Fall 2017

Statistical Properties and Applications of Empirical Mode Decomposition

Mahdi Hameed Al-Badrawi
University of New Hampshire, Durham

Statistical Properties and Applications of Empirical Mode Decomposition

BY

Mahdi H. Al-Badrawi

Bachelor of Science in Electrical Engineering, University of Baghdad, 2005

Master of Science in Communications and Computer Engineering, National University of Malaysia, 2011

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

September, 2017

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering by:

Dissertation Director, Nicholas J. Kirsch, Ph.D.
Associate Professor (Electrical and Computer Engineering)

W. Thomas Miller III, Ph.D.
Professor (Electrical and Computer Engineering)

Andrew L. Kun, Ph.D.
Associate Professor (Electrical and Computer Engineering)

Sudharman K. Jayaweera, Ph.D.
Professor (University of New Mexico, Electrical and Computer Engineering)

Bessam Z. Al-Jewad, Ph.D.
Group Manager (Globus Medical, Inc.)

on 6/16/2017.

Original approval signatures are on file with the University of New Hampshire Graduate School.

DEDICATION

I dedicate this dissertation to my large family and my small family.

ACKNOWLEDGEMENT

First and foremost, glorification and gratitude to Allah, the Almighty, for the blessings, health, and strength to complete my research successfully.

I would like to express my deep and sincere thanks to my dissertation advisor Prof. Nicholas J. Kirsch for all support during the past years starting from the first moment when he believed in me to be his Ph.D student. His guidance established a strong foundation and high standard of thinking for my outcome. I would like to acknowledge his spiritual and financial support without them it was difficult to process the doctoral program.

I would like to thank Dr. Bessam Al-Jewad, my committee member, the great friend and the truthful mate for all the years standing beside me in all good and hard moments. My thanks are dedicated to my committee members Prof. W. Thomas Miller III, and Prof. Andrew L. Kun for their valuable comments and their nice attitude for all years I know them. Special thanks to Prof. Sudharman K. Jayaweera from the University of New Mexico for being part of my dissertation committee and for his valuable comments and collaboration.

I also appreciate the efforts made by the department of Electrical and Computer Engineering to support my research through the teaching assistantship, mentioning here the department chair Prof. Kent A. Chamberlin and the administrative assistant Kathy Reynolds and the Engineering Technician Jim Abare.

I would like to say many thanks to my research colleagues, teachers, and friends from home country for being always supportive and standing when need them.

I'm grateful to my large family including my dad and mom and my sisters and brothers and to my cousin and her wonderful family and my relatives for all invaluable commitment and being always the best backup in this life, thank you for standing to each other in all times.

Finally, I would thanks and appreciate highly my wife efforts over the past four years, you have been always there when I needed, thanks for your patience and your emotional support. Also would like to thank my daughters, you are guys the light of my road.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Statistical Properties and Applications of Empirical Mode Decomposition

by

Mahdi H. Al-Badrawi

University of New Hampshire, September, 2017

Signal analysis is key to extracting information buried in noise. The decomposition of signal is a data analysis tool for determining the underlying physical components of a processed data set. However, conventional signal decomposition approaches such as wavelet analysis, Wagner-Ville, and various short-time Fourier spectrograms are inadequate to process real world signals. Moreover, most of the given techniques require *a prior* knowledge of the processed signal, to select the proper decomposition basis, which makes them improper for a wide range of practical applications. Empirical Mode Decomposition (EMD) is a non-parametric and adaptive basis driver that is capable of breaking-down non-linear, non-stationary signals into an intrinsic and finite components called Intrinsic Mode Functions (IMF). In addition, EMD approximates a dyadic filter that isolates high frequency components, e.g. noise, in higher index IMFs. Despite of being widely used in different applications, EMD is an ad hoc solution. The adaptive performance of EMD comes at the expense of formulating a theoretical base. Therefore, numerical analysis is usually adopted in literature to interpret the behavior.

This dissertation involves investigating statistical properties of EMD and utilizing the

outcome to enhance the performance of signal de-noising and spectrum sensing systems. The novel contributions can be broadly summarized in three categories: a statistical analysis of the probability distributions of the IMFs and a suggestion of Generalized Gaussian distribution (GGD) as a best fit distribution; a de-noising scheme based on a null-hypothesis of IMFs utilizing the unique filter behavior of EMD; and a novel noise estimation approach that is used to shift semi-blind spectrum sensing techniques into fully-blind ones based on the first IMF. These contributions are justified statistically and analytically and include comparison with other state of art techniques.

CHAPTER 1

Introduction

1.1 Motivation

A signal is the result of a sensor measurement or readings and it consists of useful information embedded in some level of noise. Signal examples include but are not limited to biomedical, electrical faults, health monitoring, Global Positioning System (GPS), seismic, climate, wireless communication, and mechanical vibrations. The signal received from the aforementioned examples is the link to understand the underlying features of a given process, in which signal analysis is the interpretation key of that link. On the other hand, the use of a proper signal analysis technique will result in a significant support in making a valid decision such as estimation, detection, and recognition [3]. Further, signal interpretation is necessary whether the particular measurement carry useful information or just a noise. For example, analyzing a wireless signal at a certain frequency band can help greatly to decide if that band is carrying any information or not and that is the basic concept of spectrum sensing techniques.

Practically, signals are multi-component, non-linear and non-stationary in nature, therefore, decomposition process is the way to extract its physical meaningful components. In that sense, separating, or decomposing, such multicomponent to their basic scales is rather essential to reveal the underlying physical characteristics of these measurements. Conventional decomposition techniques such as wavelet analysis, Wagner-Ville, and various short-time Fourier spectrograms are incapable of processing such signals efficiently. In addition, conventional techniques can either work effectively for non-linear signals or for non-stationary signals, but none are able to accommodate for both [4]. The robustness of a decomposition method can be examined by its ability to:

- effectively decompose non-linear and non-stationary signals.
- adaptively derive the decomposition basis out of the signal itself.
- extrat the signal features in the presence of noise.
- represent the signal frequencies locally, i.e. at their time of occurence.

Generally, Fourier Spectral analysis is the most widely used approach in data interpretation due to its simplicity in both theory and applicability. Nevertheless, the performance of Fourier analysis approaches is captivated by several limitations such as linearity and stationarity of the processed data. Moreover, and in terms of signal representation, Fourier approaches, such as Short-Time Fourier Transform STFT, (an extension of the basic Fourier transform with short time sliding window), lack the adaptive behavior as it is based on *a priori* basis and suffer from energy leakage. The Wigner-Ville function is another example of Fourier analysis in which the window is shifted copy of the processed signal itself. Because this method (Wigner-Ville) is derived from Fourier analysis, thus it suffers from all constraints that Fourier analysis exhibit [5].

On the other hand, wavelet approaches represent an advanced version of Fourier Spectral methods with a predefined window in which each window (wavelet function) can be adjusted for a specific application. Although wavelets have been proved with mathematical rigor, versatile for different applications, and have the ability to analyze non-stationary signal, wavelet methods are non-adaptive in which the decomposition basis are selected *a priori*. Empirical Orthogonal Function (EOF) is an example of *a posteriori* basis decomposition example where the basis are derived from the data [6]. In this method, the basis are obtained by calculating the eigenvectors from the covariance matrix of the processed data. Despite of the orthogonality of EOF basis, its output modes do not guarantee a meaningful representation of the data in which even random processes might reveal falsely a physical pattern [7].

Thus, in order to alleviate the drawbacks of the aforementioned decomposition methods, the Empirical Mode Decomposition (EMD), which is the key component of Hilbert-Huang Transform (HHT), is adopted in this work [8]. EMD is an adaptive data driven tool with basis that are defined *a posteriori* and a capability to analyze non-linear and non-stationary signals effectively. Beside its ability to extract signal features, EMD has the advantage of decomposing signals locally and that is becomes significant when the signal varies with time. The adaptive behavior of EMD comes at the expense of formulating a mathematical frame, therefore this approach is empirical.

Fig. 1.1 illustrates a comparative diagram of different data analysis methods with respect to their adaptivity, robustness and complexity. The term, adaptivity, refers to ability of a method to derive its basis from the processed signal itself and not *a priori*. On the other hand, the term, robustness, refers to the ability of a method to extract signal features regardless of the non-linearity and non-stationarity of the processed signal. Further, the term, complexity, refers to the arithmetic operations of an algorithm to process N samples. From Fig. 1.1, the complexity level for Fourier transform [9], and Wavelet transform [10] are shown to be less than the EOF [11] and EMD [12] and close to Wigner-ville distribution [13].

1.1.1 Significance of EMD Statistical Properties

In recent years, EMD gained a lot of attention and has been applied in different areas because of the unique properties and the superior performance. Despite its use, successfully, in various applications, EMD still lacks a theoretical basis and is defined as an algorithm with no analytical interpretation [14]. Therefore, it is essential to understand its statistical properties and analyze its behavior to be utilized in real world applications.

Applying EMD algorithm on a time-series signal will result in a complete and finite set of components called Intrinsic Mode Functions (IMFs). The statistical characteristics of these IMFs are studied in literature and it has been conducted to have a significant influence on the way EMD is applied for various applications [3]. One of the important aspects of

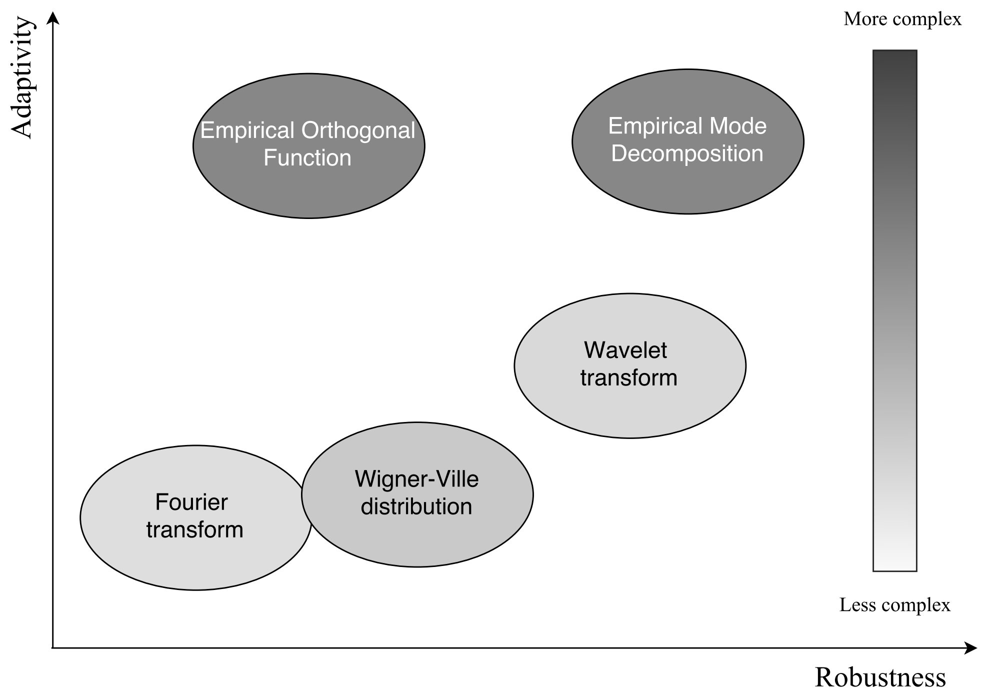

Figure 1.1: A comparative diagram to classify different data analysis methods based on their adaptivity, robustness and complexity level.

understanding the behavior of EMD process led to conclude that the EMD resembles a dyadic filter (similar to wavelet) in which IMFs play the role of overlapping bandpass filters [15]. Thereafter, the filter bank properties make it more obvious that the findings of other filter bank techniques can be applied in EMD effectively. On the other hand, the investigation of the probability distribution functions of the IMFs supports such filtering features and provides insight to develop more advanced adaptive filtering methods [16].

1.1.2 Significance of Signal De-noising

The aim of de-noising is to recover useful information buried in noisy measurement processes while retain its features unaltered. However, noise is an inevitable part of any real collected data and might vary from highly correlated to uncorrelated with the corrupted signal. Different filters have been proposed such as linear filters and transform-based methods. However, these techniques are vulnerable to signal changing characteristics and/or prior filter parameter assumptions [16].

Filtering through EMD comes quite natural, in which the relevant modes corresponding to the signal components (almost noise free) are separated during the sifting operation. However, selecting the relevant IMFs is a critical process especially in the presence of high noise power. Unfortunately, analytical expressions of the signal IMFs are not available. Thus, most of the important results on the EMD are all based on the empirically determined findings from numerical experiments.

1.1.3 Significance of Signal Detection

Signal detection is the process of deciding whether a certain measure (either in time or frequency domain) is occupied with information or if it comprises noise only. Cognitive radio is one of the most contemporary approaches that is relied on the concept of detection theory. Cognitive radio enables the use of radio spectrum that is underutilized due to the fix allocation of frequency bands. In the physical layer of cognitive radio, spectrum sensing plays a key role in detecting the holes (vacant channels) for a given band. Further, energy detection (ED) is used widely for spectrum sensing purposes due to its non-coherent nature and low computational complexity [17].

More recently, Empirical Mode Decomposition (EMD) has been proposed as a detection method for wireless applications. EMD works adaptively and blindly to decompose time-series signals into a set of modes (IMFs) that can be utilized for detection purposes. Similar to ED, EMD behaves non coherently towards the received signals in which it requires no prior information about the signal characteristics.

1.1.4 Significance of Noise Power Estimation

In practical scenarios, signals are corrupted with noise unavoidably. Further, the noise power characteristics is unknown at the receiver and range from very mild (high SNR) to very severe (low SNR). If the noise power is incorrectly assumed, the spectrum energy detector for example, may result in a significant degradation in performance and might yield in false decisions. Therefore, noise power estimation is required to mitigate the impact of the presumptions and its consequences.

Empirical Mode Decomposition algorithm can be used for noise power estimation because it has a unique ability to process noise. In wireless communications, an accurate estimation of the noise power could enhance the quality of the communication link by tuning of the transmitter and receiver parameters. Further, the knowledge of the estimated noise power can be used to calculate the SNR which is an indication of the communication channel reliability.

1.2 Dissertation Contributions

In this work, the aim is to investigate the EMD sifting process and represent the IMF's probability distribution functions in more general representation in which the existing fitting distributions are seen as special cases. The understanding of IMF probability distributions is used for de-noising purposes in which an adaptive filter based on EMD characteristics is designed. In addition, the first IMF statistical features is utilized to propose a blind noise estimation approach that can be applied to shift semi-blind spectrum sensing methods to fully-blind ones. In this sense, the contributions of our research can be outlined as follows:

1. Modeling the Intrinsic Mode Functions of input signal distributions using Generalized Gaussian distribution (GGD) and validate that model via null hypothesis test [18].

2. Investigating of the IMFs statistical properties when random variables of different distributions are decomposed by EMD. The outcome of this investigation led to an important result that for an unknown distributed signal, the generated IMFs will be a set of GGD modes and residual.

3. Proposing a novel scheme for signal de-noising based on the null hypothesis of the Generalized Gaussian distribution of Intrinsic Mode Functions [18]. This includes an

evaluation of IMF models including the well-known Gaussian, Laplacian and the GGD in different scenarios and compare these models from different perspectives.

4. Proposing a multi-channel detection methods based on the behavior of Intrinsic Mode Function energies in frequency domain [1, 2]. Further, analyzing the properties of the first Intrinsic Mode Function and use its unique features for signal detection [19, 20].

5. Proposed a novel approach for noise estimation using the characteristics of the first Intrinsic Mode Function and apply that approach to enhance the performance of noise-dependent spectrum sensing techniques such as energy detectors [21].

1.3 Dissertation Layout

This dissertation is outlined as follows:

- Chapter 2 addresses the theoretical background of the Empirical Mode Decomposition algorithm associated with sifting process parameters. Further, examples on the ability of EMD to sift different scales out of the composite signal is presented. On the other hand, a literature review of the EMD statistical properties is outlined followed by review of signal de-noising and spectrum sensing methods.

- Chapter 3 includes a statistical analysis of the Intrinsic Mode Functions and an investigation of its probability distribution functions. From the statistical findings, a de-noising scheme based on the null hypothesis of IMFs is proposed and examined in different scenarios for various types of signals.

- Chapter 4 covers the use of EMD properties in filtering to propose an adaptive spectrum sensing technique. The proposed technique is blind in which it does not required a predefined parameters or a knowledge of the received signal characteristics such as the noise variance. The performance of the proposed technique is evaluated under low signal to noise ratio regimes and different sampling rates.

- Chapter 5 includes a novel noise power estimation method based on the scaling of the first IMF and use that knowledge of the noise power to shift the semi-blind spectrum sensing techniques to blind ones. The proposed method is explained analytically and demonstrated to exhibit good performance in contrast to the other noise power estimation techniques.

- Chapter 6 covers the conclusion remarks of this dissertation including the findings and their importance in addition to the future work.

CHAPTER 2

Background

In this chapter, the pertinent background to the performed and the proposed contributions are presented. The theoretical definition of the Empirical Mode Decomposition algorithm and its sifting process is addressed with illustration examples. A literature review of all recent and relevant findings in the EMD domain and it applications in both de-noising and detection is outlined.

2.1 Empirical Mode Decomposition: Theoretical Background

2.1.1 EMD Algorithm

Empirical Mode Decomposition is an adaptive approach that is used to analyze non-linear and non-stationary signals. This technique decomposes time-series signals, through the sifting process, into a complete and finite set of amplitude and frequency modulation (AM-FM) oscillatory components called Intrinsic Mode Functions (IMFs). Sifting process is the key concept of the EMD in which the iteration will continue over the processed signal until a stoppage criteria is satisfied. However, the number of iterations over each sifted IMF is a function of the signal length and its smoothness. The IMF must satisfy the following two conditions: (1) the number of extrema and the number of zero-crossings must be either equal or differ at most by one for the entire data set. (2) The mean value of the envelope defined by the local maxima and by the local minima is zero, at any point.

If $y(n)$ was the processed signal, where n is the sample index, thus the EMD algorithm procedure can be summarized as follows:

1. Initialize the input $r(n)$ as $y(n)$ (the residue signal).

2. Identify extrema points of $r(n)$: maxima and minima.

3. Interpolate maxima and minima points to form the upper and lower envelopes $e_{\max}(n)$ and $e_{\min}(n)$ respectively.

4. Evaluate the mean: $m(n) = (e_{\min}(n) + e_{\max}(n))/2$.

5. Extract the detailed signal: $h(n) = r(n) - m(n)$.

6. If $h(n)$ does not satisfy the stoppage criteria, then the process is repeated and $h(n)$ is the input to step (2).

7. If $h(n)$ satisfies the stoppage criteria, then $h(n)$ is the j^{th} IMF. The residue is $y(n) = r(n) - \text{IMF}_j(n)$. If the number of zero crossings of the residue <2, then break the process and keep the last collected signal as a trend. Otherwise, back to step (1) with the residue as the input.

The original input signal can be reconstructed as the sum of the IMFs and the trend such that:

$$\hat{y}(n) = \mathcal{T}(n) + \sum_{i=1}^{M} \text{IMF}_i(n) \tag{2.1}$$

where, $\hat{y}(n)$ is the reconstructed signal, $\mathcal{T}(n)$ is the trend of $y(n)$ (or residual) and M is the number of sifted IMFs.

2.1.2 Sifting process parameters

EMD is an algorithmic-based technique that does not hold a rigor mathematical definition. However, the sifting process represents a unique and powerful phase of EMD that compensated the lacking of theoretical framework.

In the sifting process, two parameters are mainly investigated in the literature and they are given as follows:

Stoppage criteria

The stoppage criteria is suggested to halt the sifting process at a point that ensure a completeness of the physical meaning for the extracted IMF. In that sense, several stoppage criteria have been proposed [8, 22, 23].

In [8], Cauchy convergence test, which is defined as a normalized squared difference (SD) between two successive sifting operations, is used such that:

$$SD_k = \frac{\sum_{n=0}^{N} |h_{k-1}(n) - h_k(n)|^2}{\sum_{n=0}^{N} h_{k-1}^2} \tag{2.2}$$

where h_k is the k^{th} extracted signal through the sifting, and k is the iteration number. If SD_k is smaller than the predetermined value ($0.2 \sim 0.3$), the sifting process will be stopped.

In [22], Huang *et al.* suggested another stoppage criteria called S-number. S-number is a pre-selected parameter that is is used to stop the sifting process whenever the number of zero-crossings and extrema stay same or almost differ by one after S-consecutive times. Generally, S-Number selection is ad-hoc, however, Huang *et al.* established an empirical guide, and he found that for optimal sifting, the range of S-number should be set between 4 and 8.

In [23], the authors proposed two thresholds (θ_1 and θ_2) approach that endeavor to retain locally large deviations and globally small mean oscillations of the extracted IMF. For that purpose, the authors introduced two terms namely: mode amplitude $a(n)$, and evaluation function $\eta(n)$, such that:

$$a(n) = \frac{e_{max}(n) - e_{min}(n)}{2}$$
$$\eta(n) = |\frac{m(n)}{a(n)}|$$

Herein, the sifting process will iterate until the following condition is satisfied:

$$\begin{cases} \eta(n) < \theta_1 & : (1-\vartheta) \;\text{ of total IMF duration} \\ \eta(n) < \theta_2 & : \vartheta \;\text{ of total IMF duration} \end{cases}$$

where the authors set $\vartheta \approx 0.05$, $\theta_1 \approx 0.05$ and $\theta_2 \approx 10\theta_1$ as a default values.

Interpolation approaches

In EMD domain, the interpolation is the process of connecting the identified extrema (maxima/minima) to form the upper and lower envelopes respectively. To this end, different interpolation approaches have been proposed aiming for more efficient decomposition process. In [8], Huang *et al.* suggested the use of cubic spline interpolation to fit all maxima (resp. minima) data points. However, Huang *et al.* conducted that higher order splines required extra computational time and more parameters have to be determined and that is discrediting the adaptivity of the approach. In [24], the authors proposed the rational splines that include the cubic spline as a special case, however, this method is performing better if only the pole parameter that controls the tautness of the spline, is tuned carefully. In [25], raised cosine interpolation is proposed to provide more efficient implementation as the proposed approach is based on fast Fourier transform. However, this method of interpolation is dependent on the roll-off factor and the sampling period which in turns is signal-dependent leading to not a fully adaptive EMD approach.

2.1.3 Examples on EMD sifting and IMFs extraction

To further illustrate the mechanism of sifting process that led to extract IMFs, the following two examples are presented:

Example 2.1.1. In the first example, the EMD algorithm is applied on a multicomponent signal that consist of two tones of equal amplitudes, $\mathbb{A}$, and zero phase shifts in which the frequency of the first tone (f_1) is double of that in the second tone (f_2) such that:

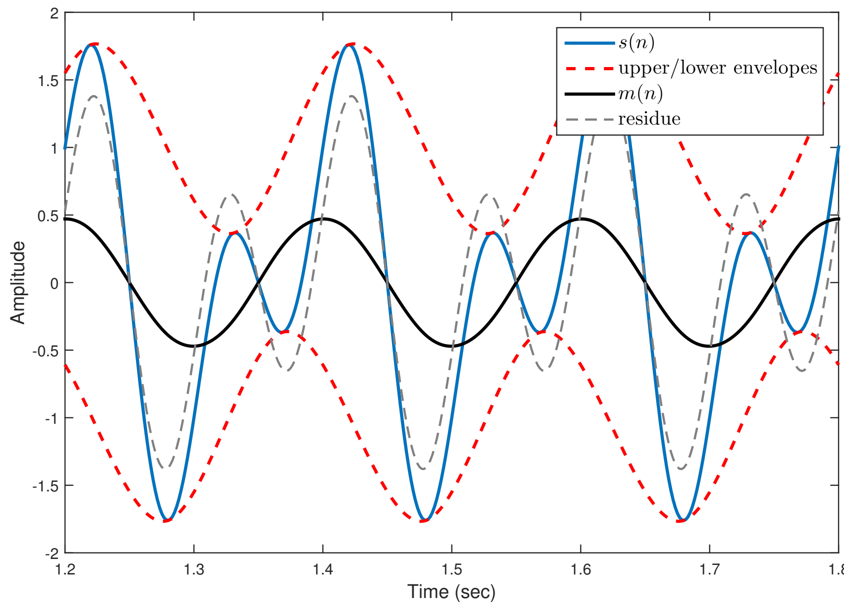

Figure 2.1: The first iteration of sifting process after applying EMD algorithm on $\mathbf{s}(n)$

$$\mathbf{s}(n) = \mathbf{s}1(n) + \mathbf{s}2(n)$$

where $\mathbf{s}1(n)$ and $\mathbf{s}2(n)$ are defined as follows:

$$\mathbf{s}1(n) = \mathbb{A}\sin(2\pi f_1 t) \tag{2.3}$$
$$\mathbf{s}2(n) = \mathbb{A}\sin(2\pi f_2 t)$$

The sifting process starts by identifying the local maxima and the local minima of $\mathbf{s}(n)$, then a cubic spline interpolation (3^{rd} order) is applied to connect these extrema (maxima/minima) to form the upper and lower envelopes respectively. Subsequently, the mean, $m(n)$, of the upper and lower envelopes are calculated and subtracted from $\mathbf{s}(n)$ to obtain the residue in which the residue is tested for stoppage criteria condition to decide whether it is IMF or not. Fig. 2.1 illustrates the sifting process after applying EMD on $\mathbf{s}(n)$.

Example 2.1.2. In the second example, the signal $\mathbf{s}(n)$, defined in Example (2.1.1), is contaminated with a white Gaussian noise (wGn) as follows:

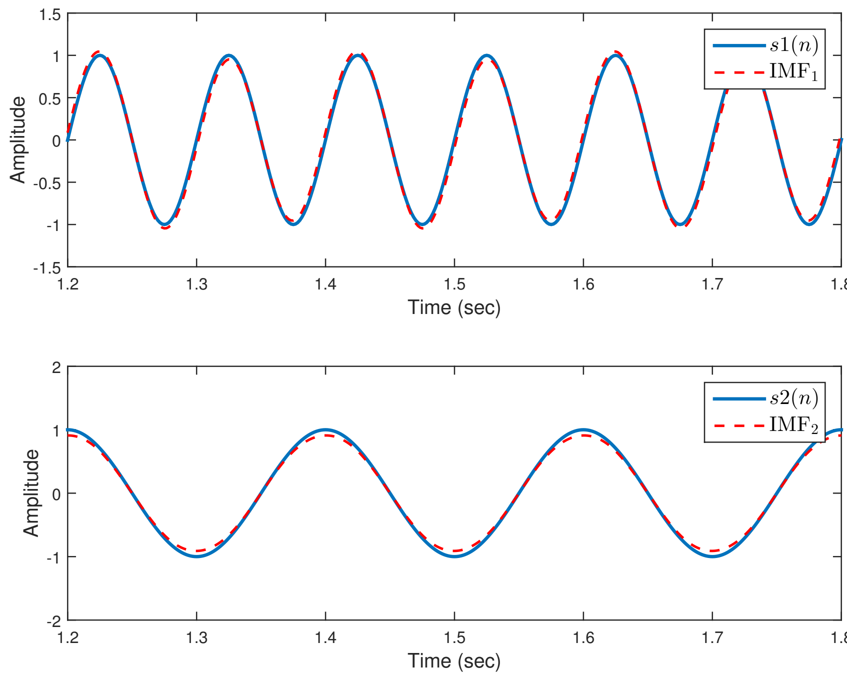

Figure 2.2: The upper figure shows the first sifted IMF (IMF_1) in contrast to the tone $s1(n)$ and the lower figure shows the second IMF (IMF_2) in contrast to the tone $s2(n)$

$$y(n) = s(n) + w(n)$$

where $y(n)$ is the noisy version of $s(n)$ and $w(n)$ is independent and identically distributed ($i.i.d.$) white noise with zero-mean and variance σ_w^2 $i.e.$ $w(n) = \mathcal{N}(0, \sigma_w^2)$.

Similar to Example (2.1.1), the EMD algorithm is applied on $y(n)$ and the sifting process is shown in Fig. 2.3. From that figure, it is obvious that IMF_1 carries the highest frequency of the processed signal which is in this case the white Gaussian noise, $w(n)$. Similarly, IMF_2 and IMF_3 sift the residual of the highest frequency components (mostly noise). However, it is noticeable that IMF_4 and IMF_5 sifted the frequency of $\mathbf{s}1(n)$ and $\mathbf{s}2(n)$ reflecting the behavior of dyadic filter. Herein, it becomes clear that EMD works as a natural and adaptive signal separation technique with an advantage of extracting the signal features in noisy environments.

Another illustration is given in Fig. 2.4 in which the Power Spectral Density (PSD)

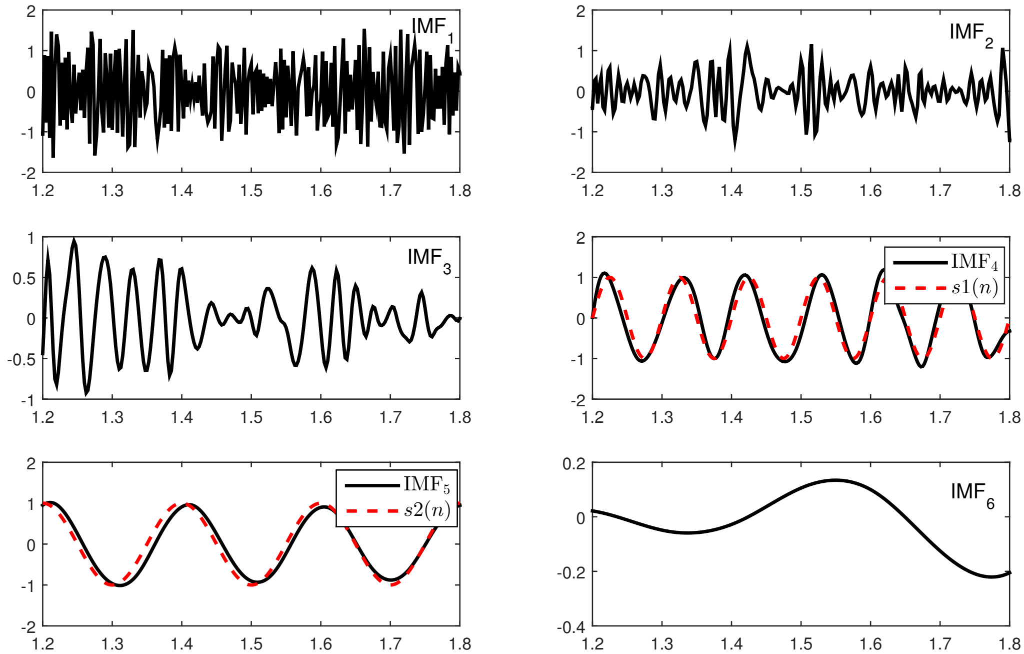

Figure 2.3: The first 6 IMFs and the corresponding $\mathbf{s}1(n)$ and $\mathbf{s}2(n)$ tones

of the processed signal, $y(n)$, and the corresponding IMFs (IMF_4 and IMF_5) are presented in frequency domain. From that figure, under sufficient sampling rate (>Nyquist rate), the EMD decomposed the processed signal from highest to lowest frequency resembling the behavior filter bank.

2.1.4 Sampling Rate Effect on EMD Performance

Originally, Empirical Mode Decomposition was designed to process continuous signals and hence any discontinuity can be misinterpreted as a signal component, thus oversampling is required for practical applications [3]. In [26], experiments have been performed to establish an EMD sampling limit. It was conducted that when the frequency of the processed signal approaches the sampling frequency, the EMD performance will degrade due to low amplitude resolution. However, it was found that a sampling frequency needs to be 5 times the Nyquist rate ($2 \times f_{max}$, where f_{max} is the highest frequency in the processed signal) in order to sift all possible features of the processed signal.

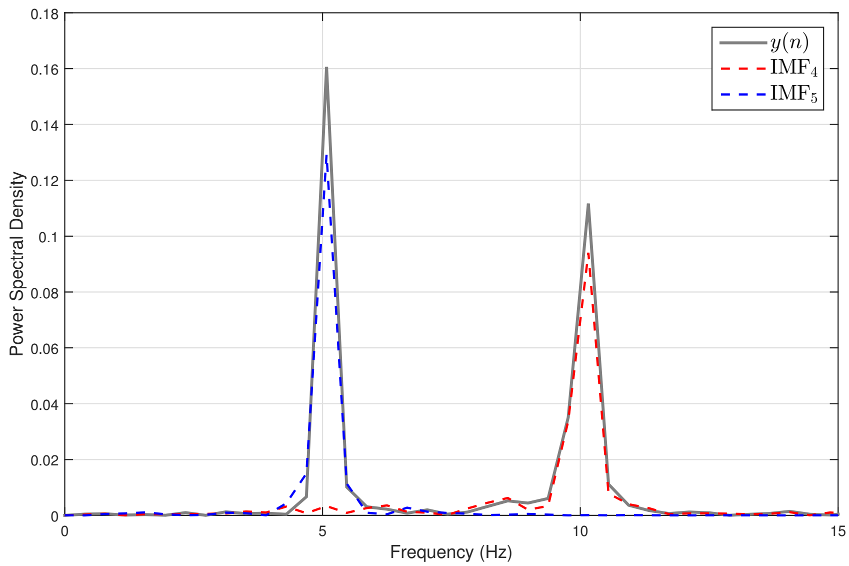

Figure 2.4: The Power Spectral Density of $y(n)$ and the corresponding IMFs

In [27], the authors studied the influence of the sampling on the EMD performance and illustrated possible sampling rate requirements. Moreover, the authors stated that the oversampling requirement is independent from the Nyquist criterion and is more related to reducing discontinuities in the signal. If the input to the EMD process is sampled at the Nyquist rate and an extremum (maximum/minimum pair) is lost during the sampling process thus, a local oscillation for the EMD process will not be determined in the sifting process. In [28], the proper sampling bound is selected based on a derived distribution model of white Gaussian noise whereas the deviation from this model indicates that the sampling rate is not sufficient enough to extract the information of the processed signal. In [29], the authors used a sampling rate of 50 times faster the Nyquist rate to fulfill the oversampling condition of EMD process.

2.1.5 EMD Applications

Despite of its empirical bases, EMD has gained a significant importance due to its robustness and superiority over other analogues. EMD has been applied into different data analysis fields of research, some of these fields are summarized as follows:

- **Biomedical measurements:** include features extractions [30,31], de-noising [16,18], and seizure detection [32]

- **Wireless communications:** include radar systems [33,34], satellite faults diagnosis [35], overlay communications [36,37], and spectrum sensing in cognitive radio networks [1,2,21,38]

- **Faults analysis:** include Power quality assessment [39], Damage detection [40,41], and Structural health monitoring [42]

- **Speech processing:** include speech noise suppression [43], speech recognition [44], and speech watermarking [45]

- **Image processing:** include image analysis [46], image watermarking [47], and image compression [48]

2.2 Literature Review

In this section, a review of the recent works that are related to our contributions and the proposed research is presented. This section is categorized in three different; however, coherent; areas begin with a review of the statistical analysis of the IMF properties. The de-noising problem is reviewed and the focus will be mainly on the EMD-based techniques in addition to other recent researches. In the detection category, the cognitive radio field with its key part, spectrum sensing, will be reviewed by addressing the most important related techniques.

2.2.1 Empirical Mode Decomposition: Statistical Properties Review

The EMD output mode, IMFs, and the decomposition sifting process have been investigated and analyzed statistically to understand its associated properties and underlying features [15, 49, 50]. Wu and Huang studied the statistical characteristics of IMFs of white Gaussian noise (wGn) and they concluded that the IMFs will approximately follow a Gaussian distribution [49]. Flandrin *et al.* analyzed fractional Gaussian noise (fGn) and came to a similar conclusion [15]. The phrase 'approximately Gaussian distributed' is used by both research groups loosely and has no strict mathematical meaning.

In addition to the aforementioned attempts to determine the statistical characteristics of IMFs, G. Schlotthauer *et al.* investigated the effect of the number of samples and the number of sifting iterations on the properties of the IMFs [50]. The later research group remarked that the probability density functions, *pdf*, of the IMFs are strongly dependent on the data length and on the maximum allowed number of iterations. However, unlike the findings in [15, 49], they observed (after applying two normality test methods) that many of the collected IMFs can have a Laplacian or a multi-modal rather than a Gaussian distribution except for few cases. Tseng and Lee proposed a two step entropic method to study IMF's properties by using a Bayesian interpretation of the resulting probabilities [28]. The two step entropic method involves in its first step clustering the resulted IMFs into two groups in which the first group contains the first set of IMFs, and the second group consists of the rest. The second step includes calculating the relative entropy between the input signal and each cluster in which minimum relative entropy indicates the presence of information.

2.2.2 Signal De-noising Review

Signal de-noising is a widely used term in signal processing which refers to removing the noise corruption to measurements such as biomedical, seismic, faults, and speech signals to name some. In this review, the focus will be on EMD-based de-noising techniques and some of the peer techniques from different approaches. Donoho *et al.* proposed a method, based on wavelet filtering, to de-noise signals by applying a soft thresholding approach on the obtained empirical wavelet coefficients and then invert the filter coefficients to recover the smoothed signal [10]. Sameni *et al.* suggested the use of non-linear Bayesian filter framework to process a single channel electrocardiogram using a modified non-linear dynamic model [51]. The modified model is utilized in several filters includes Extended Kalman Filter, and Unscented Kalman Filter with a use of automatic parameter selection method for better model adaptation.

The advantage of understanding the statistical properties of IMFs is that signals can be de-noised more effectively, in which the relevant modes corresponding to the signal components (almost noise free) are separated during the sifting operation. However, selecting the relevant IMFs is a critical process especially in the presence of high noise power. Flandrin *et al.* suggested the use of first Intrinsic Mode Function to set a noise-only model in order to use it as a threshold to decide whether the processed signal is carrying information or not [52]. In [53], the discrimination between different modes is approached using a correlation based threshold. In [54], the energy between two consecutive reconstructions of the signal is measured and the IMF index is identified when the first significant energy change occurs. The statistical characteristic of IMFs for noise is used to calculate energy spread as a function of various percentiles; filtering is then achieved by reconstructing the IMF energies out of the spreading bounds [55]. In [56], the Hurst exponent is adopted to select the IMFs that contribute to the signal more than noise, and partial reconstruction is used to produce an enhanced speech version. In [57], EMD is exploited in cooperation with HOS for speech-stream detection in low SNR region. The statistical similarity distance between the probability distribution of IMFs is used to decide which IMFs should be used for reconstruction to reduce noise [16].

2.2.3 Signal Detection: Spectrum Sensing Review

There is a great demand for bandwidth in wireless communications due to the dramatic shift in data usage from voice only to multimedia applications. Cognitive radio systems (CR) were proposed as a solution to satisfy that demand by making under utilized spectrum available. Spectrum sensing is the key part of CR as these systems must adapt to use different spectrum and bandwidths based on a primary user's channel usage. The efficacy of these systems relies on the ability to detect and monitor the primary users signal to avoid interference [58].

Generally, the performance of the sensing techniques is evaluated using two statistical measurements namely, the probability of detection (P_d) and the probability of false alarm (P_{fa}). In definition, P_d is the probability to decide the presence of signal when it is truly exist, and can be written as follows:

$$P_d = Prob(\mathcal{H}_1|\mathcal{H}_1) \tag{2.4}$$

On the other hand, P_{fa} is the probability to decide incorrectly the presence of signal when it is indeed absent, and can be formulated as follows:

$$P_{fa} = Prob(\mathcal{H}_1|\mathcal{H}_0) \tag{2.5}$$

Moreover, Receiver Operating Characteristic (ROC) curves is used to provide a comprehensive performance measure by providing a link between different P_d and P_{fa} values at certain Signal-to-Noise Ratio (SNR).

CR consists of a broad range of spectrum sensing techniques, each of which have advantages and disadvantages. These techniques range from low to high computation complexity and have varying levels of ability to determine the presence of signals in noise. An ideal spectrum sensing technique would sense any class of signals quickly, require no *a priori* information and operate at low SNR walls [17].

Energy detectors (ED) are the most popular among all spectrum sensing techniques due to its low computational complexity; in which, the energy metric of the received signal is compared to a predefined threshold to decide the occupancy of a given channel [59]. However, lacking of the accurate estimation of the noise variance may result in increased missed detection and false alarms On the other hand, noise uncertainty is additional factor that affects the efficiency of energy detectors in which noise starts to fluctuate during the sensing duration.

In [60], an adaptive multi-band energy detector accompanied with the nonparametric Kolmogorov-Smirnov (K-S) test are proposed through an iterative exploration process that reduces the number of nominee channels. Zeng and Liang proposed a spectrum sensing methods based on the sample covariance matrix of the received signal [61]. The eigenvalues of the sample covariance matrix is calculated and thresholds are derived for the probability of detection and the probability of false alarm. An example of the eigenvalue-based detector is called Maximum-Minimum Eigenvalue detector (MME) which is independent of the noise power of the received signal and perform well in low SNR regimes. Further, another example of eigenvalue-based detectors is called Maximum-Eigenvalue Detector (MED) in which is dependent on the noise power and still performs well under low SNR scenarios [62].

In terms of spectrum sensing, EMD was used as a component of detection methods and wireless applications [63]. Roy and Doherty used EMD, in general, to enhance the detection of weak signals in the presence of noise [64]. This technique is dependent on a characteristic of EMD that would require calculations that may not make it practical for real-time sensing. In an effort to detect signals under noise uncertainty, Bektas *et al.* proposed a spectrum sensing algorithm based on relative entropy [38]. However, this method also requires a large number of calculations to determine the classifier to separate a signal from noise. Gunturkun *et al.* used the Bivariate Empirical Mode Decomposition (BEMD) to facilitate radar scene analysis for cognitive radar [33]. The proposed method takes advantage of the distinct response of intrinsic mode functions energies to the fractional Gaussian character of coherent sea clutter returns. That response is used then to set a null hypothesis that distinguish the presence or the absence of the target under test. The use EMD in wireless communication is limited to the baseband signals due to the high complexity (large number of samples) that arise with the application of bandpass or wideband signals.

2.2.4 Noise Power Estimation Review

Noise power estimation is widely used in speech, image, and wireless communications [65–67]. The estimation can be performed either in time or frequency domain based on the associated application. In this review, we'll focus on the noise power estimation in wireless communications and more specifically the applications to spectrum sensing in cognitive radio networks.

In spectrum sensing, the noise power of the received signal is the key part to calculate the threshold that is used to make the detection decision. The importance of the noise power estimation becomes more obvious with the techniques such as energy detectors that are widely used in spectrum sensing due to its simplicity and efficacy. Techniques such as energy detectors and maximum eigenvalue detectors are considered to be semi-blind in which they require the knowledge of the noise power to calculate the detection threshold. However, the knowledge of the noise power is not available in practical situations and hence assuming the noise power might deteriorate the detection rate significantly [17]. Further, the noise in wireless communications is not stationary and might change over time due to the non-stationarity of channel characteristics [68]. Therefore, accurate noise power estimation can shift semi-blind techniques to blind region and to use their detection capabilities more efficiently.

The estimation of the noise power in spectrum sensing has been studied in several works [69–71]. In [69], multiple signal classification algorithm is used to separate the signal and noise subspaces. In [70], the noise power is estimated by sacrificing sub-bands of the channel to be used only for the purpose of noise estimation. In [71], minimum descriptive length is used to separate the noise eigenvalues of the covariance matrix to be used in the noise power estimation. Further, the effect of the noise power estimation on the energy detectors performance is analyzed and investigate the conditions for SNR wall, which is the highest detection rate that can be achieved regardless of the length of the observation interval [72].

CHAPTER 3

Intrinsic Mode Functions: Statistical Analysis and Applications

Empirical Mode Decomposition is a non-linear, local and fully data-driven algorithm that breaks signals into a set of modes called Intrinsic Mode Functions (IMF). Yet, there is no mathematical basis for EMD and thus an interpretation of the statistical characteristics is required to understand the non-linear nature of the algorithm. Analytically, EMD acts as a dyadic filter in which the highest oscillations embedded in the signal, typically noise, are sifted in the first IMFs. This filtering property becomes certainly advantageous by knowing that EMD is non-parametric and derive its filtering basis from the signal itself (signal dependent). In this chapter[1], an analytic analysis of the EMD-sifting process is presented and a probability distribution function is suggested as a best fit model for IMFs. The statistical properties of the IMFs are traced analytically based on the EMD process of a range of random variables with different probability distributions. The statistical properties of IMFs are used to develop two applications; de-noising scheme and SNR estimation based on null-hypothesis test of IMFs probability distribution. In both applications, partial reconstruction is applied after identifying the reference IMF which carries more information than noise. To demonstrate the validity of the proposed methods, a comparison is made with other comparable approaches in addition to previously proposed EMD-based de-noising technique.The results showed that the proposed method offers a larger gain in both the de-noising and SNR estimation methods specifically at low signal to noise ratio scenarios.

[1]Part of contents of this chapter is published in IEEE Signal Processing Letters [18].

3.1 Introduction

Data analysis is an essential part of any sensing or measurement process to extract information buried in noise from the measured values. Conventional data analysis methods are typically limited by their statistical stationary assumptions made in advance about the processed signals. Methods such as the wavelet analysis, Wagner-Ville, and the various short-time Fourier spectrograms all share this disadvantage which limits the full potential for signal analysis [4].

Most of the aforementioned techniques require *a priori* knowledge of the processed signal or part of its statistical properties to be applied properly for a wide range of applications. Empirical Mode Decomposition (EMD) was proposed to alleviate the conventional techniques drawbacks. EMD is an adaptive non-parametric signal-oriented algorithm that decomposes time-series signal into a set of AM-FM modes called Intrinsic Mode Functions (IMF). These IMFs are by definition zero-mean oscillations and produced by an iterative operation called the sifting process. The key feature of IMFs is that they are derived from the signal itself and are not predefined like the basis functions for wavelet or Fourier transformations. Since the IMFs carry the physical features of the processed signal, understanding the statistical properties of these modes is essential to establish an analytical interpretation of their relationship to the signal.

The probability distribution of IMFs yielded after processing a white Gaussian noise, wGn, are investigated in [49]. The authors concluded that the IMFs will approximately follow a Gaussian distribution. Similarly, the probability distribution of IMFs of processing a fractional Gaussian noise, fGn, is studied in [15] in which the authors came with a same conclusion. The phrase "approximately Gaussian distributed" is used by both research groups loosely and has no strict mathematical meaning.

In addition to the aforementioned attempts to determine the statistical characteristics of IMFs, G. Schlotthauer *et al.* investigated the effect of the number of samples and the number of sifting iterations on the properties of the IMFs [50]. The later research group remarked that the distribution of the IMFs are strongly dependent on the data length and on the maximum allowed number of iterations. However, unlike the findings in [15, 49], they observed (after applying two normality test methods) that many of the collected IMFs can have a Laplacian distribution or a multi-modal distribution rather than a Gaussian distribution except for few cases. These previous investigations were also restricted to Gaussian distributed signals and little attention was given to signals with different distributions. In [18], the authors proposed that the generalized Gaussian distribution (GGD) would represents a better fit of IMFs probability distribution than the suggestions in [15, 49].

The advantage of understanding the statistical properties of IMFs is that signals features can be extracted, de-noised, more effectively [16, 18, 55, 56]. In [18], the null hypothesis of the GGD for each IMF is utilized to discard the modes which follow GGD and use the other modes to reconstruct the de-noised signal. The statistical similarity distance between the probability distributions of IMFs is used to decide which IMFs should be used for reconstruction to reduce noise [16]. The statistical characteristic of IMFs of noise is used to calculate energy spread as a function of various percentiles; de-noising is achieved by reconstructing the IMF energies out of the spreading bounds [55]. In [56], the Hurst exponent is adopted to select the IMFs that contribute to the signal more than noise, and use the reconstruction to produced an enhance speech version.

The purpose of this chapter is to provide a new statistical understanding of the sifting process which is the key operation of generating the IMFs. The contributions of this chapter include: a) a study of the statistical properties of the IMFs generated by EMD-processed random variables of different signal distributions, b) tracing the statistical analysis of the IMFs to propose a better probability distribution, and c) two EMD-based applications schemes, de-noising and SNR estimation, that leverage the statistical knowledge of the IMFs.

3.2 Statistical Analysis of The IMFs

In this section, the behavior of EMD process for different distributions of random variables inputs contaminated with/without white Gaussian noise (wGn) is studied. This analysis is done by tracing the statistics of the IMFs, starting with IMF_2 through the sifting process.

3.2.1 IMFs statistical properties for different random variable distributions

In previous work, it was concluded that the EMD of wGn will yield "approximately" Gaussian distributed IMFs for the low IMF indices [49]. However, the probability distribution of IMFs for non-Gaussian distributed input signals is still unclear. In order to fill this knowledge gap, extensive simulations were performed to investigate the probability distributions of IMFs resulting from signals with different random variable distributions. The simulation includes signals distributed with excess Kurtosis ranging from 3 to -1.2, which covers the following distributions; Laplacian, hyperbolic secant, logistic, Gaussian, raised cosine, Wigner, and uniform [73]. The chosen distributions[2] also have a variety of different statistical properties.

EMD is applied on each of the random variables, a set of IMFs is generated, and the distribution of each IMF is analyzed[3]. Figure 3.1 illustrates an example of a particular IMF distribution (IMF_3) of different signal distributions. From this figure, IMF_3 can be seen to follow an "approximately Gaussian distribution" with zero-mean and different standard deviations. Similar conclusion can be drawn when the same random variables are contaminated with wGn[4] as it is shown in Fig. 3.2. At this point, the results do not contradict the previously reported findings. This result motivates an investigation to determine if there is a distribution that fits better than "approximate Gaussian."

Additionally, the variance of the produced IMFs, for different signal distributions with-

[2]This analysis was also done for different asymmetric and non-bell shape distributions including exponential, Gamma, Beta, log-normal, and Arcsine. Similar observations to the chosen distributions have been conducted.

[3]In this section, different random variable distributions are generated each of length of 5000 samples. The results are the average of 100 runs based on Monte-Carlo simulations.

[4]As real world measurements are vulnerable to noise corruption, and because the wGn is the most common used model, we suggested to study the effect of adding wGn to signal distributions.

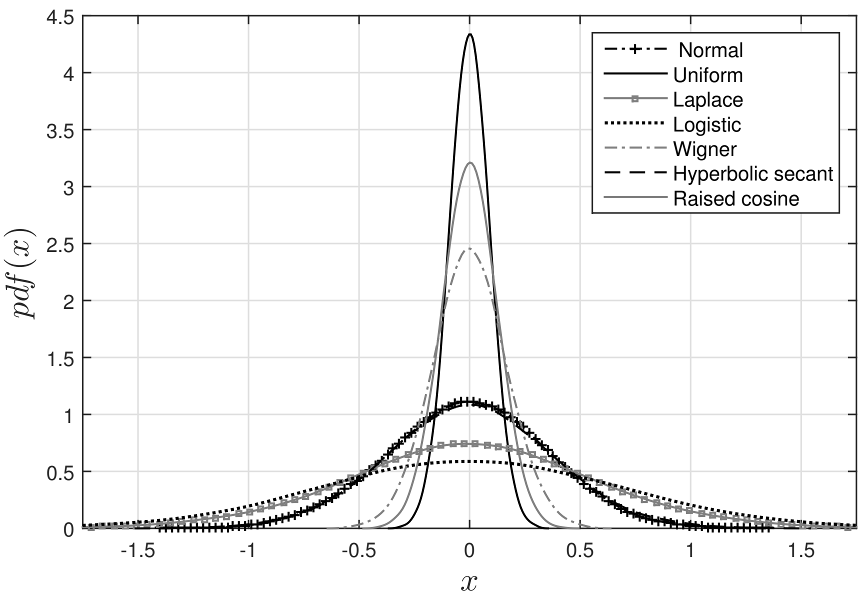

Figure 3.1: The distributions of IMF_3 different random variable distributions.

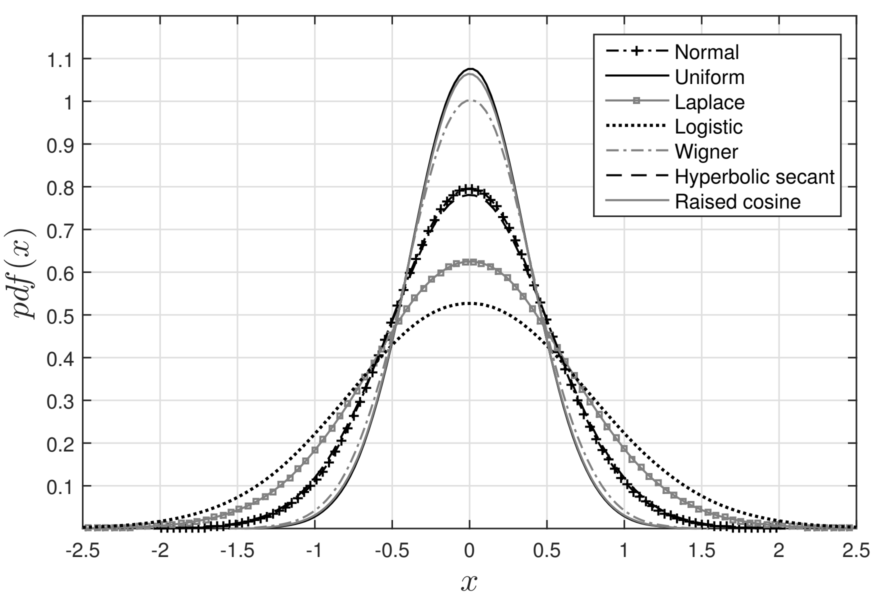

Figure 3.2: The distributions of IMF_3 different random variable distributions corrupted with wGn.

out wGn, is investigated and shown in Fig. 3.3. From that figure, it is shown that the variances decrease as the IMF indices increases because the iterative process (sifting) lowers the standard deviation of the distribution of an IMF during each iteration. Another way to understand that is the sifting process shifts the IMF distributions from platykurtic to lep-

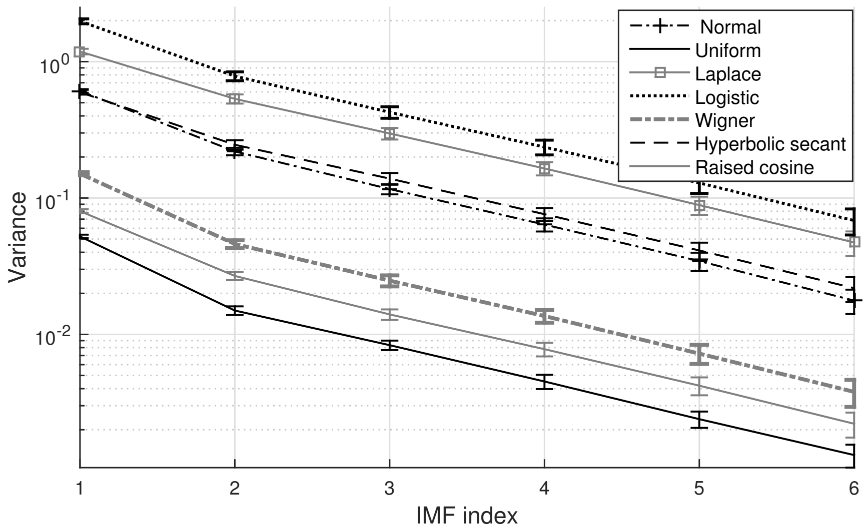

Figure 3.3: The IMF variances of different random variable distributions.

tokurtic densities and that affects the tail weight of each IMF distribution recursively which results in smaller standard deviations (variances). On the other hand, Fig. 3.4 shows the effect of corrupting the signal distributions, presented in Fig. 3.3, on the variance behavior. From that figure, it is clear that the variance values are increased in contrast to the previous figure and that is due to the contributions of the wGn variance.

Further, the excess kurtosis of the produced IMFs is evaluated for different signal distributions with and without wGn as shown in Fig. 3.5. Specifically for low IMF indices, the excess kurtosis is about **zero** which indicates that the produced IMFs follow roughly a Gaussian distribution regardless the input signals distributions. It is worthy to mention that the excess kurtosis of IMF_1 for all distributions is negative (close to -1) and that is because the first mode is bi-modal which is known to have negative excess kurtosis [74].

Based on the aforementioned results of IMFs statistical properties, it can concluded that, for different input distributions, whether contaminated or not with wGn, the resulting IMF probability distributions will follow approximately Gaussian distribution.

On the other hand, by using excess kurtosis as a rough measure for tail weight, it is clear that the IMFs excess kurtosis demonstrates tails that can be heavier or lighter than Gaussian for different distributions [75]. Therefore, one can conclude that under varying conditions

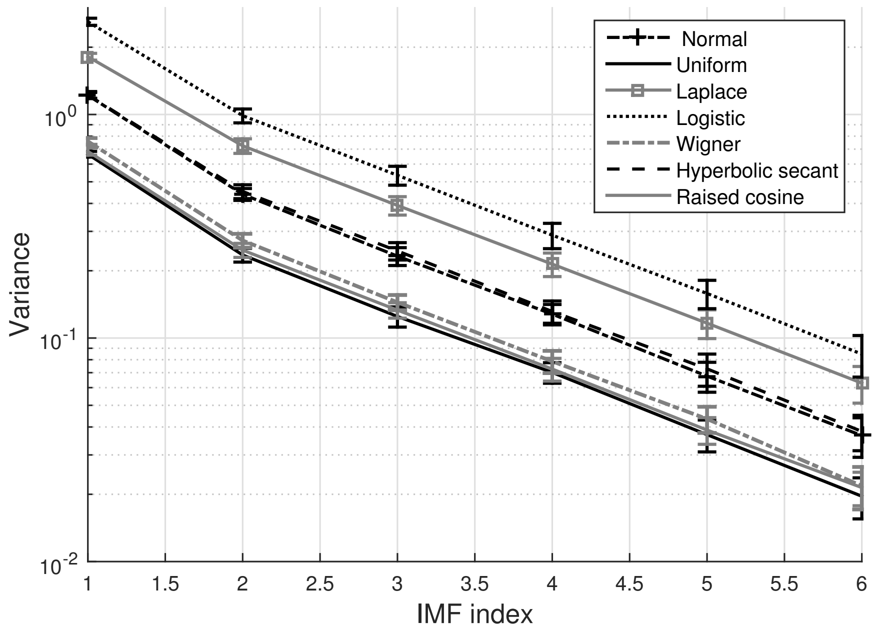

Figure 3.4: The IMF variances of different random variable distributions of corrupted with wGn.

the IMFs will follow a Gaussian or other distributions. Previous work concluded that some conditions lead to Gaussian distributed IMFs and in some cases Laplacian, which is partly true. However, there is still a knowledge gap as to what the distribution of the IMFs is for all cases. Thus, in order to verify the probability distribution of these IMFs, an analysis is presented, followed by a detailed methodology to support the proposed claim.

3.2.2 IMFs probability distribution statistical analysis

In this section, the IMFs probability distribution is derived based on the analysis of the EMD sifting process. More specifically, the distribution of the envelope mean and the residual signal are traced statistically. Then, an IMF candidate distribution is determined from the EMD process of the envelope mean and the residual signal. Further, a statistical distance measurement is used to verify the candidate probability distribution has a better fit than previously assumed distributions. Finally, a null-hypothesis test is used to justify the suggested fitting distribution.

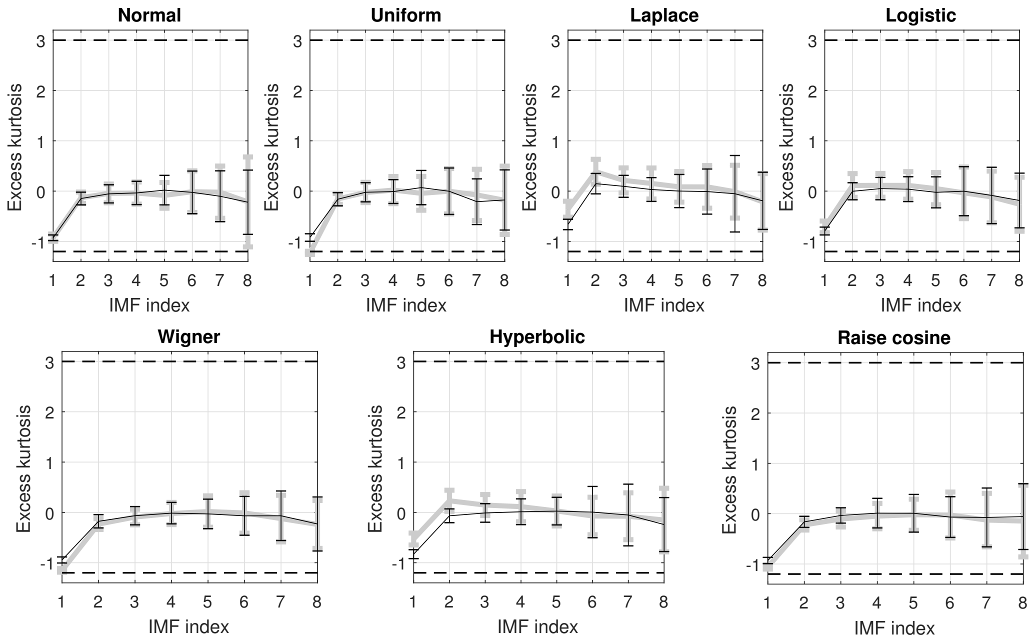

Figure 3.5: The excess kurtosis of IMFs for different random variable distributions with and without corruption of wGn. The grey solid lines and the black solid lines, along with their error bars, are the result of signal distributions without and with wGn respectively. The dashed lines represents the excess kurtosis boundaries (-1.2 and 3) of the uniform and Laplace distributions respectively.

For our analysis, we assume that the input, $z(n)$ is a Gaussian random variable with zero-mean and unity variance. The EMD algorithm is applied to $z(n)$ and set of IMFs is generated, IMF_j. The focus of the analysis is on the second IMF as the first IMF is deduced to follow bi-modal distribution [52]. The first iteration of the second IMF starts by subtracting the envelope mean $m_i(n)$ from $r_i(n)$ to yield $h_i(n)$ which is the i^{th} sifted signal of $\mathrm{IMF}_2(n)$. Here, $m_i(n)$ is the mean of the upper and lower envelopes of the residual signal $r_i(n)$ which is resulting from subtracting the $\mathrm{IMF}_1(n)$ from the input signal, $z(n)$.

To understand how the EMD process affects the distribution of the IMFs, we consider the distribution of the residual signal $r_i(n)$ and mean of the envelopes $m_i(n)$ with a quantile plot to assess their deviations from the theoretical normal distribution quantile. Figure 3.6 shows that the tail of the envelope mean becomes heavier (grey circles) with reference to the standard normal distribution (solid black). This is due to the sifting process changing the

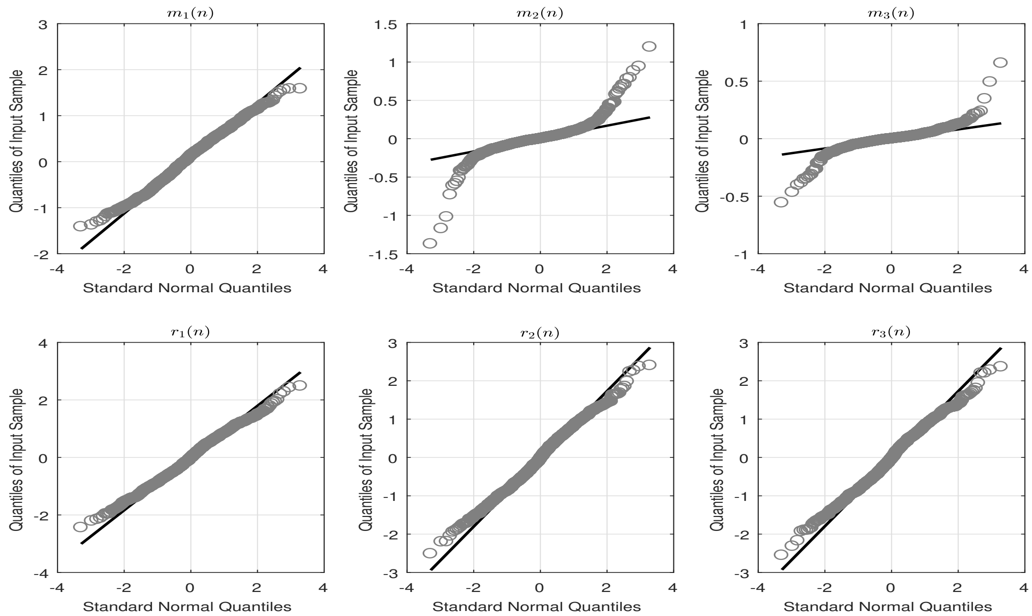

Figure 3.6: The QQ plot of the envelope mean and residual signal for three sifting iterations of IMF_2.

data density from platykurtic to leptokurtic. On the other hand, the residual signal exhibits tail that resembles the Gaussian distribution. Therefore, the distribution of $m_i(n)$ and $r_i(n)$ follow $student - t$ and Gaussian distribution respectively; this conclusion is supported by the best fit probability distribution simulation[5]. For simplicity, we'll denote $m_i(n)$, $r_i(n)$, and $h_i(n)$ by m_i, r_i, and h_i respectively.

From Sec. 2.1.1, the detailed signal, h_i, is the convolution of r_i and m_i in frequency-domain. However, being derived from the same random variable $z(n)$, there is a possibility that m_i and r_i are statistically dependent. For dependent random variables, a copula joint distribution is used to depict the dependency between two random variables [78]. The next step is to understand the degree of dependency in order to find the distribution of the candidate IMF. The candidate IMF, h_i, is given in terms of the joint probability distribution

[5] Different well-known distributions are utilized to best fit the data, $m_i(n)$ and $r_i(n)$. Bayesian information criterion (BIC) and Akaike information criterion (AIC) are used to decide the fitting parameters that can model the underlying data. The distribution that results in smallest the BIC or AIC is assumed the best fitting distribution [76, 77].

f_{m_i,r_i} by

$$f_{h_i}(h_i) = \int_{-\infty}^{\infty} f_{m_i,r_i}(m_i, h_i + m_i) dm_i \tag{3.1}$$

Clearly if m_i and r_i are independent then $f_{m_i,r_i}(m_i, r_i) = f_{m_i}(m_i) f_{r_i}(r_i)$ and (3.1) becomes a convolution integral.

If m_i and r_i are assumed to be dependent variables, thus the multivariate cumulative distribution function $F(m_i, r_i) = C\big(F_{m_i}(m_i), F_{r_i}(r_i)\big)$, where C is the copula, $F_{m_i}(m_i)$ and $F_{r_i}(r_i)$ are the cumulative marginal distributions of m_i and r_i respectively. In this respect, two statistical nonparametric measures of dependence; Kendall's tau (τ) and Spearman's rho (ρ) are defined such that $\tau(m_i, r_i) = 4E[C(U, V)] - 1$ and $\rho(m_i, r_i) = 12E[C(U, V)] - 3$, where U,$V\sim$ Uniform(0,1), $U = F(m_i)$ and $V = G(r_i)$, F and G are the marginal distributions of m_i and d_i respectively [79]. In this work, the ratio, denoted by ξ, of $\rho(m_i, r_i)$ to $\tau(m_i, r_i)$ is used as an indicator to show the degree of dependency of the joint distribution between m_i and r_i. In [80], it was conducted that the ratio ξ approaches 3/2, for infinite number of samples, as the joint distribution approaches that of two independent random variables.

In Fig. 3.7, an analysis of the ratio ξ for different signal[6] distributions of $\text{IMF}_2(n)$ is illustrated. From this figure, the value of ξ is approaching 3/2 over all iterations. Therefore, the dependency can be assumed weak and approaching independence for which the convolution integral is a good approximation for h_i [80].

Given that m_i and r_i can be assumed to be independent and they are convoluted, then the IMF candidate distribution can be traced statistically. Nason *et al.* derived a general closed-form expression for the convolution of $student - t$ and a Gaussian distribution [81],

[6]Random variables of different distributions with 5000 samples each are averaged from 1000 trails using Monte-Carlo simulations.

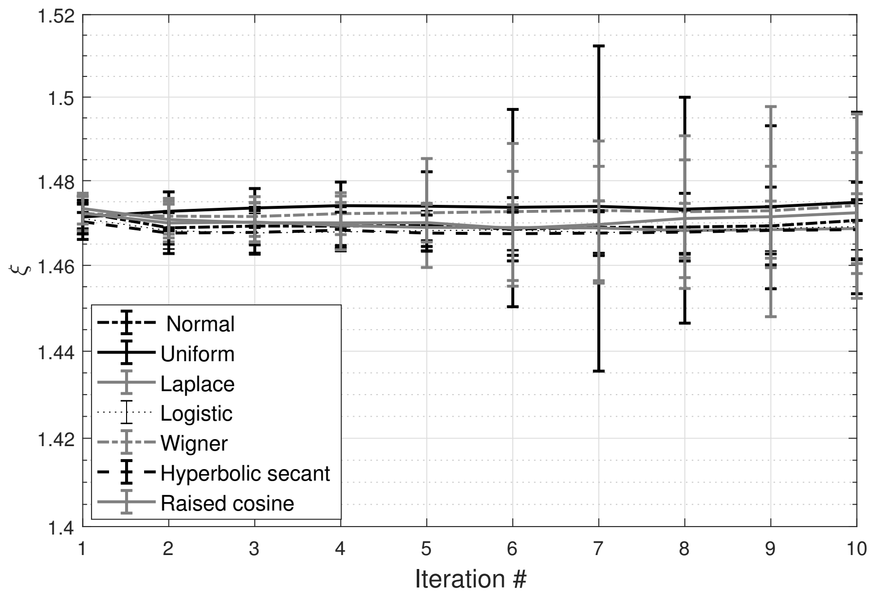

Figure 3.7: The ratio ξ for different signal distributions between the envelope mean and residual signal for 10 sifting iterations of IMF_2.

which is given as follows:

$$\begin{aligned}\psi(x) = \frac{b}{\pi}exp(a/2)\Bigg[& \left\{(1-a)cos(\mu a + \frac{pb}{2}sin(\mu a)\right\} I_{C1}(p,d) \\ & + \left\{((1-a)sin(\mu a) - \frac{pb}{b}cos(\mu a)\right\} I_{S1}(p,d) + be^{-d^2}/2\Bigg]\end{aligned} \tag{3.2}$$

where $a = \sigma^{-2}$, $b = \sqrt{2}/\sigma$, $d = a/b$, $p = \mu b$,

$$I_{C1}(p,d) = \int_d^\infty cos(px)e^{-x^2}dx = \frac{\sqrt{\pi}}{2}e^{-p^2/4}[1 - Re\left\{erf(c) - erf(c-d)\right\}],$$

and

$$I_{S1}(p,d) = \int_d^\infty sin(px)e^{-x^2}dx = \frac{\sqrt{\pi}}{2}e^{-p^2/4}Im\left\{erf(c-d)\right\},$$

where $c = ip/2$, $i = \sqrt{-1}$, Re and Im extract the real and imaginary parts, $erf(.)$ is the error function.

However, Nason *et al.* did not specify the resulting distribution of his derived closed-form expression. A simple method is to determine a best fit Gaussian distribution to the closed-form expression given in [81]. The probability distribution function of a Gaussian distribution is given as follows,

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \tag{3.3}$$

where μ and σ are the mean and the standard deviation respectively [82].

Intuitively, replacing the shape parameter of the Gaussian distribution ($\kappa = 2$) by a variable creates a statistical family called Generalized Gaussian Distribution (GGD) [82]. This family includes the Laplacian and Gaussian distributions, κ =1 and 2 respectively, as a special cases. The probability distribution function of a symmetric GGD is,

$$f(x) = \frac{\kappa}{2\rho\Gamma(1/\kappa)} e^{-(|x-\mu|/\rho)^\kappa} \tag{3.4}$$

where κ is a parameter that controls the distribution tail, ρ is defined as $\sqrt{\Gamma(\frac{1}{\kappa})/\Gamma(\frac{3}{\kappa})}\sigma$, and $\Gamma(.)$ is gamma function. Further, the GGD is suggested to model non-Gaussian processes, where the tail weights could be heavier or lighter than Gaussian based on the shaped parameter [83].

In order to show that GGD is a better fitting distribution than Gaussian with respect to the closed-form (3.2) solution, the Hausdorff distance measure is calculated for the two distributions. The Hausdorff distance measure is calculated by,

$$\mathcal{HD}(A, B) = max\big(\mathfrak{D}(A, B), \mathfrak{D}(B, A)\big) \tag{3.5}$$

where $\mathfrak{D}(A, B) = \underset{a\in A}{max}\ \underset{b\in B}{min}||a - b||$ and $\mathfrak{D}(B, A) = \underset{b\in B}{max}\ \underset{a\in A}{min}||b - a||$ [84]. To calculate Hausdorff distance for this comparison, the distribution of closed-form expression (3.2) and GGD (3.4) will be substituted into equation (3.5) instead of A and B respectively, as follows:

$$\mathcal{HD}\Big(\psi(x), f(x)\Big) = max\bigg(\mathfrak{D}\Big(\psi(x), f(x)\Big), \mathfrak{D}\Big(f(x), \psi(x)\Big)\bigg) \tag{3.6}$$

The two distributions are compared and the κ of GGD was varied between 1.4 and 2

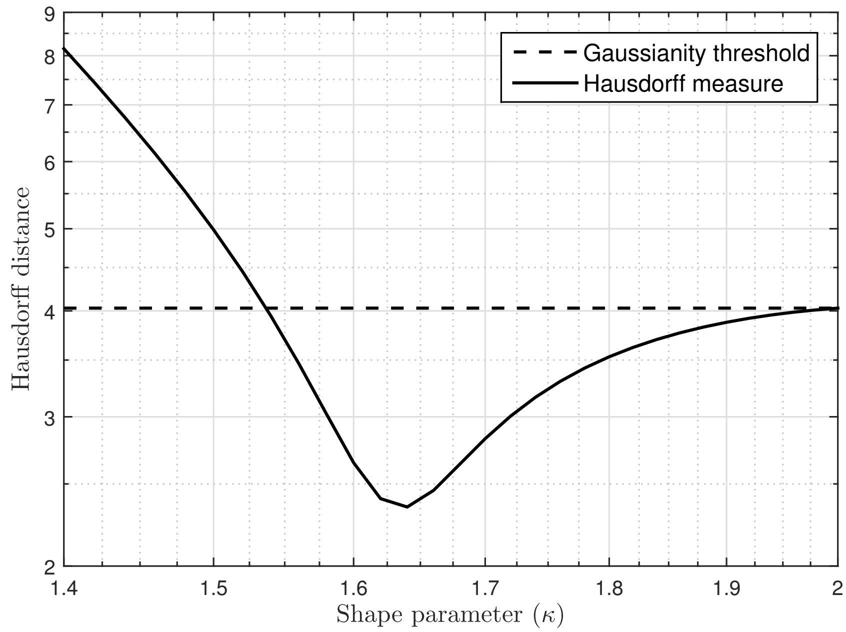

Figure 3.8: Hausdorff distance measure of the GGD for values of κ with respect to the closed-form expression

to determine whether there is a distribution that reveals better fitting to the closed-form expression (3.2) than the Gaussian. From Fig. 3.8, it is clear that at ($\kappa = 1.64$), the shortest distance or the highest similarity of GGD with respect to closed-form expression is obtained. Further, it is obvious that the κ values in the region under the "Gaussianity threshold" represent the closed-form expression better than the best fit Gaussian model.

The generalization of GGD includes platykurtic densities that span from the normal density ($\kappa = 2$) to the uniform density ($\kappa = \infty$) and a leptokurtic densities that span from the Laplace ($\kappa = 1$) to the normal density ($\kappa = 2$) [83]. Thus, the GGD successfully explains the previous findings [15, 49, 50], where Laplace and Gaussian distributions are just special cases in GGD family.

3.2.3 The null hypothesis test of GGD

The analytical analysis in Sec. 3.2.2, suggests that the EMD-processing of Gaussian distributed random variable leads to a GGD. Therefore, to verify the validity of the suggested distribution (GGD), a null hypothesis test is evaluated for random variables with different distributions including the Gaussian distribution. In [18], a null hypothesis test of GGD is suggested and will be revisited in this work with more details.

In order to apply a GGD null hypothesis test, the given distribution must be transformed first to Gaussian. In this respect, there are different normality tests in which the Shapiro-Wilk parametric hypothesis test of composite normality is used widely [85]. The Shapiro-Wilk test and the Shapiro-Francia test reveal better performance for a platykurtic and leptokurtic samples respectively [85,86]. Therefore, an initial kurtosis check on the samples set is performed before deciding the method used to determine its Gaussianity. Both tests return a single value (H) where the value of zero indicates that the hypothesis is not rejected within the predefined confidence interval (α), and the value of one represents that the hypothesis is rejected. This is the formal definition of the null hypothesis of composite normality and is the best reasonable assumption regarding the population distribution of a random sample.

To apply the null hypothesis of GGD, the given random variable is transformed to the cumulative distribution function (*cdf*) of the GGD. If the data samples indeed follow a GGD then, this will convert the distribution of the random variable to become uniform. The *cdf* for the GGD defined in (3.4) is given by,

$$F(x) = \frac{1}{2} + \mathrm{sgn}(x-\mu)\frac{\hat{\gamma}[1/\kappa, (\frac{|x-\mu|}{\rho})^{\kappa}]}{2\Gamma(1/\kappa)} \tag{3.7}$$

where $\hat{\gamma}$ denotes the lower incomplete gamma function [87], $\rho = \sqrt{\Gamma(\frac{1}{\kappa})/\Gamma(\frac{3}{\kappa})}\sigma$ and σ is the standard deviation of the distribution [88]. We will restrict[7] our search for GGD of κ between 1 and 3, since the initial analysis indicated that the distribution of all IMFs for any random signal will change from a Laplace distribution $\kappa = 1$ in the extreme case to a more round top distribution. The complete procedure to test if any random variable follows GGD

[7]Practical experimentations showed that if x follows a GGD with κ more than 3, then another distribution, like the "*Beta*" or "generalized *Gamma*" distribution would be a better fit. This is why we chose to test GGD for κ limit not exceeding 3 specifically.

is given as:

1. Assign the given random variable to x.

2. Find $F(x)$ of the GGD from (3.7), with a predefined κ (start from $\kappa = 1$).

3. Convert x to a uniform distribution by applying the transformation $y = F(x)$. Now if x follows GGD with the predefined κ then y will be uniformly distributed between 0 and 1.

4. Convert y to a Gaussian distribution by applying the transformation $z = erf^{-1}(2y-1)$ where erf and erf^{-1} is the error function and the inverse error function respectively; if y is uniformly distributed between 0 and 1 then, z will follow Gaussian distribution.

5. Check the kurtosis of the obtained distribution to determine whether the Shapiro-Wilk or the Shapiro-Francia null hypothesis test should be used.

6. Evaluate the null hypothesis of Gaussian distribution, z, and test the normality of z. If z is normally distributed then y is uniformly distributed between 0 and 1. Therefore, x follows GGD with the given κ. Otherwise, increment the value of κ by $d\kappa = 0.01$ and as long as κ is less than 3 go to step (3)

7. If the steps above failed to reject the hypothesis of x following GGD then, we need to check the confidence level of this conclusion[8].

Table-3.1 illustrates the null hypothesis test for different random variable distributions with 5000 samples. In that table, H = 0 indicates that the corresponding IMFs follow GGD. On the other hand, H = 1 (excluding IMF_1) indicates that the corresponding IMFs do not follow GGD (or we failed to show the IMF does follow a GGD). The corresponding κ values

[8] The confidence level of α in the original null hypothesis test would mean that the hypothesis of y not being uniformly distributed failed with a confidence level of $0.5erf(\alpha)$. Further, this would mean that, the hypothesis that x did not follow GGD failed with a confidence level of $k[\Gamma_{inc}^{-1}(0.5erf(\alpha))]^{1/\kappa}$ where Γ_{inc}^{-1} is the inverse incomplete Gamma function and $k = \sqrt{\Gamma(\frac{1}{\kappa})/\Gamma(\frac{3}{\kappa})}$.

Table 3.1: Null hypothesis (H) and shape parameter (κ) of different random variable distributions.

pdf	IMF index																			
	1		2		3		4		5		6		7		8		9		10	
	H	κ	H	κ	H	κ	H	κ	H	κ	H	κ	H	κ	H	κ	H	κ	H	κ
Normal	1	3	0	2	0	2	0	1	0	1.17	0	1.7	0	1	0	1.29	0	2.6	1	3
Uniform	1	3	0	2	0	2	0	2	0	1	0	1.06	0	1	0	1.26	1	3	0	1.24
Laplace	1	3	0	1.01	0	1	0	1	0	1.04	0	1	0	1.3	1	3	0	1.21	1	3
Logistic	1	3	0	2	0	2	0	2	0	1	0	1.13	0	1.58	0	1.12	0	1.22	0	1
Wigner	1	3	0	1.08	0	1.14	0	1	0	1.1	0	1.01	0	1.07	0	2.62	1	3	1	3
Hyperbolic secant	1	3	0	1	0	1.02	0	2	0	2	0	1.27	0	1	0	1.06	0	1.23	1	3
Raised Cosine	1	3	0	1	0	2	0	1.17	0	1.01	0	1.07	0	1	0	1.55	0	1.14	1	3

give an evidence that GGD distribution could fit the resulting IMFs other than the Gaussian or Laplacian distributions.

3.3 EMD-H de-noising

The null hypothesis (H) test in Sec. 3.2.3 is utilized to propose a de-noising method (EMD-H) [18]. Assuming an EMD was applied to a signal with noise and a set of IMFs have been generated, a GGD null hypothesis (H) test of GGD is evaluated for each IMF. The IMFs that carry mainly noise will follow a GGD rather than a Gaussian distribution. A deviation from GGD distribution would indicate the presence of a signal as it signifies a change in statistical characteristics.

The GGD null hypothesis value (H) is used to select the IMF reference, IMF_R, in which at this IMF the information components will dominate the noise components. The change of H value from 0 to 1 (excluding IMF_1) is an indication of the reference IMF and is also the start point for the partial reconstruction. The proposed EMD-H de-noising method can be summarized as follows:

1. Apply the EMD algorithm on a given signal contaminated with noise to produce a set of IMFs.

2. Apply the GGD null hypothesis (H) for each IMF (excluding IMF_1).

3. Select the IMF index that corresponds to $(H = 1)$ and consider that IMF as a reference, IMF_R.

4. Apply the partial reconstruction as follows:

$$\mathbf{y}(n) = \mathcal{T}(n) + \sum_{i=R}^{M} \text{IMF}_i(n) \tag{3.8}$$

where $\mathbf{y}(n)$ is the reconstructed signal and R is the index of the reference IMF

3.4 Simulations and Results

In this section, three distribution models are compared with their corresponding null hypothesis test on a synthetic ECG signal to show the validity of our proposed model (GGD) in light of different amounts and kinds of noise. In order to understand the impact of sampling rate, we compare the SNR after de-noising for different values of input SNR and sampling rate. Finally, we compare the EMD-H de-noising technique with other techniques on three different types of signals. The sampling rate used in this section is $8N_q$ and Monte Carlo simulations are carried out for all obtained results by averaging 1000 runs.

First, the effects, of varying the sampling rate, N_q, and SNR levels, on the quality of the received signal (in terms of SNR_o) is illustrated in Fig. 3.9. From this figure, it is clear that the SNR_o is enhanced by increasing the sampling rate and increasing the input SNR levels. This follows the EMD characteristic stated in Sec. 2.1.4, where oversampling is required to capture all signal features.

The next result is an analysis of EMD-H using GGD as the proposed distribution model compared to the null hypothesis tests using the Gaussian [15] and Laplacian [50] distributions. The simulations were performed on an ECG contaminated with noise for various input SNRs. Unlike white noise, the spectral density of real noise is not flat and the samples are correlated. Therefore, we also include colored noise in our comparison to highlight the performance under more realistic conditions [51] in which positively correlated colored noise

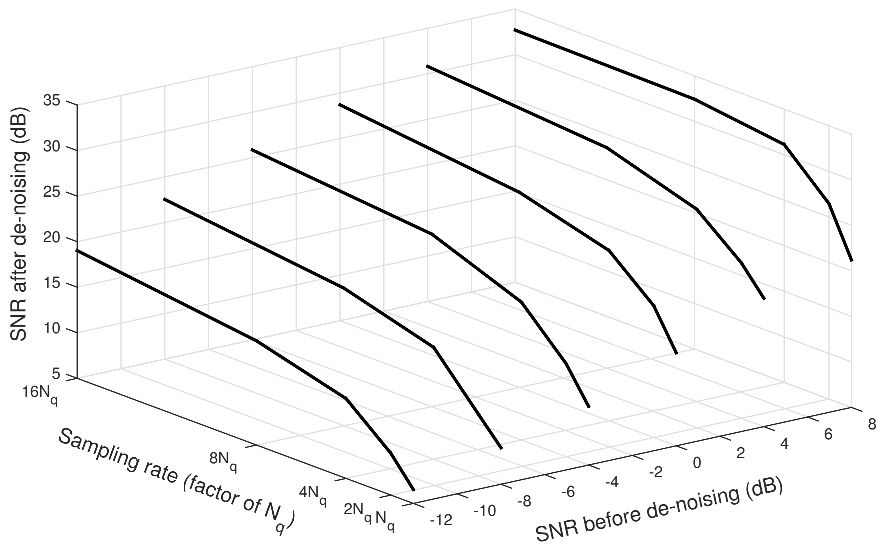

Figure 3.9: The effect of sampling rate and SNR variation of GGD model for an ECG signal.

is generated based on the work in [89]. The output SNR or SNR after de-noising, SNR_o, and Mean Square Error, MSE, are used for the performance evaluation as these metrics demonstrate the relevant improvement in SNR and difference from the original signal.

The results of this analysis are shown in Table-3.2. It is clear that GGD performs better than Gaussian and Laplacian specifically at low SNR values (< 4 dB) both in terms of MSE and SNR_o. However, at high SNR levels (> 4 dB), the Gaussian model starts to show slightly better results as the presumably noise-dominated IMFs will carry more signal power causing the GGD hypothesis to be rejected. Hence, the Gaussian model will slightly fit better than GGD as it will require less transformations with a slightly broader confidence interval. In comparison, the performance of the Laplacian distribution fit shows a much weaker performance.

For further validation, an ECG signal contaminated with wGn is used to compare the EMD-H with two conventional de-noising techniques; wavelet de-noising and Savitzky-Golay (SG) filtering. To emphasize the adaptive behavior of EMD, default settings[9] have been

[9] A Daubechies mother wavelet of order 8 is used for the Wavelet de-noising approach, the decomposition level is set to 5, and the soft thresholding is adopted. Also, a 3^{rd} order filter and frame size equal to 41 is used for the SG filtering approach.

Table 3.2: A comparison of the SNR after de-nosing and MSE for three distribution models in a null hypothesis test for white and color noise over various SNR input values before de-noising, denoted by SNR_i.

Dist. model	SNR_i (dB)	Colored		White		SNR_i (dB)	Colored		White	
		SNR_0	MSE	SNR_0	MSE		SNR_0	MSE	SNR_0	MSE
Laplacian	−12	2.66	1.51	5.26	0.60	−6	6.27	0.386	10.17	0.15
Gaussian		6.12	0.505	16.09	0.044		10.98	0.129	21.45	0.0127
GGD		**8.62**	**0.29**	**17.11**	**0.032**		**12.55**	**0.084**	**22.15**	**0.0095**
Laplacian	0	11.48	0.096	15.70	0.038	4	15.26	0.038	19.55	0.015
Gaussian		16.44	0.032	26.44	0.0034		19.87	0.014	**29.74**	**0.0016**
GGD		**17.25**	**0.026**	**26.5**	**0.0033**		**20.28**	**0.013**	29.37	0.0018
Laplacian	8	19.22	0.015	23.49	0.0062					
Gaussian		**23.57**	**0.006**	**32.92**	**0.0008**					
GGD		23.54	0.0072	32.54	0.001					

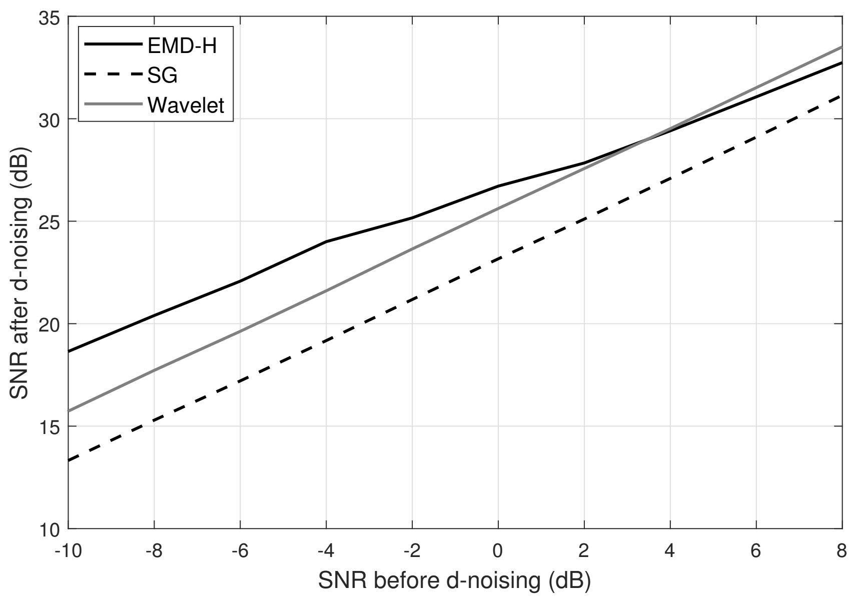

Figure 3.10: Illustrates the performance of EMD-H method in contrast with other conventional de-noising techniques

used for the conventional approaches. From Fig.3.10, it is clear that the EMD-H de-noising method performs better than the other techniques for SNR $<$ 4 dB. At high SNRs, $>$ 4 dB, the wavelet de-noising method starts slightly to overcome the EMD-H as the wavelet starts to capture the signal structure more accurately.

Finally, a comprehensive test is conducted on an ECG signal and the Blocks and Doppler

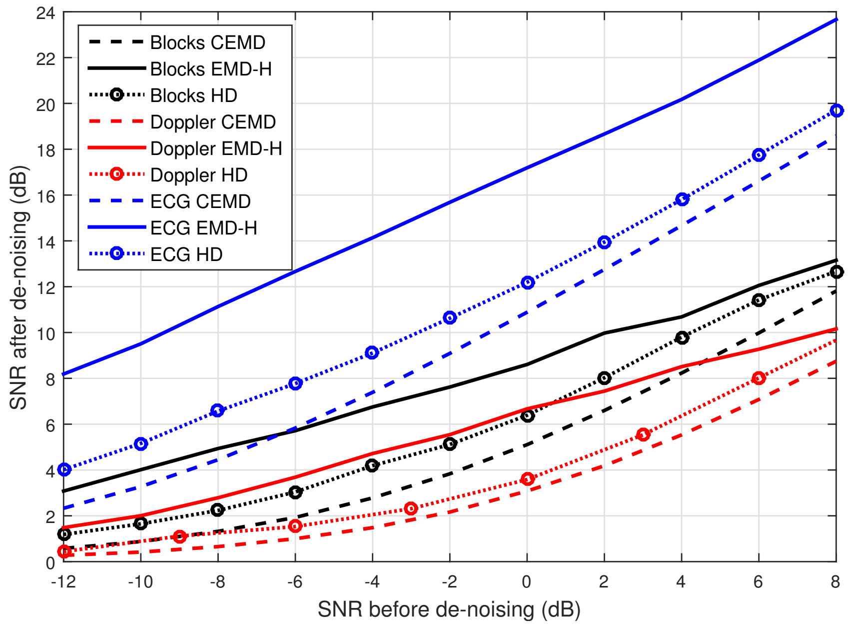

Figure 3.11: SNR comparison between EMD-H and other EMD-based de-noising methods for various signals with $8N_q$ sampling rate

signals [90], which are all contaminated by colored noise. The proposed method, EMD-H, is compared to other EMD-based de-noising techniques such as Conventional EMD (CEMD) and EMD-based Hausdorff distance (HD) [16,52]. The results in Figure 3.11 show that EMD-H outperforms all of the other methods. At -10 dB, EMD-H has a gain of 1.5, 2.5, and 5 dB for the Doppler, Blocks, and ECG signals over the best EMD-based peer. EMD-H has better de-noising performance compared to the other techniques due to the proper selection of the reference IMF, IMF_R. The performance of CEMD is limited because the selection of IMF_R is based on a hard threshold set by the energy of IMF_1, which in turn is only dominated by noise under strict sampling rate and SNR conditions [52]. EMD-H also performs better than HD because the latter method is highly sensitive to temporal changes in the noise, which leads to the wrong estimation of IMF_R [16]. Generally, EMD-based de-noising techniques are sensitive to signal types as seen with the worse de-noising performance of the high-pass Doppler signal. The high frequency content is sifted by EMD into the first modes, which are discarded as noise-only modes.

3.5 Conclusion

In this chapter, the statistical analysis of the Intrinsic Mode Functions (IMFs) and its application to de-noising and SNR estimation are evaluated. The first contribution is the study of the IMFs statistical properties when random variables of different signal distributions are decomposed by EMD. Further, it is shown that the distribution shape for all IMFs (except IMF_1 which is likely to be bi-modal) is Gaussian-like. At this point, there was no guarantee that these distributions were Gaussian or have any other distribution related to the noise or signal distribution. The second contribution is a statistical analysis to prove whether these IMFs are Gaussian distributed or not. The analysis of the IMFs probabilities, through extensive experiments, led to infer that the resulted IMFs of Gaussian distributed random variable follow Generalized Gaussian Distribution functions. Based on that finding, it appears that the claims in [15, 49, 50] were only partially true, whereas the GGD includes the Gaussian, Laplacian distributions as a special cases. The proposal that the IMF distribution follows the GGD method was successfully validated with a null-hypothesis test.

The EMD-H method is relying on the partial reconstruction of the relevant IMFs to generate a noise-free version from the noisy processed signal. In the de-noising scheme, a synthetic and real signal are used in which both signals are assessed with quality measure metrics. The proposed de-noising method requires no predefined parameters and makes no assumptions of the signal type. The proposed de-noising scheme was compared to conventional techniques: wavelet de-noising, non-linear Savitzky-Golay filtering, and other EMD-based de-noising techniques and demonstrated good ability for noise removal.

CHAPTER 4

EMD-Based Energy Detector for Spectrum Sensing in Cognitive Radio

Sec. 2.1.5 lists different applications for EMD. In this chapter[1], we apply EMD to spectrum sensing in cognitive radio. The filtering characteristics of Empirical Mode Decomposition (EMD) are used to create a blind and adaptive energy detector for single or multi-channel spectrum sensing. Due to EMD filtering behavior, the first IMF (IMF_1) is mostly contaminated by noise from the received noisy signal. The proposed approach takes advantage of Cell Averaging Constant False Alarm Rate (CA -CFAR) as an optimal detector to enhance the probability of detection. Alternative to conventional CA-CFAR (which requires at least one nearby vacant channel for good noise estimation), the first IMF will be used as a training function for noise estimation purposes. Based on the IMF_1 characteristics in frequency domain, the noise floor of the received signal is estimated and a threshold is derived for a given false alarm rate. Simulations show the supremacy of the proposed detector in comparison with other conventional detectors.

4.1 Introduction

Cognitive radio (CR) systems are proposed to cope with bandwidth deficiency by making underutilized spectrum available. Spectrum sensing is the key part of CR as these systems must adapt to use different spectrum windows and bandwidths based on a primary user's (PU) channel usage. The efficacy of these systems relies on the ability to detect and monitor the primary users signal to avoid interference [58, 91].

[1]The contents of this chapter is published in IEEE International Conference on Computing, Networking and Communications [20].

There is a broad range of spectrum sensing techniques each with it's own advantages and disadvantages. These techniques range from low to high computational complexity and have varying levels of ability to determine the presence of signals embedded in noise. Energy detectors (ED) are the most popular among all spectrum sensing techniques due to low computational complexity; in which, the energy metric of the received signal is compared to a predefined threshold to decide the occupancy of a given channel [17]. However, lacking an accurate estimation of the noise power may result in increased missed detection and false alarm rates. Furthermore, since ED threshold is a function of noise power, it is influenced by the temporal changes (uncertainty) in noise profile that is a result of both wireless channel and receiver characteristics such as thermal noise [92].

Different approaches are proposed to alleviate the threshold sensitivity towards inaccurate noise power estimation [60, 93, 94]. In [93], noise power estimation is achieved by using an auxiliary radiometer that operates over an adjacent spectral region in parallel with a radiometer detecting a signal. In [94], signal classification is performed by utilizing spectral techniques to separate the noise samples and detect outliers for a given false alarm rate. Moreover, forward method via Cell Averaging Constant False Alarm Rate (CA-CFAR) is utilized to detect and locate the outliers. In [60], an adaptive multi-band energy detector accompanied with the nonparametric Kolmogorov-Smirnov (K-S) test are proposed through an iterative exploration process that reduces the number of nominee channels.

More recently, Empirical Mode Decomposition (EMD) is used as a component of detection methods for spectrum sensing and wireless applications [63]. Gunturkun *et al.* used the Bivariate Empirical Mode Decomposition (BEMD) to facilitate radar scene analysis for radar detection [33]. The proposed method takes advantage of the distinct response of intrinsic mode functions energy to the fractional Gaussian character of coherent sea clutter returns. Roy and Doherty used EMD, in the general sense, to enhance the detection of weak signals in the presence of noise [64].

In an effort to detect non-stationary and non-linear signals, Bektas *et al.* proposed a spectrum sensing algorithm using relative entropy [38]. However, this method also requires a large number of calculations to determine a classifier to separate the signal from noise. In [1], a non-parametric threshold is derived from the power spectral density of Intrinsic Mode Functions (IMF), the EMD output, for multi-channel detection. In [2], the noise is estimated and two thresholds are calculated by applying CFAR method on a set of IMFs in the frequency domain. All these techniques depend on EMD's unique characteristics for noise estimation. However, in most cases they require a lot of calculations deeming them impractical for real-time sensing.

In this chapter, the first Intrinsic Mode Function (IMF_1) is used to estimate the noise floor power of the received signal for a given false alarm rate inspired by the CA-CFAR approach. The proposed method takes its robustness from the nature of EMD as a fully data-driven tool with no predefined parameters. In the proposed technique, a received single or multi-channel band is compared to double thresholds derived from IMF_1 for a given false alarm probability.

4.2 EMD-based detector model

The philosophy of spectrum sensing for CR is to determine if a certain channel is occupied or not. The detection problem can be converted (as shown below) into a binary hypothesis test where, $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the absence and presence of the PU respectively.

$$y(n) = \begin{cases} w(n) & : \mathcal{H}_0 \\ s(n) + w(n) & : \mathcal{H}_1 \end{cases}$$

where $n = 1 \ldots N$ and N is the total number of samples, and $y(n)$, $s(n)$, and $w(n)$ are the received signal at cognitive radio (CR) user, transmitted signal from PU, and additive white Gaussian noise respectively. However, $w(n)$ is independent and identically distributed white Gaussian noise which is modeled by a zero-mean process with variance σ_w^2 *i.e.* $w(n) = \mathcal{N}(0, \sigma_w^2)$, $s(n)$ is the transmission from a PU or another CR user and assumed to be independent from $w(n)$.

In this chapter, the EMD processor is required to sift only the first IMF, denoted by $c_1(n)$, out of the received signal $y(n)$. An estimate of the power spectral density (PSD) of IMF_1 is obtained using the Welch spectrum estimator. The Welch estimation is performed by dividing the data samples N into K segments, where each segment is windowed before calculating the periodogram. Thus, the Welch PSD estimator will be the result of averaging all windowed periodograms, and the averaging will yield lower variance of estimated PSD [95]. The PSD estimation of IMF_1 is written as:

$$C_1(f) = \frac{1}{K} \sum_{k=1}^{K} \phi_k(f) \tag{4.1}$$

where $k = 1 \dots K$, ϕ_k is the windowed periodogram and is defined as:

$$\phi_k(f) = \frac{1}{LJ} \left| \sum_{n=1}^{L} v(n) c_{1k}(n) e^{-j2\pi f n} \right|^2 \tag{4.2}$$

where L is the number of samples of the k^{th} segment, $v(n)$ is the window function (Hamming), and J is the power of the temporal window $v(n)$ and can be calculated as $J = \frac{1}{L}\sum_{n=1}^{L} |v(n)|^2$.

Similarly, the PSD estimation of the received signal $r(n)$ (an energy metric to make the final detection decision) can be given as follows:

$$S_r(f) = \frac{1}{K} \sum_{k=1}^{K} \hat{\phi}_k(f) \tag{4.3}$$

where $\hat{\phi}_k(f) = \frac{1}{LP} \left| \sum_{n=1}^{L} v(n) r_k(n) e^{-j2\pi f n} \right|^2$ is the windowed periodogram of the k^{th} segment for the received signal $r(n)$ and similarly P is the power of the temporal window $v(n)$.

4.3 IMF_1 Probability distribution analysis and Proposed approach

In this section, the distribution of IMF_1 is investigated for both time and frequency domains. In addition, the proposed method is presented and a threshold is derived for a given probability of false alarm.

4.3.1 IMF_1 Probability Distribution Analysis

The first Intrinsic Mode Function (IMF_1) of EMD algorithm, exhibits special statistical characteristics over other IMFs. Due to the nature of EMD sifting process, IMF_1 contains the finest scale and is defined as a high-band highpass filter [15]. Furthermore, and under a sampling rate bigger than Nyquist, IMF_1 will be mostly dominated by noise [52], [14]. The probability distribution of IMF_1 for Gaussian processes is a mix of two distributions which give a single bi-modal distribution [96]. The justification of the bi-modality for such a distribution relies in the large discrepancy values (in case of noisy or noise only signals) yielded by the maxima and minima envelopes [97].

In order to use IMF_1 in the spectrum detection model, its statistical properties must be investigated. For a white Gaussian noise (wGn) signal, the distribution of upper and lower envelopes produced by IMF_1 are denoted by $p_u(x)$ and $p_l(x)$ respectively. However, since the wGn is an uncorrelated and independent random process, $p_u(x)$ and $p_l(x)$ are assumed to be random, independent from each other, with each one following a different distribution. It must be noted that samples of the upper or the lower envelopes are not mixed or added as each sample of IMF_1 belongs either to an upper or to a lower envelope. Thus, IMF_1 can be thought of as the addition of two mutually exclusive random variables. The resultant of these two distributions will be a mixture distribution such that [98]:

$$p_{mix}(x) = \frac{max\big(p_u(x), p_l(x)\big)}{\mathcal{A}} \tag{4.4}$$

where $p_u(x) = \left[\frac{1}{2} + \left(\frac{1}{2} - erf(x)\right)\right]^2 \gamma(x)$, $p_l(x) = \left[\frac{1}{2} - \left(\frac{1}{2} - erf(x)\right)\right]^2 \gamma(x)$, $\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $erf(x)$ is the error function, and $\mathcal{A}$ is the total area under the entire curve.

The mixture distribution $p_{mix}(x)$ has two different modes and weights. Fig. 4.1 shows the distribution of both upper and lower envelopes of IMF_1 for processed wGn signal. In the frequency domain, the Fourier transform coefficients of IMF_1 can be considered independent and follow a scaled Gaussian distribution for both upper and lower envelopes denoted by $P_u(x)$ and $P_l(x)$.

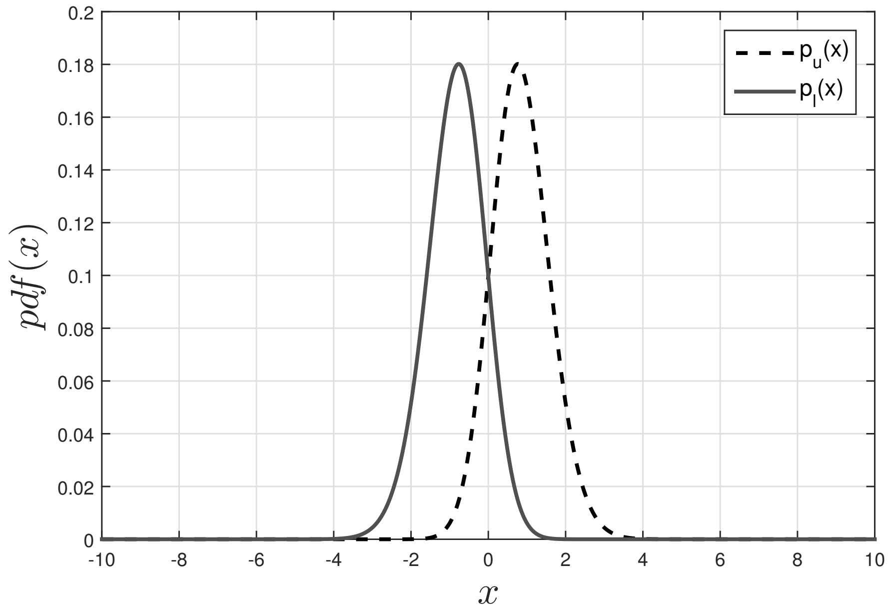

Figure 4.1: The probability distribution of both upper and lower envelops of IMF_1 for a processed wGn signal

For a unimodal Gaussian distribution variable, each periodogram bin follows an exponential distribution [99]. Therefore, the periodogram bin distributions for both $P_u(x)$ and $P_l(x)$ will follow exponential distributions and since they are mutually exclusive, the same applies to IMF_1. Fig. 4.2 illustrates that the distribution of IMF_1 periodograms for a wGn signal (with 95% confidence interval) follows an exponential distribution. The square-law output of the IMF_1 periodogram follows an exponential distribution and can be modeled as follows [100]:

$$f(x) = \frac{1}{2\hat{\alpha}} e^{-x/2\hat{\alpha}} \tag{4.5}$$

where $\hat{\alpha}$ is the scale parameter that measures the variation of the data.

4.3.2 The Proposed Approach

When the noise power is not known *a priori*, the averaging technique CA-CFAR is used to estimate the noise power of the received signal by assuming some portions (cells) of the spectrum (called "training cells") to be dominated by noise only in which these training cells

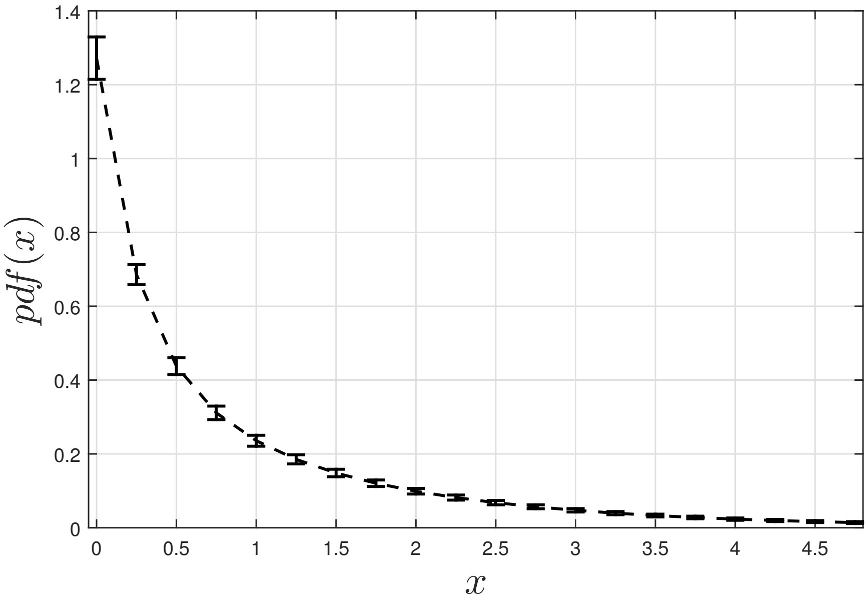

Figure 4.2: The exponential distribution of each periodogram for IMF_1 with 95% confidence interval

surround the cell under test [101]. In this chapter, the first IMF is assumed to be dominated by noise which can thus serve here as a "training cell". It has been shown previously [52] that such an assumption can almost always be made valid by proper oversampling. However, α in Eq. (4.5) can be replaced by σ^2_{W1} to yield:

$$g(x) = \frac{1}{2\sigma^2_{W1}} e^{-x/2\sigma^2_{W1}} \tag{4.6}$$

where σ^2_{W1} is the expectation of $C_1(f)$, such that:

$$\sigma^2_{W1} = E\{C_1(f)\} \tag{4.7}$$

Here, the quantity $(2\sigma^2_{W1})$ represents the noise floor of the received signal in the frequency domain and is used as a lower threshold (denoted by λ_L) such that:

$$\lambda_L = 2\sigma^2_{W1} \tag{4.8}$$

However, the proposed detector threshold can be derived based on (4.6) for a given probability of false alarm as follows:

$$P_{fa} = \int_{\lambda_U}^{\infty} g(x)\,\mathrm{d}x = e^{-\lambda_U/2\sigma_{W1}^2} \tag{4.9}$$

From (4.9), the proposed detector threshold λ, will be:

$$\lambda_U = -\ln(P_{fa}) \times 2\sigma_{W1}^2 \tag{4.10}$$

The derived threshold in (4.10) can be seen as a scaled version of the noise floor ($2\sigma_{W1}^2$). IMF_1 is the only IMF needed to estimate the noise floor of the received signal and hence determine the optimal detector threshold. Although the threshold in (4.10) seems to be a function of noise power σ_{W1}^2, the later in this case is estimated from IMF_1 which is assumed to contain noise-only samples. The detection methodology is shown in Fig. 4.3 and can be summarized as follows:

1. Decompose the received signal $r(n)$ to obtain IMF_1 which is denoted by $c_1(n)$.

2. Transform $c_1(n)$ to the frequency domain then calculate its PSD denoted by $C_1(f)$ using (4.1).

3. Determine the noise floor ($2\sigma_{W1}^2$) and denote it by λ_L as a lower threshold, and similarly λ_U in (4.9) as the upper threshold, where both thresholds λ_L & λ_U are used for detection and localization.

4. Apply both thresholds on the PSD of the received signal given in (4.3) for both coarse and fine detection.

5. In the coarse detection, each consecutive group of samples that exceeds λ_L is defined as a i^{th} nominated channel and denoted by $\mathbf{X}_i$.

6. In the fine detection, if any sample or set of samples of the i^{th} nominated channel $\mathbf{X}_i$ exceeds λ_U, then it is decided that the channel is occupied.

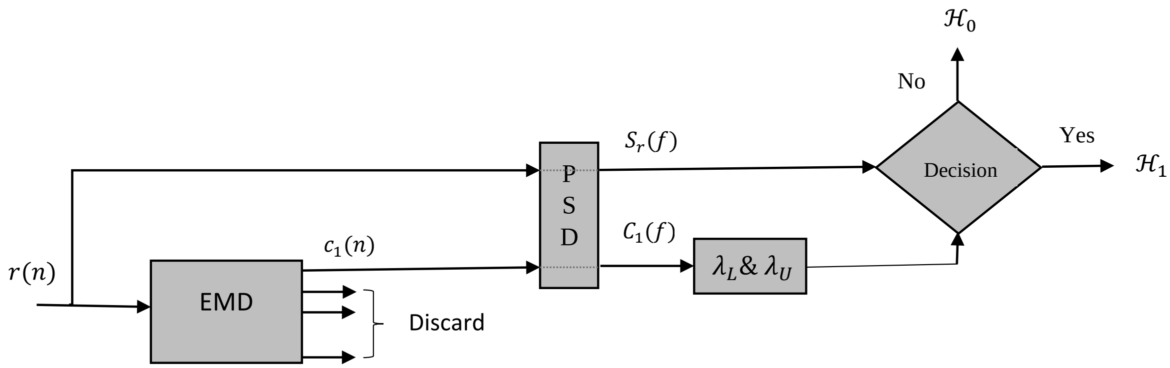

Figure 4.3: The schematic diagram of the proposed method

4.4 Simulation and results

In this section, the proposed IMF_1-based detector is evaluated using a synthetic baseband oversampled OFDM modulated time-domain signal with an observation period (T). In the analysis, Monte-Carlo simulations were carried out, where each result is the average of 2000 runs.

4.4.1 The Evaluation of the Proposed Detector Performance

To evaluate the performance of the proposed sensing technique, two statistical measurements are used namely, the probability of detection (P_d) and the probability of false alarm (P_{fa}) over different SNR values and sampling sizes. In Fig. 4.4, the receiver operating characteristic (ROC) is performed to examine the proposed detector efficacy for a single channel of 200 KHz bandwidth. From Fig. 4.4, it is shown that the proposed detector starts to reveal a good detection rate when the SNR ≥ -10 dB with a sample size of 5000. The behavior of the proposed detector is further assessed through ROC curves by varying the sample size for SNR $= -12$ dB as it is illustrated in Fig. 4.5. It is noticeable that increasing the sample size (>5000) will enhance the detection rate and that is justified by knowing that more features can be extracted by increasing the sample size or oversampling the received signal.

In Fig. 4.6, multi-channel detection is performed where three out of sixteen channels are

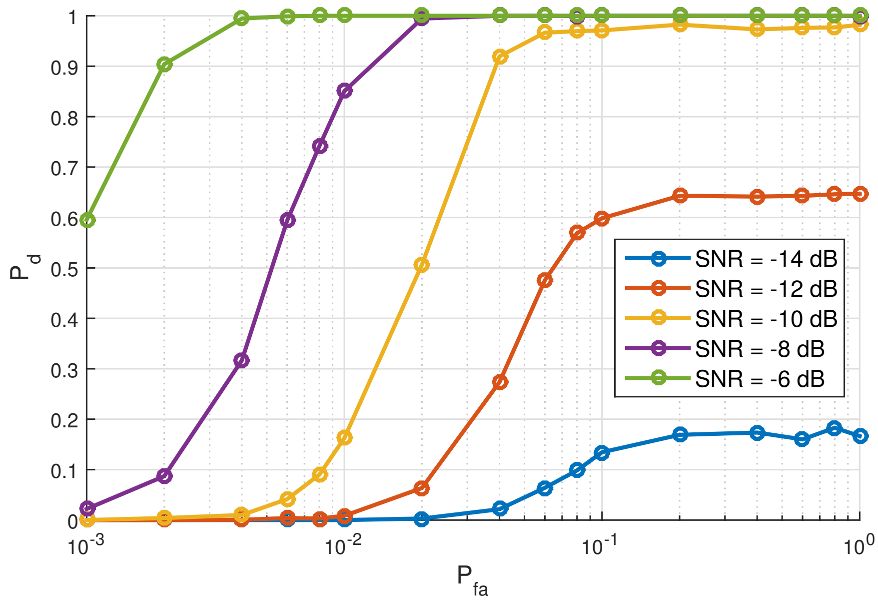

Figure 4.4: ROC of the proposed approach for different SNR values and $N = 5000$

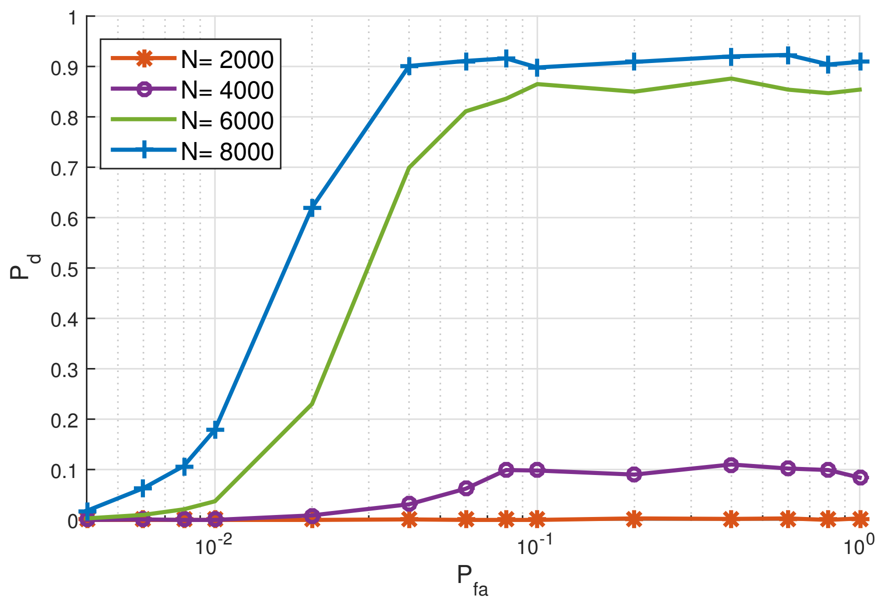

Figure 4.5: ROC of the proposed approach for different sample sizes and SNR $= -12$ dB

set to be randomly occupied with each channel of 200 KHz bandwidth. In this case, the probability of detection (P_D) is referred to how many correct channels are detected along with their correct locations. As it is expected, increasing the false alarm rate (from 0.001 to 0.1) will result in higher probability of detection with the increase of the sample size to 24000 samples accordingly as the number of channels (and hence overall bandwidth) is increased.

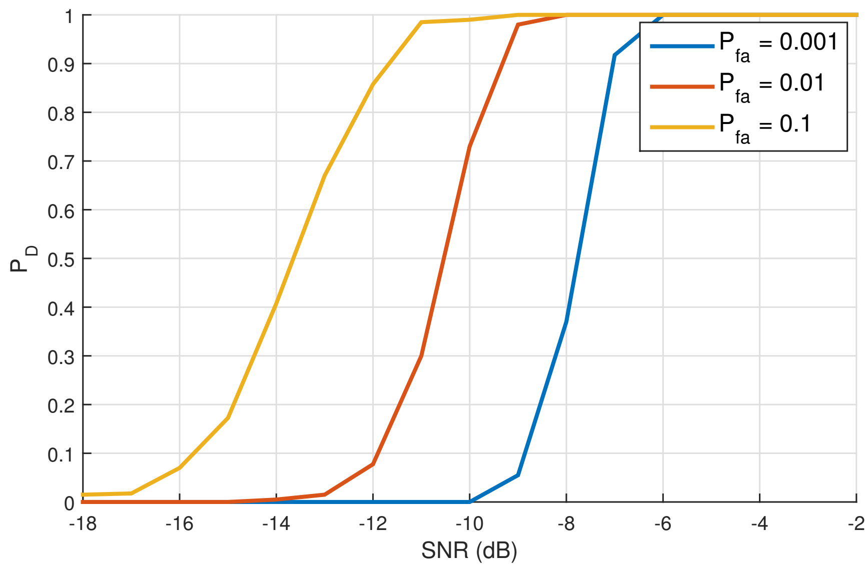

Figure 4.6: A multi-channel detection rate for different false alarm probabilities

Finally, the proposed method (for a single channel) is compared to other known techniques such as the Energy detector model (ED) [102], and Maximum Eigenvalue Detector (MED) [62]. In this comparison, the probability of detection is evaluated with $P_{fa} = 0.1$ and $N = 5000$. The proposed method is adaptive in terms of setting the detection threshold, unlike the ED and MED which are sensitive to noise uncertainty [103]. In Fig. 4.7, the comparison between the aforementioned techniques shows that ED and MED outperform the proposed method when the noise power is known a priori. However, in the presence of noise uncertainty, their performance deteriorates significantly. However, under a noise uncertainty of 2 dB, the proposed method exhibits a better performance than both MED-2 (with a smoothing factor = 8) and the ED-2.

4.4.2 The Evaluation of the Proposed Detector Computational Complexity

EMD suffers from a relatively high computational complexity due to its iterative nature (sifting process) which turns to become more significant with large number of processed samples. Conventional EMD approach uses cubic spline interpolation (to connect the maxima/minima points) which is known to be a relatively slow computational tool because it works by

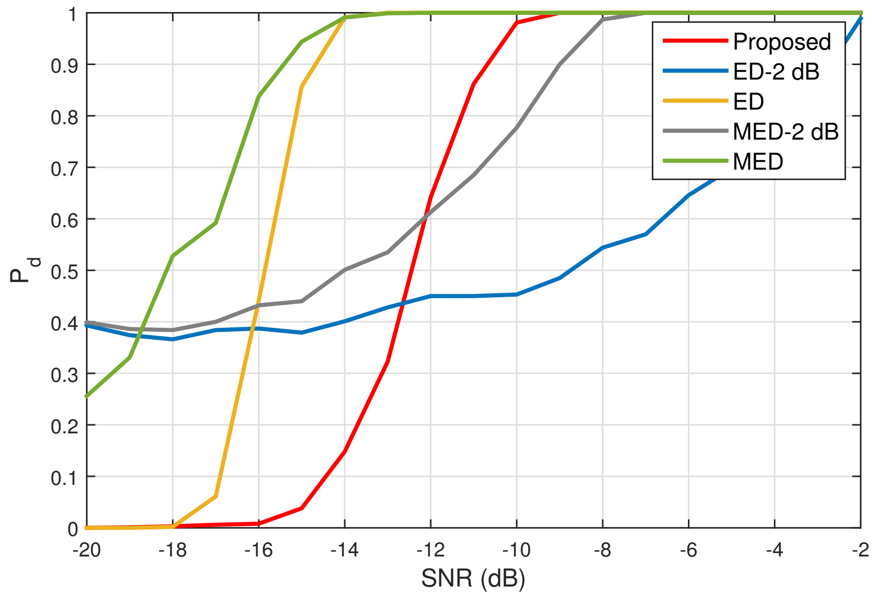

Figure 4.7: A comparison of the proposed method to the ED and MED under noise uncertainty effect with $N = 5000$

solving linear systems. In this section, the computational complexity (time-based) of the proposed approach is discussed and a comparison is made with other EMD-based techniques given in [1, 2, 64].

Wang *et al* showed that the complexity of EMD process is equivalent to Fourier transform but scaled by a factor due to its iterative behavior [12]. For the case where all IMFs are required to be used for detection decision, the computational complexity denoted by $\mathbb{C}^{All}$, can be given as follows:

$$\mathbb{C}^{All} = \mathbf{S} \times \mathcal{O}(N \log N) \tag{4.11}$$

where N is the number of samples, and $\mathbf{S}$ is the sifting multiplier and depends mainly on the processed data length and its smoothness (corresponding to extrema density).

On the other hand, the proposed method needs only to extract IMF_1 for noise power calculations. The computational complexity of IMF_1 sifting process will be denoted by

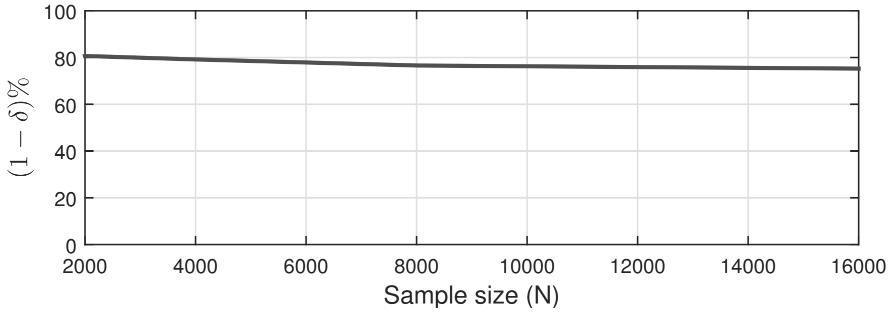

Figure 4.8: The enhancement percentage of using IMF_1 over the use of all IMFs in terms of EMD processing computational complexity

$\mathbb{C}^{First}$, and is given as follows:

$$\mathbb{C}^{First} = \hat{\mathbf{S}} \times \mathcal{O}(N \log N) \tag{4.12}$$

where $\hat{\mathbf{S}}$ is the sifting multiplier of IMF_1, and $\hat{\mathbf{S}} < \mathbf{S}$.

In order to evaluate the computational complexity enhancement of using IMF_1 over the use of all IMFs, an indicator, namely, δ is defined as follows:

$$\delta = \frac{\mathbb{C}^{First}}{\mathbb{C}^{All}} = \frac{\hat{\mathbf{S}}}{\mathbf{S}} \tag{4.13}$$

where $(1-\delta)\%$ is the enhancement percentage (time saving) of using the proposed method over the methods in [1, 2, 64]. Fig. 4.8 illustrates the computational gain, where an average of 75% can be obtained through the use of the proposed method.

In Fig. 4.9, the processing time[2] (to make a detection decision) for the proposed method, ED, MED, and the methods in [1, 2] are compared. From Fig. 4.9, it is obvious that ED outperforms all other techniques in which ED known to be as a fast detector. However, MED exhibits better sensing time than the proposed method and that performance might deteriorate as the smoothing factor increased (smoothing factor in this figure was 8). Despite of ED and MED superiority, the proposed method can be used in sensing as it requires less

[2]The simulations are carried out using Matlab R2015a installed on a computer with the following specifications: Operating system: Windows Server 2008 R2 standard, Processor: Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz, RAM: 32.0 GB, System type: 64-bit

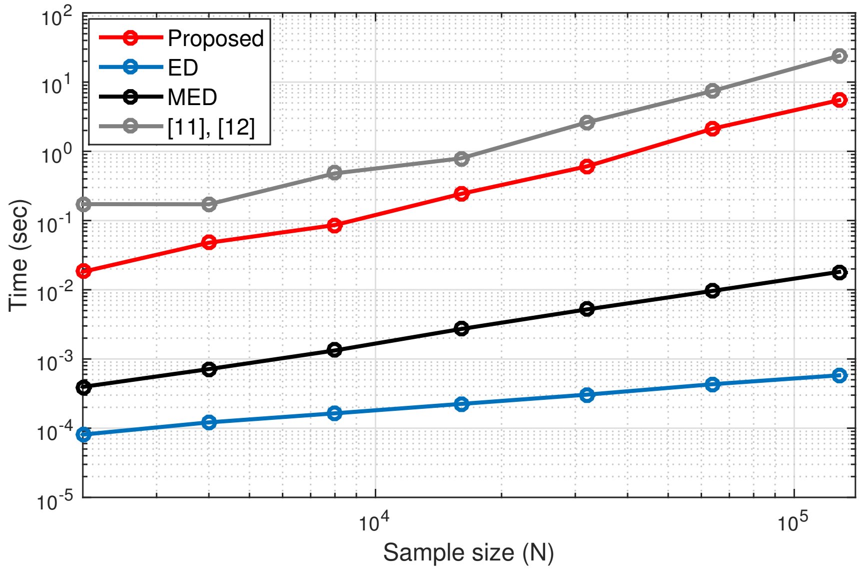

Figure 4.9: Complexity comparison of the proposed method versus ED, MED and conventional EMD-spectrum sensing techniques given in [1, 2]

than 6 *sec* to perform the detection process for $N = 128000$ and that's within the specs of FCC [104].

4.5 Conclusion

In this paper, an IMF_1-based double threshold detector is proposed for single or multi-channel detection in CR networks. The proposed method is adaptive, requires no prior information about the received signal or its noise power and is robust to noise uncertainty. The proposed method, inherits the advantage of CA-CFAR techniques to maintain a constant probability of false alarm while using the first IMF to estimate the noise floor adaptively. Due to the use of one IMF in setting the threshold, the proposed technique is much faster than other conventional EMD-based techniques. The method is evaluated in terms of probability of detection and false alarm rates over different SNR and sampling size values to demonstrate its performance. The results show an enhancement in the probability of detection over other techniques such as ED and MED in the presence of noise uncertainty.

CHAPTER 5

Intrinsic Mode Function-based Noise Power Estimation

In this chapter[1], an algorithm is proposed to blindly estimate the noise power embedded in a received signal and use the knowledge of the noise power to upgrade semi-blind spectrum sensors to blind sensors. The proposed noise estimator is based on a relationship between the noise power of a received signal and the power of the first mode of the Empirical Mode Decomposition algorithm. Simulations and results demonstrate an enhancement to the detectors performance in contrast with their conventional semi-blind versions.

5.1 Intrinsic Mode Functions Statistical Analysis

The first Contribution will be an expansion for the EMD statistical investigations presented in Sec. 3.2.2 and Sec. 4.3. Our investigation will be steered to a thorough analytical study of the Intrinsic Mode Function probability distribution and provide a statistical proof of the best fitting distribution model. On the other hand, the first Intrinsic Mode Function properties is further evaluated by examining the relationship between its probability second moment and a Gaussian random process.

5.2 Introduction

Spectrum sensing is a key part of cognitive radio (CR) functionality in which the secondary users (SU) detect spectrum owned by a licensed or primary user (PU) [17]. Different spectrum sensing methods have different levels of "blindness". The term "blindness" refers to the

[1]This work is accepted for publication in IEEE Signal Processing Letters 2017 [105] and some results are published in [21]

ability of the spectrum sensor to work without *a priori* knowledge of PU channel statistics, such as noise power [106]. There is a wide range techniques for non-blind (Waveform), semi-blind (Energy Detector) and fully-blind spectrum (Maximum-Minimum Eigenvalue) sensing [61].

Typically, the noise power of the received signal is used to calculate the detection threshold, which is in turn utilized to decide the occupancy status of the channel under test. Practically, the knowledge of noise power is not available unless sacrificing part of the spectrum for measuring the noise; noise estimation is a substantial step to shift the semi-blind spectrum sensing methods to the blind region. However, an inaccurate estimation of the noise power may result in an increase in the miss-detection and false alarm rates [107]. Noise uncertainty is an additional factor that affects the efficiency of semi-blind detectors in which noise starts to fluctuate during the sensing duration [103].

Several approaches are proposed to estimate the noise power of the received signal and hence improve the detector performance [70, 108–110]. In [108], the noise level of the receiver, including noise sources from radio circuitry, is calibrated prior the detection process. However, the noise variance is assumed to be precisely known to the receiver therefore, this approach is not practical and needs the incorporation of noise uncertainty. In [70, 109], it was assumed that some sub-bands are vacant and hence these sub-bands are used as noise references in which unnecessary sacrificing of the spectrum must be adopted. In [110], forward consecutive mean excision (FCME) and forward cell averaging (CA) methods are used to estimate the level of the noise for a certain false alarm rate however, this approach is sensitive to the designed level of false alarm rate. In [71], the eigenvalue groups of the sample covariance matrix are split using minimum descriptive length (MDL); a goodness of fit for probability distribution of the noise eigenvalues is used to estimate the noise power. This approach requires high number of samples and depends on *ad hoc* parameters such as smoothing factor and bandwidth ratio and hence its implementing is practically challenging. In addition to the aforementioned approaches, Wavelet de-noising can be exploited to estimate the noise power by subtracting the de-noised version from the received noisy signal [10]. However, this method, although shows good noise estimation, suffers from the non-adaptivity in terms of selecting wavelet basis, decomposition level, and number of coefficients [111].

This chapter presents the analytical justification of an Empirical Mode Decomposition-based algorithm that was previously proposed for noise power estimation [21]. The proposed noise estimator is non-parametric and adaptive to the signal in which no predefined parameters or knowledge about the received signal is required. This work includes a detailed performance analysis in comparison to other comparable methods.

5.3 System Model

Similar to the system model introduced in 4.2, we assume a received signal, $y(n)$, carries either Gaussian noise only, $w(n)$, or PU/SU signal components, $s(n)$, contaminated additively with, $w(n)$. Subsequently, a binary hypothesis test can be formulated as follows:

$$y(n) = \begin{cases} w(n) & : \mathcal{H}_0 \\ s(n) + w(n) & : \mathcal{H}_1 \end{cases} \tag{5.1}$$

where $\mathcal{H}_0$ represents the absence, or the failure to prove the presence of PU/SU and $\mathcal{H}_1$ represents the indication of PU/SU presence. It is assumed that $r(n)$ is an over-sampled signal representing a snapshot of the channel with a duration T representing the sensing cycle and samples $n = 1 \ldots N$, where N is the total number of samples. $w(n)$ is modeled as independent and identically distributed additive white Gaussian noise with zero-mean and variance σ_w^2 i.e. $w(n) \sim \mathcal{N}(0, \sigma_w^2)$ and assumed to be independent from $s(n)$.

From frequency-domain perspective, EMD acts like a dyadic filter bank, where the subsequent IMFs (except for the first IMF) behave like overlapping bandpass filters [15]. The core part of the sifting process relies on interpolating the extrema (maxima/minima) points. Therefore, oversampling is required to extract all the local oscillations through the sifting procedure [27].

In this chapter, the first IMF, IMF_1, is used for noise estimation. This IMF exhibits distinct statistical characteristics and was previously exploited for different applications [14, 112, 113]. The EMD sifting process captures the highest frequencies in IMF_1 [15]. However, for noisy signals, IMF_1 might include low-band frequencies (possibly PU/SU signals) when the sampling rate is not sufficient and/or the noise power is too low.

The probability distribution of IMF_1 for an input Gaussian processes is a mix of two normal distributions represented by the Gaussian mixture (bi-modal) distribution [96]. The justification of the bi-modality in such a distribution lies in the large discrepancy of values (in case of noisy or noise only signals) yielded by the maxima and minima envelops [97]. The first IMF, denoted by $c_1(n)$, follows a bi-modal distribution [114]:

$$f\big(c_1(n)\big) = \frac{\varepsilon}{\sqrt{2\pi\sigma_u^2}} e^{-\frac{(c_1(n)-\mu_u)^2}{2\sigma_u^2}} + \frac{1-\varepsilon}{\sqrt{2\pi\sigma_l^2}} e^{-\frac{(c_1(n)-\mu_l)^2}{2\sigma_l^2}} \tag{5.2}$$

where $\mu_u, \sigma_u, \mu_l, \sigma_l$ are the mean and the standard deviation for upper and lower mode distributions respectively, and $\varepsilon \in [0, 1]$ represents the mode distribution weight.

5.4 Proposed Noise Estimation Method for Semi-blind Spectrum sensors

According to Sec. 5.3, IMFs can be interpreted as a dyadic filter bank that resembles the behavior of wavelets. However, unlike the filtering properties of wavelets, the EMD non-linear decomposition process introduces different cutoff frequencies [98]. Figure 5.1 empirically illustrates the normalized frequency response for the power spectral density (PSD) of $c_1(n)$ of a noise only input signal $w(n)$ for 5000 averaged trials. This figure shows the adaptive high-pass filtering characteristic of IMF_1 and how the sampling rate impacts the cutoff frequency. The first IMF, under sufficient sampling rates, is dominated by noise and thus can be exploited to estimate the noise power of a received signal [52]. Next, we propose a method to determine the total noise power from these observations.

An empirical ratio of the first IMF power to the total noise power of the received signal denoted by β is:

$$\beta = \frac{\hat{\sigma}^2_{c_1}}{\sigma^2_w} \tag{5.3}$$

where $\hat{\sigma}^2_{c_1}$ is the estimated variance of $c_1(n)$ (see Appendix 5.A). Further, an analytical explanation of the proposed noise estimation ratio is described in the Appendix 5.B.

The ratio, β, plays the role of a scaling factor that can be used to estimate σ^2_w. Figure 5.2 shows a comparison of β^2, dashed line, to the best fit linear regression model, solid line, when $r(n) = w(n)$. It was found that β can be approximated by a simple function of the sample size, N, of the received signal [21]:

$$\beta(N) = \mathbb{S}\log_2(N) + \beta(1) \tag{5.4}$$

where $\beta(1)$ is the y-intercept of the best fit linear model using polynomial least-squares and $\mathbb{S}$ is the linear fit slope. From Fig. 5.2, the sample size, N, is represented logarithmically in order to linearize the trend of β. These β values maintain a linear trend over different sample size values and validate the model in (5.4). According to (5.4), and for a sample size N, the estimated noise power, $\hat{\sigma}^2_w$, of the received signal $r(n)$ can be given as:

$$\hat{\sigma}^2_w = \frac{\hat{\sigma}^2_{c_1}}{\beta(N)} \tag{5.5}$$

In this chapter, the noise power estimation in (5.5) is used to upgrade two semi-blind methods, energy detector (ED) [103] and the maximum eigenvalue detector (MED) [62], into blind ones. The noise power of these semi-blind methods is replaced by the proposed estimated noise power, $\hat{\sigma}^2_w$.

[2] In Monte-Carlo simulation, a Gaussian random process of each sample size, N, is generated for 5000 trials. The estimated noise power, $\hat{\sigma}^2_{c_1}$, is obtained for each trial and the β value of each sample size is the result of averaging of all trials.

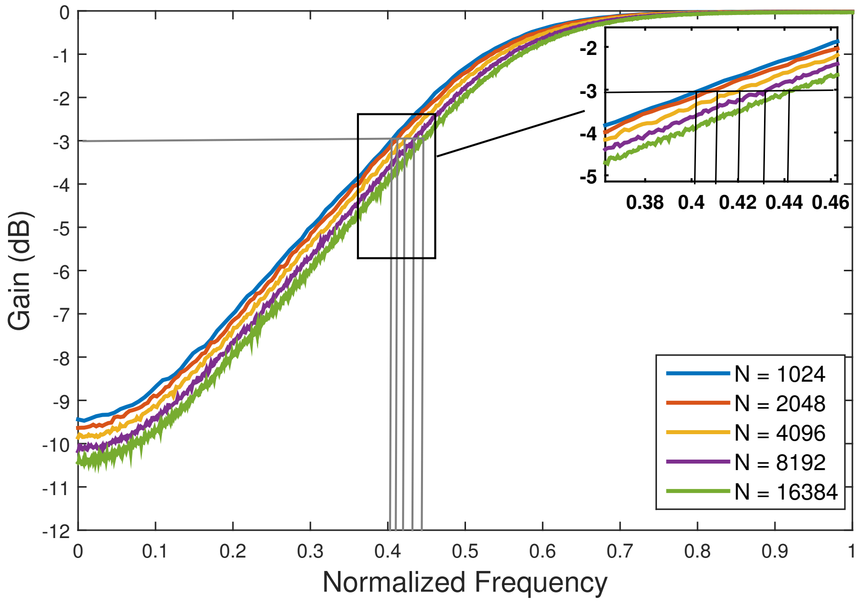

Figure 5.1: Normalized frequency response of IMF_1 for different values of N

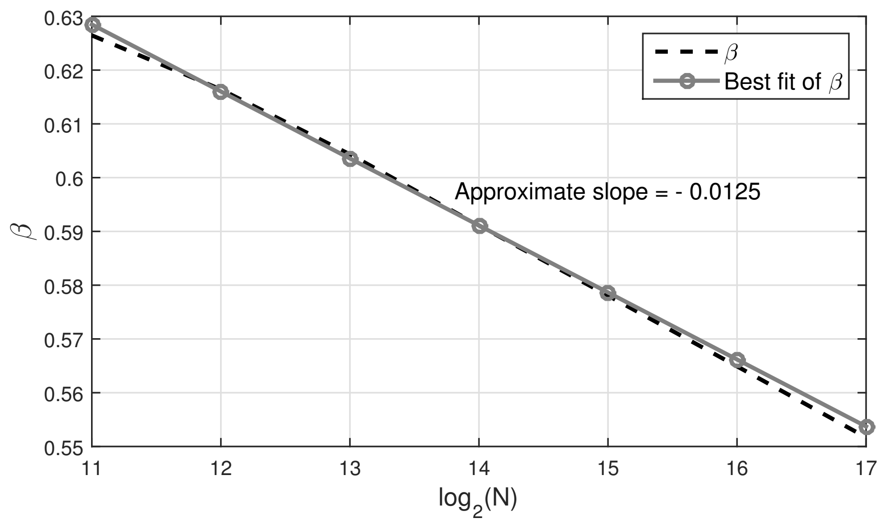

Figure 5.2: The relationship between β and $\log_2(N)$

5.5 Simulation and Results

We assume a scenario of one PU/SU node with an additive white Gaussian noise (AWGN) channel and a single channel spectrum scanner with a band pass filter (BPF). An OFDM-modulated communication signal is synthesized with a known noise power. The results are carried out through a Monte-Carlo simulation by averaging 2000 runs.

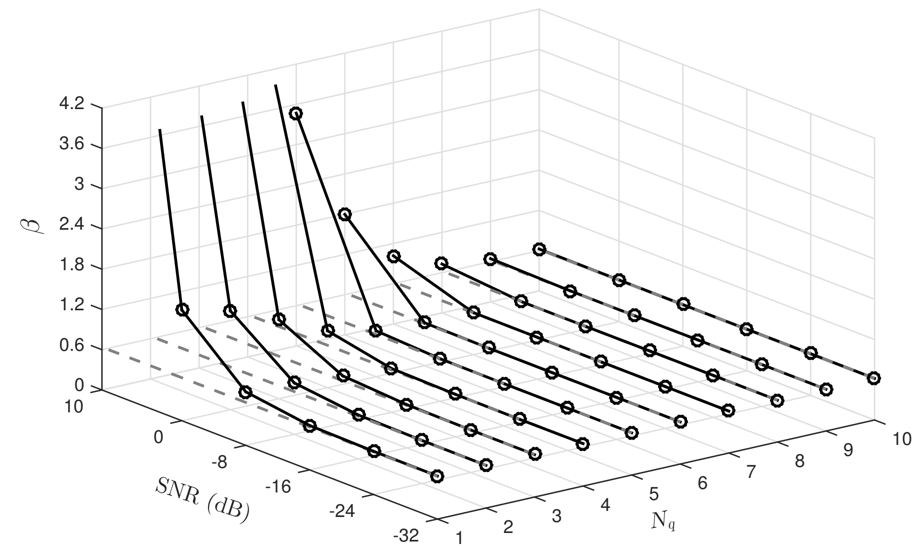

Figure 5.3: β vs. number of samples and SNR values

First, we show the boundaries of the proposed noise estimation method under different values of Signal-to-Noise Ratio (SNR) and sampling rates, Nyquist rate N_q. In Fig. 5.3, a received Gaussian noise-only signal (dotted gray line) is used as a reference to compare it with the scenario of a noisy signal. It is obvious that, for low sampling rate, β, from (5.4), will perform well only at low SNR values (<-8 dB) as the noise will dominate the signal resembling the case of noise-only scenario. However, for high sampling rate ($\geq 8N_q$), β will behave in a very similar way to the noise-only case. The deviation of β at low sampling rate and high SNR values occurs because IMF_1 sifts part of the signal in addition to the noise and increases the corresponding variance. Therefore, to ensure the proper functionality of the proposed β model at wide range of SNR values, oversampling is required.

The proposed noise estimation model (5.4) is further evaluated using a percentage error metric in which the true noise, σ_w^2, of the received signal is used as a reference. Table 5.1 shows a comparison between the proposed β model and two other transform-based estimation techniques; Forward consecutive mean excision (FCME)[3] and Wavelet[4] noise estimation

[3]While the FCME algorithm is in the frequency-domain, we transform the noise samples to the time-domain for comparison [94].

[4]In this chapter, a *Symlets* mother-wavelet of order 16, 7^{th} decomposition level, heuristic variant hard

Table 5.1: Percentage error for transform-based noise estimation methods with 2000 samples, or $8N_q$, and different SNR values

Estimation method	**SNR (dB)**				
	-30	**-20**	**-10**	**0**	**10**
	Estimation error (%)				
Wavelet	1.49	**0.90**	6.53	18.4	21.81
Forward	**1.00**	1.10	9.86	11.78	13.61
β	2.75	2.85	**2.83**	**2.84**	**9.65**

methods. Table 5.1 shows that at low SNR values (<-10 dB) both the FCME and wavelet methods performed slightly better than the proposed method. However, both methods exhibit a rapid degradation in the noise estimation performance as the SNR increases (≥ -10 dB) and that is expected due to the domination of the signal features over the noise in which these techniques fail to differentiate the signal from the noise. In comparison, β method has steady performance over all SNR values except at SNR $\geq$ 10 dB where IMF_1 starts to sift signal components in addition to the noise. The degradation of the proposed model can be mitigated by increasing the sampling rate, which guarantees that IMF_1 will be dominated by noise samples only.

In Fig. 5.4, the probability of detection, P_d, is used as a metric to evaluate the performance of ED modified with our proposed β technique in contrast to FCME, Wavelet, and ED with the known or true noise[5] with probability of false alarm, P_{fa} at 0.1 and SNR $= -12$ dB. From this figure, the β method performs nearly as good as the true noise especially at low number of samples.

The Receiver Operating Characteristic (ROC) is used as a performance metric for both ED and MED using different noise estimation methods in which 2000 samples represent a sensing cycle of 625 μs at $8N_q$; the SNRs in Figs. 5.5 and 5.6 were chosen to best reflect the performance of each method. In Fig. 5.5, the performance of the β modified ED technique

threshold and first level coefficients threshold rescaling is utilized to estimate the noise power [10].

[5]The true noise is the variance of the noise, $w(n)$ from (5.1).

is illustrated, at SNR of -10 dB, for different transform-based noise estimation methods as well as the model-based method given in [71] with smoothing factor of 50. Figure 5.6 illustrates the performance of the β modified MED technique at SNR of -14 dB for the same noise estimation methods. For comparison, a fully-blind spectrum sensing technique, Maximum-Minimum Eigenvalue (MME), computed with 6000 samples is added to Fig. 5.6.

From Figs. 5.5 and 5.6, it is demonstrated that the proposed β noise technique outperforms other estimators in addition to the MME technique which requires a much higher sampling rate. The β method performs close to true noise for the assigned SNR value and number of samples. In addition, the degradation of other method's performance is due to the estimation error as SNR values increase. While Wavelet and FCME show slightly better noise estimation error (Table 5.1) compared with the β method, that does not mean they can be used for very low SNR detection. The reason is that for the number of samples used in this work each detector will have an SNR wall that can not be exceeded even if the number of samples is increased infinitely [103].

Finally, the performance of the proposed noise estimation method is compared to the noise estimation given in [115] using ED. Figure 5.7 illustrates the detection probability of the ED methods using the DVB-T model parameters with $P_{fa} = 0.01$ [115]. In addition, the noise uncertainty (1 dB) is presented to reveal the effect of noise fluctuation on the detector performance. The ED with the true noise exhibits the best performance in contrast to the worst performance of 1-dB noise uncertainty. Our proposed noise estimation method shows a slightly better performance compared to the method in [115]. Unlike the noise power estimation in [115], which is specifically designed for DVB-T signals, the proposed method can work for a wider range of modulation schemes.

5.6 Conclusion

In this chapter, a signal driven noise power estimation method is proposed and used to make two conventional semi-blind detectors blind. The proposed algorithm takes the advantage

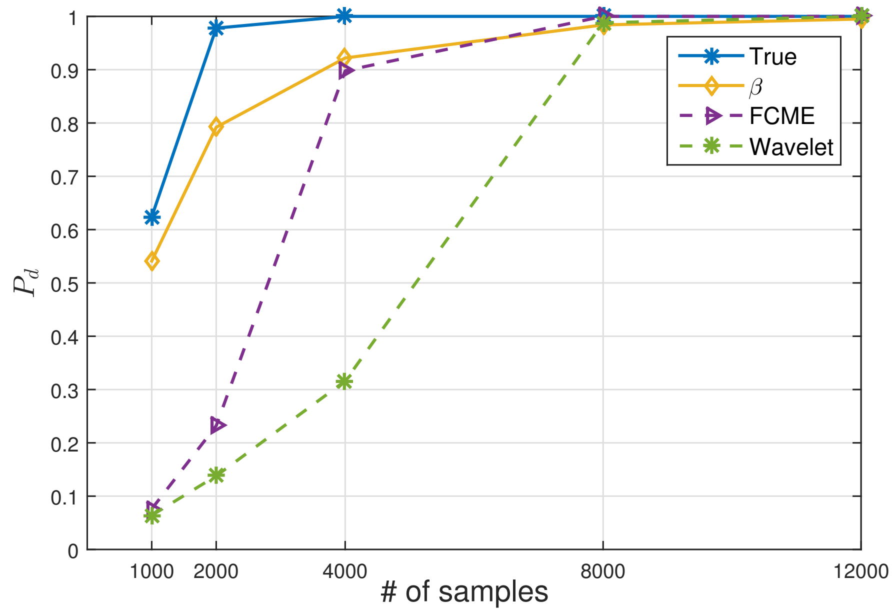

Figure 5.4: No. of samples in respect to P_d

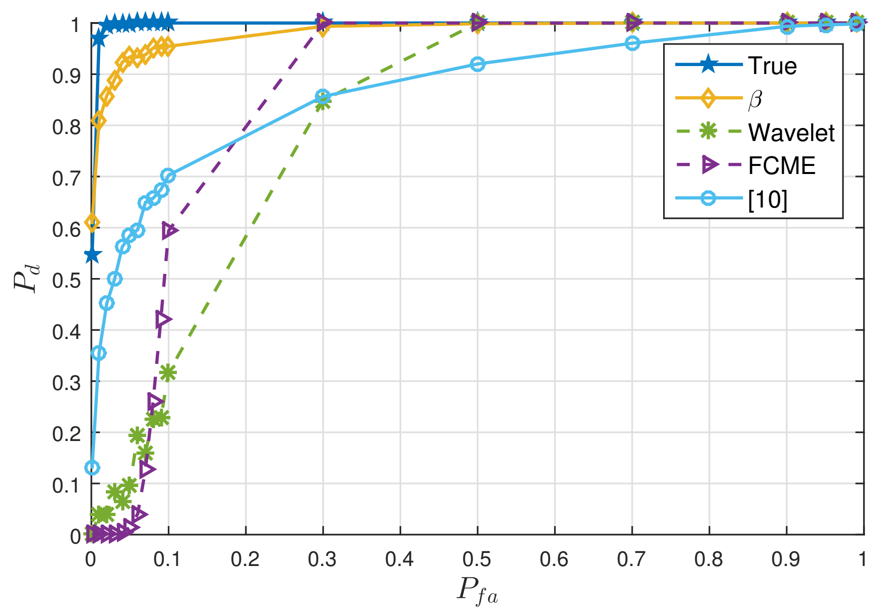

Figure 5.5: ROC curves of ED

of an observed unique ratio, β, between the IMF_1 power and the total noise power in the received signal. The performance of the proposed method is further tested through upgraded detectors and the results are compared to other estimation schemes. The proposed noise estimation method outperforms other schemes at low SNR.

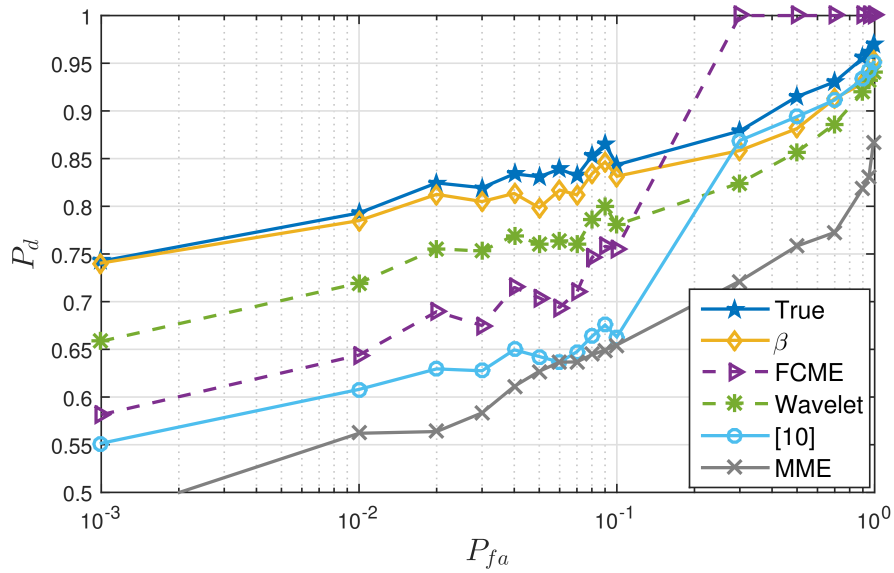

Figure 5.6: ROC curves of MED and MME

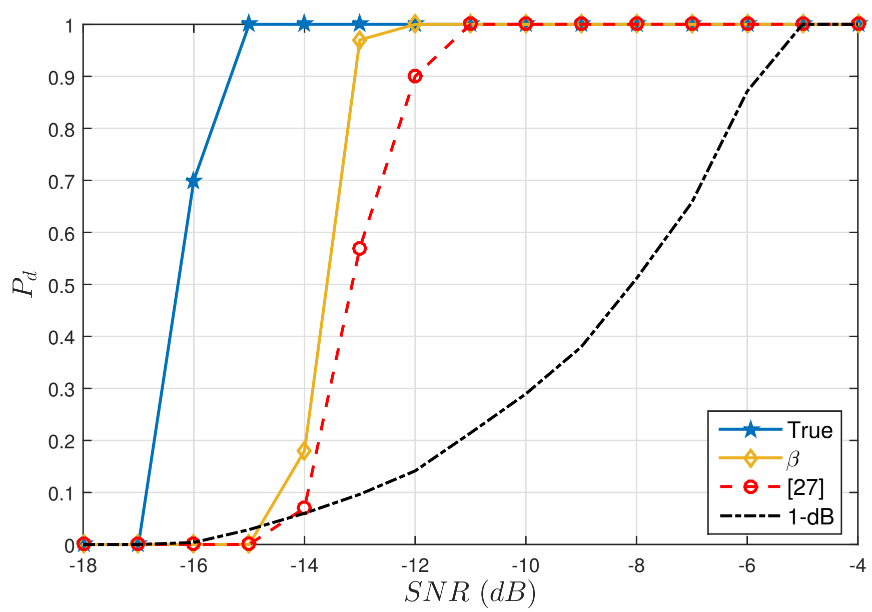

Figure 5.7: P_d vs SNR

Appendix

5.A

The first IMF, $c_1(n)$, can be modeled as a bi-modal zero-mean normal process with a variance of $\sigma^2_{c_1}$ i.e. $c_1(n) \sim \mathcal{N}(0, \sigma^2_{c_1})$. The variance of each mode distribution in (5.2) is estimated using maximum-likelihood parameter estimation via Expectation Maximization (EM) algorithm [116]. The estimated overall variance of IMF_1, denoted by $\hat{\sigma}^2_{c_1}$, can be given as [117].

$$\hat{\sigma}^2_{c_1} = \varepsilon\hat{\sigma}^2_u + (1-\varepsilon)\hat{\sigma}^2_l + \varepsilon(1-\varepsilon)(\hat{\mu}_u - \hat{\mu}_l)^2 \tag{5.6}$$

where $\hat{\sigma}^2_u$, $\hat{\sigma}^2_l$, $\hat{\mu}_u$ and $\hat{\mu}_l$ are the estimated variances and means of the upper and lower mode distributions. The bi-modality of $c_1(n)$ can be attributed to the inherent switching between two mutually exclusive Gaussian processes of different means. For simplification, the mode distribution weight ε is assumed to be 0.5 and that assumption is rationalized by the fact that maxima and minima of the upper and lower envelopes are almost equally likely and symmetrically distributed around the zero overall mean of the signal.

5.B

Analytically, the scaling factor β can be expressed as the ratio of integrating the PSD of IMF_1 and the received signal, $r(n) = w(n)$, in terms of extrema (maxima/minima) distribution. The extrema are equally spaced with the maxima being located at integer time instants and the minima at half the distance between a pair of consecutive maxima. For the case of cubic spline interpolation, the frequency response of the unit spaced knots $I(\nu)$ is given as [118]:

$$I(\nu) = \left(\frac{\sin \pi\nu}{\pi\nu}\right)^4 \frac{3}{2 + \cos 2\pi\nu} \tag{5.7}$$

The PSD of the first IMF, $S_{c_1}(\nu)$, is given by:

$$S_{c_1}(\nu) = |(1 - I(3\nu))|^2 \; S_w(\nu) \quad (0 \leq \nu < 1/2) \tag{5.8}$$

where $S_w(\nu)$ is the PSD of the received signal $w(n)$. Thus, the corresponding ratio $\hat{\beta}$ is given by:

$$\hat{\beta} = \frac{\int_0^{1/2} S_{c_1}(\nu)d\nu}{\int_0^{1/2} S_w(\nu)d\nu} \tag{5.9}$$

The scaling factor, $\hat{\beta}$, as given in (5.9) is the result of the first iteration (through the sifting process) to obtain IMF_1. The $\hat{\beta}$ in (5.9) is not as generic as the one given in (5.4) however, it is presented here to provide further validating evidence for the proposed noise estimator.

CHAPTER 6

Conclusion Remarks

6.1 Conclusion

Empirical Mode Decomposition is an adaptive, signal-dependent algorithm that shown to be promising in different area of research due to its superiority against the most widely used signal analysis method. The lacking of a mathematical bases encourage the researchers to understand its underlying behavior and leverage that into important application problems. Despite its iterative nature, which leads to computational complexity, EMD is still one of the most widely used in signal decomposition relying on the robustness provided by the adaptive and blind sifting process.

In this dissertation, the statistical properties of the Empirical Mode Decomposition (EMD) algorithm is investigated and analyzed to have better understanding of the empirical structure of the algorithm. The understanding of EMD behavior, under different scenarios, is useful to leverage its capabilities into practical applications. These application ranging from signal de-noising in time-domain to spectrum sensing development and not ending at noise power estimation. The success of EMD lies in the domination over the other well-known conventional signal analysis techniques such as Fourier and Wavelet tools although its lacking for a rigorous mathematical framework.

In Chapter 1, we summarized the motivation behind using the EMD and briefly explained the importance of understanding its statistical features. Further, the significance of the proposed de-noising and detection problems is emphasized and concluding the chapter by outlining the contributions of the dissertation.

Chapter 2 introduced the background of the EMD algorithm with an explanation of the sifting process which is illustrated using two examples. Because EMD is iterative process, different stoppage criteria from literature is summarized in addition to the interpolation methods that is used to connect the extrema points during the sifting. As EMD designed analyzed continuous signal, the effect of sampling rate on the algorithm performance is presented. The chapter is ended by a summary of EMD applications in order to underline and justify the selection of this algorithm over other widely used ones to be the main topic of this dissertation.

The first novel work is covered in Chapter 3, in which a generalized fitting distribution is proposed to model the IMFs. Chapter 3 starts by investigating the effect of changing the signal distribution with respect to IMF distributions where two statistical measures, variance and excess kurtosis, are used to understand that effect. Further, these signal distributions are corrupted by white Gaussian noise to study the influence of this inevitable component (i.e. noise). The conclusion of the aforementioned experiments was that, for the distributions we examined, the the first set of IMFs will follow a distribution rather than the Gaussian which is well-known in literature. The next was to trace the statistical distributions of the sifting process, particularly $\mathrm{IMF}_2(n)$, in order to suggest a better fitting distribution. The mean of envelopes, m_i and the residue, r_i, signals are chosen to be under statistical consideration as these two signals are the main components during the iterative sifting process. The tracing of m_i and r_i distributions show that they can be modeled by generalized Gaussian distribution, GGD, and that model is verified through a null hypothesis test. The use of different signal distribution with and without wGn corruption lead to a conclusion that the GGD-based IMFs can be used effectively to separate the noise from useful signal components or simply called de-noising. The proposed novel de-noising scheme is working by partially reconstructed the IMFs that are not following GGD and discard the others. The proposed scheme is evaluated under different scenarios and noise models to validate its efficacy. The simulation and results show that the proposed de-noising scheme outperforms other de-noising techniques under low

SNR regimes.

In chapter 4, we presented the second application of EMD for signal detection in cognitive radio networks. In that chapter, the first IMF is utilized to propose a spectrum sensing method that can make a decision of channel vacancy without a predefined parameters or knowledge of the received signal noise characteristics. Further, it is robust for noise uncertainty which is a crucial factor in signal detection due to the temporal fluctuation of noise power. The proposed method involves the advantage of cell-averaging constant false alarm rate, CA-CFAR, techniques to maintain a constant probability of false alarm. The CA-CFAR requires a vacant channels (noise-only) to estimate the noise floor therefore we suggested to use the first IMF to estimate the noise floor adaptively. Due to the use of one IMF in setting the threshold, the proposed technique is much faster than other conventional EMD-based techniques. The method is evaluated in terms of probability of detection and false alarm rates over different SNR and sampling size values to demonstrate its performance. The results show an enhancement in the probability of detection over other techniques such as energy detectors, ED, and maximum eigenvalue detectors, MED, in the presence of 2 dB of noise uncertainty.

Chapter 5 presents our original work for the noise power estimation based on the noise power of the first Intrinsic Mode Function, $\text{IMF}_1(n)$. The filtering characteristics of EMD encouraged the authors to investigate the ratio of the $\text{IMF}_1(n)$ to the total noise power in the received noisy signal. The purpose was to shift the most widely used energy detector from semi-blind to blind region where the conventional ED require the knowledge of the received signal noise power. The ratio of the two noise powers, denoted by β, is modeled in terms of number of samples and the validity boundaries are evaluated numerically. The power (variance) of the first IMF is estimated based on its probability distribution. Further, an analytical explanation of the proposed ratio is provided. The simulation and results reveal an enhancement in the detectors performance in contrast with other peer noise estimation methods.

6.2 Limitations

The capabilities of EMD come on the expense of deriving a mathematical model. Although the empirical nature of EMD keeps the doors open for further research, the lacking for a strict mathematical frame can be considered as the main limitation. We summarized our observations for the EMD limitations as follows:

1. The computational complexity: the main complexity contribution comes from the iterative use of spline interpolation during the sifting process. The complexity level increase with the increase of the number of samples of the processed signal.

   (a) As the EMD is designed to process continuous signals, the conventional version of EMD requires high sampling rate to capture all features of the underlying components. Therefore, the level of complexity will be significant to extract all IMFs.

   (b) The required high sampling rate might limit the ability of EMD to process wide-band frequency signals signals and therefore the conventional EMD is impractical to use.

2. The ability of EMD to derive its decomposition bases adaptively (signal-dependent), leads to variant cut-off frequencies and that make the frequencies separation more complicated.

6.3 Future Work

For the future work, the developed EMD-based techniques in this dissertation can be further improved and new applications can be involved. more specifically, an optimization of the shape parameter, κ, search during the GGD null hypothesis test can be investigated. In addition, the related de-noising technique can be even more enhanced by replacing the hard thresholding by soft thresholding in which each relevant IMF can be filtered for better de-noising. On the other hand, the EMD algorithm can be further developed to be able to process signals in bandpass region rather than the baseband for spectrum sensing purposes. That can be achieved by using different sifting methods and/or replace the spline interpolation with less complex method to reduce the latency of EMD to process large data samples and retain the adaptive behavior in same time. Finally, the EMD-based noise estimation can be evaluated to include signals under different channels other than the AWGN.

At the end, we want to emphasize that, being empirical, our investigation and contributions to EMD opens further doors to study more aspects of this interesting field. Our motivation is reinforced by the success of the algorithm in real-world applications and its simple structure compared to other complex, though famous, techniques.

LIST OF REFERENCES

[1] Mahdi H Al-Badrawi and Nicholas J Kirsch. An emd-based spectrum sensing technique for cognitive radio networks. In *Vehicular Technology Conference (VTC Fall), 2015 IEEE 81st*, pages 1–5. IEEE, 2015.

[2] Mahdi H Al-Badrawi and Nicholas J Kirsch. An emd-based double threshold detector for spectrum sensing in cognitive radio networks. In *Vehicular Technology Conference (VTC Fall), 2015 IEEE 82nd*, pages 1–5. IEEE, 2015.

[3] Samuel S.P. Shen. Norden E. Huang. *The Hilbert-Huang Transform and its applications.* World Scientific Publishing Co. Pte. Ltd., 2005.

[4] Norden E. Huang and Zhaohua Wu. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics*, 46(2), 2008.

[5] David Havelock, Sonoko Kuwano, and Michael Vorländer. *Handbook of signal processing in acoustics*. Springer Science & Business Media, 2008.

[6] Adam H Monahan, John C Fyfe, Maarten HP Ambaum, David B Stephenson, and Gerald R North. Empirical orthogonal functions: The medium is the message. *Journal of Climate*, 22(24):6501–6514, 2009.

[7] A Hannachi. A primer for eof analysis of climate data. *Department of Meteorology, University of Reading*, pages 1–33, 2004.

[8] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 454, pages 903–995. The Royal Society, 1998.

[9] Pierre Duhamel and Martin Vetterli. Fast fourier transforms: a tutorial review and a state of the art. *Signal processing*, 19(4):259–299, 1990.

[10] David L Donoho. De-noising by soft-thresholding. *IEEE transactions on information theory*, 41(3):613–627, 1995.

[11] Paul Fieguth. *Statistical image processing and multidimensional modeling*. Springer Science & Business Media, 2010.

[12] Yung-Hung Wang, Chien-Hung Yeh, Hsu-Wen Vincent Young, Kun Hu, and Men-Tzung Lo. On the computational complexity of the empirical mode decomposition algorithm. *Physica A: Statistical Mechanics and its Applications*, 400:159–167, 2014.

[13] John OToole, Mostefa Mesbah, and Boualem Boashash. A discrete time and frequency wigner-ville distribution: properties and implementation. In *Proc. Int. Conf. on Digital Signal Processing and Comm. Systems*, volume 400, pages 19–21, 2005.

[14] Yannis Kopsinis and Stephen McLaughlin. Development of emd-based denoising methods inspired by wavelet thresholding. *IEEE Transactions on signal Processing*, 57(4):1351–1362, 2009.

[15] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *Signal Processing Letters, IEEE*, 11(2):112–114, 2004.

[16] A. Komaty, A.-O. Boudraa, B. Augier, and D. Dare-Emzivat. EMD-based filtering using similarity measure between probability density functions of IMFs. *Instrumentation and Measurement, IEEE Transactions on*, 63(1):27–34, Jan 2014.

[17] T. Yucek and H. Arslan. A survey of spectrum sensing algorithms for cognitive radio applications. *Communications Surveys Tutorials, IEEE*, 11(1):116–130, 2009.

[18] M. H. Al-Badrawi, B. Z. Al-Jewad, W. J. Smith, and N. J. Kirsch. A de-noising scheme based on the null hypothesis of intrinsic mode functions. *IEEE Signal Processing Letters*, 23(7):924–928, July 2016.

[19] M.H. Al-Badrawi, A.M Nasr, B.Z. Al-Jewad, and N.J. Kirsch. An adaptive energy detection scheme using EMD for spectrum sensing. *Consumer Communications and Networking Conference (CCNC 2017).* to be published.

[20] Mahdi H Al-Badrawi, Nicholas J Kirsch, and Bessam Z Al-Jewad. An intrinsic mode function based energy detector for spectrum sensing in cognitive radio. In *Computing, Networking and Communications (ICNC), 2017 International Conference on*, pages 131–136. IEEE, 2017.

[21] Mahdi H Al-Badrawi, Nicholas J Kirsch, and Bessam Z Al-Jewad. Noise estimation for spectrum sensing schemes. In *Vehicular Technology Conference (VTC-Fall), 2016 IEEE 84th*, pages 1–2. IEEE, 2016.

[22] Norden E Huang, Man-Li C Wu, Steven R Long, Samuel SP Shen, Wendong Qu, Per Gloersen, and Kuang L Fan. A confidence limit for the empirical mode decomposition and hilbert spectral analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 459, pages 2317–2345. The Royal Society, 2003.

[23] Gabriel Rilling, Patrick Flandrin, Paulo Goncalves, et al. On empirical mode decomposition and its algorithms. In *IEEE-EURASIP workshop on nonlinear signal and image processing*, volume 3, pages 8–11. IEEER, 2003.

[24] Geoffrey GS Pegram, Murray C Peel, and Thomas A McMahon. Empirical mode decomposition using rational splines: an application to rainfall time series. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 464, pages 1483–1501. The Royal Society, 2008.

[25] A Roy and JF Doherty. Improved signal analysis performance at low sampling rates using raised cosine empirical mode decomposition. *Electronics letters*, 46(2):176–177, 2010.

[26] Nathan Stevenson, Mostefa Mesbah, and Boualem Boashash. A sampling limit for the empirical mode decomposition. In *ISSPA 2005: The 8th International Symposium on Signal Processing and its Applications, Vols 1 and 2, Proceedings*, volume 2, pages 647–650. IEEE, 2005.

[27] G. Rilling and P. Flandrin. On the influence of sampling on the empirical mode decomposition. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 3, pages III–III, May 2006.

[28] Chih-Yuan Tseng and HC Lee. Entropic interpretation of empirical mode decomposition and its applications in signal processing. *Advances in Adaptive Data Analysis*, 2(04):429–449, 2010.

[29] Yannis Kopsinis and Steve McLaughlin. Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach. *IEEE Transactions on Signal Processing*, 56(1):1–13, 2008.

[30] P. Kumari and A. Vaish. Information-theoretic measures on intrinsic mode function for the individual identification using eeg sensors. *IEEE Sensors Journal*, 15(9):4950–4960, Sept 2015.

[31] P. C. Petrantonakis and L. J. Hadjileontiadis. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Transactions on Affective Computing*, 1(2):81–97, July 2010.

[32] S. M. S. Alam and M. I. H. Bhuiyan. Detection of seizure and epilepsy using higher order statistics in the emd domain. *IEEE Journal of Biomedical and Health Informatics*, 17(2):312–318, March 2013.

[33] Ulas Gunturkun. Bivariate empirical mode decomposition for cognitive radar scene analysis. *Signal Processing Letters, IEEE*, 22(5):603–607, 2015.

[34] I. Mostafanezhad, E. Yavari, O. Boric-Lubecke, V. M. Lubecke, and D. P. Mandic. Cancellation of unwanted doppler radar sensor motion using empirical mode decomposition. *IEEE Sensors Journal*, 13(5):1897–1904, May 2013.

[35] Y. Sherr, Y. Zhang, and Z. Wang. Satellite fault diagnosis method based on predictive filter and empirical mode decomposition. *Journal of Systems Engineering and Electronics*, 22(1):83–87, Feb 2011.

[36] A. Roy and J. F. Doherty. Overlay communications using empirical mode decomposition. *IEEE Systems Journal*, 5(1):121–128, March 2011.

[37] A. Roy and J. F. Doherty. Covert communications based on signal overlay. *Advances in Adaptive Data Analysis*, 02(03):295–311, 2010.

[38] C. Bektas, A. Akan, S. Kent, and S. Baykut. Spectrum sensing using empirical mode decomposition and relative entropy. In *Signal Processing and Communications Applications Conference (SIU), 2013 21st*, pages 1–4, April 2013.

[39] Stuti Shukla, Sukumar Mishra, and Bhim Singh. Empirical-mode decomposition with hilbert transform for power-quality assessment. *IEEE transactions on power delivery*, 24(4):2159–2165, 2009.

[40] SJ Loutridis. Damage detection in gear systems using empirical mode decomposition. *Engineering Structures*, 26(12):1833–1841, 2004.

[41] YL Xu and J Chen. Structural damage detection using empirical mode decomposition: experimental investigation. *Journal of engineering mechanics*, 130(11):1279–1288, 2004.

[42] Darryll Pines and Liming Salvino. Structural health monitoring using empirical mode decomposition and the hilbert phase. *Journal of sound and vibration*, 294(1):97–124, 2006.

[43] T. Hasan and M. K. Hasan. Suppression of residual noise from speech signals using empirical mode decomposition. *IEEE Signal Processing Letters*, 16(1):2–5, Jan 2009.

[44] S. T. Pan and X. Y. Li. An fpga-based embedded robust speech recognition system designed by combining empirical mode decomposition and a genetic algorithm. *IEEE Transactions on Instrumentation and Measurement*, 61(9):2560–2572, Sept 2012.

[45] K. Khaldi and A. O. Boudraa. Audio watermarking via emd. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(3):675–680, March 2013.

[46] Jean Claude Nunes, Yasmina Bouaoune, Eric Delechelle, Oumar Niang, and Ph Bunel. Image analysis by bidimensional empirical mode decomposition. *Image and vision computing*, 21(12):1019–1026, 2003.

[47] Ning Bi, Qiyu Sun, Daren Huang, Zhihua Yang, and Jiwu Huang. Robust image watermarking based on multiband wavelets and empirical mode decomposition. *IEEE Transactions on Image Processing*, 16(8):1956–1966, 2007.

[48] Anna Linderhed. Compression by image empirical mode decomposition. In *IEEE International Conference on Image Processing 2005*, volume 1, pages I–553. IEEE, 2005.

[49] Zhaohua Wu and Norden E Huang. A study of the characteristics of white noise using the empirical mode decomposition method. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 460, pages 1597–1611. The Royal Society, 2004.

[50] Gaston Schlotthauer, Maria Eugenia Torres, Hugo L. Rufiner, and Patrick Flandrin. EMD of gaussian white noise: effects of signal length and sifting number on the statistical properties of Intrinsic Mode Functions. *Advances in Adaptive Data Analysis*, 01(04):517–527, 2009.

[51] Reza Sameni, Mohammad B Shamsollahi, Christian Jutten, and Gari D Clifford. A nonlinear bayesian filtering framework for ecg denoising. *IEEE Transactions on Biomedical Engineering*, 54(12):2172–2185, 2007.

[52] Patrick Flandrin, Paulo Gonçalves, and Gabriel Rilling. EMD equivalent filter banks, from interpretation to applications. In *Hilbert-Huang Transform: Introduction and Applications*, volume 5. World Scientific Publishing Co. Pte. Ltd., 2005.

[53] Albert Ayenu-Prah and Nii Attoh-Okine. A criterion for selecting relevant intrinsic mode functions in empirical mode decomposition. *Advances in Adaptive Data Analysis*, 02(01):1–24, 2010.

[54] A.-O. Boudraa and J.-C. Cexus. Emd-based signal filtering. *Instrumentation and Measurement, IEEE Transactions on*, 56(6):2196–2202, Dec 2007.

[55] A. Karagiannis and Philip Constantinou. Noise-assisted data processing with empirical mode decomposition in biomedical signals. *Information Technology in Biomedicine, IEEE Transactions on*, 15(1):11–18, Jan 2011.

[56] L. Zao, R. Coelho, and P. Flandrin. Speech enhancement with EMD and hurst-based mode selection. *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, 22(5):899–911, May 2014.

[57] Xueyao Li, Wu Wang, and Rubo Zhang. Speech-stream detection with low signal-to-noise ratios based on empirical mode decomposition and fourth-order statistics. In *Computer and Computational Sciences, 2007. IMSCCS 2007. Second International Multi-Symposiums on*, pages 187–191, Aug 2007.

[58]

[59] Harry Urkowitz. Energy detection of unknown deterministic signals. *Proceedings of the IEEE*, 55(4):523–531, April 1967.

[60] Yan Feng and Xiaodong Wang. Adaptive multiband spectrum sensing. *IEEE Wireless Communications Letters*, 1(2):121–124, 2012.

[61] Yonghong Zeng and Ying-Chang Liang. Spectrum-sensing algorithms for cognitive radio based on statistical covariances. *IEEE transactions on Vehicular Technology*, 58(4):1804–1815, 2009.

[62] Yonghong Zeng, C.L. Koh, and Ying-Chang Liang. Maximum eigenvalue detection: Theory and application. In *Communications, 2008. ICC '08. IEEE International Conference on*, pages 4160–4164, May 2008.

[63] Chang-Jun Yu, Yuan-Yuan He, and Tai-Fan Quan. Frequency spectrum prediction method based on EMD and SVR. In *Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on*, volume 3, pages 39–44, Nov 2008.

[64] A. Roy and J.F. Doherty. Weak signal sensing using empirical mode decomposition and stochastic data reordering. In *Military communication conference, 2011 - MILCOM 2011*, pages 37–41, Nov 2011.

[65] Timo Gerkmann and Richard C Hendriks. Unbiased mmse-based noise power estimation with low complexity and low tracking delay. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(4):1383–1393, 2012.

[66] Jan Sijbers, AJ Den Dekker, Johan Van Audekerke, Marleen Verhoye, and Dirk Van Dyck. Estimation of the noise in magnitude mr images. *Magnetic Resonance Imaging*, 16(1):87–90, 1998.

[67] Hüseyin Arslan and Sharath Reddy. Noise power and snr estimation for ofdm based wireless communication systems. In *Proc. of 3rd IASTED International Conference on Wireless and Optical Communications (WOC), Banff, Alberta, Canada*, 2003.

[68] Simon Haykin. Cognitive radio: brain-empowered wireless communications. *Selected Areas in Communications, IEEE Journal on*, 23(2):201–220, Feb 2005.

[69] Marc P Olivieri, Greg Barnett, Alex Lackpour, Albert Davis, and Phuong Ngo. A scalable dynamic spectrum allocation system with interference mitigation for teams of spectrally agile software defined radios. In *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pages 170–179. IEEE, 2005.

[70] Abbas Taherpour, Saeed Gazor, and Masoumeh Nasiri-Kenari. Invariant wideband spectrum sensing under unknown variances. *IEEE Transactions on Wireless Communications*, 8(5):2182–2186, 2009.

[71] Mohamed Hamid, Niclas Björsell, and Slimane Ben Slimane. Energy and eigenvalue based combined fully blind self adapted spectrum sensing algorithm. *IEEE Transactions on Vehicular Technology*, 65(2):630–642, 2016.

[72] Andrea Mariani, Andrea Giorgetti, and Marco Chiani. Effects of noise power estimation on energy detection for cognitive radio applications. *IEEE Transactions on Communications*, 59(12):3410–3420, 2011.

[73] Kevin P. Balanda and H. L. MacGillivray. Kurtosis: A critical review. *The American Statistician*, 42(2):111–119, 1988.

[74] Lawrence T DeCarlo. On the meaning and use of kurtosis. *Psychological methods*, 2(3):292, 1997.

[75] Chris Brooks. *Introductory econometrics for finance*. Cambridge university press, 2014.

[76] David Vose. *Risk analysis: a quantitative guide*. John Wiley & Sons, 2008.

[77] Suzanne Lacasse and Farrokh Nadim. Risk and reliability in geotechnical engineering. 1998.

[78] Giorgio Dall'Aglio, Samuel Kotz, and Gabriella Salinetti. *Advances in probability distributions with given marginals: beyond the copulas*, volume 67. Springer Science & Business Media, 2012.

[79] Filip Lindskog et al. *Modelling dependence with copulas and applications to risk management*. PhD thesis, Master Thesis, ETH Zürich, 2000.

[80] Gregory A Fredricks and Roger B Nelsen. On the relationship between spearman's rho and kendall's tau for pairs of continuous random variables. *Journal of statistical planning and inference*, 137(7):2143–2150, 2007.

[81] Guy P Nason. On the sum of t and gaussian random variables. *Statistics & probability letters*, 76(12):1280–1286, 2006.

[82] Saralees Nadarajah. A generalized normal distribution. *Journal of Applied Statistics*, 32(7):685–694, 2005.

[83] G.R. Arce. *Nonlinear Signal Processing: A Statistical Approach*. Wiley, 2005.

[84] R.T. Rockafellar and R.J.B. Wets. *Variational Analysis*. Springer, 2009.

[85] Nornadiah Mohd Razali and Yap Bee Wah. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1):21–33, 2011.

[86] Davtyan Azat. GMM estimation and Shapiro-Francia normality test: A case study of CEE economies. *International Journal of Economic Sciences*, 2014(1):12–26, 2014.

[87] J.J. Shynk. *Probability, Random Variables, and Random Processes: Theory and Signal Processing Applications*. Wiley, 2012.

[88] S. Banerjee and M. Agrawal. Underwater acoustic noise with generalized gaussian statistics: Effects on error performance. In *OCEANS - Bergen, 2013 MTS/IEEE*, pages 1–8, June 2013.

[89] T. Neckel and F. Rupp. *Random Differential Equations in Scientific Computing*. De Gruyter, 2013.

[90] David L Donoho and Jain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[91] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris. Spectrum assignment in cognitive radio networks: A comprehensive survey. *IEEE Communications Surveys Tutorials*, 15(3):1108–1135, Third 2013.

[92] A. Ghasemi and E.S. Sousa. Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs. *Communications Magazine, IEEE*, 46(4):32–39, April 2008.

[93] Don Torrieri. The radiometer and its practical implementation. In *Military Communications Conference, 2010-MILCOM 2010*, pages 304–310. IEEE, 2010.

[94] J. J. Lehtomaki, J. Vartiainen, M. Juntti, and H. Saarnisaari. CFAR outlier detection with forward methods. *IEEE Transactions on Signal Processing*, 55(9):4702–4706, Sept 2007.

[95] P. Stoica and R. Moses. *Spectral Analysis of Signals*. Pearson Prentice Hall, 2005.

[96] Patrick Flandrin and Paulo Goncalves. Empirical mode decomposition as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, 02(04):477–496, 2004.

[97] Zhaohua Wu and Norden E. Huang. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 01(01):1–41, 2009.

[98] Patrick Flandrin and Paulo Goncalves. Empirical mode decompositions as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(04):477–496, 2004.

[99] David R Brillinger. *Time series: data analysis and theory*, volume 36. Siam, 1981.

[100] P.P. Gandhi and S.A. Kassam. Analysis of CFAR processors in homogeneous background. *Aerospace and Electronic Systems, IEEE Transactions on*, 24(4):427–445, July 1988.

[101] P.P. Gandhi and S.A. Kassam. Optimality of the cell averaging CFAR detector. *Information Theory, IEEE Transactions on*, 40(4):1226–1228, Jul 1994.

[102] A. Sonnenschein and P.M. Fishman. Radiometric detection of spread-spectrum signals in noise of uncertain power. *Aerospace and Electronic Systems, IEEE Transactions on*, 28(3):654–660, Jul 1992.

[103] R. Tandra and A. Sahai. SNR walls for signal detection. *Selected Topics in Signal Processing, IEEE Journal of*, 2(1):4–17, Feb 2008.

[104] Carlos Cordeiro, Kiran Challapali, Dagnachew Birru, and Nov Sai Shankar. Ieee 802.22: the first worldwide wireless standard based on cognitive radios. In *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pages 328–337. IEEE, 2005.

[105] M. H. Al-Badrawi, N. J. Kirsch, and B. Z. Al-Jewad. Intrinsic mode function based noise power estimation with applications to semiblind spectrum sensing methods. *IEEE Signal Processing Letters*, 24(7):1088–1092, July 2017.

[106] M. Hamid, S. B. Slimane, W. V. Moer, and N. Bjorsell. Spectrum sensing challenges: blind sensing and sensing optimization. *IEEE Instrumentation Measurement Magazine*, 19(2):44–52, April 2016.

[107] Anant Sahai, Niels Hoven, and Rahul Tandra. Some fundamental limits on cognitive radio. In *Allerton Conference on Communication, Control, and Computing*, pages 1662–1671. Monticello, Illinois, 2004.

[108] Danijela Cabric, Artem Tkachenko, and Robert W Brodersen. Experimental study of spectrum sensing based on energy detection and network cooperation. In *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, page 12. ACM, 2006.

[109] Zhuan Ye, Gokhan Memik, and John Grosspietsch. Energy detection using estimated noise variance for spectrum sensing in cognitive radio networks. In *2008 IEEE Wireless Communications and Networking Conference*, pages 711–716. IEEE, 2008.

[110] Janne J Lehtomaki, Johanna Vartiainen, Markku Juntti, and Harri Saarnisaari. Spectrum sensing with forward methods. In *MILCOM 2006-2006 IEEE Military Communications conference*, pages 1–7. IEEE, 2006.

[111] L. Satish and B. Nazneen. Wavelet-based denoising of partial discharge signals buried in excessive noise and interference. *IEEE Transactions on Dielectrics and Electrical Insulation*, 10(2):354–367, April 2003.

[112] V.K. Rai and A.R. Mohanty. Bearing fault diagnosis using FFT of intrinsic mode functions in hilberthuang transform. *Mechanical Systems and Signal Processing*, 21(6):2607 – 2615, 2007.

[113] N. Senroy, S. Suryanarayanan, and P.F. Ribeiro. An improved hilbert-huang method for analysis of time-varying waveforms in power quality. *Power Systems, IEEE Transactions on*, 22(4):1843–1850, Nov 2007.

[114] Erhard Reschenhofer. The bimodality principle. *J Stat Educ*, 9(1):1–16, 2001.

[115] Ming Jin, Qinghua Guo, Jun Tong, Jiangtao Xi, and Youming Li. Energy detection of DVB-T signals against noise uncertainty. *IEEE Communications Letters*, 18(10):1831–1834, 2014.

[116] Jeff A Bilmes. A gentle tutorial of the EM algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. *International Computer Science Institute*, 4(510):126, 1998.

[117] David Ruppert. *Statistics and finance: an introduction*. Springer Science & Business Media, 2004.

[118] G. Rilling and P. Flandrin. One or two frequencies? the empirical mode decomposition answers. *Signal Processing, IEEE Transactions on*, 56(1):85–95, Jan 2008.