ESTIMATION OF THE SPECTRAL EXPONENT OF $1/f^{\gamma}$ PROCESS CORRUPTED BY WHITE NOISE

Süleyman Baykut[1], *Melike Erol*[2], *Tayfun Akgül*[3]

[1,3]Depts. of Electronics and Communications Engineering, [2]Computer Engineering, Istanbul Technical University
Maslak, 34469, Istanbul, Turkey
[3]TUBITAK-MRC, Earth and Marine Sciences Research Institute, Gebze, Kocaeli, Turkey
Tel: +(90) 212 285 3568, E-mail: baykut@ehb.itu.edu.tr, Web: www.ehb.itu.edu.tr/∼baykut

ABSTRACT

$1/f^{\gamma}$ noise is used to model a large number of processes; such as network traffic data, GPS (Global Positioning System) noise, financial and biological data. However, observations on real data have shown that assumption of a purely $1/f^{\gamma}$model may be inadequate, as the measured data may contain trend, periodicity or noise. These are considerable factors effecting the estimation of γ. In this work, we examine real data from GPS noise and network traffic data and apply a wavelet based method for the removal of the effect of white noise in these data sets.

1. INTRODUCTION

A wide range of natural and man-made processes from different fields show self-similar behavior. Geophysical events like variation of temperature, rainfall records, flood level variation of rivers, sunspot variability are examples of natural events whereas network traffic and financial data are examples of man-made processes [1-4]. The power spectra of such self-similar processes obey a power-law:

$$S_x(f) \sim \sigma_x^2 |f|^{-\gamma} \tag{1}$$

where σ_x^2 is the variance, and γ is the spectral exponent [5].Therefore, they are also called as $1/f^{\gamma}$ processes. Note that, γ is the slope of the power spectrum in logarithmic scale. Using this single parameter for modeling such complex processes is desirable for many applications. However, estimation of this parameter is not straightforward, especially for real data since they are prone to measurement errors (e.g., errors originating from the electronic devices in GPS antenna [6]), periodic activities (e.g., daily or weekly cycles in network data [4]), shift-in-the-mean events (e.g., high-low load transitions during busy-idle hours in networks) and etc. These factors form additive white noise, periodicity, shifting-means which may effect the estimation [7, 8].

In this work we choose two types of real data: GPS noise as an example of natural processes and network traffic data as an example of man-made processes. GPS noise is known to be the superposition of two processes: i) $1/f^{\gamma}$ noise resulting from the monument instability of the GPS antenna, meteorological events, etc., ii) white noise resulting from the characteristics of measurement devices. The existence of white noise term introduces a knee in the spectrum [6, 9] which misleads the estimators.

For the network traffic, a recent study analyzes traffic traces taken from an IP backbone [10] and a country-wide ISP [17] where the authors mention that the network traffic shows evidence of possessing two different regimes with two different γ values. We observe the same phenomena in different trace files taken from different days in 2-3 January 2006 of the same network. These observations suggest a model where $1/f^{\gamma}$noise and white noise coexist.

The measured power spectrum ($S(f)$) of $1/f^{\gamma}$noise under additive white noise becomes:

$$S(f) = \frac{\sigma_x^2}{|f|^{\gamma}} + \sigma_w^2 \tag{2}$$

Here f is the frequency, σ_w^2 and σ_x^2 are the variances of the white noise and $1/f^{\gamma}$noise, respectively. There are several studies on the estimation of spectral exponent of $1/f^{\gamma}$ processes corrupted by white noise [6, 11-13]. Here, we use a wavelet based approach to filter out the effect of white noise and estimate the γ. Note that we are not filtering out the white noise portion, we are removing its effect on the estimation. We show the results of this technique over GPS noise and network traffic data.

In Section 2, we give a brief summary of $1/f^{\gamma}$ processes. In Section 3, we introduce the wavelet based γ estimation method and discuss the effects of white noise on the estimation of γ and next, we suggest a method to remove this effect. In Section 4, we present our observations on GPS noise data and network traffic data. Finally, in Section 5 we conclude the paper.

2. $1/f^{\gamma}$ NOISE

$1/f^{\gamma}$noise has a self-similar structure and the degree of self-similarity is measured by γ parameter. In general $1/f^{\gamma}$ processes can be modeled by fGn (fractional Gaussian Noise) when γ is in the interval (-1, 1) whereas they can be modeled by fBm (fractional Brownian motion) when γ is in the interval (1, 3). fBm and fGn are normally distributed, zero-mean processes, fGn is the stationary incremental processes of fBm [2]. For $\gamma = 0$ the process becomes white noise having a flat spectrum and for $\gamma = 1$ it is called flicker noise.

3. WAVELET BASED γ ESTIMATION

The estimation of γ is an important issue since it identifies the behavior of the $1/f^{\gamma}$ process. There are a handful of methods proposed for the estimation of γ [14]. We use the wavelet based method because the performance analysis of various estimators show that it is comprehensive [15]. The estimation of γ via wavelet based method uses the wavelet transform of a process $x(t)$;

$$x_n^m = \int_{-\infty}^{\infty} x(t)\,\psi_n^m(t)\,dt \tag{3}$$

where x_n^m are the wavelet coefficients and $\psi_n^m(t)$ is the normalized dilations (m) and translations (n) of the mother wavelet, $\psi(t)$. For $1/f^\gamma$ processes the variances of these coefficients follow a power-law relationship, i.e., they are given by [5]:

$$var x_n^m = \sigma^2 2^{-\gamma m} \tag{4}$$

Here σ^2 is a positive real constant. γ is estimated as the slope of the straight line achieved from the logarithms of both sides of (4).

3.1 The Effect of White Noise over γ Estimation

In order to demonstrate that the performance of the spectral estimation methods are affected by white noise, we give a visual example using the wavelet based γ estimation method. Fig.1-a gives the wavelet based γ estimator plot of a process containing flicker noise $c(t)$ where $\gamma = 1$ and white noise $w(t)$ where $\gamma = 0$. The plot of this compound signal forms a knee-like shape with broken line around scale 7 which means that the higher scales are affected much more than the lower scales. The energies at high frequencies are smaller than that of lower frequencies for $1/f^\gamma$ processes. Therefore, at lower frequencies (scales) colored noise[1] seems dominant over white noise and at high frequencies (scales) white noise masks the colored noise.

In Fig. 1-b, we show the estimated slopes ($\gamma = 0.818$ and $\gamma = 0.107$) by wavelet based method. Neither of the estimated slopes give the original γ. It is clear that white noise affects the estimation, therefore it is essential to remove the effect of the white noise from the colored noise before computing the γ parameter. Here, we simply apply a difference operator to cancel out this effect.

Lets define the compound process as the summation of colored $c(t)$ and white noise $w(t)$:

$$v(t) = c(t) + w(t) \tag{5}$$

Since $c(t)$ and $w(t)$ are statistically independent, the wavelet coefficients of $v(t)$ is obtained as:

$$v_n^m = c_n^m + w_n^m \tag{6}$$

and the variances of these coefficients are:

$$(\sigma_v^m)^2 = \sigma^2 2^{-\gamma m} + \sigma_w^2 \tag{7}$$

To estimate γ, σ^2, and σ_w^2, we first take the difference of both sides of (7) and cancel out the constant term, σ_w^2:

$$\triangle(\sigma_v^m)^2 = (\sigma_v^m)^2 - (\sigma_v^{m+1})^2 \tag{8}$$

$$\triangle(\sigma_v^m)^2 = \sigma^2\left(1 - 2^{-\gamma}\right)2^{-\gamma m} \tag{9}$$

Here, γ and σ^2 can be estimated by fitting a straight line to the logarithm of (9). Afterwards, σ_w^2 can also be estimated by substituting γ and σ^2 in (7).

[1] In this paper, we use colored noise and $1/f^\gamma$ process interchangeably.

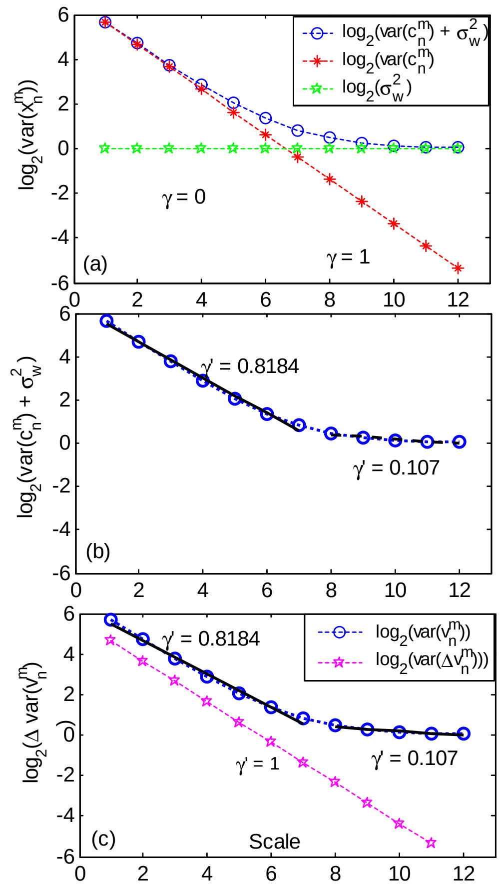

Figure 1: The mimic plots of the wavelet based estimation method for flicker noise + white noise, (a) variance progression of $c(t)$, $w(t)$ and $c(t) + w(t)$, (b) the estimated slopes, (c) the difference sequence of the variances of the wavelet coefficients.

If we apply the above operations on the process of Fig. 1-a we obtain the γ estimate given in Fig. 1-c. Here, γ is estimated as 1 which is equal to the γ of the underlying process (in this case, flicker noise).

4. EXAMPLES OF REAL DATA

In this section we analyze two real data sets: GPS noise as an example of natural processes and network traffic data as an example of man-made processes.

4.1 GPS Noise Data

GPS systems are used for the estimation of surface displacement and strain velocity. Correct estimation of these requires accurate error analysis. It is already known that GPS error is best modeled by white noise ($\gamma = 0$) plus flicker noise ($\gamma = 1$) [3, 6, 9]. Therefore, it is appropriate to apply the method of Section 3.1 to GPS noise.

We analyze the GPS data collected in various stations in Turkey between 1999-2004. These data sets are recorded on

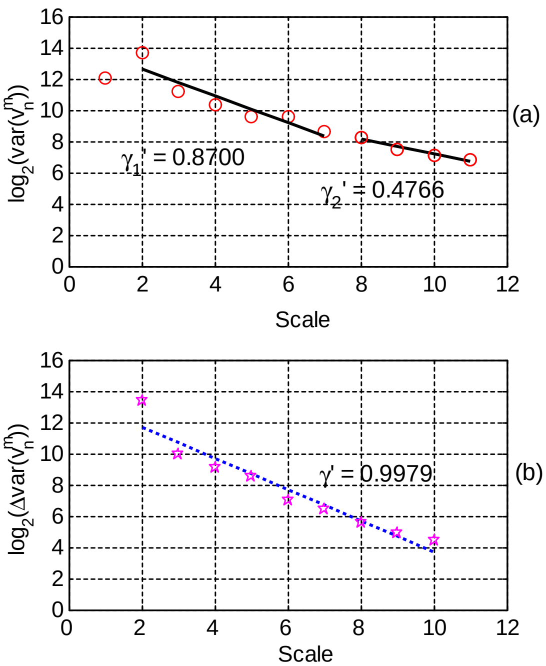

Figure 2: Wavelet based estimation of GPS noise (from station 1), (a) standart wavelet based estimator, (b) wavelet based estimator with difference operator.

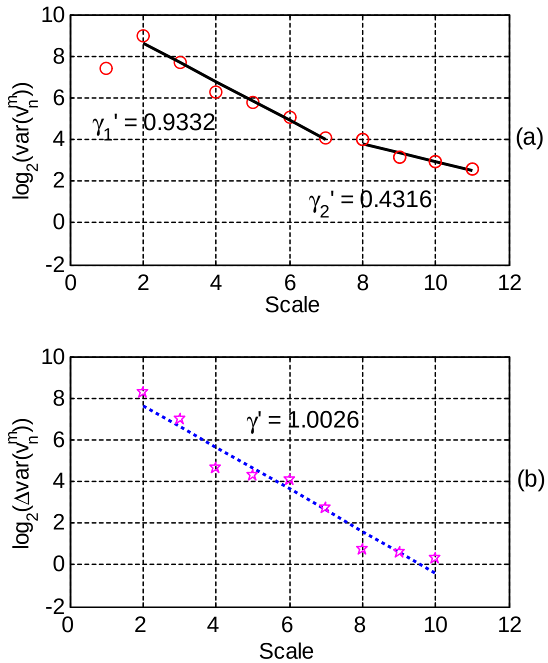

Figure 3: Wavelet based estimation of GPS noise (from station 2), (a) standart wavelet based estimator, (b) wavelet based estimator with difference operator.

a daily basis and the data length is $N = 1600$. Here, we provide only the data sets taken from two different GPS stations. However, the results for the other data sets are similar. In Fig. 3-a and Fig. 4-a, we give the result of the wavelet based γ estimation method. Due to the white noise component, a knee is observed around the 7^{th} scale. In Fig. 3-b and Fig. 4-b, we give the resulting γ after applying the difference operator to the variances of the wavelet coefficients. The estimated γ values are quite close to 1 ($\gamma = 0.9979$ and $\gamma = 1.0026$) which shows that the corresponding colored noise is flicker noise.

4.2 Network Traffic Data

Network traffic is known to be self-similar since 1990s [16] and it is widely modeled by fGn. In a recent work a behavior similar to GPS noise (knee in the wavelet plot) is observed in measurements from an IP backbone [10] and a country-wide ISP [17]. We use the same traffic archive of [10] (MAWI Archive [18]) and examine the trace files collected on 2-3 January 2006. The traces are collected from a trans-pasific link between 14:00 and 14:15 spanning an interval of 900s. We aggregate the data over 10ms intervals to form the byte count process. In Fig.5-a and Fig.6-a, we give the result of the wavelet based γ estimation method of these byte count processes. A knee is observed around the 9^{th} scale. In Fig. 5-b and Fig. 6-b, we give the resulting γ after applying the difference operator to the variances of the wavelet coefficients. The γ values are estimated as $\gamma = 0.6793$ and $\gamma = 0.802$.

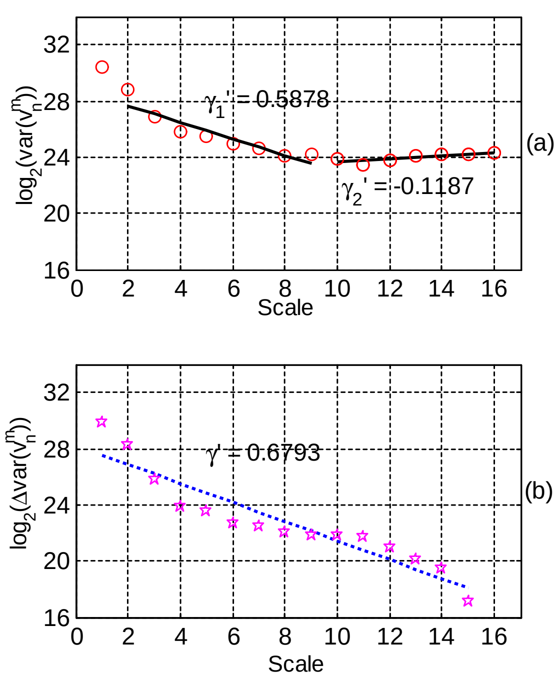

Figure 4: Wavelet based estimation of network traffic data (2 Jan. 2006), (a) standart wavelet based estimator, (b) wavelet based estimator with difference operator.

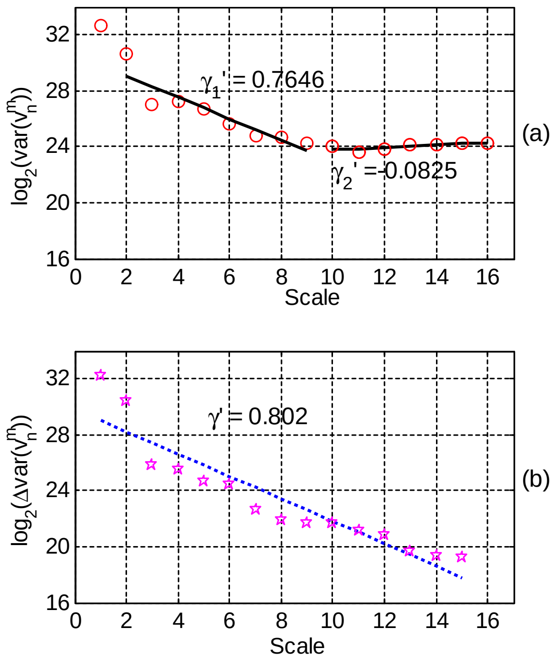

Figure 5: Wavelet based estimation of network traffic data (3 Jan. 2006), (a) standart wavelet based estimator, (b) wavelet based estimator with difference operator.

5. CONCLUSION

When colored noise and white noise coexist, the estimation of γ is affected. White noise introduces a knee in the plots of wavelet based γ estimation method yielding two different slopes (estimates). Since neither of the slopes give the exact γ parameter, γ should be estimated after removing the effect of white noise. In this paper we apply a difference operator to the variances of the wavelet coefficients and cancel out this effect. We show the results of this technique on two real data examples; namely GPS noise and network traffic data.

Acknowledgement: The authors would like to thank Sema Oktuğ for her valuable suggestions, Semih Ergintav for providing GPS data and Kenjiro Cho for providing traffic traces.

REFERENCES

[1] G. W. Wornell, *Signal Processing with Fractals: A wavelet Based Approach*, Prentice Hall, NJ, 1996.

[2] B. B. Mandelbrot, J. W. Van Ness, "Fractional Brownian Motions, Fractional Noise and Applications," *SIAM Review*, vol. 10, no. 4, pp. 422-437, 1968.

[3] D. C. Agnew, "The Time-domain Behavior of Power-law Noises", *Geophys. Res. Lett.*, vol. 19,pp. 333-336, 1992.

[4] J. Beran, *Statistics for Long-memory Processes*, Chapman&Hall, 1994.

[5] G. W. Wornell, "Wavelet-Based Representations for the $1/f$ Family of Fractal Processes," *Proc. of IEEE*, vol. 81, no. 10, pp. 1428-1450, 1993.

[6] J. Langbein, and H. Johnson, "Correlated Errors in Geodetic Time Series: Implications for Time-Dependent Deformation," *Journal of Geophysical Research*, vol. 102,pp. 591-604, 1997.

[7] M.S. Taqqu, V. Teverovsky, "Testing for Long-range Dependence in the Presence of Shifting Means or a Slowly Declining Trend, Using a Variance-type Estimator", *Journal of Time Series Analysis*, vol. 18, pp. 279-304,1997.

[8] S. Baykut, T. E. Özkurt, M. Erol, T. Akgül , "The Influence of a Single Tone Sinusoid Over Hurst Estimators", *Proc. of European Signal Processing Conference (EUSIPCO)*, Antalya, September 4-8, 2005.

[9] A. Mao, C. G. A. Harrison, T. H. Dixon, "Noise in GPS Coordinate Time Series", *J. Geophys. Res.*, vol.104, pp. 2797-2826, 1999.

[10] S. Bregni, L. Primerano, "The Modified Allan Variance as Time-Domain Analysis Tool for Estimating the Hurst Parameter of Long-Range Dependent Traffic", *Proc. of IEEE GLOBECOM*, Dallas, USA, Nov. 29 - Dec. 3, 2004.

[11] G. W. Wornell, A. V. Oppenheim, "Estimation of Fractal Signals from Noisy Measurements Using Wavelets," *IEEE Transactions on Signal Processing*, vol. 40, no.3, 1992.

[12] L. Du, Y. Zhuang, Y. Wu, "$1/f^{\gamma}$ Noise Separated from White Noise with Wavelet Denoising," *Microelectronic Reliability*, vol. 42,pp. 183-188, 2002.

[13] B. Ninness, "Estimation of $1/f$ Noise", *IEEE Transactions on Information Theory*, vol. 44(1), pp. 32-46, 1998.

[14] M. S. Taqqu, V. Teverovsky, W. Willinger, "Estimators for Long-range Dependence: An Empirical Study," *Fractals*, vol. 3, no. 4, pp 785-798, 1995.

[15] P. Doukhan, G. Oppenheim, M. S. Taqqu, *Theory and Applications of Long-Range Dependence*, Birkäuser, 2003.

[16] W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson, "On the Self-Similar Nature of Ethernet Traffic", (extended version), *IEEE/ACM Trans on Networking*, vol. 2, no.1, pp. 1-15, 1993.

[17] S. Bregni, W. Erangoli, "Fractional Noise in Experimental Measurements of IP Traffic in a Metropolitan Area Network", *Proc. of IEEE GLOBECOM*, St. Louis, USA, Nov. 28 - Dec. 2, 2005.

[18] MAWI Project Traffic Archive, http://tracer.csl.sony.co.jp/mawi/