LUND
UNIVERSITY

FACULITY OF SCIENCE

CENTRE FOR MATHEMATICAL SCIENCES

MATHEMATICAL STATISTICS

TIME-FREQUENCY SPECTRAL REPRESENTATION OF AUDITORY BRAINSTEM RESPONSE (ABR) DATA

BY

AMARE TEREFE GASHAYE

A Thesis Submitted to Professor Maria Sandsten

in Partial Fulfillment of the Requirement for the Degree of Master of Science

in Mathematical Statistics

August 17, 2012

ACKNOWLEDGEMENTS

I have nothing but deepest and cordial honor to my advisor Professor Maria Sandsten, without her thoughtful and friendly approach I would have had a lame product. Being an advisee of her is not only learning research but also sharing wisdom, experience and strength. The thesis would not exist without her scientific knowledge and unselfish devotion, in particular the enormous efforts and time-spending during the preparation of this thesis.

I would like to extend my thanks to Johan Kallstrand-SensoDetect for his cooperation in explaining problems and providing necessary information including data keenly.

My sincere gratitude goes to my family for their persistent love and encouragement. I do not believe words can express my feelings to my aunt Mulu Gulelat who has been my devotee since my earliest child hood. She is the ever special person I saw in my life. It is a pride for me to express my appreciation and love to my sweetheart Zenebu Hailu for her support, generous care and love. You were always beside me during the happy and hard moments to push me and motivate me. I can see the good shape of my study because of your help and encouragements. A journey is easier when you travel together, your personal support and great patience at all times are always my strength.

Especial thanks goes to my honest friends Abebaw Alemayeh, Selam Geremew and Peter Ohlsson who have been beside me through all the ups and downs of my journey. My heartfelt thanks goes to Aurelia Vogel, our department's x-secretary who has been keen and helpful all the times that she were in the university.

Above all I want to express my praise to the Most Gracious and the Most Merciful Heavenly Father, God, for his showers of blessings and powers throughout my life to overcome all the complications I faced.

ABSTRACT

The time-frequency (TF) spectral representation of Auditory Brainstem Response (ABR) signal data provides information about their spectral contents. We apply the Spectrogram, Thomson Multitaper and Peak Matched Multiple Window (PM MW) spectral estimation methods to four different number of clicks per average (i.e., 1313, 300, 100 and 50 number of clicks per average) of a simulated signal data. For the purpose of model selection we simulate sinusoidal signal data which have the same trend as the empirical ABR signal data, and then apply the selected model to ABR data from 17 healthy, normal hearing individual ears as recorded using SD-BERA, SensoDetect-Brainstem Evoked Response Audiometry. The root mean square error (RMSE) is the main tool used to compare the proposed spectral estimation methods. The Spectrogram is found to be an appropriate method of spectral estimation for signals with relatively low disturbance. In particular, for signals with a white disturbance with standard deviation, σ, value in the interval $[0,15.0]$, it is found to be best of the three methods. For $15.0 \leq \sigma \leq 30.0$, the PM MW method performs as good as the spectrogram, if not better. Finally, for $\sigma \geq 30.0$ the PM MW continues to be the best of the three methods where as the Spectrogram turns out to be worst of them.

Key Words: TF, Spectrogram, Thomson Multitaper, PM MW, ABR, SensoDetect

Contents

1. INTRODUCTION

1.1 Background of the Study

The field of digital signal processing (DSP) has been in explosive growth during the past four decades, as phenomenal advances both in research and application have been made . This is due to the fact that signals play an important role in our daily life activities. Speech, music, picture and video are a few examples of signals that we encounter frequently [6].

The nature of signal processes is diverse such that often characterizations and classifications of them are challenges. Most of the time, scholars differ among them depending on the nature of the independent variables and the value of the functions defining the signals. Mainly, signals could be classified as deterministic and stochastic in which deterministic signals are explicitly known and stochastic signals are random processes which are represented by a realization and characterized by distributions. Signals could also be classified as stationary and non-stationary where a stationary signal has time-invariant properties and a non-stationary signal has time-variant properties. A signal process can be either a continuous or a discrete function of the independent variable. Moreover, a signal can be either a real-valued or a complex-valued function; they can be also either analytical or non-analytical [6].

The fundamental variable we use to study a signal is time. However, we can also consider other variables. One of the most important other variable that we should consider is frequency. In general, time-frequency analysis of non-stationary time-varying analytic signals have been a focus of researchers for many years [1]. Our study bestows a due emphasis for a such type of signal process, i.e., the human auditory brainstem response (ABR) signals.

The ABR consists of far-field evoked potentials from the auditory nervous system that occur during 10 milliseconds (ms) after the presentation of a transient sound such as a click sound [21]. According to Christopher Emrick, 2008, ABR data are the most frequently used auditory evoked potentials. They consist of a series of $5-7$ peaks in the time-averaged waveform to click stimuli observed in an approximately 10 ms post-stimulus recording. This response reflects auditory activity from the cochlea to the midbrain. By convention, the waves are designated by Roman numerals. Clinically, the most significant ABR peaks are designated by wave I, III, and V. In reality, waves III and V are often combinations of wave II-III and waves IV-V respectively [20]. Typical human ABR recording is shown in **Figure 2.1**.

A recent increase of interest in spectral analysis has been prompted by the suggestion that spectral components of the ABR may contribute in psychiatric diagnostic investigation [22]. Hence, knowledge of the effects of sound stimuli on a spectral representation of the ABR is essential.

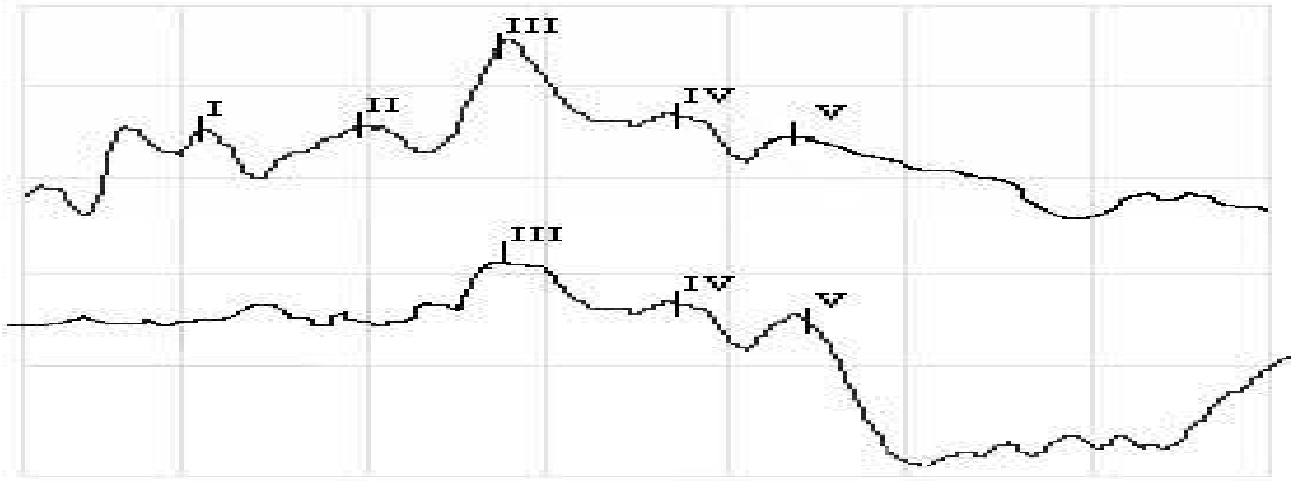


Figure 2.1 *ABR waveform recorded in a normal-hearing adult.*

The aim of spectral analysis is to extract relevant information from a signal process by transforming it. Among the classical methods of spectral estimation, Fourier based method are well known. These approaches are computationally efficient. In spite of this advantage, they have many inherent drawbacks in case of short signals. The most prominent limitation is poor frequency resolution, i.e., low quality of the ability to distinguish the spectral responses of two or more signals. Other limitation is windowing leakage in the spectral domain where energy in the main lobe of a spectral response leaks into the side-lobes, obscuring and distorting other spectral responses that are present [2], [3]. In attempting to overcome these limitations many alternative spectral estimation methods have been developed. Hence, the main purpose of this project is to discuss two different approaches: Spectrogram and multitaper spectrograms (the Peak Matched Multiple Windows and Thomson methods) as estimators of a spectrum for ABR data.

1.2 Statement of the Problem

The brainstem activity can be explained in two-dimensional curves: frequency versus time. The main reason to use such types of plots is to find frequencies in different parts of the curve for different number of clicks per average. This study tries to use some registered electrophysiological reaction signals (or ABR signal data) that originate in the brainstem when click stimuli are presented. Elucidating the ABR data, from different people, by means of some non-parametric time-frequency signal process analysis techniques, which perhaps is a new approach for SensoDetect, will be the sole purpose of our work. We will focus on fitting, comparison and selection of Spectrogram, Thomson Multitaper and Peak Matched Multiple Window (PM MW) methods.

1.3 Objectives

1.3.1 General Objective

The general objective of this study is to give a better time-frequency representation of the ABR data using one or more of the spectrogram, Thomson Multitaper or PM MW techniques.

1.3.2 Specific Objectives

The specific objectives are:

- to simulate a sinusoidal signal data that looks like the actual data;
- to determine a reasonable interval for a phase shift in the simulated function;
- to fit the simulated data for the three target models at four different number of clicks per averages;
- to discuss an appropriate modeling approach for some non-parametric time frequency models;
- to discuss some properties of the spectrogram and multitaper methods;
- to make some evaluation measures for the fitted models;
- to select the best model for the simulated data;
- to make suggestion on the level of disturbance that should be set in the simulated data to have feasible results;
- to express the actual ABR data in time-frequency representation using the once chosen model and investigate the time-frequency composition;
- creating a model that can be tested and used by SensoDetect and other bodies.

1.4 Significance of the Study

The results of the research work could be helpful in:

- reviewing model construction processes in some non-parametric spectral estimation mechanisms like Spectrogram and Multitaper techniques;
- understanding the importance of spectral analysis in explaining ABR data;
- investigating the ABR signal data represented in a two-dimensional spectrum;
- managing frequency content analysis in connection with timing properties in a healthy, and normal-hearing recorded ABR data;
- clinical psychiatric diagnostic investigation;
- helping in classification of individuals based on their responses to different sound stimuli;
- providing vital information in relation with improvement of SD-BERA;
- extending a new approach signal analysis for SensoDetect;
- helping as a spring board for further studies.

1.5 Thesis Outline

The thesis is organized in sequential chapters with different focuses. In chapter two, we give some literature reviews and state the previous works in time-frequency analysis of Auditory Brainstem Response (ABR) data and related methods. Special attention is given for the Spectrogram, Thomson Multitaper and Peak Matched Multiple Window methods.

Chapter three provides the essential methodological tools that would be handy for the data analysis. In this chapter we supply the necessary technical schemes for time-frequency analysis targeting towards our goal.

This chapter explains the need for time-frequency analysis. We then discuss the development of Spectrogram method from the short time Fourier transform followed with the formulation of Thomson Multitaper and later introduce the key concepts of Peak Matched Multiple Window methods.

In chapter four, we describe the implementation of our methodology. First, we discuss the nature of our observed and simulated data , followed by the spectral estimation of the simulated data using the three methods. Second, the strategy for model selection is given. Third, an attempt for application of the selected model on our observed data is presented.

We give a conclusion and summarize our analysis results together with some recommendations in chapter five.

2. LITRATURE REVIEW

In an effort to extract more diagnostic information from the ABR, some researchers are beginning to look beyond the traditional measures of peak latency and amplitude and are turning their attention to the spectral content of the ABR [22].

Spectral analysis separates the ABR waveform into its time and frequency components allowing the investigator to analyze the ABR in the time-frequency domain rather than in the time domain only. Precise description and quantification of the normal ABR spectrum has become important as a result of a recent interest in the use of ABR spectral analysis tool. A report by Hall (1986) suggested that ABR spectral analysis may be useful in differential diagnosis of auditory pathology [22], [23].

Mainly, the spectral analysis of the ABR is based on the investigation of frequency components. previous studies like Urbach and Pratt (1986) and Malinoff and Spivak (1990) which have examined the spectral content of the normal ABR revealed that there are three major energy peaks [22], [24]. Other studies such as Laukli and Mair (1981) indicate that the greatest amount of spectral energy is concentrated in the low-frequency region [25].

Under our study, we are trying to apply some spectral estimators of ABR signal data. The simplest spectral estimate of the ABR signals is the Periodogram. This spectral estimate suffers from two problems: bias and variance. Bias arises because signals at different frequencies are mixed together and “blurred” unless the data length is infinite. Even if the data length were infinite, the periodogram spectral estimate would simply square the data without averaging. As a result, it would never converge to the correct value and would remain inconsistent which is the consequence of variance. Recordings of neural signals are often sufficiently limited so that bias and variance can present major limitations in the analysis [26].

One of the universally applicable solution to solve these two problems is the use of multitapers. Pesaran (2008) concludes neural signals, as many other signals, face the problem of bias and variance in their periodograms. Finally he suggested applying the multitaper techniques and spectral averages will resolve the problems to a considerable level [26].

Our main target in this study is to see the performance of a Spectrogram and two multitaper spectral representations of the ABR. Several researchers and scholars have been giving some reviews on the performance of multi-taper and single-taper spectral estimation procedures.

In the choice of an appropriate estimator for a spectrum of a stationary time series from a finite sample of the process, the two chronic problems are bias control and consistency, or "smoothing". In attempting to overcome these problems, D.J. Thomson (1982) introduced a method called Thomson Multitaper. This method is the same as using weighted average of a series of direct spectrum estimates (Periodograms) based

on orthogonal data windows. In this paper he successfully showed that it is possible to control the bias and variance and smooth the spectrum to have a better resolution [10].

In a research conducted in 1987, J. Park stated that spectral estimation procedures which employ several tapers have been shown to yield better results than standard single-taper spectral analysis when used on a variety of signal data. In particular, he confirmed that multitaper spectral analysis techniques offer a seismologist formal and practical advantages over single-taper techniques [19].

The main reason scholars use multiple window methods is that they decrease the variance of the spectrum estimate by smoothing it. Particularly, if we consider a peaky spectrum, the variance and bias are relatively high. To solve this problem, many researchers have been trying different approaches. For example, M. Hansson and G. Salomonsson (1997), showed that a peak matched multiple window method for peaked spectra is suitable. The resulting spectrum estimate has low variance and bias in the neighborhood of the peak. The proposed method shows, however, a large bias at a notch. This is due to large side-lobes of the windows that cause leakage from frequencies outside the resolution width. This leakage is suppressed with the use of a penalty function. Finally, they concluded that this method is suitable not only for peaked spectra but also for spectra with notches [17].

3. METHODOLOGY

3.1 Signal Representation as a General Case

Time frequency signal analysis and processing concerns the analysis and processing of signals with time-varying frequency content. Such signals are best represented by time-frequency distribution, which is intended to show how the energy of the signal is distributed over the two-dimensional time-frequency space. Processing the signal may exploit the features produced by the concentration of signal energy in two dimensions (time and frequency) instead of only one (time or frequency) [1].

The two classical representations of a non-stationary signal process $\{x\}$ are the time-domain representation $x(t)$ and the frequency-domain representation $X(f)$. In both forms, the variables t and f are treated as mutually exclusive: to obtain a representation in terms of one variable, the other variable is "integrated out". Consequently, each classical representation of the signal is non-localized with respect to the excluded variable. But, in time-frequency distribution $S(t,f)$ the variables t and f are not mutually exclusive rather they are present together. Hence, the time-frequency distribution is localized both in t and f [1].

Naturally, any signal can be expressed as a function of time t. This representations results in instantaneous power given by $|x(t)|^2$, which shows how the energy of the signal is distributed over time. The total signal energy is:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad \cdots\cdots (3.1)$$

The time-domain representation tends to obscure information about frequency, because it assumes that the two variables t and f are mutually exclusive.

Similarly, any signal $x(t)$ can be represented in frequency domain by its Fourier transform $X(f)$, given by

$$X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t) e^{-i2\pi f t} dt \quad \cdots\cdots (3.2)$$

where the symbol $\mathcal{F}$ stands for the term 'Fourier transform'. The Fourier transform (FT) is in general complex; its magnitude, $|X(f)|$, is called the magnitude spectrum and its phase is called the phase spectrum. The square of the magnitude spectrum,$|X(f)|^2$ is the energy spectrum and shows how the energy of the signal is distributed over the frequency domain; the total energy of the signal will be

$$E = \int_{-\infty}^{\infty} |X(f)|^2 df = \int_{-\infty}^{\infty} X(f) X^*(f) df \quad \cdots\cdots (3.3)$$

where $X^*(f)$ is the complex conjugate of $X(f)$. The time-domain representation can be recovered from the FT by taking the inverse Fourier transformation (IFT):

$$x(t) = \mathcal{F}^{-1}\{X(f)\} = \int_{-\infty}^{\infty} X(f)e^{i2\pi ft}df \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.4)$$

where the symbol $\mathcal{F}^{-1}$ stands for IFT. However, the frequency-domain representation given by (3.2) "hides" the information about timing as $X(f)$ integrates out time t.

Therefore, the need for a time-frequency distribution as a representation of time-varying signal processes is a must in the real world observable facts. For example, our sense of hearing readily interprets sounds in terms of variations of frequency with time because representations of a sound signal in the time domain or frequency domain only do not facilitate an appealing interpretation. Hereafter, we will discuss some time-frequency domain representations of signals. But, first it is advisable to make clear what type of signals we are going to deal with.

3.2 Analytic Signals

A non-stationary signal process $\{x(t): t\epsilon\mathbb{R}\}$, is real-valued if and only if

$$X(-f) = X(f) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.5)$$

where $X(f)$ is the Fourier transform of $x(t)$. This indicates that a real signal is one that shows Hermitian Symmetry between the negative- and positive-frequency components, allowing the former to be deducted from the latter. Hence, the negative-frequency components of a real signal may be eliminated from the signal representation without losing information. In the case of a real low pass signal, removal of negative frequencies has a twofold advantage:

- it halves the total bandwidth, allowing the signal to be sampled at half of the usual Nyquist rate without aliasing [3].
- it avoids the appearance of interference terms resulted by the interaction of positive and negative components in quadratic time-frequency distributions like Wigner-Ville distribution.

A signal $z(t)$ is said to be analytic if and only if

$$Z(f) = 0 \; for \; f < 0, \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.6)$$

In other words, an analytic signal has no negative frequencies; it may have a spectral component at zero frequency.

The signal

$$z(t) = x(t) + iy(t) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.7)$$

where $x(t)$and $y(t)$ are real, is analytic with a real spectral component at zero frequency if and only if

$$Y(f) = (-sgnf)X(f) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.8)$$

where $X(f)$ and $Y(f)$ are the Fourier transforms of $x(t)$and $y(t)$, respectively, and

$$sgnf = \begin{cases} -1 & if\ f < 0 \\ 0 & if\ f = 0 \\ 1 & if\ f > 0 \end{cases} \cdots \quad (3.9)$$

If the Fourier transforms of $x(t)$and $y(t)$ are related according to (3.8), we say $y(t)$ is the Hilbert transform of $x(t)$ and denoted as

$$y(t) = \mathcal{H}\big(x(t)\big) \cdots \quad (3.10)$$

Hence, if one is given a real signal $x(t)$, he/she is able to construct a complex signal as

$$z(t) = x(t) + i\mathcal{H}\big(x(t)\big) \cdots \quad (3.11)$$

such that $z(t)$ is analytic, where the Hilbert transform is evaluated from

$$y(t) = \mathcal{H}\big(x(t)\big) = \mathcal{F}^{-1}\big\{(-isgnf)\mathcal{F}\{x(t)\}\big\} \cdots \quad (3.12)$$

3.3 Spectrogram and Wigner-Ville Distribution

3.3.1 Spectrogram

Consider a signal $x(\tau)$ and a real, even window $h(\tau)$, whose FTs are $X(f)$ and $H(f)$, respectively. To obtain a localized spectrum of $x(\tau)$ at time $\tau = t$, multiply the signal by the window $h(\tau)$ centered at time $\tau = t$ to obtain

$$x(t,\tau) = x(\tau)h(\tau - t) \cdots \quad (3.13)$$

and then take the FT with respect to τ, resulting the short-time Fourier transform (STFT) of $x(\tau)$ given as,

$$X(t,f) = \mathcal{F}\{x(t,\tau)\} = \int_{-\infty}^{\infty} x(t,\tau)e^{-i2\pi f\tau}d\tau = \int_{-\infty}^{\infty} x(\tau)h(\tau - t)e^{-i2\pi f\tau}d\tau \cdots \quad (3.14)$$

The squared magnitude of the STFT, given by

$$S_x(t,f) = |X(t,f)|^2 \cdots \quad (3.15)$$

is called the spectrogram, which is used very frequently for analyzing time-varying and non-stationary signals.

Fast implementation using fast Fourier transform (FFT) and easy for interpretation are among the appealing advantages of a spectrogram. However, it has also limitations; theoretically, the time slots should be very short to give us clear information about when the frequencies start and stop. But, a short time slot gives a bad frequency resolution, which causes the problem that we cannot differ closely spaced frequencies. The length of the window function $h(\tau)$ is very important as it determines the resolution in time and frequency: a short window gives better resolution in time than in frequency and vice versa.

Hence, finding an optimal window length of a spectrogram is another task that we should take into account. In addition, to overcome the limitation, we can try other advanced spectral estimation levels like Wigner-Ville distribution, reassigned spectrogram and multitaper reassigned spectrogram. However, either finding an optimal window length or trying other advanced levels is out of the purpose of this study.

Since we need the Wigner-Ville distribution without noise and phase shift in the case of spectral estimates' comparison, it is indispensable to discuss about its theoretical formulation. Note that we are not going to use it as a spectral estimate instead it will serve us as an actual spectrum of a simulated sample signal data. This is the reason why we do not impose a disturbance and phase change when we compute it from the simulated data.

3.3.2 Wigner-Ville Distribution

For a non-stationary analytic signal process $\{x(t): t \in \mathbb{R}\}$, the Wigner-Ville spectrum (WVS) is always mentioned first among the various possibilities of defining a time-varying spectrum. The time-varying spectral density is defined as

$$S_x(t,f) = \int_{-\infty}^{\infty} E\left\{x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\right\} e^{-i2\pi f\tau} d\tau \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.16)$$

where

$$E\left\{x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\right\} = r_x(t,\tau) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.17)$$

is a symmetrical instantaneous autocorrelation function ,IAF, under the two variables t and τ and $E\{.\}$ stands for the expectation operator. This formulation is called the Wigner-Ville spectrum (WVS). For a nonrandom or deterministic signal, we have

$$E\left\{x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\right\} = x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2}) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.18)$$

Substituting (3.18) into (3.16) results the Wigner-Ville distribution (WVD) defined as

$$W_x(t,f) = \int_{-\infty}^{\infty} x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\, e^{-i2\pi f\tau} d\tau \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.19)$$

Under normal conditions, from (3.16) and (3.19), we can conclude that the WVS is the ensemble average of the WVDs,

$$S_x(t,f) = E\{W_x(t,f)\} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.20)$$

The WVD has a better time-frequency (TF) resolution compared to a spectrogram which suffers from a poor TF resolution. Especially, it is optimal to analyze signals constituted by a single component. By definition, the WVD does present the advantage of depending only on the signal. Its role is central in TF analysis and it possesses a number of theoretical properties among which one can mention its ability to be perfectly localized in the case of linear frequency-modulated signals.

Despite of these advantages, it is not well-suited for application in multi-component signals since, in such cases; it will be followed by the presence of interference- or cross-terms. Hence, for multi-component

signals, it has weak time and frequency supports, i.e., if $x(t) = 0$ when $t < t_1$ and $t > t_2$ then $W_x(t,f) = 0$ for $t < t_1$ and $t > t_2$, and similarly, if $X(f) = 0$ when $f < f_1$ and $f > f_2$ then $W_x(t,f) = 0$ for $f < f_1$ and $f > f_2$. The cross terms are the most troublesome in the representation of time-frequency analysis such that outcomes of an analysis could be easily misinterpreted [2], [4].

3.3.3 WVD as Actual Spectrum

One serious disadvantage of the WVD is cross-term interference. Cross-terms are artifacts that appear in the WVD representation between auto-terms, which correspond to physically existing signal components. These cross-terms falsely indicate the existence of signal components between auto-terms. They are strongest at the midpoints between the auto-terms. In general, as the number of auto-terms increases, the auto-terms and the cross-terms overlap. Consequently, distinguishing the auto-terms from cross-terms can be challenging.

The time-frequency plane includes positive frequencies and negative frequencies. Signal components present at positive frequencies in real-valued signals symmetric components at negative frequencies. The cross-terms appear between auto-terms at positive frequencies, between auto-terms at negative frequencies, and between auto-terms at positive and negative frequencies.

If we convert a real-valued signal into a complex-valued analytic signal by removing the auto-terms at negative frequencies before we apply the WVD, we can reduce the number of cross-terms in the WVD. The analytic frequency signal has the same spectral content at positive frequencies as the original real-valued signal but has no spectral content at negative frequencies. By converting the real-valued signal to an analytic signal, one removes the cross-terms between auto-terms at negative frequencies and the cross-terms between auto-terms at positive frequencies and negative frequencies. However, it is hard to eliminate the cross-terms between auto-terms at positive frequencies.

Suppose we have a non-stationary signal process $\{z(t): t \in \mathbb{R}\}$ with two components given as

$$z(t) = x(t) + y(t) + e \quad \cdots\cdots (3.21)$$

where $x(t)$ and $y(t)$ are other signals which may be the same but differ in frequencies and phase change only, and e is an identically and independently distribute white noise with mean 0 and σ^2.

For illustration, suppose we ignore the white disturbance and assume the analytic signal is only the sum of the two main signal components. The Wigner-Ville distribution will be given as

$$W_z(t,f) = W_x(t,f) + W_y(t,f) + 2\Re\left[W_{x,y}(t,f)\right] \quad \cdots\cdots (3.22)$$

where $W_x(t,f)$ and $W_y(t,f)$, called the auto-terms, are the Wigner distributions of $x(t)$ and $y(t)$, respectively; and $2\Re\left[W_{x,y}(t,f)\right]$ is called the cross-term, which always presents and is located in the midway between the auto-terms oscillating proportionally to the distance between the auto-terms with a

direction of oscillation orthogonal to the line connecting the auto-terms. Above this, we can imagine how the problem will be aggravated by the addition of the white disturbance term. Generally, the overall readability of the WVD is hampered by the cross-terms oscillatory interference phenomena.

However, considering a theoretical case (for simulated data), we can create a WVD which is clearly readable. Suppose we assume that $x(t)$ and $y(t)$ are from the same function but differ by frequencies only, that is, there is no phase shift and white noise imposed on either of them. It is possible to compute the individual WVD of $x(t)$ and $y(t)$ separately and take the sum as :

$$WVD_{xy} = WVD_x + WVD_y \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{(3.23)}$$

where WVD_x and WVD_y, the WVD of $x(t)$ and $y(t)$ computed indipendently, when the signals are free of phase changes and disturbances, respectively. WVD_{xy}, is considered as an actual (observed) spectrum which helps in spectral estimate validation. This is to mean that we can compare a spectral estimate with WVD_{xy} to check whether it is a good fit or not.

3.4 Thomson and Peak Matched Multitapers

3.4.1 Thomson multitapers

The thought of multiple windows or multitapers came to existence for the first time by David Thomson in 1982 [8],[10]; in fact, the idea of Multitapers had been exploited earlier in the form of one window shifted in time in a method called Welch or WOSA by Welch [11]. The central idea of Multitapers is to diminish the variance of a periodogram by averaging several uncorrelated periodograms. The time-shifted window by Welch provide uncorrelated periodograms since the window makes non-overlap parts of different data sequences for the same window. But, Thomson's approach was to use the same data sequence for all periodograms, that is, utilizing the whole data sequence, except changing the shape of the window for different periodograms in a way that gives uncorrelated periodograms and thereby reducing variance [12].
In Thomson's Multitapers approach, if we are given a power spectrum $S_x(f)$, of a real-valued stationary discrete-time random process $x(n)$, we estimate the spectrum from N samples $\boldsymbol{x} = [x(0) \cdots x(N-1)]^T$ of the process by using the estimator,

$$\hat{S}(f) = \sum_{k=1}^{K} \alpha_k \hat{\phi}_k(f) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{(3.24)}$$

where

$$\hat{\phi}_k(f) = \left| \sum_{n=0}^{N-1} x(n) \mathrm{h_k(n)} \mathrm{e}^{-\mathrm{i2\pi fn}} \right|^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{(3.25)}$$

and $\alpha_k = \frac{1}{K}$. Equation (3.22) is a windowed periodogram obtained by using the Slepian sequences $\boldsymbol{h_k} = [h_k(0) \cdots h_k(N-1)]^T$ as data window [13]. Hence, (3.24) provides a weighted sum of K Periodograms for

the Thomson multitaper estimate. The value of K in the Thomson estimator, reduces the variance in the estimate by a factor of K if the averaged periodograms are uncorrelated. This implies that K should be as large as possible. However large values of K increase bias [14].

To fully understand the limitations of the Thomson multitaper method, it is advisable to evaluate the variance of the estimated spectrum. The variance of the is given by all combinations of the different periodogram covariances,

$$Variance\left(\hat{S}(f)\right) = \sum_{k=1}^{K}\sum_{p=1}^{K} \alpha_k \alpha_p cov(\hat{\phi}_k(f), \hat{\phi}_p(f)) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.26)$$

Assuming $\boldsymbol{x}$ to be Gaussian and making some simplifications give,

$$cov\left(\hat{\phi}_k(f), \hat{\phi}_p(f)\right) = |\boldsymbol{h}_k^T A^H(f) \boldsymbol{R_x} A(f) \boldsymbol{h_l}|^2 + \left|\boldsymbol{h}_k^T A(f) \boldsymbol{R_x} A(f) \boldsymbol{h_p}\right|^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.27)$$

where $A(f) = diag[1 \;\; e^{-i2\pi f} \;\cdots\; e^{-i2\pi(N-1)f}]$ and $\boldsymbol{R_x}$ is the covariance matrix of $\boldsymbol{x}$.

According Walden [14], [15] the second term of (3.27) is large only for frequencies near to 0 and 0.5, hence one can omit it and have

$$cov\left(\hat{\phi}_k(f), \hat{\phi}_p(f)\right) \approx \left|\boldsymbol{h}_k^T A^H(f) \boldsymbol{R_x} A(f) \boldsymbol{h_p}\right|^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.28)$$

If we choose an evaluation frequency f_0 at the center of the base band filter, the variance will be given as,

$$Variance\left(\hat{S}(f)\right) \approx \sum_{k=1}^{K}\sum_{p=1}^{K} \alpha_k \alpha_p \left|\boldsymbol{h}_k^T \boldsymbol{R_x}^{\boldsymbol{0}} \boldsymbol{h_p}\right|^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.29)$$

where $\boldsymbol{R_x}^{\boldsymbol{0}}$ represents the covariance function of an identical spectrum modulated so that the evaluation frequency is 0 and modulating the spectra according to

$$\boldsymbol{R_x} = A(f_0) \boldsymbol{R_x}^{\boldsymbol{0}} A^H(f_0) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.30)$$

The variance will be small if we get rid of every cross-covariance term, that is, if $\boldsymbol{h}_k^T \boldsymbol{R_x}^{\boldsymbol{0}} \boldsymbol{h_p} = 0$, for all $k \neq p$. This is more or less true for multitapers that are the eigenvectors of the covariance matrix $\boldsymbol{R_x}^{\boldsymbol{0}}$, by choosing a covariance function with properties similar to the spectrum shape one can reduce the cross-covariance terms. In other words, this property is true for processes of similar shape as the baseband filter, that is, band-limited white noise and white noise for the Slepian windows. In such a case, the variance will be reduced by a factor K compared to the periodogram [14].

However, the Thomson Multitapers are not appropriate for spectra with peaks and notches. This is because, in such a case, the cross-covariance is no longer zero, and hence then the resulting variance increases.

3.4.2 Peak Matched Multitapers

For spectra with peaks and notches other multitaper methods such as the sinusoidal multitapers [16] have better properties. The sinusoidal Multitapers have the analytic expression given by

$$h_k(n) = \sqrt{\frac{2}{N+1}} \sin\left(\frac{\pi k(n+1)}{N+1}\right), n = 0, \cdots, N-1 \; and \; k = 1, \cdots, K \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.31)$$

The main reason that these windows have low cross-covariance is that they are orthogonal to the frequency plane.

In a similar manner, using a covariance function corresponding to a template spectrum with a peaked shape results in windows that are perfectly orthogonal at the peaks of such spectra. This is the concept of the Peaked Matched Multiple Windows (PM MW) [17]. Unfortunately, the property of frequency localization is not fulfilled.

In the case of PM MW, we are expected to design windows $h_k, k = 1, \cdots, K$ giving a small bias as well as a low variance estimate of $S_x(f)$ in the neighborhood of the peak frequency. Low bias is obtained by matching the windows to the peak of $S_x(f)$, where as reduction of the variance is established with uncorrelated Periodograms $\hat{\phi}_k(f)$ at the peak. Thus, a frequency local estimate is desired. To prevent leakage from regions outside a predetermined interval of width β, the Fourier transforms $H_k(f)$ of $h_k, k = 1, \cdots, K$ have to be band-limited to the interval $(-\frac{\beta}{2}, \frac{\beta}{2})$. The main-lobes of the windows should be inside this band, and the side-lobes of each window should be as low as possible [17].

The multiple window estimation method can be considered to be a filtering procedure in a filter bank. The impulse responses of the sub-filters are $h_k, k = 1, \cdots, K$. Given the input signal $x(n)$, the power of the output signal within the frequency interval $(-\frac{\beta}{2}, \frac{\beta}{2})$ is

$$P_\beta = \sum_{k=1}^{K} \int_{-\frac{\beta}{2}}^{\frac{\beta}{2}} |H_k(f)|^2 S_x(f) df = \sum_{k=1}^{K} \alpha_k \boldsymbol{h}_k^T \boldsymbol{R}_\beta \boldsymbol{h}_k \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.32)$$

The $(N \times N)$ Toeplitz covariance matrix $\boldsymbol{R}_\beta$ has the elements

$$r_\beta(l) = r_x(l) * \beta sinc(\beta l), 0 \leq |l| \leq N-1, \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.33)$$

where $r_x(l)$ is the covariance function of $x(n)$, $sinc(x) = \frac{sin(\pi x)}{\pi x}$, and $*$ denotes the convolution operator. In (3.29), P_β is the power of $x(n)$ within the main-lobe of the windows. Now, we have to find the K window functions h_k that maximize P_β. The optimization is performed subject to the constraint

$$P_Z = \sum_{k=1}^{K} \alpha_k \int_{-\frac{1}{2}}^{\frac{1}{2}} |H_k(f)|^2 S_z(f) df = \sum_{k=1}^{K} \alpha_k \boldsymbol{h}_k^T \boldsymbol{R}_Z \boldsymbol{h}_k = 1 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.34)$$

where $S_z(f)$ with the corresponding Toeplitz covariance matrix $\boldsymbol{R}_Z$ is chosen for suppression of the side-lobes of the windows. The solution with respect to $\boldsymbol{h}_k$ is the set of eigenvectors of the generalized eigen-value problem

$$\boldsymbol{R_B q}_k = \boldsymbol{\lambda}_k \boldsymbol{R}_Z \boldsymbol{q}_k, k = 1, \cdots, N \quad \cdots\cdots (3.35)$$

where $\lambda_1 \geq \lambda_2 \cdots \lambda_N$. The eigenvectors corresponding to the K largest eigen-values are used as windows $\boldsymbol{h}_k = \boldsymbol{q_k},\ k = 1 \cdots K$; and are known as peak matched multiple windows (PM MW's). The windows are orthogonal to the covariance matrix $\boldsymbol{R_B}$, implying uncorrelated Periodograms $\hat{\phi}_k(f)$ at the peak frequency. The covariance matrix $\boldsymbol{R_Z}$ is chosen in order to grant certain properties to the estimate $\hat{\phi}(f)$. Usually we take $\boldsymbol{R_Z} = \boldsymbol{R_G}$, where $\boldsymbol{R_G}$ corresponds to a penalty frequency function. Any window for spectral estimation should have low-level side lobes otherwise leakage from high power frequencies could hide important information in the spectrum. One possible solution to this problem is including a penalty frequency function,

$$\phi_G(f) = \begin{cases} G & |f| > {}^{\beta}/_{2} \\ 1 & |f| \leq {}^{\beta}/_{2} \end{cases} \quad \cdots\cdots (3.36)$$

is used to decrease the leakage from the side-lobes. G is such that $\boldsymbol{R_G}$ is the corresponding Toeplitz covariance matrix, and β is the corresponding bandwidth. The ideal window functions satisfy,

$$\sum_{k=1}^{K} \alpha_k |H_k(f)|^2 = \phi_G^{-1}(f), -\frac{1}{2} \leq f \leq \frac{1}{2} \quad \cdots\cdots (3.37)$$

If the power P is found to be large in value, the side lobes of $|H_k(f)|^2$ outside $|f| > \frac{\beta}{2}$ will be suppressed by this factor. The weighting factor α_k is a parameter that can be chosen arbitrarily. We study a matched spectrum approach $\sum_{k=1}^{K} \alpha_k |H_k(f)|^2 = S_x(f)$ in the local interval $-\frac{\beta}{2} \leq f \leq \frac{\beta}{2}$. The total filter function should have the same appearance as the peak to minimize bias as well as give a low variance of the power spectrum estimate in the neighborhood of the peak. The PM MW approached is fulfilled with

$$\alpha_k = \frac{\lambda_k}{\sum_{k=1}^{K} \lambda_k} \quad \cdots\cdots (3.38)$$

3.5 Measuring Model Fit

3.5.1 Nonparametric Regression

The statistical problem of nonparametric regression is to estimate a function $y = f(x)$ from data pairs (x_j, y_j), for $j = 1, \ldots, n$, with the response values y_j assumed to be observations from the "signal plus noise" model here:

$$Y_j = f(x_j) + \varepsilon_j \quad \cdots\cdots (3.39)$$

It is usually solved by some variation on least squares, which may be motivated by an assumption that the ε_j variables are independent and Normally distributed. Generalized nonparametric regression replaces the Normal distribution with some other distribution, and modifies the model accordingly.

3.5.2 (Root) Mean Squared Error

A method of (generalized) nonparametric regression produces at each x an estimate $\hat{f}(x)$. A good method is one that is likely to make the difference between $\hat{f}(x)$ and $f(x)$ small for every relevant value of x. The simplest and most widely used way to measure likely difference is mean squared error (MSE) [26]. :

$$MSE(x) = E\left(\left(\hat{f}(x) - f(x)\right)^2\right) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.40)$$

Here E stands for expectation, meaning the theoretical average (mean) over hypothetical repetitions of the data-collection process. An extremely important general relation is as follows [26].:

$$MSE = Variance + Bias^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.41)$$

The root mean-squared error (RMSE),which is the square root of the MSE, is also a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed. Empirically, to compute a RMSE, first the difference between forecast and corresponding observed values are each squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means, it is most useful when large errors are particularly undesirable. The RMSE can range from 0 to ∞. But, lower values are indications of better model fit.

Generally, the RMSE tells us about the fit of the model to the data set used [18]. The sole purpose of our study is to find a good spectral estimator of our signal data from the three proposed estimator. In our case, the WVD without noise and phase-shift is considered as actually observed spectrum. If S is the spectral estimate of WVD, then WVD is expected to be equal to some constant a times S plus an error:

$$\boldsymbol{WVD} = a\boldsymbol{S} + \boldsymbol{e} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.42)$$

where a is a constant (slope) and $\boldsymbol{e}$ is an error. The **WVD** and **S** are matrices of the same size of actual and estimated values, respectively. Hence, the same size is true for the error $\boldsymbol{e}$.

Both low variance and small bias can be evaluated with the use of the mean squared error (MSE), as this measure includes both variance and squared bias. theoretically, the MSE of $\boldsymbol{e} = \boldsymbol{WVD} - a\boldsymbol{S}$, is defined as

$$MSE(\boldsymbol{e}) = E[\boldsymbol{e}]^2 = Variance(\boldsymbol{e}) + [Bias(\boldsymbol{e})]^2. \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.43)$$

Taking into account the fact that the mathematical operations are conducted under matrix rules, the RMSE is defined as

$$RMSE = \sqrt{MSE(\boldsymbol{e})} = \sqrt{\frac{(\boldsymbol{e}.*\,\boldsymbol{e})}{N}} \cdots \cdots \cdots (3.44)$$

where N is the sample size. Here, we should understand the mathematical operation $\boldsymbol{e}.*\,\boldsymbol{e}$ stands for element-wise multiplication of the error matrix.

4. DATA AND DATA ANALYSIS

4.1 Data

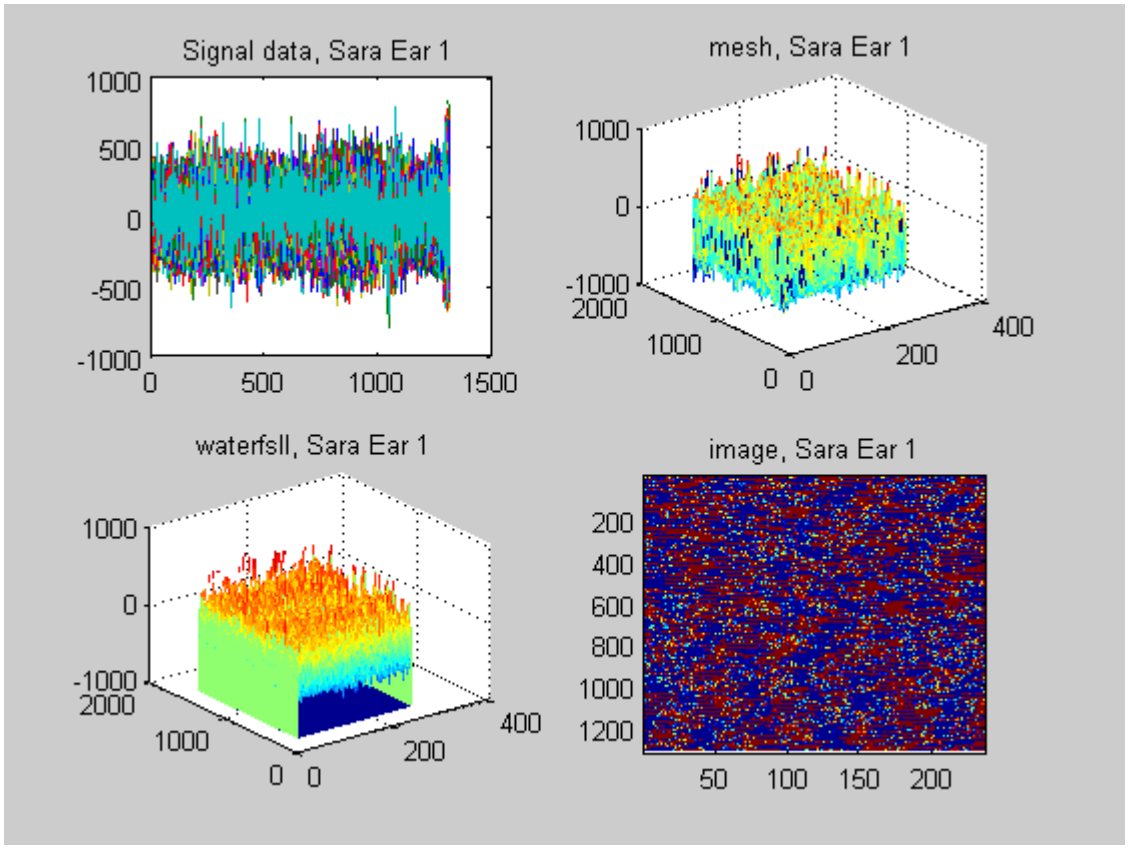
4.1.1 Actual Data

The source of our data is a research center called SensoDetect which has developed the globally patented SD-BERA (SensoDetect-Brainstem Evoked Response Audiometry) technology based on over thirty years research experiences in Lund University. The SD-BERA technology registers electrophysiological reaction patterns that emanate in the brain when specific sound stimuli are presented.

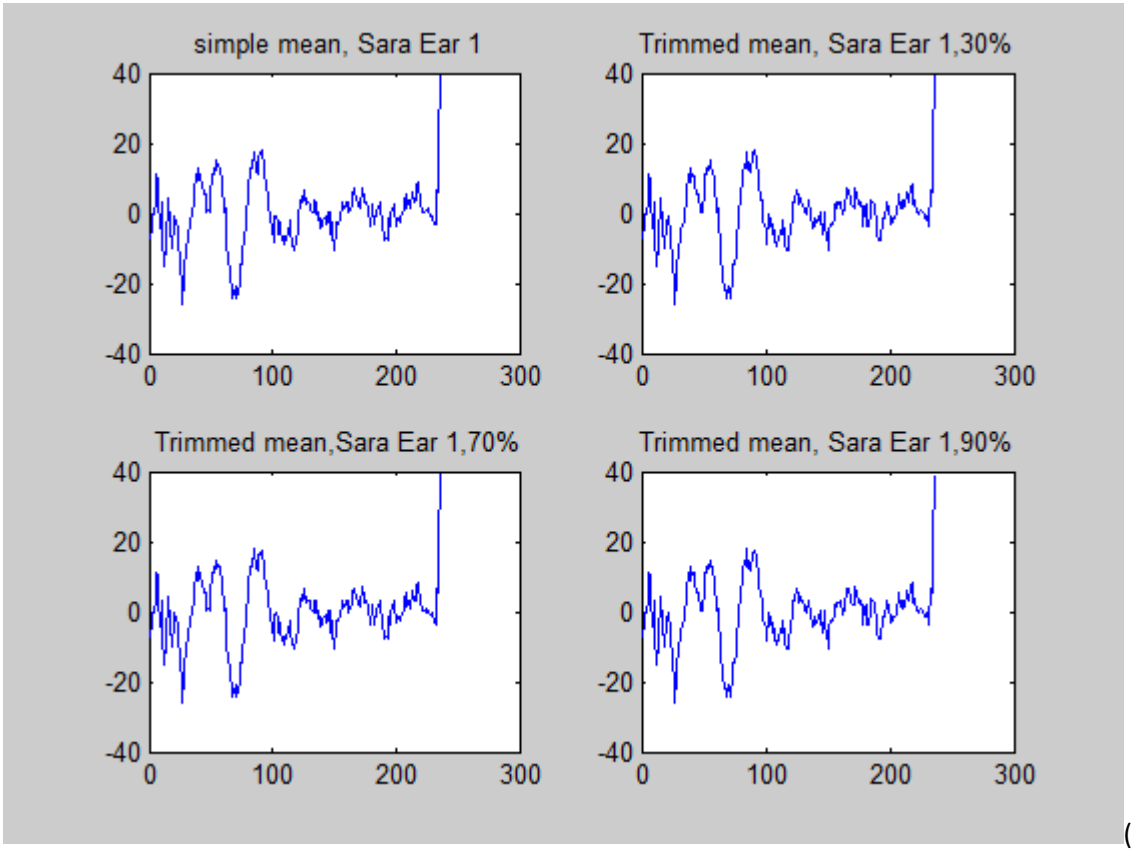
The data set consists of digital signals from Seventeen individuals provided in excel. Each signal data are arranged in 1313 rows and 255 columns such that each row represents a click and each column stands for a time point. Hence, our signal data set is a collection of a 1313×255 real-valued matrices. However, we ignored the first 20 columns as they are too noisy. Hence, we remained with a 1313×235 real-valued matrix. **Figure 4.1** shows only one sample of the Seventeen individuals just to see how the real-valued matrix looks like. From part (b) of this figure, we can observe two things: (1) The arithmetic mean vector and the three different trimmed mean vectors produce quite alike curves. Thus, in our future analysis, we can use any of these mean vectors. (2) Our ABR data has really a sinusoidal nature.

Note: Trimmed mean is a method of averaging that removes a small percentage of the largest and smallest values before calculating the mean. After removing the specified observations, the trimmed mean is found using an arithmetic averaging formula. The trimmed mean looks to reduce the effects of outliers on the calculated average. This method is best suited for data with large, erratic deviations or extremely skewed distributions. A trimmed mean is stated as a mean trimmed by $X\%$, where X is the sum of the percentage of observations removed from both the upper and lower bounds.

Before we indulge directly to the analysis of the actual signal data, it is better to see the theoretical facts using some simulated sinusoidal signal data of the same size as our actual data. A sinusoidal function which is expected to have similar trend as the observed data was developed after a rigorous investigation of the real data. To check how the sample data and the simulated data looks alike see **Figures 4.1-4.3**. Particularly, **Figure 4.3 (b)** shows in a more clear way how the empirical and simulated data look like. Hence, it will be reasonable if we fit and select an appropriate spectral model using the simulated data, and apply this model on the actual data.



(a)



(b)

Figure 4.1 *Sample plots of one individual ABR signals: (a) original (b) different means*

4.1.2 Simulated Signals

For simplification, we will consider the following uncomplicated function, which is a sum of two sinusoidal functions which differ with frequencies and a phase-shift only:

$$z(n) = \sin\left(2\pi \frac{20}{500} n\right) + \sin\left(2\pi \frac{5}{500} n + \varphi\right) + e \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (4.1)$$

where $n = 1,2,\cdots,235$, φ is a phase change which is uniformly distributed between 0 and θ, i.e., $\varphi \sim U[0,\theta]$ such that $0 < \theta \leq 2\pi$ and e is a white noise identically and independently distributed (i.i.d) with mean 0 and variance σ^2, i.e., $e \sim i.i.d(0,\sigma^2)$. **Figure 4.2**, below, shows the plots of a simulated matrix of size 1313×235 and $\theta = {}^{\pi}/_{4}$ (a) without noise, (b) with noise. It is expected that a spectrum of this simulated sinusoidal signal process has peaks at frequencies $f = 20$ and $f = 5\ Hz$.

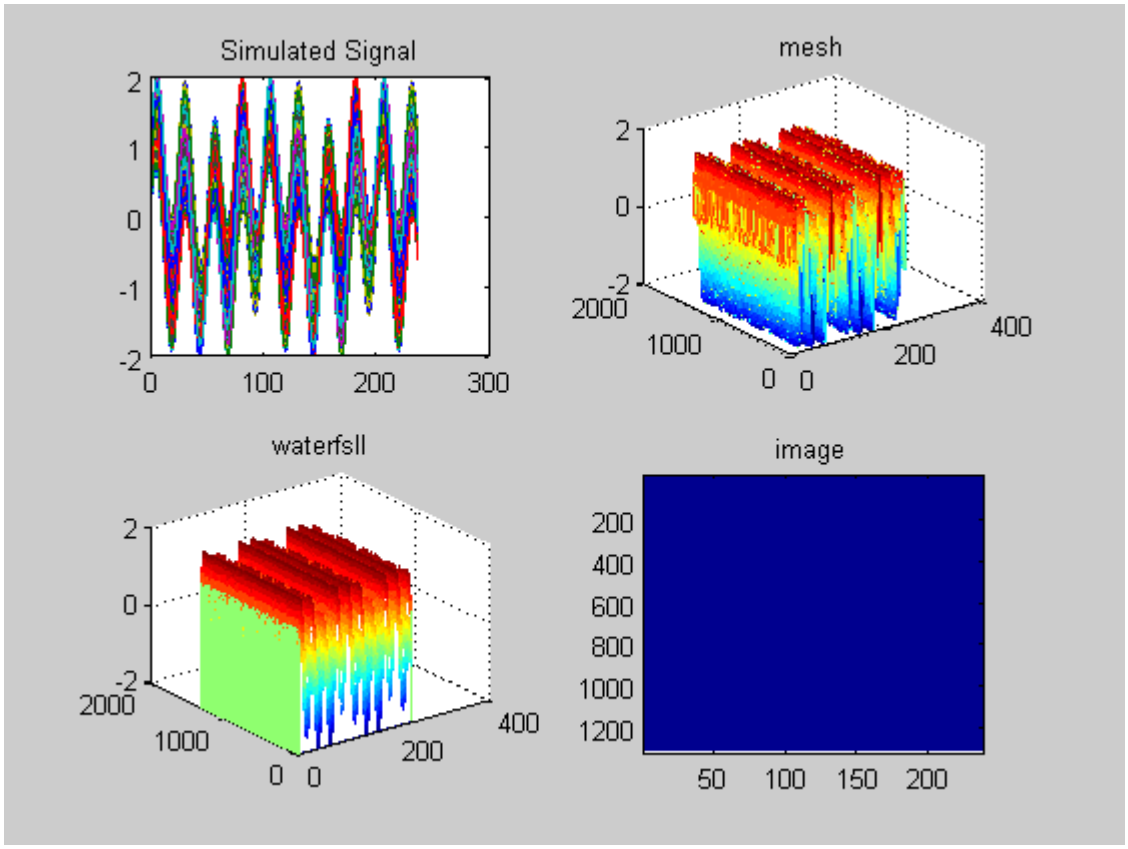
This function can be written as

$$z(n) = x(n) + y(n) + e \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (4.2)$$

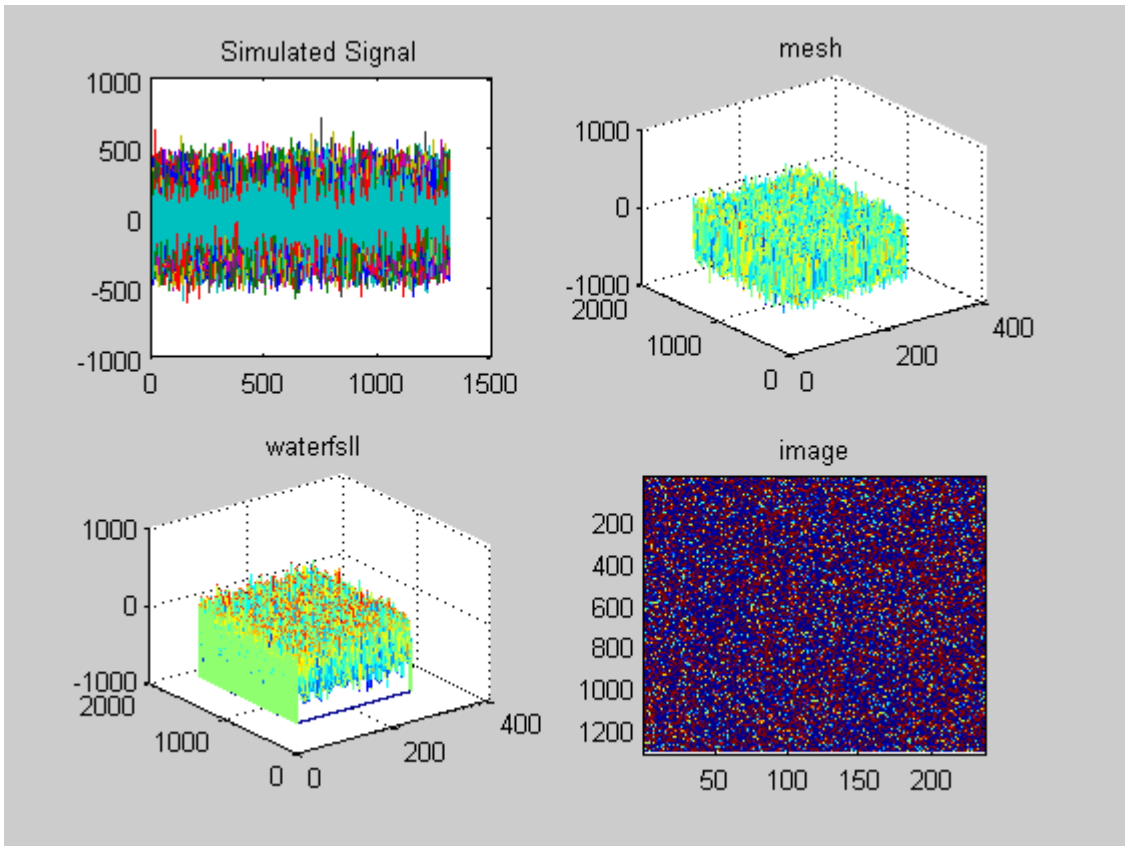
where $x(n) = \sin\left(2\pi \frac{20}{500} n\right)$ and $y(n) = \sin\left(2\pi \frac{5}{500} n + \varphi\right)$. Equation (4.2) is in a similar format with equation (3.21) of section (3.3.3). Consequently, from equation (3.23) of the same section, the actual (observed) spectrum of $z(n)$

$$WVD_z = WVD_x + WVD_y \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (4.3)$$

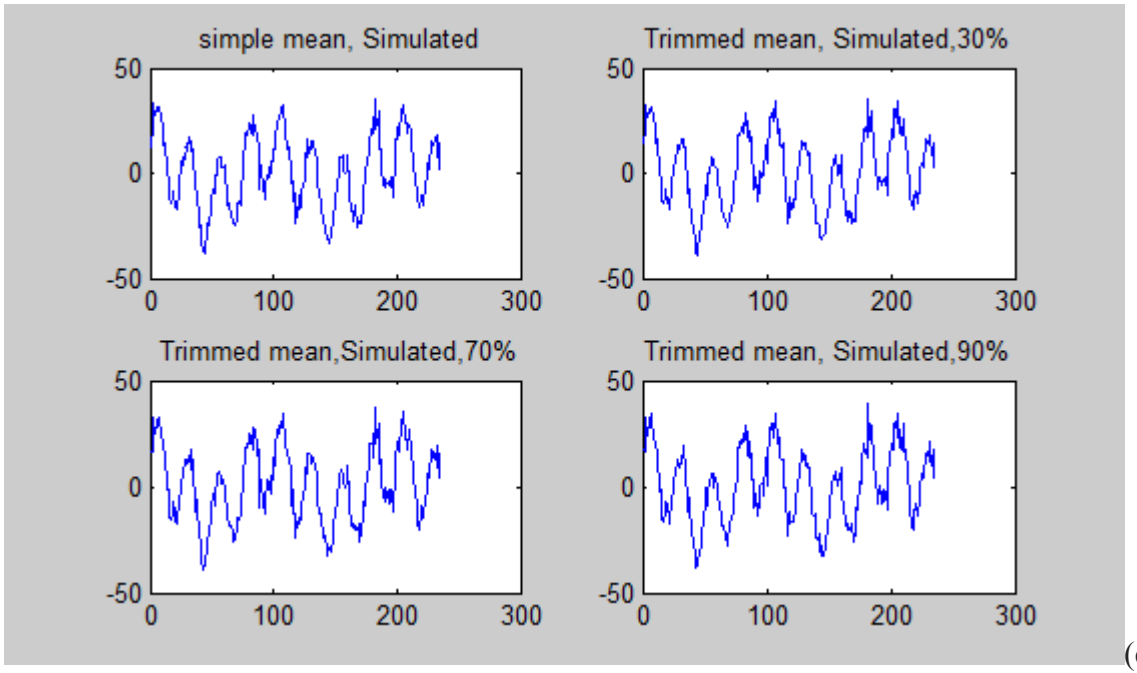
where WVD_x and WVD_y are the Wigner-Ville Distributions of $x(n)$ and $y(n)$, respectively. The reason why we call WVD_z as an actual spectrum is because we include neither a disturbance nor a phase shift in $x(n)$ and $y(n)$ before we compute WVD_x and WVD_y, and that the Wigner-Ville distribution has the best possible performance of all time-frequency methods.


Simulated Signal
mesh
waterfsll
image

(a)

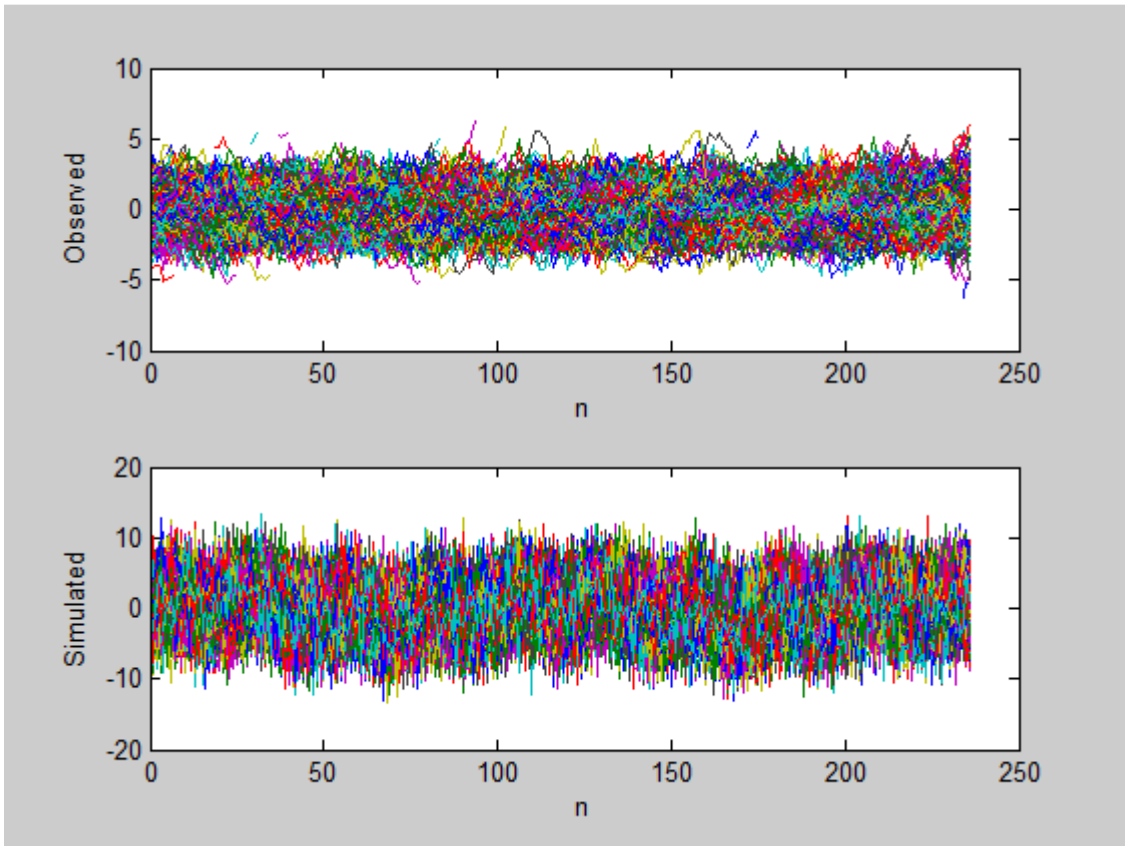

Simulated Signal
mesh
waterfsll
image

(b)

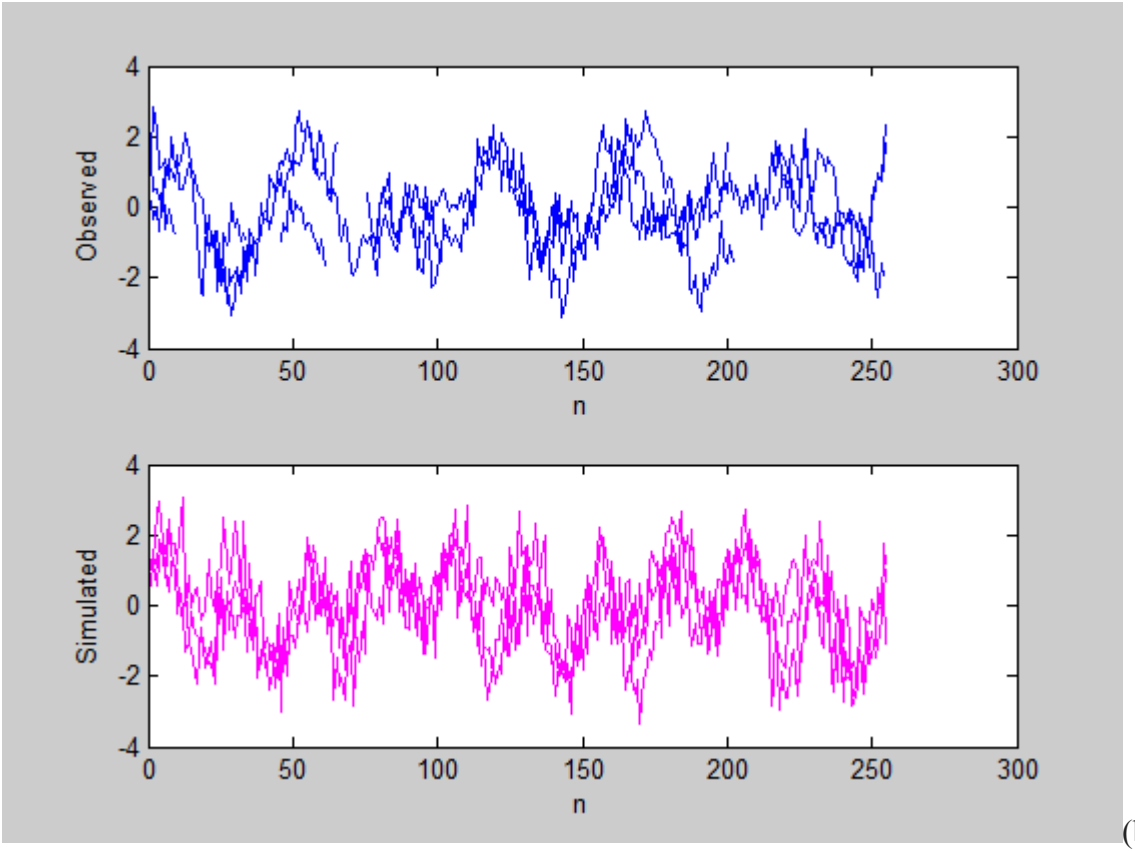


(c)

Figure 4.2 *Plots of the simulated sinusoidal signals: (a) without noise (b) with noise (c) different column means*



(a)



(b)

Figure 4.3 *Comparison between empirical and simulated data: (a) matrix (*1313×235*) (b) the first three columns*

In order to investigate the theoretical facts, it is advisable to study the simulated data in different averages. Hence, for each of the spectral estimation methods we see in the next sections, we will take four different number of clicks per average, these are: the whole(1313) clicks per average which results in a single vector of averages, 300 clicks per average which results in four vectors of averages, 100 clicks per average which results in thirteen vectors of averages, and 50 clicks per average which results in twenty-six vectors of averages, respectively. In the case of the last three number of clicks per average, after having the spectral estimate of each vector, we will take the average spectrum for further investigation.

Practically, observed signals are noisy. Therefore, even though the data set is simulated, we will use the one which embeds a white noise. Because the spectral estimate from the noisy simulated signal will help us to study the observed signals later.

4.2 Data Analysis and Discussion

Before we proceed further, it seems reasonable to put a clear interval of the phase shift to have clearly visible spectra resolution. As shown in equation (4.1), our signal data is simulated from a sinusoidal function with two components which differ only in frequency and phase. In **Figure 4.4 (a)**, we simulated three typical signals. Since the resultant data is a matrix, we changed them in to vectors by taking their corresponding means along time. In this picture, three differently colored curves are observed. These curves stand for no-shift (blue line), 2π-shift (red broken line) and $\frac{\pi}{2}$-shift (magenta line) cases, respectively. The no-shift and 2π-shift curves are totally in-phase showing that shifting the data by 2π brings nothing new but

repeating the original curve after a complete cycle. Had it not been for the color difference, no one could differentiate these two curves. However, it is quite simple to distinguish between the no-shift and $\frac{\pi}{2}$-shift curves. Hence, unlike the 2π-shift, the $\frac{\pi}{2}$-shift produces a new signal which is different from the original one in some senses.

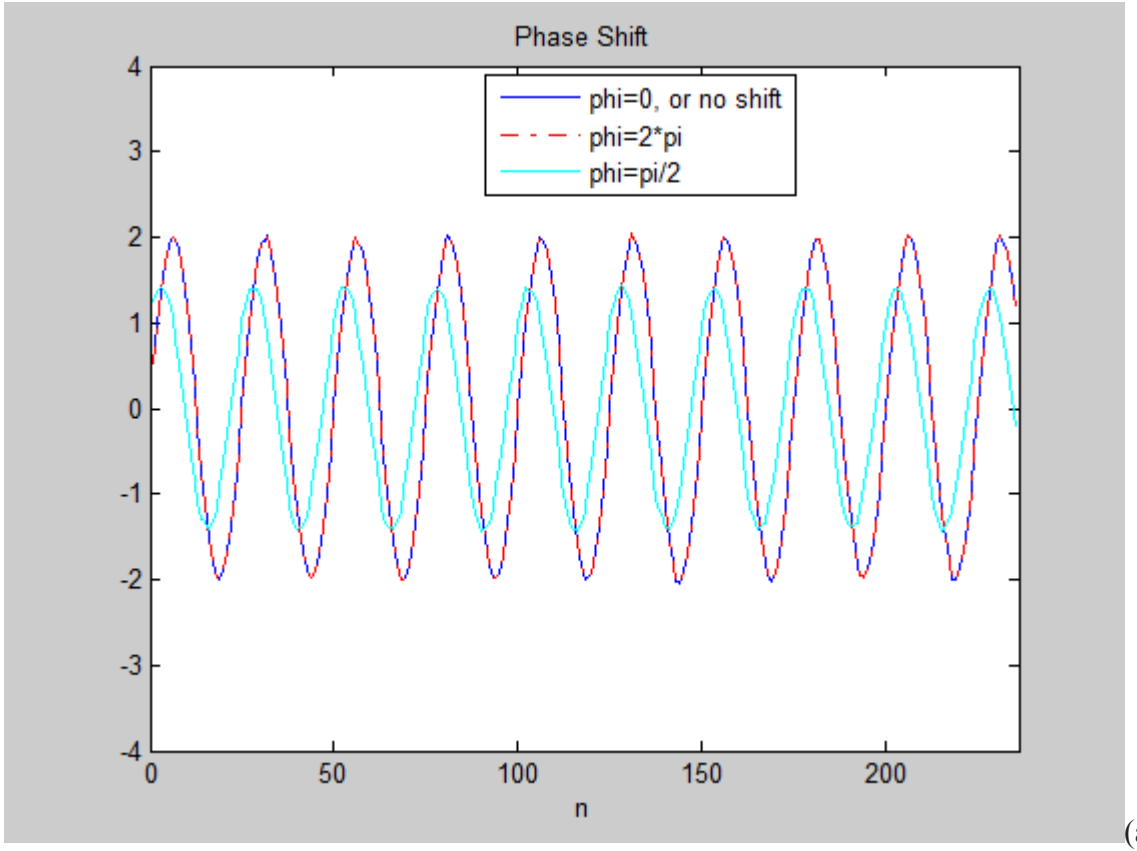
Similar figures are given in **Figure 4.4 (b)**, except here φ is treated as a random variable which is uniformly distributed in some intervals: $\varphi \sim U[0,2\pi]$ (red line), $\varphi \sim U[0,\pi/2]$ (magenta line), $\varphi \sim U[0,\pi/4]$ (green line).In these curves the phase shifts, respectively, are $5.5363\,rad\ (\approx 316.66^{\circ}), 0.5267\,rad\ (\approx 30.18^{\circ})\ and\ 0.1775\,rad\ (10.17^{\circ})$. Note that the shift corresponding to $\varphi \sim U[0,2\pi]$ is almost one cycle.

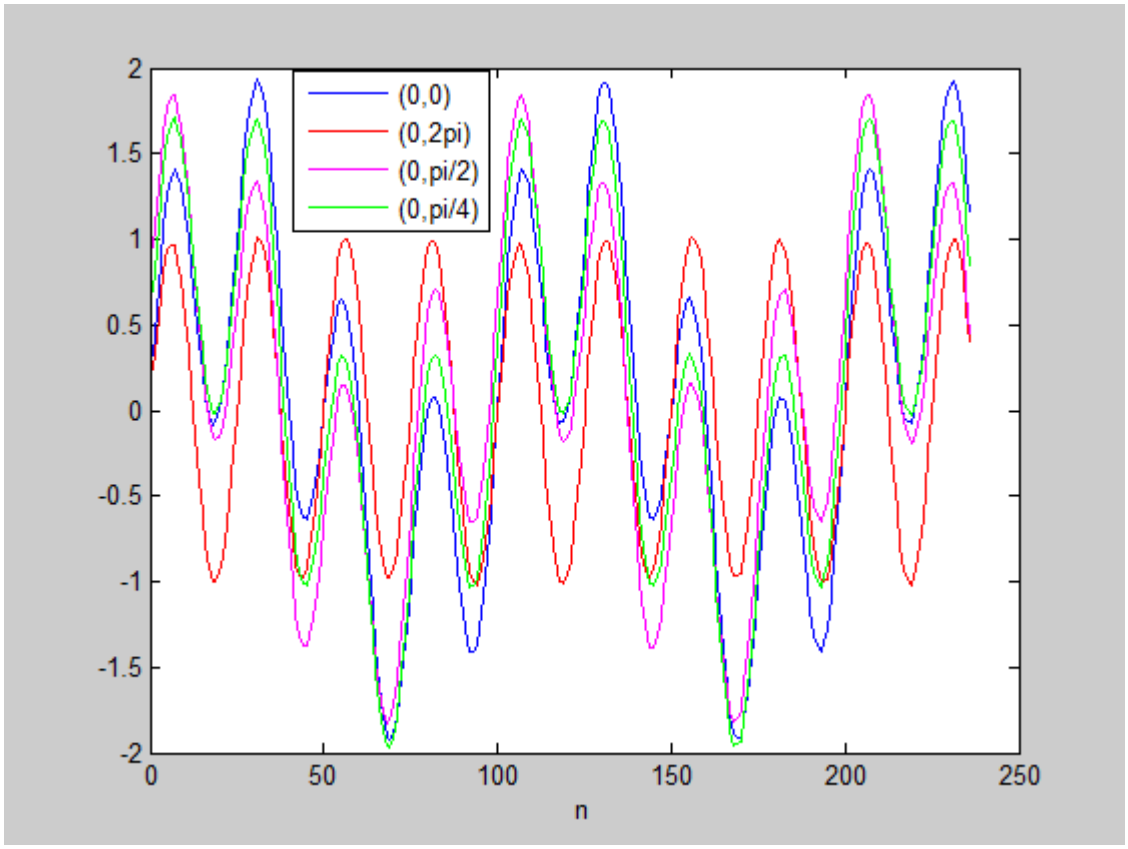
The facts that we saw under the phase-shift curves will also be seen on the corresponding spectral representations. The left-top spectrograms of **Figure 4.4 (c)** correspond to a phase shift uniformly distributed in the interval $[0,2\pi]$; in this case the spectrograms show a single component instead of two. This is due to the in-phase behavior which results almost complete overlapping of the original signal and the shifted one. In the same figure, right-top we have $\varphi \sim U[0,\pi]$, still the second component is not clearly noticeable. The left-bottom and right-bottom spectrograms correspond, respectively, to $\varphi \sim U[0,\pi/2]$ and $\varphi \sim U[0,\pi/4]$, with the latter one having best frequency resolution.

It is palpable that the narrower the interval the better is the spectrogram in frequency resolution and showing the components clearly. Hence, in our case, to have a better resolution in frequency, we should have $\varphi \sim U[0,\theta]$ such that $0 < \theta$ and $\theta \to 0$. This interval is true for all spectra that we are going to see; for our convenience, we will use $\varphi \sim U[0,\pi/4]$ for all our discussion henceforth.

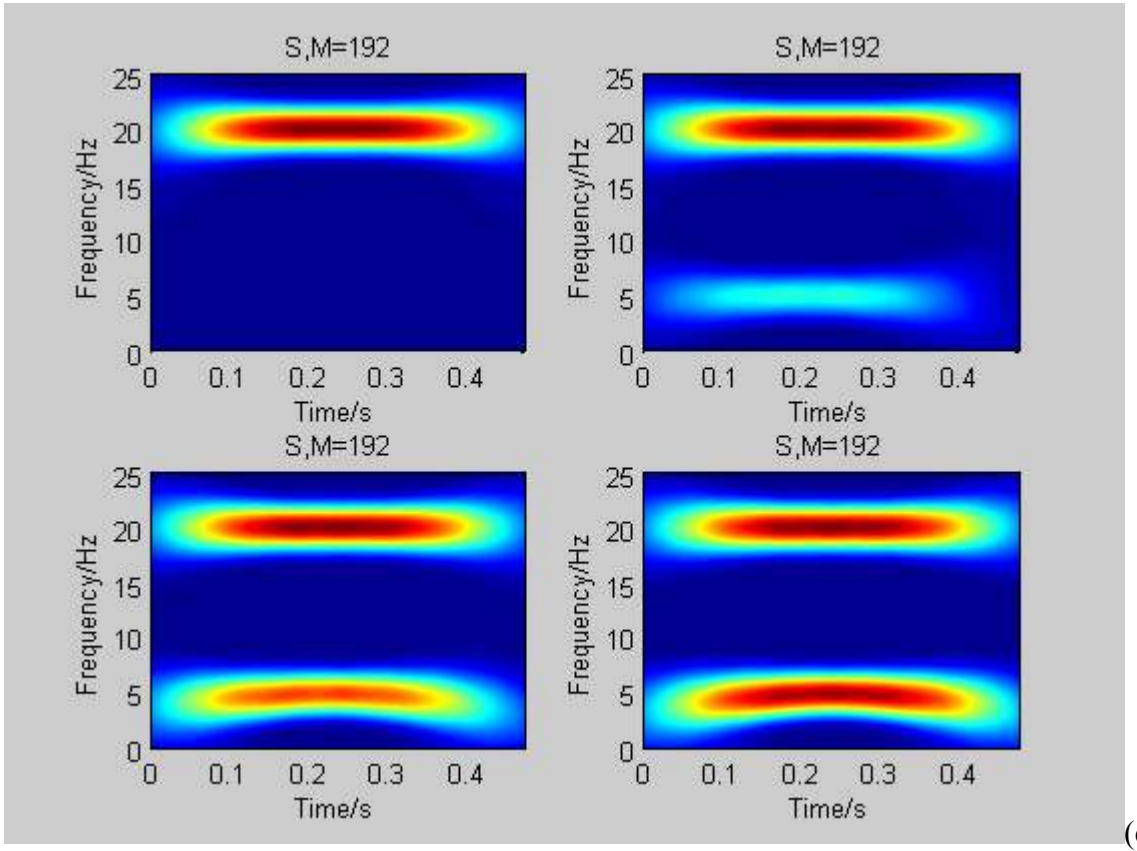
Another way to see the frequency-peaks in the spectrogram is to take individual spectrograms of the component signals and adding them irrespective of the value of the phase shift. **Figure 4.4 (d) left**, represents the spectrogram of $w_1(n) = z_1(n) + z_2(n)$ with $z_1(n)$ and $z_2(n)$ defined by Equation (4.1), including the phase shifts $\varphi_1 = 0$ and $\varphi_2 = 2\pi$, respectively. **Figure 4.4 (d) right**, presents the sum of the individual spectrograms of $z_1(n)$ and $z_2(n)$, i.e., $S_{w_2}(t,f) = S_{z_1}(t,f) + S_{z_2}(t,f)$, where $S_{z_1}(t,f)$ and $S_{z_2}(t,f)$, respectively, are the spectrograms of $z_1(n)$ and $z_2(n)$, showing the advantage in averaging spectrograms for noisy sequences of shifting phase.

Having found an appropriate interval for the phase shift, an important issue that we have to consider under the analysis of spectrogram is the window length, as we discussed in chapter three, the shorter the window length the better resolution in time is and the longer the window length the better resolution in frequency is. As shown in **Figures 4.5-4.16**, **Appendix A**, the time-resolution is getting worse when the window length (M) runs from $M = 64$ to $M = 192$; conversely, the frequency-resolution is becoming better when the window length varies from $M = 64$ to $M = 192$. This fact will be seen more clearly if we take too small or too large window lengths.

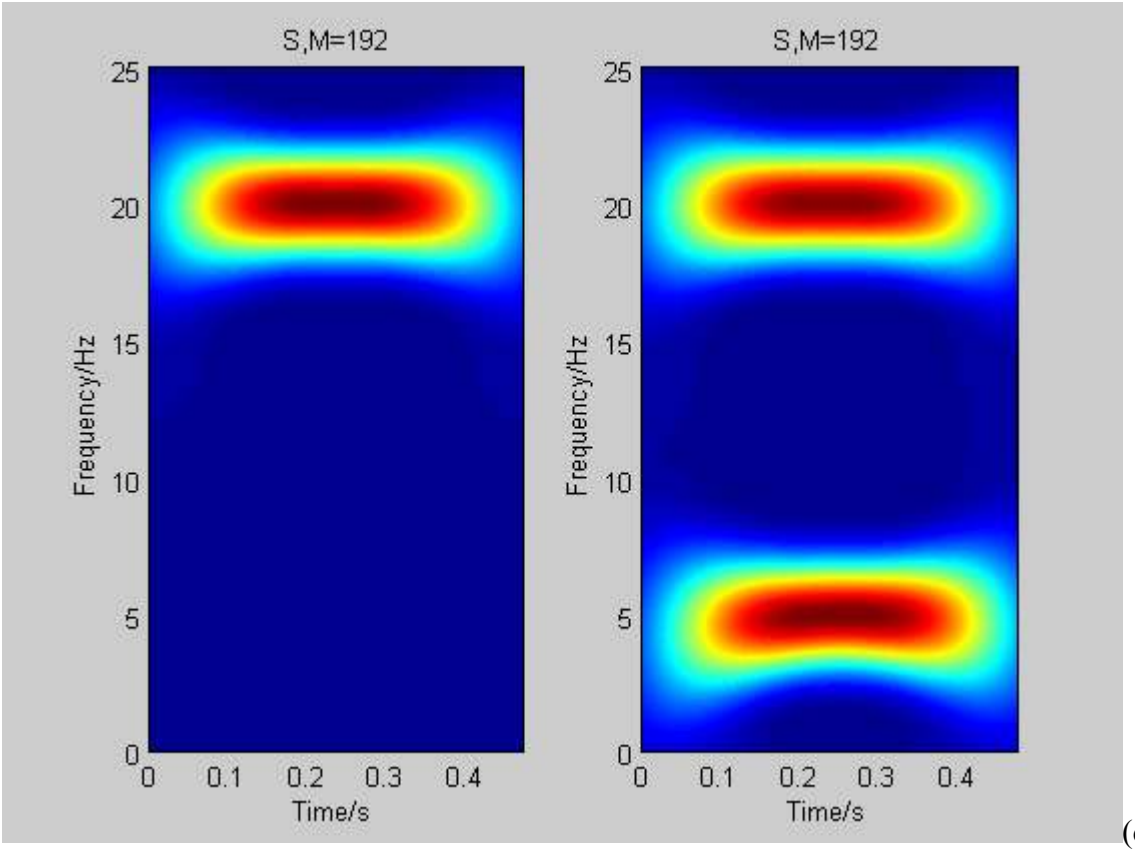

Phase Shift
phi=0, or no shift
phi=2*pi
phi=pi/2
4
3
2
1
0
-1
-2
-3
-4
0
50
100
150
200
n
(a)



(b)



(c)



(d)

Figure 4.4 *(a) plots of the means of the simulated signals for* $\varphi = 0, \varphi = 2\pi$ *and* $\varphi = \pi/2$*, (b) plots of the means of the simulated signals for* $\varphi = 0, \varphi \sim U(0,2\pi)$*,* $\varphi \sim U\left(0, \frac{\pi}{2}\right)$ *and* $\varphi \sim U\left(0, \frac{\pi}{4}\right)$*, (c) Spectrograms differ due to changes of phase shift* φ*: left-top* $\varphi \sim U[0,2\pi]$*, right-top* $\varphi \sim U[0,\pi]$*, left-bottom* $\varphi \sim U[0,\pi/2]$ *and right-bottom* $\varphi \sim U[0,\pi/4]$ *for1313 clicks per average, (d)(left) spectrogram of sum of two components with* $\varphi = 2\pi$ *and right sum of spectrogram of the components.*

A more careful observation of the different spectral estimates in **Figures 4.5-4.16**, **Appendix A**, confers better resolutions in some and bad ones in others. Basically, the figures are of three groups: Figures 4.5-4.8, Appendix A.1, for Spectrograms, Figures 4.9-4.12, Appendix A.2, for Thomson Multitapers and Figures 4.13-4.16, Appendix A.3, for PM MW. Better resolutions are observed in the cases of Spectrogram and PM MW than that of Thomson Multitapers. This may be a clue towards the good fit of the simulated data, that is either the Spectrogram or the PM MW or else both estimates fit well the sinusoidal simulated data. On the other hand, this probably means that the Thomson Multitaper estimate is not a good fit for the mentioned data. However, we will know the decision in the next section.

If this is going to be true, one can raise questions like, "why Spectrogram or PM MW or both but not Thomson Multitaper?". It is possible to put a seemingly reason over here. Theoretically, to employ Thomson Multitaper estimator we assume that the underlying sample data come from a stationary family. Yet, the simulated sinusoidal signal sample data are expected to come from a non-stationary family. Consequently, this might be one reason which hinders the Thomson Multitaper to look like a good fit of the data. Albeit this assumption is also valid for the PM MW estimator, the PM MW method is suitable for peaky spectra

which is the case for the simulated data which has high peaks at frequencies $f = 5$ and $f = 20$. Hence, this might help for the PM MW to have a better resolution compared to the Thomson Multitaper estimator. Finally, it is obvious that the Spectrogram is used very frequently for analyzing time-varying and non-stationary signals. This could be the reason it has better resolution compared to the Thomson Multitaper.

4.3 Comparison of Estimators

In section 4.2 above we presented the data analysis along with its discussion based on the three spectral estimators namely, Spectrogram, Thomson Multitaper and PM MW. However, it is natural that different models outperform differently in diverse situations. For this reason, in this section, we are going to make some comparisons amongst these three estimators and select the one which we think fits 'best' the simulated sample signal.

The simplest technique to check whether a spectral estimate fits an actual spectrum or not is to plot both side by side or one over the other. A complete similarity (or fit) in shape of the two is a perfect agreement of the actual and estimated spectra. In fact, practically, there is no a perfect match between a model and its estimate. So, the more they look alike the better the fit will be.

A cautious surveillance of **Figures 4.17-4.19, Appendix B.1**, leads one to say that the Spectrogram method creates a better fit between the actual spectrum ($\boldsymbol{WVD}$) and the estimated spectrum ($\boldsymbol{S}$) (multiplied by a constant) than the other two methods. It seems also obvious that the Thomson Multitaper method gives the worst fit.

Once again, in section 4.2, we suggested that the Spectrogram seems to give as equal resolution as the PM MW does. Nevertheless, here, we should not be tempted to say the Spectrogram and PM MW equally perform just by giving a quick glance at the plots; if we examine their corresponding plots, we can easily come out with a conformity that the Spectrogram appears to be better.

At this time, we need to have a numerical evidence to support our suggestion stating that "the Spectrogram method outperforms than the other two," which is based on only the visual plots. Because such visual plots may not be accurate. Often, the MSE or RMSE is used to make comparison among different spectral estimates. In **Tables 4.1-4.3, Appendix C.1**, we present the RMSE values for the four different number of clicks per average of the three estimates. Generally, we can see easily that the RMSE values of the Spectrogram for the first fourteen or fifteen standard deviations of the noise are very low compared to the corresponding values of the other two methods. Particularly, the overall mean RMSE of the Spectrogram, Thomson Multitaper, and PM MW are 14.30, 21.85 and 16.55, respectively, confirming the fact that the spectrogram still has the chance to be best. Yet, there is a doubt whether the supremacy of the spectrogram is uniform throughout all values of σ, the standard deviation values of the disturbance. This hesitation arises

from a careful observation of the RMSE and average RMSE for higher values of σ. For illustration, let us look the RMSE and mean RMSE corresponding to $\sigma = 15,30,$ and 40, respectively:

- Spectrogram

σ	clicks	1313	300	100	50	mean
15		10,3399	15,53	29,6501	36,5119	**23,01**
30		15,6375	37,5103	39,4501	45,4913	**34,52**
40		19,6915	47,9764	45,234	48.9517	**37,63**
	mean	**15,22**	**33,67**	**38,11**	**41,00**	**32,00**

- Thomson Multitaper

σ	clicks	1313	300	100	50	mean
15		19,583	22,7697	29,3849	36,2598	**27,00**
30		23,0915	30,7269	38,3026	45,7474	**34,47**
40		26,9587	37,074	41,7168	42,9826	**37,18**
	mean	**23,21**	**30,19**	**36,47**	**41,66**	**32,88**

- PM MW

σ	clicks	1313	300	100	50	mean
15		13,3143	15,6242	29,4521	35,786	**23,54**
30		19,7392	33,2098	44,3804	40,2838	**34,40**
40		25,6733	37,9338	41,1894	40,2633	**36,26**
	mean	**19,58**	**28,92**	**38,34**	**38,78**	**31,40**

From the above tabular values, it is possible to see how the performance of the Spectrogram method is degraded as the σ values get higher. For $15.0 \leq \sigma \leq 30.0$, the Spectrogram and the PM MW methods have more or less the same (mean) RMSE values with a bit lower in PM MW for fewer number of clicks per average. This means that around $15.0 \leq \sigma \leq 30.0$ these two methods perform nearly the same if not PM MW better. Note that, for these values of σ, the Thomson Multitaper method persist to be bad performer. But, if we approach to higher σ values like $\sigma \geq 30.0$, gradually, the Spectrogram method is acquainted with higher (mean) RMSE than the rest two methods. For such values of σ lower values of (mean) RMSE are associated with PM MW. This might be an indication that for highly noisy simulated signals of our case, either the Thomson Multitaper or the M MW method is preferable than the Spectrogram method. The highest preference might be given to the PM MW.

To make things easily and more clear, it is better to plot the mean RMSE values. Making an intense glimpse on **Figure 4.20, Appendix C.1**, provides a concrete evidence in a more obvious way to see the performances of the estimators. The strength of the spectrogram estimate lasts only for some interval of the noise-standard deviation values. In this graph, we plot the average RMSE vectors of the three methods of

estimation against the noise-standard deviation values. For all values of the standard deviation less or equal to 15.0, i.e., $0 \leq \sigma \leq 15.0$, the average RMSE for the Spectrogram are the lowest. The highest values are assigned to the corresponding values of the Thomson Multitaper. Thus, this picture gives unambiguous evidence that the Spectrogram performs top in this interval.

Further investigation of the same graph certifies that for $15.0 \leq \sigma \leq 30.0$, the PM MW performs as equal as the Spectrogram does, if not better. Finally, for $\sigma \geq 30$ the PM MW performs best and the Spectrogram worst. These observations can be summarized as below,

Interval for noise-standard deviation (σ)	Model Performance		
	Best	Medium	Worst
$0 \leq \sigma \leq 15$	Spectrogram	PM MW	Thomson Multitaper
$15 \leq \sigma \leq 30$	PM MW & Spectrogram are nearly equal, if not PM MW	slightly the Spectrogram compared with PM MW	Thomson Multitaper
$\sigma \geq 30$	PM MW	Thomson Multitaper	Spectrogram

Before we windup this section, we prefer to say a few on the impact of the intensity (or strength) of the white noise within the simulated signal on the spectral resolution. This strength could be measured using the standard deviation values of the noise. Obviously, high-valued standard deviation is followed with strong disturbance. This truth is seen if we compare **Figure 4.21 (a)** and **(b), Appendix C.2.** The plots in (a) are obtained from a standard deviation value 0.5 and the plots in (b) correspond to 40.0 standard deviation value. The deviation of the estimated spectrum from the actual one is very high in the latter case. This designates that we should impose a fairly small amount of the white disturbance to the simulated signal. A vigilant observation of **Figures 4.22-4.24, Appendix C.2** and **Figure 4.20, Appendix C.1** shows the preferable standard deviation values for the white noise may range from 0.0 to 2.0 to have a good spectral resolution. To this end, we would like to inform that all the other analysis results were done for the white noise standard deviation values from this range.

4.4 Application to Empirical Data

So far we have been trying to select an appropriate model. Now, we will climb to the last stage of this chapter-exploiting the selected model to figure out some facts from an empirical sample signal. The model that we are going to use is the spectrogram. This is reasonable because the amount of noise in the empirical signals seems as strong as the white disturbance we impose on the simulated signal with σ less than 15.0, for which the spectrogram is best. In addition to this we use different averages which smooth out disturbance.

We will begin our discussion with experimental signal responses as heard by the first ear (as put in the data set from the company) of an individual called Sara. **Figure 4.25, Appendix D**, exhibits Spectrograms with four different window lengths (i.e., $M = 64, 96, 128\ and\ 192$) of the four number of clicks per average (i.e., $1313, 300, 100$ and 50 clicks per average, respectively).

Though showing time-frequency resolution tradeoff, we can guess that the highest peak is located at around frequency point $f = 10$ and time point $t = 0.15$. But, we cannot see the side-lobes clearly, so that information leakage my lead us to unwanted conclusions. In order to compensate this problem, we tried to have plots of frequency (or time) versus smoothed spectral estimates separately in a such away many peaks can be visible easily. **Figure 4.26, Appendix D**, contains such plots. In doing so, first, we took higher window length for better frequency resolution and then lower window length for better time resolution. Finally, we use a robust peak detection mechanism in noisy signals separately for frequency and time. **Table 4.4, Appendix D**, demonstrates the four highest peak points along time and frequency of the spectral estimate corresponding the four number of clicks per average at appropriate window lengths in relation with time and frequency resolutions. In average, the highest peak is situated at $t = 0.152$ and $f = 11.141$ along time and frequency, respectively, which is quite near to the guess we made from the spectral pictures.

An overwhelming trend is seen as we decrease the number of clicks per average. In the first two number of clicks per average (i.e., 1313 and 300), the strong peak(s) situated in the same position mentioned above. But, in the other number of clicks per average (i.e., 100 and 50), there are more strong peaks at around $f = 10$ and $t = 0.45$. This might be one way of exposing hidden information in the larger number of clicks per average.

Apparently, employing the estimated Spectrogram to produce all the plots and discussions made for Sara's first ear is time and space taking. Above all, if we do the same procedures for all the Seventeen individual ears in the data set, this will not be only time and space taking, but also it will be tedious and unmanageable. However, in order to see the consistency and performance of the model for our ABR data, we will discuss two other ears: Sara's second ear and Stellan's first ear.

In a similar manner with Sara's first ear case, we can have all the plots and arguments for her second ear too. **Figure 4.28, Appendix D**, reveals the spectral estimates of the four number of clicks per average for different window lengths in case of the second ear. These spectrograms have analogous characteristics with the first ear corresponding plots except a little shift in peak locations both in time and frequency, in average the highest peak is located at around $f = 11.1388$ and $t = 0.187$, see **Figures 4.29 and 4.30,** and **Table 4.5**, **Appendix D**. The same is true for Sellan's first ear, with highest peak at around $f = 0.2215$ and $f = 10.4215$, see **Figures 4.31-4.33** and **Table 4.6**, **Appendix D**.

From the spectral representations of these three individual ears we have observed some common trends. Concerning peaks in relation with frequencies, higher peaks are commonly situated at the lower frequency values, often below $50\ Hz$. In addition, in most cases, the highest peak corresponds to the lowest frequency value of frequencies values with peaks, often $9 - 12\ Hz$, watch **Figures 4.26**, **4.29**, and **4.32**, **Appendix D**. Plots in connection with peak along time slots of Sara's both ears situated at the lower time values. But, for Stellan's first ear, the same peaks are firmed towards the middle time values, look **Figures 4.27**, **4.30**, and **4.33**, **Appendix D**. One last common observation is that the number of peaks along time points increases as the number of clicks per average decreases. Number of peaks along time ranges from 4 to 10 as the number of clicks per average decreases from 1313 to 50, **Figures 4.27**, **4.30**, and **4.33**, **Appendix D**. In general, what we understand from the three ear cases is that the selected model could be applicable in time-frequency spectral representation of the ABR signal data from SensoDetect.

5. CONCLUSIONS AND RECOMMENDATIONS

The main objective of this study is to fit a time-frequency representation of the ABR data registered by SD-BERA from SensoDetect by comparing three spectral estimation methods: Spectrogram, Thomson Multitaper, and Peak Matched Multitaper. In doing so, we have passed through several course of actions.

One of our accomplishments is proposing an apt interval of values for the phase in the function we used for simulation. The function we used for simulation is a sum of two sine functions which differ in frequencies and a phase only. The spectral estimate from this simulation is expected to have peaks at the two frequencies. Nevertheless, these peaks will be noticeably seen for some values of the phase. Hence, it is shown that such phase φ exists if $\varphi \sim U[0, \theta]$ such that $0 < \theta \leq 2\pi$ and $\theta \to 0$.

Verdict on the interval of the phase value was followed by the core part of the analysis. This stage includes all the sections which were established for the purpose of model selection. Broadly viewed, the two sections, i.e., sections $4.2\ and\ 4.3$ contributed some aids for the purpose of model selection out of the three speculated models. What we see in these sections could be summarized as, (i) judging against frequency resolutions, Figures 4.5-4.16, Appendix A, (ii) creating visual plots, Figures 4.17-4.19, Appendix B and (iii) measuring goodness of fitness numerically/statistically, Tables 4.1-4.3 and Figure 4.20, Appendix C.1. We stated that the Spectrogram and PM MW methods have better frequency resolution than the Thomson Multitaper technique. But this will not place us in a position to distinguish the frequency resolutions of the first two methods. Visual plots are plots such that the estimated spectra is drawn together with (or over) the actual ones. In some cases, they give adequate meaning to give an opinion on the fit of an estimate and its estimand. We noticed that the Spectrogram method gives finer fits than the rest two methods, even better than the PM MW. Sometimes, conclusions based on figures only may misguide us to unwanted result. Therefore, what we observe from graphical facts should be tested statistically for its reliability. The RMSE was the tool we exploited for this purpose. Based on this tool, we found out that the Spectrogram method outshines over the other two methods for disturbances with $0 \leq \sigma \leq 15.0$. Note that this last conclusion is in a harmony with the results we encounter at the frequency resolution and visual plot levels. But for higher disturbances, $15.0 \leq \sigma \leq 30.0$, the PM MW performs as good as the Spectrogram, if not better. Finally, for even higher disturbances, $\sigma \geq 30.0$, the PM MW becomes to be the best and the Spectrogram turned down to the ground level, even exceeded by Thomson Multitaper method. Hence, under our study, we conclude that all the three methods could work based on the intensity of the disturbance imposed on the signal data.

Hand in hand with the model selection process, we suggestion on the possible intensity of the white disturbance embedded within the simulated data. The amount of noise imposed on a simulated signal could have immense impact on the estimated spectral frequency resolution. Thus, a precaution should be taken to the degree of noise we need to inflict on the signal. Usually, the level of white disturbance could be quantified using its standard deviation. A higher standard deviation is followed by a poor (frequency)

resolution in the spectral estimate. That is, as low as possible standard deviation is required to get a good resolution. In our case, we succeeded to show that the favorable range of standard deviation for the white nose we apply on the signal is more or less between 0.0 and 2.0, Figures 4.21-4.24, Appendix C.2 and Figure 4.20, Appendix C.1.

A best fitted model using a simulated data should be tested for functionality via a real observed sample. To accomplish this, we applied the Spectrogram on observed experimental signals heard through three ears (2 from the first individual and 1 from the second one) of individuals-Sara and Stellan, Figures 25-33**,** and Tables 4.4-4.6**, Appendix D**. The reason we use the spectrogram for the empirical signal is that the data have comparatively low disturbance. In average, highest peaks for Sara's 1st and 2nd ears and Stella's first ear are located at (0.1415,11.124), (0.187,11.1388) and (0.2215,10.4215), respectively, where the first coordinate stands for time and the second one for frequency. Some common observations seen in all the three empirical signals are: (1) higher peaks are located at lower frequency values with peaks, $f \leq 50\,Hz$, (2) the highest peaks are put in the $9-12\,Hz$ frequency ranges, etc. To sum up, we dare to say that the Spectrogram is the best one of the three methods under the study in representing the ABR data registered by SD-BERA.

Finally, we would like to forward some recommendations concerning future improvement of this modeling procedure for ABR data. As it is obvious that there is a tradeoff between resolution and window length of the Spectrogram, an improvement should be made. Though it is not part of this study, there is a method of finding optimal window length for the spectrogram which gives us better resolution both in time and frequency simultaneously. It is also advisable to look for other techniques like time-frequency reassignment and multitaper time-frequency reassignment which are recent and efficient in smoothing out the spectral estimate to increase readability and decrease variability. Instead of sticking ourselves to nonparametric approaches only, sometimes, it might be wise to test parametric approaches too.

REFERENCES

[1]. Boualem Boashash, "Time Frequency Signal Analysis and Processing," A Comprehensive Reference, Elsevier 2003

[2]. J. Xiao and P. Flandrin, "Multitaper Time-Frequency Reassignment for Nonstationary Spectrum Estimation and Chirp Enhancement," IEEE Transactions on Signal Processing, Vol 55, No. 6.

[3]. L.Cohen, " Time-Frequency Distributions-A Review," Proc IEEE, vol 77, pp. 941-981, July 1989, Invited Paper.

[4]. M.B.I. Raez, M.S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," January 18, 2006.

[5]. P. Flandrin, F. Auger and E. Chassande-Mottin, " Time–frequency reassignment: from principles to algorithms," http://perso.ens-lyon.fr/patrick.flandrin/0065-Ch05.pdf

[6]. S. K. Mitra, "Digital Signal Processing-A Computer-based approach," McGraw-Hill series in electrical and computer engineering, 2006.

[7]. Z.Leonowicz, T. Lobos, and J. Rezmer, "Spectrum Estimation of Non-Stationary Signals in Power Systems," (2003), International Conference on Power Systems Transients-IPST 2003 in Hong Kong.

[8]. A.T. Walden, "A Unified View of Multitaper Multivariate Spectral Estimation," Biometrika, Vol. 87, No. 4 (Dec., 2000), pp. 767-788

[9]. A.A.M. Teamah, and H.S. Bakouch, "Multitaper Multivariate Spectral Estimators of Time Series with Distorted Observations," International Journal of Pure and Applied Mathematics, Volume 14 No. 1 2004, 45-57

[10]. D. J. Thomson, "Spectrum Estimation and Harmonic Analysis," Proceedings Of The IEEE, Vol. 70, No. 9, September 1982

[11]. P. D. Welch, " The Use of Fast Fourier Transform for the Estimation of Power Spectra," A Method Based on Time Averaging Over Short, Modified Periodograms, IEEE Trans. on Audio Electro acoustics, AU-15(2):70-73, June 1967

[12]. T.P. Bronze, "On the Performance Advantage of Multitaper Spectral analysis," IEEE Trans. on signal Processing, V. 40, Issue 12, pp. 2941 - 2946, Dec 1992.

[13]. L. Varshney, "On the Use of Discrete Prolate Spheroidal Windows for Frequency Selective Filter Design," Applications of Signal Processing, School of Electrical and, Computer Engineering, Cornell University, February 23, 2004.

[14]. G. A. Prieto, R. L. Parker, D. J. Thomson, F. L. Vernon and R. L. Graham, "Reducing the bias of multitaper spectrum estimates," Accepted 2007 August 20. Received 2007 August 20; in original form 2007 February 23.

[15]. M. A. Wieczorek and F. J. Simons, "Minimum-Variance Multitaper Spectral Estimation on the Sphere," The Journal of Fourier Analysis and Applications.

[16]. K. S. Riedel and A. Sidorenko," Minimum Bias Multiple Taper Spectral Estimation," Courant Institute of Mathematical Sciences, New York University, New York, I.E.E.E. Trans. on Signal Processing, vol. 43, pp. 188-195, Jan. 1995.

[17]. M. Hansson and G. Salomonsson, "A Multiple Window Method for Estimation of Peaked Spectra," IEEE Transactions, On Signal Processing, Vol. 45, No. 3, 778-781,March 1997.

[18]. Ö, Yeniay and A. Göktas, "A Comparison Of Partial Least Squares Regression With Other Prediction Methods," Hacettepe Journal of Mathematics and Statistics, Volume 31 (2002), 99-111, Received 25. 06. 2002.

[19]. J. Park, " Multitaper Spectra analysis of High-frequency Seismograms," Journal of Geophysical Research, Vol. 92, pp. 12,675-12,684, November 10, 1987.

[20]. C. Emrick, " Estimation of Hearing Sensitivity using the Auditory Brainstem and Auditory Steady State Responses," A Senior Honors Thesis, June 2008.

[21]. A. R. Möller, "Hearing: Its Physiology and Pathophysiology," San Diego, London, Academic Press, 2000.

[22]. R. L. Malinoff and L. G. Spivak, " Effects of Stimulus Parameters on Auditory Brainstem Response Spectral Analysis," New York Eye and Ear Infirmary, Audiology, 1990, vol. 29, pp. 21-28.

[23]. J. W. Hall, "Auditory Brainstem Response Spectral Content in Comatose Head-injured Patients," Ear and Hearing: Electrophysiological Techniques in Audiology, vol 7, issue 6, pp. 383-389, Dec. 1986.

[24]. D. Urbach and H. Pratt, " Application of Finite Impulse Response Digital Filters to Auditory Brain-Stem Evoked Potentials," Electroencephalogr Clin Neurophysiol, 1986 Sep; 64(3):269-73.

[25]. E. Laukli and I.W.S. Mair, " Early Auditory-evoked Responses: Filter Effects," Audiology 20 (4): 300-312(1981).

[26]. P. Mitra, "Short Course III Neural Signal Processing: Quantitative Analysis of Neural Activity: Spectral Analysis for Neural Signals ," Washington, DC: Society for Neuroscience, 2008, pp. 3-11.

APPENDICES

A. Spectral Estimates

A.1. Spectrogram

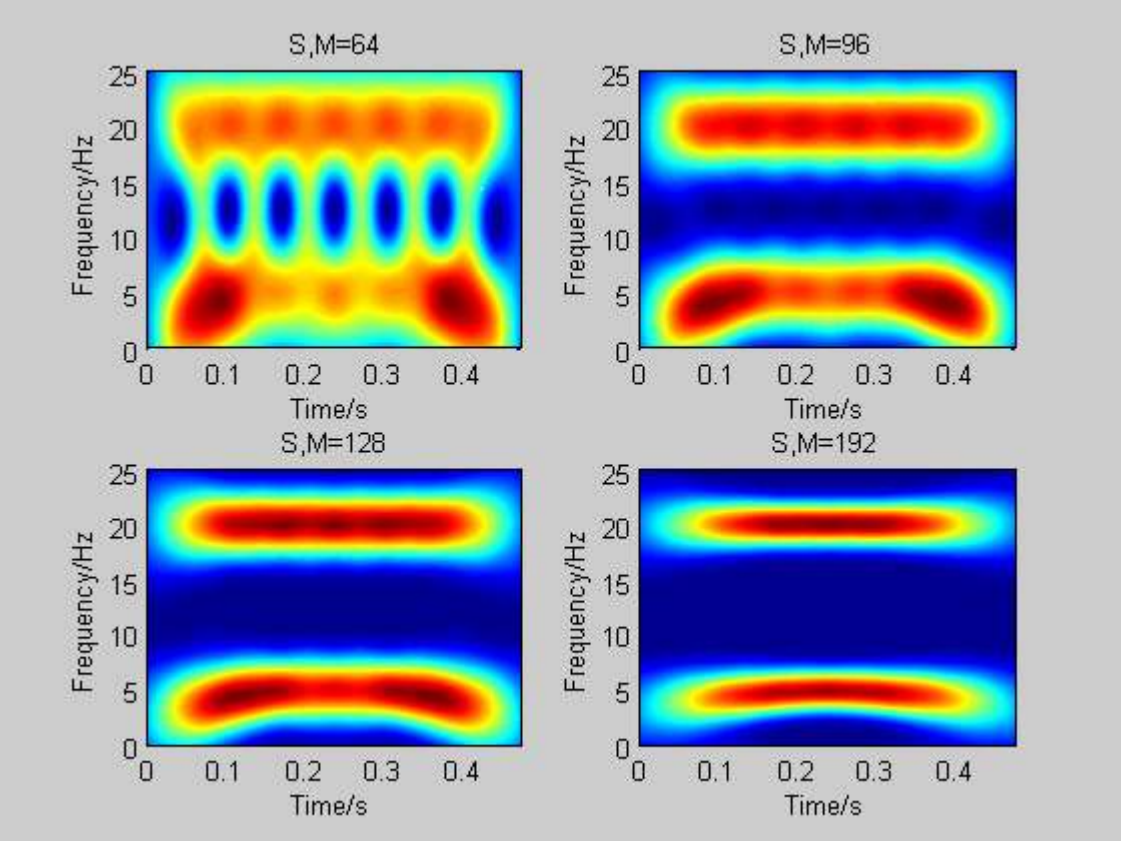


Figure 4.5 *Spectrograms of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the whole clicks per average.*

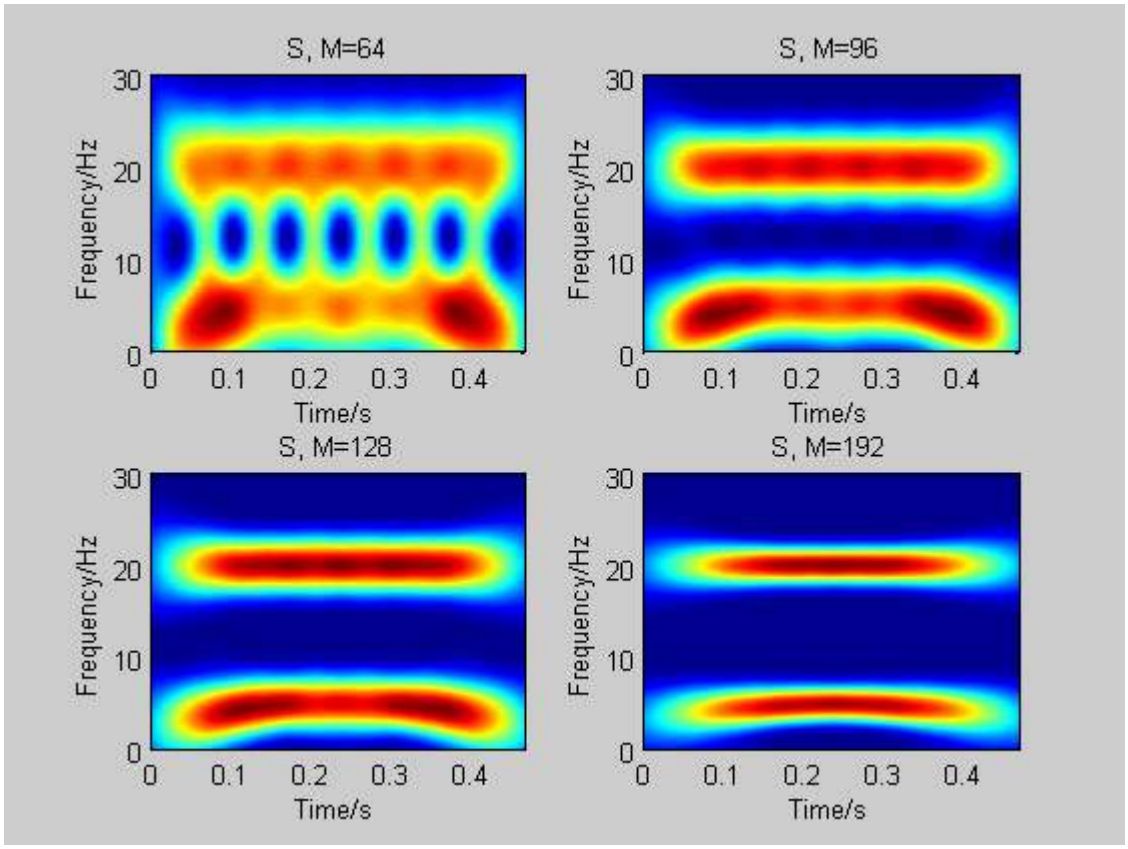


Figure 4.6 *Spectrograms of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the three hundred clicks per average .*

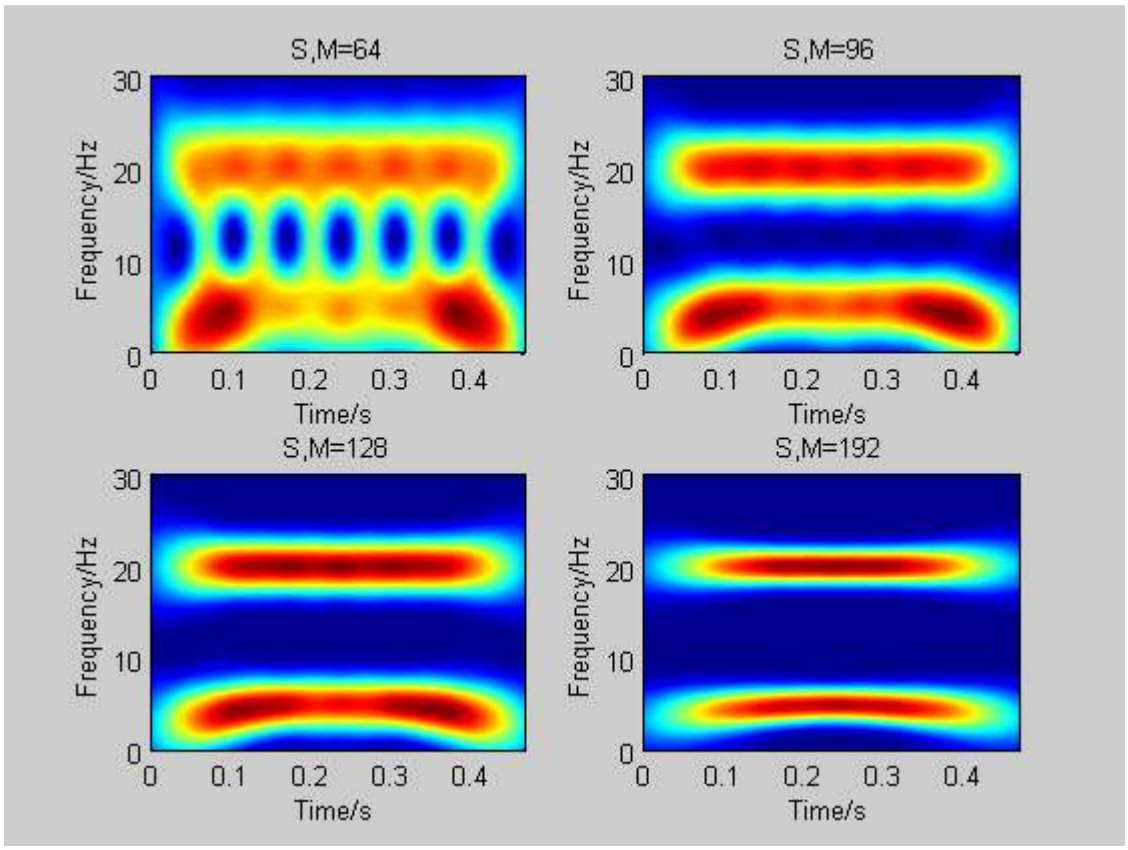


Figure 4.7 *Spectrograms of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the one hundred clicks per average.*

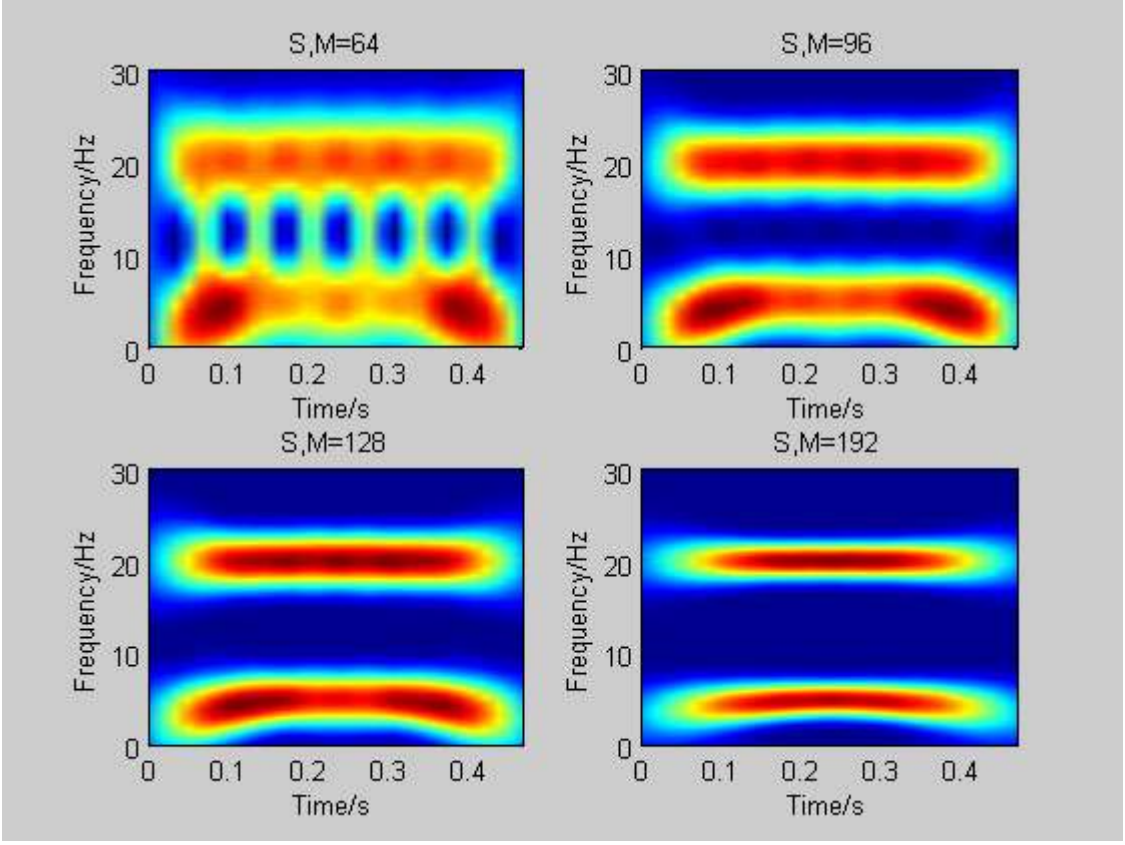


Figure 4.8 *Spectrograms of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) the fifty clicks per average.*

A.2. Thomson Multitaper

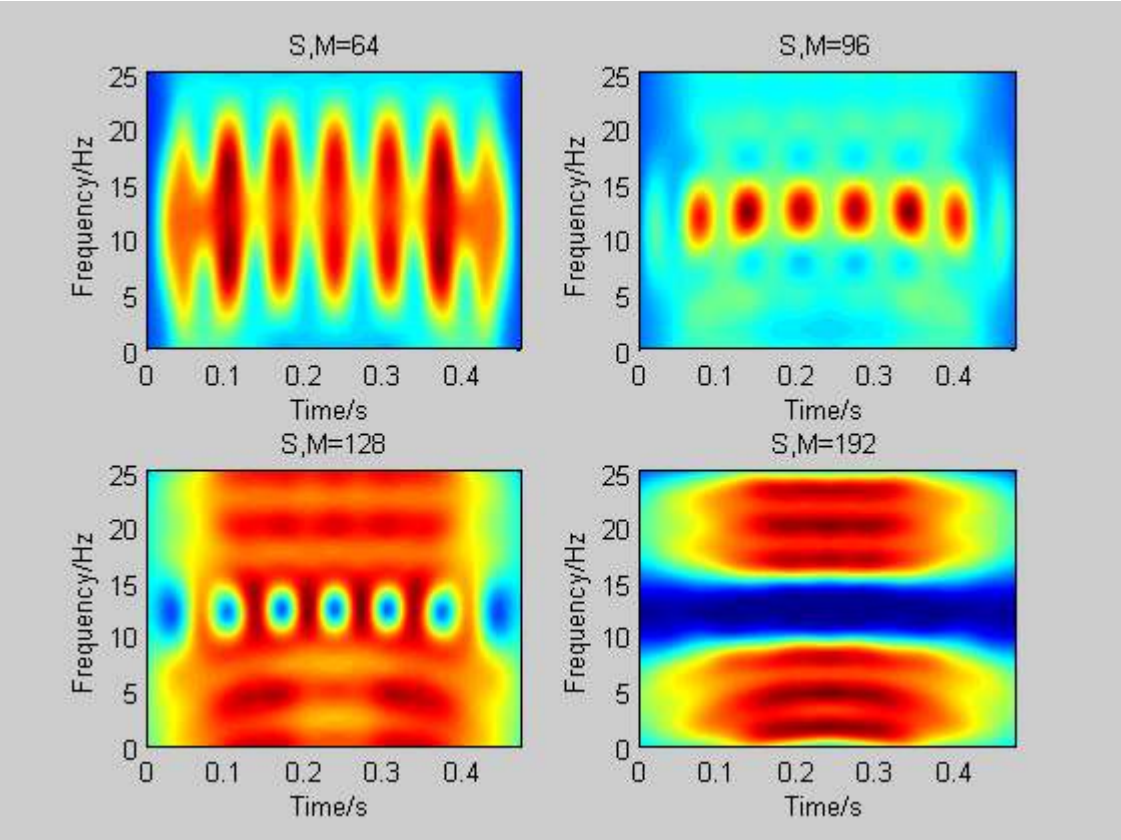


Figure 4.9 *Thomson Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the whole clicks per average.*

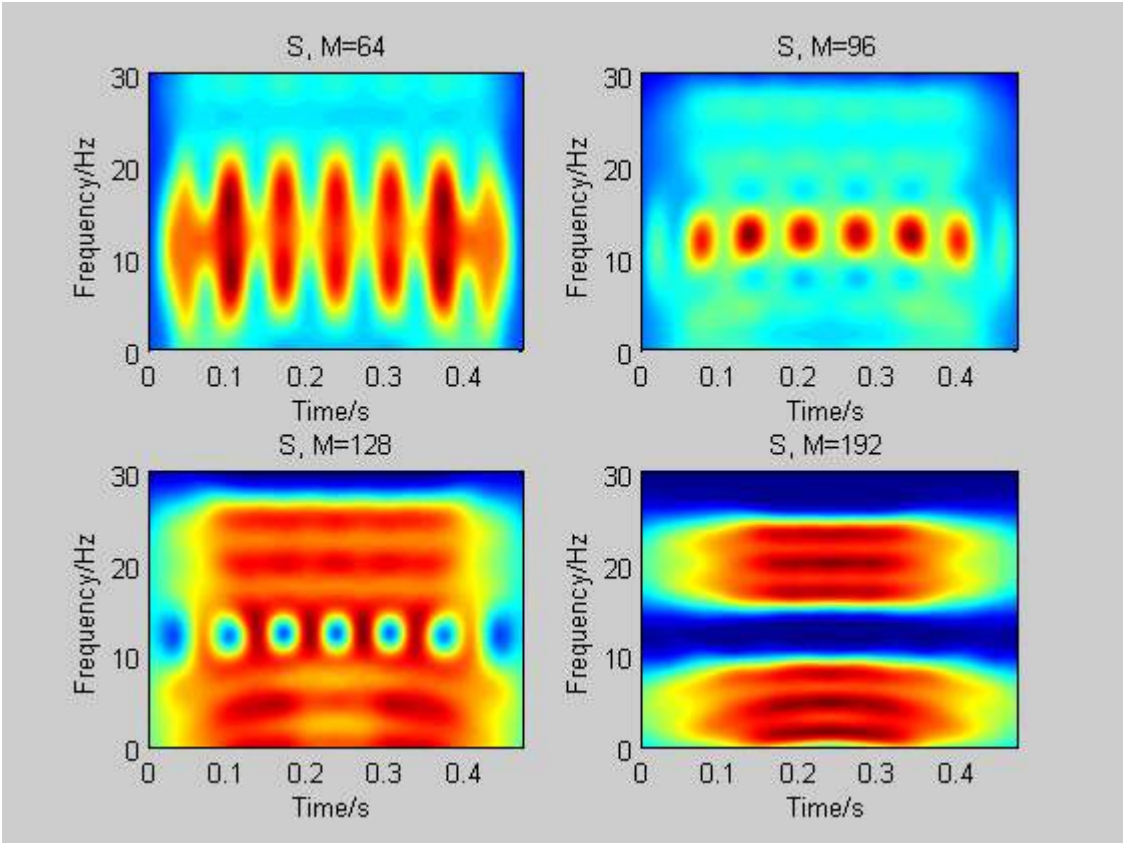


Figure 4.10 *Thomson Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the three hundred clicks per average.*

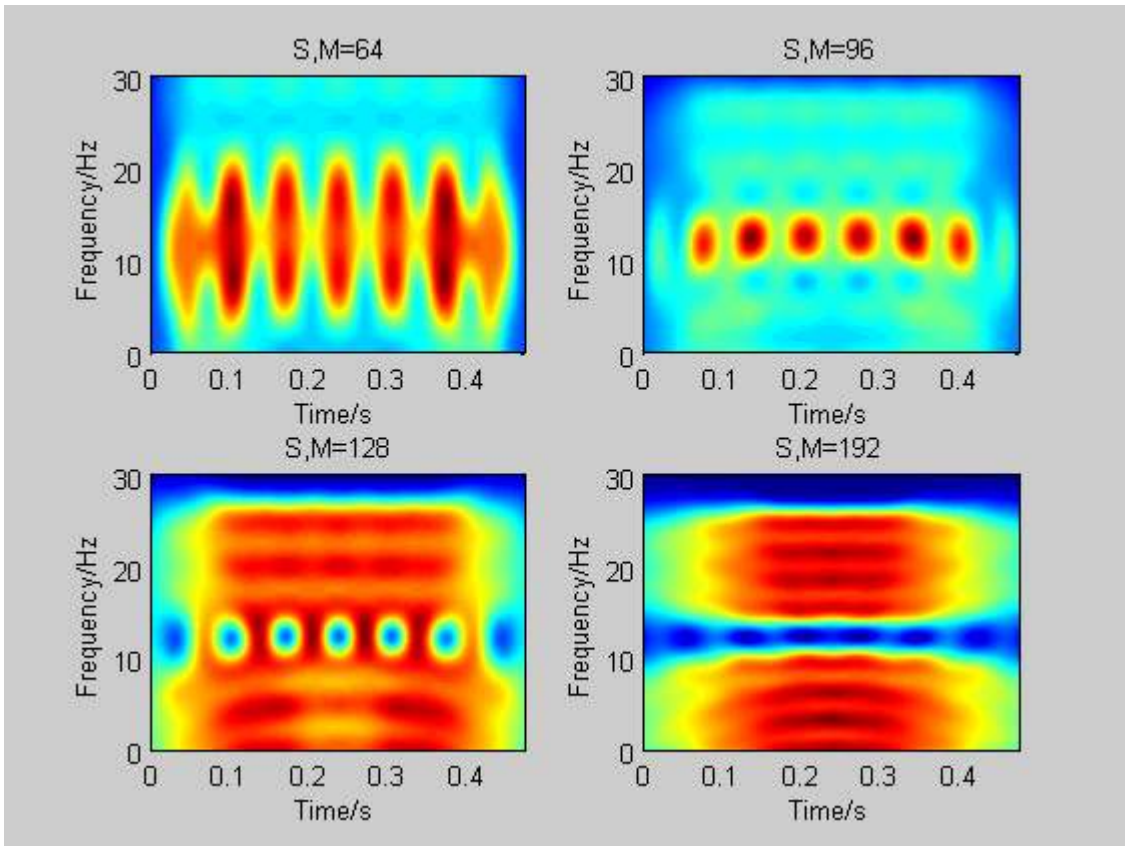


Figure 4.11 *Thomson Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the one hundred clicks per average.*

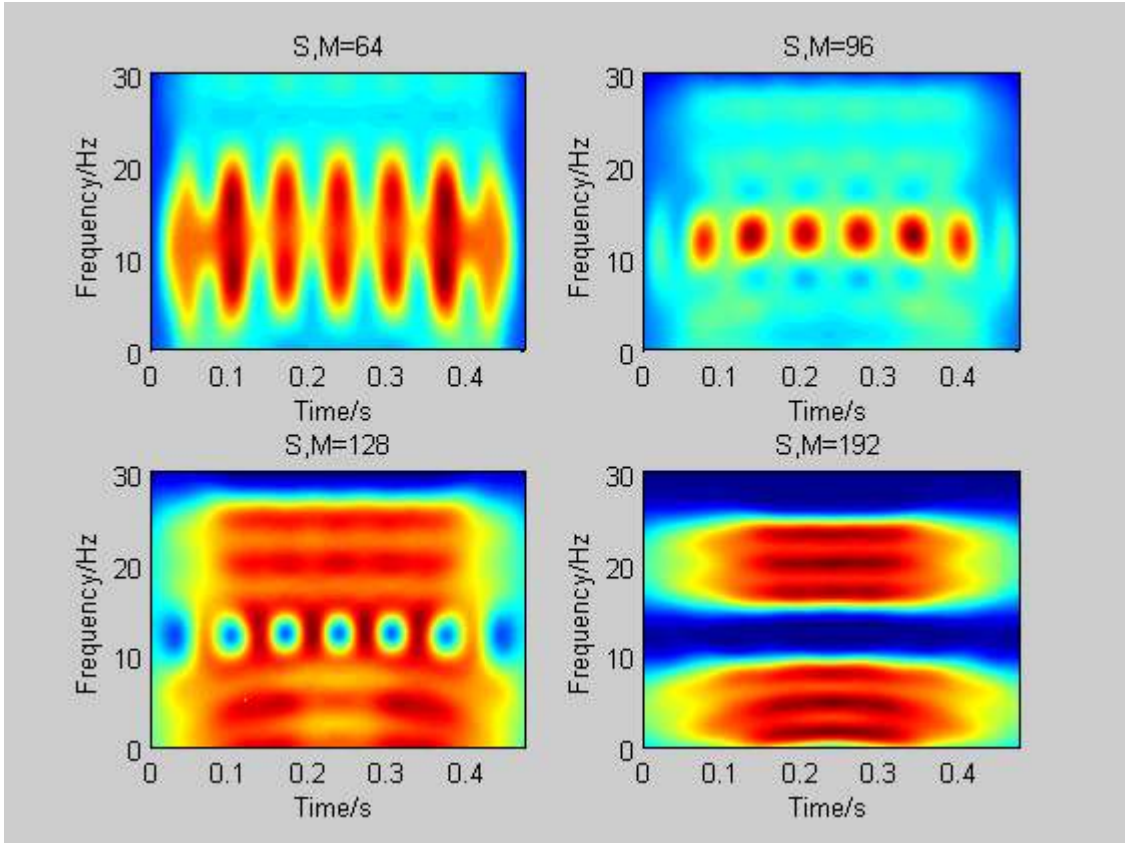


Figure 4.12 *Thomson Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the Fifty clicks per average.*

A.3. Peak Matched Multitapers

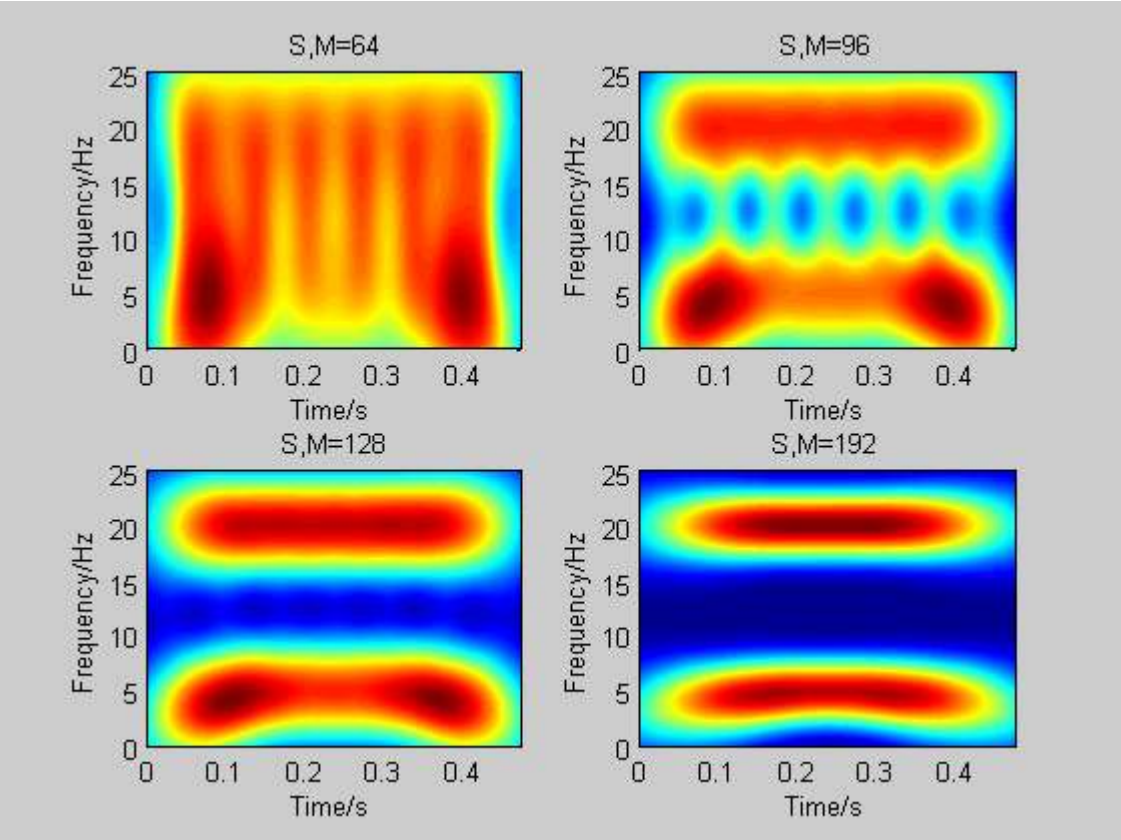


Figure 4.13 *Peak Matched Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the whole clicks per average.*

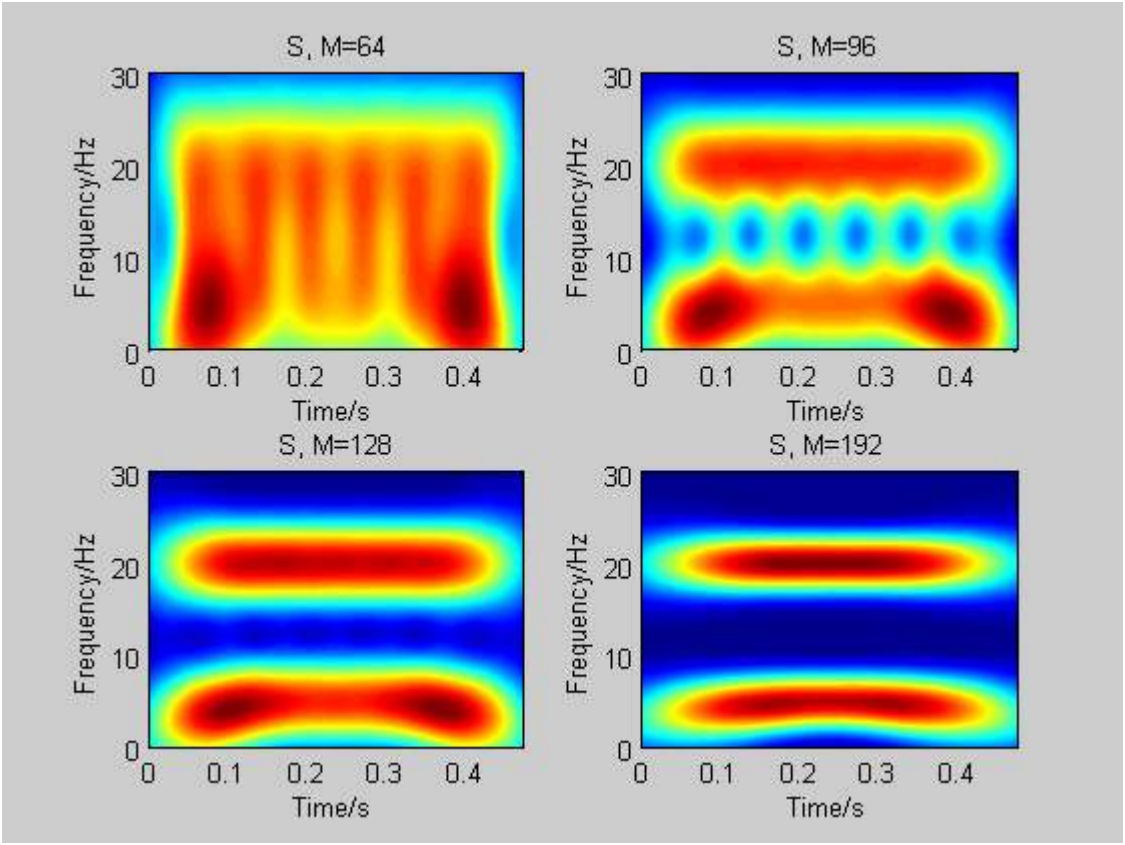


Figure 4.14 *Peak Matched Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the three Hundred clicks per average.*

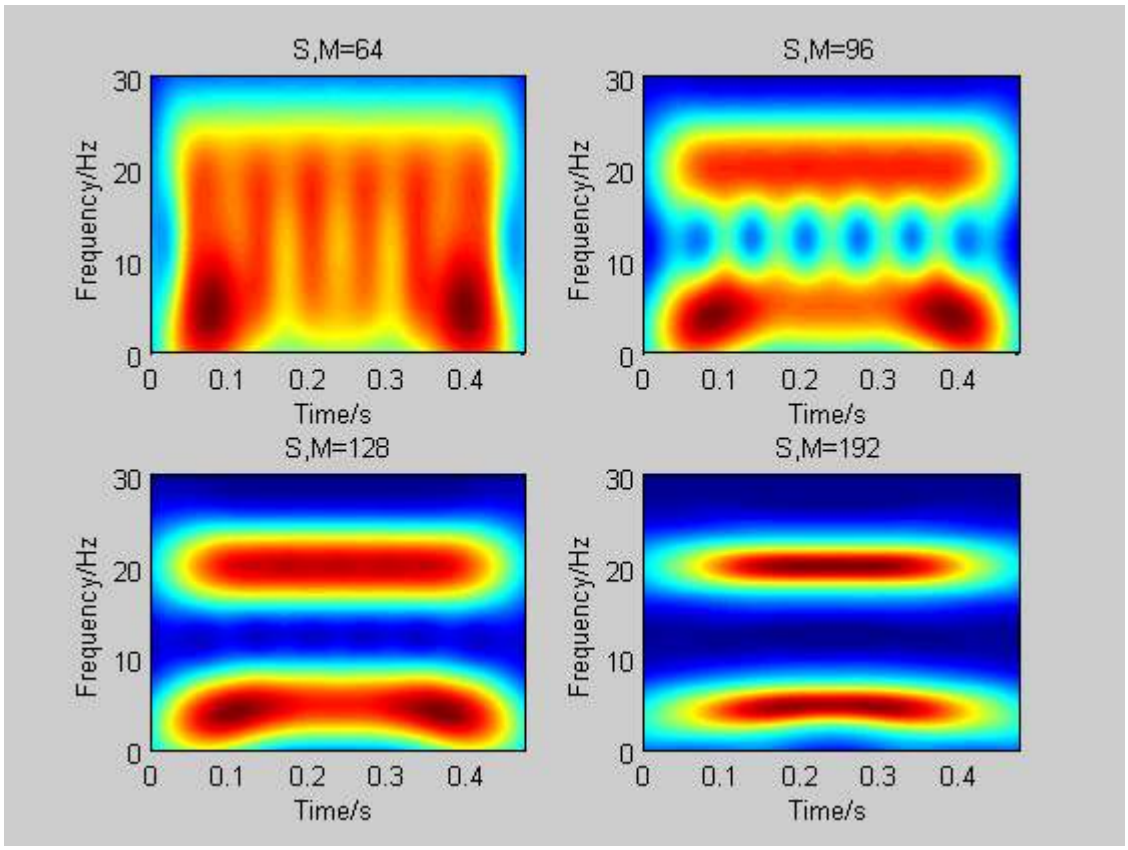


Figure 4.15 *Peak Matched Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the one Hundred clicks per average.*

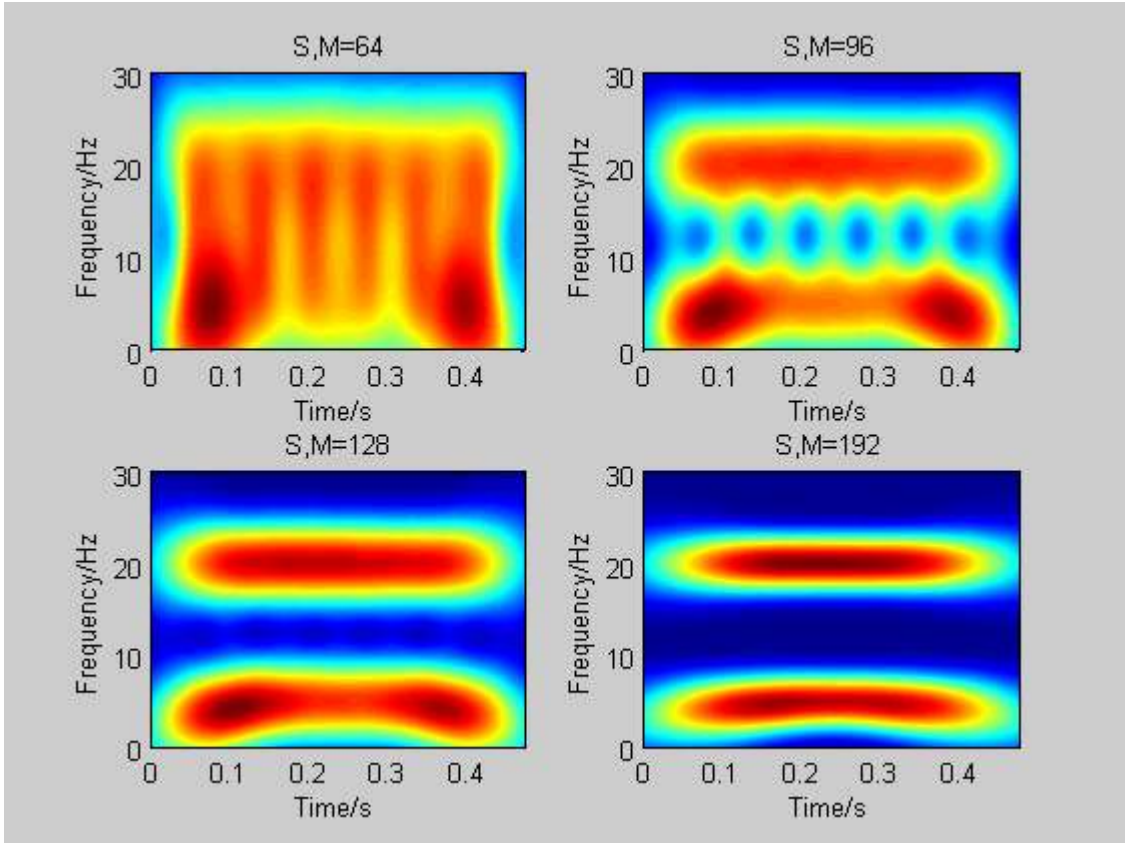
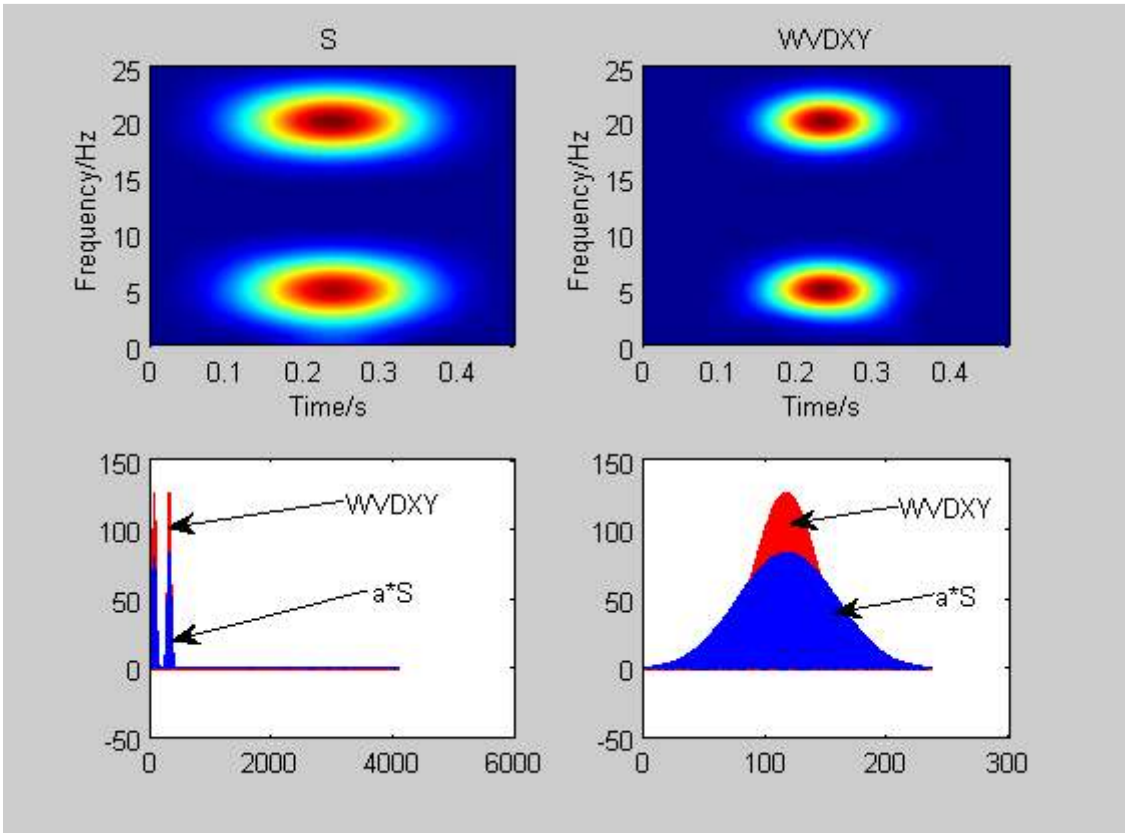


Figure 4.16 *Peak Matched Multitapers of different window lengths: M=64 (left-top), M=96 (right-top), M=128 (left-bottom) and M=192 (right-bottom) of the Fifty clicks per average.*

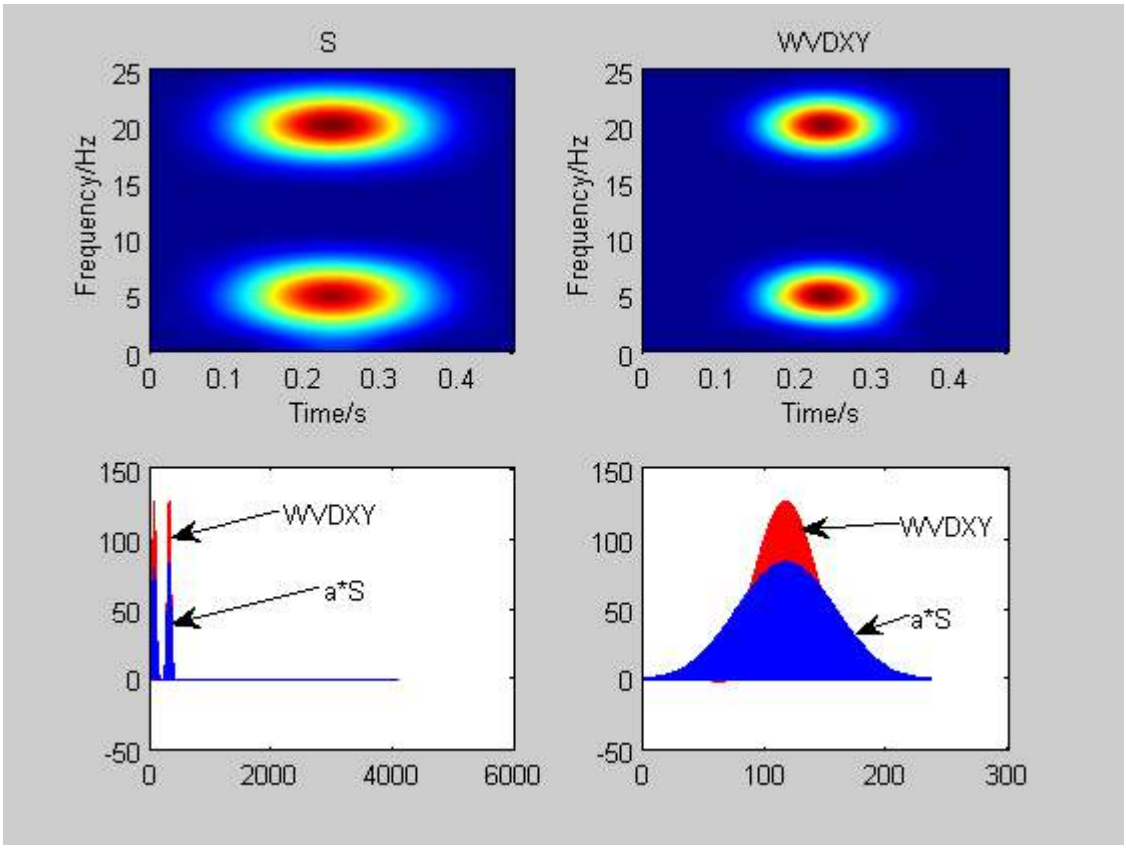
B. Measure of Goodness-of-fit

B.1. Visual Plots

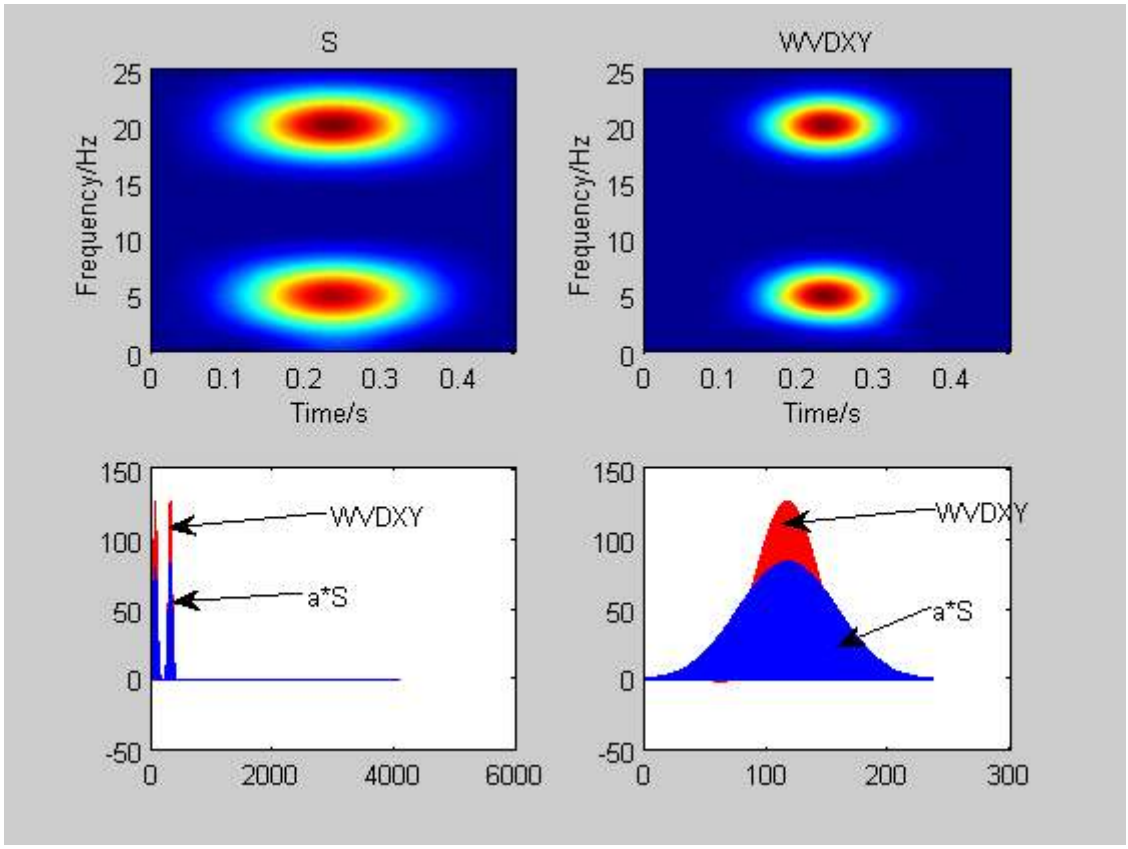
- **Spectrograms**



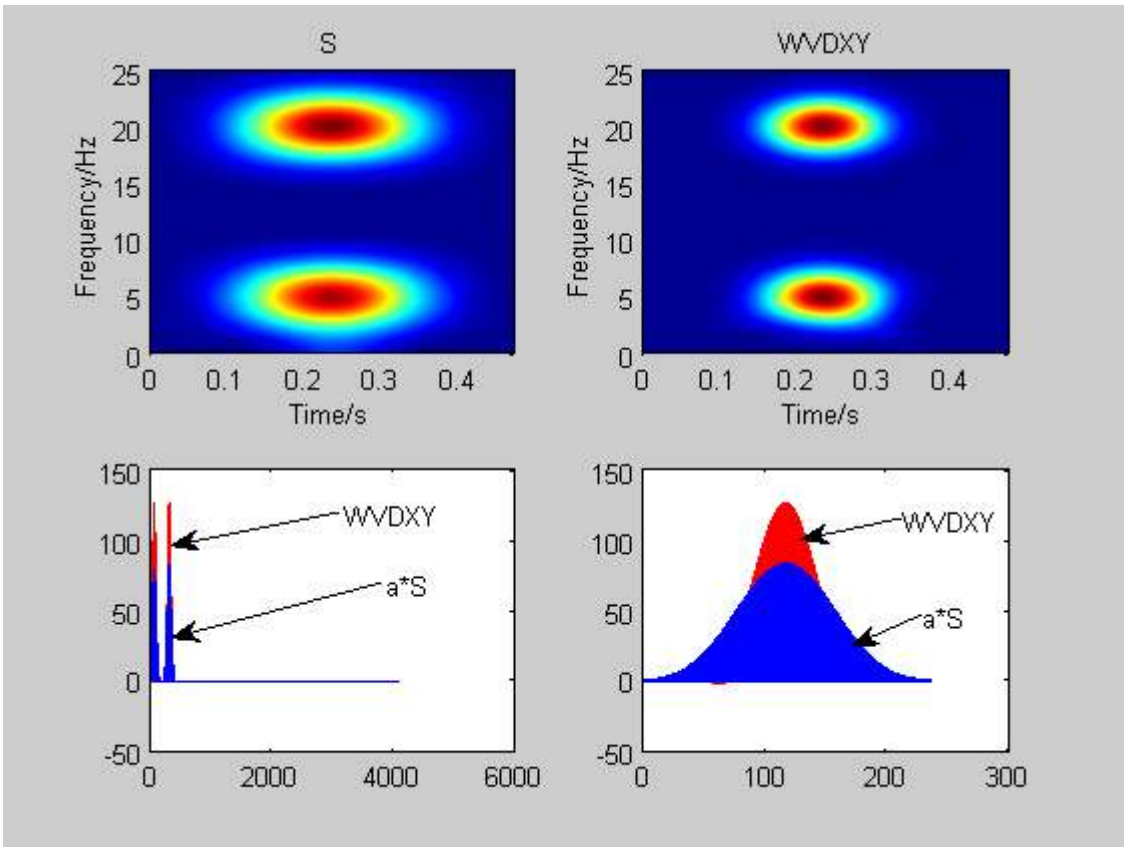
(a)



(b)



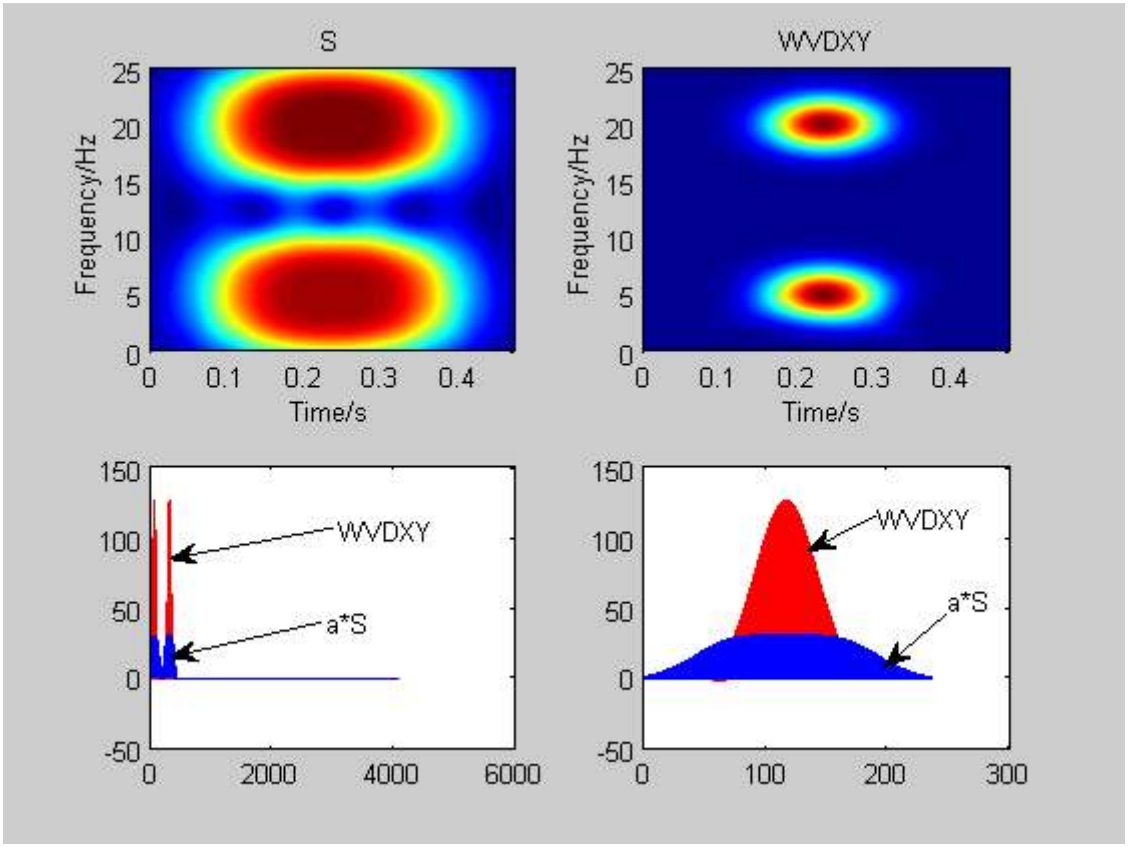
(c)



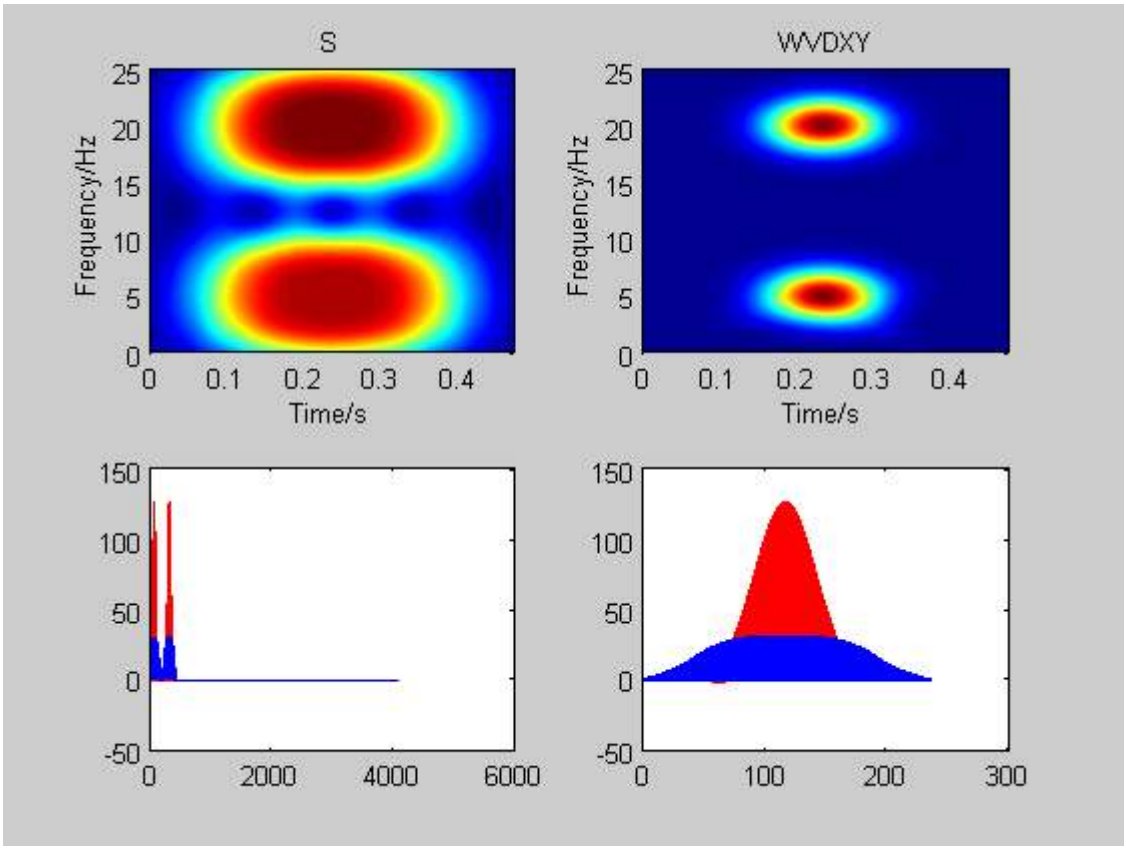
(d)

Figure 4.17 *Comparison between WVD and spectrogram in different forms of the average of (a) total data, (b) three hundreds, (c) hundreds and (d)fifties in Spectrograms.*

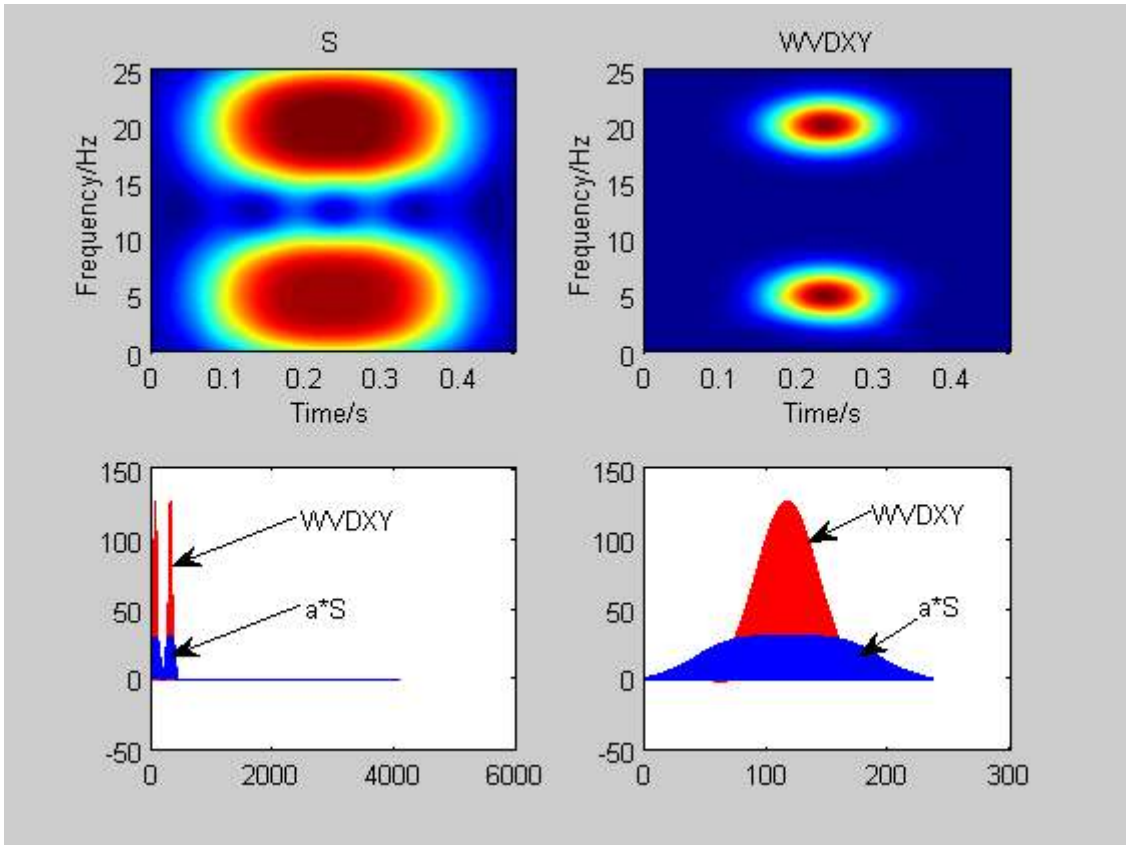
- **Thomson Multitapers**



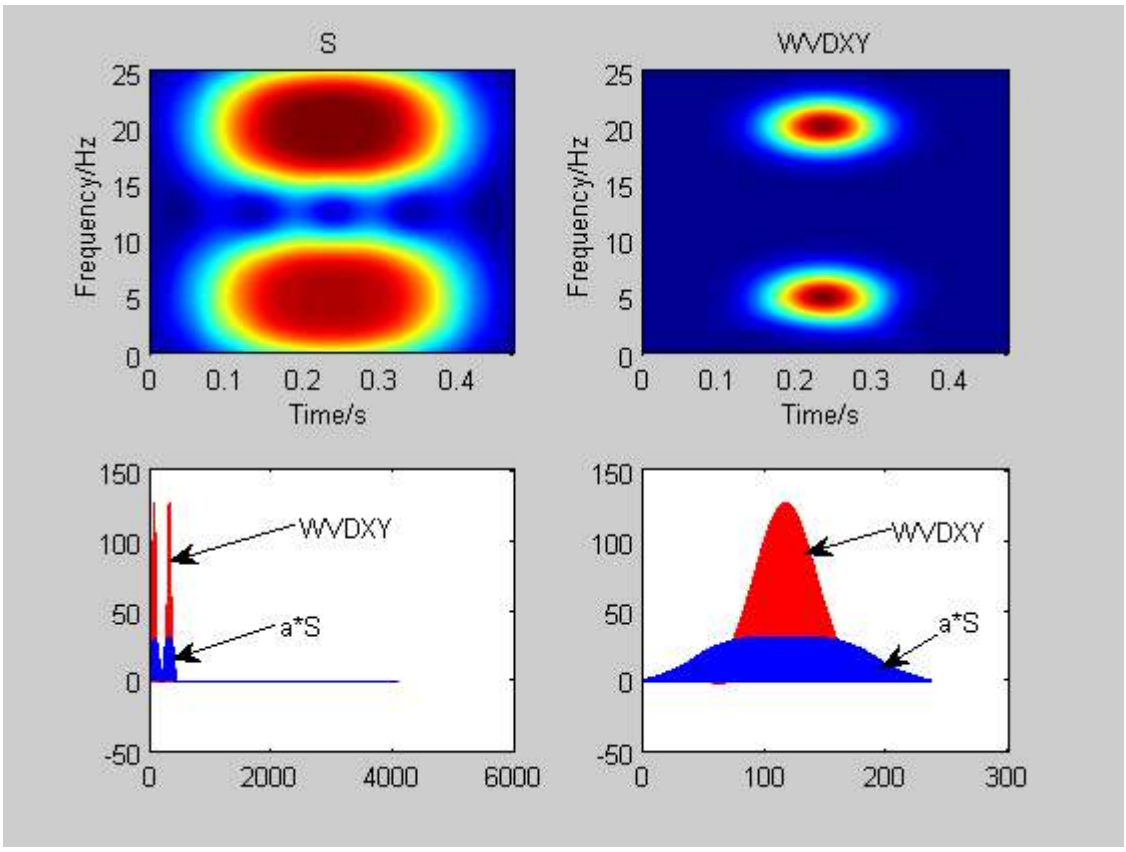
(a)



(b)



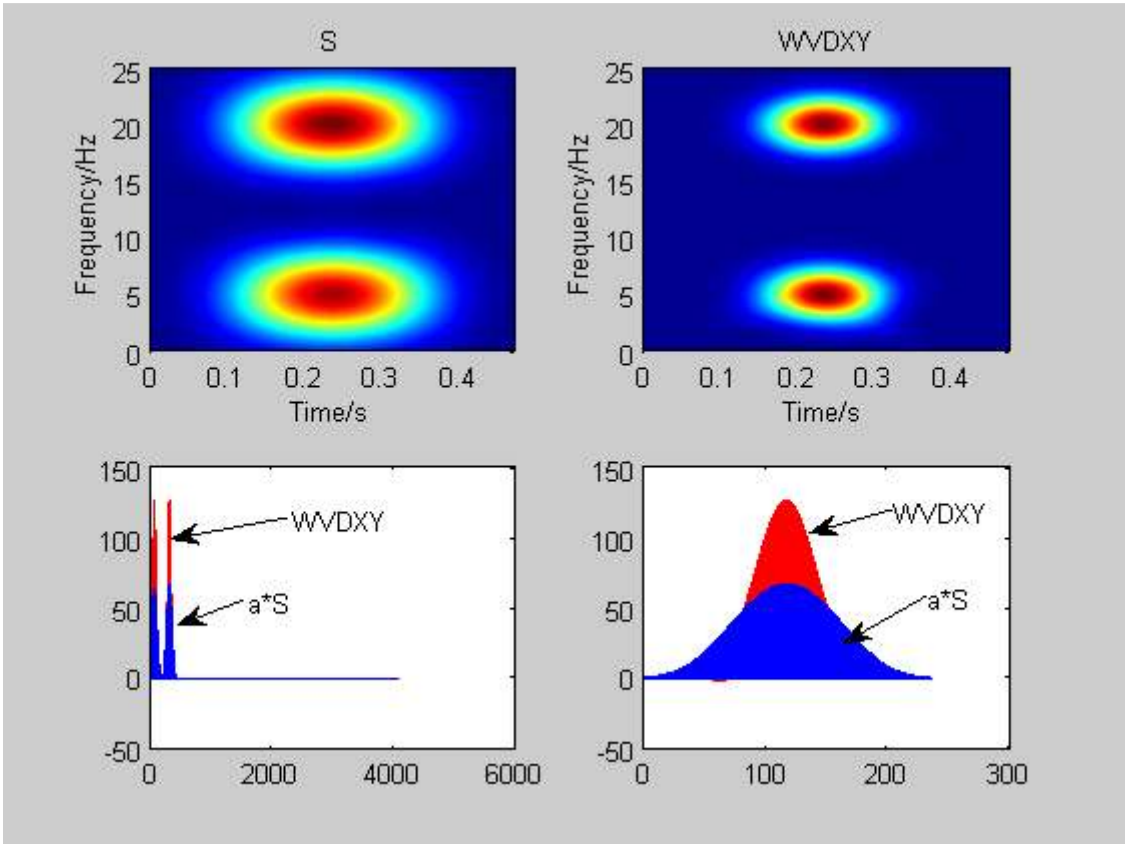
(c)



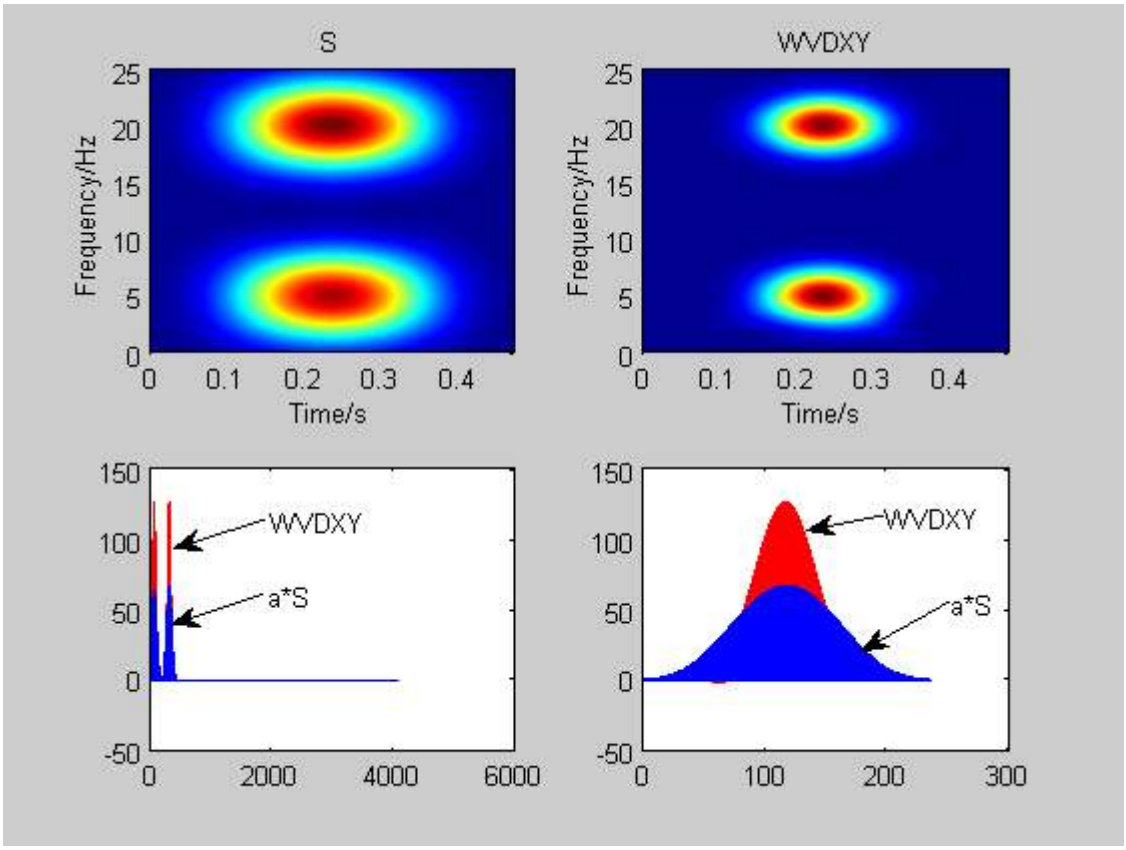
(d)

Figure 4.18 *Comparison between WVD and spectrogram in different forms of the average of (a) total data, (b) three hundreds, (c) hundreds and (d)fifties in Thomson Multitapers.*

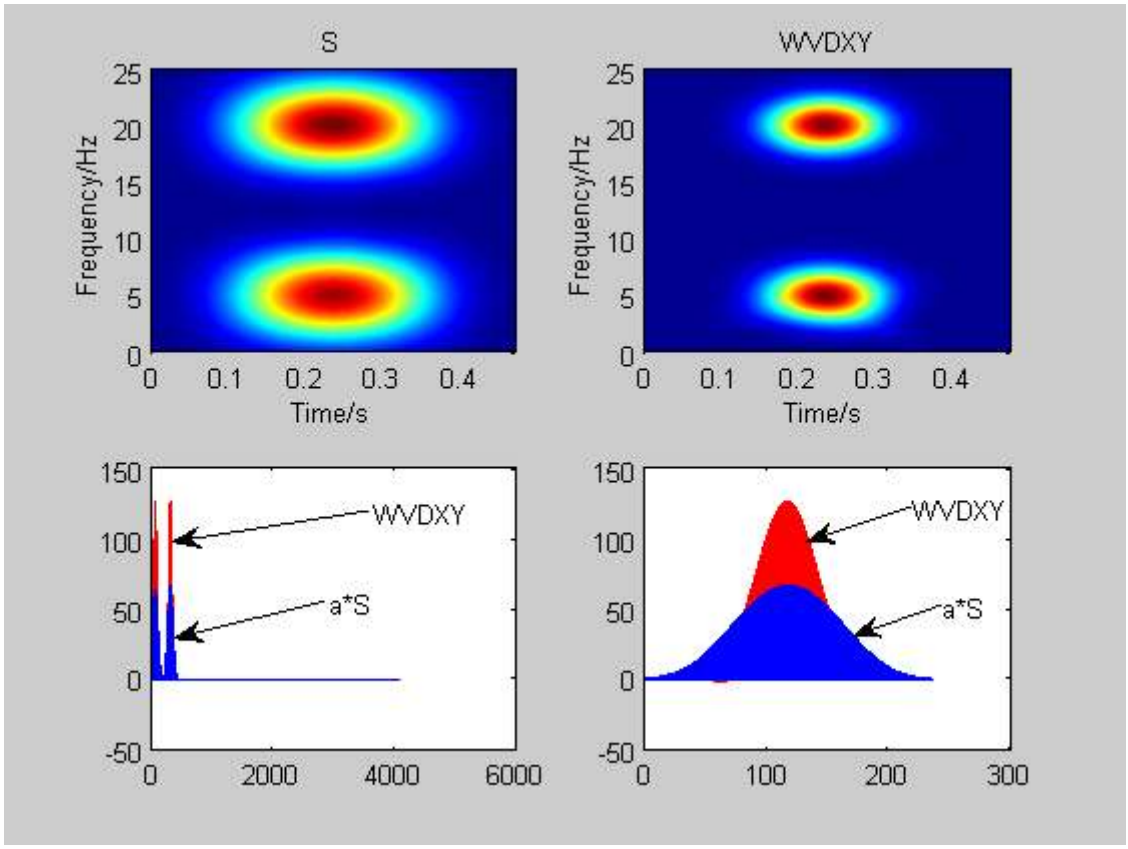
- **Peak Matched Multitapers**



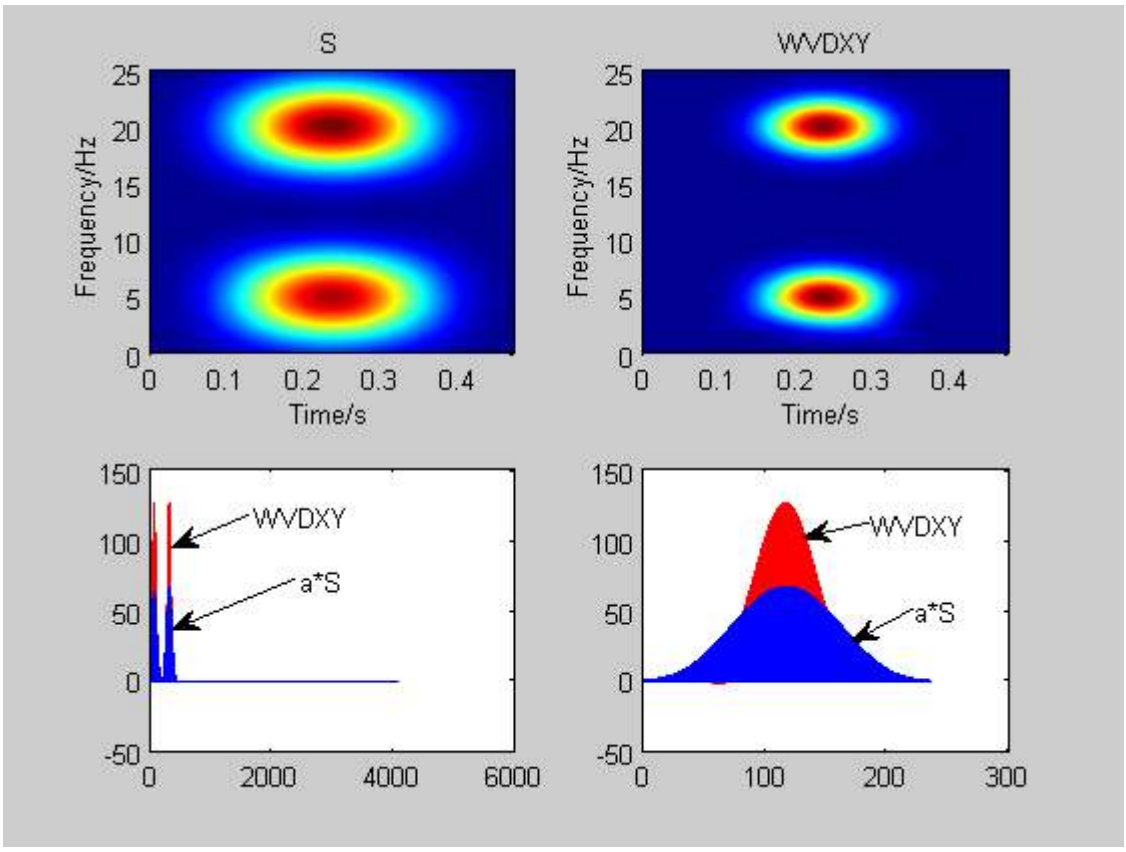
(a)



(b)



(c)



(d)

Figure 4.19 *Comparison between WVD and spectrogram in different forms of the average of (a) total data, (b) three hundreds, (c) hundreds and (d)fifties in Peak Matched Multitapers.*

C. Estimator Comparisons

C.1. RMSE

Table 4.1 *Different RMSE values of Spectrogram for different number of clicks per average.*

σ	clicks	1313	300	100	50	mean
0		9,4136	9,4141	9,4133	9,3834	**9,41**
0,5		9,4013	9,4229	9,4178	9,3958	**9,41**
1		9,404	9,4285	9,465	9,4214	**9,43**
2		9,4368	9,4087	9,4444	9,6257	**9,48**
3		9,4687	9,4851	9,7225	10,145	**9,71**
4		9,4959	9,5913	10,1254	11,1283	**10,09**
5		9,505	9,7826	10,6582	12,4978	**10,61**
6		9,6531	9,9809	11,6401	14,8772	**11,54**
7		9,691	10,1774	12,4554	17,3322	**12,41**
8		9,589	10,5874	14,3607	19,5042	**13,51**
9		10,4305	11,1626	14,3828	23,4474	**14,86**
10		9,9689	11,3172	17,3924	26,1213	**16,20**
11		10,6525	11,7476	19,0002	24,5224	**16,48**
12		10,295	12,9462	21,6981	27,923	**18,22**
15		10,3399	15,53	29,6501	36,5119	**23,01**
30		15,6375	37,5103	39,4501	45,4913	**34,52**
40		19,6915	47,9764	45,234	48.9517	**37,63**
	mean	**10,15**	**12,34**	**15,52**	**19,21**	**14,30**

Table 4.2 *Different RMSE values of Thomson Multitaper for different number of clicks per average.*

σ	clicks	1313	300	100	50	mean
0		19,0905	19,0903	19,09	19,0905	**19,09**
0,5		19,0904	19,0918	19,0952	19,0963	**19,09**
1		19,0968	19,0974	19,1013	19,1062	**19,10**
2		19,0912	19,1222	19,1279	19,1939	**19,13**
3		19,0721	19,1216	19,2083	19,4602	**19,22**
4		19,0992	19,1302	19,3297	19,818	**19,34**
5		19,2367	19,2146	19,6067	20,6988	**19,69**
6		19,1195	19,2367	19,978	21,4417	**19,94**
7		19,1225	19,3627	20,5898	23,0456	**20,53**
8		19,1882	19,505	21,2185	24,7699	**21,17**
9		19,1551	19,6897	22,2892	25,4293	**21,64**
10		19,2023	19,8803	23,449	27,64	**22,54**
11		19,2652	20,1729	24,1109	29,9563	**23,38**
12		19,6959	20,6015	25,4211	31,4519	**24,29**
15		19,583	22,7697	29,3849	36,2598	**27,00**
30		23,0915	30,7269	38,3026	45,7474	**34,47**
40		26,9587	37,074	41,7168	42,9826	**37,18**
	mean	**19,45**	**20,36**	**22,46**	**25,14**	**21,85**

Table 4.3 *Different RMSE values of Peak Matched Multitaper for different number of clicks per average.*

σ	clicks	1313	300	100	50	mean
0		12,3052	12,3059	12,3064	12,3053	**12,31**
0,5		12,3091	12,3027	12,3165	12,3088	**12,31**
1		12,2997	12,3156	12,3392	12,3259	**12,32**
2		12,3079	12,3331	12,4527	12,5208	**12,40**
3		12,3234	12,386	12,5652	12,9251	**12,55**
4		12,3244	12,4194	12,8628	14,0022	**12,90**
5		12,3965	12,4247	13,3352	15,063	**13,30**
6		12,4572	12,3733	14,1292	16,8701	**13,96**
7		12,3703	12,8747	15,1721	18,5474	**14,74**
8		12,5555	13,1737	15,5975	21,9181	**15,81**
9		13,2572	13,6331	17,5504	24,5578	**17,25**
10		13,0032	14,1952	18,3457	25,0516	**17,65**
11		12,8281	14,2923	20,7561	28,7086	**19,15**
12		12,8294	15,6783	21,1231	30,8818	**20,13**
15		13,3143	15,6242	29,4521	35,786	**23,54**
30		19,7392	33,2098	44,3804	40,2838	**34,40**
40		25,6733	37,9338	41,1894	40,2633	**36,26**
	mean	**13,04**	**14,47**	**17,79**	**20,88**	**16,55**

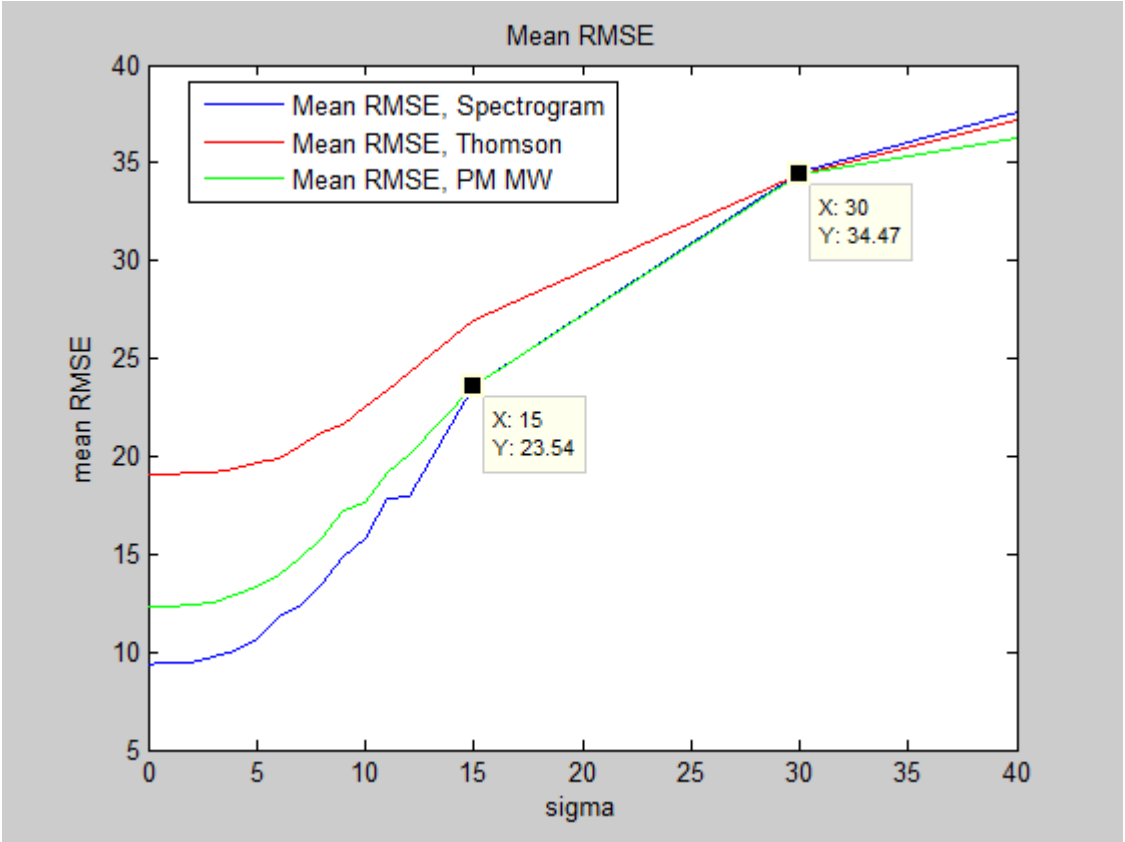
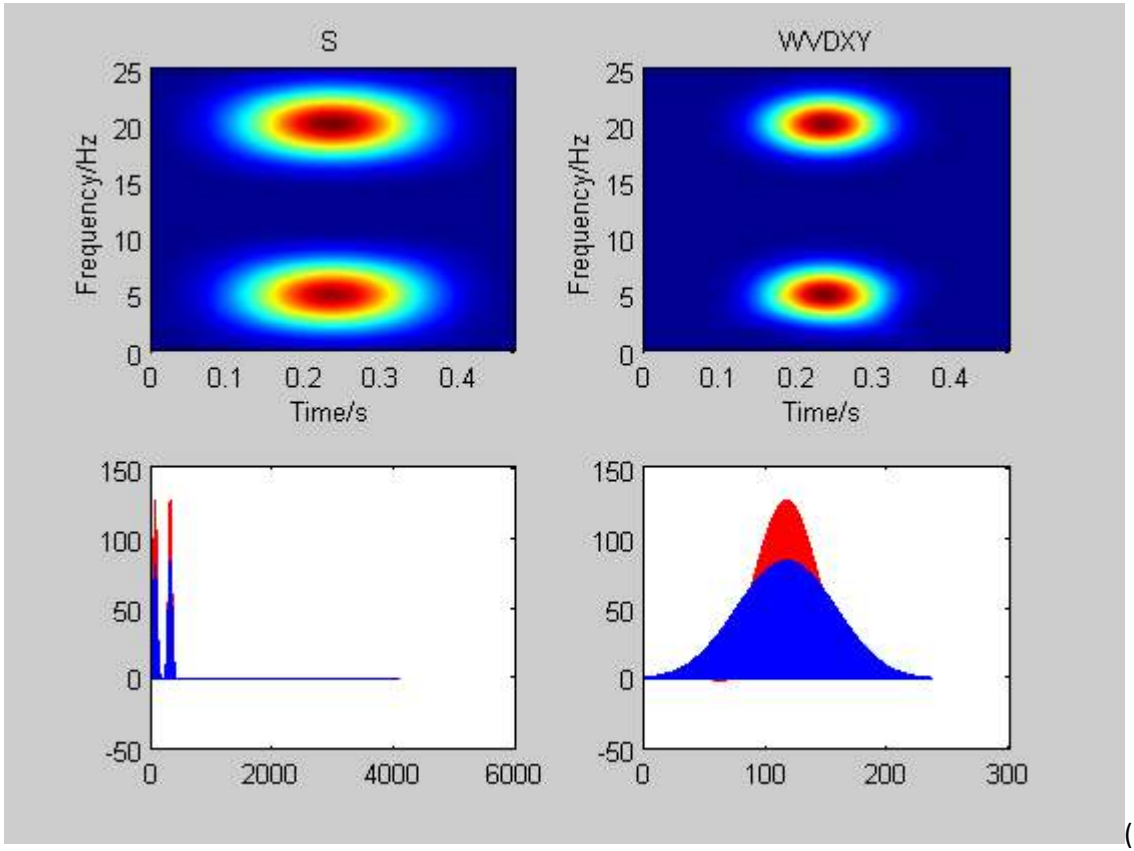
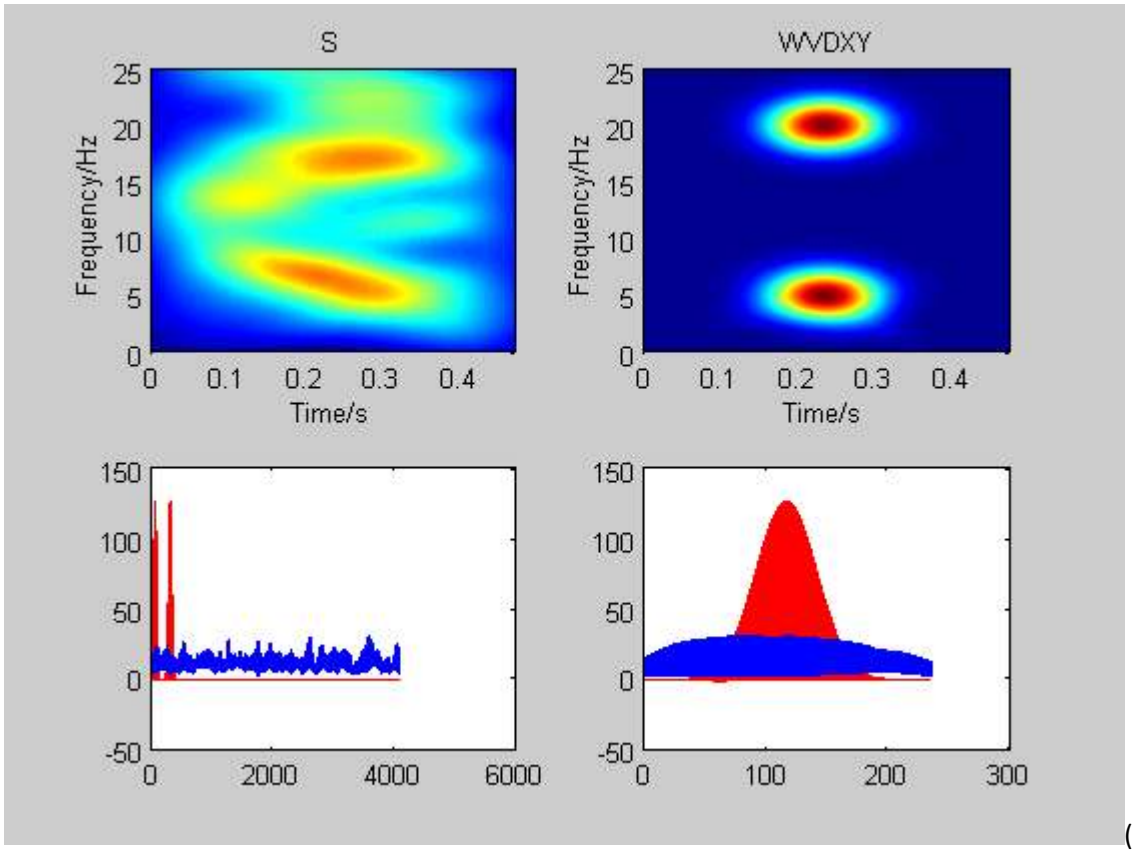


Figure 4.20 *vectors of mean RMSE in Spectrogram (blue), PM MW (green) and Thomson Multitaper (red), respectively*

C.2. Noise Standard Deviation



(b)



(b)

Figure 4.21 *Comparison between WVD and spectrogram in 50 clicks per average of (a)* $\sigma = 0.5$*, (b)* $\sigma = 40$.

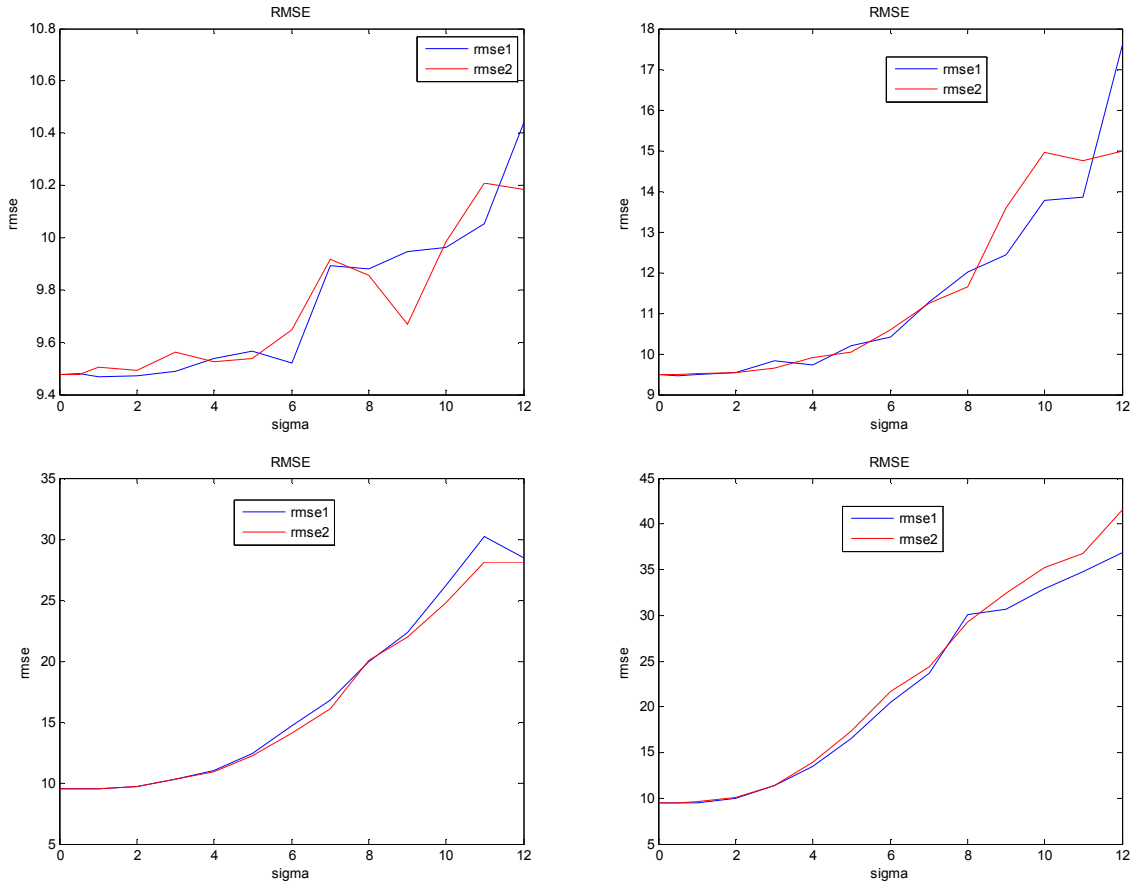


Figure 4.22 *Two simulated vectors of RMSE in Spectrogram from averages of total data (Left-top), three hundreds (Right-top), hundreds(Left-bottom) and fifties(Right-bottom, respectively.*

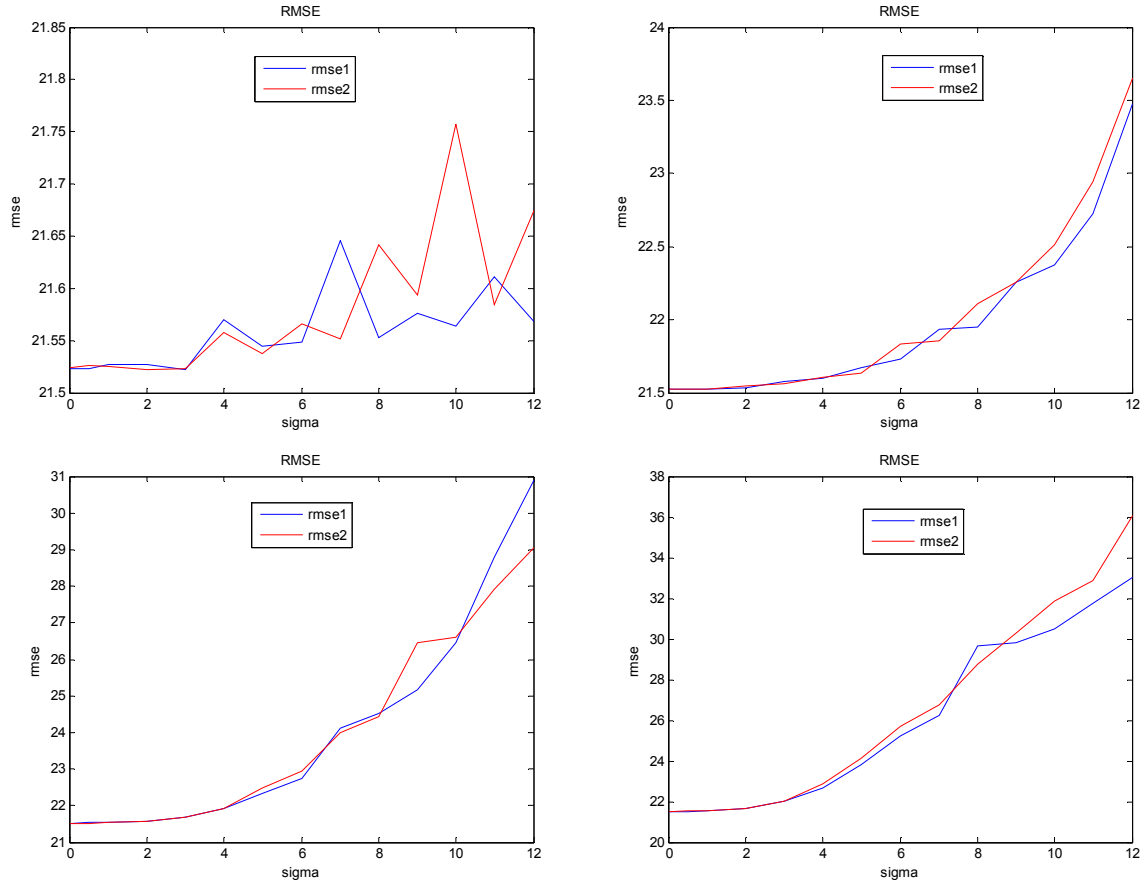


Figure 4.23 *Two simulated vectors of RMSE in Thomson Multitapers from averages of total data (Left-top), three hundreds (Right-top), hundreds(Left-bottom) and fifties(Right-bottom, respectively.*

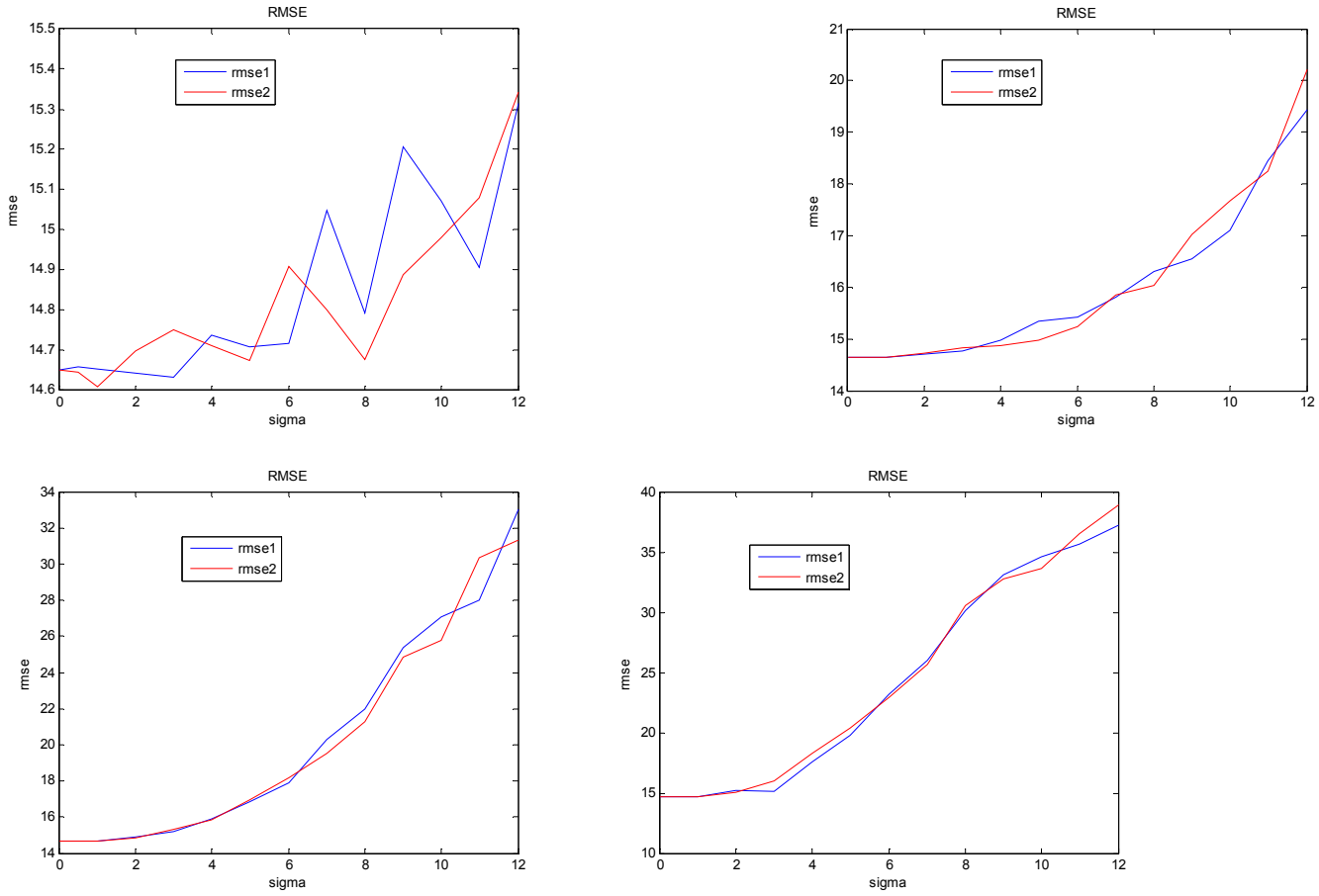
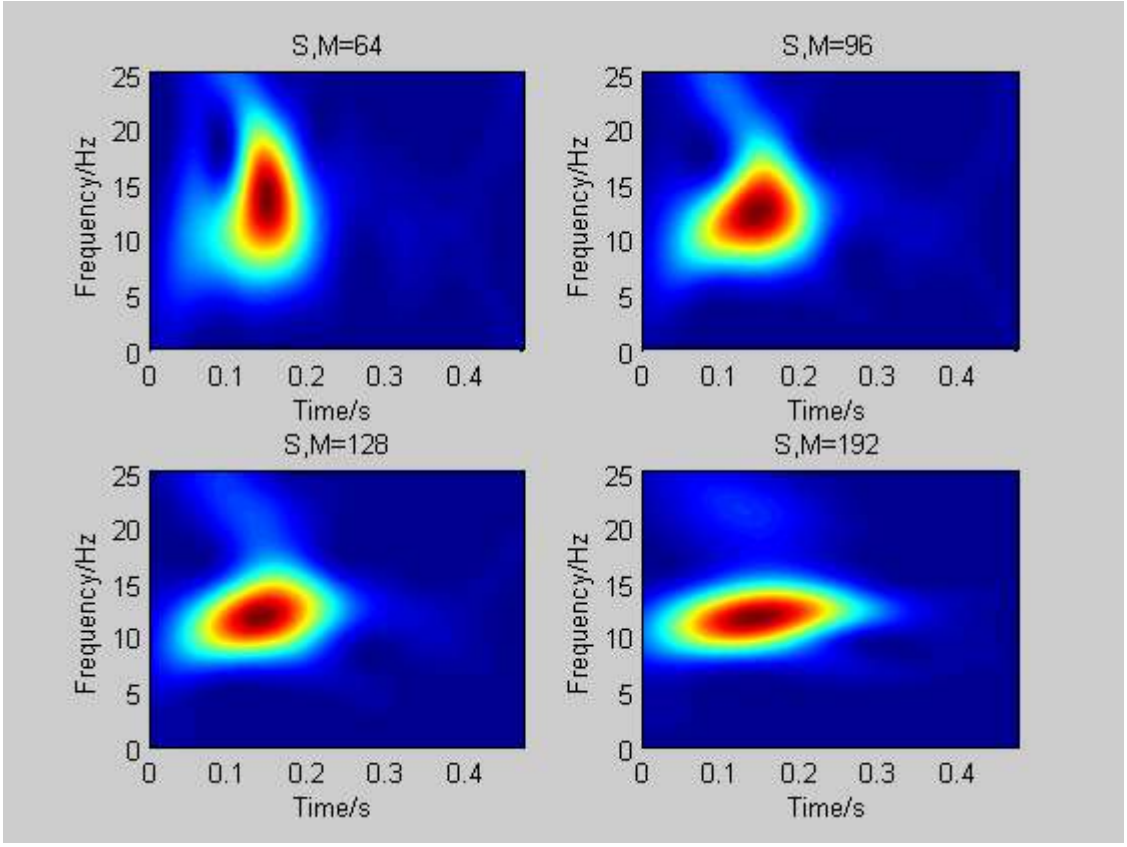
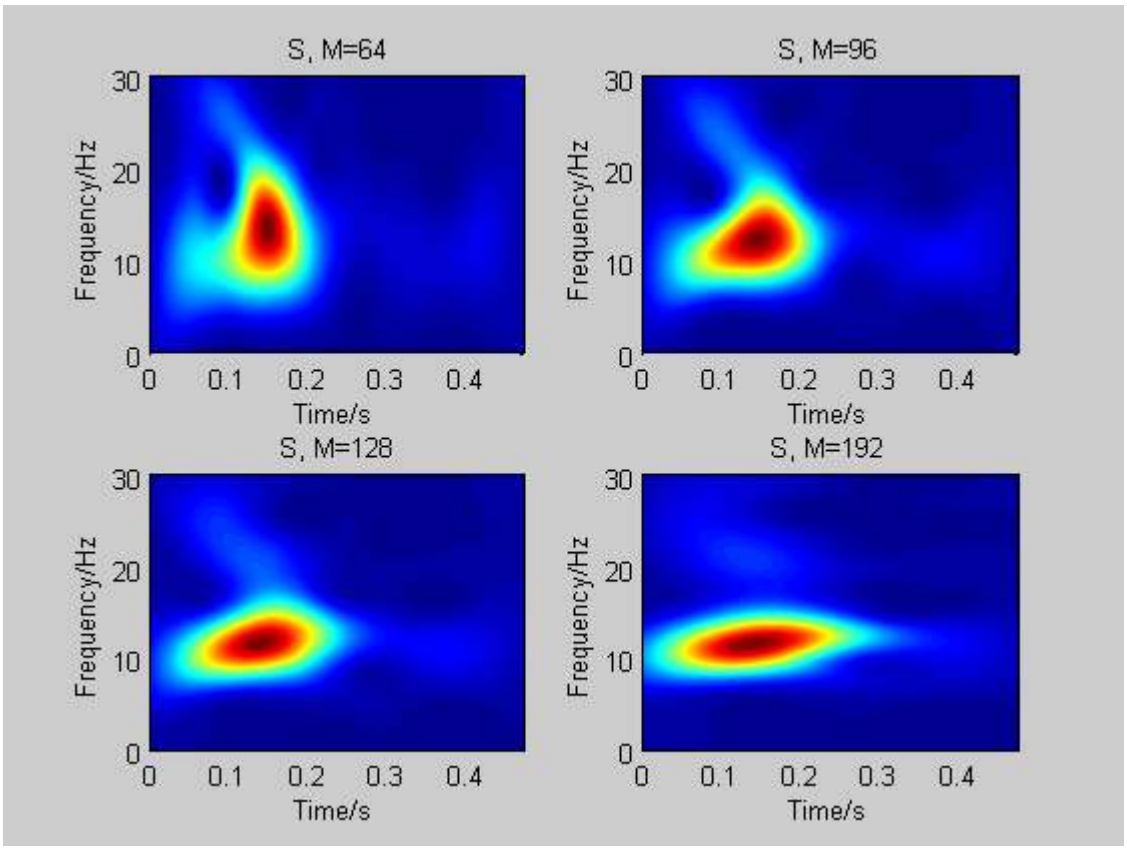


Figure 4.24 *Two simulated vectors of RMSE in Peak Matched Multitapers from averages of total data (Left-top), three hundreds (Right-top), hundreds(Left-bottom) and fifties(Right-bottom, respectively.*

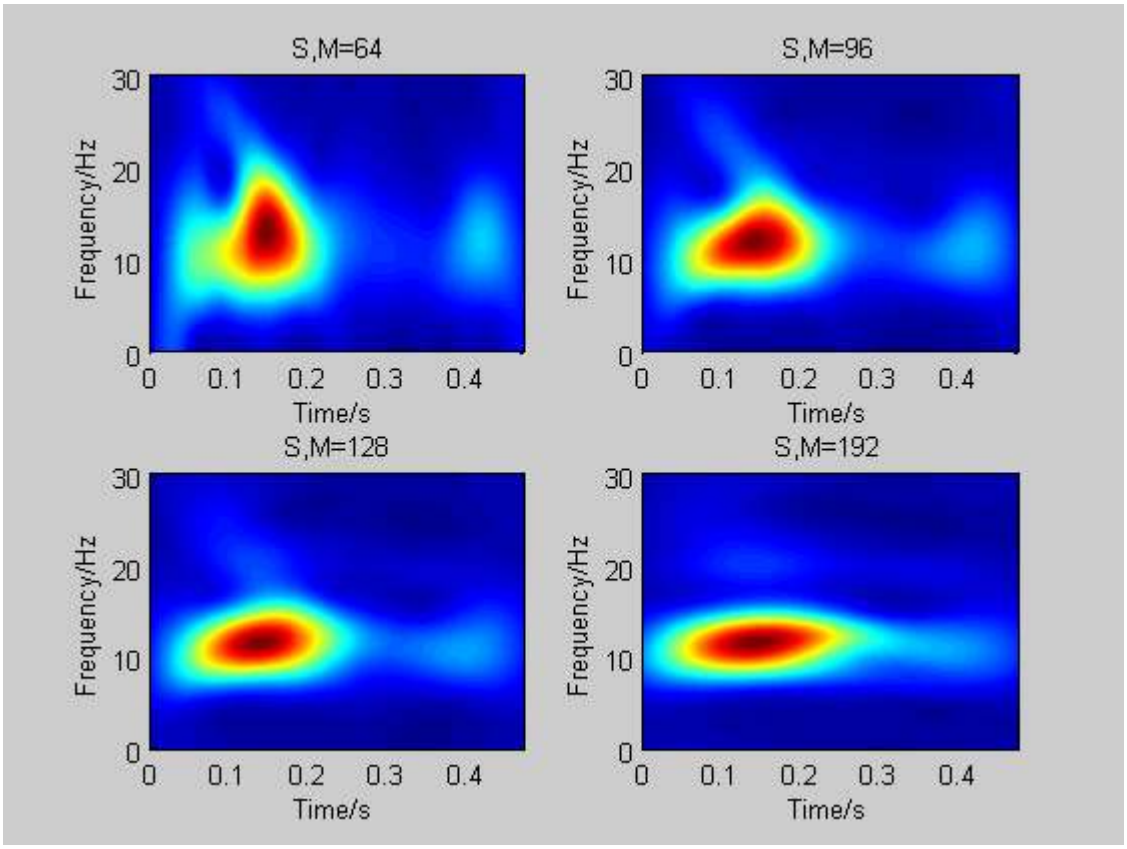
D. Application to Real Data



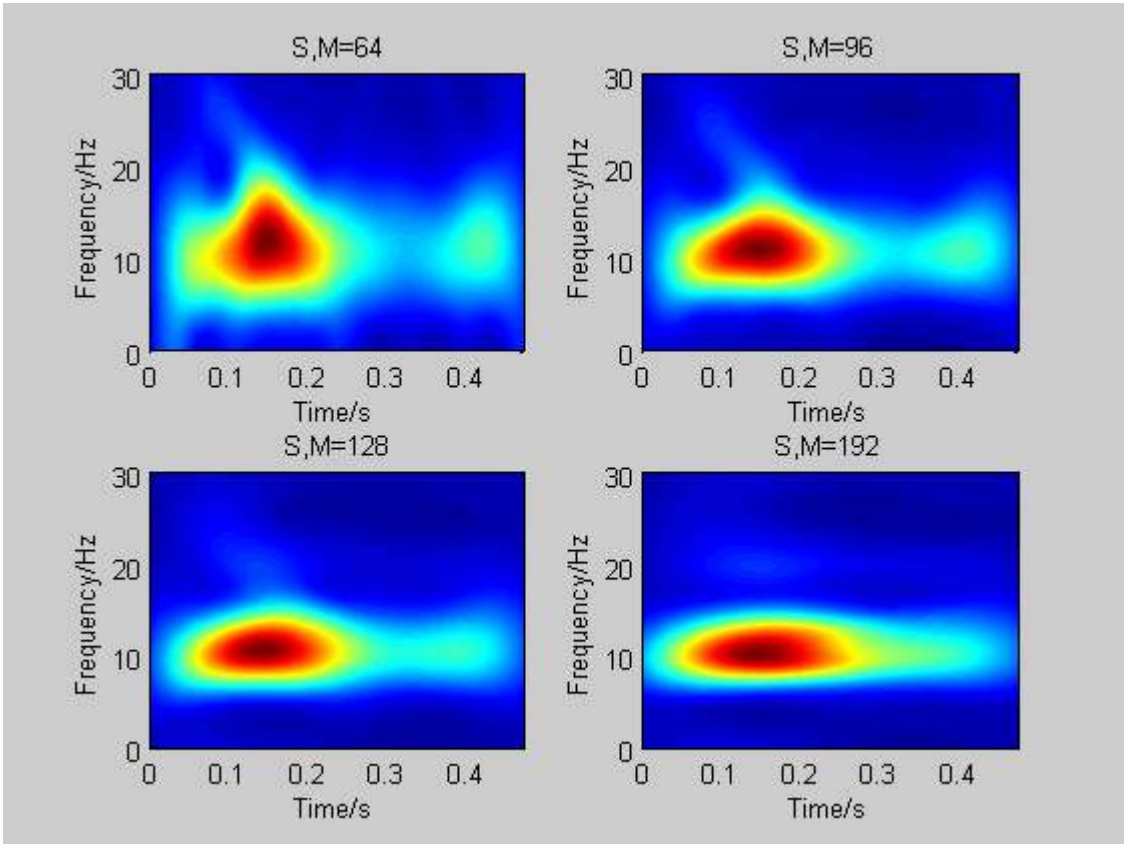
(a)



(b)

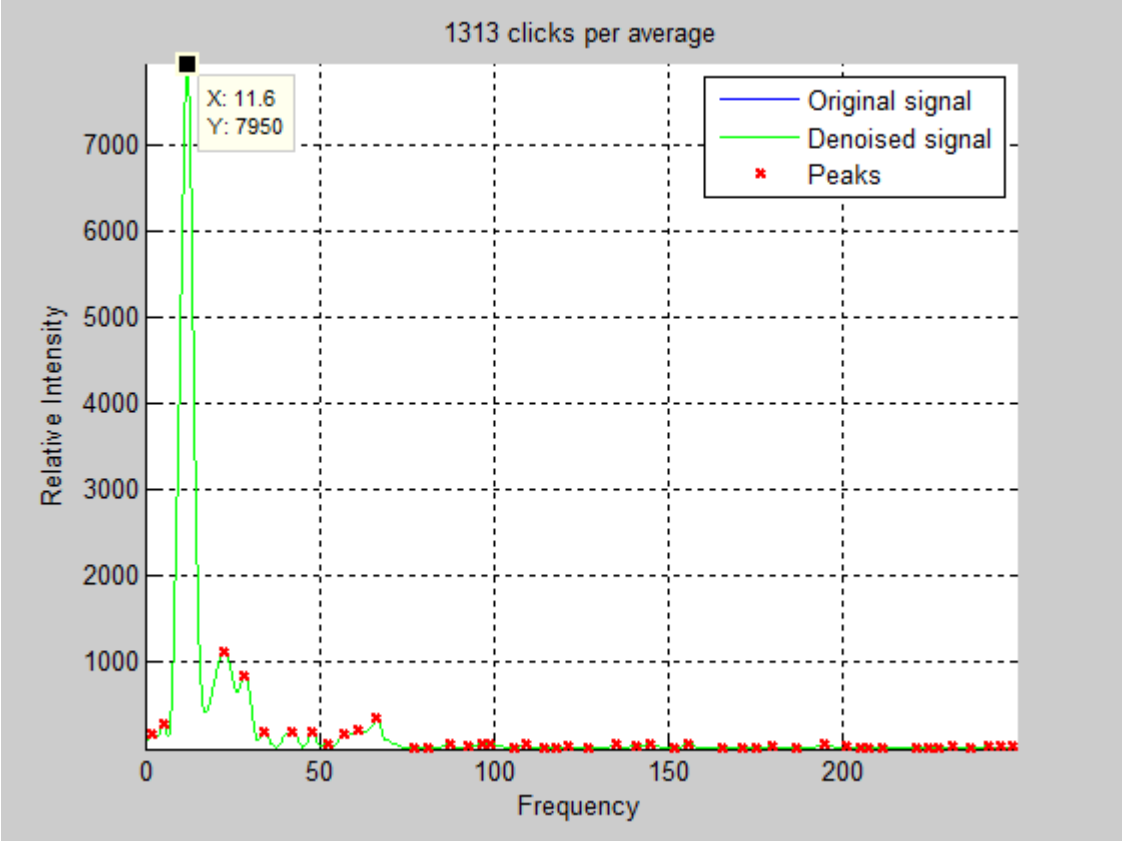


(c)

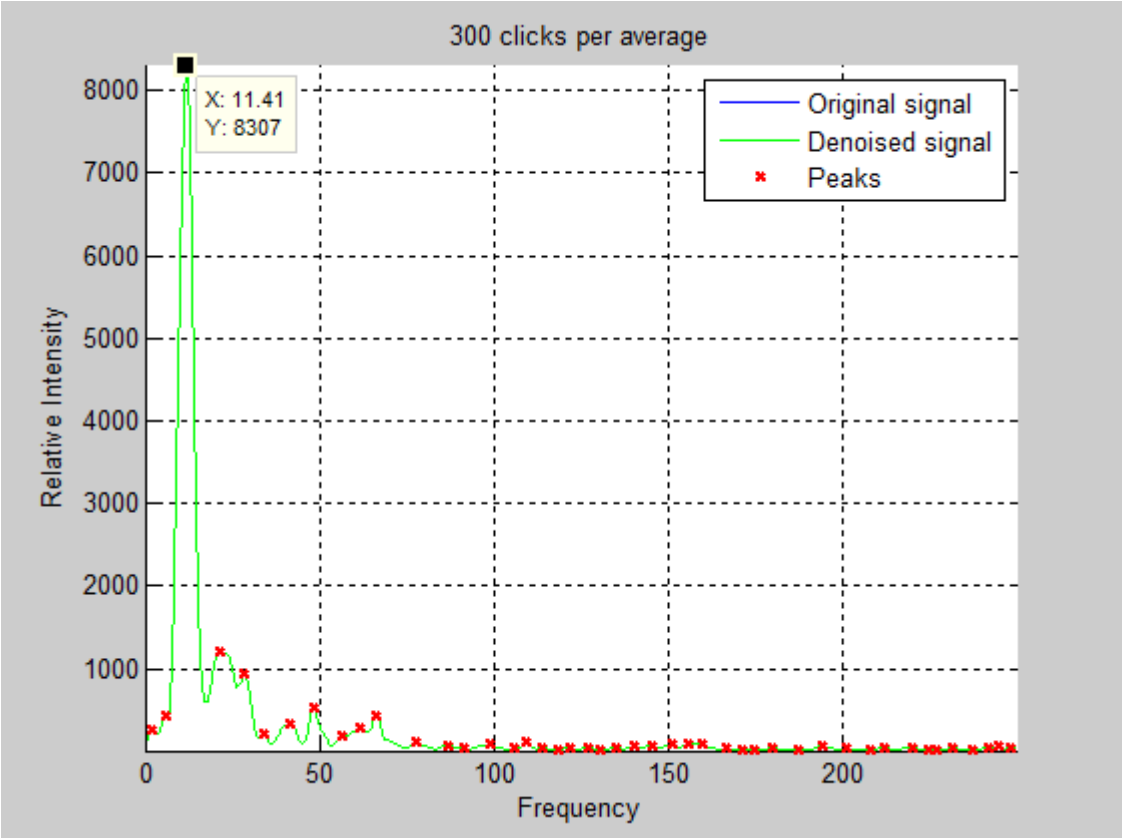


(d)

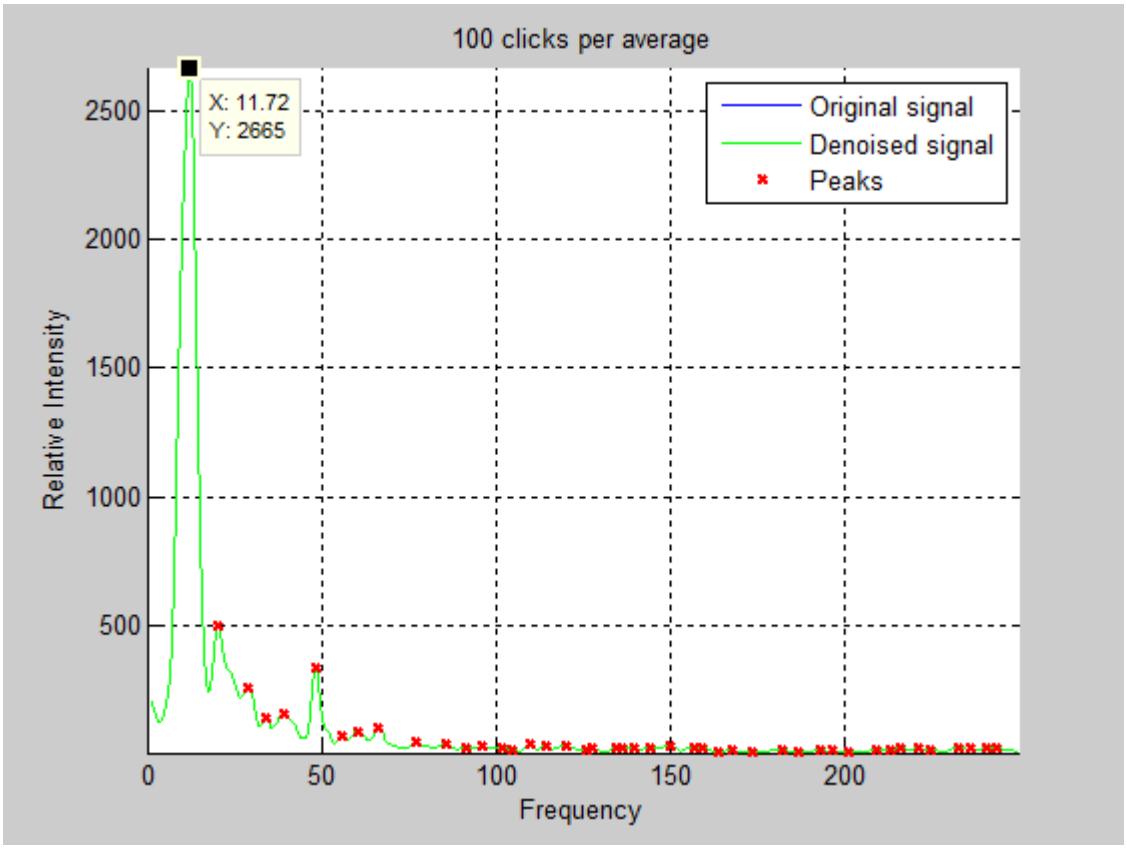
Figure 4.25 *Spectrograms with different window lengths (i.e., M=64-top-left, M=96-top-right, M=128-bottom-left, and M=192-bottom-right) of (a) whole (1313) clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average of Sara Ear_1.*


1313 clicks per average
X: 11.6
Y: 7950
Original signal
Denoised signal
Peaks
Relative Intensity
Frequency
0
50
100
150
200
1000
2000
3000
4000
5000
6000
7000

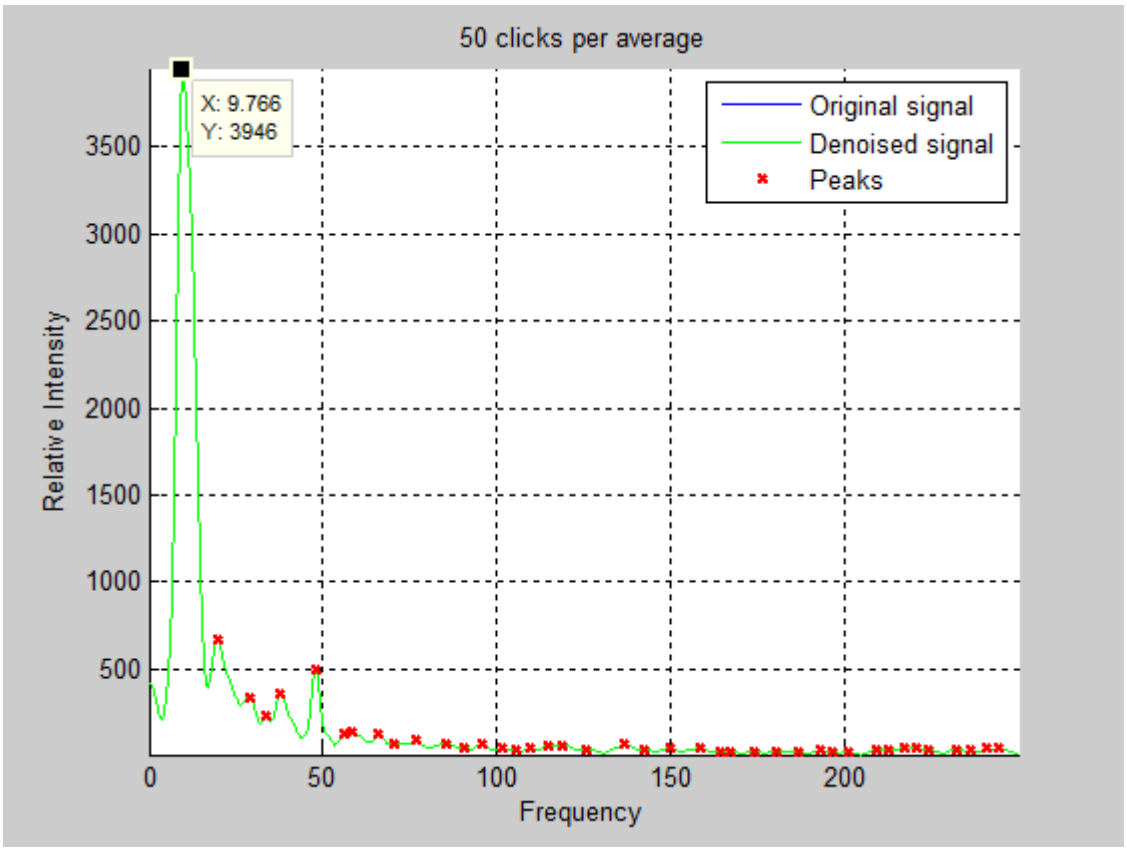
(a)


300 clicks per average
X: 11.41
Y: 8307
Original signal
Denoised signal
Peaks
Relative Intensity
Frequency
0
50
100
150
200
1000
2000
3000
4000
5000
6000
7000
8000

(b)

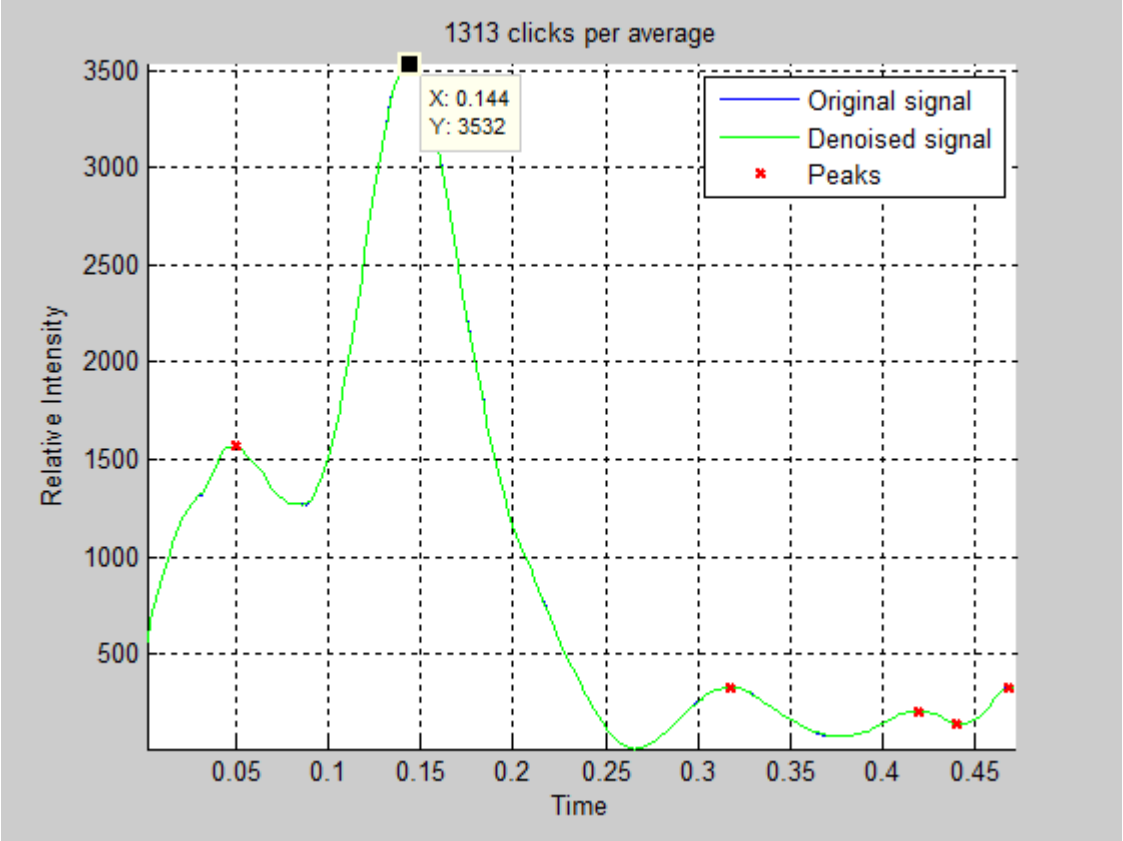


(c)

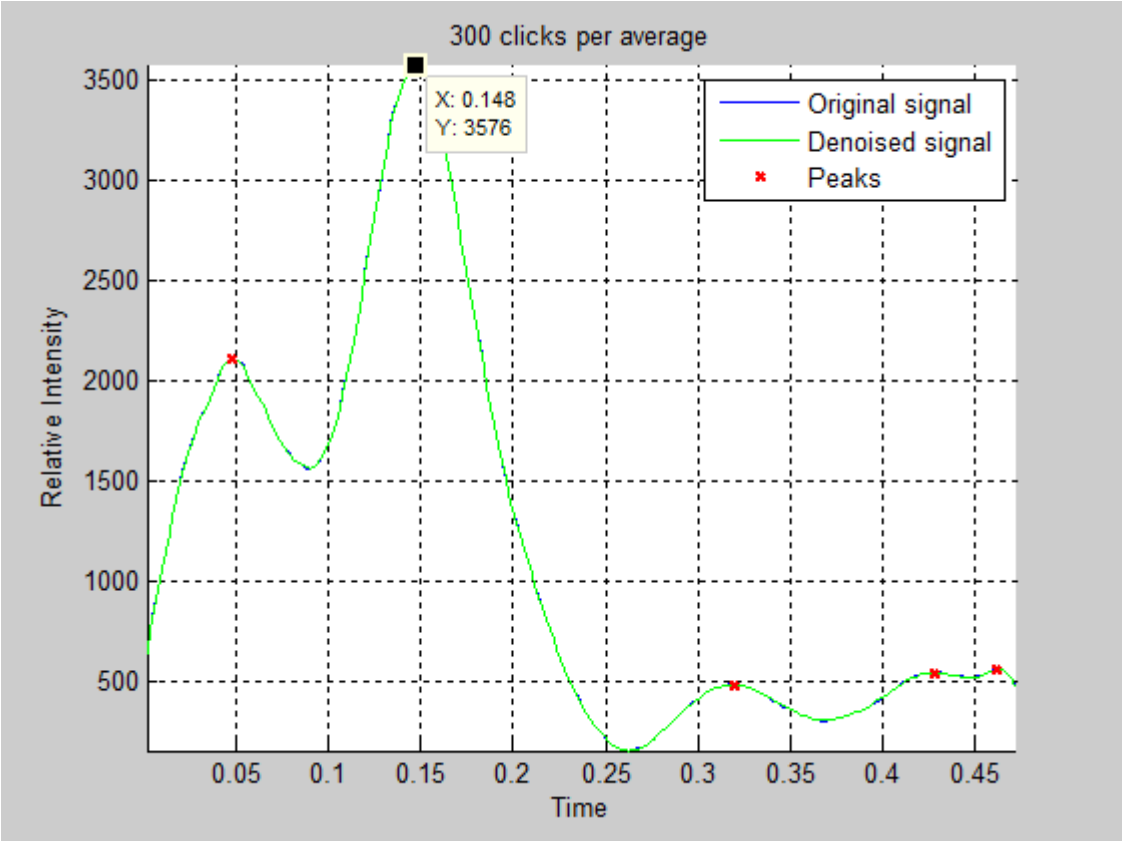


(d)

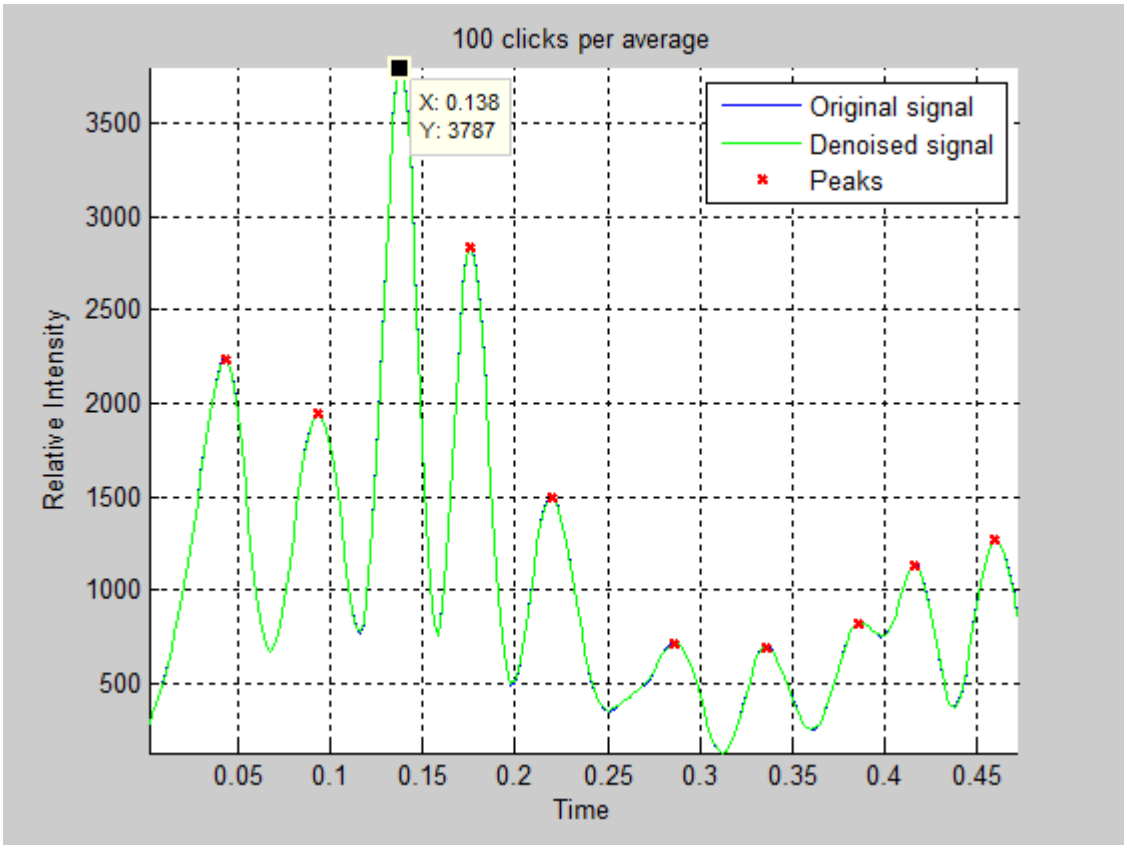
Figure 4.26 *Peaks of spectral estimates along frequencies with window length M=384 and (a) whole clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average for Sara Ear_1.*


1313 clicks per average
X: 0.144
Y: 3532
Original signal
Denoised signal
Peaks
Relative Intensity
Time
3500
3000
2500
2000
1500
1000
500
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

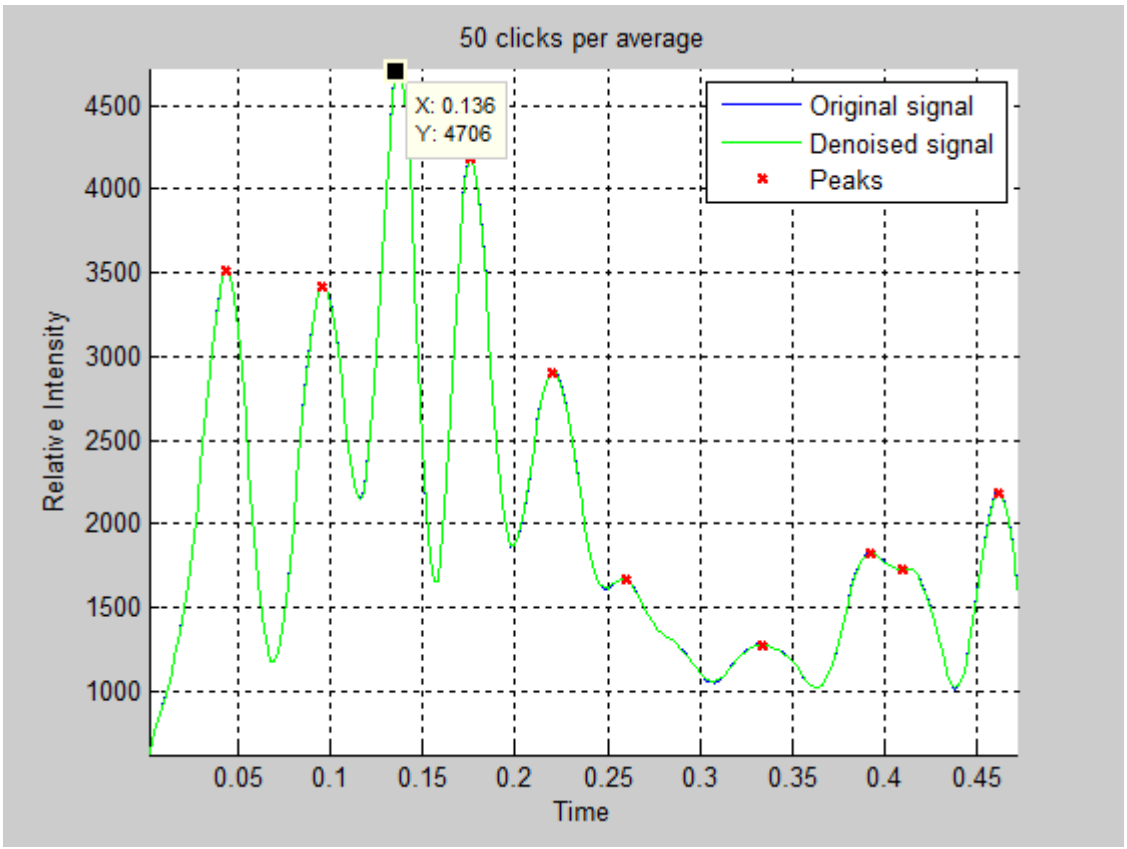
(a)


300 clicks per average
X: 0.148
Y: 3576
Original signal
Denoised signal
Peaks
Relative Intensity
Time
3500
3000
2500
2000
1500
1000
500
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

(b)



(c)

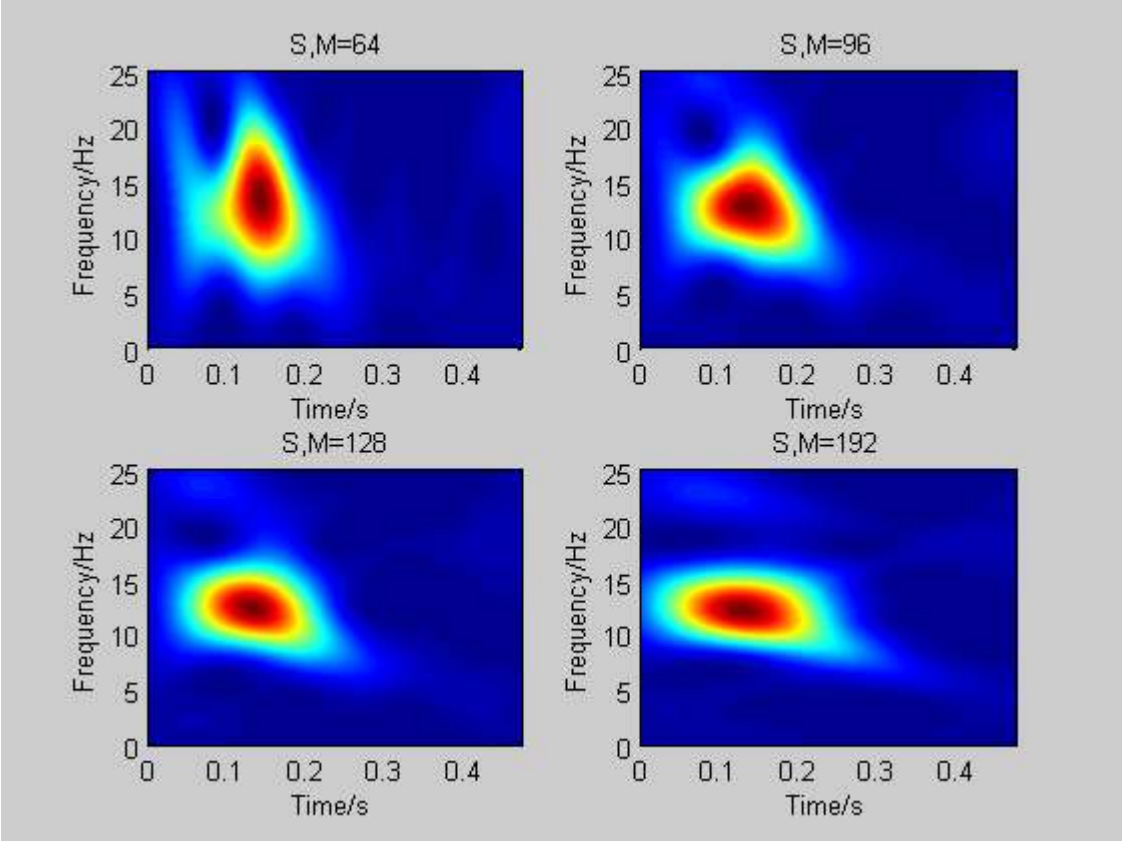


(d)

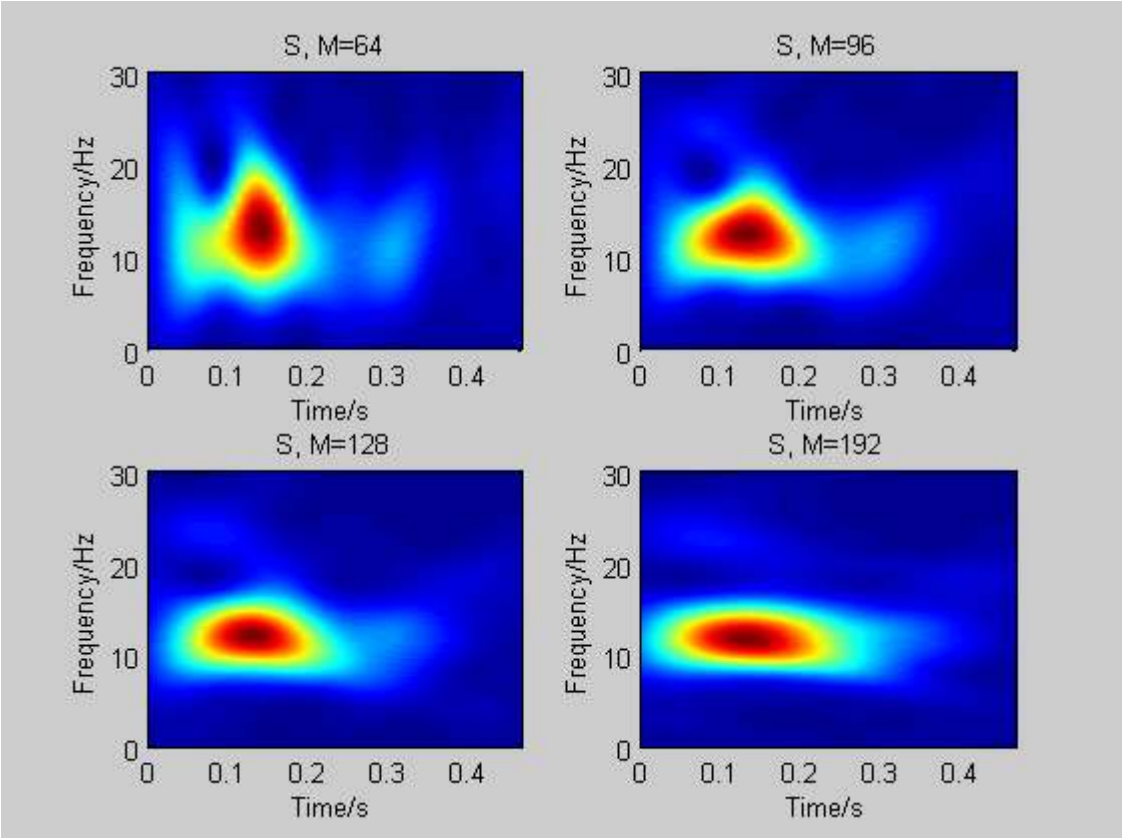
Figure 4.27 *Peaks of spectral estimates along times with window length M=32 and (a) whole clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average for Sara Ear_1.*

Table 4.4 *Highest Peak points of the spectral estimate of Sara Ear_1.*

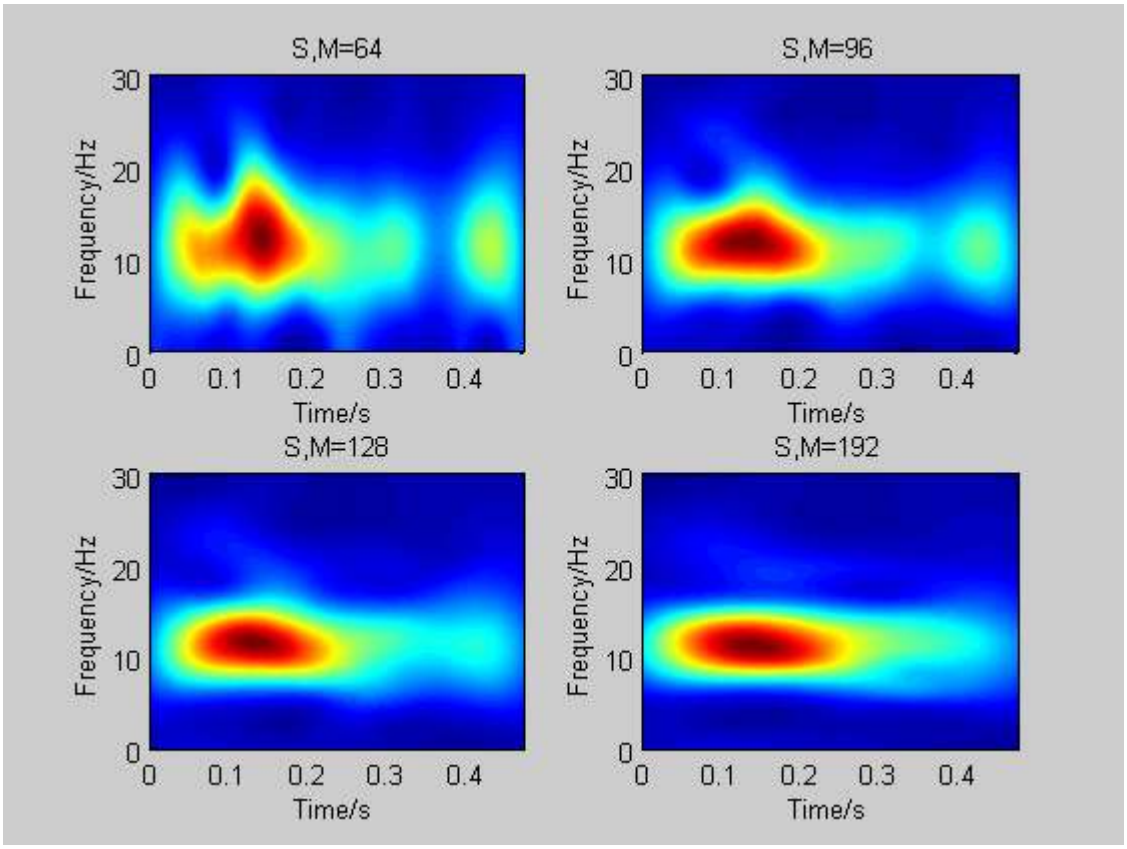
$\boldsymbol{Number\ of\ Clicks\ Per\ Average}$ $\leftrightarrow$					$\boldsymbol{Mean}$ $\downarrow$
$\boldsymbol{Peak\ at}$ $\downarrow$	$\boldsymbol{Whole}$ $(\mathbf{1313})$	$\mathbf{300}$	$\mathbf{100}$	$\mathbf{50}$	
$\boldsymbol{Time}$ $(\boldsymbol{t})$	0.144	0.148	0.138	0.136	0.1415
$\boldsymbol{Frequency}$ $(\boldsymbol{f})$	11.600	11.410	11.720	09.766	11.124


S,M=64
S,M=96
S,M=128
S,M=192
Frequency/Hz
Time/s
0
5
10
15
20
25
0.1
0.2
0.3
0.4

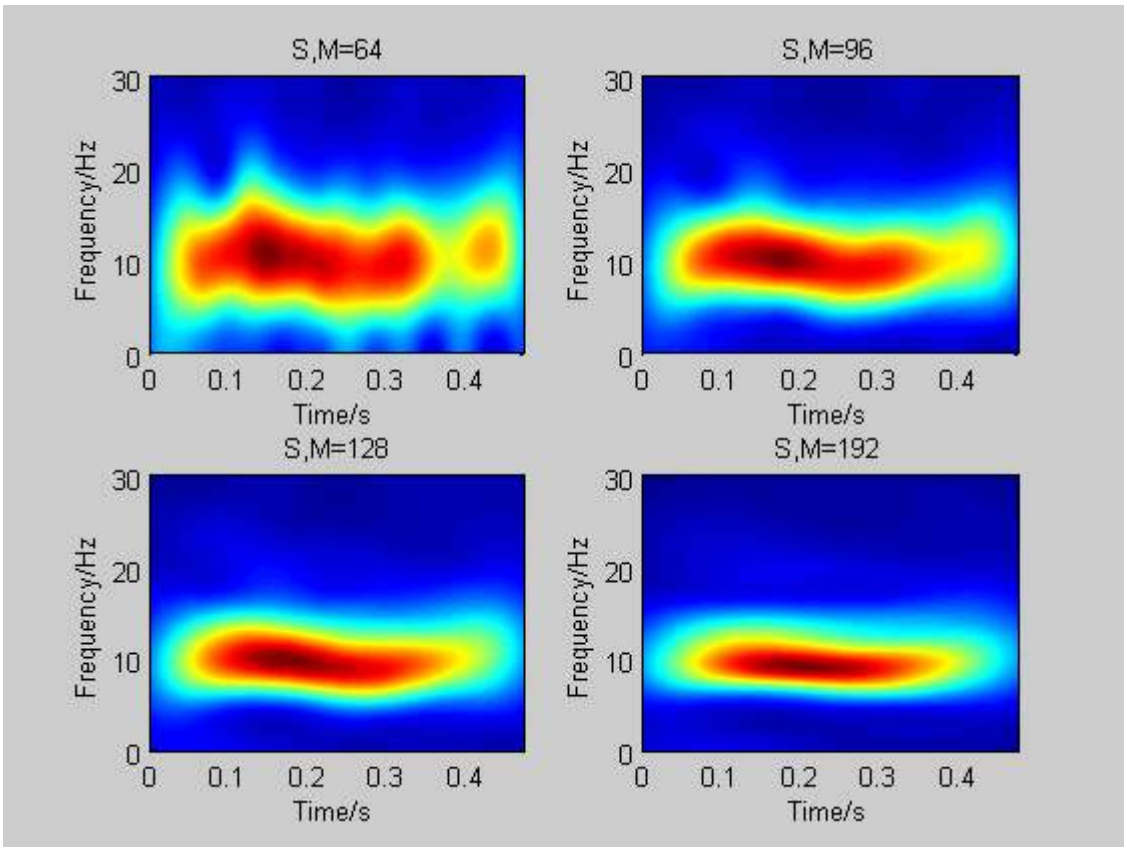
(a)


S, M=64
S, M=96
S, M=128
S, M=192
Frequency/Hz
Time/s
0
10
20
30
0.1
0.2
0.3
0.4

(b)

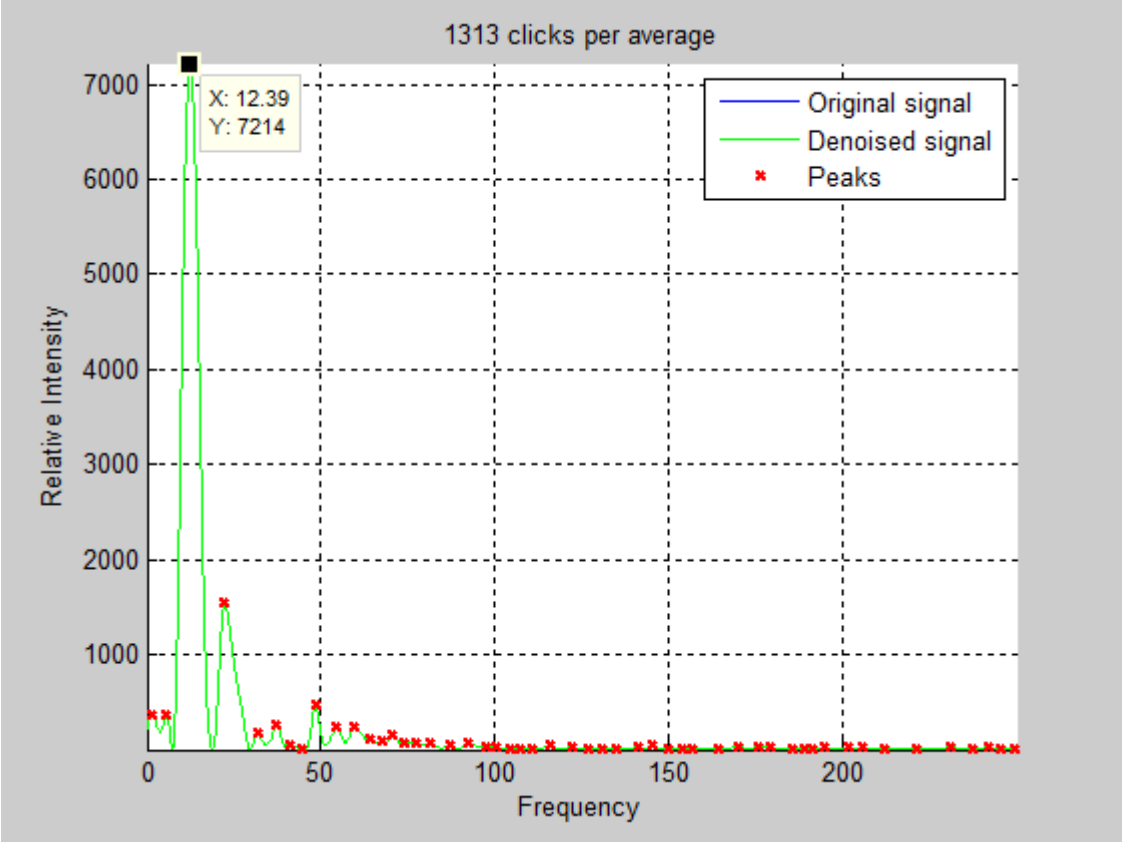


(c)

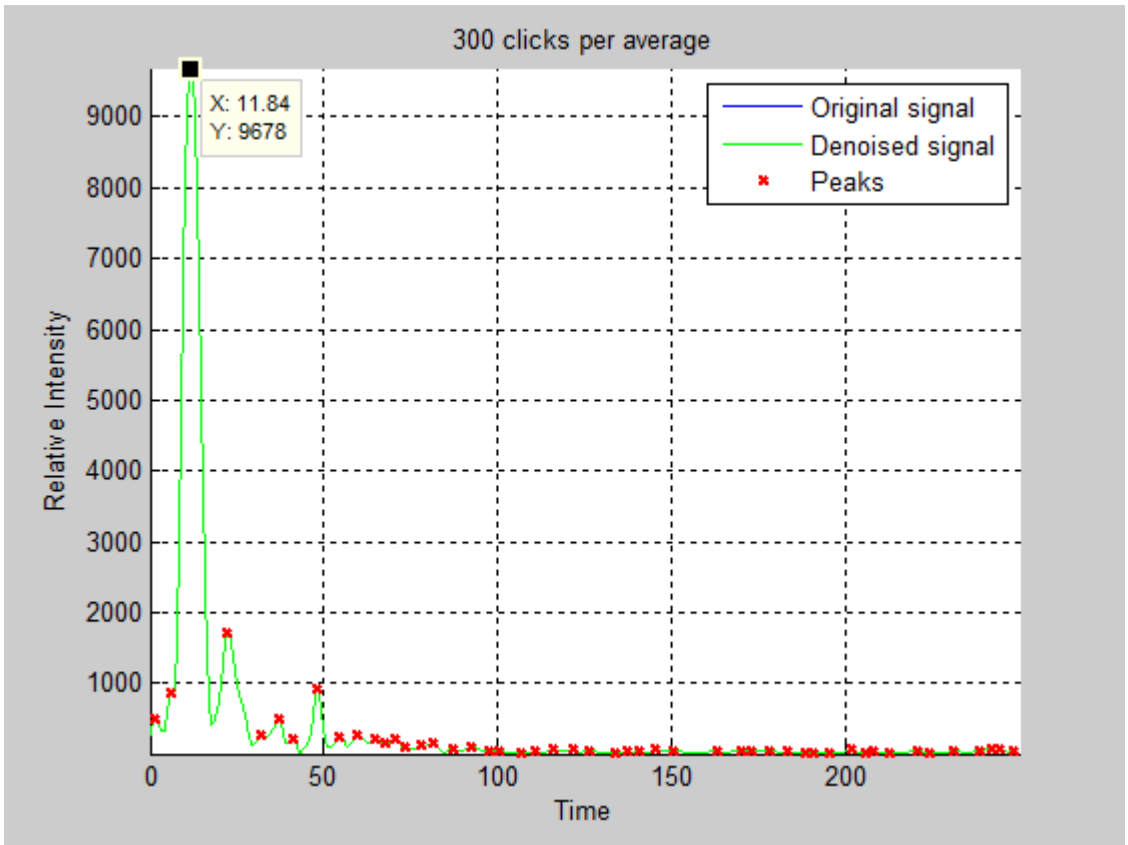


(d)

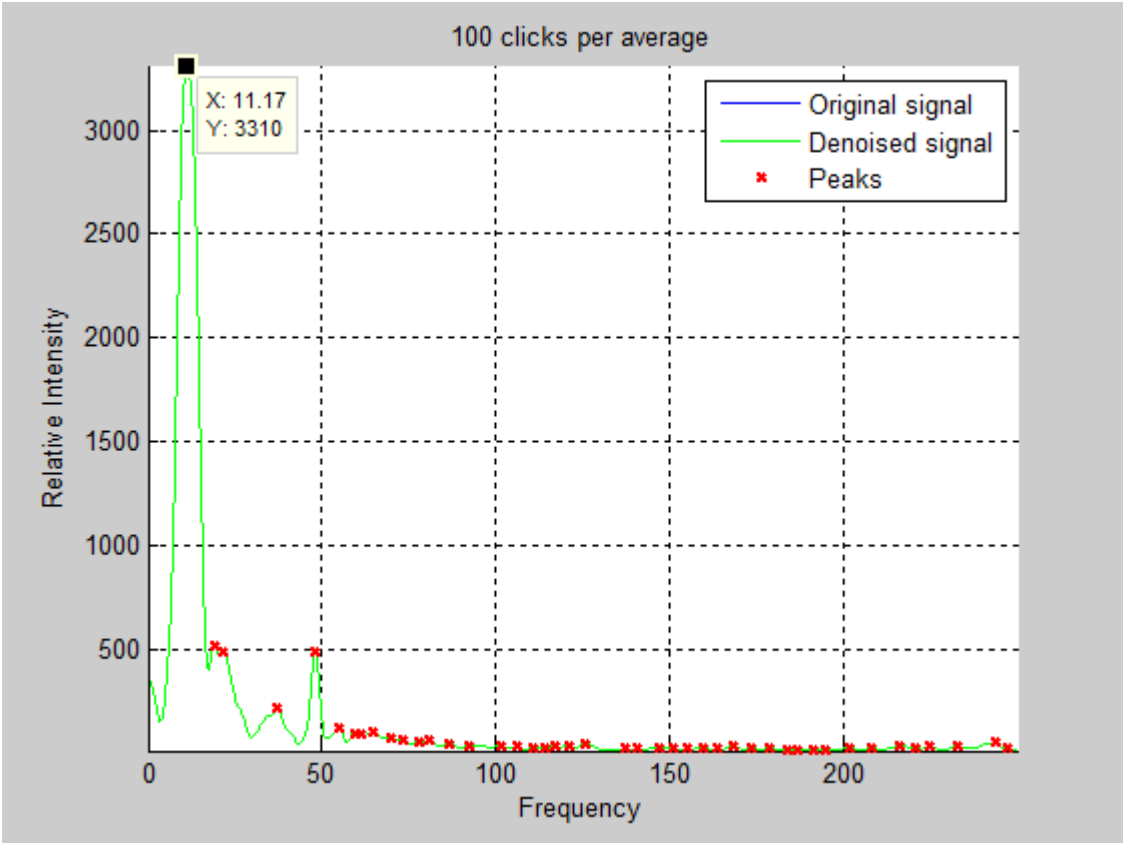
Figure 4.28 *Spectrograms with different window lengths (i.e., M=64-top-left, M=96-top-right, M=128-bottom-left, and M=192-bottom-right) of (a) whole (1313) clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average of Sara Ear_2.*


1313 clicks per average
X: 12.39
Y: 7214
Original signal
Denoised signal
Peaks
Relative Intensity
Frequency
0
50
100
150
200
1000
2000
3000
4000
5000
6000
7000

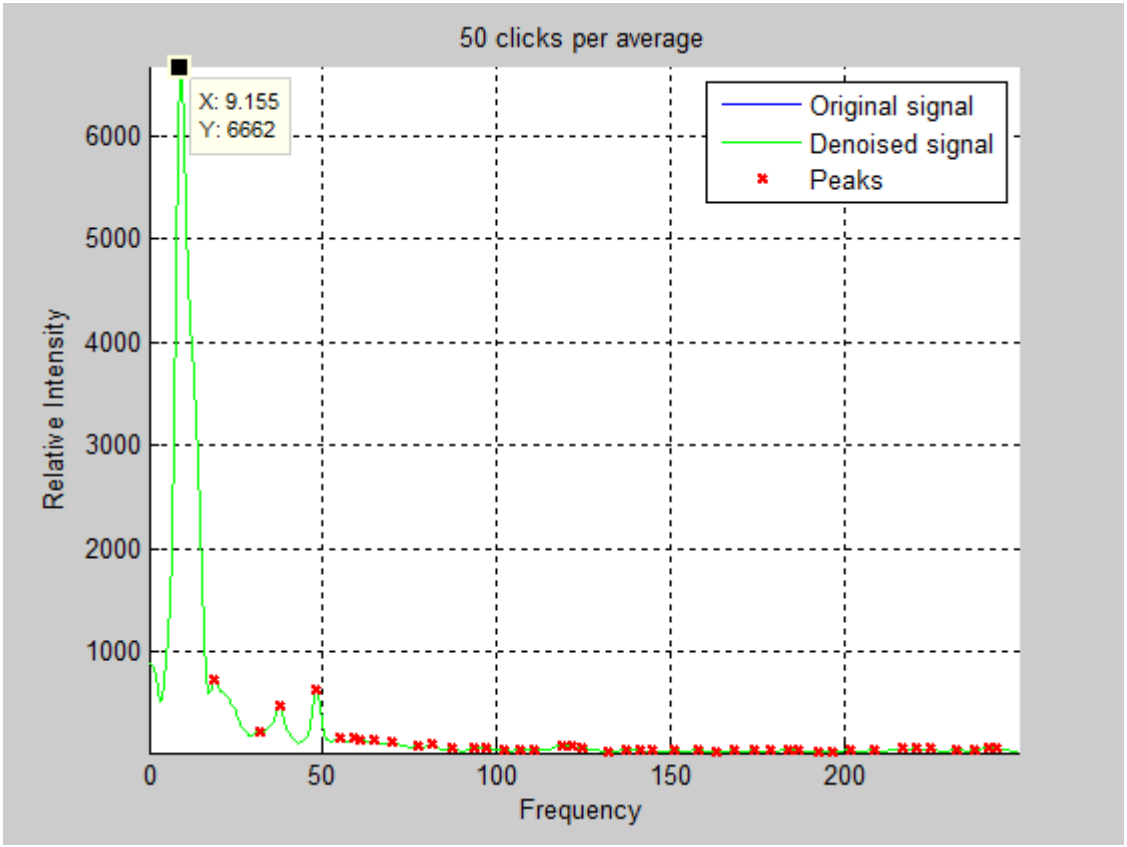
(a)


300 clicks per average
X: 11.84
Y: 9678
Original signal
Denoised signal
Peaks
Relative Intensity
Time
0
50
100
150
200
1000
2000
3000
4000
5000
6000
7000
8000
9000

(b)

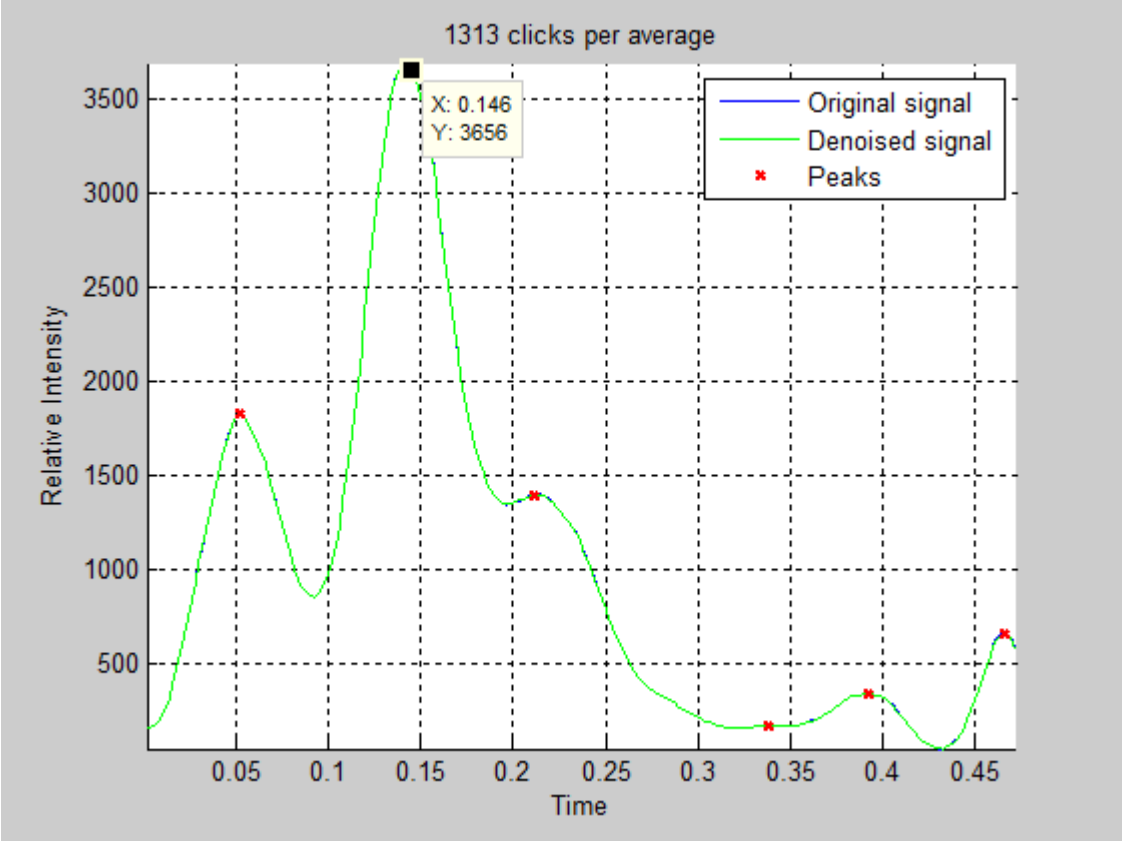


(c)

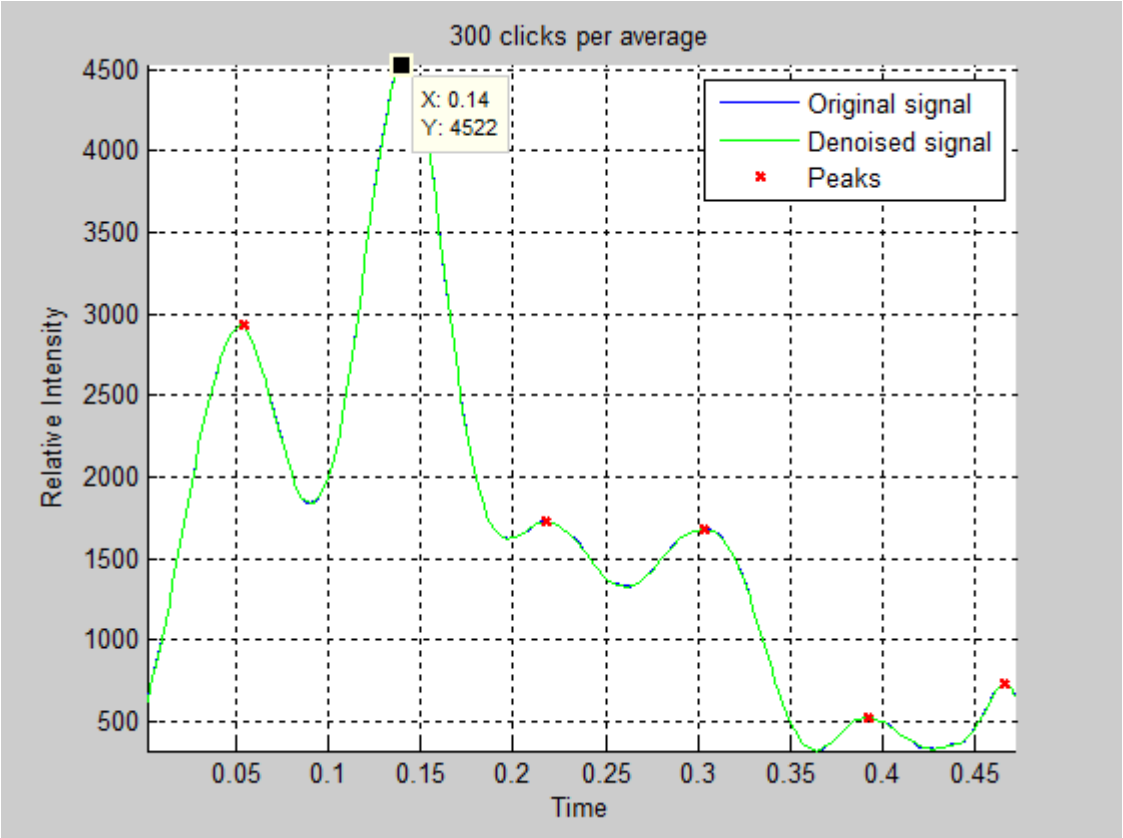


(d)

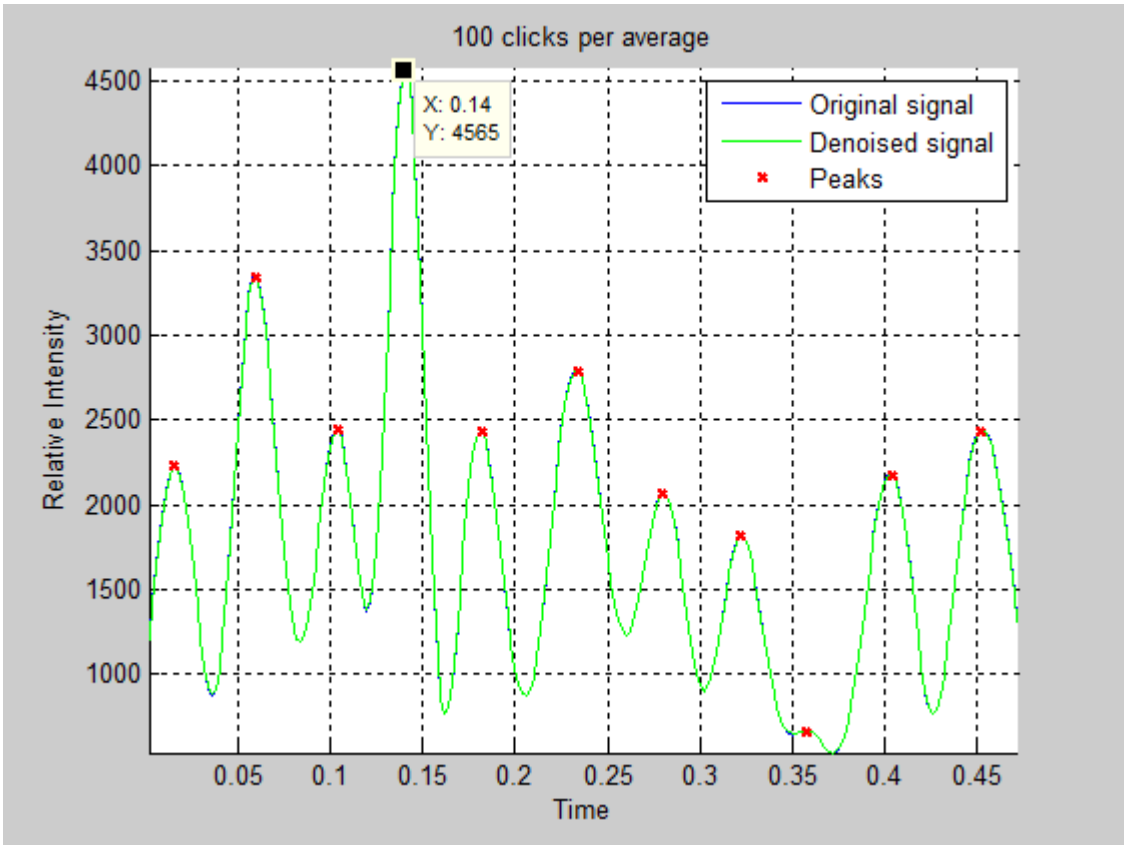
Figure 4.29 *Peaks of spectral estimates along frequencies with window length M=384 and (a) whole clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average for Sara Ear_2.*


1313 clicks per average
X: 0.146
Y: 3656
Original signal
Denoised signal
Peaks
Relative Intensity
Time
3500
3000
2500
2000
1500
1000
500
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

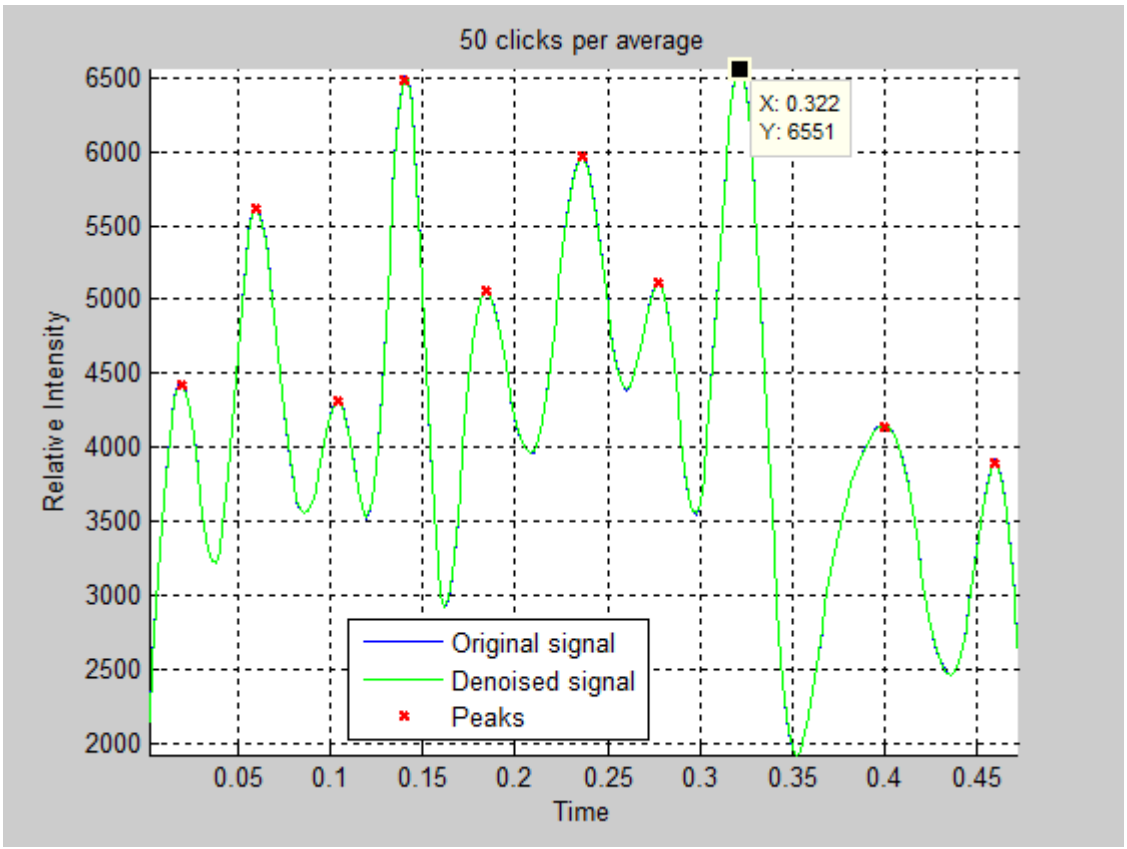
(a)


300 clicks per average
X: 0.14
Y: 4522
Original signal
Denoised signal
Peaks
Relative Intensity
Time
4500
4000
3500
3000
2500
2000
1500
1000
500
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

(b)



(c)

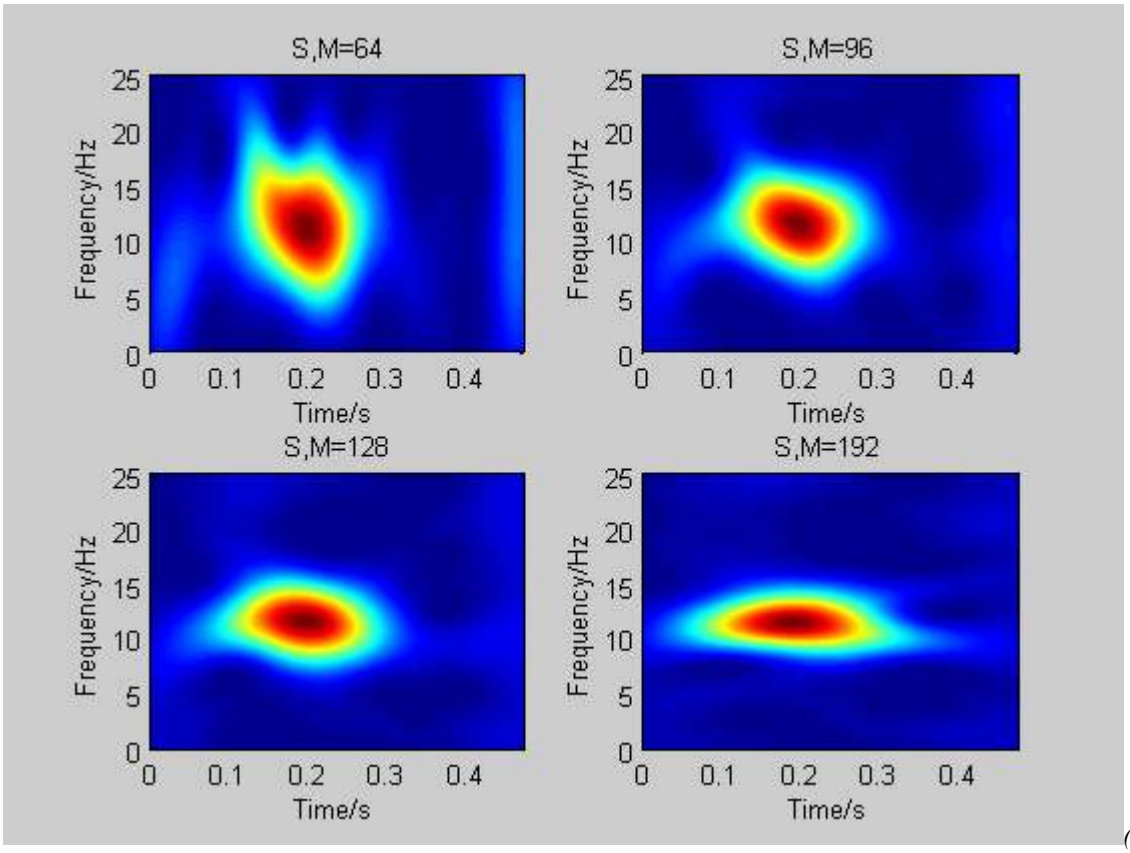


(d)

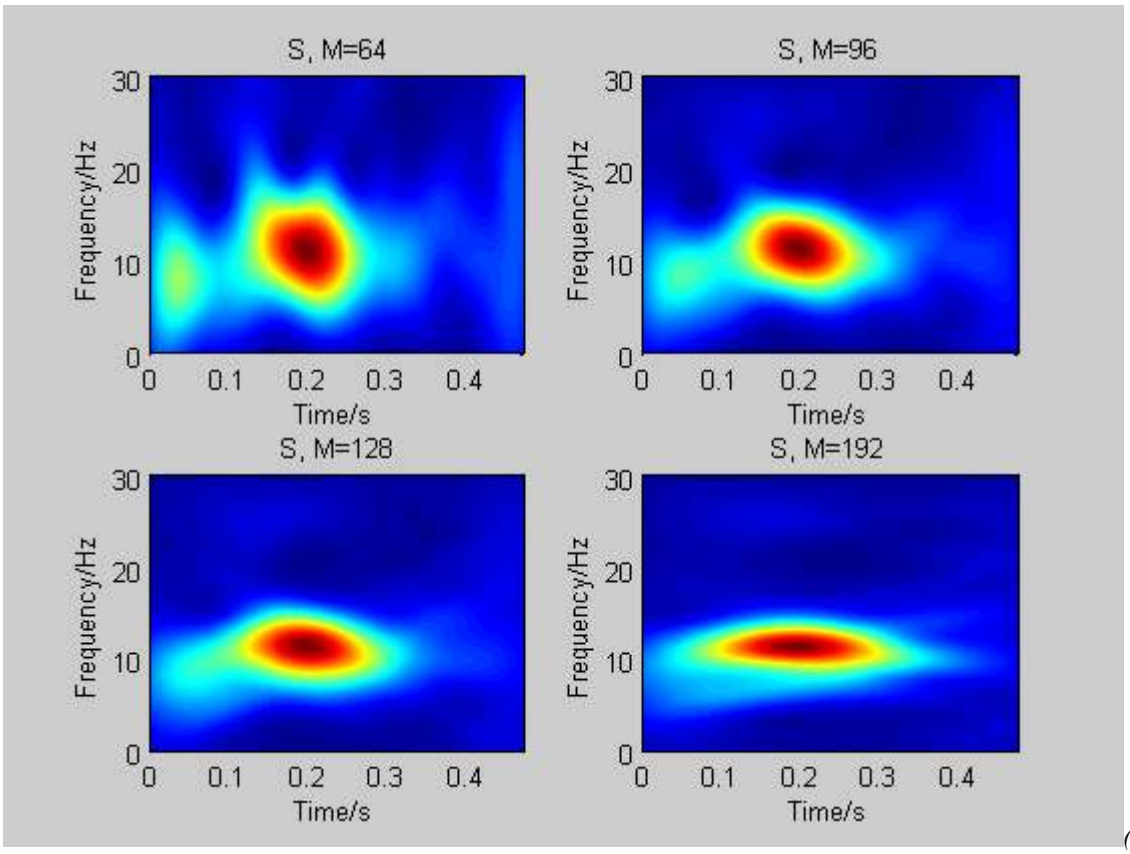
Figure 4.30 *Peaks of spectral estimates along times with window length M=32 and (a) whole clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average for Sara Ear_2.*

Table 4.5 *Highest Peak points of the spectral estimate of Sara Ear_2.*

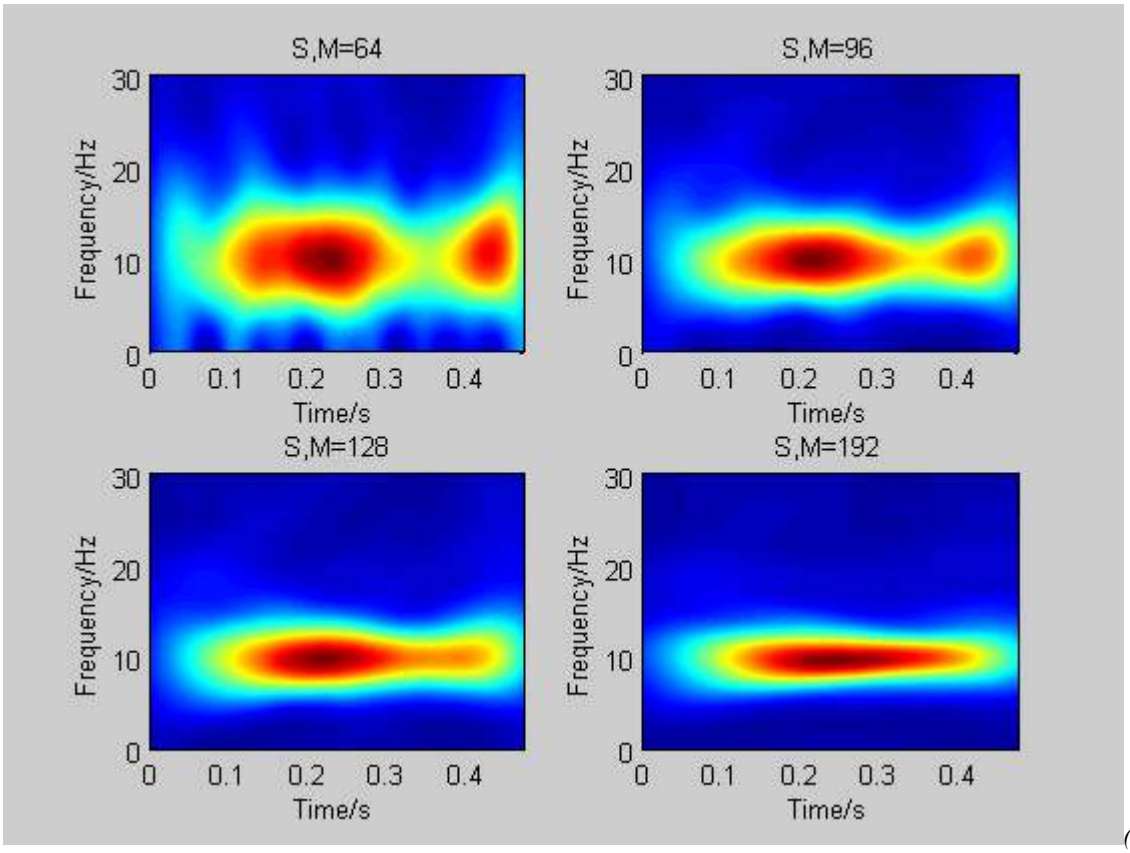
$\boldsymbol{Number\ of\ Clicks\ Per\ Average\ \leftrightarrow}$					$\boldsymbol{Mean\downarrow}$
$\boldsymbol{Peak\ at\downarrow}$	$\boldsymbol{Whole\ (1313)}$	**300**	**100**	**50**	
$\boldsymbol{Time\ (t)}$	0.146	0.140	0.140	0.322	0.187
$\boldsymbol{Frequency\ (f)}$	12.390	11.840	11.170	09.155	11.1388


S,M=64
S,M=96
S,M=128
S,M=192
Frequency/Hz
Time/s
0
5
10
15
20
25
0.1
0.2
0.3
0.4

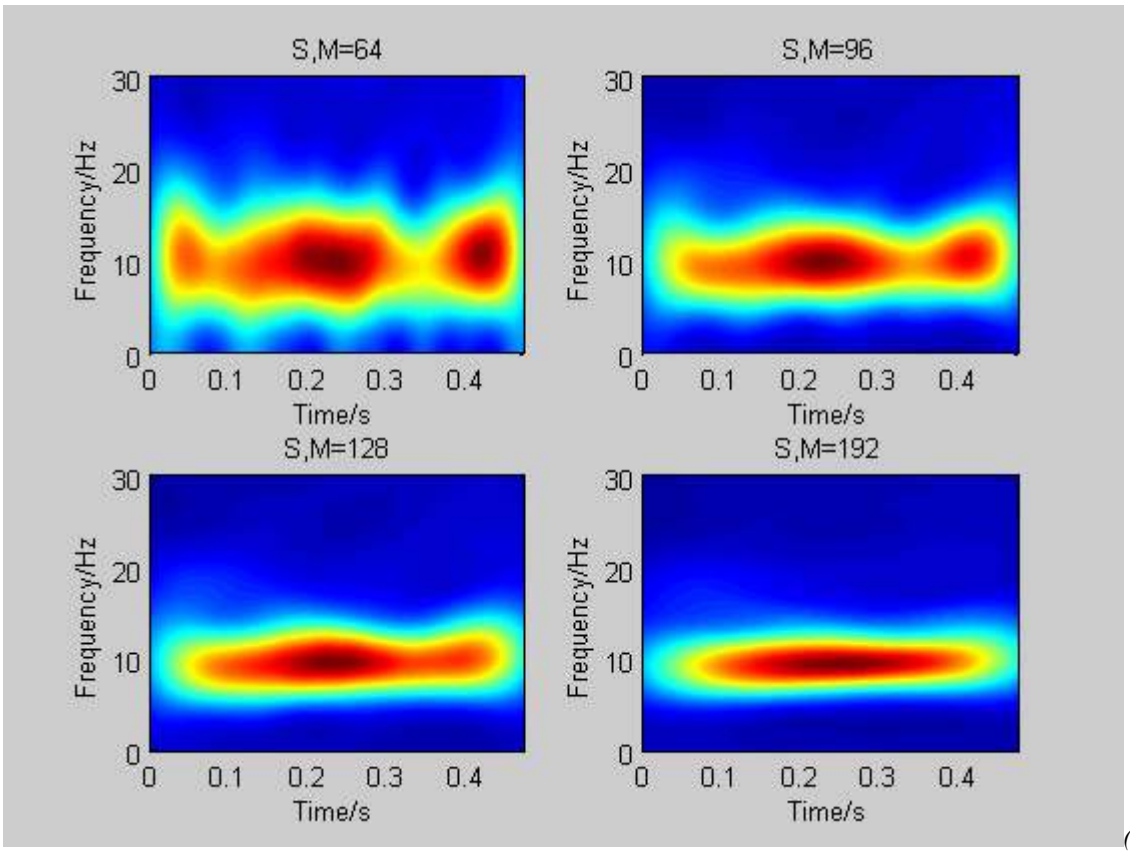
(a)


S, M=64
S, M=96
S, M=128
S, M=192
Frequency/Hz
Time/s
0
10
20
30
0.1
0.2
0.3
0.4

(b)

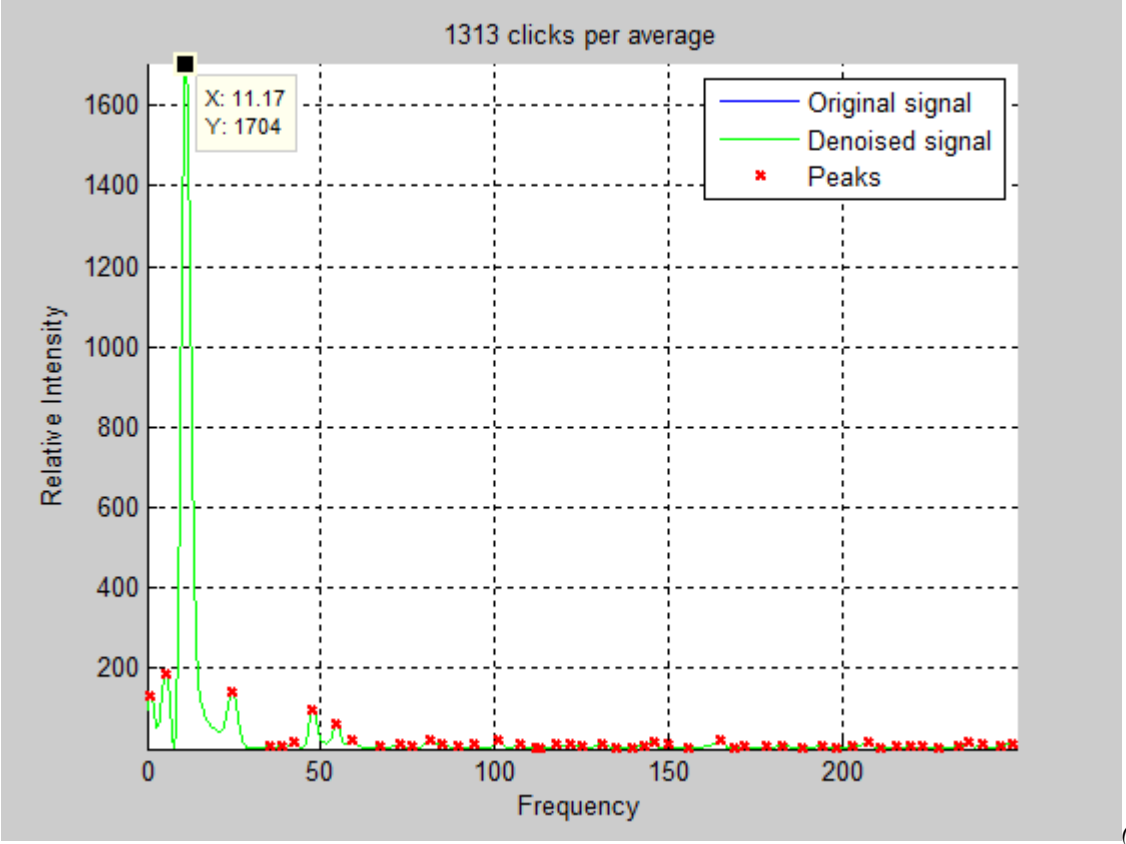


(c)

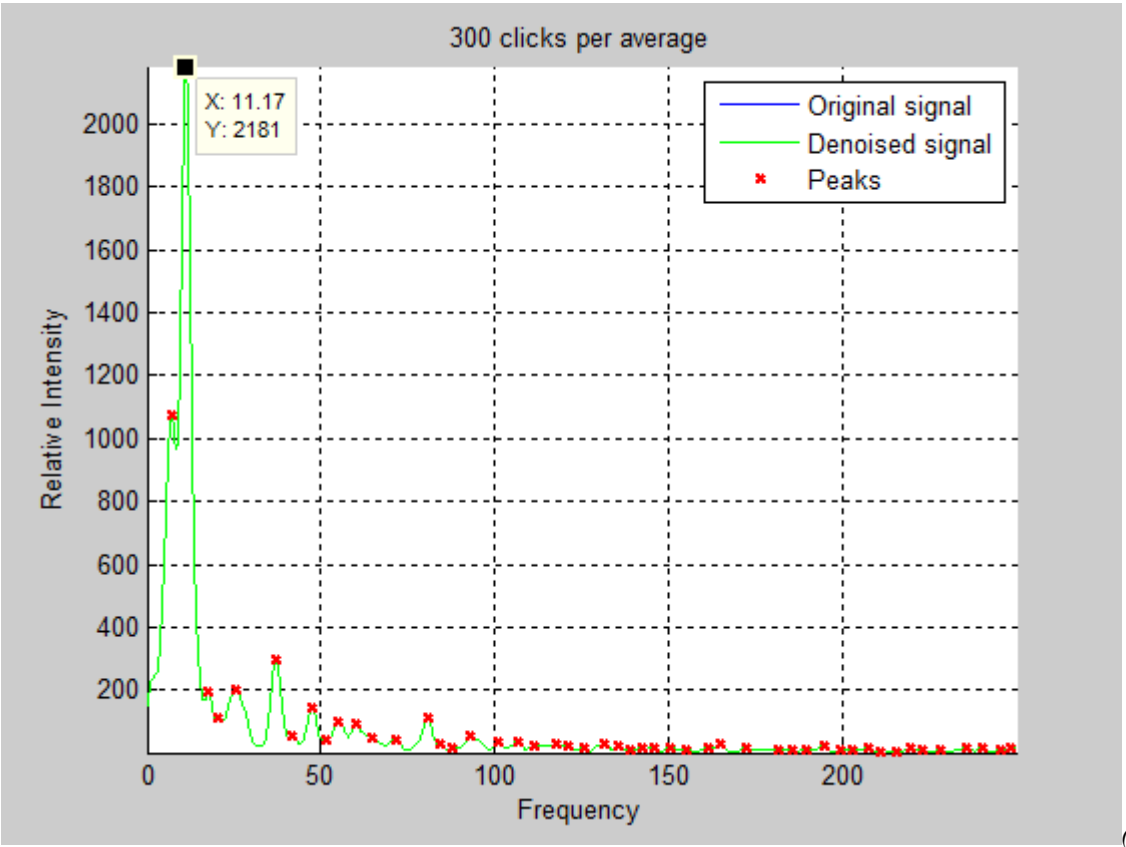


(d)

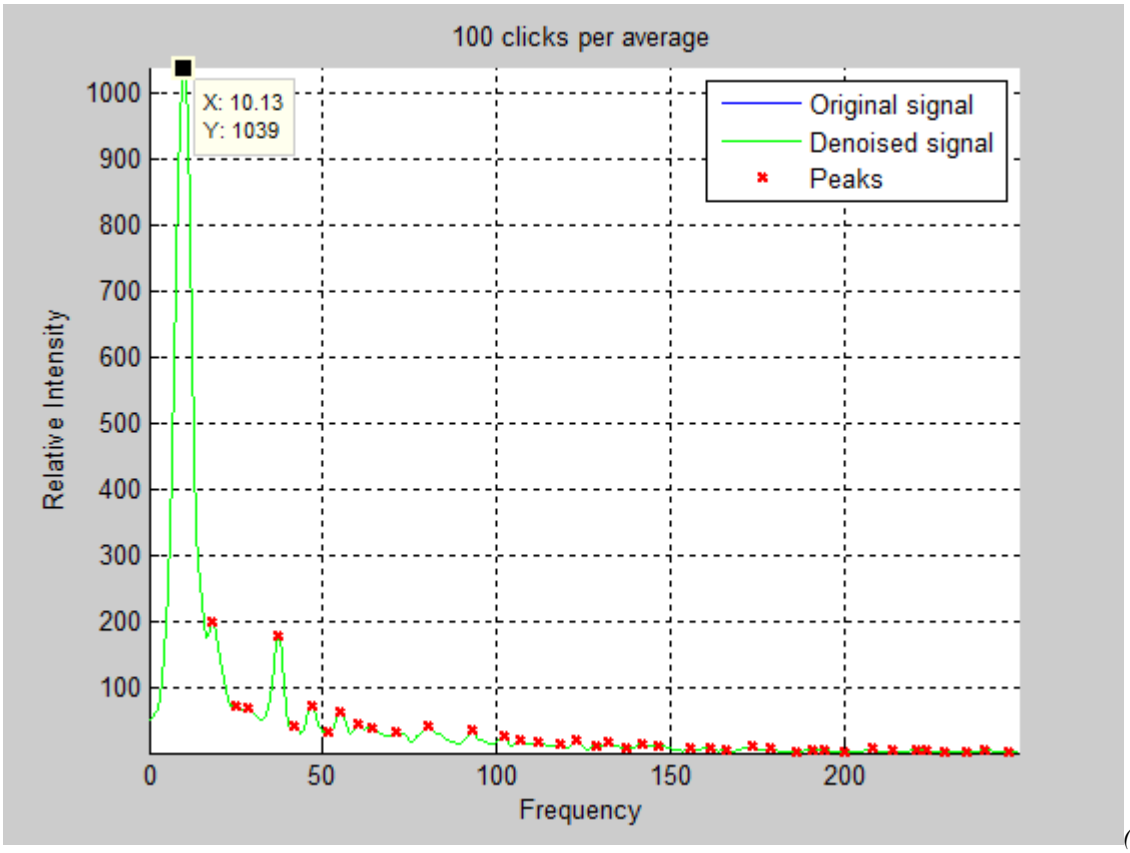
Figure 4.31 *Spectrograms with different window lengths (i.e., M=64-top-left, M=96-top-right, M=128-bottom-left, and M=192-bottom-right) of (a) whole (1313) clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average of Stellan Ear_1.*


1313 clicks per average
X: 11.17
Y: 1704
Original signal
Denoised signal
Peaks
Relative Intensity
Frequency
1600
1400
1200
1000
800
600
400
200
0
50
100
150
200

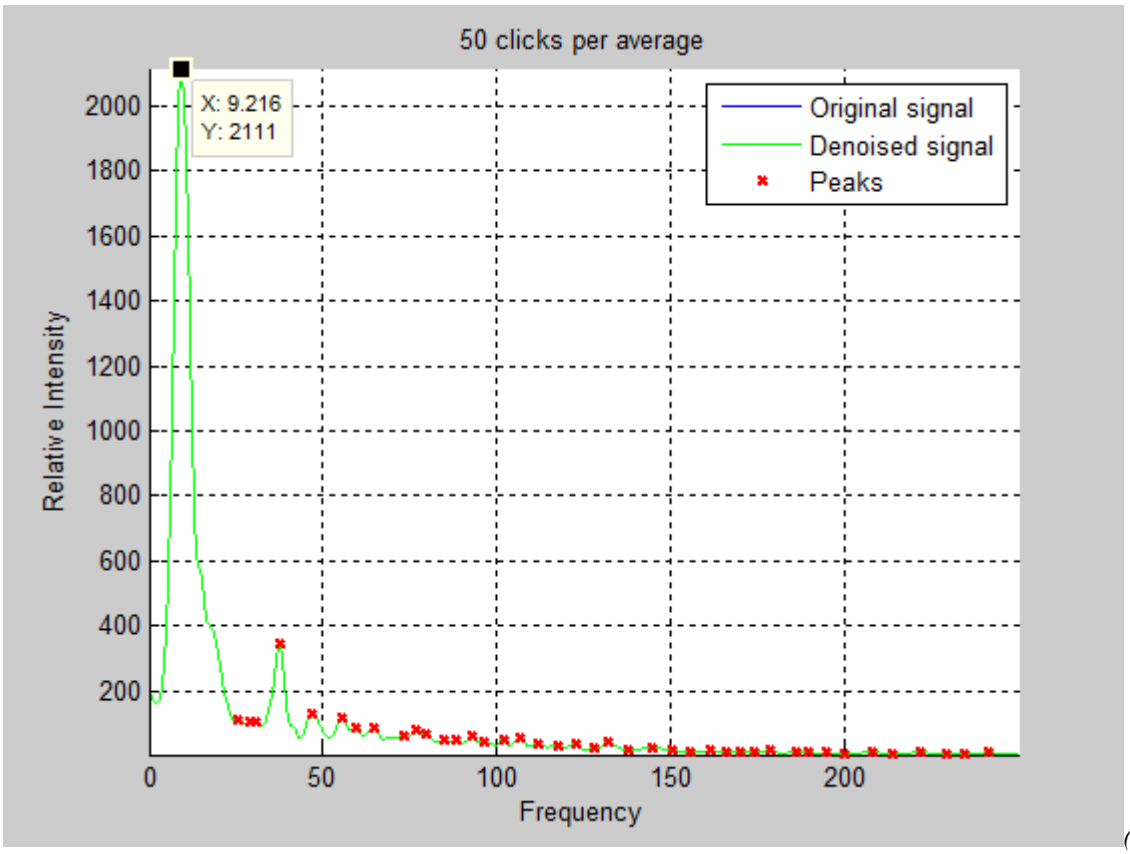
(a)


300 clicks per average
X: 11.17
Y: 2181
Original signal
Denoised signal
Peaks
Relative Intensity
Frequency
2000
1800
1600
1400
1200
1000
800
600
400
200
0
50
100
150
200

(b)

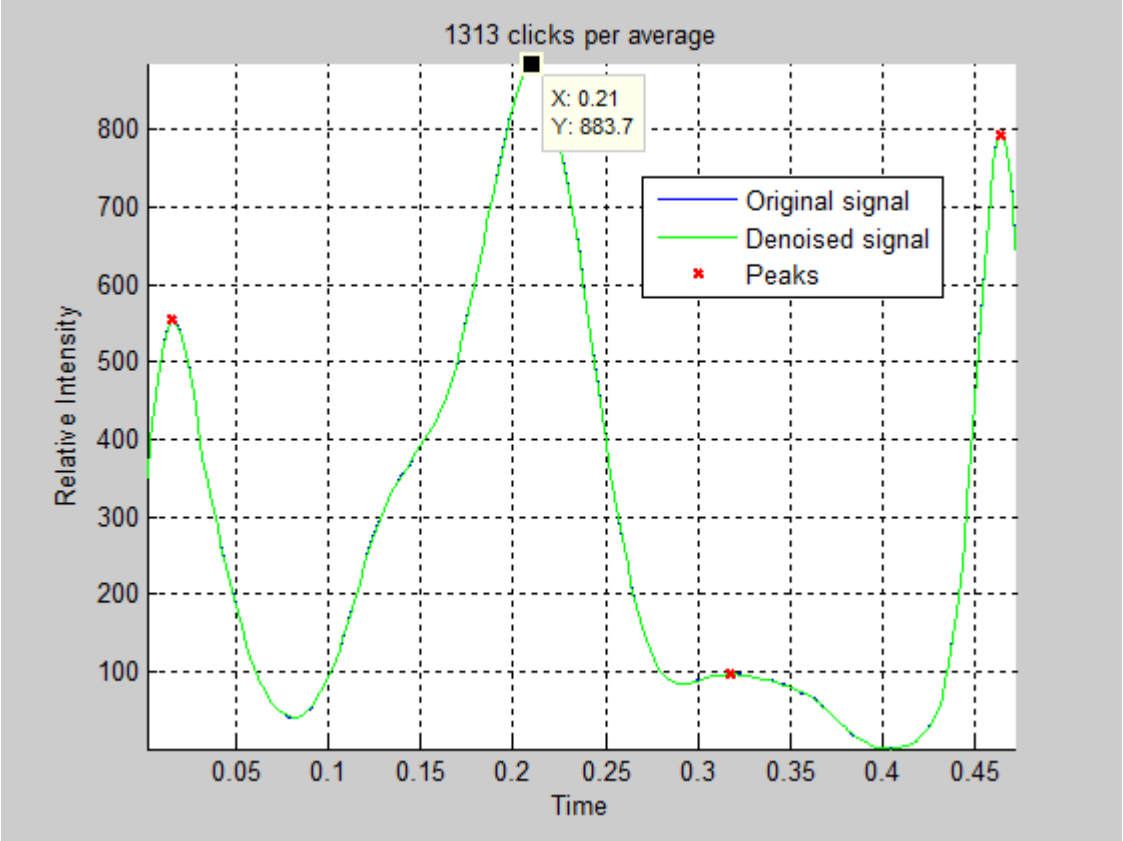


(c)

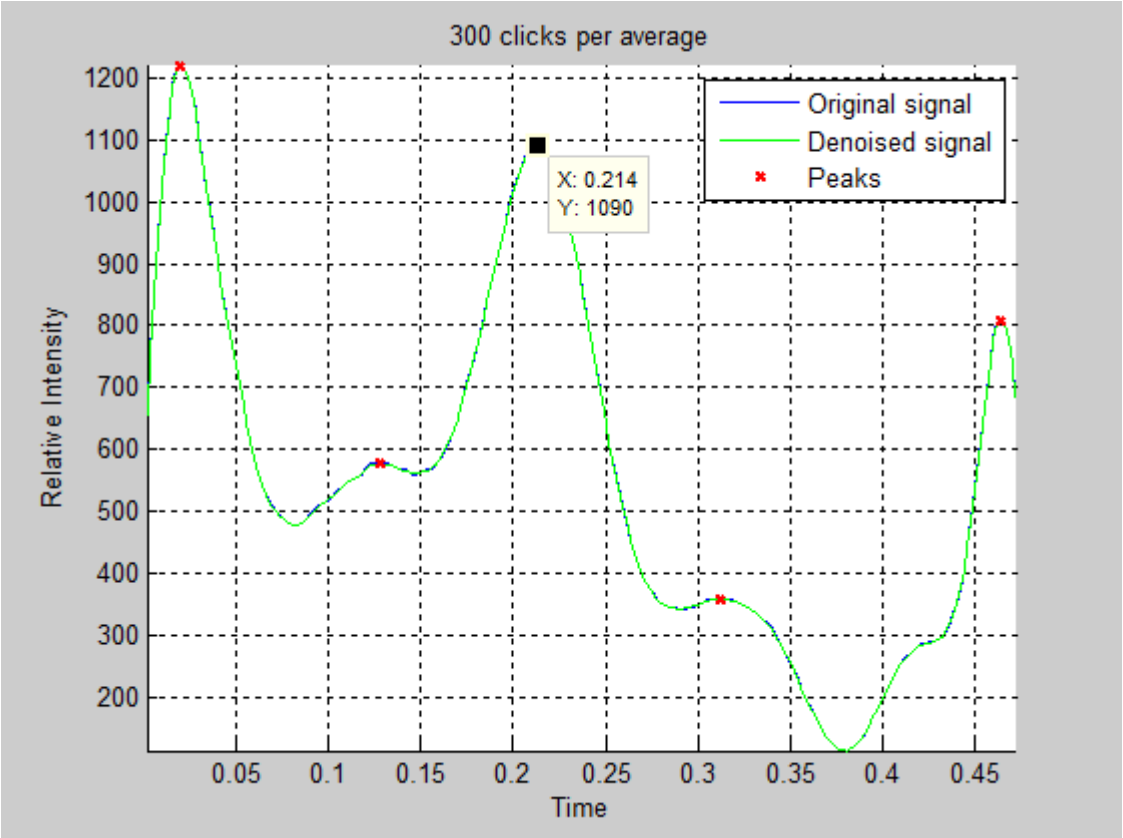


(d)

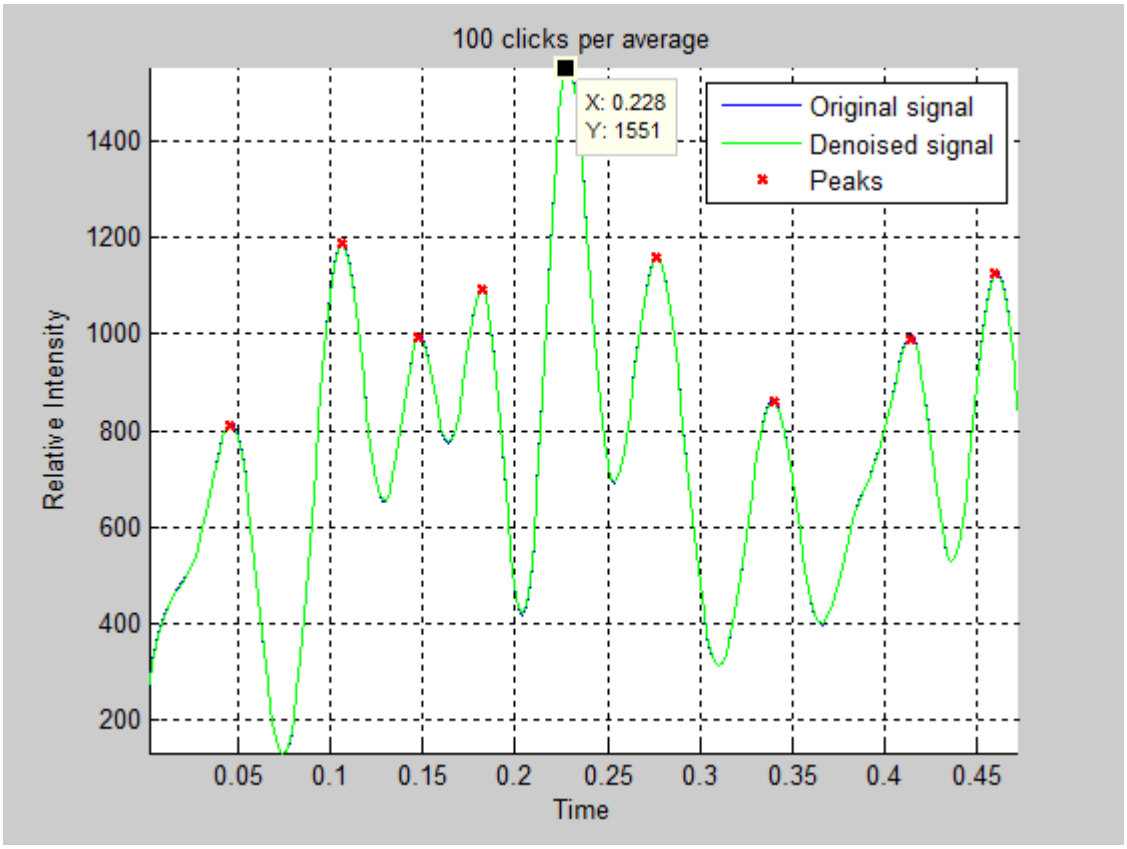
Figure 4.32 *Peaks of spectral estimates along frequencies with window length M=384 and (a) whole clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average for Stellan Ear_1.*


1313 clicks per average
X: 0.21
Y: 883.7
Original signal
Denoised signal
Peaks
Relative Intensity
Time
100
200
300
400
500
600
700
800
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

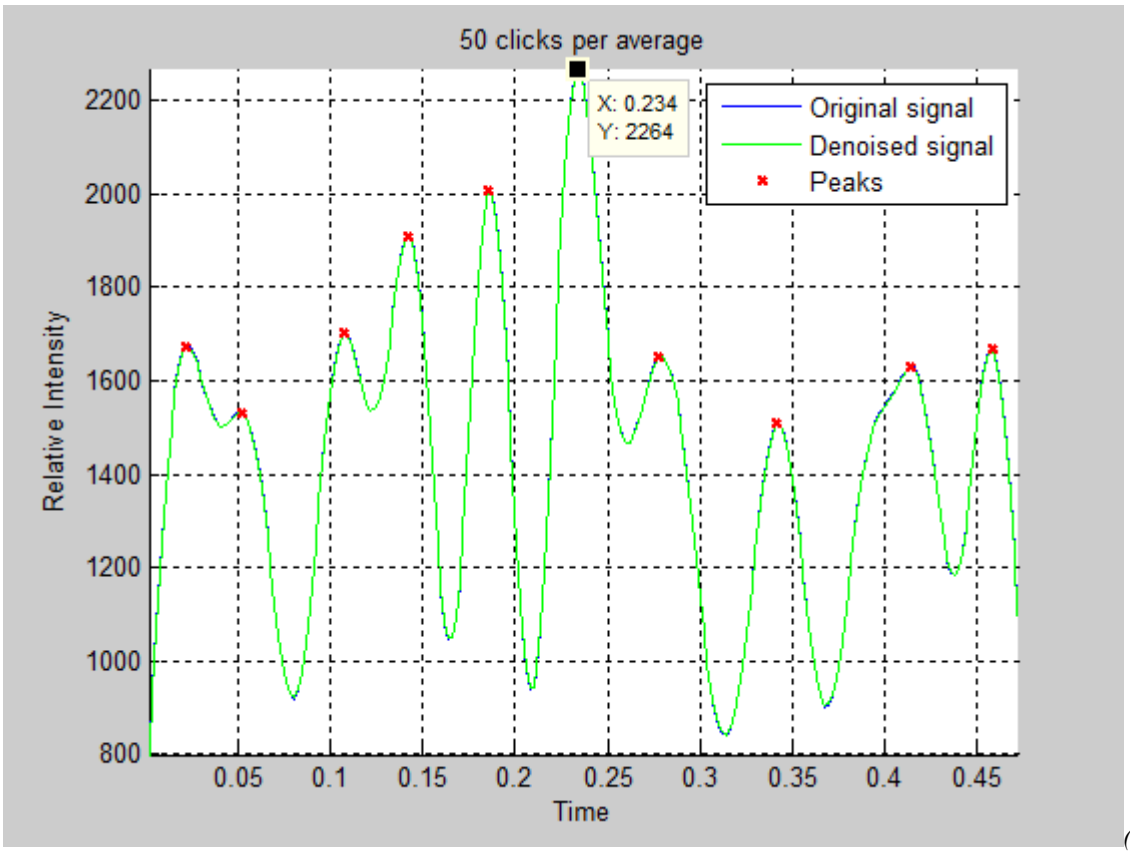
(a)


300 clicks per average
X: 0.214
Y: 1090
Original signal
Denoised signal
Peaks
Relative Intensity
Time
200
300
400
500
600
700
800
900
1000
1100
1200
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45

(b)



(c)



(d)

Figure 4.33 *Peaks of spectral estimates along times with window length M=32 and (a) whole clicks per average, (b) 300 clicks per average, (c) 100 clicks per average and (d) 50 clicks per average for Stellan Ear_1.*

Table 4.6 *Highest Peak points of the spectral estimate of Stellan Ear_1.*

$\boldsymbol{Number\ of\ Clicks\ Per\ Average}$ $\leftrightarrow$					$\boldsymbol{Mean}$ $\downarrow$
$\boldsymbol{Peak\ at}$ $\downarrow$	$\boldsymbol{Whole}$ $(\mathbf{1313})$	$\mathbf{300}$	$\mathbf{100}$	$\mathbf{50}$	
$\boldsymbol{Time}$ $(\boldsymbol{t})$	0.210	0.214	0.228	0.234	0.2215
$\boldsymbol{Frequency}$ $(\boldsymbol{f})$	11.170	11.170	10.130	09.216	10.4215