Signal Denoising Using Empirical Mode Decomposition and Higher Order Statistics

George Tsolis and Thomas D. Xenos
Aristotle University of Thessaloniki, Greece, Department of Electrical and Computer Engineering.
strantis@gmail.com, tdxenos@eng.auth.gr

Abstract

A hybrid denoising method is presented as a combination of Empirical Mode Decomposition (EMD) and Higher Order Statistics (HOS). EMD, an adaptive data-driven method, is used for effective decomposition of a noisy signal into its functional components. Then Kurtosis and Bispectrum operate as Gaussianity estimators, supplemented by Bootstrap techniques, ensuring detection and removal of the signal's Gaussian components. Thresholding techniques are used at the final step for maximum suppression of signal noise, where thresholds are set by estimating the long-term correlation of the corrupting colored noise in the form of the Hurst exponent. Experimental results prove the applicability of the method in signal denoising. Specifically, EMD-HOS outperforms similar denoising techniques based on Wavelets for the most types of test signals.

Keywords: *Empirical Mode Decomposition, Higher Order Statistics, Hurst exponent, Bootstrap Statistics, Signal Thresholding.*

1. Introduction

Detection of signals in a noisy environment is a classic problem in Signal Processing. Noise is considered as any unwanted signal that interferes with the information signal; it can be an acoustic signal, a biomedical signal, a RADAR signal, or any other type of signal conveying useful information for the respective application. The goal of any signal denoising method is to effectively reduce noise level, in order for a useful signal to emerge, while minimizing the information loss. The simpler method for signal denoising is to employ some type of linear filter to remove the noise from the signal. The problem with this technique is that the spectrum of the noise must be known for the appropriate filter to be constructed and the signal needs to be stationary, which is usually not the case for real signals. Time-frequency analysis methods, such as Short Time Fourier Transform (STFT) or Wavelets are considered more appropriate to handle non-stationary signals. Thus they are frequently used for real signal denoising either in time or frequency domain. However, STFT suffers from its inability to precisely localize the signal in time, while simultaneously maintaining adequate frequency resolution [1]. Moreover, STFT requires piece-wise stationarity of the data, and also assumes that the stationarity scales coincide with the sliding window length used for the decomposition, something that is practically impossible to guarantee, [2]. Wavelets, although achieve satisfactory time-frequency resolution, need a basis function to be specified, the choice of which is not always a trivial problem. Empirical Mode Decomposition is a relatively new method in signal processing proposed by Norden Huang of NASA, [2]. Despite the fact that it is primarily an empirical method, it has proved very effective for the analysis of non-stationary and non-linear signals. The biggest advantage of the method is that it is totally adaptive and data driven, without the need for a-priori basis function selection (i.e., mother wavelet) for signal decomposition.

The EMD denoising properties were addressed soon after the introduction of the method, and they are reported in [3], [4], [5], [6]. Subsequently, EMD was effectively used for signal denoising in a wide range of applications, such as biomedical signals, [7], acoustic signals, [8], and ionospheric signals [9]. It was also used as a hybrid method in conjunction with a fractal dimension filter for denoising of biomedical signals, [10]. More recently, a wavelet-type thresholding approach was applied to EMD, where thresholds were decided depending on the estimated energy of the noise in the signal, [11].

In this study we propose the use of EMD combined with Higher Order Statistics in a method denoted as EMD-HOS hereafter, for Gaussian noise removal. HOS methods were used in several studies for Gaussian noise classification, together with other signal analysis methods, notably Wavelets or STFT [1], [12], [13] [14]. In our case EMD is used for the decomposition of the signal in its functional elements, followed by HOS which aims to detect and remove any Gaussian scales from the signal. The accuracy of the HOS parameters is verified by applying the Bootstrap technique, a powerful statistical method especially effective when the number of signal samples is too small to ensure stable results. EMD thresholding techniques introduced by [11] are also used, to further suppress the remaining noise energy in the signal and improve the output Signal to Noise Ratio (SNR). The proposed method was tested on various signal types giving special emphasis to RADAR signals.

2. Empirical Mode Decomposition

Empirical Mode Decomposition constitutes the first stage of an algorithm known as Hilbert-Huang transform, where a real or complex signal is decomposed in a series of structural components, known as Intrinsic Mode Functions (IMF). An IMF is defined as any function having the same number of zero-crossings and extrema, and also having symmetric envelopes defined by the local maxima and minima respectively. Since IMFs admit well-behaved Hilbert transforms, the second stage of the algorithm is to use the Hilbert transform to provide instantaneous frequencies as a function of time for each one of the IMF components. Depending on the application, only the first stage of the Hilbert-Huang Transform may be used, which is the case in this particular study.

For a discrete time signal x(n) the EMD starts by defining the envelopes of its maxima and minima using cubic splines interpolation. Then, the mean of the two envelopes is calculated as

$$m_1(n) = \frac{E_{max\ (n)} + E_{min\ (n)}}{2} \tag{1}$$

Accordingly, the mean $m_1(n)$ is then subtracted from the original signal

$$h_1(n) = x(n) - m_1(n) \tag{2}$$

and the residual $h_1(n)$ is examined for the IMF criteria completeness. If it is an IMF then the procedure stops and the new signal under examination is expressed as

$$x_1(n) = x(n) - h_1(n) \tag{3}$$

However, if $h_1(n)$ is not an IMF, the procedure, also known as ''sifting" is continued k times until the first IMF is realized. Thus,

$$h_{11}(n) = h_1(n) - m_{11}(n) \tag{4}$$

where the second subscript index corresponds to sifting number, and finally

$$\mathrm{IMF}_1 = h_{1k}(n) = h_{(k-1)}(n) - m_{1k}(n) \tag{5}$$

The sifting process is continued until the last residual is either a monotonic function or a constant. It should be mentioned that as the sifting process evolves the number of the extrema from one residual to the next drops, thus guaranteeing that complete decomposition is achieved in a finite number of steps. The final product is a wavelet-like decomposition going from higher oscillations to lower oscillations, with the frequency content of each mode to decrease as the order of the IMF increases. The big difference however, with the wavelet analysis is that while modes and residuals can intuitively be given a "spectral" interpretation in the general case, their high versus low frequency discrimination applies only locally and corresponds in no way to a predetermined subband filtering. Selection of modes instead corresponds to an automatic and adaptive (signal-dependent) time-variant filtering, [5].

After completion of EMD the signal can be written as follows,

$$x(n) = \sum_{i=1}^{k} \mathrm{IMF}_i + r(n) \tag{6}$$

where k is the total number of the IMF components and $r(n)$ is the residual.

3. Higher Order Statistics

Higher Order Statistics is a broad term encompassing statistical descriptors of order greater than two. The Cumulants and Polyspectra are just higher order generalizations of the Correlation and the Spectrum, [15]. Among the unique properties that HOS incorporate, is their ability to classify signals as Gaussian or not. This is because the HOS of Gaussian signals are equal to zero, which is not the case for non-Gaussian ones. Therefore, the degree of a signal non-Gaussianity is practically equivalent to measuring its cumulants deviation from zero. A good choice for Gaussianity measurement using HOS is the Kurtosis, which is defined as the normalized version of the fourth order Cumulant, [1]. Assuming a zero mean signal, the normalized Kurtosis is expressed as:

$$K_4 = \frac{E[x(t)^4]}{(E[x(t)^2])^2} - 3 = N\frac{\sum_{n=1}^{N} x(n)^4}{\left(\sum_{n=1}^{N} x(n)^2\right)^2} - 3 \tag{7}$$

where N is the number of signal samples, and $n = 1,2,\ldots,N$. If the signal is Gaussian then the bias and variance of the Kurtosis is estimated in [1] and [12] as

$$B(K_4) = -\frac{6}{N}, \mathrm{Var}(K_4) = \frac{24}{N} \tag{8}$$

and a Kurtosis estimator can take values as follows,

$$|K_4| \leq \frac{\sqrt{24/N}}{\sqrt{1-a}} \tag{9}$$

where is an authorized confidence percentage value, with a numerically estimated optimum equal to 90%, [1,12]. However, even though the Kurtosis of a Gaussian signal is restricted by equation (9), in fact Kurtosis estimation may still be invalid, especially when the samples of the signal are not numerous enough to ensure convergence. Here, the solution comes in the form of the Bootstrap technique. Bootstrap is a statistical method to increase the accuracy of the estimator, and it is very effective in cases when the available signal samples are limited. Bootstrap does exactly what a scientist would do in practice if it were possible: it repeats the experiment many times. Bootstrap randomly reassigns the observations, re-computes the estimates many times and treats these reassignments as repeated experiments, [16]. In this study Bootstrap is used to evaluate the Kurtosis of each of the signal IMFs right after EMD. The number of the IMF samples reassignments is limited to 1000 for computational load purposes. The Boostrap algorithm gives a maximum and minimum Kurtosis estimation of the Kurtosis value for each of the signal IMFs, which in turn is compared with the theoretical Kurtosis limit as it is calculated by equation (9).

Another candidate to test Gaussianity is the signal Bispectrum. The Bispectrum is defined as the third order Spectrum of the signal and is calculated either as the Fourier transform of its third order Cumulant, or as the triple product of its Fourier coefficients, [17]. That is,

$$B_2^x(\omega_1,\omega_2) = E\{X(\omega_1)X(\omega_2)X^*(\omega_1+\omega_2)\} = m_3 X(\omega_1)X(\omega_2)X^*(\omega_1+\omega_2) \quad (10)$$

where $|\omega_1| \leq \pi, |\omega_1| \leq \pi, |\omega_1+\omega_2| \leq \pi,$ $X(\omega)$ is the Fourier transform of the signal $x(n)$ and m_3 is its third order moment. The Bispectrum of a Gaussian process is zero. However, there exist statistical processes where the Bispectrum is zero despite deviating from gaussianity. In other words a non Gaussian signal may admit zero Bispectrum, but if a signal is Gaussian its Bispectrum has to be equal to zero. To this end, using the Bispectrum criterion to ensure Gaussianity after the signal is classified as Gaussian by the Kurtosis test, appears to be a good choice. Although the computational cost increases by using two Gaussianity estimators, by doing this we ensure that the IMF under examination can be safely classified as Gaussian and excluded from the signal reconstruction process as explained in section 5.

Various methods exist in literature for Bispectrum estimation, either based on Fourier transform known as conventional methods, or parametric ones based on statistical models (Autoregressive, Moving Average, ARMA). Conventional methods are categorized also as direct or indirect depending on the choice of the technique used for the Bispectrum calculation. In this work the indirect method proposed by [17] was adopted, that is, data segmentation for the third order moment sequence estimation, and then calculation of the Bispectrum by taking the two-dimensional Fourier transform. For the Gaussianity criterion, the magnitude of each IMF Bispectrum is compared with the Bootstrap estimated Bispectrum value, of a simulated Gaussian noise signal having the same number of samples with the noisy signal.

4. Gaussian Noise Model and EMD Signal Denoising

Fractional Gaussian noise (fGn) is the generalization of white noise. fGn exhibits a flat spectrum, and its statistical properties are determined solely by a scalar parameter H, known as the Hurst exponent, [5]. In fact, if a zero mean Gaussian stationary process is classified as fGn, then its autocorrelation function is expressed as in [5],

$$r_H[k] = \frac{\sigma^2}{2}\left(|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H}\right) \quad (11)$$

where σ is the variance of the process, H is the Hurst exponent and k is the lag of the correlation. For the special case where $H = 0.5$, the process is reduced to uncorrelated white noise, whereas for other values of H it expresses colored Gaussian noise. Extended simulations with fGn and white noise verified that EMD acts on fGn as a dyadic filter bank, [5]. Moreover, the log-variance of the IMFs follows a simple linear model which is completely controlled by the Hurst exponent as,

$$\log_2 V_H[k] = \log_2 V_H[2] + 2(H-1)(k-2)\log_2 \rho_H \quad (12)$$

for $k \geq 2$ and $\rho_H \approx 2$. Then the energy of each of the IMFs can be parameterized as a function of the first IMF energy as,

$$E_k = (E_1/\beta_H)\rho_H{}^{-2(1-H)k}, k \geq 2 \tag{13}$$

where β_H is experimentally estimated in [5] for three values of the Hurst exponent as $(H = 0.2, \beta_H = 0.487), (H = 0.5, \beta_H = 0.719), (H = 0.8, \beta_H = 1.025)$, and the energy of the first IMF is given by the equation:

$$E_1 = \frac{1}{N}\sum_{n=1}^{N}(IMF_1)^2 \tag{14}$$

A simple denoising method based on EMD is proposed by [5] by applying the above model. The method suggests decomposition of the noisy signal into IMFs, and then comparing the IMF energies with the theoretical estimated noise-only IMF energies as they are calculated by equation (13). Then, the signal reconstruction is implemented by summing the IMFs, whose energy deviates from the theoretical noise model. Although this technique works very well when the noise is in a different frequency band from the signal, hence noise is captured in specific IMFs, it doesn't give reasonable results when signal and noise share the same bandwidth. When this is the case, IMFs will carry noise, while also maintaining a significant amount of the signal energy, thus deviating from the theoretical noise model. This prohibits their exclusion from the signal reconstruction process. To alleviate this problem, wavelet-like thresholding techniques are proposed by [11], where the signal IMFs are thresholded using variable thresholds specified by the above theoretical noise model. The threshold is a modified version of the universal threshold proposed by Dohono [19], expressed as

$$T = c\sqrt{V_i 2\ln(N)} \tag{15}$$

where Vi is the noise variance estimated by the noise model for the i^{th} IMF, $(i \geq 2)$, N the number of signal samples and c a constant experimentally found to take values 1 to 0.7 depending of the type of signal. The assumption also that the total noise energy is captured by the first IMF is not valid in the general case, therefore the noise variance for the first IMF is estimated using a better estimator as

$$V_1 = \left(\frac{\text{median } |IMF_1|}{0.6745}\right)^2 \tag{16}$$

An alternative approach for the noise variance estimator is proposed in [6], where the absolute median deviation of the first IMF is taken into account as

$$V_1 = \left(\frac{\text{median } (|IMF_1 - \text{median } (IMF_1)|)}{0.6745}\right)^2 \tag{17}$$

A series of simulations concluded that the second of these two versions of noise variance estimator, performs better for all types of signals. For space purposes, the results of these trials are not presented in this paper. Nevertheless, the denoising method presented in this study was implemented by using the noise variance estimator expressed by the equation (17).

Having now determined the thresholds for each IMF, the wavelet-like denoising method would necessitate zeroing the portion of the IMF which is below the threshold. However, the IMF nature, requires setting to zero the IMF portion between two adjacent zero crossings, when the absolute maximum of the IMF in this interval is below the predefined threshold. This fact is based on the assumption that if the extremum which lies inside two adjacent zero crossings interval exceeds the threshold, the interval is signal dominant, otherwise it is noise dominant, [11]. Hence the thresholding operation for the EMD case and for every two successive zero crossings interval, $z_i^j = \left[z_i^j,\ z_i^{j+1}\right]$, can be expressed as

$$\tilde{z}_i^j = \begin{cases} z_i^j, & \left|r_i^j\right| > T_i \\ 0, & \left|r_i^j\right| \le T_i \end{cases} \tag{18}$$

where $\tilde{z}_i^j$ denotes the thresholded interval, is the IMF order index, r_i^j is the j^{th} extremum of the i^{th} IMF, and $j = 1,2,..(M_i - 1)$, with M_i being the number of the zero crossings of the i^{th} IMF. Similarly, for the soft thresholding case, the z_i^j is given as

$$\tilde{z}_i^j = \begin{cases} z_i^j \frac{\left|r_i^j\right| - T_i}{r_i^j}, & \left|r_i^j\right| > T_i \\ 0, & \left|r_i^j\right| \le T_i \end{cases} \tag{19}$$

Consequently, the thresholded IMF is formed by concatenating the thresholded intervals,

$$\mathrm{I\widetilde{M}F}_i = \left[\tilde{z}_i^1, \tilde{z}_i^2, \tilde{z}_i^3, \ldots, \tilde{z}_i^j\right] \tag{20}$$

In the denoising scheme explained above, the noise is assumed to be white and Gaussian. However, the noise model of [5], allows accounting for the colored Gaussian noise case, if there is only available information for its Hurst exponent. There are various methods to estimate the Hurst exponent in data, mainly by examining their long range correlations. The Aggregated Variance method or Dispersional Analysis is one good candidate, since it is a time domain method useful for non-stationary time series. The method performs a multi-scale analysis by aggregation of adjacent points and measuring the similarity in terms of variance, [20]. The variance is calculated by averaging the time over bins of width τ. Given a fGn series $\xi_H(i), i = 1,2.3, \ldots, N$, the algorithm is as follows, [21],

a. Set the bin size $\tau = 1$
b. Calculate the standard deviation of N data points and record the point $(\tau, \tau \cdot \sigma_\tau)$.
c. Average the neighboring data points and store in the original data set as
   $\xi_H(i) \leftarrow \frac{1}{2}[\xi_H(2i-1) + \xi_H(2i)]$
d. Rescale appropriately as $N \leftarrow \frac{N}{2}, \tau \leftarrow 2\tau$.
e. When $N > 4$, step two is repeated.
f. Plot the log-log graph and perform linear regression. The Hurst exponent is the slope of the log-log plot as $\log(\tau, \tau \cdot \sigma_\tau) = H\log(\tau) + L$, where L is a constant.

The obvious question that arises at this point is which data will be used for the Hurst exponent estimation. Apparently, the Hurst exponent of the original signal is not necessarily equal to the Hurst exponent of the noise, since various correlations in the data may affect the Hurst estimation procedure mentioned. However, when the signal to noise ratio (SNR) is relatively low, as is the case for RADAR signals, then the noise dominates in the signal-noise combination and the Hurst exponent estimated using the noisy signal as input, approximates the Hurst exponent of the corrupting noise. In fact, in a series of trials conducted using various types of simulated signals corrupted with fGn of different Hurst exponent values, the noise Hurst exponent estimation using the signal itself proved correct in 80% of the cases, when the SNR was -5dB or less. Since the Hurst exponent information is only needed to set the thresholds for the IMFs (as will be explained in the next section), this approach is adopted in the absence of a more accurate one.

5. EMD-HOS Method

EMD-HOS is a hybrid method which integrates the techniques explained in the previous sections, in an effort to effectively suppress noise in the signal and to increase the output SNR. The procedure starts by breaking down the noisy signal into IMF components. Then the HOS criteria are applied to detect the IMFs capturing only Gaussian noise and therefore they can be safely excluded from the final signal reconstruction process. However, even after removing the Gaussian IMFs, we generally assume that significant amount of noise power still exists in the remaining IMFs. This noise cannot be detected using HOS methods, due to mode mixing with the IMF components during the decomposition. Therefore, the only way of removing it is by applying thresholding techniques to the IMFs. Consequently, the aggregated variance method is used to get an estimate of the noise Hurst exponent, which in turn is applied to the noise model to evaluate the IMF energies, a necessary step for the IMF threshold estimation. For consistency with the noise model, Hurst exponent is constrained in three preset values as follows. When a value of 0.3 or below is estimated by the aggregated variance algorithm, a Hurst exponent value equal to 0.2 is input to the noise model. For calculated values inside the margin 0.3 - 0.7 the Hurst exponent is set to 0.5 and for every value over 0.7 the Hurst exponent is set to 0.8. The steps of the EMD-HOS method are listed below and depicted in flow chart form in Figure 1.

a. Apply the EMD and decompose the noisy signal in IMFs.
b. Get an estimate for the Hurst exponent of the signal noise, by applying the aggregated variance algorithm.
c. Calculate the Kurtosis values of the IMFs.
d. Use the Bootstrap technique to evaluate the minimum and maximum Kurtosis value for each IMF. The number of iterations is limited to 1000 for computational load purposes.
e. Examine the IMFs for gaussianity using the Kurtosis criterion. Specifically, check if the minimum and maximum estimated Kurtosis value is in the range specified by equation (9).
f. If the IMF passes the initial Kurtosis test, check its Bispectrum. For the IMF to be finally classified as Gaussian, its Bispectrum magnitude should lie inside the range of the Bootstrap estimated Bispectrum values of a simulated Gaussian noise signal having equal length with the original signal.
g. Scrap the IMF if it is classified as Gaussian.
h. Evaluate the noise variance using equation (17).
i. Evaluate the thresholds for the remaining IMFs using equations (13) and (15) and the estimated Hurst exponent value from step 2.
j. Apply the thresholding technique explained in section 4 using the estimated thresholds to the non-gaussian IMFs.
k. Reconstruct the signal adding the thresholded IMFs.

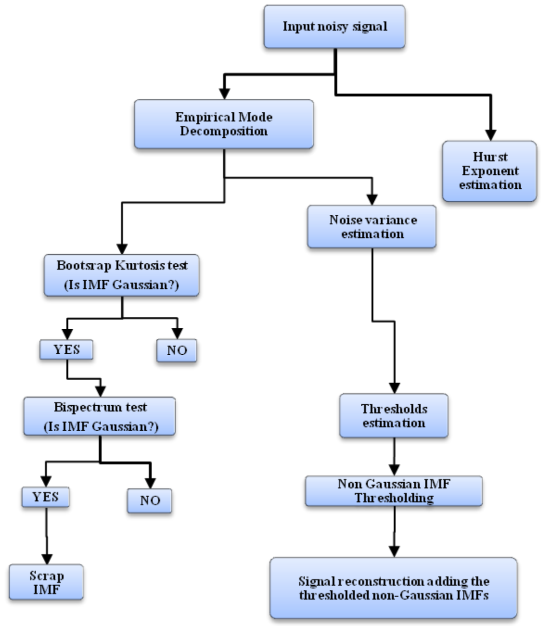


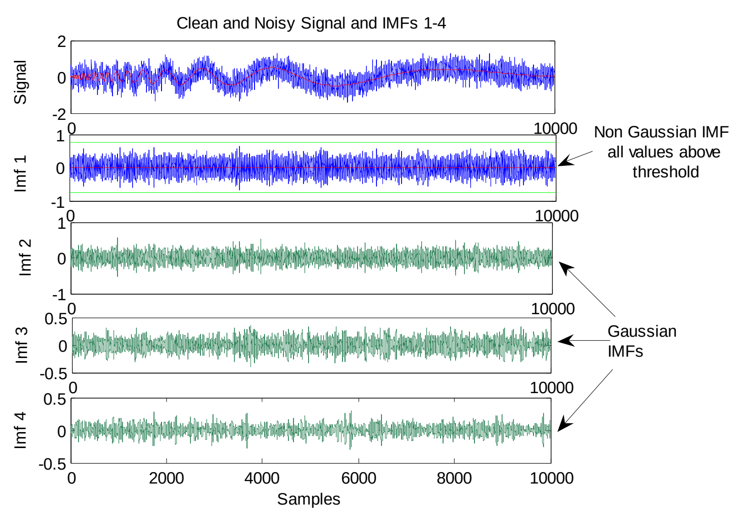
Figure 1: EMD-HOS Denoising Scheme

6. Experimental Results and Discussion

The new method was initially tested using two basic signals, ‘‘Bumps’’ and ‘‘Heavisine’’, since these signals are usually used in similar studies. Consequently three RADAR signals were used for testing; a plain Pulse signal, a Doppler signal, and a Linear FM (LFM) signal. The signals were tested in SNR values -10, -5, 0, 5 and 10dB, in an artificial noise environment. The number of signal samples for the Bumps, Heavisine Doppler and LFM signals was kept constant in all trials and equal to 10,000. Similarly, the number of signal samples for the pulse signal was 40,000. The Gaussian noise was simulated by differentiating a Fractional Brownian Motion (fBm) time series of different Hurst exponent values, that is; 0.5 for white noise and 0.2 and 0.8 for colored. The method was evaluated versus MATLAB’s Wavelets denoising menu, where various thresholding techniques are implemented as build-in functions, [22]. More specifically, the test signals were denoised applying HOS-EMD hard and soft thresholding, and Wavelets ‘‘rigrsure,’’ ‘‘heursure,’’ ‘‘sqtwolog’’and ‘‘minimax’’ rules, where 8th degree Symlets were used in a five level decomposition. To assess the results the output SNR, the Mean Squared Error (MSE) and the correlation coefficient value between the clean signal and the denoised version were used as figures of merit. The experimental results are portrayed in tabular form in Tables 1-9. In particular, Tables 1-5 depict the results for each one of the test signals when white noise of various SNR values was added. Consequently, Tables 6-9 show the results for the Doppler, Heavisine and Pulse signals in the presence of two different Hurst exponent correlated noise values of 0.2 and 0.8. For space purposes similar results for the rest of the test signals are omitted. In all tables fonts in bold denote best performance. Figure 2 presents an example of the algorithm application to a Doppler signal where, after EMD the signal IMFs are either removed or thresholded depending on the Gaussianity criteria. In Figures 3-7 the test signals are illustrated together with their denoised versions for selected SNR values, for both EMD-HOS and the best performing Wavelets technique for each particular signal. As shown, EMD-HOS outperforms Wavelets methods in most types of signals and combinations of SNR and

Hurst exponents. The only exception is in the LFM signal, where for SMRs lower than 0dB Wavelets perform better than EMD-HOS. It should be noted also that EMD-HOS hard thresholding performs better than EMD-HOS soft thresholding, with only a few exceptions to that.

Considering the tests results, EMD-HOS seems to be a promising technique for signal noise removal. However, the following particularities and shortfalls of the method should be addressed. EMD-HOS is slower than any of the Wavelets techniques. The EMD procedure itself is by nature computationally bulky, something that is further deteriorated by the overall complexity of the EMD-HOS algorithm. This problem can only be alleviated by restricting the signal samples to a considerably small number. The next drawback of the method is related with the significant amount of oversampling requirement, much more than the Nyquist limit for the EMD to perform properly. Although this issue has been addressed in [23] where the EMD sampling error is bounded, as a general rule, a fair amount of oversampling is required for the EMD to work effectively, especially in heavy noise environment. Unfortunately this problem amplifies the previous one, that is, if the signal samples are increased in order to achieve better EMD results, then the execution time of EMD-HOS will be also increased. The last shortfall of the EMD-HOS method is its relative inaccuracy on the Hurst exponent estimation. As mentioned when the signal is not noise dominant, which is the case for high SNR values, the Hurst exponent estimation using the noisy signal as input gives erroneous results, regarding the true Hurst exponent of the corrupting noise. In cases like this the type of noise in the signal should be a-priori estimated by experience and/or observation, and the Hurst exponent should be manually input to the algorithm. When there is no knowledge of the noise type, then a white noise case should be assumed in the signal and the thresholds should be manually adjusted by modifying the value of the threshold constant c in the equation (15), and the best value selected by inspection of the output signal in several trials. Nevertheless, understanding that EMD-HOS is not a panacea, the potential of the method in noise removal can be useful, when the aforementioned weaknesses are bounded within an acceptable limit depending on the application.



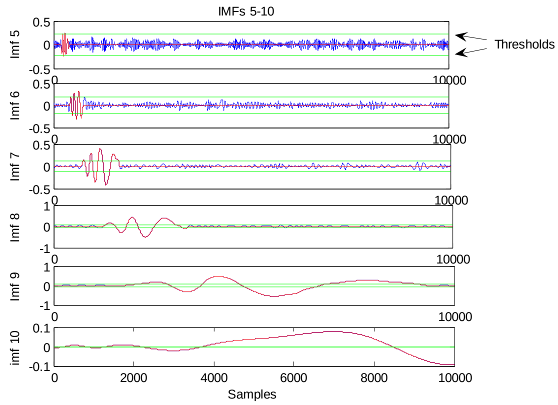


Figure 2: Clean and Noisy signal and the EMD-HOS denoising process. Original signal and IMFs are depicted in blue, Gaussian IMFs in dark green and thresholded IMFs in red. The thresholds are shown in light green.

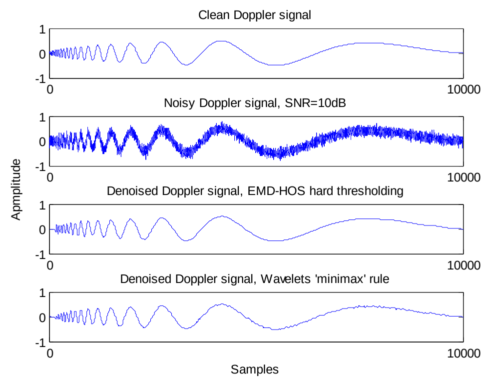


Figure 3: Doppler signal denoised with EMD-HOS and Wavelets ‘‘rigrsure’’ rule. Initial SNR=10dB

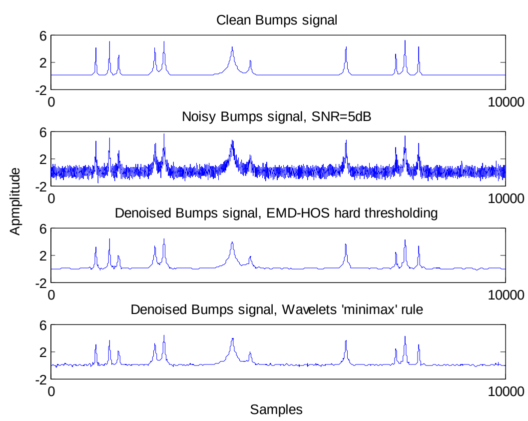


Figure 4: Bumps signal denoised with EMD-HOS and Wavelets Rigorous SURE rule. Initial SNR=5dB.

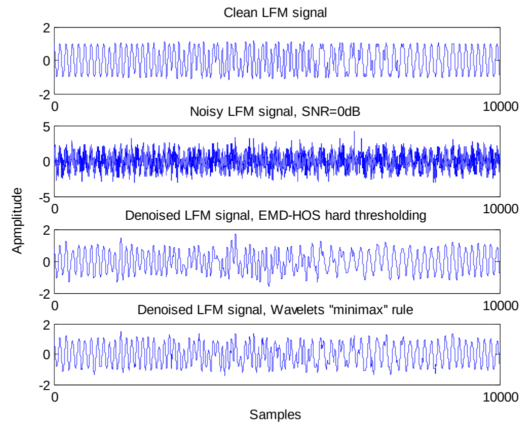


Figure 5: LFM signal denoised with EMD-HOS and Wavelets Minimax rule. Initial SNR=0dB.

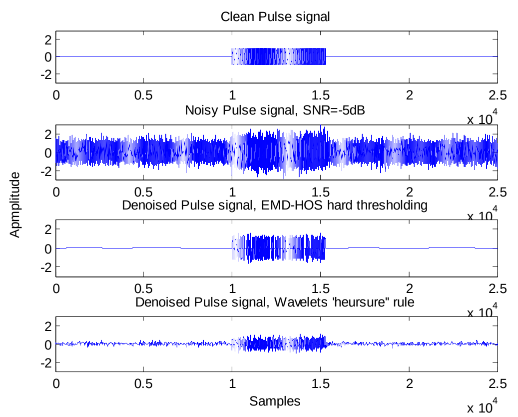


Figure 6: RADAR pulse signal denoised with EMD-HOS and Wavelets ''heursure'' rule. Initial SNR=-5dB.

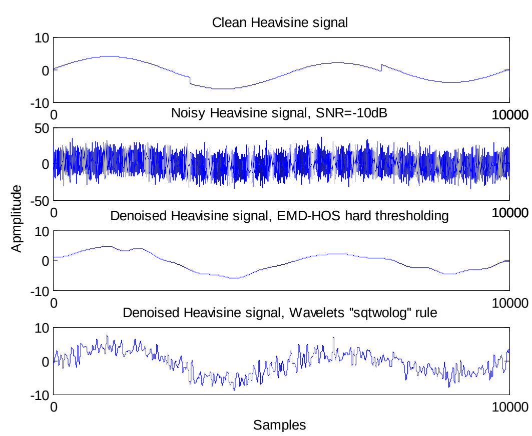


Figure 7: Heavisine signal denoised with EMD-HOS and Wavelets ''sqtwolog'' rule. Initial SNR=-10dB.

TABLE 1 Denoising results for Bumps signal corrupted by white noise in various SNR values.

Bumps Signal, White noise case							
		Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	10dB	0.0116	0.0148	0.0064	0.0047	**0.0036**	0.0099
	5dB	0.0157	0.178	0.0132	0.0105	**0.0093**	0.0257
	0dB	0.0247	0.029	0.0285	0.0225	0.0221	**0.0218**
	-5dB	0.7774	0.227	0.9383	0.0634	0.0605	**0.0519**
	-10dB	5.1362	3.7865	0.1848	0.189	**0.1458**	0.1649
SNR out	10dB	16.53	15.4	19.1	20.4	**21.5**	17.19
	5dB	15.21	14.66	15.96	16.97	**17.47**	13.06
	0dB	13.23	12.53	12.62	13.1	13.6	**13.78**
	-5dB	-1.73	3.06	9	9.14	9.34	**10**
	-10dB	-9.93	-8.61	4.49	4.4	**5.53**	4.99
Cor	10dB	0.9857	0.9857	0.9939	0.9955	**0.9966**	0.9914
	5dB	0.9848	0.9828	0.9873	0.9899	**0.991**	0.9793
	0dB	0.9762	0.9792	0.9725	0.9754	0.9782	**0.9793**
	-5dB	0.6203	0.8234	0.9383	0.9402	0.9432	**0.9494**
	-10dB	0.3086	0.3521	0.8569	0.8547	**0.8612**	0.8295

TABLE 2 Denoising results for Heavisine signal corrupted by white noise in various SNR values.

Heavisine Signal, White noise case							
		Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	10dB	0.0324	0.0328	0.0331	0.0332	**0.0146**	0.0157
	5dB	2.9814	1.648	0.0961	0.982	**0.0308**	0.0354
	0dB	9.7227	8.4981	0.2882	0.2913	0.0789	**0.0691**
	-5dB	30.4	29	0.9532	0.9509	**0.117**	0.2
	-10dB	94.63	93.35	3.06	3.09	0.4925	**0.4119**
SNR out	10dB	24.6	24.6	24.59	24.58	**28.1**	27.83
	5dB	5	7.6	19.96	19.8	**24.9**	24.3
	0dB	-0.0887	0.495	15.18	15.14	20.81	**21.39**
	-5dB	-5	-4.84	10	10	**19.1**	16.77
	-10dB	-9.97	-9.91	4.92	4.88	12.86	**13.64**
Cor	10dB	0.9983	0.9983	0.9983	0.9983	**0.9992**	0.9992
	5dB	0.8729	0.9235	0.995	0.9949	**0.9984**	0.9982
	0dB	0.7027	0.7262	0.9851	0.985	0.9959	**0.9964**
	-5dB	0.4831	0.4912	0.9532	0.9523	**0.9944**	0.9895
	-10dB	0.3011	0.303	0.8689	0.8678	0.9746	**0.9781**

TABLE 3 Denoising results for Doppler signal corrupted by white noise in various SNR values

Doppler Signal, White noise case							
		Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	10dB	0.0004	0.0004	0.0004	0.00037	**0.00021**	0.0009
	5dB	0.0009	0.0009	0.0009	0.0009	**0.0005**	0.002
	0dB	0.0031	0.0031	0.0031	0.0031	**0.0014**	0.0034
	-5dB	0.0089	0.0089	0.0089	0.0092	**0.0037**	0.0079
	-10dB	0.0298	0.0286	0.0286	0.0295	**0.0066**	0.0112
SNR out	10dB	22.57	22.57	23.28	23.59	**25.96**	19.76
	5dB	19.37	19.37	19.45	19.53	**22.24**	16.35
	0dB	14.43	14.43	14.43	14.36	**17.79**	14
	-5dB	9.83	9.83	9.83	9.67	**13.62**	10.35
	-10dB	4.6	4.77	4.77	4.63	**11.14**	8.85
Cor	10dB	0.9972	0.9972	0.9977	0.9978	**0.9988**	0.9957
	5dB	0.9942	0.9942	0.9943	0.9944	**0.997**	0.9914
	0dB	0.9832	0.9832	0.9832	0.983	**0.9919**	0.9815
	-5dB	0.95	0.95	0.95	0.94	**0.9781**	0.953
	-10dB	0.8589	0.8634	0.8634	0.8599	**0.9608**	0.9334

TABLE 4 Denoising results for Pulse signal corrupted by white noise in various SNR values

Pulse Signal, White noise case							
		Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	10dB	0.0158	0.0181	0.0143	0.0079	**0.0029**	0.0167
	5dB	0.0181	0.0221	0.0324	0.0207	**0.0063**	0.0322
	0dB	0.0284	0.0338	0.0697	0.0459	**0.0225**	0.0735
	-5dB	0.0656	0.0598	0.1576	0.1053	**0.0412**	0.1247
	-10dB	0.0929	0.0767	0.1339	0.1133	**0.0716**	0.0726
SNR out	10dB	11.21	10.62	11.65	14.23	**18.57**	10.96
	5dB	10.63	10.63	8	10	**15.23**	8.12
	0dB	8.66	7.9	4.7	6.58	**11.12**	4.53
	-5dB	5	5.43	1.22	2.9	**7**	2.24
	-10dB	0.51	1.5	-1	-0.35	**1.63**	1.58
Cor	10dB	0.9615	0.96	0.9765	0.9868	**0.9931**	0.9904
	5dB	0.956	0.9527	0.9465	0.9626	**0.985**	0.9754
	0dB	0.9319	0.9234	0.8696	0.9084	**0.9623**	0.9455
	-5dB	0.8538	0.8457	0.5031	0.7219	**0.9011**	0.793
	-10dB	0.6274	0.6378	0.0251	0.2314	**0.6391**	0.6300

TABLE 5 Denoising results for LFM signal corrupted by white noise in various SNR values

LFM signal, White noise case							
		Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	10dB	0.0197	0.0197	0.0099	0.0067	**0.0051**	0.0161
	5dB	0.0234	0.0234	0.0192	0.0152	**0.0138**	0.0221
	0dB	0.0326	0.0399	0.0337	0.0379	**0.0306**	0.0664
	-5dB	0.0926	**0.063**	0.0667	0.0673	0.1089	0.1575
	-10dB	4.9748	3.7842	**0.1875**	0.1896	0.3042	0.3408
SNR out	10dB	13.98	13.98	16.9	18.64	**19.88**	14.86
	5dB	13.22	13.22	14.09	15.11	**15.51**	13.47
	0dB	11.61	11.62	11.65	11.26	**11.96**	8.7
	-5dB	7.26	**8.72**	8.68	8.65	6.55	4.95
	-10dB	-10	-8.8	**4.19**	4.14	2.094	1.6042
Cor	10dB	0.9798	0.9798	0.9899	0.9931	**0.9949**	0.991
	5dB	0.9761	0.9761	0.9846	0.9846	**0.986**	0.978
	0dB	0.9652	0.9653	0.9655	0.9624	**0.9679**	0.9374
	-5dB	0.9368	**0.9368**	0.9362	0.9358	0.8944	0.8564
	-10dB	0.2918	0.329	**0.8373**	0.8359	0.7039	0.5758

TABLE 6 Denoising results for Heavisine signal corrupted by colored noise SNR=-5dB

Heavisine Signal SNR=-5dB, Colored noise case, SNR=-5dB							
	Hurst	Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	0.2	0.0023	0.0023	0.0023	0.0028	0.0014	**0.0011**
	0.8	0.0139	0.0139	0.0141	0.0151	**0.0024**	0.006
SNR out	0.2	19.63	19.63	19.63	18.81	22	**23**
	0.8	1.21	1.21	1.17	0.87	**8.89**	4.48
Cor	0.2	0.9946	0.9946	0.9946	0.9935	0.9969	**0.9975**
	0.8	0.6996	0.6996	0.6978	0.6851	**0.9343**	0.8666

TABLE 7 Denoising results for Doppler signal corrupted by colored noise SNR=-5dB

Doppler Signal SNR=-5dB, Colored noise case, SNR=-5dB							
	Hurst	Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	0.2	0.003	0.003	0.003	0.0033	**0.002**	0.006
	0.8	0.0135	0.0135	0.0137	0.0147	**0.0047**	0.108
SNR out	0.2	19	19	19	18.67	**20**	16
	0.8	1.28	1.28	1.22	0.9	**7.11**	3.5
Cor	0.2	0.9938	0.9938	0.9938	0.9933	**0.9961**	0.9908
	0.8	0.7822	0.7822	0.7804	0.769	**0.8978**	0.7786

TABLE 8 Denoising results for Pulse signal corrupted by colored noise SNR=-5dB

Pulse Signal, Colored Noise case, SNR=-5dB							
	Hurst	Wav Heursure Hard	Wav Heursure Soft	Wav Universal	Wav Minimax	HOS-EMD Hard	HOS-EMD Soft
MSE	0.2	0.0256	0.0321	0.1826	0.1127	**0.024**	0.075
	0.8	0.034	0.034	0.0206	0.0206	**0.0121**	0.0167
SNR out	0.2	9.34	8.34	0.8	2.89	**9.61**	4.66
	0.8	-2.43	-2.43	-0.64	-0.26	**2**	0.64
Cor	0.2	0.9413	0.9381	0.555	0.8279	**0.9454**	0.8895
	0.8	-0.0009	-0.0009	0.4275	0.5302	**0.6956**	0.3845

7. Conclusions

A hybrid signal denoising technique is presented based on Empirical Mode Decomposition and Higher Order Statistics. The adaptive nature of EMD in signal decomposition is combined with the applicability of HOS in detecting gaussianity. Wavelets thresholding techniques recently applied in EMD are also used for maximum noise suppression. The experimental results are in favour of the new method, when compared with traditional Wavelets methods. Despite the fact that it is computationally demanding and lacks accuracy when applied in signals corrupted by colored Gaussian noise, EMD-HOS should be considered as a promising method for signal denoising.

References

[1]. Galal E. Nadim, “Detection and Localization of Radar Pulses in Severe Noisy Environment Using Wavelet Packets Transform Combined with Higher-Order-Statistics Thresholding Technique’’, GSPx 2006, The 4th International Signal Processing Conference, 30 October – 2 November 2006, Santa Clara, California, USA.

[2]. N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, Q. Shin, H.H. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, “The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis’’, Proc. Royal Soc. London A, vol. 454, pp. 903-995, 1998.

[3]. Z. Wu, and N. E. Huang, “A study of the characteristics of white noise using the empirical mode decomposition metho’’, Proc. Roy. Soc. London A, vol. 460, pp. 1597–1611, Jun. 2004.

[4]. P. Flandrin, P. Gonçalvès, and G. Rilling, “Detrending and denoising with Empirical Mode Decompositions’’, Eusipco, 12th European Signal Processing Conference, September 6-10 Vienna, Austria, 2004.

[5]. P. Flandrin, G. Rilling, and P. Gonçalvès, “EMD equivalent filter banks, from interpetation to applications, in Hilbert-Huang Transform and Its Applications’’, N. E. Huang and S. Shen, Eds., 1st ed. Singapore: World Scientific, 2005.

[6]. A. O. Boudraa, and J. C. Cexus, “Denoising via empirical mode decomposition’’, in Proc. ISCCSP 2006, 2006.

[7]. B. Weng, M. Blanco-Velasco, and K.E. Barner, “ECG Denoising Based on the Empirical Mode Decomposition’’, Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006.

[8]. K. Khaldi, M. Turki-Hadj Alouane, A.-O. Boudraa, "A new EMD denoising approach dedicated to voiced speech signals'', Signals, Circuits and Systems, 2008. SCS 2008. 2nd International Conference on Volume, Issue 7-9 Page(s): 1 – 5, Nov. 2008.
[9]. G.S Tsolis, T.D. Xenos, "Seismo-ionospheric coupling Correlation Analysis of earthquakes in Greece'' using Empirical Mode Decomposition, Nonlin. Processes Geophys., 16, 123-130, 2009.
[10]. L. Hadjileontiadis, "Empirical mode decomposition and fractal dimension filter'', IEEE Eng. Med. Biol. Mag., pp. 30–39, Jan. 2007.
[11]. Y. Kopsinis, S. McLaughlin, "Development of EMD Based Denoising Methods Inspired by Wavelet Thresholding'', IEEE Transactions on Signal Processing 57, pp. 1351-1362, 2009.
[12]. P. Ravier and P. O. Amblard, "Wavelet packets and denoising based on higher-order-statistics for transient detection'', Signal Processing, Vol. 81, No. 9, 1909–1926, August 2001.
[13]. W. Zhang, X.H. Zhao, "Wavelet Thresholding using Higher Order Statistics for Signal Denoising'', Info-tech and Info-net, 2001. Proceedings. ICII 2001 - Beijing. 2001 International Conferences on, page(s): 803-808 vol.2, 2001.
[14]. H.E. Kan, R.D. Hippenstiel, M.P. Fargues, "Bootstrap-based signal denoising'', page(s): 958- 962 vol.1, Nov. 2002.
[15]. A. Swami, G. B. Giannakis, G. Zhou, "Bibliography on higher-order statistics'', Signal Processing, 60 (1), pp. 65-126, 1997
[16]. A.M. Zouhir, and B. Boashash, "The Bootstrap and its Application in Signal Processing'', IEEE Signal Processing Magazine, Vol. 15, No. 1,pp. 56-76, Jan. 1998.
[17]. A. Al-Smadi, "Tests for Gaussianity of a Stationary Time Series'', World Academy of Science, Engineering and Technology 10, 2005.
[18]. C.L. Nikias and M.R. Raghuveer, "Bispectrumn Estimation: A Digital Signal Processing Framework'', Proc. IEEE, Vol. 75, No.7, pp. 869-891, July 1987.
[19]. D.L. Donoho, "De-noising by soft-thresholding'', IEEE Trans. Inform. Theory, vol. 41, no. 3, pp. 613-627, 1995.
[20]. F. Esposti, M. G. Signorini, "Evaluation of a Blind Method for the Estimation of the Hurst's exponent in time series'', 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006.
[21]. H.J Blok, "On the nature of the stock market: Simulations and experiments'', PhD THESIS on Statistical Mechanics, arxiv.org/abs/cond-mat/0010211v1, Oct. 2000.
[22].MATLAB's Wavelet Toolbox 4, Users Guide, http://www.mathworks.com/ access/helpdesk/help/ toolbox/wavelet/waveleta.html
[23]. G. Rilling, P. Flandrin, "On the influence of sampling on the Empirical Mode Decomposition'', in Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-06, Paper SPTM-P5.11, Toulouse (F), 2006.

Authors

George S. Tsolis received a Diploma from the Pilot School Department of Helenic Air Force Academy, and a M. Sc. degree in Electronic Systems Engineering from Naval Postgraduate School, CA, USA. He is currently pursuing a Ph.D. degree at Electrical Engineering Department of the Aristotle University of Thessaloniki, Greece. His research interests include, RADAR signal processing, and Electronic Warfare systems engineering.

Thomas D. Xenos received a Diploma in Electrical Engineering from the University of Patras and a Ph.D in Radio Wave Propagation from the Aristotle University of Thessaloniki. He is an Electrical Engineering Professor at the Aristotle University of Thessaloniki. He took part in many national and international projects and he has authored over 90 scientific papers. His research interests include ionospheric prediction methods and models, wireless communications, radiowave propagation, non-ionizing electromagnetic radiation measurements and electromagnetic compatibility.