NOTE TO USERS

This reproduction is the best copy available.

A NEW APPROACH TO NONPARAMETRIC SPECTRAL DENSITY ESTIMATION

by

Sudeshna Pal

B.Sc. in Electrical Engineering, Queen's University, Kingston, ON, 2005

A thesis
presented to Ryerson University
in partial fulfillment of the
requirement for the degree of
Master of Applied Science (MASc)
in the Program of
Electrical and Computer Engineering.

Toronto, Ontario, Canada, 2007

Author's Declaration

I hereby declare that I am the sole author of this thesis.
I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Signature

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Signature

Borrower's Page

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Name	Signature	Address	Date

Abstract

A New Approach to Nonparametric Spectral Density Estimation

Master of Applied Science (MASc)
Department of Electrical and Computer Engineering
Ryerson University

A novel approach to nonparametric spectral density estimation has been proposed. The approach is based on a new evaluation criterion called autocorrelation mean square error (AMSE) for power spectral density (PSD) estimates of available finite length data. Minimization of this criterion not only provides the optimum segmentation for existing PSDE approaches, but also provides a new optimum windowing within the segments that can be combined additionally to the existing methods of nonparametric PSDE. Furthermore, the problem of frequency resolution in existing PSDE methods for noisy signals has been analyzed. In the existing approaches, the additive noise and the finiteness of data which are the causes of the original loss of the frequency resolution are not treated separately. The suggested new approach to spectrum estimation takes advantage of these two different causes of the problem and tackles the problem of resolution in two steps. First, the method optimally reduces noise interference with the signal via minimum noiseless description length (MNDL). The new power spectrum estimation MNDL-Periodogram of the denoised signal is then computed via conventional indirect periodogram to improve frequency resolution.

Acknowledgments

First of all, I am very much thankful to my *parents* for being the biggest reason of all my accomplishments to this day. They have not only supported me financially but also provided me with much moral support to allow me to get through this stage of my life with least trouble as possible. Thank You!

Thank you, *Kaberi*, my dear sister who has always believed in me more than I do and always pulled me right up when things were too overwhelming or stress drained me down. Thank you for always being there for me.

I cannot express my appreciation enough to my best friend *Leo*, who has been there for me since my undergraduates. Thank you for making the transition easier for me and always ready to lend me a helping hand or a shoulder to lean on during tough times. With much confidence I can now say that I am a stronger person thanks to you. I have learnt so many valuable things from you throughout the course of my studies - your friendship is a priceless treasure.

It would have been quite impossible to go through my masters if *Sumandeep* hasn't been there for me at every step of my way. From the very first day till the last - I cannot thank you enough for your generous help and delightful company.

I would also like to acknowledge and thank *Sina* for his uplifting conversations and consolations. Thank you, *Mohammad Reza* for your warm support since the first day. Also, thank you *Omid, Alon, Raymond, Huma, Sarah, Azadeh, Sanjay, Matija, Matt* and *ENG325* - many cherished memories reside with you.

Many thanks to the *Department of Electrical and Computer Engineering*; especially to the *Committee Members* for their precious time, valuable reviews and suggestions.

Also, sincere thanks to the faculty supervisor, *Dr. Beheshti*, for her guidance and advise throughout the course of my masters.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the most important application areas of digital signal processing is the power spectral density estimation (PSDE) of finite length data. Power spectral analysis is a signal processing method that characterizes the autocorrelation and frequency information of a measured signal [1].

The power spectrum represents information contained within the signal, which may be important to modeling and analyzing the phenomena which generated the signal. As such, this information has many useful purposes in engineering and related fields [3].

Spectral analysis applications cover a wide range of problems; one of the most important areas includes speech analysis. Speech analysis is performed for a variety of reasons, including phonetics research, understanding the speech production process, and speech modeling. Thus, it serves many applications in speech recognition and speech coding. Some problems of interest in speech analysis involve the construction of time series, spectral models and the detection of changes in the speech spectrum [9, 15].

Spectrum estimation also plays an important role in signal detection and tracking. In many applications, much interest lies in narrow-band signal detection which may be recorded in very noisy environment. Therefore, signal detection and frequency estimation become non-trivial problems that require robust, high-resolution spectrum estimation techniques [2].

Other applications of spectrum estimation include harmonic analysis and prediction, time-series extrapolation and interpolation, spectral smoothing, bandwidth compression,

beam-forming and direction finding [2, 9, 10].

Power spectrum estimation is statistically based and covers a variety of digital signal processing concepts, some of which are briefly summarized in chapter 2.

A signal encountered in practice is usually a sampled version of the original signal, hence its spectrum is most often a discrete function of frequency calculated via Discrete Fourier Transform (DFT) or Fast Fourier Transform (FFT). Among other factors, the resolution and accuracy of this discrete spectrum depends on the amount of data included in the DFT. Because the amount of data is ultimately finite, it is impossible to calculate the exact spectral representation of the signal and therefore it must be estimated. The finite length of a signal can introduce statistical uncertainty if the signal has been truncated and is not known for all time [3]. These uncertainties will be briefly introduced in the next chapter.

There are two methods available for the power spectral estimation of noisy signals: non-parametric and parametric power spectrum estimation methods. The non-parametric approaches or classical methods make no assumption on the structure of the signal. The well known periodogram is the fundamental method of this type which was first introduced by Schuster in 1898 citeproakis. However, due to the limited ability of the periodogram to produce an accurate estimate of the power spectrum, a number of modifications to the method have been proposed to improve its statistical properties. These methods were developed by Bartlett (1948), Blackman-Tukey (1958) and Welch (1967) [1, 2, 5]. These methods are averaging, modified versions of the periodogram as well as windowed autocorrelation estimates. On the other hand, parametric power spectral estimation uses finite number of parameters to model the spectral density of signals. The main focus of this thesis research is on the classical methods and some common existing approaches to parametric and nonparametric PSDE.

The main motivation of this research is the essential need to compare the nonparametric PSDE approaches not only by their asymptotic properties, but also by their properties based on the observed finite length data.

This report briefly introduces some of the statistical digital signal processing concepts

which are important to the computation of power spectrum estimation. Overview and limitations of the existing nonparametric PSDE methods are provided and some of the common parametric PSDE methods are introduced. Chapter 3 presents a new approach to PSDE via optimally windowing autocorrelation estimates of existing PSDE methods. The problem of frequency resolution is considered in Chapter 4 and a new approach to signal denoising prior to computation of PSDE is offered. Chapter 5 provides further insight into the application areas of PSDE and suggests some future research.

Chapter 2

Background

This chapter reviews general background information on spectrum analysis methods, focusing particularly on those areas pertaining to non-parametric methods.

2.1 Random Processes

Generally, signals and their associated spectra are usually categorized as random signals, due to the fact that either analytical models may not be available to determine the signal's structure or the underlying phenomena may be too complex. Most of the concepts and practice in spectrum analysis originated from a statistical analysis of random signals. Random signals are best represented in the correlation domain and spectral domain.

A random process is an ensemble collection of random signals. These random signals in the time domain are expressed in terms of statistical properties such as mean and variance which can be used to describe the probability density function (PDF) for the underlying random process. Because the entire set of random functions which could define a random signal is not usually available as this would require analysis of its joint probability distribution, the goal is to find the second order statistics of the random signal. Furthermore, these statistical properties cannot be computed exactly, but are rather estimated from finite length data.

In many applications, the statistical averages are determined by using time averages from

a single realization of the random process. If the entire signal is not available, then estimates of the true values, such as the mean and variance, are made using the available portion of the random signal. Correlation and spectral domain analysis of random signals are likewise based on the estimation of the true correlation function and spectrum of the random process. Spectral estimates are realized either from the magnitude of Fourier transforms applied to the finite length time signal or to the correlation function estimate.

Assuming that a stationary random sequence, $x[n]$ of finite length, where $-N \leq n \leq N$ and N is the length of the signal, is generated from a random process X. Therefore, its *mean* for all n is computed as

$$m_{x[n]} = E[x[n]] = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x[n]. \tag{2.1}$$

Since (2.1) is a precise representation of the mean value in the limit as N approaches infinity then

$$\hat{m}_{x[n]} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \tag{2.2}$$

may be regarded as a fairly accurate estimate of $m_{x[n]}$ for sufficiently large N [11].

Similarly, for this stationary random process, where its statistical properties do not vary with time, the *autocorrelation* is the expected value of the product of the random signal realization with a time-shifted version of itself, such that

$$R_{xx}[n, n+l] = R_{xx}[l] = E[X[n]X[n+l]] \tag{2.3}$$

and its *power spectrum density estimation* (PSDE) is simply the Fourier transform of the autocorrelation

$$P_{xx}(f) = \int_{l=-\infty}^{l=\infty} R_{xx}[l] e^{-j2\pi fl} dl. \tag{2.4}$$

However, for a majority of cases, sufficient information will not be available to build a complete function of the random signals for the above analysis. If this is the case, the known information about the function can be treated as a finite discrete signal in order to

estimate the autocorrelation. Representation of the discrete-time autocorrelation and PSD function for finite length real signals introduces the first problem of nonparametric PSDE which is explained further in Section 2.2. A few common properties used to characterize the estimator are defined below which include *bias*, *variance* and *consistency*. These properties are important to describe the *periodogram* in Section 2.2.

Assuming that autocorrelation function $r_{xx}^N[l]$ estimate is generated from finite length data $x[-N], ..., x[N]$ where $-N \leq n \leq N$, the following properties are defined:

2.1.1 Bias

The difference between the mean or expected value $E[\hat{r}_{xx}^N]$ of an estimate $\hat{r}_{xx}^N$ and its true value r_{xx}^N is called the bias.

$$B_{\hat{r}_{xx}^N} = r_{xx}^N - E[\hat{r}_{xx}^N] \tag{2.5}$$

Thus, if the mean of an estimate is equal to the true value, it is considered to be unbiased and having a bias value equal to zero. At any particular value of frequency, bias error describes how much the mean value of the spectral estimate differs from the true spectrum of the signal because of finite resolution in the dimensions of magnitude and frequency. The bias error is significantly influenced by truncation in the time domain, which results in reduced resolution in frequency and reduced accuracy in magnitude [3].

2.1.2 Variance

The variance of an estimator effectively measures the width of the probability density and is defined as

$$var_{\hat{r}_{xx}^N} = E[(\hat{r}_{xx}^N - E[\hat{r}_{xx}^N])^2] \tag{2.6}$$

A good estimator should have a small variance in addition to having a small bias suggesting that the probability density function is concentrated about its mean value.

2.1.3 Consistency

If the bias error and variance both approach zero as the limit approaches infinity, or the number of observations become large, the estimator is said to be consistent. Thus,

$$\lim_{N\to\infty} var_{\hat{r}_{xx}^N} = 0 \tag{2.7}$$

and,

$$\lim_{N\to\infty} B_{\hat{r}_{xx}^N} = 0. \tag{2.8}$$

This implies that the estimator converges in probability to the true value of the quantity being estimated as N becomes infinite [4].

2.2 Nonparametric PSDE: The Periodogram and its Variance Estimation

As mentioned earlier, non-parametric methods make no assumption on the structure of the signal. These methods are generally based on the mathematically equivalent paths of estimating spectra by either the indirect or the direct methods. The indirect method computes the spectral estimate by applying the Fourier Transform (F.T.) to the estimate of the autocorrelation function of the input signal, i.e. from the time domain signal, the method obtains the correlation domain signal and then transforms it into the spectral domain. In the direct method, the F.T. is applied to the time signal and the spectral estimate is computed by multiplication of the Fourier transform with its own complex conjugate, i.e. from time domain transform into the Fourier domain which is the spectral domain.

Going back to (2.3), since signals of finite duration are only available, the following biased autocorrelation function and its corresponding power spectrum are:

$$r_{xx}^{P'}[l] = \frac{1}{N-l}\sum_{n=0}^{N-l-1} x^*[n]x[n+l] \tag{2.9}$$

and,

$$P_{xx}^{P'}(f) = \sum_{l=-N+1}^{l=N-1} r_{xx}^{P'}[l]e^{-j2\pi fl}; \tag{2.10}$$

When computing the statistical characteristics of the above, the following is observed: The expected value $E[r_{xx}^{P'}[l]] = R_{xx}[l]$ which is the true autocorrelation derived in (2.3). Therefore, $r_{xx}^{P'}[l]$ is an unbiased estimate of $R_{xx}[l]$. Also, since

$$\lim_{N\to\infty} \text{var}[r_{xx}^{P'}[l]] = 0, \tag{2.11}$$

$r_{xx}^{P'}[l]$ is a consistent estimate of $R_{xx}[l]$. But for large l the variance of the above time-average autocorrelation function becomes very large, therefore a modified autocorrelation function is computed as follows:

$$r_{xx}^{P}[l] = \frac{1}{N} \sum_{n=0}^{N-l-1} x^*[n]x[n+l]. \tag{2.12}$$

and the corresponding power spectrum is the Fourier transform of the above such that

$$P_{xx}^{P}(f) = \sum_{l=-N+1}^{l=N-1} r_{xx}^{P}[l]e^{-j2\pi fl} \tag{2.13}$$

$$= \frac{1}{N}|\sum_{N-1}^{n=0} x[n]e^{-j2\pi fn}|^2 \tag{2.14}$$

$$= \frac{1}{N}|X(f)|^2 \tag{2.15}$$

But it is noted that,

$$E[r_{xx}^{P}[l]] = (1 - \frac{|l|}{N})R_{xx}[l] \tag{2.16}$$

however,

$$\lim_{N\to\infty} E[r_{xx}^{P}[l]] = R_{xx}[l] \tag{2.17}$$

Therefore $r_{xx}^{P}[l]$ is asymptotically unbiased and since

$$\lim_{N\to\infty} \text{var}[r_{xx}[l]] = 0, \tag{2.18}$$

$r_{xx}[l]$ is a consistent estimate of $R_{xx}[l]$.

So, now a windowed autocorrelation function results

$$\tilde{R}_{xx}[l] = (1 - \frac{|l|}{N})R_{xx}[l] \tag{2.19}$$

The mean of the estimated spectrum is

$$E[P_{xx}^P(f)] = \sum_{-\infty}^{\infty} \tilde{R}_{xx}[l]e^{-j2\pi f} \tag{2.20}$$

$$= \int_{-1/2}^{1/2} P_{xx}(\alpha)W_B(f-\alpha)d\alpha \tag{2.21}$$

which illustrates that the mean of the estimated spectrum is the convolution of the true power density spectrum $P_{xx}(f)$ with the Fourier transform $W_B(f)$ of the Bartlett window.

However, we observe that the estimated spectrum is asymptotically unbiased, such that

$$\lim_{N\to\infty} E[\sum_{-N+1}^{N-1} r_{xx}^P[l]e^{-j2\pi fl}] = \sum_{-\infty}^{\infty} R_{xx}e^{-j2\pi fl} \tag{2.22}$$

$$= P_{xx}(f) \tag{2.23}$$

But the variance of the estimate $P_{xx}^P(f)$ does not decay to zero as $N \to \infty$ since,

$$\lim_{N\to\infty} \text{var}[P_{xx}^P(f)] = P_{xx}^2(f) \tag{2.24}$$

and therefore the *periodogram* is not a consistent estimate of the true power density spectrum. In conclusion, the estimated autocorrelation $r_{xx}^P[l]$ is a consistent estimate of the true autocorrelation function $R_{xx}[l]$. However, its Fourier Transform $P_{xx}^P(f)$, the periodogram, is not a consistent estimate of the true power density spectrum. We observed that $P_{xx}^P(f)$ is an asymptotically unbiased estimate of $P_{xx}(f)$, but for a finite duration sequence, the mean value of $P_{xx}^P(f)$ contains a bias, which is evident as a distortion of the true power density spectrum. Thus the estimated spectrum suffers from the smoothing effects and the leakage embodied in the Bartlett window [1].

2.3 Nonparametric Methods for PSDE: Modifications to the Periodogram

Since the periodogram is not a consistent estimate of the true power density spectrum, several modifications to the periodogram have been proposed to improve its statistical properties due to its limited ability to produce an accurate estimate of the power spectrum. There are three well known methods available for the non-parametric approach for power spectrum estimation. All of these estimation techniques described in the following subsections reduce the variance of the periodogram's spectral estimate at the expense of decreasing the frequency resolution.

2.3.1 The Bartlett Method

The Bartlett method, also known as the averaging Periodogram, reduces the variance in the periodogram by subdividing the N-point sequence into **K** non-overlapping segments of length M. For each resulting K data segments $x_i[n]$ of length M, where

$$x_i[n] = x[n + iM], \quad i = 0, 1, ..., K-1, \quad n = 0, 1, ..., M-1 \tag{2.25}$$

the periodogram is computed such that,

$$P_{xx}^{(i)}(f) = \frac{1}{M} |\sum_{n=0}^{M-1} x_i[n] e^{-2\pi f n}|^2 \tag{2.26}$$

Finally, the Bartlett power spectrum estimate is obtained by averaging the periodograms for the K segments

$$P_{xx}^{B}(f) = \frac{1}{K} \sum_{i=0}^{K-1} P_{xx}^{(i)}(f) \tag{2.27}$$

The effect of reducing the length of the data from N to $M = N/K$ results in a window whose spectral width has been increased by a factor of K. Consequently, although variance is reduced, the frequency resolution has been reduced by a factor of K [1].

2.3.2 The Welch Method

Similar to the Bartlett method, the Welch method, also known as the averaging modified periodograms, allows data segmentation with overlap. For an allowance of 50% overlap, then $2 * K = L$ segments are obtained. The data segments of length M are represented such that

$$x_i[n] = x[n + iD] \quad i = 0, 1, ..., M-1 \;\; n = 0, 1, ..., L-1, \tag{2.28}$$

where, if $D = M/2$ for the 50% overlap between successive data segments, then $L = 2K$ segments are obtained. Furthermore, another modification made by Welch is to window the data segments prior to computing the periodogram such that

$$\tilde{P}_{xx}^{(i)}(f) = \frac{1}{MU}|\sum_{n=0}^{M-1} x_i[n]w[n]e^{-2\pi fn}|^2 \qquad i \;\; = 0, 1, ..., L-1, \tag{2.29}$$

where U is a normalization factor for the power in the window function, selected as

$$U = \frac{1}{M}\sum_{n=0}^{M-1} w^2[m]. \tag{2.30}$$

Finally, the Welch power spectrum estimate is the average of these modified periodograms,

$$P_{xx}^{W}(f) = \frac{1}{L}\sum_{i=0}^{L-1} \tilde{P}_{xx}^{(i)}(f). \tag{2.31}$$

2.3.3 The Blackman-Tukey Method

In this method, the sample autocorrelation is windowed first, and the Fourier transform is applied to yield the power spectrum estimate. The effect of windowing the autocorrelation is to smooth the periodogram estimate, therefore decreasing the variance in the estimate at the expense of reducing the resolution. The following computation shows the aforementioned effects:

The Blackman-Tukey estimate is

$$P_{xx}^{BT}(f) = \sum_{l=-(M-1)}^{M-1} r_{xx}[l]w[l]e^{-j2\pi fl} \tag{2.32}$$

where the window function $w[l]$ has length $2M-1$ and is zero for $|l| \geq M$. Again, the effect of windowing the autocorrelation is to smooth the periodogram estimate, thus decreasing the variance in the estimate at the expense of reducing the resolution.

2.4 Parametric Methods for PSDE

Parametric methods, unlike non-parametric, require some a *priori* information on how data is generated for extrapolation. This modeling approach eliminates the need for window functions and therefore avoids the problem of leakage and provides better frequency resolution than nonparametric methods. The two most common methods for parametric power spectrum estimation include the Yule-Walker method and Burg method for the Auto Regressive (AR) on model parameters [1]. Further explanation on these methods are not included since the problem in this research only deals with nonparametric approach.

2.5 Review Conclusion

As introduced in the previous section, existing PSDE approaches are only concerned with the periodograms inconsistent estimate of the power spectrum and hence offer different kinds of averaging and modifications to reduce the variance of the estimate. However, the main motivation of this work is the essential need for comparison of the nonparametric PSDE approaches not only by their asymptotic properties such as consistency, but also by their properties based on the observed finite length data. For example, it is known that the Welch approach is preferred over the Periodogram approach for its asymptotic properties. However, it is not known that for which data segmentation the performance using Welch method is optimum.

Furthermore, the problem of frequency resolution in non-parametric PSDE for noisy signals is considered. As seen previously, the finite length of data as well as the additive noise, both contribute to a decreased frequency resolution. The existing PSDE approaches offer different forms of averaging and windowing on the available data only to improve statistical properties of the estimates, however, at the expense of reducing frequency resolution. In these approaches, the additive noise and the finiteness of data which are the causes of the original loss of the frequency resolution are not treated separately. In chapter 4, it is illustrated that treating both noise corruption as well as the observed finite length data, improves frequency

resolution of PSDE.

Chapter 3

A New PSDE Approach: Optimum Windowing Autocorrelation Estimates

3.1 Introduction

From Chapter 2, it is observed that all existing nonparametric PSDE approaches are evaluated based on their asymptotic properties such as variance and consistency. However, it is also important to classify these methods by their properties based on the observed finite length data [10]. For example, it is known that the Welch approach is preferred over the periodogram approach for its asymptotic properties, i.e., variance of estimated spectrum reduces to zero as number of segments approaches infinity. However, it is not known for which data segmentation the performance of Welch method reaches optimum performance. This evokes the main motivation for the essential need of a new comparison criterion of the nonparametric PSDE methods. In this chapter we suggest using the autocorrelation mean square error (AMSE), which is the same as the PSD mean square error (PSD-MSE), as the comparison criterion. A novel method for estimating the desired criterion is provided by using the available observed data. The calculation approach follows the same fundamentals that are used for signal denoising in [7, 22] and for optimum order selection in parametric PSDE in [8, 21]. The introduction of this criterion also enables us to show that the averaging PSDE methods are improved through windowing where the optimum window length is obtained by minimizing the AMSE [14, 16].

3.2 Problem Statement

Consider a wide-sense stationary (WSS) random process X which has the following structure

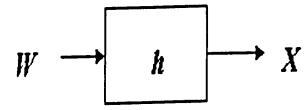

Figure 3.1: Random Process X

$$X = h * W \tag{3.1}$$

where W is a unit variance, white Gaussian process $N(0,1)$ and h is an LTI filter with real coefficients ('*' denotes the convolution operator).

The true autocorrelation $R_{xx}[l]$ and power density spectrum $P_{xx}(f)$ of this random process are

$$R_{xx}[l] = \sum_{n=-\infty}^{\infty} h[n]h[n+l]; \tag{3.2}$$

$$P_{xx}(f) = \sum_{l=-\infty}^{\infty} R_{xx}[l]e^{-j2\pi fl} \tag{3.3}$$

The new PSDE problem is to find the best estimate of the autocorrelation and PSD of X by using an available *finite* length sample of X of length N, $x[0], x[1], \cdots, x[N]$.

3.2.1 Existing Non-parametric PSDE Approaches

We already know that the well known nonparametric approaches include periodogram, modified periodogram, Bartlett, Welch, and Blackman-Tukey methods [1]. In these approaches the PSD is the Fourier Transform of windowed or segmented versions of their autocorrelation estimates. The autocorrelation estimate is obtained by summing the averages of the

deterministic autocorrelation of windowed versions of the observed sequence. For example, in periodogram the window co-efficients are all equal to one, where no averaging is involved, and thus the autocorrelation estimate is the following biased estimate

$$r_{xx}^{p}[l] = \frac{1}{N}\sum_{n=0}^{N-l-1} x^{*}[n]x[n+l], 0 \leq l < N \tag{3.4}$$

$$P_{xx}^{p}(f) = \sum_{-N+1}^{l=N-1} r_{xx}^{p}[l]e^{-j2\pi fl} \tag{3.5}$$

where $P_{xx}^{p}(f)$ is the PSD estimate. On the other hand, in the Welch method the data is divided into overlapping segments and the autocorrelation estimate is the average of the autocorrelations of those windowed segments. Details for choosing the biased periodogram estimate over the unbiased case are explained in Appendix A.

3.2.2 Evaluation of the Existing Methods

As mentioned earlier, the performance of the existing PSDE approaches is evaluated by the asymptotic behavior of the method's statistical properties such as variance and consistency. Important estimation errors in time and frequency domain are defined as

$$e_{xx}[l] = r_{xx}^{p}[l] - R_{xx}[l], \tag{3.6}$$

$$eP_{xx}(f) = P_{xx}^{p}(f) - P_{xx}(f) \tag{3.7}$$

It is observed that as the length of the data grows (i.e. as data length $\rightarrow \infty$), the bias and variance of the estimation error approaches zero. As such, consistency is proven. For example, the periodogram is asymptotically unbiased, i.e., $\lim_{N\rightarrow\infty} \mathrm{E}(e_{xx}[l]) = 0$. However, with this approach its error variance does not go to zero even as the length of data grows. On the other hand, methods such as Welch or Blackman-Tukey are not asymptotically unbiased. However, the error variance of these approaches goes to zero as the data length grows. This property of the latter approaches makes them much better PSDE method candidates than the periodogram approach in applications.

3.3 New PSDE Approach

The objective of the new PSDE approach is to estimate an optimum windowed version of existing averaging PSD estimate such as Welch, and Bartlett methods. However, to find the optimum window length, it is necessary to first introduce a new evaluation criterion for comparison of the PSDE methods.

3.3.1 New Evaluation criterion

The asymptotic behavior of estimation errors in (3.6, 3.7) have been studied extensively. However, no method is available to compare the methods for the available *finite* data non-asymptotically. For example, the choice of optimum segmentation in Bartlett and Welch methods is not available and we are not able to compare the performance of these methods for a given segmentation. To answer these questions, evaluate and compare PSDE methods, it is suggested to use a comparison criterion in the form of the mean-square of the estimation error denoted by autocorrelation mean square error (AMSE)

$$\mathrm{E}||e_{xx}^{L}||^2 = \mathrm{E}||eP_{xx}^{L}[e^{j\omega}]||^2 \tag{3.8}$$

where

$$e_{xx}^{L} = [e_{xx}[-L+1], \cdots, e_{xx}[0], \cdots, e_{xx}[L-1]] \tag{3.9}$$

and eP_{xx}^{L} is the MSE of the FFT of e_{xx}^{L}. The equality of AMSE and PSD-MSE domain holds true due to Parseval's Theorem. Note that the error, e_{xx}^{L}, is generated by segmented available data of length L where $0 \leq L \leq N$ with N being the total length of the data.

3.3.2 Optimum Windowing

In order to find the optimum window length L^*, the goal is to find an estimate of the introduced evaluation criterion in (3.8) for each window length. We divide the observed data to S segments of length L. The estimate of $R_{xx}[l]$ in (3.2) for the i^{th} segment is denoted by $\hat{r}_{xx}^{L}[l,i]$.

The evaluation criterion's estimate is provided with a new adaptive method based on the observed *finte* data. That is,

$$e_m^L[n] = \begin{cases} e_{xx}^L[n] & \text{if } -m \leq n \leq m, \\ 0 & \text{otherwise.} \end{cases} \tag{3.10}$$

Now consider a window of length $2m-1$ which is multiplied by autocorrelation estimate of observed data of length L. The goal is to find the optimum m for $1 \leq m \leq L$. In this case,

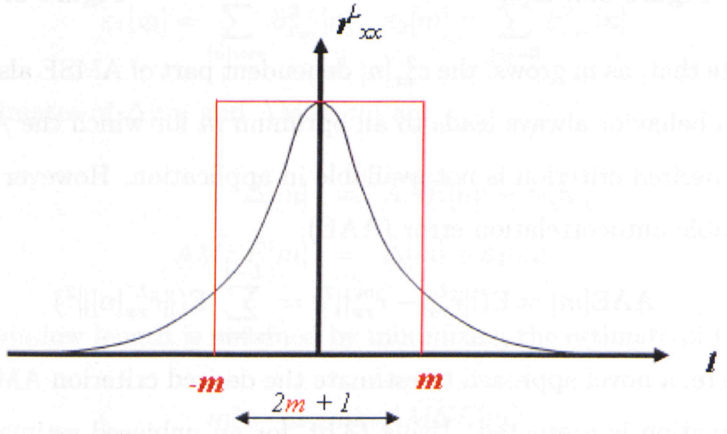

Figure 3.2: Autocorrelation Estimation $\hat{r}_{xx}^L$ windowed by $2m+1$ of $\hat{r}_{xx}^m$

the desired evaluation criterion, AMSE in (3.8), for each windowed autocorrelation estimate is

$$\text{AMSE}[m] = E(||R_{xx}^L - \hat{r}_{xx}^m||^2) = \varepsilon[m] + \Delta[m] \tag{3.11}$$

where $\hat{r}_{xx}^m$ is the m-windowed autocorrelation estimate and

$$\varepsilon[m] = \sum_{|n|=0}^{m-1} E(e_{xx}^2[n]), \tag{3.12}$$

$$\Delta[m] = \sum_{|n|=m}^{L-1} \|R_{xx}[n]\|^2. \tag{3.13}$$

We observe the following trends,

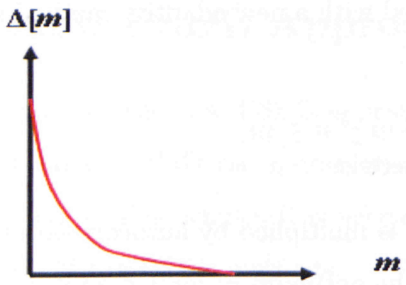

Figure 3.3: $\Delta[m]$

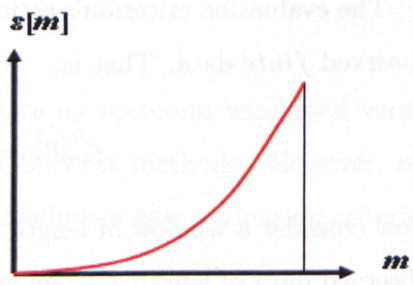

Figure 3.4: $\varepsilon[m]$

Here we note that, as m grows, the $e_{xx}^2[n]$ dependent part of AMSE also grows, while $\Delta[m]$ decreases. This behavior always leads to an optimum m for which the AMSE is minimized. Note that this desired criterion is not available in application. However, using the observed data, the available autocorrelation error (AAE):

$$\mathrm{AAE}[m] = \mathrm{E}(||\hat{r}_{xx}^{L} - \hat{r}_{xx}^{m}||^2) = \sum_{|n|=m}^{L-1} \mathrm{E}(\|\hat{r}_{xx}^{L}[n]\|^2) \tag{3.14}$$

is available. Here, a novel approach to estimate the desired criterion $\mathrm{AMSE}[m]$ by using this available information is presented. Using (3.6), for an unbiased estimator, $(E[e_{xx}^2[l]] = 0)$, the $\mathrm{AAE}[m]$ in (3.14) is

$$\mathrm{AAE}[m] = \varepsilon'[m] + \Delta[m] \tag{3.15}$$

where

$$\varepsilon'[m] = \sum_{|n|=m}^{L-1} E(e_{xx}^2[n]). \tag{3.16}$$

In the following section, an estimate of the e_{xx}^2 dependant parts in (3.12) and (3.16) are shown and can be provided based on the observed data. Therefore, an estimate of unavailable $\Delta[m]$ can calculated from (3.15) and substituted in (3.11).

3.3.3 Estimation of AMSE

Using the segmented data, the estimate of expected value and variance of $\hat{r}_{xx}^{L}$ are

$$\bar{\hat{r}}_{xx}[l] = \frac{1}{S}\sum_{i=1}^{S} \hat{r}_{xx}^{L}[l,i], \tag{3.17}$$

$$\bar{\sigma}^2_{\hat{r}_{xx}[l]} = \frac{1}{S}\sum_{i=1}^{S} \|\hat{r}^L_{xx}[l,i] - \bar{\hat{r}}_{xx}[l]\|^2 \tag{3.18}$$

and an estimate of AAE in (3.14) is

$$\widehat{\mathrm{AAE}}[m] = \frac{1}{S}\sum_{i=1}^{S} \|\hat{r}^L_{xx}[l,i] - \hat{r}^m_{xx}[l,i]\|^2 \tag{3.19}$$

The estimate of the e^2_{xx} dependent parts of AMSE and AAE in (3.15,3.11) are

$$\varepsilon_1[m] = \sum_{|n|=m}^{L-1} \bar{\sigma}^2_{\hat{r}_{xx}}[n], \quad \varepsilon_2[m] = \sum_{|n|=0}^{m} \bar{\sigma}^2_{\hat{r}_{xx}}[n] \tag{3.20}$$

Finally, the estimates of $\Delta[m]$ and $\mathrm{AMSE}[m]$ are

$$\hat{\Delta}[m] = \widehat{AAE}[m] - \varepsilon_1[m], \tag{3.21}$$

$$\widehat{AMSE}[m] = \hat{\Delta}[m] + \varepsilon_2[m] \tag{3.22}$$

The optimum window length is obtained by minimizing the estimate of the desired criterion

$$m^* = \arg\min_m \widehat{AMSE}[m] \tag{3.23}$$

Optimizing the desired criterion not only provides the optimum window length, but can also provide the optimum segmentation. This will be evident in the simulation result section.

3.4 Simulation Results and Discussion

To demonstrate the the effectiveness of the new approach, two sets of data, each of length 10,000 are used. The data is generated with (3.1). The first data set, x_1, is generated with a one-pole filter h_1 and the second data set, x_2 is generated with FIR Hamming filter h_2, of length 40:

$$h_1[n] = 0.75^{(n-1)}u[n] \tag{3.24}$$

$$h_2[n] = \text{Hamming filter of length 40} \tag{3.25}$$

where $u[n]$ is the unit step function.

3.4.1 Optimum Windowing for a Given Segmentation

Figure 3.5 shows both AMSE and $\widehat{AMSE}$ using Bartlett Windowing PSDE approach. The estimates are with 100 and 20 non-overlapping segments of x_2 with length 100 and 500 respectively.

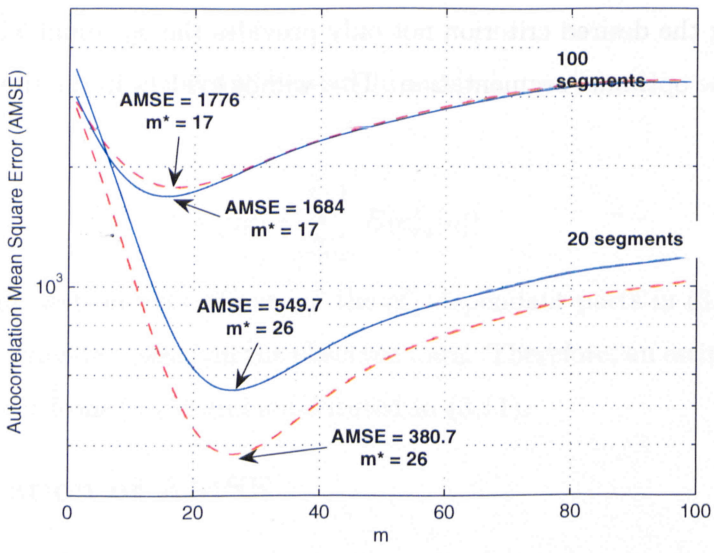

Figure 3.5: AMSE(dashed) and $\widehat{AMSE}$ (solid) for 100 and 20 segments of available data x_1 as a function of m where $2m-1$ is window's length

As the figure shows, the AMSE estimate in (3.22) is a successful estimate of AMSE in (3.11) which is the unknown desired criterion. More importantly, the optimum window length obtained by $\widehat{AMSE}$ is the same as the optimum m for the true AMSE. Note that while the optimum m for 20 segments is $m^* = 26$, the Bartlett method with no windowing is the same as Bartlett windowing when the window length is the segment's length, i.e., $m^B = L = 100$. Figure 3.6 shows the results of Welch windowing.

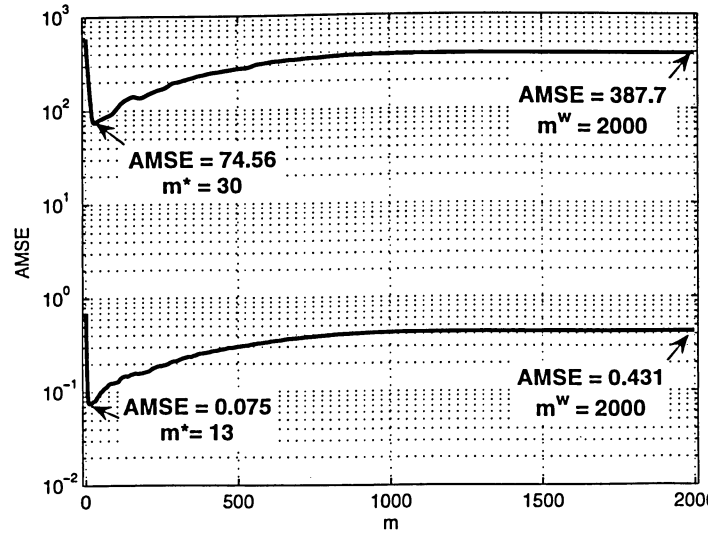

Figure 3.6: Optimum Hamming windowing of Welch method with 9 segments and 50% overlap. Upper curve is the AMSE using x_2 and lower curve using AMSE of x_1.

In this case, the AMSE estimate is minimized for $m^* = 13$ and $m^* = 30$ for x_1 and x_2 respectively. The figure shows the importance of windowing by providing the AMSE for different values of m. For example, the Welch method with no extra windowing, is equivalent to the choice of a window with $m^W = 2000$ which results in a very high AMSE of 387.7 for x_2. On the other hand, the minimum value of AMSE at $m^* = 30$ is 74.56. A window of length $2m^* - 1$ is applied to taper the r_{xx} before obtaining its PSD estimate. This provides a better autocorrelation and PSD estimate than Welch method alone with its original window length $(2(2000) - 1)$.

3.4.2 Optimum Windowing and Optimum Segmentation

Figure 3.7 illustrates $\widehat{AMSE}$ using hamming windowing and Welch approach for segments of "different length" with 50% overlap.

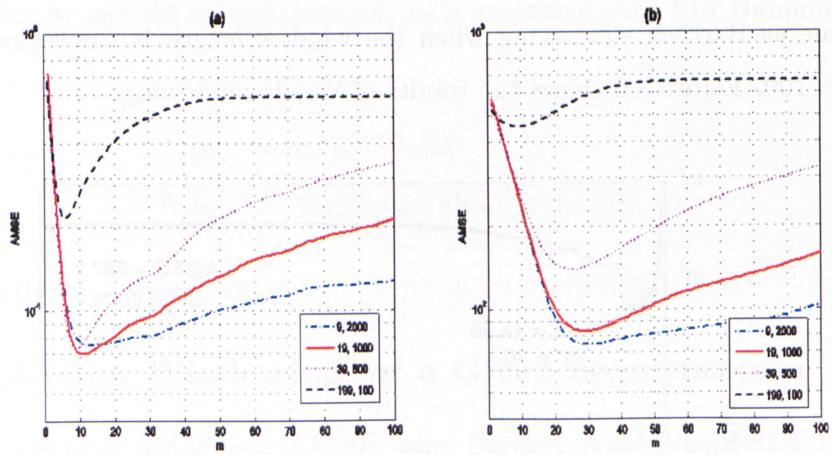

Figure 3.7: Welch with 50% overlap for x_1 (left) and x_2 (right) : displaying Segments: 9(solid), 19(dot-dashed), 39 (dotted) and 199(dashed)

As the figure shows, by minimizing the AMSE estimate, the optimum segmentation for x_1 is 19 with $m^* = 11$ and the optimum segmentation for x_2 is 9 with $m^* = 30$. Table 3.1 below illustrates all the results with the optimum segmentation.

Methods	(S, L)	AMSE$[L]$	m^*	AMSE$[m^*]$
Bartlett(x_1)	(20, 500)	2.754	12	0.2563
Welch(x_1)	(19, 1250)	0.4879	11	0.0698
Bartlett(x_2)	(20, 500)	2832	26	549.7
Welch(x_2)	(9, 2000)	387.7	30	74.56

Table 3.1: (S, L): Number of segments and length of segments. AMSE$[L]$: MSE of the Bartlett or Welch approach alone. AMSE$[m^*]$: MSE of the optimally windowed PSDE approach.

As it can be observed from the table, it is illustrated that optimum windowing method outperform the Bartlett and Welch approaches alone. Figure 3.8 shows the autocorrelation estimate of x_2 with the optimum Bartlett windowing.

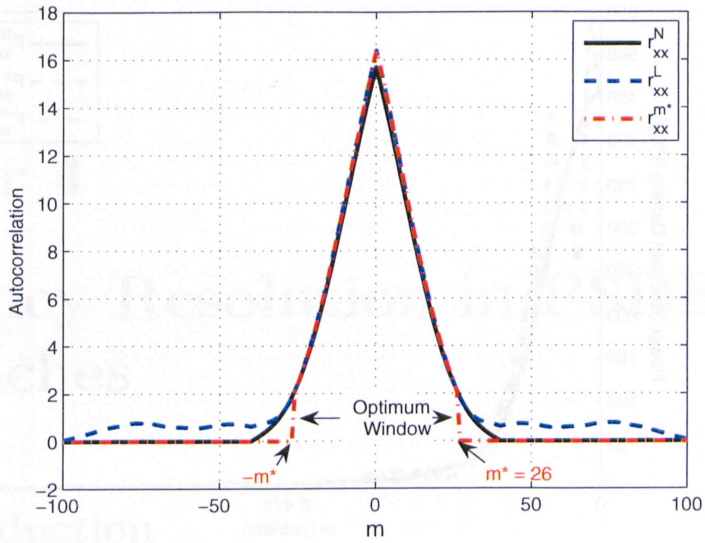

Figure 3.8: Autocorrelation of random process X_2 (solid line), Bartlett method with 20 segments (dashed line) and Bartlett optimum windowing $m^* = 26$.

From the above figure, it can be observed that only peak information around the centre is retained by the window while the rest of the tail of the autocorrelation estimate is truncated. Figure 3.9 shows the PSD estimate of x_2 with Bartlett windowing.

Note that the AMSEs in Table 3.1 are also equivalent PSD-MSE. Therefore, the MSE in the PSD of results in Figure 3.9 is the same as what the table displays. Therefore, in both Figure 3.8 and Figure 3.9, the AMSE and PSD-MSE of Bartlett method is 2832 where the PSD-MSE of the optimally windowed Bartlett is only 549.7!

3.5 Review Conclusion

In this chapter, a new evaluation criterion is introduced for non-parametric PSDE methods, other than the existing asymptotic criteria. The AMSE criterion is not only based on finite observation, but can also evaluate the consistency of the method as the data grows in size. The criterion guarantees minimizing the AMSE and PSD-MSE simultaneously. The new

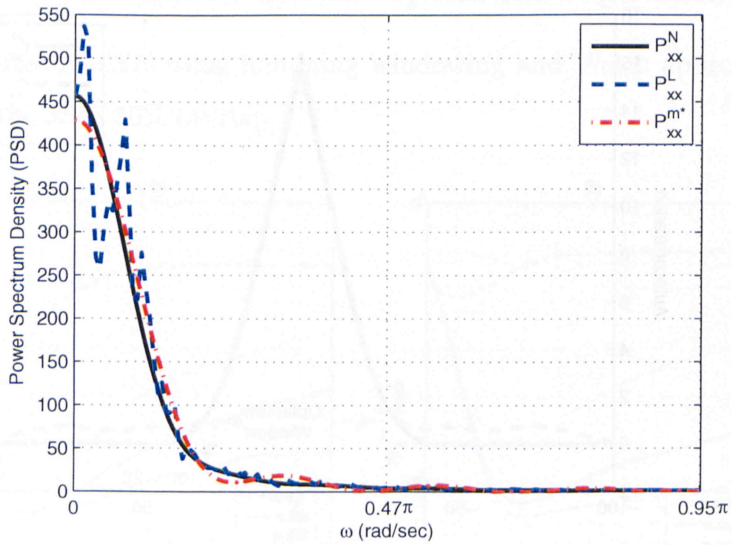

Figure 3.9: PSD of random process X_2 (dashed line), Bartlett and Bartlett optimum windowing PSD estimates using x_1.

PSDE approach based on this criterion provides not only the optimum windowing of the autocorrelation estimate, but also the optimum data segmentation for the averaging PSDE methods. For any averaging PSD approach such as the Welch or Bartlett methods, the new method performs better since the averaging PSD method is a special case of windowing with a maximum window length. The significance of the new PSDE approach is in the consistency of its evaluation criterion, the AMSE, and the use of the available finite length data to estimate the criterion.

Chapter 4

Frequency Resolution in PSDE Approaches

4.1 Introduction

Spectrum estimation plays an important role in signal detection and tracking. For example, sonar arrays are placed on an ocean floor to listen for narrow-band acoustic signals that are generated by the rotating machinery or propellers of a ship. Once a narrow-band signal is detected, the problem of interest is to estimate its center frequency in order to determine the ships direction or velocity [2]. Hence, in many applications such as these, much interest lies in narrow-band signal detection where the signal may have been recorded in very noisy environment. Therefore, signal detection and frequency estimation become non-trivial problems that require robust, high-resolution spectrum estimation techniques [2]. In this chapter, the PSDE problem for noisy signals is considered and a different approach to this problem is discussed [13].

Figure 4.1 illustrates the spectrum of an available noiseless sinusoid superimposed on the synthesized noisy signal. The frequency resolution of closely spaced spectra in the figure is affected by the noise and therefore imposes a need to overcome this disturbance.

From Chapter 2 it is known that in application, finite-length data is always dealt with and this already constitutes in a decreased frequency resolution of any signal of this property. Further windowing of the signal-data, as done by the existing modified versions of

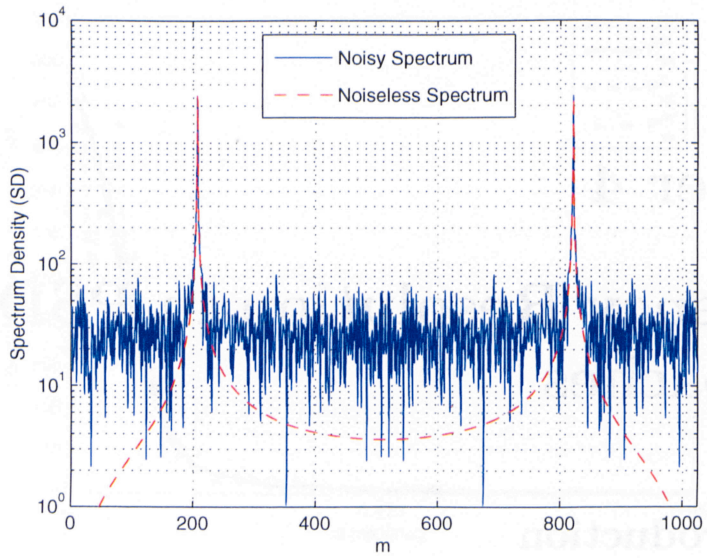

Figure 4.1: Noisy $Y_1(e^{j\omega_o n})$ and Noiseless $\bar{Y}_1(e^{j\omega_o n})$ PSD for $y_1[n]$

the periodogram (Bartlett, Welch and Blackman-Tukey) for improved spectrum estimation, consequently result in further deteriorated frequency resolution of the signal. The additive noise embedded in noisy signals is not considered distinctly in the existing PSDE methods which impacts frequency resolution of these noisy signals as well.

The classical methods introduced in Section 2.3 emphasize on obtaining a consistent estimate of the power spectrum through some averaging or smoothing operations performed directly on the periodogram or on the autocorrelation of the noisy data. Although the variance of the modified periodogram estimates is decreased, the effects of these operations are performed at the expense of reducing the frequency resolution. In section 4.2.1 these effects on frequency resolution for the existing nonparametric methods will be explored in greater detail.

Since frequency resolution of noisy signals is the main focus of this chapter, the problem of PSDE becomes two-fold. First, it is necessary to denoise the signal from its interfering background. Secondly, it is required to compute its power spectrum estimation, such that frequency resolution is not decreased due to further windowing of data as is the case in

existing methods. The proposed method chooses a new spectrum from the noisy spectrum via the minimum noiseless description length (MNDL) and fixed thresholding, similar to [7]. In section 4.3, a detailed calculation approach which follows the same fundamentals used for signal denoising in [7] and for optimum order selection in [8] will be provided. The denoised spectrum is then used to find the new PSDE of the signal via MNDL-Periodogram.

We further support our finding by illustrating frequency resolution reduction in classical averaging or modified periodograms for estimating power spectrum density. It will be observed that the novel approach for power spectrum estimation proposed maintains frequency resolution close to the original spectrum, which is also demonstrated in the simulation result section.

4.2 Problem Statement

In noisy signals, the additive noise, as well as the broadening of the spectrum being estimated due to windowing, are particularly a problem when it is desired to resolve signals with closely spaced frequency components.

Consider a noisy signal of the following form

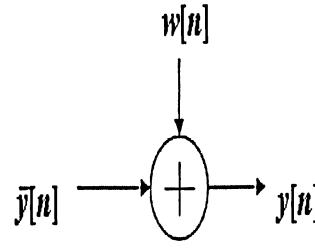

Figure 4.2: Noisy signal

$$y[n] = \bar{y}[n] + w[n] \tag{4.1}$$

for an available finite N length sample of Y, $\bar{y}[1], \bar{y}[2], \cdots, \bar{y}[N]$, where $w[n]$ is white Gaussian noise with variance σ_w^2 $(N(0,1))$. It should be noted that $\bar{y}(n)$ is the noiseless data.

The noisy spectrum $Y(f)$ is obtained via FFT such that

$$Y(e^{j\omega}) = \frac{1}{N}\sum_{n=0}^{N} y[n]e^{-j2\pi fn} \tag{4.2}$$

where $Y(f)$ is the Fourier Transform of noisy signal $y[n]$.

Here, the new PSDE problem is to first find the best denoised spectrum of $Y(f)$ and to use it to estimate PSD of the denoised signal such that frequency resolution is not affected by any additional windowing or modifications.

4.2.1 Review of Existing PSDE Methods

This section briefly summarizes the existing modifications to the periodogram, pertaining to the problem of frequency resolution, which have been proposed to improve only the statistical properties of the spectrum estimate. The effects of these modifications on frequency resolution of spectra are explained as follows: The Bartlett method, also known as averaging periodogram, allows data to be subdivided into smaller segments prior to computing the periodogram. The effect of reducing the length of data into shorter segments results in a window whose spectral width has been increased by a certain factor. Consequently, the frequency resolution is reduced by the same factor. Similarly, the Welch method, known as modified periodogram, allows data segments to not only overlap but also applies a window for variance reduction. Resolution in this case is not only window dependent, but also suffers from the same effects as the Bartlett method due to data segmentation. In the Blackman-Tukey method, the autocorrelation estimate is windowed first, prior to spectrum estimation computation. The effect of windowing the autocorrelation is to smooth the periodogram estimate and thus decreasing the variance in the estimate as a result. However, this is done at the expense of reducing the resolution since a smaller number of estimates are used to form the estimate of the power spectrum [1].

It is realized that the periodogram is only modified to improve its statistical properties at the cost of deteriorating frequency resolution, and therefore propose the new method in

the following section to overcome the problem of noise and windowing data via optimally denoising the signal first to improve the spectrums frequency resolution.

4.3 A New PSDE Approach: MNDL-Periodogram

In this approach, a new PSD estimate is obtained in the following two steps:

4.3.1 Spectrum Denoising

The considered noisy signal is first optimally denoised prior to its PSD computation. Consider the noisy signal $y[n]$ in (4.1). To evaluate the noisy spectrum, the FFT error for each n-point is an important unavailable factor, that is:

$$e[e^{j\omega_o n}] = Y(e^{j\omega_o n}) - \bar{Y}(e^{j\omega_o n}) \tag{4.3}$$

where $\bar{Y}(e^{j\omega_o n})$ is the noiseless spectrum and $\omega_o = \frac{2\pi}{N+1}$. After obtaining the N-point FFT of the signal, the absolute value of this FFT coefficients are sorted and thus the sorted versions are denoted $Y^s[n]$ and its associated denoised coefficients by $\bar{Y}^s[n]$.

The tail of the sorted spectrum is more affected by the noise than the FFT points with the highest absolute values. Therefore, for the denoising step, the goal is to choose the optimum number of these sorted noisy spectrum. For each value m, $0 \leq m \leq N$ the chosen noiseless FFT is

$$Y_m[n] = \begin{cases} Y^s[n] & \text{if } 0 \leq n \leq m, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

that represents the choice of the first m estimates and sets the rest of the FFT values to zero.

Now, the new error criterion becomes:

$$e_m Y[n] = \bar{Y}^s[n] - Y_m[n]. \tag{4.5}$$

Note that n in this error criterion and n in (4.3) possibly represent two different frequencies. The goal here is to obtain the optimum value of m which results in the minimum spectrum

mean-square estimation error (SMSE). This criterion's estimate is provided with a new adaptive method based on the observed noisy spectrum such that

$$\begin{aligned} \text{SMSE}[m] &= E(||\bar{Y}^s[n] - Y_m[n]||^2) && (4.6) \\ &= \sum_{n=0}^{m-1} E(||e_m Y[n]||^2) + \Delta[m] && (4.7) \end{aligned}$$

where

$$\Delta[m] = \sum_{n=m}^{N-1} ||\bar{Y}^s[n]||^2. \tag{4.8}$$

SMSE is similar to the MNDL criterion introduced in [7] where the objective is to minimize the error or noise between noisy and noiseless data or spectrum in this case. Note that as m grows, the $e_m Y[n]$ dependent part of SMSE grows, while $\Delta[m]$ decreases. This always leads to an optimum m for which the SMSE is minimized. This desired criterion is not available; however, from the observed data, an estimate of the following available spectrum error (ASE) is available:

$$\text{ASE}[m] = \sum_{n=0}^{N-1} \text{E}(||Y^s[n] - Y_m[n]||^2). \tag{4.9}$$

Here, we present a novel approach which uses the available spectrum error to provide an estimate of the desired criterion SMSE$[m]$.

Using (4.6) for an unbiased estimator ($E[e_m Y[n]] = 0$), then $ASE[m]$ in (4.9) becomes

$$\text{ASE}[m] = \sum_{n=m}^{N-1} E[||e_m Y[n]||^2] + \Delta[m] \tag{4.10}$$

It will now be shown that an estimate of the $e_m Y[n]^2$ dependent parts in (4.8) and (4.10) can be provided, based of the observed data. Therefore, an estimate of unavailable $\Delta[m]$ can be calculated from (4.8) and substituted back in (4.6).

The estimate of the $e_m Y[n]^2$ dependent parts of SMSE and ASE in (4.8, 4.10) are calculated as

$$\varepsilon_{1u}[m] = \sum_{|n|=m}^{N-1} (N - |n|)\sigma_w^2, \quad \varepsilon_{2u}[m] = \sum_{|n|=0}^{m} |n|\sigma_w^2 \tag{4.11}$$

Furthermore, variance of the additive noise $w[n]$ is also calculated and sorted in descending order, keeping large noise variance at the front and low variance towards the tail such that:

$$\sigma_{ws}^2[n] = \frac{1}{N}\sum_{i=1}^{S}(||w_i[n] - \bar{w}_i[n]||^2) \tag{4.12}$$

where S is the number of additive noise samples and $\bar{w}_i[n]$ represents the mean of that i^{th} noise sample. Sorted estimates of the dependent parts in (4.11) now become

$$\varepsilon_{1s}[m] = \frac{1}{N}\sum_{|n|=m}^{N}\sigma_{ws}^2, \quad \varepsilon_{2s}[m] = \sum_{|n|=0}^{m}|n|\sigma_{ws}^2 \tag{4.13}$$

Finally, the estimates of $\Delta[m]$ and SMSE$[m]$ are

$$\hat{\Delta}[m] = \widehat{ASE}[m] - \varepsilon_{1x}[m], \tag{4.14}$$

$$\widehat{SMSE}[m] = \hat{\Delta}[m] + \varepsilon_{2x}[m] \tag{4.15}$$

where $x = u$ for unsorted estimates and $x = s$ for the sorted estimates.

The optimum window length for obtaining an acceptable noiseless spectrum in this case is

$$m^* = \arg\min_m \widehat{SMSE}[m] \tag{4.16}$$

Minimization of the estimated desired criterion above provides optimum denoising of noisy signal for this case.

Unknown Noise Variance

So far, the estimation calculation is considered for a noisy signal of known variance. However, in application, this information is not available. It is observed that the estimates of ASE and Δ for different noise variance which is illustrated in Figure 4.3 below

It is observed that by increasing the noise variance over 1.5, this allows the estimate to become negative, thereby setting a threshold for the unknown noise variance.

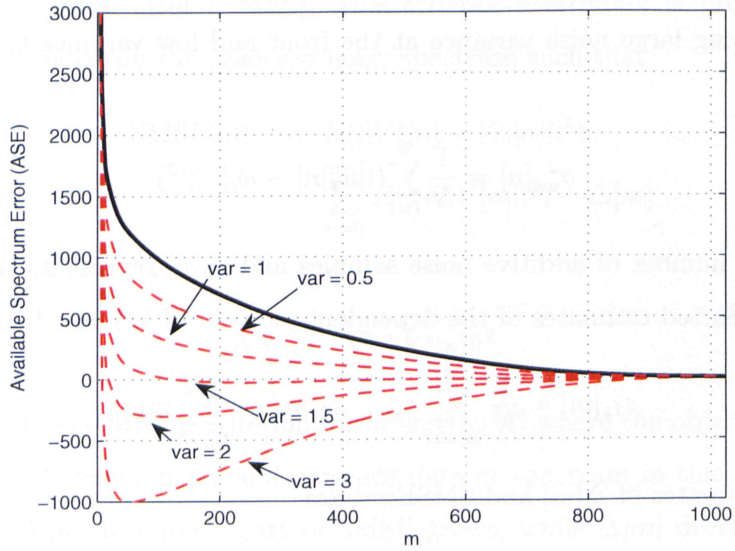

Figure 4.3: True ASE (solid) and $\hat{ASE}$ (dashed) for difference noise variances.

4.3.2 PSDE of Denoised Spectrum

The PSD estimate of the denoised spectrum is obtained via the conventional indirect periodogram. That is, the denoised time-domain signal $y^{MNDL}[n]$ is obtained via m-point IFFT of the denoised spectrum, $Y^{MNDL}(e^{j\omega_o})$, by choosing m^* frequency components. Computation of the new MNDL-Periodogram ($P^{MNDL}(e^{j\omega_o})$) is as follows:

$$P^{MNDL}(e^{j\omega_o}) = \sum_{m=-(N-1)}^{N-1} r_{yy}^{MNDL}[m]e^{-j\omega_o m} \tag{4.17}$$

where $r_{yy}^{MNDL}[m]$ is the autocorrelation of denoised time-domain signal $y^{MDNL}[n]$.

4.4 Simulation Results

For the results, two sets of noisy sinusoidal signals of length $N = 1024$ are generated with (4.1) as follows:

$$y_1[n] = 5sin[0.4\pi n] + w[n] \tag{4.18}$$

$$y_2[n] = 5sin[0.4\pi n] + 5sin[0.41\pi n] + w[n] \tag{4.19}$$

where $w[n]$ is the added white Gaussian noise with unit variance ($\sigma_w^2 = 1$).

For both the $y_1[n]$ and $y_2[n]$ noisy signals above, the noisy spectrum, ($Y_1(e^{j\omega_o n})$ and $Y_2(e^{j\omega_o n})$), and corresponding noiseless spectrum, ($\bar{Y}_1(e^{j\omega_o n})$ and $\bar{Y}_2(e^{j\omega_o n})$) are obtained. Note that here the noiseless spectrums, which are obtained by simply taking the spectrum of pure sinusoids, is computed only as a comparison criterion to verify against the estimated noiseless spectrums in the simulations. Figure 4.4 shows the power spectrum for $y_1[n]$ noisy signal $Y_1(e^{j\omega_o n})$ and its corresponding noiseless spectrum $\bar{Y}_1(e^{j\omega_o n})$.

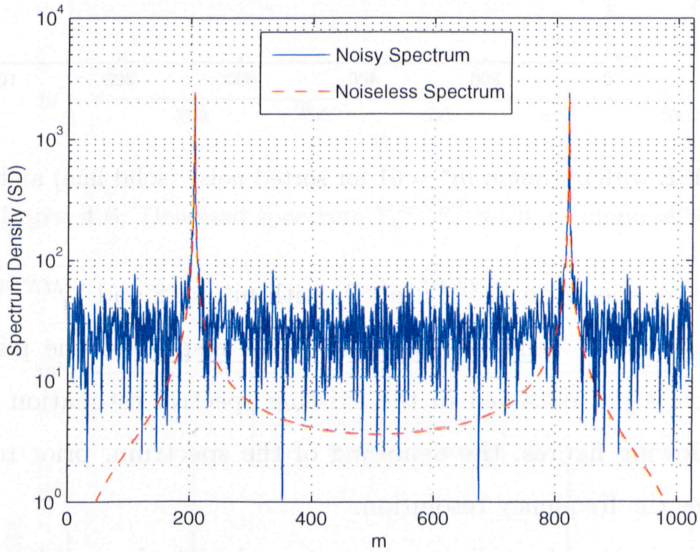

Figure 4.4: Noisy $Y_1(e^{j\omega_o n})$ and Noiseless $\bar{Y}_1(e^{j\omega_o n})$ PSD for $y_1[n]$

4.4.1 Optimum Spectrum Denoising

In Figure 4.5the estimated $\widehat{\text{SMSE}}$ criterion for $y_1[n]$ is shown. Also, it is observed that the optimum minimum occurs at $m^* = 61$ for sorted noise variance and $m^* = 437$ for unsorted noise, while the desired SMSE in (4.6) selects m^* at 18.

Similarly, for $y_2[n]$ the optimum minimum occurs at $m^* = 105$ for sorted noise variance while the desired SMSE in 4.6 selects m^* at 26. However, in Figure 4.6 and 4.7 it is noted that

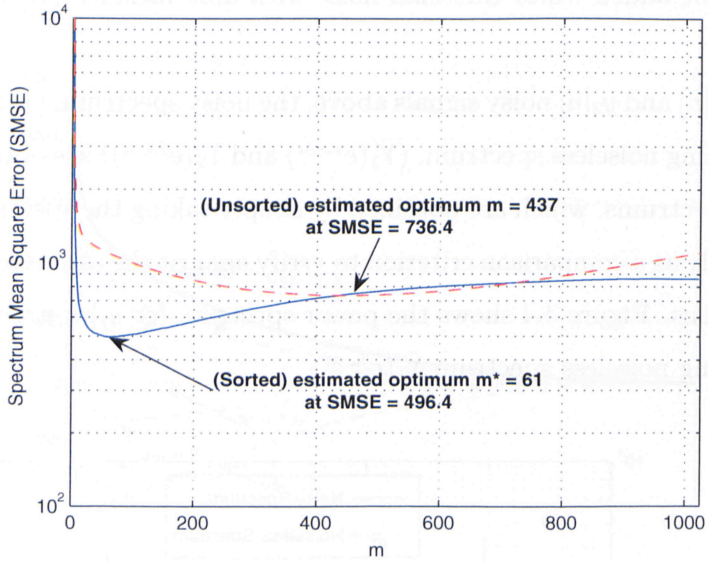

Figure 4.5: $\widehat{\text{SMSE}}$ with optimum $m^* = 61$ for sorted noise (solid line) and $m^* = 437$ for unsorted noise (dashed line).

the denoised power spectrum $Y_1^{MNDL}(e^{j\omega_o n})$ with $m^* = 18$ and $Y_1^{MNDL}(e^{j\omega_o n})$ with $m^* = 61$ are successful estimated denoised spectrums with respect to the pure $\bar{Y}_1(e^{j\omega_o n})$ noiseless spectrum, which is not obtained in the existing spectral estimation methods. As will be observed in following figures, the denoising of the spectrum, prior to spectral estimation, greatly enhances the frequency resolution.

The thresholding in the above figures is compared with the well-known hard thresholding approach, Donoho and Johnstone thresholding of $\sigma_w\sqrt{2\log N} = 2373$ [18, 19] and MDL thresholding of $\sigma_w\sqrt{\log N} = 1678$ [20]. However, these thresholds are much worse than the estimated threshold obtained by $\widehat{\text{SMSE}}$ as they pick up most of the noise. Therefore, the optimum m from (4.16) provides the best denoised spectrum.

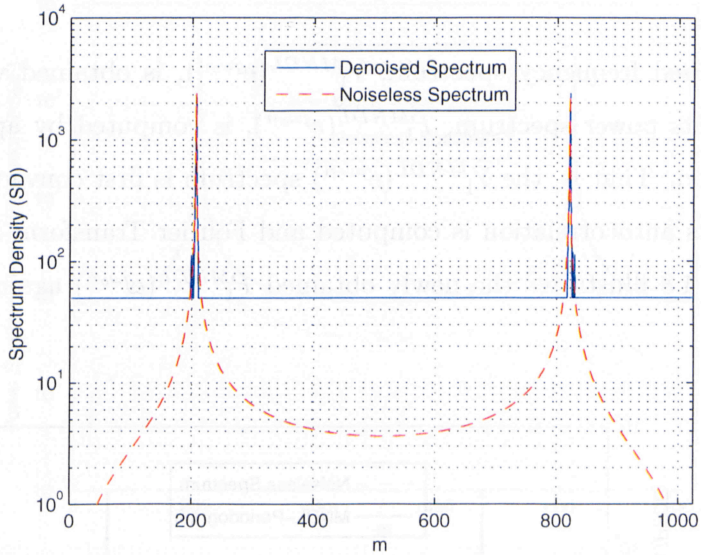

Figure 4.6: Denoised spectrum Y_1^{MNDL} with optimum m = 18.

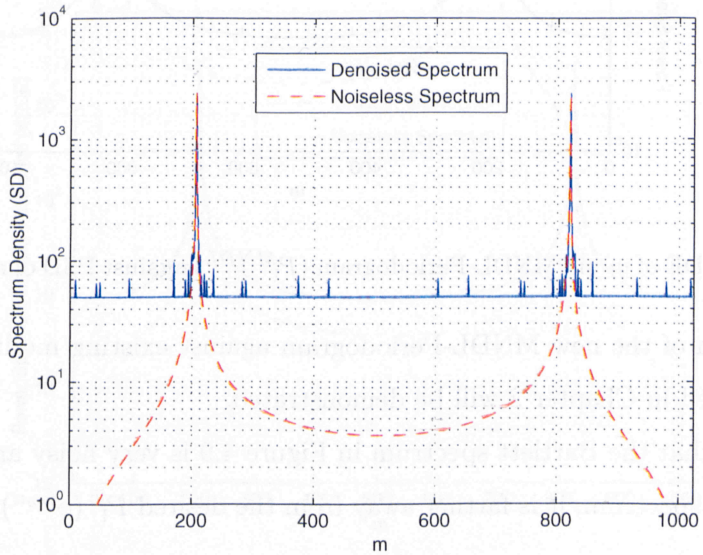

Figure 4.7: Denoised spectrum Y_1^{MNDL} with optimum m = 61.

4.4.2 PSDE comparison: MNDL-Periodogram vs. existing methods

After the denoised frequency spectrum, $Y_1^{MNDL}(e^{j\omega_o n})$, is obtained via the minimum m^* of $\widehat{SMSE}[m]$, its power spectrum, $P_1^{MNDL}(e^{j\omega_o n})$, is computed by applying indirect periodogram method. That is, the $Y_1^{MNDL}(e^{j\omega_o n})$ spectrum is first converted into time-domain signal. Then its autocorrelation is computed and Fourier Transform is found according to (4.17). Figure 4.8 illustrates the newly obtained $P_1^{MNDL}(e^{j\omega_o n})$ against the noiseless PSD $P_{\bar{y}_1}(e^{j\omega_o n})$.

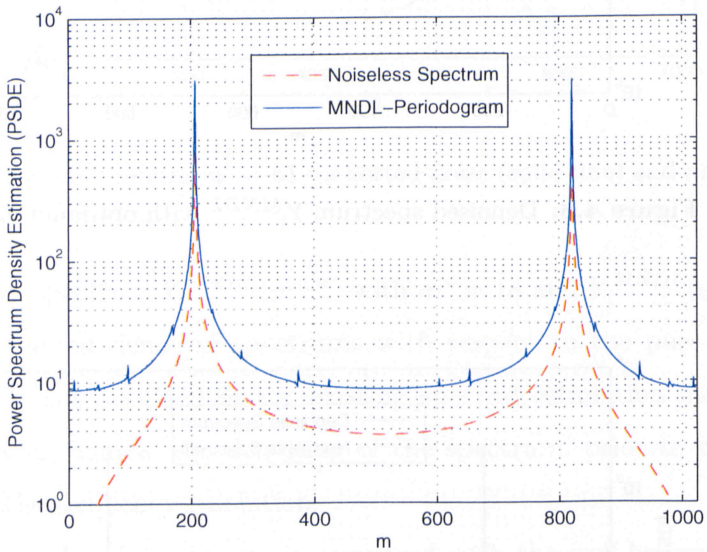

Figure 4.8: PSDE of y_1 via MNDL-Periodogram (P_1^{MNDL}) against PSD of noiseless signal ($P_{\bar{y}_1}(f)$)

Comparison of the new MNDL-Periodogram against existing modified periodogram, already mentioned in Chapter 2 will be demonstrated.

It is observed that the Bartlett spectrum in Figure 4.9 is very noisy and relative to MNDL-Periodogram's spectrum, it is farther away from the desired $P_1^{\bar{y}_1}(e^{j\omega_o n})$ noiseless power spectrum.

The Welch spectrum in Figure 4.10 is even further away from the desired $P_1^{\bar{y}_1}(e^{j\omega_o n})$ noiseless power spectrum. However its noise level is better than the Bartlett spectrum and therefore

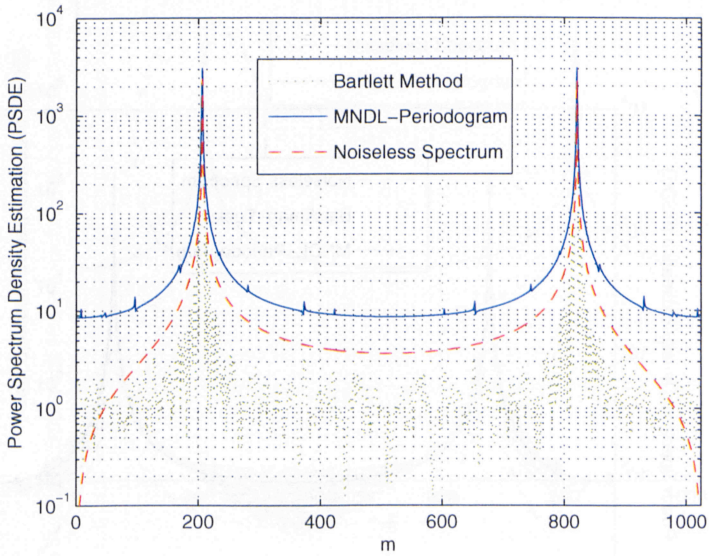

Figure 4.9: PSDE of y_1 via Bartlett (P_1^B) and MNDL-Periodogram (P_1^{MNDL}) against PSD of noiseless signal ($P_1^{\tilde{y}_1}$)

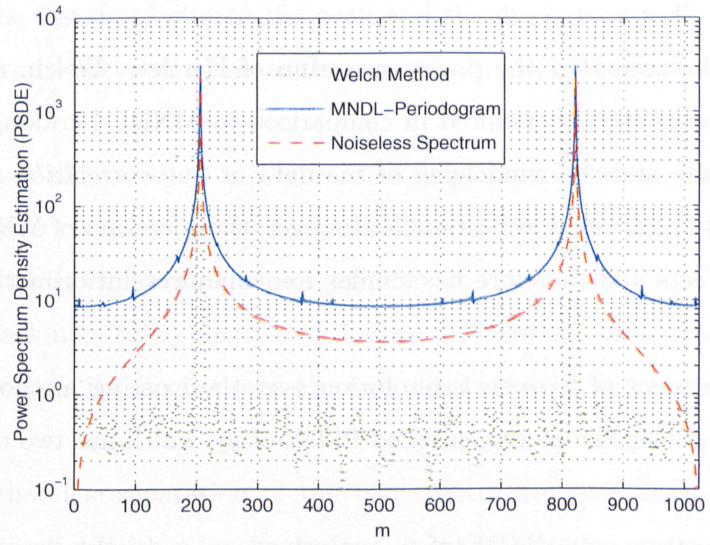

Figure 4.10: PSDE of y_1 via Welch (P_1^W) and MNDL-Periodogram (P_1^{MNDL}) against PSD of noiseless signal ($P_1^{\tilde{y}_1}$)

may provide better frequency resolution.

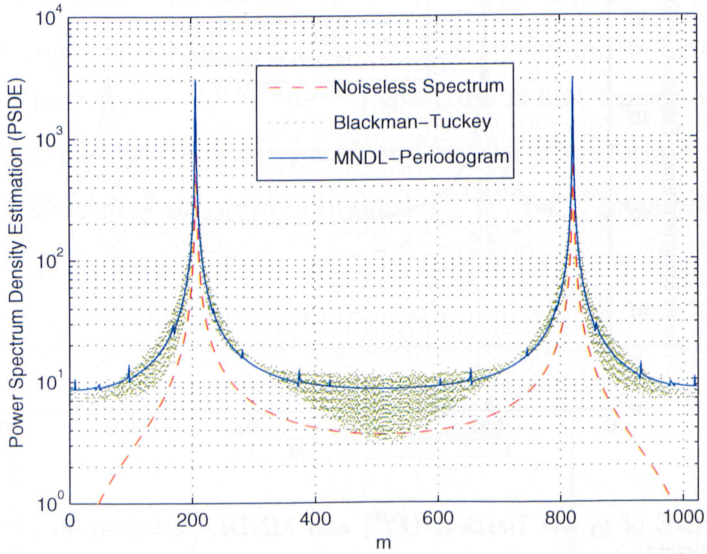

Figure 4.11: PSDE of y_1 via Blackman-Tuckey (P_1^{BT}) and MNDL-Periodogram (P_1^{MNDL}) against PSD of noiseless signal ($P_1^{\bar{y}_1}$)

It can clearly be seen that the power spectrum of Bartlett, Welch, and Blackman-Tukey are definitely broadened/ corrupted in comparison to MNDL-Periodogram. This is due to the existing noise and the windowing of the data or autocorrelation function of the noisy signal as already mentioned in Section 2. However, the spectrum of MNDL-Periodogram has no modified effects and therefore has similar frequency resolution with respect to noiseless spectrum.

The improvement of frequency resolution over the existing methods is further demonstrated by computing the aforementioned PSDE of y_2, which has two closely spaced spectra in the signal.

Figure 4.12 illustrates the PSDE of y_2 performed on y_1 in the previous figure. The two closely spaced spectra are encircled and zoomed-in versions of the two peaks are shown in figures below.Further examination of the two closely spaced spectra in the following zoomed

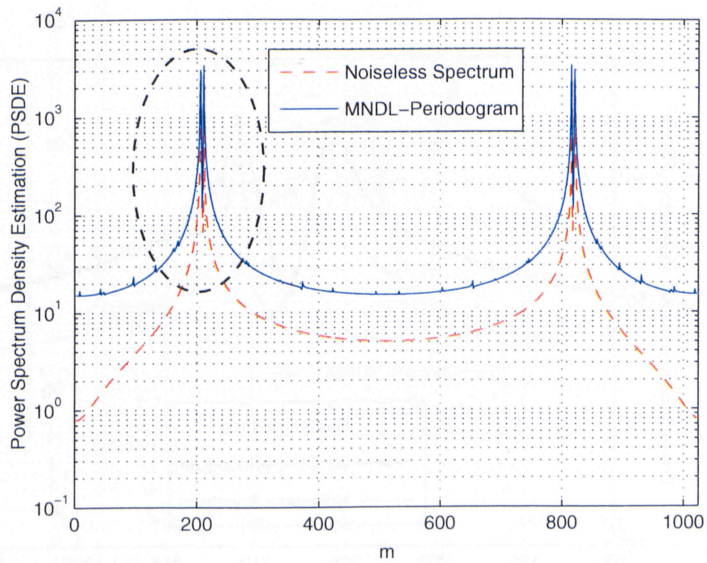

Figure 4.12: PSDE of y_2 via MNDL-Periodogram(P_2^{MNDL}) against PSD of noiseless signal ($P_2^{\tilde{y}_2}$)

in figures display the Bartlett, Welch and Blackman-Tukey methods, each against MNDL-Periodogram and available noiseless PSD.

In Figure 4.13 the distance between the peak and the dip is very well distinguished in the MNDL-Periodogram spectrum where-as the same resolution is not obtained with the Bartlett approach.

Furthermore, the Welch method in Figure 4.14 displays even worse results where the frequency resolution is completely lost due to same data segmentation as in Bartlett and in addition to a hamming windowing of the segments.

Although the Blackman-Tukey method in Figure 4.15 obtains closer resolution as MNDL-Periodogram, however its dip between the two closely spaced spectra is not as distinct as the latter.

We observe that the separation of the two spectra are best distinguished in MNDL-Periodogram estimation with a longer dip and closer to the expected noiseless power spectrum than the existing PSDE shown by Bartlett, Welch, and Blackman-Tukey. Clearly this novel approach can be used for better signal detection and spectrum estimation, especially

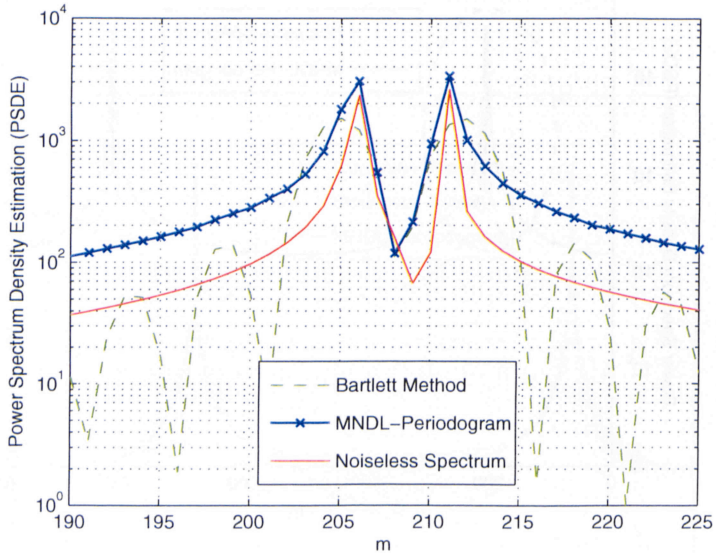

Figure 4.13: Bartlett spectrum with zoomed in two closely spaced spectra

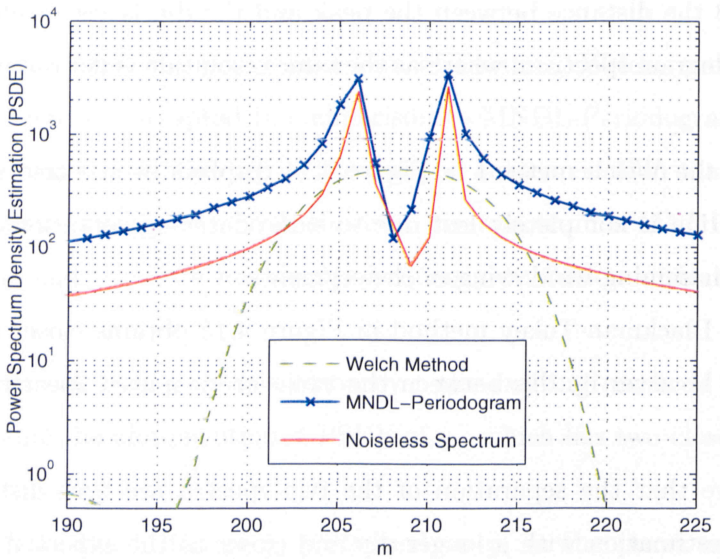

Figure 4.14: Welch spectrum with zoomed in two closely spaced spectra

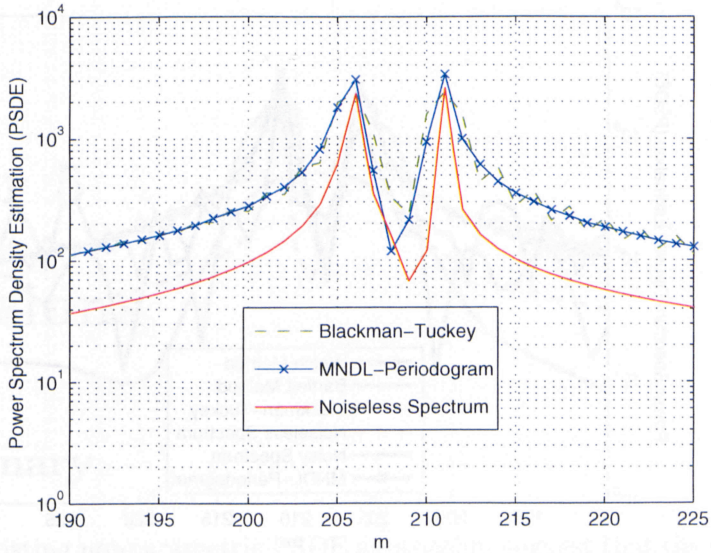

Figure 4.15: Blackman-Tukey spectrum with zoomed in two closely spaced spectra in noisy environments.

4.5 Review Conclusion

In this chapter, the imposed problem of all existing modifications on the periodogram that have been proposed in order to improve only the statistical properties of the spectrum estimate at the cost of frequency resolution were acknowledged. The PSDE problem with presence of additive noise was also addressed. In this case, windowing the available data in the existing nonparametric methods clearly demonstrates the decreased resolution in PSD estimates of noisy signals. This decrease, which is due to both the finiteness of the data and the presence of the additive noise, is improved by the MNDL-Periodogram. The approach first denoises the data and then estimates the PSD. In the first step, minimizing the estimated spectrum mean square error (SMSE) provides the best denoised spectrum in comparison to other thresholding criteria. Next, a simple periodogram approach is used for PSD estimation. This novel approach maintains frequency resolution close to the original noiseless

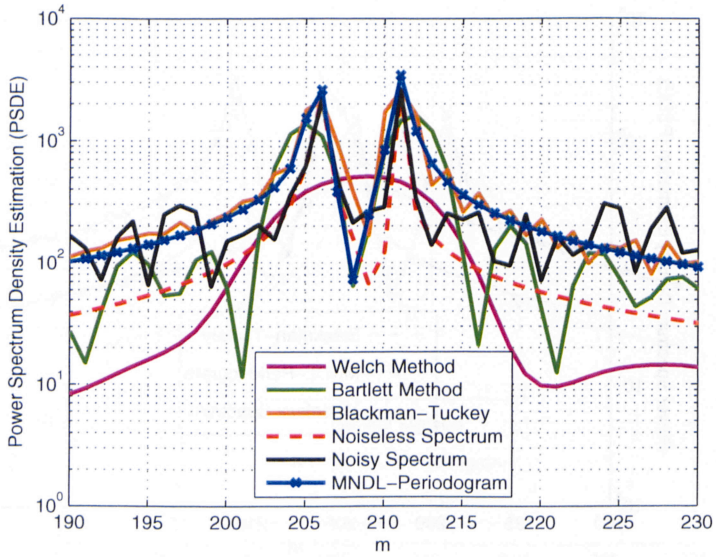

Figure 4.16: Comparison with all existing methods

spectrum as no additional windowing of signal is applied and, as a result, outperforms the existing approaches.

Chapter 5

Conclusions

5.1 Summary

Analysis of the existing nonparametric PSDE approaches suggest that the comparison of performance of these methods is only based on the asymptotic behavior of the obtained spectral estimates. Furthermore, there is no defined procedure to obtain the optimum number of segmentation for the averaging PSDE methods. Moreover, the existing spectrum estimation for noisy signal do not account for the loss of frequency resolution of the estimate which is caused by the averaging and windowing effects during the estimation process.

In this work we have provided a new approach to PSD estimation. A new evaluation criterion for non-parametric PSDE is estimated based on the available observed data. The criterion guarantees minimizing the autocorrelation MSE and PSD-MSE simultaneously. The new PSDE approach based on this criterion not only provides the optimum windowing of the autocorrelation estimate, but also the optimum segmentation for the averaging PSDE methods. For any of the averaging PSD approaches such as the Welch or Bartlett methods, the new method is observed to perform better since the existing averaging PSD methods are a special case of windowing with a maximum window length versus the optimum window length obtained via AMSE criterion.

The problem of frequency resolution reduction in the existing PSDE approaches for noisy signals is imposed by the modifications on the periodogram which have been proposed to improve the statistical properties of the spectrum estimate. For example, windowing the

available data as done in the existing nonparametric methods clearly demonstrates the decreased resolution in PSD estimates of noisy signals. This decrease is not only due to the finiteness of the data but also because of the presence of additive noise. Therefore, the new approach minimizes the loss via MNDL-Periodogram. This approach first denoises the noisy data and then estimates the PSD. In the first step, the estimated spectrum mean square error (SMSE) is minimized and provides the best denoised spectrum in comparison to other thresholding criteria. Next, a simple periodogram approach is used for PSD estimation. This novel approach maintains frequency resolution close to the original noiseless spectrum as no additional windowing of signal is applied and, as a result, outperforms the existing approaches.

5.2 Future Research

For the existing averaging PSDE methods such as the Bartlett and Welch methods, the number of data segmentation has been optimized in order to obtain good estimates of the power spectrum. However, for the Blackman-Tukey method, which applies a window to its autocorrelation estimate prior to obtaining the PSD estimate, a definite or optimum size of that window is still not definite and can be the subject of future research. This optimum window size will aid to ideally smooth the periodogram estimate and thus, decrease the variance of the estimate. Furthermore, although illustratively where the improvement in frequency resolution of PSD estimates via the new approach MNDL-Periodogram has been shown, there have been no closed-form expressions of the resolution provided which would help better comprehend and compare the performance of the new and existing PSDE methods with each other.

Appendix A

Appendix A

A.1 Biased and Unbiased Variance Estimation

The standard normal distribution is defined as having a mean of zero and standard deviation equal to one. Over many trials, it is observed that the mean of a random normal deviates indeed from the expected mean of zero. The standard deviations of the random normal also deviates from the expected standard deviation of one.

However, as the sample size becomes very small, the standard deviations of the random normal deviates are consistently less than one, even though their means correctly approximated the expected mean of zero. The true variance, also known as the biased variance, is defined as

$$\sigma_b^2 = \frac{1}{n}\sum x^2 \tag{A.1}$$

A question naturally arises whether some index other than the true variance could approximate the expected value better. Since the expected standard deviation is consistently underestimated only for the small values, a prime candidate for a new (unbiased) variance index is defined as

$$\sigma_u^2 = \frac{1}{n-1}\sum x^2 \tag{A.2}$$

since division by a smaller value makes the value of the fraction larger. A minute decrement of the n by the -1 seemed logical, since for large n, division by n or by $n-1$ makes for a

small, negligible increase of the value of the fraction. However, for the small n, the increase of the value of a fraction with n decremented by 1, can be large.

By changing the definition of the variance in such a way that the sum of the deviation scores is divided by $n-1$, the standard deviations of the random normal deviates start to approximate the expected values of one even for the small sample sizes. This new index is called the unbiased variance [?].

For example, the distribution of variances of n samples from a normal distribution has a "chi-square" distribution with n-1 degrees of freedom. Thus, in order to have a measure of variance that converges precisely on σ^2 for a normal distribution, we have to divide by $n-1$ instead of n. In other words, we have to use the unbiased estimate given by equation A.2 [?]

Bibliography

[1] J.G. Proakis, D.G. Manolakis. *Digital Signal Processing, Principles. Algorithms, and Applications*, Prentice-Hall, Inc, 2002.

[2] J.G. Hayes. *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons Inc, 2003.

[3] J.F. Dahl. Time Aliasing Methods for Spectrum Estimation (Thesis), Brigham Young University, 2003.

[4] National Semiconductor Corporation. Power Spectra Estimation, Application Note (AN) 255., 1980.

[5] S.K. Mitra. *Digital Signal Processing*, McGraw-Hill, 2006.

[6] I.S. Konvalinka. Iterative Nonparametric Spectrum Estimation IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-32, No.1, February 1984

[7] S. Beheshti, M.A. Daleh. A New Information Theorectic Approach to Signal Denoising and Best Basis Selection, IEEE Trans, on Signal Processing, Vol 53, No. 10, pp. 3613-3624, October 2005

[8] S. Beheshti. A New Approach to Order Selection and Parametric Spectrum Estimation, Acoustics, Speech and Signal Processing Conference Proceedings (ICASSP) Vol. 3, pp. 21-24, May 2006

[9] G.M. Jenkins, D.G. Watts. *Spectral Analysis and its Applications* Holden-Day, Inc, San Francisco, 1968.

[10] N. Sandgren, P. Stoica. On nonparametric estimation of 2-D smooth spectra IEEE Signal Processing Letters, Vol.13 (10), pp. 632-635, October 2006.

[11] D.S. Kim, S.Y. Lee, K. Lee. Emperical conditional mean: Nonparametric estimator for comparametric exposure compensation., Acoustics, Speech and Signal Processing Conference Proceedings (ICASSP) Vol. 2, pp. 957-960, May 2006

[12] A. V. Oppenheim, A. S. Willsky, S. Hamid Nawab. *Signals & Systems*, Prentice Hall, 1996.

[13] S. Pal, S. Beheshti. A new look at frequency resolution in power spectral density estimation, to appear in proceeding of IEEE Computational Intelligence in Image and Signal Processing (CIISP), April 2007.

[14] S. Beheshti, S. Pal. Optimum Segmentation and Windowing in Nonparametric Power Spectral Density Estimation, submitted to 15th International Conference on Digital Signal Processing, July 2007.

[15] T. Lobos, Z. Leonowicz, J. Rezmer, P. Schenger. High-Resolution Spectrum-Estimation Methods for Signal Analysis in Power Systems., IEEE Transactions on Instrumentation and Measurement, vol. 55 (1), pp. 219-225, February 2006

[16] M.G. Amin. Power Spectrum Estimation with Partially Known Autocorrelation Function., Proceedings of the IEEE, vol. 79, No.1, Jan. 1988.

[17] J. Mathews. A Unified Approach to Nonparametric Spectrum Estimation Algorithms., IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSO-35, No.3, February 1987

[18] D. Donoho. Denoising by soft thresholding., IEEE Trans. Inf. Theory, vol. 41, pp. 613 627, 1995.

[19] D. Donoho, I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage., Biometrika, pp. 425455, 1994

[20] J. Rissanen. Minimum description length denoising., IEEE Trans. Inf. Theory, vol. 46, no. 7, pp. 2537 2543, November 2000.

[21] S. Beheshti, M.A. Dahleh. LTI systems, additive noise and order estimation, Conference on Decision and Control., Conference on Decision and Control, Volume 6, pp.6491-6496, 2003.

[22] S. Beheshti, M.A. Dahleh. A new information theoretic approach to order estimation problem., Proceedings of the IFAC Symposium on System Identification, 2003