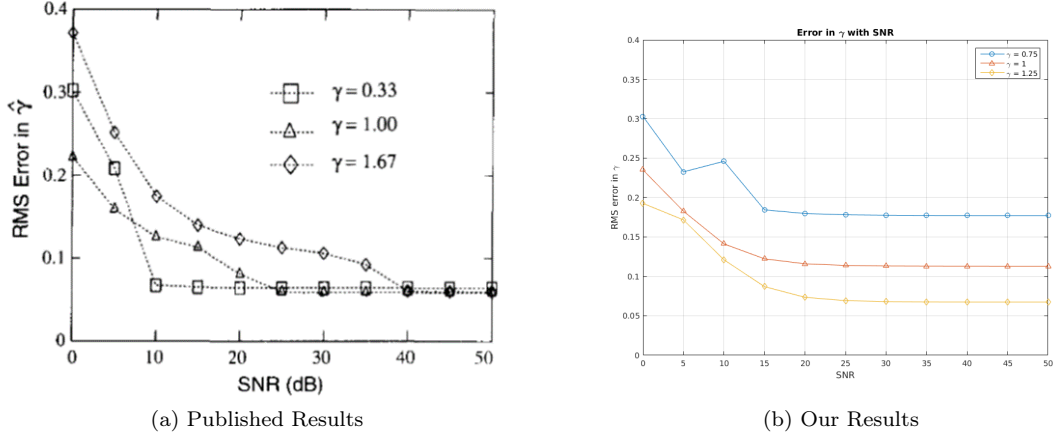

(a) Published Results
(b) Our Results

Figure 9: Error in γ with SNR

Comments and Inference:

- The error convergence with SNR is seen in both the results.

- An interesting observation is made in our results that the error values we obtained were higher for lower values of γ as opposed to the published results. (Possible reasons mentioned in the conclusion)

4.2.5 Absolute Error in estimation of γ for varying Sample Size (Fig.10)
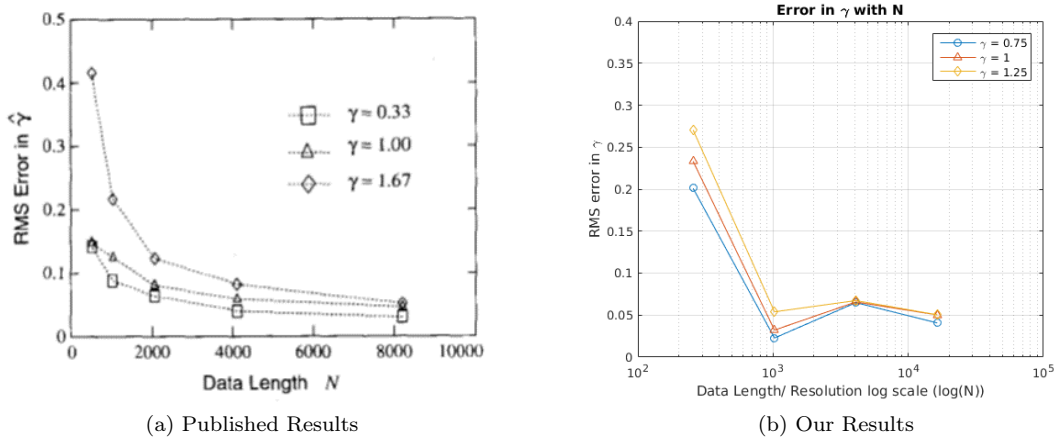
SNR = 20 dB; Wavelet Basis = Daubechies 5th order

(a) Published Results
(b) Our Results

Figure 10: Error in γ with Data Length (N)

Comments and Inference:

- The error convergence with increasing sample size is seen in both the results.

- The trend with varying γ is also similar in both the results.

- An observation is made in our results that the error values anomalously drop for the data point at N = 1024. (Possible reasons mentioned in the conclusion)

4.2.6 Signal Reconstruction (See Fig.?? and Fig. 12)

Sample Size = 65536 ; SNR = 0 dB ; Wavelet basis = DB-5; $\gamma = 1.67$ for published results; $\gamma = 1$ for our results.

Comments and Inference:

- As can be seen in both the results that the reconstructed signal estimate loses a large amount of its high frequency variations in the process of de-noising.

- We used values of $\gamma = 1$ (or close to 1) in all our simulations because of the high deviation from $1/f$ behaviour at higher or lower values of γ. This was seen in the linear fitting curve mentioned in **??**. This is mostly attributed to the signal generating process used by us which is a standard MATLAB function [$dsp.ColoredNoise()$] which is mainly suited for standard values of γ only.
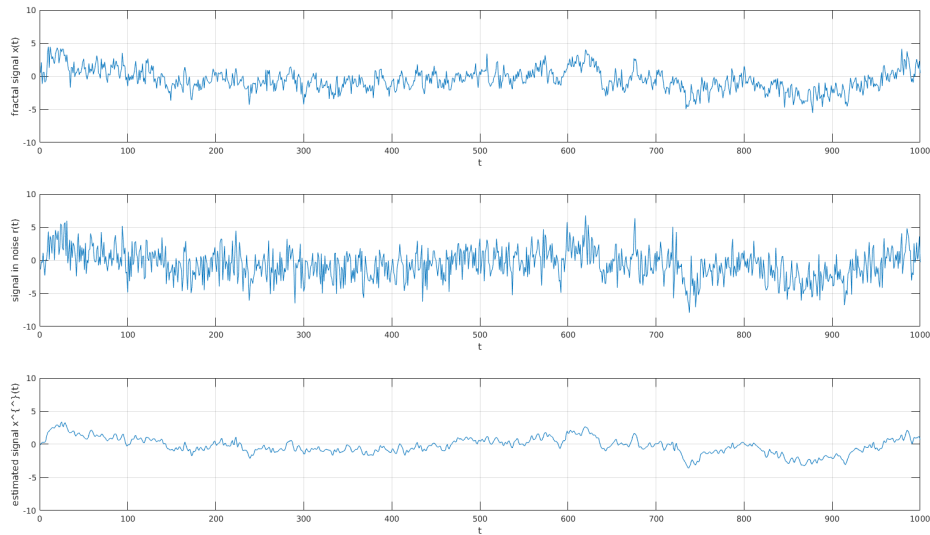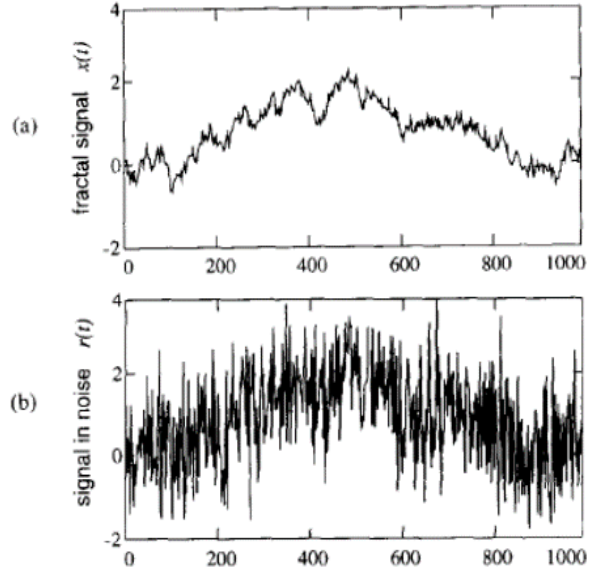
Figure 12: Signal reconstruction from our simulations

4.2.7 Additional Results on Effect of Choice of Wavelet Basis (Fig. 13)

Comments and Inference:

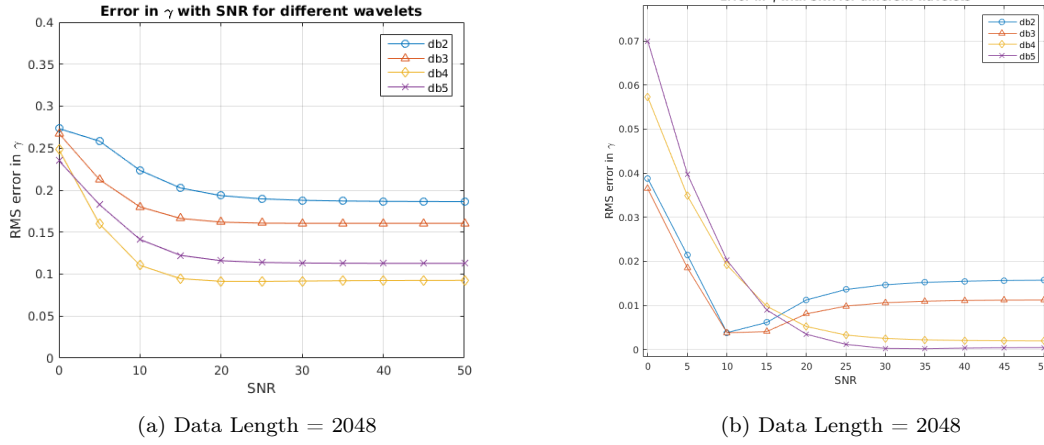- This data was not published in the paper and was additionally obtained from our simulations for $\gamma = 1$.

10

(a) Data Length = 2048 (b) Data Length = 2048

Figure 13: Error in γ with Different Wavelet Basis

- It was generally observed that the error was lower for wavelet basis with higher vanishing moments.

- Another interesting observation was made that for certain values of N, different wavelets seemed to perform the best as can be seen for N = 2048, DB-4 has a better error convergence than DB-5. Similar results were seen for other values of N. (Possible reasons explained in conclusion)

4.3   Conclusion and Future Work

- The results of the paper were replicated to a good extent. Code snippets can be found in Appendix of the report

- The deviations and anomalies, as observed above, can be accounted for by the following explanations:
    - We used a MATLAB based signal generator to replicate the 1/f process. This signal did not necessarily have a zero mean. The results published in the paper is for zero-mean signals only. Hence a lot of the anomalies can be accounted for due to this.
    - The non-zero mean in the signal lead to a greater error in the detail coefficients at coarser scales as was observed in the linear fitting curve.
    - The results published were an aggregate for Monte Carlo simulations of 64 trials. Our results were published for a signal generated by a default Random Number Generator in MATLAB. This reduced the room for variations but maintained replicability in multiple runs or re-runs of the simulation.

- This implementation can be used a base to extend the algorithm to f-ARIMA processes in future.

- The effect of wavelet basis can be studied theoretically to account for the anomalies, as seen in the results 7

References

[1] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *IEEE Trans. Information Theory*, vol. 38, no. 2, pp. 904–909, 1992. [Online]. Available: https://doi.org/10.1109/18.119750

[2] L. Li, J. Hu, Y. Chen, and Y. Zhang, "PCA based hurst exponent estimator for fbm signals under disturbances," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2840–2846, 2009. [Online]. Available: https://doi.org/10.1109/TSP.2009.2016877

[3] D. Delignieres, "Correlation properties of (discrete) fractional gaussian noise and fractional brownian motion," vol. 2015, p. Article ID 485623, 08 2015.

[4] H. Friedman, Sally; Weisberg, "Interpreting the first eigenvalue of a correlation matrix," *Educational and Psychological Measurement*, vol. 41, no. 1, pp. 11–21, 1981. [Online]. Available: http://hdl.handle.net/2027.42/67830

[5] H. van Zanten, "Comments on "pca based hurst exponent estimator for fbm signals under disturbances," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4466–4467, Aug 2010.

[6] L. Li, "Response to comments on pca based hurst exponent estimator for fbm signals under disturbances," 05 2008.

[7] Z. Wang, D. Guo, X. Li, and Y. Fei, "Estimating hurst exponent with wavelet packet," in *2006 7th International Conference on Computer-Aided Industrial Design and Conceptual Design*, Nov 2006, pp. 1–4.

[8] G. W. Wornell and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Trans. Signal Processing*, vol. 40, no. 3, pp. 611–623, 1992. [Online]. Available: https://doi.org/10.1109/78.120804