INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

INVESTIGATION INTO STEADY-STATE AUDITORY BRAINSTEM RESPONSE DETECTION: WEIGHTED TIME AVERAGING AND AUTOREGRESSIVE SPECTRAL ESTIMATION

by

Tony R. Orsi

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
and the Institute of Biomedical Engineering
University of Toronto

INVESTIGATION INTO STEADY-STATE AUDITORY BRAINSTEM RESPONSE DETECTION: WEIGHTED TIME AVERAGING AND AUTOREGRESSIVE SPECTRAL ESTIMATION

Master of Applied Science 1998

Tony R. Orsi

Graduate Department of Electrical and Computer Engineering

and the Institute of Biomedical Engineering

University of Toronto

ABSTRACT

Steady-**S**tate **A**uditory **B**rainstem **R**esponses (SSABRs) are auditory evoked potentials which provide objective, non-invasive, audiometric information. Unfortunately, test time is too long to warrant clinical use. Thus, the purpose of this thesis is to reduce test time. We determined that the long test time is due to high EEG background noise, idiosyncratic physiological artifacts, the use of Artifact Rejection under noisy recording conditions and periodogram based spectral estimation which suffers from poor resolution.

To improve detection, we compared Weighted Time Averaging with Artifact Rejection and a detector using Autoregressive spectral estimation with the **M**agnitude **S**quared **C**oherence (MSC) method (currently the best SSABR detector). Both comparisons used data recorded under quiet and noisy conditions. We found that Weighted Time Averaging performed as well as Artifact Rejection meanwhile reducing recording time by 36% for noisy conditions. Furthermore, the MSC method had 5% greater ROC Area (on average) compared to the Autoregressive based detector.

ACKNOWLEDGMENTS

I would like to thank Professor Kunov for his guidance, friendship and financial support. I am also grateful to Professor Madsen for his friendship, technical assistance and constant words of wisdom. Last but not least, Professor Hatzinakos has also helped me a great deal with shaping my research and with sorting out various signal processing issues in my work. I would also like to thank NSERC for the PGS A Scholarship that they awarded me.

Many thanks also go out to the staff and students at the Institute of Biomedical Engineering and Biomaterials and the Department of Electrical and Computer Engineering. Thanks go out to Professors Dolan and Eizenman, Anne Mitchell and Sarah Cherian for their support and guidance. Thanks to Thas Yuwaraj and Dave Purcell for computer assistance, Steve Davies and Nathanael Kuehner for signal processing assistance, Nic Schiopu and Xinde Li for hardware assistance and George Fung, Peter Picton, Elad Sagi, Taha Jaffer and Martin Pienkowski for fun times at the acoustics lab.

Finally and certainly not least, I would like to thank my family for their constant support and encouragement. They helped make my life easier while I worked on my research. I would also like to thank God for giving me the strength and courage to accomplish my research goals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TERMS AND ABBREVIATIONS

Abbreviation	Meaning
α	false positive level
A_c	Amplitude of carrier waveform
A/D	Analog to Digital (Conversion)
ABR	Auditory Brainstem Response
ACF	Autocorrelation Function
ACNS	Auditory Central Nervous System
AEP	Auditory Evoked Potential
AFR	Artifact Rejection
AIC	Akaike Information Criteria AR model order estimator
AM	Amplitude Modulated
AR	Autoregressive
AR-SSABR	AR-Based SSABR Detector
ARMA	Autoregressive Moving Average
ARPE	AR Parameter Estimator
BIC	Modified Akaike Information Criteria AR model order estimator
BM	Basilar Membrane
BPF	Bandpass Filter
CMMR	Common Mode Rejection Ratio
CSM_{crit}	threshold used for CSM detection statistic
CT^2	Circular T^2 statistic
CSM	Component Synchrony Measure
D/A	Digital to Analog Conversion
df	degrees of freedom
DC	Direct Current (or zero frequency)
DRMSE	Difference in RMSE between WTA and STA methods
DSNR	Difference in SNR between WTA and STA methods
ECG	Electrocardiogram

Abbreviation	Meaning
EEG	Electroencephalogram
EP	Evoked Potential
f_c	carrier frequency in SSABR stimulus
f_m	modulation frequency in SSABR stimulus
f_s	sampling frequency
$F_{critical}$	detection threshold for F test
FAR	False Alarm Rate
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
HL	Hearing Level
HOS	Higher Order Statistics
HR	Hit Rate
HR(0.05)	Hit Rate at a False Alarm rate of 5%
Hz	Hertz
IHCs	Inner Hair Cells
IIR	Infinite Impulse Response
KS	Kolmogorov-Smirnov
L_c	Stimulus Intensity for SSABR stimulus
LPF	Lowpass Filter
MSC_{WTA}	Magnitude Squared Coherence with WTA *within each* subaverage
$MSC\text{-}WA_{WTA}$	Magnitude Squared Coherence with Weighted Averaging *on each* subaverage and WTA *within each* subaverage
MA	Moving Average
MCM	Modified Covariance Method AR parameter estimator
MDL	Minimum Description Length AR model order estimator
ME	Minimum Energy WTA method
MMO	Maximum Model Order
MO	Magnitude Only
MOE	Model Order Estimator
MP	Mean of the Power of the subaverages in the MSC method

Abbreviation	Meaning
MSC	Magnitude Squared Coherence
MSC-WA	Magnitude Squared Coherence with Weighted Averaging
MSE	Mean Squared Error
naps	the number of time averages for each subaverage in the MSC method
nHL	normal Hearing Level
NIH	National Institute of Health (a U.S. based organization)
NPFFWC	No pre-filtering for weight calculation
OAEs	Otoacoustic Emissions
OHCs	Outer Hair Cells
pdf	probability distribution function
Pa	Pascals
PC	Phase Coherence
PFFWC	Pre-filtering for Weight Calculation
PM	Power of the Mean of the subaverages in the MSC method
PSE	Parametric Spectral Estimation
PV	Phase Variance
RA	Roc Area
ROC	Receiver Operator Characteristic
RMS	Root Mean Square
RMSE	Root Mean Square Error
SNRp	Pseudo-SNR AR model order estimator
SC	Successive Correlation
SE	Standard Error
SL	Sensation Level
SNR	Signal to Noise Ratio
SPL	Sound Pressure Level
SSABR	Steady State Auditory Brainstem Response
SSABR-Burg	AR-based SSABR detector based on the Burg ARPE
SSABR-MCM	AR-based SSABR detector based on the MCM ARPE

Abbreviation	Meaning
SSABR-PER	SSABR detector based on the Periodogram
STA	Synchronous Time Averaging
SVD	Singular Value Decomposition
TM	Tectorial Membrane
μ	modulation index used in SSABR stimulus
V^{-1}	Inverse Variance WTA method
VEP	Visual Evoked Potential
VNF	Vector Nonlinear Filtering WTA method
WA	Weighted Averaging
WTA	Weighted Time Averaging
YW2	Yule-Walker (Autocorrelation) AR parameter estimator
YW4	HOS-based Yule-Walker AR parameter estimator based on 4^{th} order cumulants

1.0 INTRODUCTION

This thesis will attempt to improve the signal processing techniques used in the detection of **S**teady **S**tate **A**uditory **B**rainstem **R**esponses (SSABRs) in order to reduce the amount of recording time needed to detect the SSABR. SSABRs are one type of auditory evoked potential in which an acoustical stimulus elicits a neural response that is recorded using 3 scalp electrodes. SSABRs are important because they provide an indication of the integrity of the peripheral and central auditory systems and can be used to assess an individual's auditory threshold at a specific audiometric frequency.

One major application area for SSABRs is neonatal screening. Approximately 100 in 100,000 children are born with a significant hearing loss (defined as a hearing loss in both ears averaging greater than 50 dB across 0.5 to 4 kHz) each year and 30 more will develop a hearing loss within the first few years of life (Lins et al., 1996). This will compromise the child's ability to learn and develop his/her communication skills. For this reason universal screening of all newborn infants is desired (NIH Consensus Statement, 1993).

These screening tests must be objective (since a newborn can not yet communicate), non-invasive (does not require surgery) and fast so that many children can be screened. Because behavioral audiometry is not accurate before six months of age and click evoked **A**uditory **B**rainstem **R**esponse (ABR) testing is not reliable for low frequencies, SSABRs have been proposed for hearing screening (Lins et al, 1996). However, the testing procedure must be fast and herein lies the problem; SSABR detection with current signal processing techniques requires long recording times.

This chapter will further discuss the motivation behind this thesis, define the problems which must be overcome and clearly state the research objectives. An outline of the chapters and appendices in this thesis will be given. To conclude, the research contributions of this thesis will be briefly discussed.

1.1 Why are SSABRs Important?

SSABR measurement is likely to become an important tool for audiology clinics and hospitals. This audiometric test is frequency specific, allowing a particular portion of the audiometric range (250 to 8000 Hz) to be tested. In addition, SSABR measurement will

complement other existing audiometric tests since SSABRs depend on the status of the cochlea and the auditory brainstem, meanwhile other tests such as Otoacoustic Emissions depend only on the status of the outer hair cells of the cochlea. SSABR testing is also objective and non-invasive making this test suitable for screening infants and school children for hearing loss. The test may even be applied to workers at risk for hearing loss.

1.2 Obstacles in SSABR Testing

SSABRs typically have a very low **S**ignal-to-**N**oise **R**atio (SNR). Pseudo-real data using real EEG background noise and simulated sinusoids (to represent the SSABR; see Appendix L) indicate that 40 Hz SSABRs have an SNR range of -6 dB to -30 dB and 80 Hz SSABRs have an SNR range of -15.4 dB to -27.4 dB. These measurements were taken over 1 second data intervals using a 10 Hz bandwidth for the noise power (i.e. 35 to 45 Hz for 40 Hz SSABRs and 75 to 85 Hz for 80 Hz SSABRs). In addition, the recording is very susceptible to contamination from physiological artifacts which is noise generated by the test subject due to movement. To detect the response, techniques based on artifact rejection and time averaging are employed, however in the case of a restless subject, artifact rejection requires long recording times due to the excessive amount of recorded data which is rejected due to artifact contamination. Furthermore, the method of time averaging requires substantial data acquisition (i.e. must record for a long time) due to the low SNR. For example, at moderate to high stimulus intensities the SSABR will be detectable within 10 to 30 seconds for some subjects, meanwhile at low stimulus intensities all subjects will require recording times on the order of 3 to 4 minutes (at least). This long recording time hinders the viability of SSABRs for use in clinical settings. Thus it is desirable to detect the response as fast as possible with a low probability of false detection (i.e. detecting noise as a response).

1.3 SSABR Research Objectives

The theme of this research is to reduce the amount of time that it takes to correctly identify an SSABR in background EEG noise. To facilitate this goal the following sub-objectives have been identified:

1. Identify an alternative method to Artifact Rejection which will eliminate artifact

contamination while not requiring an excessive amount of recording time.

2. Examine the SSABR signal and noise properties and current SSABR detection algorithms to determine techniques which can be used for improved SSABR detection. Design a new SSABR detector based on these techniques.
3. Compare the new detector with Magnitude Squared Coherence (the current benchmark in SSABR detection) using **R**eceiver **O**perator **C**haracteristic (ROC) analysis.

1.4 Thesis Outline

Chapter 2 will discuss the relevant background information including a brief introduction to auditory physiology, hearing disorders and current audiometric tests for hearing loss. Auditory Brainstem Responses will then be briefly discussed followed by a more in depth analysis of SSABRs. Chapter 3 then examines the noise in SSABR recordings as well as current SSABR detection algorithms. Based on this analysis, an SSABR model will be put forth and new techniques to improve detection will be proposed: Weighted Time Averaging and AR spectral estimation. In addition, modifications to the MSC technique will be proposed.

Chapter 4 then examines the use of **W**eighted **T**ime **A**veraging (WTA) as a possible data preprocessing replacement for **A**rtifact **R**ejection (AFR) and **S**ynchronous **T**ime **A**veraging (STA). To facilitate this examination, four WTA algorithms from the literature will be compared on pseudo-real data recorded under noisy and quiet conditions from 5 subjects. The following questions will then be answered: Can WTA be successfully applied to SSABRs? If so, how much of a performance improvement can WTA provide over the combination of Artifact Rejection and Synchronous Time Averaging? Furthermore, can WTA improve the SSABR SNR more than STA with AFR can on data recorded under quiet conditions?

In Chapter 5 we will investigate the use of AR spectral estimation for SSABR detection. To begin, background will be given on parametric time series modeling with emphasis on **A**utoregressive (AR) Modeling. A detection algorithm based on AR parameter estimation will then be proposed. ROC analysis will be used to choose the different parameters of the detector in order to optimize detection performance. In Chapter 6, the performance of the modified MSC methods will be investigated and compared to the AR-based SSABR detector. In addition, the computational complexity of the different SSABR detectors will be compared. Conclusions and recommendations for future work will then be presented in Chapter 7.

The Appendices also contain relevant information, some of which is referred to many times throughout the thesis. Appendix A provides a brief epidemiological study on the occurrence of retrocochlear lesions (a further reason why SSABRs are needed) in different subject groups. Appendix B describes work in the literature on characterizing the EEG probability density function. Appendix C describes the experimental setup and instrumentation used to acquire experimental data for this thesis. Appendix D describes the calculations done in obtaining the computational complexity of the different WTA methods which were investigated in this thesis. Appendix E describes ROC theory and the statistical tests used in this thesis for ROC comparison of different detectors. Appendices F and G describe the parameters chosen for synchronous time averaging and bandpass filtering. Appendix H briefly describes the sources and magnitudes of different instrumentation noise in our experimental setup. Appendix I describes the parameters chosen for the MSC detectors used in this thesis and Appendix J provides more in depth information on the ROC results for the AR-based SSABR detectors. Appendix K shows the calculations done to obtain the computational complexity for each SSABR detector. Appendix L discusses the SNR range for 40 Hz and 80 Hz SSABRs.

1.5 Research Contributions

One contribution of this thesis involves the investigation of different **W**eighted **T**ime **A**veraging (WTA) algorithms (Chapter 4) as possible data preprocessing replacements for the combination of **Art**ifact **R**ejection (AFR) and **S**ynchronous **T**ime **A**veraging (STA). Results showed that for 40 Hz SSABRs (section 4.3.6) and for 80 Hz SSABRs (section 4.3.7), WTA based on the Inverse Variance method performed just as well as the method of STA with AFR while yielding a 36% reduction in recording time when testing restless subjects (section 4.3.10). The other contributions of the thesis were the modifications of the MSC method (section 6.2) and the development of an AR-based SSABR detector (Chapter 5). ROC analysis (section 6.4) showed that AR-based SSABR detection is inferior to the MSC based detectors in all test conditions. Furthermore, we found that the best detector under both quiet and noisy recording conditions, was the modified MSC-WA method with WTA done *within each* subaverage.

2.0 BACKGROUND

This chapter begins with a brief overview on human auditory physiology followed by a discussion on the different types of hearing loss and the current audiometric tests used to detect them. Auditory Brainstem Responses will then be briefly discussed followed by SSABRs which will be examined in greater detail.

2.1 Auditory System Physiology

The auditory system has a dynamic range of 20 μPa to 20 Pa and a frequency range of 20 Hz to 20 kHz which is remarkable. The auditory system consists of the peripheral auditory system (the ear) and the **A**uditory **C**entral **N**ervous **S**ystem (ACNS) which comprises the auditory brainstem and the auditory cortex. The ear can be further partitioned into the outer ear, the middle ear and the inner ear (see Figure 2.1). The auditory brainstem provides a neural pathway between the inner ear and the auditory cortex. Each component of the auditory system will now be discussed briefly.

2.1.1 The Outer Ear

The outer ear consists of the pinna (auricle), concha and the external auditory meatus (ear canal). The pinna is made mostly of cartilage and is the most visible part of the ear. The concha is the indented portion of the pinna leading to the auditory canal which is cylindrical in shape and lined with wax producing glands that protect and lubricate the ear. The ear canal leads to the tympanic membrane (ear drum) which is the beginning of the middle ear (Newby et al., 1992).

The function of the pinna and the concha is to direct and concentrate incoming sound waves to the auditory canal which then directs the sound to the tympanic membrane. The outer ear also aids in sound localization, the perception of sound source elevation and protects the middle and inner ear from foreign objects. In addition, the concha and ear canal have natural resonant frequencies at 5 and 2.5 kHz respectively which lead to an increase in sound pressure level by a factor of 10 to 12 dB (Bess et al., 1995).

2.1.2 The Middle Ear

The middle ear or tympanum (see Figure 2.2) consists of the tympanic membrane (ear drum), the eustachian tube and the ossicles (malleus, incus and stapes). The tympanic membrane is cone shaped and is made stiffer by radial and circular fibers. Attached to its center is the handle of the malleus. The other end of the malleus connects with the incus which then connects to the stapes. The footplate of the stapes connects to the oval window on the cochlea of the inner ear. The Eustachian tube connects the middle ear with the back of the nasal cavity which allows the middle ear to adjust to variations in outside air pressure and drain fluid to the nasopharynx.

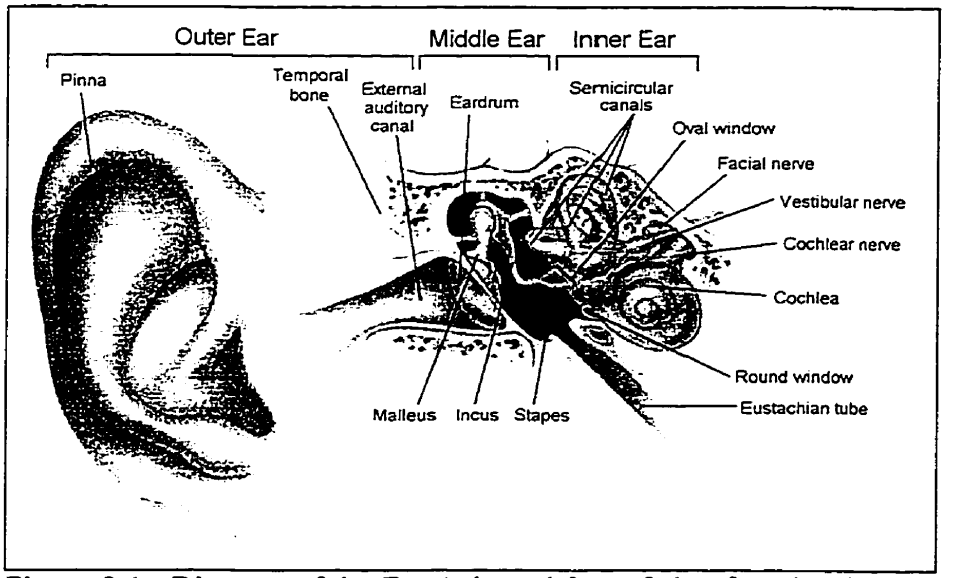

Figure 2.1. Diagram of the Ear (adapted from Schauf et al., 1990).

The tympanic membrane completely seals the middle ear from the outer ear. This sets up pressure differences between the outer and middle ear which causes the tympanic membrane to vibrate in reaction to incoming sound waves. Tympanic membrane vibration causes the ossicles to vibrate. The motion of the ossicles then causes the stapes to move in and out thus pulling and pushing on the oval window and transferring acoustic energy to the inner ear.

To further facilitate sound conduction, the middle ear acts as an impedance transformer between the fluid (high acoustic impedance[1]) of the inner ear and the air (low acoustic impedance) of the middle ear. The impedance transformation is due to the ratio of the larger

[1] Acoustic impedance indicates the ease of transference of an acoustic wave.

tympanic membrane area to the much smaller surface area of the stapes. This area ratio provides a 22% increase in the transmitted sound wave pressure. However, the ability of the middle ear to amplify sound pressure varies with frequency. There is little amplification for frequencies below 100 Hz or above 2 to 2.5 kHz, however taking the outer and middle ear into account there is an amplification of approximately 20 to 25 dB in the range of 100 Hz to 5 kHz (Bess et al., 1995).

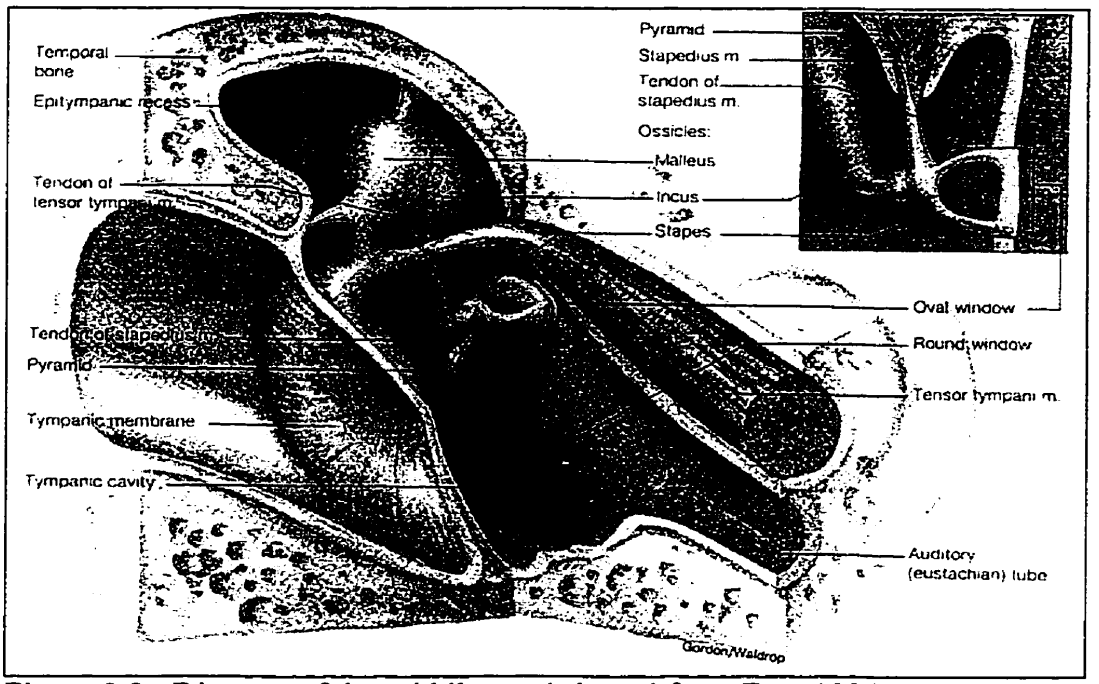

Figure 2.2. Diagram of the middle ear (adapted from Fox, 1996)

2.1.3 The Inner Ear

The inner ear (see Figure 2.3) consists of the semicircular canals (labyrinth), the vestibule and the cochlea. The vestibule and semicircular canals are used for balance. The cochlea is a system of tubes, 3.5 cm in length, coiled circularly about 2 ¾ times and partitioned into three sections: the scala vestibuli, the scala tympani and the scala media. The scala vestibuli (tympani) is the top-most (bottom-most) cochlear compartment and the scala media is the central cochlear compartment. This cochlear partitioning is accomplished by the vestibular membrane which separates the scala vestibuli and scala media and the basilar membrane which separates the scala media and scala tympani (Fox, 1996). Further, the scala vestibuli and the scala tympani share fluid called perilymph and are actually connected at the apical end of the

cochlea via a small opening called the helicotrema. The scala vestibuli is totally isolated from the scala tympani and scala media and contains fluid called endolymph.

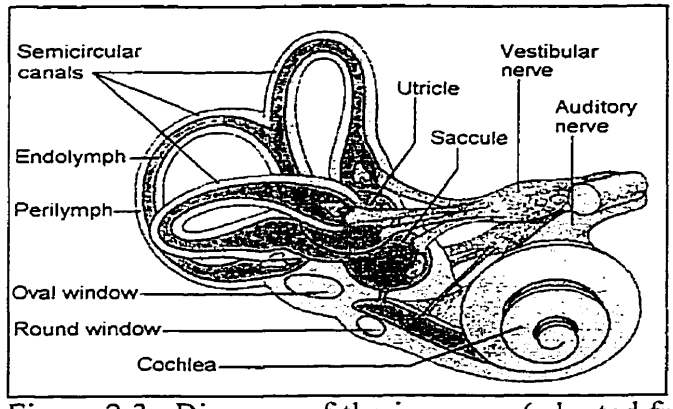

Figure 2.3. Diagram of the inner ear (adapted from Sherwood, 1993).

Another membrane, the **T**ectorial **M**embrane (TM), is in the scala media and covers the Organ of Corti (fundamental to sound transduction) which sits on the **B**asilar **M**embrane (BM) (see Figure 2.4). In addition, there are two membrane covered bumps (the oval and round windows) on the surface of the cochlea closest to the stapes. The oval window is at the basal end of the scala vestibuli and the round window is at the basal end of the scala tympani.

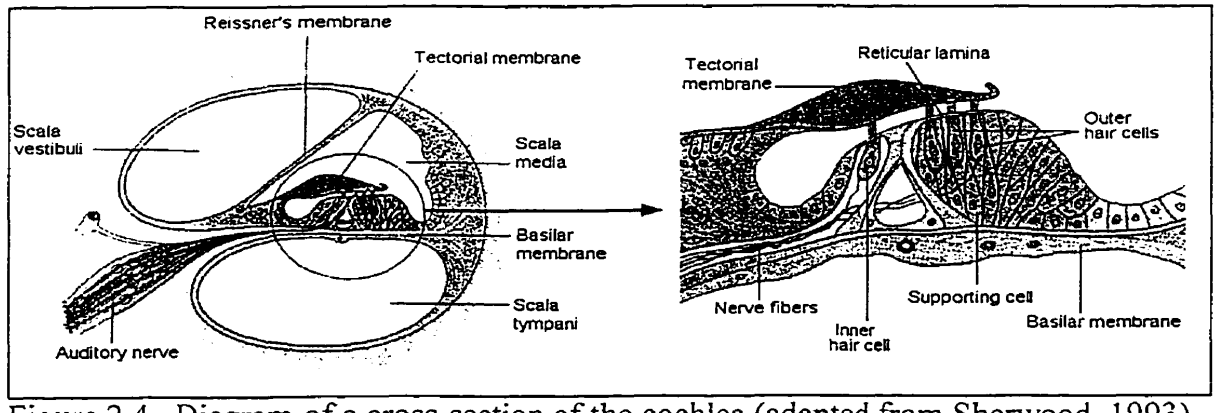

Figure 2.4. Diagram of a cross-section of the cochlea (adapted from Sherwood, 1993).

As mentioned previously, sound waves are transmitted to the cochlea by the movement of the stapes which applies pressure to the oval window. This displaces the perilymph in the scala vestibuli which causes the round window to simultaneously move in the opposite direction of the oval window. Outward motion of the oval window causes the scala vestibuli to enlarge and the scala tympani to reduce in size, meanwhile inward motion of the oval window

causes the scala tympani to enlarge and the scala vestibuli to decrease in size. These size changes produce a wavelike displacement (see Figure 2.5) in the scala media which is usually simplified by considering just the motion of traveling waves on the BM.

The BM is organized such that it responds to certain frequencies best at different locations along its length. At the basal end of the cochlea (near the stapes), the BM is narrow and stiff and responds to high frequency sound waves while at the apex of the cochlea the BM is broad and compliant and responds to low frequency sound waves (see Figure 2.6). Sound waves will diminish on the region of the BM which best responds to the frequency of the sound wave. The Organ of Corti in this region then transduces the mechanical vibrations of the BM into electrical or neural signals.

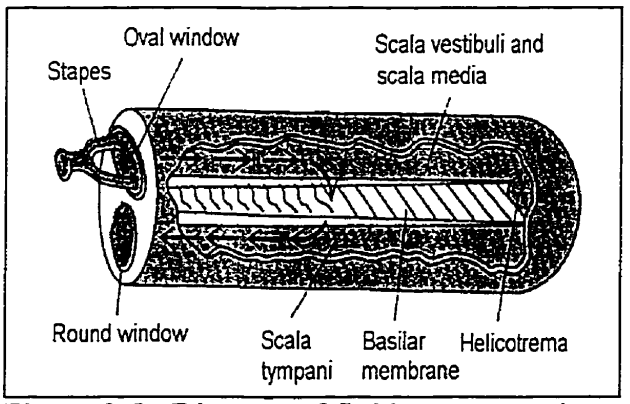

Figure 2.5. Diagram of fluid movement in the cochlea during Sound Transduction. (adapted from Guyton, 1992).

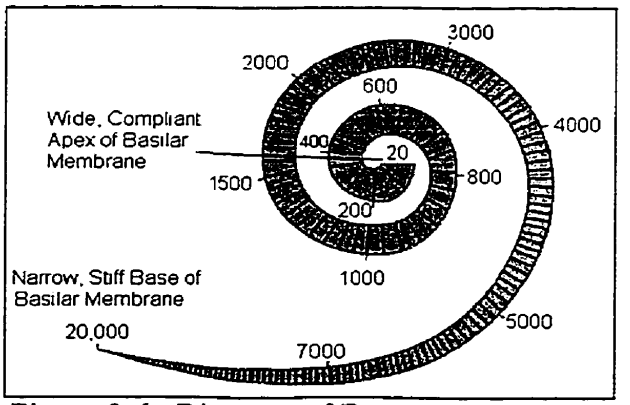

Figure 2.6. Diagram of Frequency Distribution (in Hz) along the BM. (adapted from Sherwood, 1993).

The transduction process of the Organ of Corti is facilitated by specialized sensory hair cells which are electromechanically sensitive. The hair cells are distinguished as either **O**uter **H**air **C**ells (OHCs) or **I**nner **H**air **C**ells (IHCs). There are about 3500 IHCs and 15000 OHCs (Guyton, 1992). The stereocilia of the OHCs are embedded within the TM. The stereocilia of the IHCs are not embedded in any membrane and thus the IHCs are moved by the fluid between themselves and the TM. Ninety-five percent of the afferent fibers which leave the cochlea (i.e. travel to the brain) are from the IHCs meanwhile 5% are from the OHCs. This suggests that IHCs provide sensory information to the higher ACNS levels, meanwhile the OHCs (which

receive information from higher structures and can change size) are thought to alter the relative motion of the TM and BM to change the level of IHC stimulation.

When there is a traveling wave on the BM, a shearing force is created which causes stereocilia on the IHCs to bend which leads to a depolarization (change in electrical potential) of these hair cells and the stimulation of an associated sensory neuron. The sensory neurons (nerve fibers) are connected to the auditory pathway through the vestibulocochlear (VIIIth cranial) nerve. Fibers with a high characteristic frequency are at the perimeter of the nerve and originate from the high-frequency base of the cochlea, meanwhile fibers with low characteristic frequencies are at the nerve core and originate from the low frequency end of the cochlea (Bess et al., 1995).

2.1.4 The Auditory Central Nervous System

The Auditory Central Nervous System (ACNS) refers to components of the auditory system from the VIIIth cranial nerve to the auditory cortex. The auditory brainstem (see Figure 2.7) consists of afferent (ascending) pathways which carry information up the ACNS and efferent (descending) pathways which carry information from the cortex and lower brainstem centers to the auditory periphery. The ACNS is very complex and not yet fully understood. Most nerve fibers cross over at some point along the ACNS so that activity of the right ear is represented more strongly on the left side of the cortex and vice-versa. In addition, the ACNS has tonotopic organization (an orderly mapping of frequency to place) and redundancy (Bess et al., 1995).

The afferent auditory pathway starts with the neural signals produced by the IHCs of the cochlea. These neural signals travel in neurons ipsilaterally (on the same side of the brainstem) through the cochlear nerve to the ventral and dorsal cochlear nuclei. Most of the neurons then cross through the trapezoid body of the pons and project to the contralateral superior olivary complex. Some fibers terminate on the trapezoid body. The superior olivary complex gets a large amount of ipsilateral and contralateral neural inputs which lets it sense the direction of a sound source. From the superior olivary complex, the neurons proceed in a tract called the lateral lemniscus which leads to the inferior colliculus. The auditory neurons then proceed to nuclei called the medial geniculate bodies in the thalamus where there is no crossing over. From here auditory radiations spread to the auditory cortex of the cerebrum. The auditory

pathway is organized such that neurons in different regions of the BM stimulate neurons in corresponding areas of the auditory cortex (Martin, 1991).

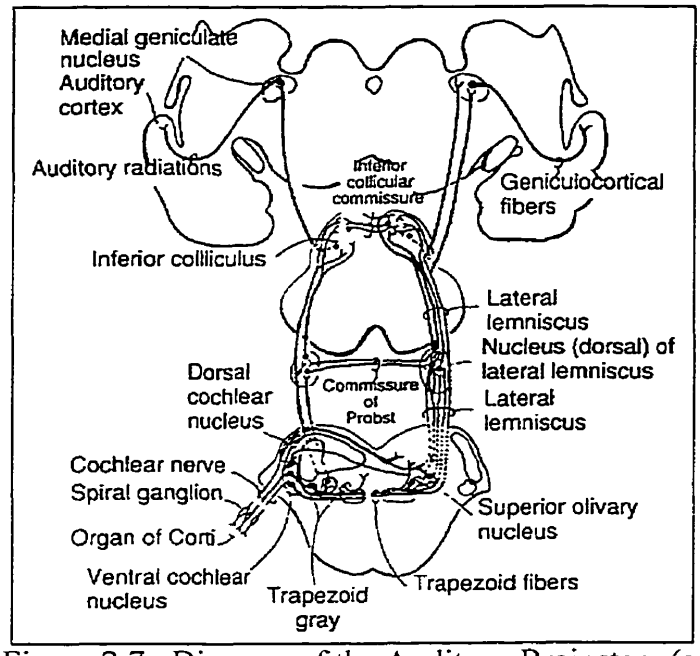

Figure 2.7. Diagram of the Auditory Brainstem (adapted from Guyton, 1992).

The descending efferent auditory pathway is more of a mystery. It comprises many of the same structures that are used in the ascending pathway, however, a different set of nerve fibers is used. It is believed that the efferent fibers may modify incoming neurally coded sensory information at any of the centers along the ACNS. In addition, there are 2 efferent fibers for every 100 afferent nerve fibers. The last nerve fibers in the descending pathway run from the superior olivary complex to the cochlea ipsilaterally or contralaterally; they are called either the crossed (more abundant; 60 to 80%) or uncrossed olivocochlear bundles. The crossed olivocochlear bundles cross over to the ipsilateral cochlea and innervate the OHCs. The uncrossed bundle fibers appear to innervate the IHCs (Bess et al., 1995).

2.2 Hearing Disorders

Hearing disorders are either conductive or sensorineural. Conductive hearing losses are due to a blockage of sound wave transmission through the ear canal and middle ear. Some forms of conductive losses are excessive ear wax, a foreign body blocking the passageway or diseases affecting the tympanic membrane or the middle ear cavity such as ossicular defects.

Conductive hearing loss impairs hearing at all frequencies (Glasscock et al., 1987). Sensorineural hearing loss occurs when there is damage in the neural auditory pathway. The conductive mechanism is usually intact, however, the electrical signals which should be produced in the cochlea are either not produced or not transmitted to the auditory cortex. The problem may be with the hair cell system (a sensory loss) or the auditory brainstem (a neural loss). Sensorineural hearing loss may be due to bacterial infection, ototoxic drugs, excessive noise exposure or simple aging (presbycusis). Neural losses include acoustic tumors along the ACNS (Newby et al., 1992). Further, Sensorineural hearing loss does not always impair the hearing of all frequencies.

2.3 Audiometric Tests for Hearing Loss

Over the years a battery of audiometric tests has been constructed to detect hearing impairment. These tests comprise pure tone audiometry, speech audiometry, acoustic immittance measurements, stapedial reflex tests and click-evoked ABRs. The audiometry tests are subjective, require participation on the part of the test subject and can test for the general integrity of the auditory system as well as aid in localizing the cause of a hearing problem. Briefly, pure tone audiometry presents pure tones at different test frequencies to detect a subject's hearing threshold (the lowest sound level that a person can correctly hear 50% of the time). This threshold information is plotted against frequency to obtain the audiogram. Hearing thresholds less than 25 dB HL[2] are considered normal. Air conducted stimuli are normally used, however, to determine if there is a conduction problem bone-conducted stimuli can be used. Speech audiometry follows the same principles except it uses speech and it can not discriminate between a conductive or a sensorineural hearing loss (Bess et al., 1995).

Acoustic Immittance measurements are sensitive, objective tests which can identify the presence of middle ear fluid, evaluate eustachian tube and facial nerve function and determine the nature of a hearing loss and assist in the diagnosis of the site of auditory lesion (Bess et al., 1995). To probe for retrocochlear lesions, the following tests are used: threshold tone decay, acoustic reflex threshold, ABR, and acoustic reflex decay. This test battery is 90 to 95% accurate in identifying retrocochlear lesions (Bess et al., 1995). These tests are usually

[2] HL is referenced to a normal hearing person's audiometric threshold across different audiometric frequencies.

performed on suspicion of an auditory physiology problem. The equipment is specialized and is contained in audiology clinics and thus can not be used in the universal screening of neonates for hearing loss. However, recent objective, noninvasive tests exist which can be used for screening.

These recent objective, noninvasive audiometric tests are based on **O**toacoustic **E**missions (OAEs) and SSABRs. OAEs, first discovered by Kemp in 1978, are sounds in the external ear canal which are generated by the OHCs within the cochlea. A strong OAE indicates the healthiness of the OHCs and that the sound conduction properties of the middle ear are normal. There are two popular types of OAEs: Transiently Evoked OAEs and Distortion Product OAEs . They are evoked in almost every normal ear and are reduced or absent in hearing impaired ears so that a distinction can be made between normal and pathological ears. They also require brief examination periods, and have high retest reliability (Probst et al., 1991).

SSABRs are electrophysiological signals which are measurable at the scalp and elicited by an acoustic stimulus. SSABRs reflect the integrity of the auditory periphery, the cochlear nerve and the auditory brainstem. SSABRs have been used to assess the physiological hearing threshold of normal and hearing impaired subjects and these measurements have matched well with behavioral hearing thresholds. Unfortunately, SSABR testing requires longer test times than those used in OAE testing, however, SSABRs provide a more thorough inspection of the auditory system (SSABRs probe structures beyond the cochlea). This can not be ignored because the prevalence of retrocochlear based hearing loss is too high to disregard in infants, school children and adults (please refer to Appendix A). Furthermore, OAEs may not detect retrocochlear based hearing losses because they only probe as far as the cochlea. This was seen in a study by Starr et al. (1996) on 10 children and young adults. In addition, Dobie et al. (1995b) stated that with widespread OAE screening recommendations, the need for SSABRs will increase in order to verify OAE test failures.

2.4 The Auditory Brainstem Response

The **A**uditory **B**rainstem **R**esponse (ABR) is the current standard in audiologic and neurologic investigations. They are a type of **A**uditory **E**voked **P**otential (AEP) meaning an acoustic stimulus induces a potential along the inner ear and ACNS which is measurable with

electrodes. Thus, ABRs monitor the status of the auditory periphery (inner ear), and the ACNS. ABRs can be classified according to the acoustic stimulus used (click, tone pip, tone burst or Amplitude Modulated (AM) wave) and the types of response evoked (transient or steady state). The conventional click evoked ABR is composed of 5 waves with certain amplitude and latency relationships. An abnormality in the amplitude or latency of any wave indicates a problem with the associated portion of the auditory pathway (Zimmerman, 1994).

Unfortunately ABRs have some shortcomings. The stimulus for click-evoked ABRs contains broadband energy which is not good for determining frequency specific hearing deficits (Durieux-Smith et al., 1991). In fact, the wideband click stimulus usually produces evoked thresholds which reflect the average or best hearing across 1 to 4 kHz, however the frequency specificity can be improved if short tones and/or notched noise are used in the stimulus (Rance et al., 1995). In addition, physiological thresholds obtained with ABRs are elevated above the behavioral thresholds and the ABR is not useful for assessing hearing below 2 kHz (Kuwada et al., 1986). Another problem with ABRs is that they have low SNRs which usually dictates recording times on the order of minutes (2 or 3) for response detection even when using high intensity stimuli (Dobie et al., 1995b).

2.5 The Steady State Auditory Brainstem Response

SSABRs[3] are ABRs in response to an AM stimulus. They result from the synchronous discharge of auditory neurons phase-locked to the modulation frequency of the AM stimulus (Aoyagi et al., 1993b). The SSABR amplitudes indicate the hearing sensitivity of the subject and the absence of an SSABR may indicate a hearing loss (Kuwada et al., 1986). SSABRs can be recorded from subjects of all ages, even infants within a few days of birth (Rickards et al., 1994). SSABRs are currently used for assessing hearing threshold (Lins et al., 1995a) however they are applicable to screening tests.

SSABRs have definite advantages over conventional click-evoked ABRs. For instance, they are evoked by frequency-specific stimuli which enables the diagnosis of a frequency specific hearing loss. In addition, the responses to several AM signals applied simultaneously can be evaluated individually which results in a significant decrease in test time. SSABRs can

[3] SSABRs are also called Amplitude Modulation or Envelope Following Responses (AMFRs or EFRs).

also be recorded at near threshold stimulus intensities to provide auditory threshold assessments. In addition, SSABRs can follow the shape of a subject's behavioral audiogram (Aoyagi et al., 1994b). Specific properties of SSABRs and testing will now be discussed.

2.5.1 SSABR Stimulus Presentation

The SSABR stimulus is generated digitally with software, then converted to an analog form via a D/A converter and finally transduced into an acoustic waveform by headphones or an insert probe phone. The benefit of using an insert probe phone is that the speakers are miniaturized and thus do not require as much current to produce the stimulus as speakers in headphones would. This reduces the possibility of electrical coupling (via a magnetic field) between the stimulus and response pathways of the instrumentation. The stimulus can also be bandpass filtered with filter settings of 0.3 kHz (to prevent electrical artifact at the modulation rate in the recording) to 10 kHz (to remove digitization noise) (Lins et al., 1995a). Furthermore, the stimulus may be presented monaurally (to one ear only) or binaurally (to both ears at once).

2.5.1.1 Stimulus Properties

The stimulus is an AM waveform which consists of sinusoidal carrier and modulating waveforms (see Figures 2.8 and 2.9) represented by the following formula:

$$st(t) = L_c(1 + \mu\cos(2\pi f_m t))\cdot\cos(2\pi f_c t) \qquad (2.1)$$

where: L_c is the stimulus intensity
μ is the modulation index
f_m is the modulation frequency
f_c is the carrier frequency

The carrier frequency is the audiometric test frequency (i.e. the SSABR stimulus assesses the cochlea at the carrier frequency (Lins et al., 1995a)) and it can be any frequency in the range of 0.25 to 8 kHz, however usually either 0.25, 0.5, 1, 2, 4 or 8 kHz is chosen. The modulating frequency is chosen to evoke a robust response and also depends on the test subject and subject state. The most commonly used modulation frequencies are in the 30 to 50 Hz and 70 to 100 Hz ranges. The modulation index is chosen to evoke a strong response and the stimulus intensity is chosen according to the test paradigm.

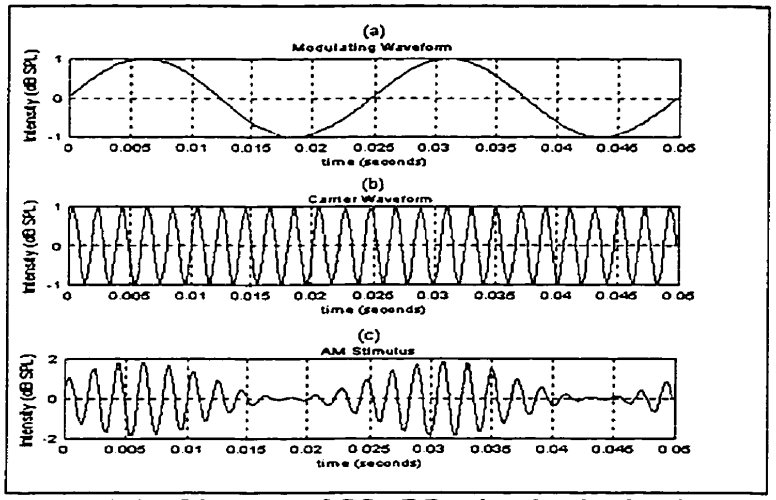

Figure 2.8. Diagram of SSABR stimulus in the time domain: (a) modulating waveform, (b) carrier waveform, (c) AM stimulus (arbitrary y axis magnitudes).

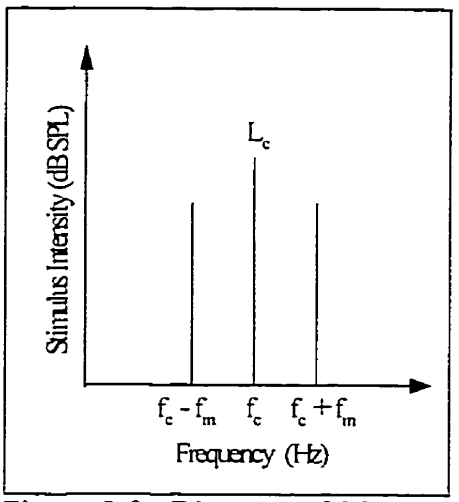

Figure 2.9. Diagram of SSABR stimulus in the frequency Domain.

2.5.1.2 The Multiple SSABR Stimulus

The Multiple SSABR stimulus comprises more than one AM waveform. For instance a Multiple SSABR stimulus might consist of AM waveforms with f_c values at 0.5, 1, 2 and 4 kHz with corresponding f_m values at 75, 80, 85 and 90 Hz. Each f_m value must be distinct so that the response to each AM waveform can be measured. Values for f_m can be chosen as close as 1.3 Hz without any effect on SSABR amplitude (John et al., 1997). Lins et al. (1995a) found that the SSABR amplitude was not very different when presenting each AM waveform by itself or all at once, however, values for f_c must be separated by an octave so that each f_c in the compound stimulus vibrates a separate part of the BM (Lins et al., 1996). Further, physiological thresholds obtained using the Multiple SSABR stimulus are not significantly different from those obtained using single SSABR stimuli. More importantly, this stimulus can greatly reduce test time.

2.5.2 SSABR Measurement

SSABR testing is usually carried out in an electromagnetically shielded, sound proof room. The shielding protects against electromagnetic contamination in the SSABR recording. The sound proof room will aid the subject in hearing the stimulus clearly. During testing, the subject needs to be quiet, relaxed and must be still to avoid adding any physiological noise to the SSABR recording. Appendix C describes the test setup used for this thesis work.

The SSABR is measured with the use of skin electrodes which are usually Silver Silver-Chloride electrodes because they have low impedance and low electrical noise for low frequencies (Neuron, 1995). Three electrodes are used; two electrodes serve for differential measurement and the other electrode serves as ground. The inter-electrode impedance should be low (less than 5 kΩ) and balanced between all three pairs of electrodes. To insure this, the skin should be abraded to reduce the electrical impedance of the skin-electrode contact. Different electrode placements are possible. Usually one active electrode is placed at the subject's vertex (crown of the head) or forehead. The other active electrode is placed on the ipsilateral mastoid (just behind the ear) and the ground electrode is placed on the contralateral mastoid (Rance et al., 1995; Levi et al., 1993). Dr. Picton (personal communication, 1997) indicated that SSABRs measured with the electrode at the forehead are 75% as large as those measured with the electrode at the vertex, however, the recordings may also be less noisy.

The electrodes lead to a differential pre-amplifier which is used to provide large gain (around 10,000 V/V). After amplification, the signal is passed through a band pass filter with a typical low frequency cutoff at 10 to 30 Hz and a high frequency cutoff at 100 to 300 Hz (Aoyagi et al., 1993c; Dobie et al., 1995b). The signal is then sent to the computer where A/D conversion takes place and the signals are processed. An FFT of a 5 seconds response duration is shown in Figure 2.10. The stimulus was at 70 dB SPL with a 1 kHz carrier frequency, 40 Hz modulation frequency and a 0.95 modulation index.

2.5.3 SSABR Generation

In order to understand how an SSABR is generated, we must understand why the SSABR frequency is at f_m while there is no stimulus component at f_m. This can be explained by cochlear processing. To begin, Griffith et al. (1991) showed that the AM stimulus stimulates the region of the cochlea that is responsive to the f_c. In this cochlear region, the stimulus undergoes compressive half-wave rectification. The IHC accounts for the compression and the ganglion cell innervating the IHC accounts for the half-wave rectification. It is the half-wave rectification that introduces a stimulus frequency component at f_m (Lins et al., 1995b).

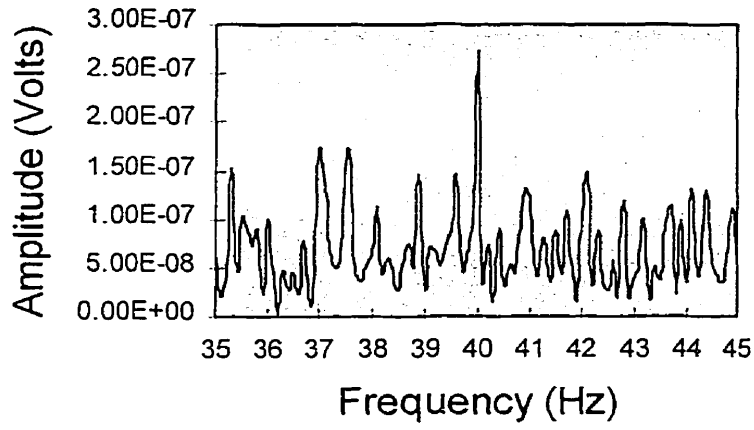

Figure 2.10. An FFT of an actual 40 Hz SSABR (5 second duration).

SSABR generation is due to sources in the ascending auditory pathway that can generate open electrical fields. Researchers are not precisely sure which structures are the generators, however, they do know that there are distinct generators for the 40 and 80 Hz SSABRs (regardless of f_c). For instance, Kuwada et al. (1986) reported that 40 Hz SSABRs have latencies of 31 ms suggesting a cortical generator meanwhile 80 Hz SSABRs have latencies of around 8 ms suggesting a midbrain generator. Furthermore, Kuwada stated that the difference in SSABR amplitude between 40 and 80 Hz SSABRs may be the result of near field and far field recording conditions due to the generators of these responses being different distances away from the scalp electrodes. Aoyagi et al. (1994a) further speculated that 80 Hz SSABRs may be generated by the inferior colliculus and cochlear nucleus and that the 40 Hz generator could be the thalamus or midbrain.

2.5.4 SSABR Properties

In this section, certain SSABR characteristics will be discussed such as the effect of different stimulus parameters, variability and properties for different subject groups.

2.5.4.1 Signal Characteristics

The SSABR is usually recognized as a sinusoid whose frequency matches the f_m used in the stimulus. However, for supra-threshold stimuli, the SSABR contains energy at the f_c and at the sidebands for low f_c (f_c = 0.5 or 1 kHz), at the f_m and at other harmonic and distortion product frequencies. For near threshold stimuli, the SSABR contains coherent energy only at f_m

(Dobie et al., 1995a, Aoyagi et al. (1993a, 1993b)). Furthermore, Picton et al. (1987) found that the 2nd harmonic was much smaller and less consistent than the response at the f_m.

2.5.4.2 Effect of Different Stimulus Parameters on the SSABR

The effect of f_c on SSABR amplitude depends on the f_m which is used. When the f_m is in the 40 Hz range, the SSABR decreases in magnitude with increasing f_c (Kuwada et al., 1986; Griffiths et al., 1991, Picton et al., 1987). However, when f_m is in the range of 70 to 100 Hz, the SSABR amplitude does not vary significantly with f_c (Lins et al.,, 1995b).

The effect of f_m on SSABR amplitude depends on subject age and state (awake or asleep). For awake adults, the largest responses occur for f_m in the 40 Hz range (Levi et al., 1993; Aoyagi et al., 1993b). Selective attention does not affect 40 Hz SSABRs (i.e. f_m = 40 Hz) (Dobie et al. (1994b), however, in sleeping adults, the 40 Hz SSABR amplitude is reduced by a factor of 2 (Lins et al., 1995a) and SSABR detection becomes more efficient at higher f_m values, particularly between 80 and 100 Hz (Aoyagi et al., 1993b; Levi et al., 1993). This is because SSABRs in the 70 to 100 Hz range are unaffected by sleep and the EEG (the background noise) in this frequency range diminishes during sleep. A comparison of 40 Hz to 80 Hz SSABR amplitudes for awake adults is shown in Table 2.1. For infants and children, the SSABR is not recordable with f_m = 40 Hz. The optimal range for f_m is in the 80-90 Hz region when the child is asleep (Aoyagi et al., 1993b). This is seen for subjects from 4 months to 15 years of age (Aoyagi et al., 1994c).

Table 2.1. SSABR Amplitude versus f_m for f_c = 1 kHz, μ = 100% and L_c = 60 dB SPL (adapted from Lins et al., 1995a).

f_m (Hz)	SSABR Amplitude (nV)
39	320 ± 200
81	80 ± 40

The modulation index (μ) also affects SSABR amplitude. The amplitude increases as μ is increased for all values of f_m and all subject groups (Lins et al., 1995b; Kuwada et al., 1986; Milford et al., 1989). In the literature, typical values for μ are between 50% and 100% (usually μ is near 100%). However with larger μ values, the response has increasing energy at the 2nd and 3rd harmonics of f_m for high stimulus intensities (Lins et al., 1995b).

The SSABR amplitude also increases with L_c. Typical values for f_c = 500 Hz and f_m = 40 Hz are 300-400 nV for L_c = 60 dB nHL and about 100 nV for L_c= 30 dB nHL (Dobie et al., 1995a). At higher values of f_c, the increase in SSABR amplitude with L_c becomes more gradual (Kuwada et al., 1986). However, Lins et al. (1995b) found that SSABR amplitude grows much larger as L_c increases above 70 dB SPL (f_m = 91 Hz was used). There is also an increase in the 2nd and 3rd harmonics with increases in L_c (Lins et al., 1995b; Rickards et al., 1984).

SSABR phase increases with f_m (Kuwada et al., 1986; Picton et al., 1987), meanwhile with decreasing L_c, the phase decreases with an average phase-intensity slope of 37°/10 dB SPL over the range of 20 to 50 dB nHL for f_m = 40 Hz (Dobie et al., 1994a). However, SSABR phase increases with f_c (Lins et al., 1995b). In addition, the SSABR phase increases as μ varies from 0 to 25% after which the SSABR phase remains relatively unchanged (Lins et al., 1995b).

2.5.4.3 Variability and Adaptation

Kuwada et al. (1986) found a 15-25% SSABR amplitude variability while testing subjects who sat quietly and read. Meanwhile, Griffiths et al. (1991) did a within-subjects test and found no significant difference in SSABR amplitude across time for any f_c. Further, Chambers et al. (1993) found strong correlation between SSABR amplitudes from two recording sessions which were 1 week apart for subjects with mild to moderate hearing loss.

In terms of adaptation, Campbell et al. (1977) noted no adaptation effect on SSABRs while using continuous stimulation (over a 4 minute period) and discontinuous stimulation (stimulating for a 25 second period and then resting for a period of 1 minute). Meanwhile, John et al. (1997) found that some subjects noted that a 35 dB SPL stimulus (f_c = 2 kHz) occasionally became inaudible during testing (the recordings were approximately 10 minutes long). Thus it appears that adaptation may occur for longer test periods.

2.5.4.4 Testing with Hearing Impaired Subjects

In normal ears, SSABRs can be recorded at low L_c (around 20 to 30 dB SPL), meanwhile in ears with a hearing loss, the response is only noticeable at higher L_c values such as 50 dB SPL. In addition, Kuwada et al. (1986) found that larger SSABRs are obtained at

frequencies where hearing sensitivity is better and that smaller SSABRs are obtained at frequencies where hearing sensitivity is poorer. Another interesting finding was that SSABR thresholds for hearing impaired subjects were closer to the behavioral thresholds than those for normal hearing subjects (Kuwada et al., 1986; Lins et al., 1996). Furthermore the literature shows that SSABR thresholds follow the behavioral audiogram of adults and children with a hearing loss (Griffiths et al., 1991; Lins et al., 1996; Aoyagi et al., 1993c; Chambers et al., 1993; Rance et al., 1995). The strength of this relationship increases with f_c, increasing degree of loss and decreasing L_c (Kuwada et al., 1986). Kuwada also found that SSABR amplitudes in the non-impaired regions of hearing impaired subjects were larger than those of normal subjects. In addition at high L_c, a subject with a hearing loss may appear normal in terms of SSABR amplitudes (Lins et al., 1996).

2.5.4.5 Testing with Infants

SSABR testing is more difficult for newborns and young children when they are awake because of physiological artifacts caused by body movement (Ayogai et al., 1994b). However, Dr. Picton (personal communication, 1997) indicated that there are fewer artifacts when infants are sleeping. Furthermore, SSABR testing on infants was found viable as shown by Rickards et al. (1994) who obtained positive responses on 90% of the infants tested (L_c = 70 dB SPL). However, Lins et al. (1996) found that infants had a harder time hearing stimuli with f_c = 500 Hz as there were responses in only 70% of 50 infant ears when L_c was 60 dB SPL. Lins also found that with L_c = 60 dB SPL, infant SSABR amplitudes were on average 45% as large as those from an adult subject group, however SSABR phase was very similar. They also found that infant SSABR thresholds were on average 10 dB above those of normal adults.

2.5.4.6 SSABR and Behavioral Threshold Levels

An SSABR threshold is defined as the lowest L_c at which an SSABR is detected at some false positive level (Dobie et al., 1995a). The literature has shown good correlation between SSABR thresholds and behavioral thresholds from pure tone audiometry tests with higher correlation at higher f_c (Lins et al., 1995b). Further, Griffiths et al. (1991) found a difference of 11.6 dB for normals at all audiometric frequencies tested. In general, for adults the SSABR thresholds are larger than behavioral thresholds ranging from within 10 to 25 dB

SL[4] (Kuwada et al., 1986; Lins et al., 1995b; Aoyagi et al., 1993c; Tucci et al., 1990). There is still some debate as to whether sleep affects SSABR thresholds (Aoyagi et al., 1993b; Jerger et al., 1986).

For normal hearing kids (aged 4 to 7), Aoyagi et al. (1994a) found physiological thresholds within 20 dB HL of behavioral thresholds. Meanwhile, Dr. Picton (personal communication, 1997) stated that infant SSABR thresholds were 20 to 30 dB higher than behavioral thresholds. Dobie et al. (1995a) also indicated that SSABR thresholds improve when test times of up to ten minutes are used.

2.5.4.7 Recording Time needed for Detection

SSABR detection time depends on SSABR amplitude and noise intensity. Most often, researchers record for a constant time length (approx. 240 seconds) in their SSABR tests regardless of response SNR (i.e. they over record in situations of high SNR). Not much work has been done in the literature to characterize the average amount of time needed to detect a response. An exception is Lins et al. (1995b) who found that with f_m = 91 Hz, f_c =1 kHz and L_c = 60 dB SPL, it takes about 80 seconds on average for a response to become statistically significant and if L_c is lowered to 30 dB SPL, response detection requires about 3 to 4 minutes.

When testing at near-threshold intensity[5], longer test times are needed to increase response detection as shown by Stappells et al. (1986). He found that with L_c = 10 dB SL only 1 out of 10 subjects showed a response after 1 minute of recording but when the recording time was increased to 8 minutes, responses were detected in 8 out of 10 subjects. Test protocol also affects test time length. For instance, the test can be done a few times at a certain L_c with different f_c's for screening purposes or the test may be repeated many times with different L_c values for the determination of physiological hearing threshold at a particular frequency.

[4] SL (Sensation Level) is referenced to the test subject's behavioral threshold at the test frequency.
[5] In threshold testing, one or two measurements can't be taken at high stimulus intensities and then extrapolated down to threshold because a linear relationship does not exist between stimulus intensity and SSABR amplitude (Campbell et al., 1977; Lins et al. 1995b).

3.0 Investigation into SSABR Detection: The Facts and what's been done so far

To improve SSABR detection, we need to investigate the SSABR signal and noise properties and the techniques that have already been used in detecting SSABRs. The SSABR signal properties have already been mentioned in section 2.5.4.1. Suffice it to say, the SSABR signal can be simply considered to be a sinusoid whose frequency is the modulating frequency used in the stimulus. With this said, this chapter will examine in more detail the noise in SSABR recordings (EEG noise and Artifacts) as well as the SSABR detection algorithms which have already been proposed. The chapter will conclude with a section on modeling the SSABR and another section with insights on possible improvements for SSABR detection.

3.1 SSABR Noise: The EEG

The predominant source of noise in SSABR recordings is due to the EEG which is the electrical activity produced by the brain. The EEG varies over different parts of the scalp and changes in a consistent, recognizable fashion when the general status of the subject changes (i.e. from relaxed to alert). The EEG comprises stochastic components which are stationary over short time periods and transient components such as wave trains, spikes and sharp waves that arise sporadically (Narasimhan et al., 1985). The EEG is usually assumed to be ergodic, stationary and gaussian, however, in practice none of these may actually hold (Gasser et al., 1979). The normal spontaneous EEG for an awake adult with open or closed eyes doesn't have significant temporal changes (Dumermuth et al., 1975) and normals do not usually exhibit spiky behavior in their EEG. Usually, the waking EEG is less than 300 μV in amplitude and becomes considerably smaller during sleep (Linden et al., 1985).

The EEG frequency range is from 0.1 to 100 Hz (Niedermeyer, 1982). In general, the EEG spectrum consists of flat white noise, pink noise and colored noise (Dumermuth et al., 1987). An example of an EEG spectrum is shown in Figure 3.1. The low frequency EEG components (less than 20 Hz) contain most of the energy. Clinically, the EEG is split into 5 frequency bands: the delta (0.5-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (14-25 Hz) and gamma bands (25 to 90 Hz or higher) (Lopes Da Silva, 1982). These EEG bands may vary from person to person which could be due to different anatomical morphologies, subject state

and subject age (Gevins, 1987). Further, when comparing adults to children, Levi et al., (1993) found physiological noise levels higher in infants than adults, however, others found similar noise levels for awake and asleep adults and asleep children (Aoyagi et al., 1993; Lins et al., 1996). The EEG also undergoes maturational changes from birth to adolescence.

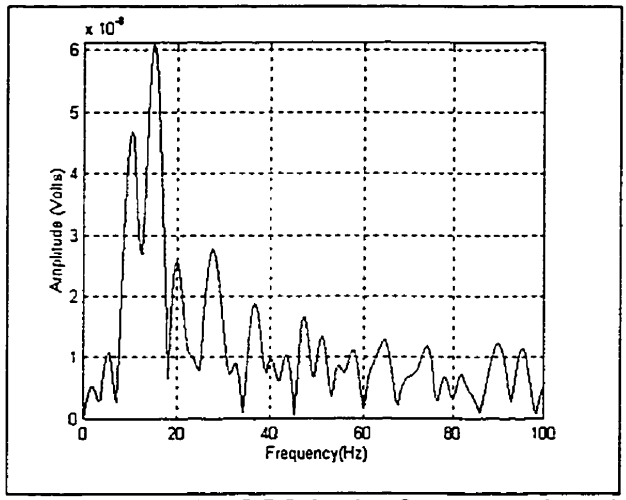

Figure 3.1. The EEG in the frequency domain.
(Note: This data has been bandpass filtered from 10 to 100 Hz).

The EEG in general is nonstationary (Lopes Da Silva, 1982), however, there are conflicting reports on the degree of nonstationarity. Hernandez et al. (1995) indicated that EEG segments longer than 10 seconds must be regarded as non stationary in most cases, meanwhile, Blanco et al. (1995) noted that the amount of time that the EEG remains stationary varies from several seconds to several minutes. In practice, to avoid nonstationarity many researchers have used 1, 2, 5 or 10 second intervals when analyzing the EEG. The nonstationarities are due to sleep cycles, changes in state, task performance, subject movement, and muscular tension. Thus to increase the length of time stationarity the subject should remain still (Dobie et al., 1996).

The amplitude probability distribution of the EEG is also a disputed topic. There is evidence supporting a gaussian probability distribution, however, deviations from gaussianity are also a common EEG feature (Dumermuth, 1977). What is most likely is that the EEG has a different amplitude probability distribution under different circumstances. For instance the normal resting EEG tends to be strongly gaussian whereas the EEG from an epileptic brain shows virtually no gaussian tendency (Barlow, 1993). Observation time also has an impact on

the gaussianity of the EEG. McGillem et al. (1987) stated that EEG gaussianity holds for periods of a few seconds or less. Further, the notion of independence between adjacent EEG samples is very important when using statistical tests to determine if the EEG is gaussian. Many original studies neglected this constraint, however, Persson (1974) pointed out that a small number of methodologically acceptable studies rejected the gaussian hypothesis.

Briefly, the test results of some researchers will be discussed (for more information refer to Appendix B). Elul (1969) conducted tests on an adult in an idle state and in a state performing mental arithmetic. Elul found the EEG to be gaussian 66% of the time in the idle state and 32% of the time in the mental arithmetic state. Later, Elul et al. (1975) conducted tests on young children and found the EEG deviated from gaussianity, however, as the age of the children increased the data became more gaussian (80% of the time). Conversely, Dumermuth (1975) performed tests on twins and sleeping adults which showed the EEG deviated from gaussianity.

In 1975, McEwen et al. analyzed EEG data using different sampling rates and found that for 2 second data segments at least 90% of the data was stationary and gaussian and for 8 second data segments at least 65% of the data was stationary and gaussian. In 1978, Sugimoto performed tests on the EEG of sleeping adults and found that 50% of 2.5 to 5 second EEG segments showed both stationarity and normality. Further, Weiss (1989) analyzed 1 second segments of EEG from a sleeping adult and found 47% of the data was non-gaussian. He also tested normal children and found 39% of the EEG data to be non-gaussian.

Gasser et al. (1986) also conducted work on EEG noise obtained from pre and post-stimulus intervals associated with **V**isual **E**voked **P**otential (VEP) testing. Two subject groups were tested; one group had low EEG noise amplitude and the other had high EEG noise amplitude. Results showed that the EEG noise from both groups had a symmetric distribution with mildly longer-tails than that expected for a normal distribution. For both groups, the assumption of normality was mildly validated (the high noise group had stronger deviations from normality but not strong enough to totally deviate from normality). Thus, Gasser concluded that EEG noise is approximately normally distributed.

3.2 SSABR Noise: Artifacts

Artifacts can be separated into 3 categories: physiological, instrumental and external. Artifacts are important because if the artifact's amplitude is sufficiently large and occurs once or twice in the EEG recording then it may be prominent even after time averaging (Regan, 1989). Physiological artifacts occur when the subject talks, swallows, clenches their teeth, blinks their eyes or moves during a recording. To reduce these artifacts, the subject is encouraged to relax by sitting in a comfortable chair with a headrest or by lying down on a bed. The most common artifacts are eye artifacts which do not exceed 50 Hz in frequency but are on the order of 100 to 200 μV in the 3 to 4 Hz region (Sauter et al., 1990). There are also ECG artifacts which look like spikes and influence the DC to 25 Hz range (Kemp et al., 1991). Further, tongue and sweat artifacts contaminate the potentials near DC (Barlow, 1986).

Muscle artifacts also come from the contraction of scalp, neck and facial muscles. These artifacts are large and sharp in contrast to the EEG with amplitudes several times greater than the spontaneous EEG activity. The muscle artifact shape and duration is variable across subjects and depends on the orientation of the muscles with respect to the electrode sites (O'Donnell et al., 1974). Johnson (1979) found several typical muscle artifacts were well approximated by a second order linear system with a center frequency at 70 Hz and a bandwidth of 70 Hz. This was also seen by O'Donnell et al. (1974) who looked at frontalis and masseter muscle contractions. O'Donnell also concluded that EEG recorded from areas not directly over a large contracting muscle appeared to reflect the activity of that muscle only slightly in the lower frequency ranges and only moderately in the mid and upper frequency ranges of the EEG. Meanwhile, Pfurtscheller et al. (1993, 1994) found that EEG activity in the 40 Hz range is strongly related to the planning of a specific movement. Using finger, hand and toe movements, they noted that the maxima of the 40 Hz oscillations occurred just before movement onset and there was a fast decrease in the 40 Hz activity after movement onset.

Instrument artifacts includes noise from electrodes, amplifiers, bandpass filters and digitization. Movement of electrode leads and people walking by the leads result in slow wave variations. This can be reduced by using Silver Silver-Chloride electrodes and lowering the scalp-electrode resistance (Kamp et al., 1982). The effects of amplifier and electrode noise can

be approximated by a scalar gaussian white noise process (Bartoli et al., 1983). For an analysis of the instrumentation noise for the test setup used in this thesis please refer to Appendix H.

External artifacts include the power line interference at 60 Hz which comes from strong sources such as power cables, lights and transformers (Barlow, 1986). This problem can be accentuated by unequal electrode impedance (Misulis, 1994). To minimize the effect of line noise, electrode impedances should be balanced and the differential amplifier input impedance must be made as high as possible. In addition a notch filter can be used to attenuate line noise. The subject may also be electromagnetically shielded to avoid interference (Kamp et al., 1982).

3.3 Current SSABR Detection Algorithms

SSABR detection algorithms consist of objective algorithms for which there is no need for a trained observer to identify responses (this is beneficial for screening purposes). A criterion is also used to control the false positive rate (the rate of detecting noise as a response) (Dobie et al., 1993). The various objective response techniques in the literature are the **M**agnitude **O**nly (MO), **C**ircular $\mathbf{T}^2$ Statistic (CT^2), **M**agnitude **S**quared **C**oherence (MSC) and **P**hase **C**oherence (PC) methods and the F test. These methods utilize conventional data preprocessing techniques such as artifact rejection (which combats the contamination of the recorded data by physiological noise) and synchronous time averaging (which reduces the noise power and increases the SNR). Furthermore, the CT^2, MSC and PC methods involve data segmentation to improve the statistical power of the test. The detection methods and data preprocessing methods will now be discussed.

3.3.1 Dealing with Artifacts: Artifact Rejection and other methods

The traditional way of dealing with EEG data contaminated with artifacts is to use **A**rtifact **R**ejection (AFR) which excludes very noisy data from further processing. This is necessary because for frequent artifact occurrence the noise reduction performance of synchronous time averaging will be degraded (Ruchkin, 1988). In AFR, a predefined voltage threshold is set and any recorded buffers with data which exceed this threshold are assumed to contain artifacts and are discarded. Furthermore, this voltage setting can be chosen so that a pre-specified amount of recorded data is discarded. One voltage threshold reported in the

literature is ±40 μV (Lins et al., 1996). The use of AFR in removing artifacts, however, has limitations such as the allowance of low-amplitude artifacts into the averaging procedure which degrades the response estimate. In addition whole data buffer rejection can significantly prolong recording time which increases the possibility of artifact generation (Bezerianos et al., 1995). All current SSABR detection algorithms use Artifact Rejection in data preprocessing.

Other artifact reduction techniques are to get the subject to relax as much as possible or to use algorithmic techniques consisting of subtraction, nulling or artifact cancellation so that there is no reduction in EEG data (Barlow, 1986). Different electrode sites can also be tried to reduce the effect of muscle artifacts on the EEG. For instance there is little muscle artifact at the vertex and the earlobe has less muscle artifact than the mastoid (Chiappa, 1990). Another technique is weighted time averaging which has been found by many researchers to be robust in processing artifact contaminated data and results in reduced recording time needed for response detection (Gasser et al., 1983; Fan et al., 1991; Gerull et al., 1996; Miskiel et al., 1987).

3.3.2 Synchronous Time Averaging

Synchronous **T**ime **A**veraging (STA) is a very common tool used in **E**voked **P**otential (EP) recordings to increase SNR. The recorded data is partitioned into data segments called sweeps such that the signal is synchronized in each sweep such that it remains constant during averaging. Meanwhile, the noise will not be synchronized in each sweep and will become small if many sweeps are averaged. STA provides an unbiased signal estimate which improves as the number of sweeps contributing to the average increases (Ruchkin, 1988). Unfortunately, if the signal changes in amplitude over the total number of sweeps recorded then there will be a bias problem which will lead to a smaller amplitude in the averaged result (Gasser et al., 1983).

STA can be performed with different statistical models assumed for the data, however, the standard assumptions required are that the EP is synchronized with stimulus presentation, the EP is identical for each stimulus presentation, the EP is statistically independent from the noise and the noise is stationary with zero mean (Regan, 1989). Given these assumptions, a statistical analysis of time averaging shows that the expected value of the noise variance ($\sigma_N^2(t)$) in the final averaged result is (Regan, 1989):

$$E\{\sigma_N^2(t)\} = E\left\{\left[\frac{1}{N}\cdot\sum_{i=1}^{N} n_i(t)\right]^2\right\} \tag{3.1}$$

$$= \frac{1}{N^2}\cdot E\left\{\left[n_1(t)+n_2(t)+\ldots+n_N(t)\right]\cdot\left[n_1(t)+n_2(t)+\ldots+n_N(t)\right]\right\} \tag{3.2}$$

$$= \frac{1}{N^2}\cdot E\left\{\sum_{k=1}^{N}\left[n_k(t)\right]^2\right\} + \frac{1}{N^2}\cdot E\left\{2\sum_{m=1}^{N-1}\sum_{p=1}^{N-m} n_m(t)n_{m+p}(t)\right\} \tag{3.3}$$

where: N is the number of sweeps averaged
n_i is the noise in the i^{th} sweep
E is the statistical expectation operator

Equation 3.3 indicates how much noise one could expect in the final average. The first term is the sum of sample noise variances for each sweep. The second term is a sum that represents the correlation between noise in two different sweeps taken across all sweeps. If the noise is uncorrelated from sweep to sweep then the second grouped term becomes zero and the noise variance in the final average is:

$$\sigma_N^2(t) = \frac{1}{N^2}\cdot N\sigma_n^2(t) = \frac{1}{N}\cdot\sigma_n^2(t) \tag{3.4}$$

Thus, there is a factor of N reduction in the noise variance which implies a factor of N increase in the SNR. This result is independent of the probability distribution of the noise (Regan, 1989). Even when some of the aforementioned assumptions are not strictly met, the SNR of the average will still usually improve with repetition, however, the most stringent assumption which must be met is that of statistical independence between the signal and noise (Childers, 1977). The above analysis also indicates that the correlation of the noise between sweeps should be kept low and thus an appropriate sweep duration should be used. Furthermore, narrowband noise has an increased degree of correlation.

3.3.3 The Magnitude Only Method

In this algorithm, after preprocessing, an FFT is performed to obtain the amplitude spectrum of the averaged result. The response amplitude is then obtained from the FFT bin which corresponds to the response frequency. An estimate of the noise amplitude at the response bin is then made by averaging across several bins on either side of the response frequency. The signal amplitude is then compared to the noise amplitude and a threshold level

determines whether a response is present. In the literature, there have been variations in the noise estimation process and the threshold level. Batra et al. (1986) used three of four bins on either side of the response bin (ignoring the bin closest to the response bin) to obtain their noise values. A signal was detected if response amplitude was larger than the mean plus three standard deviations of the noise values. Meanwhile, Chambers et al. (1993) decided a response was present if the response was larger than the average of the 4 largest noise components in the entire spectrum.

3.3.4 The Circular T^2 Statistic

After preprocessing the recorded data, the averaged result is partitioned into sixteen segments. This technique evaluates if the responses are replicable in the sixteen sections (Lins et al., 1996). For each segment, an FFT is done and the X and Y (real and imaginary) Fourier components at the response frequency are plotted on a polar plot. This test assumes that the real and imaginary variances will be equal (Dobie, 1993). The plotted values will form a cloud of points. If the recording contains pure noise then the cloud of points will cluster around the origin. If a signal is added to the noise, then the cloud of points will be displaced from the origin. A 95% circular confidence limit of the mean (of the 16 X and Y points) is calculated and if it does not contain the origin then a response is considered present (Lins et al., 1996).

3.3.5 The F Test

This test considers whether the recorded data is considerably different from noise by determining if the response power (P_R) is significantly different from the noise power (P_N) at adjacent frequencies (Lins et al., 1995a). After preprocessing, the power spectrum of the averaged result is calculated. The P_R and P_N values are then obtained (P_N is obtained by averaging noise powers from adjacent bins). The F value is then the ratio of P_R to P_N. A response is present if the F value is larger than a threshold level ($F_{critical}$). To calculate $F_{critical}$ we assume the noise is Gaussian and thus power estimates are Chi-square distributed (Zurek, 1992) and the ratio of two noise powers can be tested as an F statistic. The power at one bin will have 2 degrees of freedom (df) meanwhile the power obtained by averaging across m bins will have 2m df. Thus, $F_{critical}$ is given by the tabulated F statistic value at (2,2m) df. In the literature, John et al. (1997) used 120 neighboring bins (about 5 Hz on either side of the response bin) to

estimate P_N. Meanwhile, Dobie et al. (1996) stated that there are diminishing returns if more than 15 bins are used in the noise power estimate.

3.3.6 Phase Coherence

Phase **C**oherence (PC) involves using only phase information. Some researchers consider this beneficial because they think that phase contains signal information especially when testing near threshold. PC represents the degree of reproducibility of phase (i.e. **P**hase **V**ariance (PV)) at the response frequency between groups of averaged data segments. Small values of phase variance indicate that the Fourier component at the response frequency is more likely to represent a response rather than noise (Fridman et al., 1984).

The recorded sweeps of data are first passed through an AFR algorithm and then divided into usually 10 groups of 200 sweeps each. STA is then done on the sweeps in each group (Aoyagi et al., 1993c). Each group average is then Fourier transformed and the phase at the response frequency for each group is obtained. The degree of phase reproducibility is then assessed a **C**omponent **S**ynchrony **M**easure (CSM) based on the phase variance for the n group averages which is given by (Fridman et al., 1984):

$$PV = 1 - \left[\frac{1}{n}\sum_{i=1}^{n}\cos\Theta_i\right]^2 - \left[\frac{1}{n}\sum_{i=1}^{n}\sin\Theta_i\right]^2 \qquad (3.5)$$

where Θ_i is the phase at the response bin for the i^{th} group

The CSM is then given by:

$$CSM = 1 - PV \qquad (3.6)$$

and the phase coherence (PC) value is given by (Jerger et al., 1986):

$$PC = (CSM)^{1/2} \qquad (3.7)$$

Values of CSM vary from 0 to 1. In the case of a pure signal with no noise CSM is 1. If there is noise and no signal then CSM tends to 0. A response is considered present if the CSM value is larger than a critical value (CSM_{crit}) which is determined by theoretical values obtained when only noise is present. For the noise only case, the CSM has a mean ($\overline{CSM}$) of 1/n and a standard deviation (σ_{CSM}) of $((n-1)/n^3)^{1/2}$. For n groups of averaged data CSM_{crit} is related to $\overline{CSM}$ and σ_{CSM} by (Fridman et al., 1984):

$$CSM_{crit} = \overline{CSM} + 3\cdot\sigma_{CSM} \qquad (3.8)$$

Furthermore, in computing the FFT for each group average, an extended cosine bell window can be used to reduce the amount of spectral leakage (Aoyagi et al., 1993c).

3.3.7 Magnitude Squared Coherence

The Magnitude Squared Coherence technique measures the degree to which two stationary time series are related to each other as a function of frequency. Thus, considering an input/output system, the MSC can indicate the degree to which the output is related to the input as a function of frequency (Dobie et al., 1991). Only for the special case of a periodic input, as in EP testing, the MSC can be calculated using only the output data. The MSC value ranges between 0 and 1. If the recorded data consists of a pure signal then the MSC would equal 1. If the recorded data consisted of only noise then the MSC would tend to 0. For a mixture of signal and noise the MSC value depends on the SNR (Dobie, 1993).

The steps to calculate the MSC involve artifact rejection followed by dividing the sweeps into q groups (these groups are called subaverages) and then performing STA on sweeps in each subaverage. An FFT is then done on each subaverage STA result to obtain the response powers which are then averaged across subaverages to obtain the mean power of the subaverages (MP). The subaverage STA results are then time averaged to form a grand average. An FFT is done on the grand average and the response power (PM) is then obtained. The MSC is then:

$$MSC = \frac{PM}{MP} \tag{3.9}$$

To increase the accuracy of MSC calculation and improve detection, many subaverages should be used. For SSABR detection, the optimal number of subaverages is between 8 and 16 (Dobie et al., 1993). To determine whether a response is present in a recording, the MSC value is compared to a critical value based on no response data. Table 3.1 shows the critical values as a function of false positive rate and the number of subaverages.

MSC values rise with stimulus intensity and longer data collection periods, however, MSC values and critical values decline with an increasing number of subaverages. Further, MSC performance is unaffected by varying the sweep length (which results in an inverse change in the number of sweeps averaged in each subaverage) as long as the total test time is fixed (Dobie et al., 1995b).

Table 3.1. MSC critical values (adapted from Dobie et al., 1989).

Number of subaverages (q)	Critical Value for Desired α* Level (One-Tailed)		
	0.01	0.025	0.05
4	0.792	0.714	0.632
8	0.483	0.410	0.354
16	0.266	0.219	0.183
32	0.139	0.114	0.094
64	0.072	0.058	0.047

* α indicates the false positive rate; i.e. for 1000 no response trials and α = 0.01 we would expect 1000*0.01 = 10 trials to show a significant MSC value.

3.3.7.1 Magnitude Squared Coherence with Weighted Averaging

Weighted Averaging (WA) has been proposed to improve signal estimation when the noise is nonstationary. In this method, subaverages with high power are considered to have more noise than those with low power and so are given less weight. This is because the signal power is much smaller than the noise power for AEPs and so differences in power between subaverages can be attributed to noise. When WA is employed in MSC calculation, each subaverage is weighted by its own overall power prior to the calculation of the MSC value (Dobie et al., 1993). This technique was found to be slightly better than the regular MSC technique on real data from humans but this was not statistically significant (Dobie et al., 1994b). This method is denoted by MSC-WA.

3.3.7.2 Magnitude Squared Coherence with Phase Weighting

Phase weighting was also used to determine SSABR thresholds and was found to improve MSC performance (Dobie et al., 1994a). In phase weighting, the MSC values are multiplied by weights that are related to the phase error between the actual measured phase (from the FFT of the grandaverage) and the expected phase at a given stimulus intensity. The expected phase is obtained from averaging the slope of phase versus intensity plots across a number of test subjects and then using this slope value to predict the expected value at a lower stimulus intensity using the measured phase at a higher stimulus intensity where the last response was detected. Phase weighting is used when the first non-significant MSC test (i.e. no response detected) is seen at a given stimulus intensity. This can then be repeated for all lower intensity stimuli. The weight depends on the phase error ($\Delta\theta$) between the expected and measured phases in one of two possible ways:

Linear phase weighting: $$\text{weight} = 1 - \left|\frac{\Delta\theta}{180°}\right| \quad (3.10)$$

Cosine squared weighting: $$\text{weight} = 0 \quad \text{for} \quad |\Delta\theta| > 90° \quad (3.11a)$$

$$\text{weight} = \cos^2(\Delta\theta) \quad \text{for} \quad |\Delta\theta| < 90° \quad (3.11b)$$

For phase-weighted MSCs, the critical values are lower (based on empirical analysis of over 7500 no response MSC values). Dobie found both phase weighting methods were better than the MSC test (cosine squared weighting was better than linear weighting). The three tests were equivalent in terms of test-retest reliability. However, Dobie cautioned that the phase-intensity slope may differ for abnormally hearing subjects and may vary across subjects.

3.3.8 Relationships between SSABR Detection Algorithms

Many researchers have tried to determine which objective response detection algorithm performs the best. Champlin (1992) and Aoyagi et al. (1993b) found that phase spectral analysis was more sensitive to the detection of SSABRs than the Magnitude Only method. In addition, Dobie and Wilson (1994b) found that with human 40 Hz SSABR data, PC and MSC (with or without WA) performed very similarly in terms of ROC analysis, however, MSC-WA may have performed the best although this result was not statistically significant.

Analytically, the CT^2 statistic, MSC and PC techniques have all been shown to be related to one another. In fact, Dobie et al. (1989) showed that PC is analogous to MSC if amplitudes are set to unity. This leads to:

$$PC = (MSC)^{1/2} \quad (3.12)$$

Dobie et al. (1993) also showed that the MSC technique is an algebraic transform of the CT^2 statistic as seen in the following equation:

$$CT^2 = \frac{q-1}{q} \cdot \frac{MSC}{1-MSC} \quad (3.13)$$

Since they are algebraically related then the two must perform similarly in detecting signals buried in noise. In addition, Dobie et al. (1996) found MSC and the F test to be equivalent in statistical power based on ROC curves constructed from simulated data consisting of sinusoids added to gaussian noise using specified SNRs.

3.4 The SSABR Model

In the literature, EP models have assumed that the signal (SSABR in our case) is deterministic and the noise (ongoing EEG, non-cerebral noise, instrumental noise, etc.) is generally modeled as a stationary random process with zero mean (McGillem et al., 1987). Further, the interaction between the signal and noise can be modeled by either the popular *added power* model or the *phase aggregation* model (Dobie et al., 1994b). In the *added power* model the signal and noise are additive and uncorrelated. In the *phase aggregation* model, the external stimulus is considered to impose order on the initially random noise sources without adding energy. Thus overall power will be unchanged but the estimated noise power will decrease and the estimated signal power will increase for near-threshold stimulus intensities.

To choose the model representing the interaction between the SSABR signal and noise we can turn to the literature. Cohen et al. (1991) found that EEG noise did not vary with f_c in the SSABR stimulus. Furthermore, Champlin (1992) could not find a relation between SSABR amplitude and noise amplitude. In addition, Mast et al. (1991) studied the independence of **V**isual **E**voked **P**otentials (VEPs) and background noise and found to a first approximation that VEPs and undriven EEG components combined linearly. Further, Jervis et al. (1983) tested these interaction models for AEPs using tone burst stimuli and electrode sites at the vertex and mastoids. Jervis found all AEPs studied contained additive energy in at least 1 harmonic component which suggested that the *added power* model is more plausible.

Thus, all of these findings support the *added power* model for representing the interaction between the SSABR signal and noise. Considering this information and the fact that the SSABR has harmonics at higher stimulus intensities, the SSABR model can be represented by:

$$y(t) = s(t) + n(t) \tag{3.14}$$

where y(t) is the recorded signal
n(t) is the background noise

and the SSABR signal (s(t)) is given by[6]:

$$s(t) = \sum_{i=1}^{w} A_i \cdot \cos(2\pi f_i t + \theta_i) \tag{3.15}$$

where A_i is the unknown amplitude of the i^{th} harmonic
f_i is the known frequency of the i^{th} harmonic

[6] The SSABR is thought to vary over time and thus s(t) can be amplitude modulated to reflect this (Dobie et al., 1994b), however, this variation effect has not been proven.

θ_i is unknown phase of the i^{th} harmonic
w is the number of harmonics (w = 1, 2 or 3)

The popular background noise model is gaussian, zero mean, white noise despite the fact that the spontaneous EEG may be non-gaussian (Husar et al., 1997; Johnson et al., 1979). Further, the EEG noise can be modeled by a linear parametric model (using an AR model with order less than 10 (Tseng et al., 1995)) in which the EEG signal is described as the output of a filter which shapes input white noise to match the EEG power spectrum (Lopes Da Silva et al., 1987).

3.5 Insights into SSABR Detection

Since the MSC technique is one of the best SSABR detectors, this algorithm will be examined in a bit more detail. The MSC method involves a combination of STA and frequency averaging. The frequency averaging is used to reduce the statistical uncertainty of the periodogram because averaging K periodogram estimates will reduce the variance of the spectral estimate by a factor of 1/K (if the noise correlation between the successive segments is very small). However, this averaging occurs at the expense of frequency resolution (Challis, 1991). Periodogram based methods also suffer from other disadvantages such as poor spectral resolution for the analysis of short time length signals. The poor resolution is due to finite-width mainlobes (which is inversely proportional to the data time length) from rectangular windowing which sets a lower limit on the resolvability of certain spectral characteristics. For weak sinusoids, this windowing means an increased difficulty in detection because of the presence of sidelobes from stronger spectral components (Swingler et al., 1988). Furthermore, these two problems become troublesome for short data records.

Thus, in reducing the amount of time needed for SSABR detection, the MSC method will suffer from reduced frequency resolution and also from a reduced amount of STA within each subaverage to reduce the amount of noise. Both of these effects are troublesome due to the large amount of background noise. The noise at the lower frequencies is especially of interest due to its large magnitude relative to the rest of the spectrum and thus its spectral leakage should obscure the SSABR. To combat this effect, narrowband filtering can be used (since the signal frequency is known) before performing the FFT on each subaverage. Tang et al. (1993) used this technique (not with the MSC method) to reduce the effect of spectral

leakage on steady-state VEPs (the bandpass filter was 2 Hz wide). Tang noted that this method provided more accurate estimates of test sinusoid amplitude in white noise and in human EEG, however keep in mind that VEPs have higher SNR than SSABRs.

Another problem with the MSC method is the use of Artifact Rejection which results in a large amount of data rejection (i.e. longer recording time) when testing restless subjects. To combat this, Weighted Time Averaging can be applied *within each* subaverage for the MSC and MSC-WA detectors. This was also mentioned by Dobie et al. (1995) but never investigated; it was simply said that this method would result in increased computational overhead. Thus, both filtering and WTA modifications will be made to the MSC and MSC-WA methods and investigated later in section 6.2.

In addition, while frequency domain averaging is used to reduce the variance of the spectral estimates to increase signal detection, the noise floor is not actually reduced; only the variability of the noise floor is reduced (by a factor of N for averaging N uncorrelated spectra). This then imposes a lower limit on the level of the signal which can be detected. In this case, STA is needed to decrease the noise floor (although the variability of the noise floor is not reduced in this case). These concepts were discussed in Kulik (1995) and are shown below in Figures 3.2 and 3.3.

Thus another improvement would be to time average as many sweeps as possible. For a fixed data length, this is achieved by reducing the sweep length. However, the minimum sweep time is limited by the correlation properties of the noise in successive sweeps and the resolution (and spectral spreading) of the FFT. The reduced FFT resolution is due to the assumptions made about the data being zero outside of the measurement interval. However, due to knowledge about the data being measured, one can model the data more realistically (instead of truncating it to zero outside of the observation interval). This will result in a better spectral estimate once the parameters of the model have been estimated from the observed data (Kay et al., 1981). This is the premise of parametric spectral estimation. Thus, parametric spectral estimation can be used to overcome the frequency resolution and spectral leakage problem provided that the SNR is adequate. Parametric spectral estimation thus appears promising for SSABR detection and will be investigated in Chapters 5 and 6.

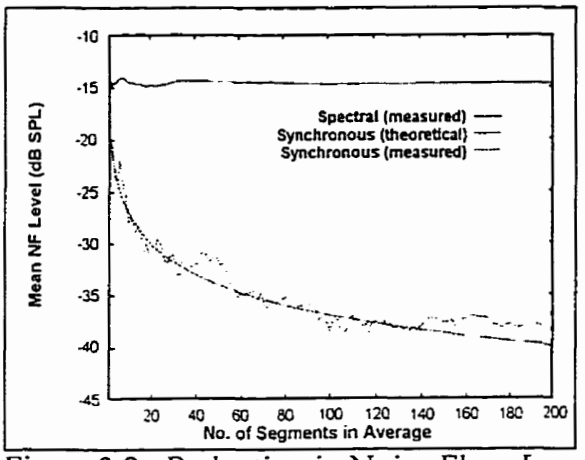

Figure 3.2. Reduction in Noise Floor Level by STA and spectral averaging (adapted from Kulik, 1995).

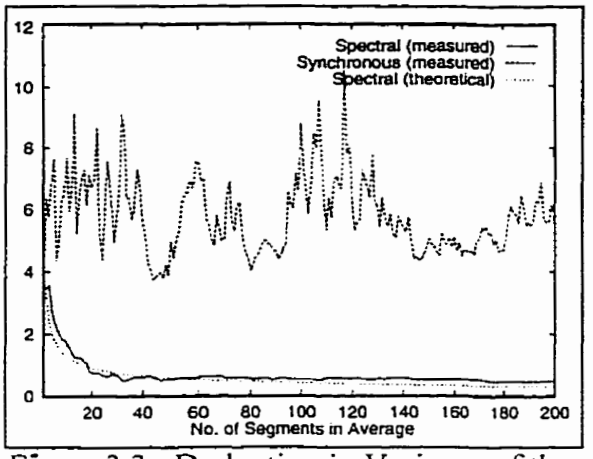

Figure 3.3. Reduction in Variance of the Noise Floor Level by STA and spectral averaging (adapted from Kulik, 1995).

In addition, current SSABR detection techniques do not take advantage of certain SSABR noise properties. In section 3.1, it was seen that the EEG noise may have a gaussian probability density function (i.e. I say 'may have' because of the conflicting reports on EEG gaussianity). If this is so then the gaussian nature of the noise can be exploited by using **H**igher-**O**rder **S**tatistics (HOS). HOS is very well known for its ability to suppress gaussian noise and thus may be able to improve the SSABR SNR. Thus HOS based parametric spectral estimation will also be examined in Chapter 5.

Finally, another area where an improvement in SSABR signal processing can be made is the replacement of Artifact Rejection. As mentioned in section 3.3.1, for particular noisy subjects, Artifact Rejection will result in excessive data elimination which will prolong test time. A possible solution to this problem is the use of Weighted Time Averaging. This will be further investigated in Chapter 4.

4.0 WEIGHTED TIME AVERAGING APPLIED TO SSABRs

This chapter investigates the use of **W**eighted **T**ime **A**veraging (WTA) as a replacement for the **A**rtifact **R**ejection (AFR) and **S**ynchronous **T**ime **A**veraging (STA) conventional preprocessing techniques. This is to overcome the limitations of AFR which are that it is susceptible to artifact contamination by low-level artifacts and that it throws away too many data buffers when testing restless individuals which can significantly prolong recording time. WTA is investigated because it has been found by many researchers to be robust in processing artifact contaminated data and it also results in reduced recording time.

WTA has only been applied once to 40 Hz SSABRs recorded from well behaved subjects using the MSC technique (see section 3.3.7.1). The weights were based on the inverse variance of the STA result in each subaverage. Each subaverage originally contained a 13 second data segment which was bandpass filtered from 30 to 100 Hz. The weighting provided marginal detection improvement (not statistically significant) over the original MSC method (Dobie et al., 1995), however, Dobie noted that if noise power fluctuations occurred primarily within time intervals shorter than the subaverage time (i.e. 13 seconds) then the weighting of individual sweeps should improve SNR much more than subaverage weighting would. This work then raises many questions. What if WTA was applied to much smaller data segments (i.e. the sweeps)? Would this provide any improvement for recordings obtained under quiet conditions (i.e. well behaved subjects)? Also, what would happen under noisy recording conditions (i.e. restless subjects)? Further, are there any other WTA algorithms which have better performance than the inverse variance technique? What happens if different bandpass filter settings were used for weight calculation to reflect the local situation in the signal region? Lastly, what kind of performance improvement can we expect over the conventional combination of AFR and STA?

To answer these questions, data was recorded from 5 adult subjects under quiet and noisy recording conditions. Four WTA algorithms were then adapted from the literature and compared on this data using the measures of **R**oot **M**ean **S**quare **E**rror (RMSE), SNR and computational complexity. These results will be shown in sections 4.3.6 (40 Hz SSABR results), 4.3.7 (80 Hz SSABR results) and 4.3.8, however, to begin with background information on WTA will be given in section 4.1. The four WTA algorithms will be discussed

in section 4.2 and evaluated in section 4.3. WTA findings in the literature will then be discussed in section 4.4 and conclusions will be made in section 4.5.

4.1 Weighted Time Averaging: Background

WTA involves a weighting algorithm that gives high weights to sweeps with low noise power (i.e. well-behaved, 'quiet' potentials) and low weights to sweeps with high noise power (i.e. 'noisy', artifact contaminated potentials) in order to minimize the contribution of artifacts to the final average. This 'noisiness' can be reflected by changes in the mean square value of the noise or the frequency structure of the noise. In addition, for stationary, well behaved noise, WTA should give an averaged result that is close to that obtained with STA.

The model of WTA assumes that the statistical noise properties are constant during one sweep, however, these properties change from sweep to sweep while the EP remains constant. WTA is analytically represented by (Gasser et al., 1983):

$$z_N(t) = \frac{1}{D} \cdot \sum_{i=1}^{N} g_i \cdot f_i(t) \tag{4.1}$$

where: $z_N(t)$ is the averaged result
$f_i(t)$ is the i^{th} sweep
g_i the weight for the i^{th} sweep
N is the number of recorded sweeps

$$D = \sum_{i=1}^{N} g_i \quad \text{is the sum of the weights} \tag{4.2}$$

WTA is not a biased signal estimator if the weights sum to unity (Bataillou et al., 1991). The performance of STA suffers with noisy potentials because the noise variance in the averaged result is (Bataillou et al., 1991):

$$\sigma_N(t) = \sum_{i=1}^{N} \frac{\sigma_i^2}{N^2} \tag{4.3}$$

where: σ_i^2 is the noise variance of the i^{th} sweep

This can be contrasted with the case of using STA on well-behaved, quiet potentials where all of the σ_i^2 are very similar which gives (note the change in summation index):

$$\sigma_N(t) = \sum_{j=1}^{N} \frac{\sigma_i^2}{N^2} = \frac{N\sigma_i^2}{N^2} = \frac{\sigma_i^2}{N} \tag{4.4}$$

Many researchers have found WTA to be more robust to 'noisy' potentials than STA and that WTA can aid in reducing the amount of recording time needed to detect a response (Gasser et al., 1983; Fan et al., 1991; Gerull et al., 1996; Miskiel et al., 1987; Lutkenhoner et al., 1985). However, while WTA has an advantage over STA for very noisy data, it was found to have no improvement in SNR over STA for well-behaved quiet potentials (Lutkenhoner et al., 1983).

4.1.1 Considerations in Weight Calculation

Care must be taken in weight calculation because blind application of WTA will result in the underestimation of signal amplitude. This underestimation depends on the degrees of freedom (df) involved in weight estimation, however, this underestimation can be neglected if the df are sufficiently increased by appropriate preprocessing methods such as data whitening (Lutkenhoner et al., 1985; Gasser et al., 1983; Gerull et al., 1996). The advantage of increasing the df is having more independent terms which can contribute to the value which is being calculated. The need for data whitening can be seen in the amplitude spectrum where each discrete frequency component contributes 2 df (1 df for each of the real and imaginary components of 1 FFT term), however, for data with low dominating frequencies (as in the case of the EEG) the effective df are small (Lutkenhoner et al., 1985). Whitening will thus flatten the spectrum and increase the df. In addition, Lukenhoner stated that the weighting factor should be found for a frequency band corresponding to the main frequency range of the signal of interest or else a nonstationarity in the desired frequency range may be masked by the stationary character of other frequency bands. The converse may also happen.

4.2 WTA Algorithms

The four WTA algorithms that were investigated were the Minimum Energy technique, the Inverse Variance technique, the Gerull technique and a vector nonlinear filtering technique. Each of these algorithms will now be described.

4.2.1 The Minimum Energy Technique

The Minimum Energy (ME) technique (Fan et al., 1991) argues that the weights should be chosen such that the energy of the average should be minimized. Fan justified this argument

by stating that since noise energy is so much larger than signal energy in EPs then a minimization of energy will mostly affect the noise energy. In their formulation[7], each sweep contains R samples and the i^{th} data sweep is represented by:

$$\bar{f}_i = \bar{n}_i + \bar{s} \tag{4.5}$$

where: $\bar{f}_i$ is the noisy EP recording (vector of size 1xR)
$\bar{n}_i$ is zero mean noise (vector of size 1xR)
$\bar{s}$ is the constant EP (vector of size 1xR)

If we record M such sweeps then the data can be represented in matrix form by:

$$\mathbf{F} = \mathbf{N} + \bar{e}\bar{s} \tag{4.6}$$

where: $\mathbf{F} = [\bar{f}_1^T, \bar{f}_2^T, \ldots, \bar{f}_M^T]^T$ (size is MxR) (4.7)

$\mathbf{N} = [\bar{n}_1^T, \bar{n}_2^T, \ldots, \bar{n}_M^T]^T$ (size is MxR) (4.8)

$\bar{e} = [1, 1, \ldots, 1]^T$ (size is Mx1) (4.9)

The weighting algorithm to minimize the energy in the final averaged waveform will produce a vector of weights ($\bar{w}_{min}$) which can be used to give the averaged result ($\bar{a}_{min}$):

$$\bar{a}_{min}{}^T = \bar{w}_{min}{}^T \mathbf{F} \tag{4.10}$$

where: $\bar{w}_{min} = [w_1, w_2, \ldots, w_M]^T$ (4.11)

This weight vector can be determined from the energy minimization of the averaged result with respect to the weight vector. This energy is given by:

$$E\{\bar{a}_{min}\, \bar{a}_{min}{}^T\} \tag{4.12}$$

Substituting equation 4.10 into equation 4.12 gives:

$$E\{(\bar{w}_{min}{}^T \mathbf{F})(\bar{w}_{min}{}^T \mathbf{F})^T\} \tag{4.13}$$

Now substituting equation 4.6 into equation 4.13 and expanding, the only energy term which depends on the weights is:

$$\bar{w}_{min}{}^T \mathbf{N}\mathbf{N}^T \bar{w}_{min} \tag{4.14}$$

Now taking the derivative of equation 4.14 with respect to the weights subject to the constraint $\bar{e}^T \bar{w}_{min} = 1$ leads to:

$$\bar{w}_{min} = \frac{(\mathbf{N}\mathbf{N}^T)^{-1}\bar{e}}{\bar{e}^T(\mathbf{N}\mathbf{N}^T)^{-1}\bar{e}} \tag{4.15}$$

[7] Symbols with arrows overhead (i.e. $\bar{v}$) represent vectors and bold-face characters represent matrices.

Since we don't have an exact expression for **N**, we can approximate **N** by **F** since the noise energy is so much larger than the signal energy. This gives:

$$\bar{w}_{min} = \frac{(\mathbf{F}\mathbf{F}^T)^{-1}\bar{e}}{\bar{e}^T(\mathbf{F}\mathbf{F}^T)^{-1}\bar{e}} \tag{4.16}$$

In the above formulation, the number of sweeps must be smaller than the number of points in a sweep (R) or else the matrix $\mathbf{F}\mathbf{F}^T$ will be singular and thus not invertible. If the number of sweeps needed to reduce the noise is greater than R then Fan recommended that the recorded sweeps be partitioned into K groups. WTA via the ME method can then be performed on each group and the resulting weighted averages from each group can be averaged using STA.

4.2.2 The Inverse Variance Technique

The Inverse Variance (V^{-1}) technique has been proposed by a number of researchers (Gasser et al., 1983; Dobie et al., 1994; Bataillou et al., 1991). In this approach, a noisy sweep will have a larger variance than a quiet sweep. To minimize the effect of the noisy sweeps to the averaged result, each weight can be based on the inverse variance of the corresponding sweep normalized by the sum of the inverse variances of all of the sweeps shown by:

$$\bar{w}_{var} = \frac{1}{\sum_{i=1}^{N}\frac{1}{\sigma_i^2}} \cdot \left[\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \ldots, \frac{1}{\sigma_N^2}\right]^T \tag{4.17}$$

Bataillou et al (1991) stated that a weight vector defined in this manner will maximize the SNR on the average and that with these weights, the variance of the averaged signal is:

$$V[z_N(j)] = E[(z_N(j) - E[z_N(j)])^2] = \frac{1}{\sum_{i=1}^{N}\frac{1}{\sigma_i^2}} \tag{4.18}$$

4.2.3 The Gerull Technique

Gerull et al. (1996) modified the V^{-1} technique because they noted that this method suffered from two problems. Firstly, the sample variance is itself a random variable whose variance is inversely proportional to the df (in variance estimation) and secondly if the signal is included in weight calculation then there is a risk of biasing which can result in a reduction of

the signal amplitude. To address these problems they increased the df for variance estimation and eliminated the signal when calculating weights. They stated that their procedure avoids signal underestimation and improves weight calculation. The procedure includes the following steps:

1) Select pairs of sweeps and form two sequences; one sequence is equal to the sweep difference and the other equals the sweep sum.
2) Whiten the sweep difference in each pair (using an adaptive whitening filter) and obtain the reciprocal sample variance which is used as the weight.
3) Weight the sweep sum in each sweep pair by the appropriate weight calculated in step 2.

In this procedure, the pair difference was used to form the variance estimate because the pair difference would not contain the signal component[8]. Adaptive whitening was used to increase the df to improve the variance estimate. Furthermore, correlation between the sweeps in each pair was made small by choosing sweeps separated by at least one sweep. The omitted sweeps were then used for another independent signal estimation and the results of both estimations were averaged to yield the final result.

4.2.4 Vector Nonlinear Filtering

This technique, proposed by Bezerianos et al. (1995), considers each sweep to be a vector with components at time instants 1, 2, ..., K ($\vec{f} = [f_{(1)}, f_{(2)}, \ldots, f_{(K)}]$) which is represented by a single point in K dimensional vector space R^K. Each sweep vector is then inputted to a nonlinear filter algorithm which computes a weight for the sweep based on its similarity to the rest of the recorded sweeps. Vectors that differ significantly from other vectors are recognized as artifacts and given small weights.

The method, termed Data Dependent Weighted Averaging based on the L2 norm (abbreviated here as VNF: **V**ector **N**onlinear **F**iltering), assigns each input vector ($\vec{f}_i$) a scalar D_i that corresponds to the aggregate distance of that vector to the rest of the N recorded vectors. Thus the smaller the D_i the more reliable the corresponding $\vec{f}_i$. The D_i value is given by:

[8] Gerull stated that the noise is assumed normally distributed and uncorrelated which makes the statistical properties of the sum and difference of the noise in the pair of sweeps equivalent.

$$D_i = \sum_{j=1}^{N} \left\| \bar{f}_i - \bar{f}_j \right\| \tag{4.19}$$

The norm in equation 4.19 is the Euclidean norm which is calculated as:

$$\left\| \bar{f}_i(n) \right\| = \sqrt{\sum_{n=1}^{K} f_i^2(n)} \tag{4.20}$$

The particular $\bar{f}_i$ for which D_i is minimum is called the vector median of the set of sweeps. The weights (C_i) are then calculated by:

$$C_i = \frac{\frac{1}{D_i^2}}{\sum_{j=1}^{N} \frac{1}{D_j^2}} \tag{4.21}$$

Inversion of D_i insures that vectors close to the vector median take high weights which is important for suppressing impulsive noise. The summation of the weighted vectors then gives the final averaged result.

4.3 Evaluation of WTA Algorithms

The previously described WTA algorithms were compared to STA using pseudo-real data consisting of a simulated sinusoid (i.e. the SSABR signal) added to recorded EEG (no AFR was performed). The pseudo-real data was then segmented and passed as input to the WTA algorithms as well as to the STA algorithm. Performance was then determined for each averaging method by the measures of RMSE and SNR, and variants thereof versus the number of sweeps averaged. The best WTA method was then compared to the combination of AFR and STA. This was done for both 40 and 80 Hz simulated SSABRs and quiet and noisy EEG data. The section outlines the methods used and the results obtained.

4.3.1 Data Acquisition: Methods, Setup and Subjects

Pure EEG noise was recorded from 5 male adult subjects aged 23 to 33. The subjects had behavioral hearing thresholds less than 20 dB HL at 1 kHz and did not have any history of hearing problems. Two recordings were made in which a 70 dB SPL pure tone at 1 kHz (would not evoke an SSABR) was delivered, using an insert probe phone, to the right ear of each

subject. Refer to Appendix C for a description of the experimental setup and the instrumentation used.

During the first recording, the subjects were instructed to remain very still to obtain a quiet EEG recording. This data will be called 'Quiet EEG Data'. During the second recording, the subjects were uneasy and made impulsive noise by clenching their teeth, raising their eyebrows, coughing, yawning and moving any other facial muscles. This movement was for a random duration with a random interval of time between two successive artifact generations ranging from one or two seconds to five seconds. This data will be called 'Noisy EEG Data'. These test cases would be the extreme situations for an EP recording. The Quiet EEG Data would represent a totally compliant subject and the Noisy EEG Data would represent a very uncooperative subject.

4.3.2 Data Preparation

A simulated sinusoid, representing the SSABR, was added to the recorded EEG to form the 'pseudo-real' test data. This was repeated twice; the first time the sinusoidal frequency was at 40 Hz and the second time it was at 80 Hz. In both occasions the sinusoidal amplitude was 300 nV peak which represents a 40 Hz SSABR of moderate amplitude and a high amplitude 80 Hz SSABR. This amplitude was used to get meaningful results with a reasonable amount of data. The data was then passed through a notch filter to remove line noise which was very prominent in the recordings. The line noise was stable at 60 Hz with amplitudes varying from 1 to 4 μV across subjects. A first order Butterworth IIR notch filter, in the non-phase shift configuration (Antoniou, 1993), with a bandstop region of 58 to 62 Hz and a 40 dB notch was used to attenuate the line noise.

The first 0.4 seconds of each recording were discarded to account for startup transients in the recording. The next 90 seconds of data were taken and partitioned into 2.4 second data blocks which were further partitioned into 0.3 second sweeps (see Appendix F for sweep length choice). This length is sufficiently sensitive to noise artifacts which are short in duration. WTA was then performed on the data starting with 2.4 seconds of data and iteratively adding 2.4 seconds of data which enabled RMSE and SNR results to be obtained every 2.4 seconds. For the cases where AFR was used, buffers of data with amplitudes greater than $\pm$ 40 μV (Lins

et al., 1995) were thrown out. In the cases where pre-filtering was used, *ONLY* for weight calculation, the filter was a 2nd order, IIR Butterworth filter in the non-phase shift configuration. The passband was 35 to 55 Hz for the 40 Hz SSABR and 65 to 95 Hz for the 80 Hz SSABR.

4.3.3 Test Cases Considered

The test cases involved WTA and STA performance analysis under the Quiet and Noisy EEG Data cases with no AFR. Since STA is usually conducted after AFR, this combination was also examined for the Quiet and Noisy EEG Data cases. Finally the analysis for the WTA algorithms was repeated with bandpass filtering applied to the data *ONLY* for weight calculation in the 35 to 55 Hz range for the 40 Hz SSABR and the 65 to 95 Hz range for the 80 Hz SSABR. All of these test cases considered are listed in Table 4.1.

TABLE 4.1. Test Cases for WTA Investigation.

Test Case	EEG data type	SSABR Amplitude (nV peak)	SSABR Frequency (Hz)	AFR?	Prefilter data for weight calculation only**
1	Noisy	300	40	No	No
2	Quiet	300	40	No	No
3*	Noisy	300	40	Yes	No
4	Noisy	300	40	No	Yes; (35 to 55 Hz)
5	Quiet	300	40	No	Yes; (35 to 55 Hz)
6*	Quiet	300	40	Yes	No
7	Noisy	300	80	No	No
8	Quiet	300	80	No	No
9*	Noisy	300	80	Yes	No
10	Noisy	300	80	No	Yes; (65 to 95 Hz)
11	Quiet	300	80	No	Yes; (65 to 95 Hz)
12*	Quiet	300	80	Yes	No

* This test case is done for STA only.

** The pre-filtering case for weight calculation does not apply to STA.

4.3.4 Implementation of WTA Algorithms

Modifications were made only to the ME and Gerull methods. For the ME method, the minimum energy weighted averages from each 2.4 second block were combined using WTA based on the V^{-1} technique instead of using STA. This was reasonable because WTA applied to sweeps in each 2.4 second block may yield results that are also very variable in variance across the blocks; i.e. one block may be completely contaminated with artifacts meanwhile other blocks can be artifact free. Thus using V^{-1} weighting should be beneficial in this situation.

The Gerull method was augmented by normalizing the weights based on all of the weights calculated; i.e. all sweeps were considered at once instead of performing the method on two separate groups of sweeps and then averaging the two results at the end. This was done for the same reason as that given for the ME method, however in this case the effect would only be prominent for a low number of sweeps averaged. Furthermore, the whitening in the method was done based on AR parameters determined for each 0.3 second sweep separately. The Yule-Walker (autocorrelation) method was used to estimate the AR parameters using the biased autocorrelation function and the MDL AR model order estimator was used based on results shown by Simpson et al., (1995), Tseng et al., (1995), Vaz et al., (1987) and Steinberg et al. (1985)[9]. Before the sweeps could be whitened, the data was upsampled by a factor of two and downsampled by a factor of five because the original data, which was sampled at 490 Hz and lowpass filtered at 100 Hz, had a 'deadband' in the 100 to 245 Hz region which adversely affected the whitening of the data between 10 and 100 Hz.

4.3.5 Performance Evaluation: Methods

For each test case, the RMSE between the WTA result and the simulated sinusoid was calculated. In addition, the SNR was calculated for each averaged result by performing the FFT and then taking the power at the response frequency and dividing it by the average of the noise power values found within 10 Hz on either side of the response frequency. Before the SNR and RMSE values were calculated, the averaged result was bandpass filtered from 25 to 55 Hz (for the 40 Hz SSABR tests) or 65 to 95 Hz (for the 80 Hz SSABR tests) to allow the amount of noise around the frequency region of interest to be directly investigated. The SNR and RMSE results were then averaged across subjects. In addition, for each subject, the difference between the RMSE for each WTA method and STA was computed. These results are referred to as DRMSE (see equation 4.22). This was also repeated using SNR values; these results are referred to as DSNR (see equation 4.23). Furthermore, the DRMSE and DSNR values were averaged across subjects.

[9] Their work basically showed that the AIC, FPE and CAT model order estimators were similar and gave disperse results meanwhile the MDL order estimator was more consistent. The work by Vaz et al. (1987) also suggested it was better to work with shorter time segments and the work by Tseng et al. (1995) showed that AR modeling was better suited than ARMA modeling in describing the EEG under a variety of different recording conditions.

$$DRMSE = RMSE_{WTA} - RMSE_{STA} \quad (4.22)$$

$$DSNR = SNR_{WTA} - SNR_{STA} \quad (4.23)$$

For comparing the performance between pre-filtering versus no pre-filtering for weight calculation, the RMSE and SNR percentage change were calculated according to:

$$\left(\frac{RMSE_{NPF} - RMSE_{PF}}{RMSE_{NPF}}\right) \cdot 100\% \quad (4.24)$$

$$\left(\frac{SNR_{PF} - SNR_{NPF}}{SNR_{NPF}}\right) \cdot 100\% \quad (4.25)$$

where: NPFFWC means **N**o **P**refiltering **F**or **W**eight **C**alculation
PFFWC means **P**refiltering **F**or **W**eight **C**alculation

These values were also averaged across the subjects. For RMSE, a positive percentage change indicates a *decrease* in RMSE when using PFFWC which is beneficial; conversely a positive percentage change for SNR means that SNR has *increased* when using PFFWC. The computational overhead for each WTA method will also be compared. The best WTA method will then be chosen based on the performance measures and computational complexity and will be compared to the combination of AFR and STA.

4.3.6 40 Hz SSABR Results

The performance results for the Noisy EEG Data case (without PFFWC) are shown in Figure 4.1. The RMSE and SNR measures indicate that WTA is considerably better than STA with no AFR for EEG data recorded under noisy conditions which is to be expected. Results also show that, in terms of RMSE and DRMSE, the ME, V^{-1} and VNF algorithms perform similarly and very well with a slight advantage to the ME method over the V^{-1} method. Meanwhile the Gerull and STA methods perform similarly and badly. In terms of SNR and DSNR, the ME, V^{-1}, and Gerull techniques all perform similarly with a slight advantage to the ME method over the V^{-1} method. Based on these results either the ME or V^{-1} method can be chosen for 40 Hz SSABRs recorded under noisy conditions.

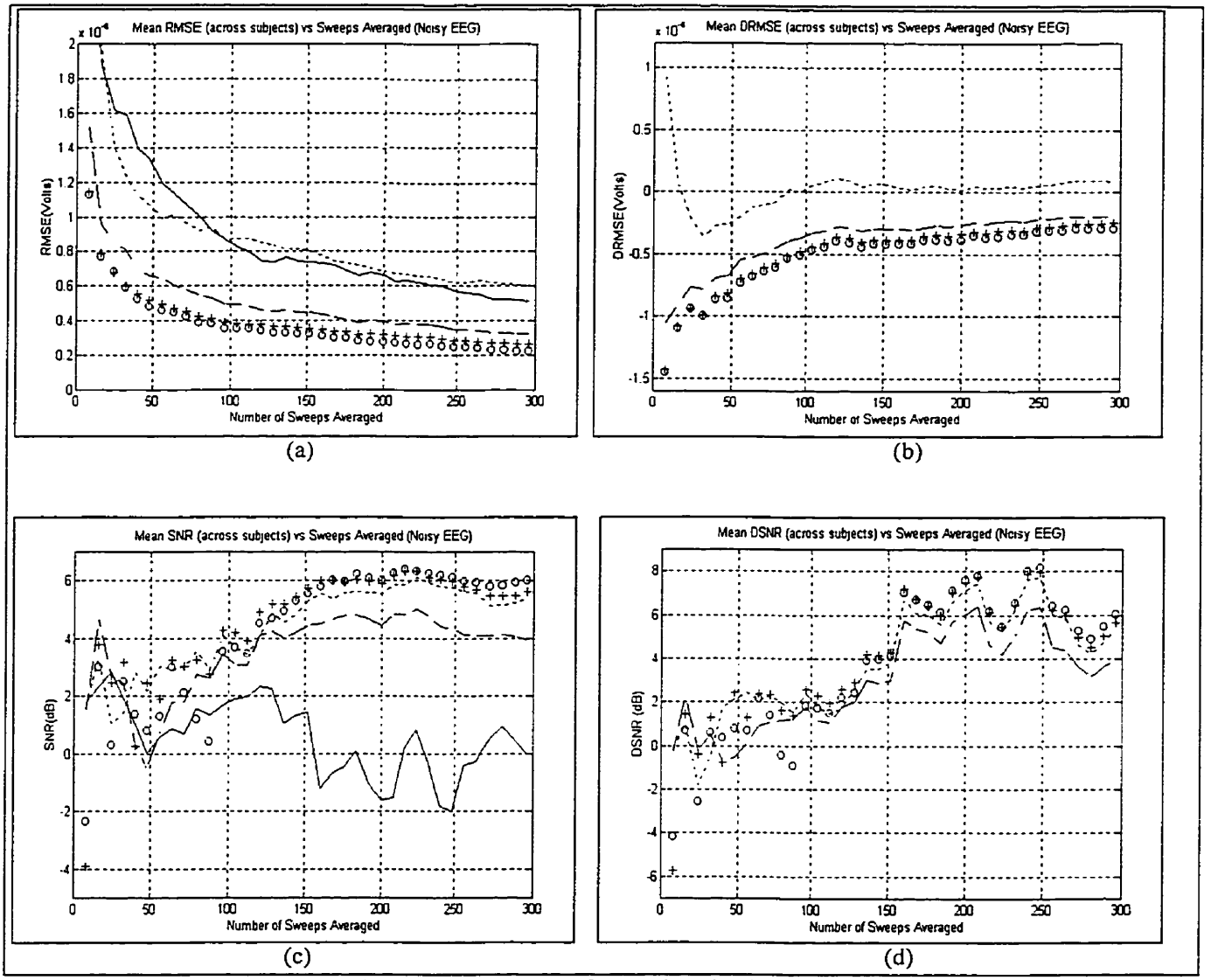

Figure 4.1. Mean WTA results across subjects for Noisy EEG Data (40 Hz SSABR) (a) RMSE (b) DRMSE (c) SNR (d) DSNR (STA '—', ME 'oo', Gerull '- -', V^{-1} '++', VNF '– –').

The above analysis was repeated based on PFFWC between 35 and 55 Hz and using these weights with the originally recorded data. The results for this test case are shown in Table 4.2 for mean percentage change in RMSE and SNR across subjects. Recall that positive values mean that PFFWC is beneficial. Looking at the RMSE percentage change, we see that the ME and Gerull methods suffer meanwhile the V^{-1} and VNF methods decrease in RMSE a bit with a larger decrease (5% on average) for the V^{-1} method. In terms of SNR, all WTA methods decrease in performance with PFFWC with the largest decreases occurring for the ME and Gerull methods. The VNF method actually increases a bit in SNR with PFFWC, however

the increase is not enough to match the SNR of the other WTA methods with NPFFWC. Based on these results PFFWC should not be used for WTA for 40 Hz SSABRs recorded under noisy conditions.

Table 4.2. Percentage Change in RMSE and SNR when using PFFWC (35 to 55 Hz) for the Noisy EEG Data case (40 Hz SSABR).

Number of Sweeps Averaged	RMSE				SNR			
	ME	Gerull	V^{-1}	VNF	ME	Gerull	V^{-1}	VNF
24	-3.52	2.92	5.87	-1.33	-845.2	88.93	12.31	19.73
48	-15.40	-3.92	3.72	1.80	-187.36	-190.29	7.85	157.59
72	-19.94	-11.01	4.34	0.31	-130.07	22.33	-31.66	48.88
96	-19.40	-12.27	2.97	0.62	-110.88	0.39	-14.79	-9.70
120	-15.25	-15.79	2.74	1.5	-111.09	2.94	-14.55	-3.39
144	-17.52	-11.38	2.40	2.38	-69.57	-12.15	-11.06	0.74
168	-8.74	-7.50	7.30	3.31	-69.07	-7.09	-8.22	1.57
192	-7.27	-12.91	8.24	3.18	-62.31	-12.74	-6.82	3.58
216	-10.32	-17.01	6.32	3.78	-69.02	-15.65	-8.47	2.44
240	-16.81	-14.36	5.22	4.29	-76.54	-13.60	-6.53	5.34
264	-25.77	-14.73	1.90	2.14	-80.65	-13.46	-4.56	5.59
288	-35.22	-13.20	0.19	2.32	-81.49	-13.33	-2.34	6.04

The performance results for the Quiet EEG Data case are shown in Figure 4.2. These results show that in terms of RMSE and DRMSE, the V^{-1}, ME and VNF methods performed similarly with STA with a slight advantage to the ME method while the Gerull method performed badly. In terms of SNR and DSNR, all WTA methods performed similarly to STA except for the ME method which performed slightly worse. In addition there was a slight advantage (early on; i.e. for number of sweeps averaged < 150) for the V^{-1}, VNF and Gerull methods in terms of SNR (see Figure 4.2d).

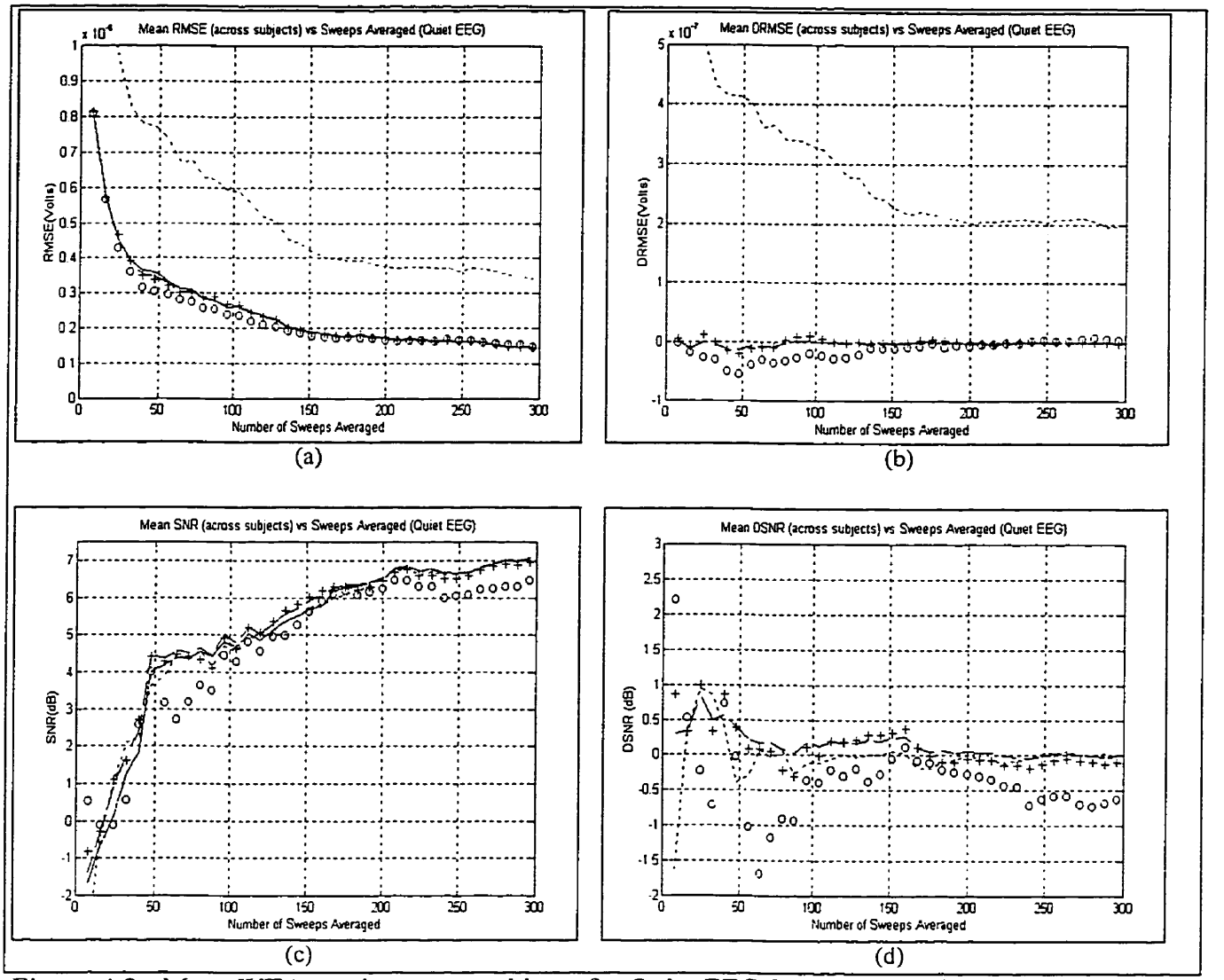

Figure 4.2. Mean WTA results across subjects for Quiet EEG Data (40 Hz SSABR) (a) RMSE (b) DRMSE (c) SNR (d) DSNR (STA '—', ME 'oo', Gerull '- -', V^{-1} '++', VNF '– –').

PFFWC (35 to 55 Hz) was also applied to the Quiet EEG Data case. The percentage change in RMSE and SNR for each WTA method is shown in Table 4.3. PFFWC increased the RMSE significantly for the ME method, moderately for the Gerull method and slightly for the V^{-1} and VNF methods. For SNR, there were decreases for each WTA method with the ME method being affected the most and the VNF and Gerull methods the least. Thus PFFWC is not suitable for 40 Hz SSABRs recorded under quiet recording conditions.

Table 4.3. Percentage Change in RMSE and SNR when using PFFWC (35 to 55 Hz) for the Quiet EEG Data case (40 Hz SSABR).

Number of Sweeps Averaged	RMSE				SNR			
	ME	Gerull	V^{-1}	VNF	ME	Gerull	V^{-1}	VNF
24	-12.62	-10.28	-0.79	-0.46	1936.27	0.66	21.36	78.45
48	-3.78	-6.42	1.50	-0.46	-121.44	-8.14	-9.89	-1.65
72	-8.55	-4.65	0.66	-1.95	-158.03	-2.48	-19.11	-1.99
96	-9.36	0.61	3.26	0.27	-125.14	4.10	-18.82	-3.02
120	-22.38	-1.78	-2.70	-1.63	-123.56	1.64	-20.82	-3.32
144	-26.12	-5.75	-5.61	-2.06	-95.97	-4.54	-16.74	-2.48
168	-30.04	-8.37	-5.35	-1.51	-79.95	-6.88	-12.90	-1.72
192	-26.13	-10.16	-5.16	-1.54	-58.71	-3.98	-8.26	-0.05
216	-24.89	-11.84	-4.65	-1.44	-51.81	-3.65	-5.75	0.17
240	-18.35	-9.49	-1.63	-1.10	-45.03	-2.39	-2.66	1.28
264	-19.47	-7.37	-2.97	-1.84	-39.67	-4.20	-3.34	0.81
288	-17.46	-9.17	-2.36	-0.61	-32.08	-4.95	-2.37	1.19

4.3.7 80 Hz SSABR Results

The performance results for the Noisy EEG Data case with NPFFWC are shown in Figure 4.3. From the RMSE and DRMSE results we can see that the best methods are the ME and V^{-1} methods with the VNF method close behind. STA is far behind once more, however the Gerull method performs better than STA (this was not seen in the 40 Hz SSABR Noisy EEG Data case). In terms of SNR and DSNR, the ME and V^{-1} methods perform very similarly with the Gerull method performing similarly at first but then dropping off a bit in performance. Conversely, the VNF method performs badly at first but then improves towards the end of the test. STA does not improve with an increase in the number of averaged sweeps as expected. We should also note that the RMSE for this case is lower than it was in the 40 Hz SSABR Noisy EEG Data case for all WTA methods and that STA performs better in this case than the corresponding 40 Hz SSABR case (i.e. compare Figures 4.3b and 4.3d with Figures 4.1b and 4.1d).

With PFFWC, from 65 to 95 Hz, the percentage changes in RMSE (see Table 4.4) indicate that the ME method improves at the beginning of the test but then halfway through it suffers. The Gerull method also suffers throughout but only very slightly. Conversely, the V^{-1} and VNF methods improve throughout the test with the V^{-1} method improving by a larger amount. In terms of SNR, we see the same trends as with RMSE. The ME method improves near the beginning but then suffers for the duration of the test, meanwhile, the Gerull method,

after an initial degradation improves slightly for the rest of the test. The VNF and V^{-1} methods both show improvement. Furthermore, the SNR improvement of the V^{-1} method with PFFWC is so large that V^{-1} with PFFWC performs better than all other methods with or without pre-filtering for the 80 Hz SSABR Noisy EEG Data case. In fact V^{-1} with PFFWC is better than the ME method by about 1 dB up to 190 sweeps averaged and then 0.5 dB afterwards. Thus PFFWC is beneficial for the V^{-1} and VNF methods for the 80 Hz SSABR recorded under noisy conditions.

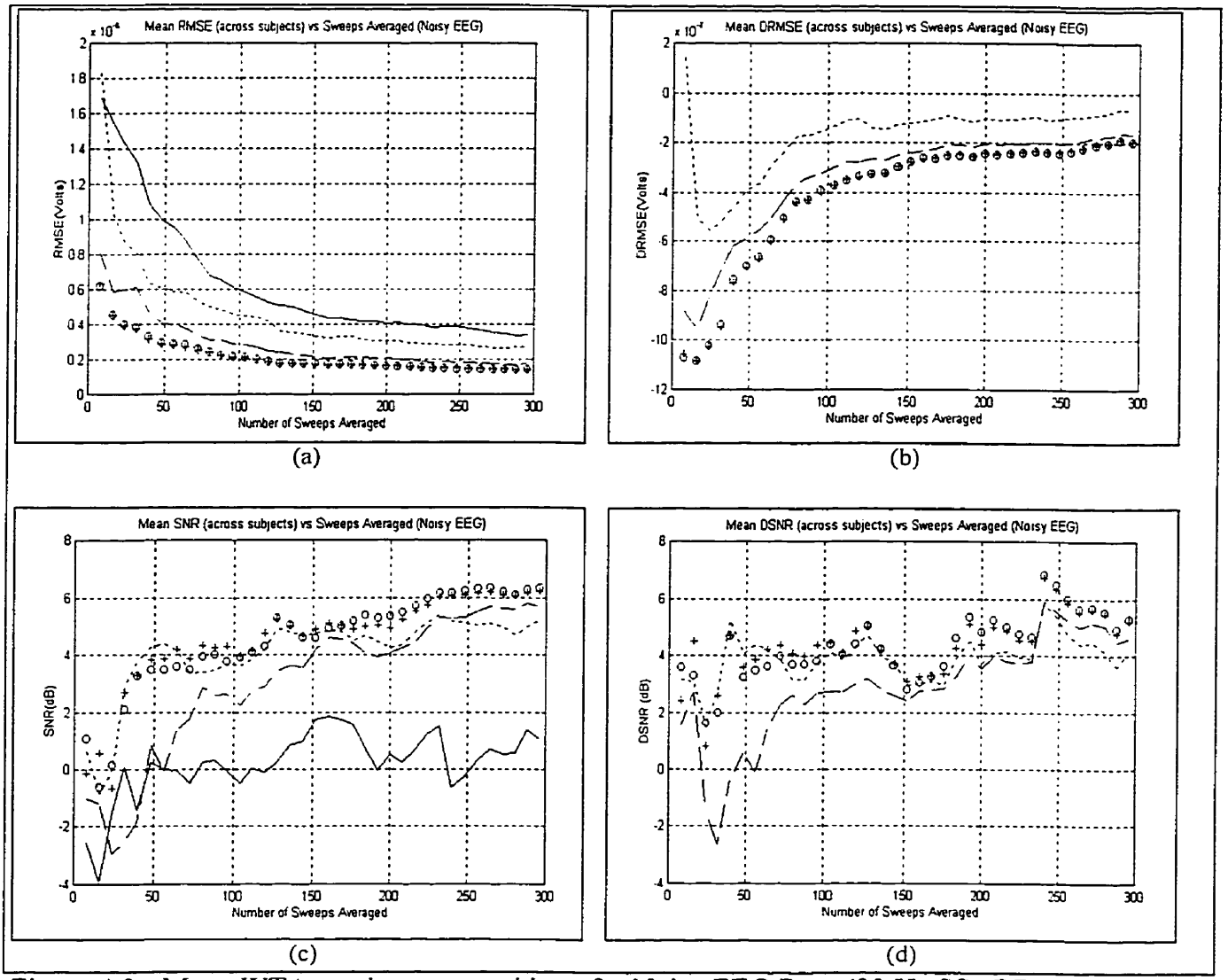

Figure 4.3. Mean WTA results across subjects for Noisy EEG Data (80 Hz SSABR) (a) RMSE (b) DRMSE (c) SNR (d) DSNR (STA '—', ME 'oo', Gerull '- -', V^{-1} '++', VNF '– –').

The performance results for the Quiet EEG Data case are shown in Figure 4.4. The RMSE and DRMSE results indicate that all WTA methods (except for the Gerull method)

perform similarly to STA with a very slight disadvantage to the ME method. In terms of SNR and DSNR, all WTA methods perform similarly to STA with an advantage to the ME method. In fact, early on in the test, the ME method outperforms the other methods by about 0.5 dB. Furthermore, comparing these results to those from the 40 Hz Quiet EEG Data case we note a reduction in the measure of RMSE and an increase in the measure of SNR.

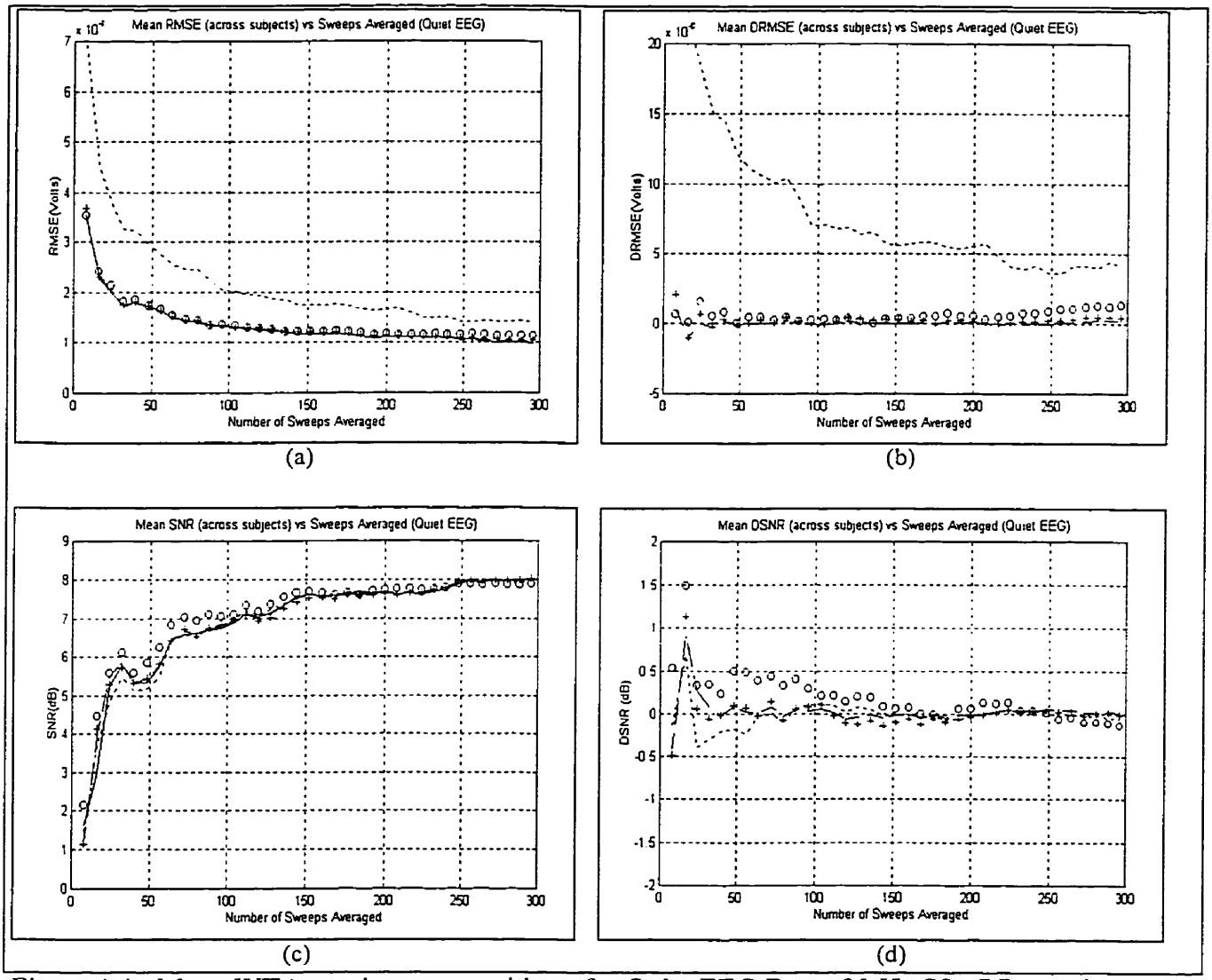

Figure 4.4. Mean WTA results across subjects for Quiet EEG Data (80 Hz SSABR) (a) RMSE (b) DRMSE (c) SNR (d) DSNR (STA '—', ME 'oo', Gerull '- -', V^{-1} '++', VNF '– –').

The results for PFFWC from 65 to 95 Hz are shown in Table 4.5. In terms of RMSE percentage change, we see that the ME and Gerull methods both degrade in performance with a bigger degradation for the ME method. Pre-filtering also causes the V^{-1} method to perform a

little bit worse meanwhile the VNF performance is hardly affected. In terms of SNR percentage change, we see that the ME method suffers badly at first but then only slightly afterwards. The Gerull, V^{-1} and VNF methods suffer slightly in performance with the first 70 or so averaged sweeps but then the performance is pretty much unchanged. Thus PFFWC offers no performance improvement for the 80 Hz SSABR recorded under quiet recording conditions.

Table 4.4. Percentage Change in RMSE and SNR when using PFFWC (65 to 95 Hz) for the Noisy EEG Data case (80 Hz SSBAR).

Number of	RMSE				SNR			
Sweeps Averaged	ME	Gerull	V^{-1}	VNF	ME	Gerull	V^{-1}	VNF
24	22.58	-1.66	25.72	15.97	-1089.98	-9215.99	-545.77	-108.86
48	21.16	-4.86	20.15	14.62	5.16	-18.50	24.74	140.39
72	16.97	-3.57	22.03	13.54	12.88	-23.84	26.18	60.87
96	6.52	-5.85	15.96	14.56	0.16	-7.59	16.64	28.30
120	-2.05	-2.94	11.44	12.45	-20.92	14.04	14.72	23.41
144	-7.18	-4.06	10.87	10.56	-26.60	8.81	25.95	11.06
168	-4.42	-6.81	9.72	10.55	-17.78	2.24	22.31	10.52
192	-6.97	-6.54	10.01	10.74	-12.45	2.08	23.50	15.26
216	-9.82	-7.17	7.20	9.41	-6.27	1.45	13.76	11.18
240	-16.80	-3.97	4.58	10.25	-8.48	1.23	8.23	10.01
264	-19.30	-4.73	1.78	9.84	-10.98	0.86	5.52	6.83
288	-20.72	-5.35	0.49	8.76	-8.74	-2.86	3.31	4.45

Table 4.5. Percentage Change in RMSE and SNR when using PFFWC (65 to 95 Hz) for the Quiet EEG Data case (80 Hz SSABR).

Number of	RMSE				SNR			
Sweeps Averaged	ME	Gerull	V^{-1}	VNF	ME	Gerull	V^{-1}	VNF
24	0.33	-9.02	2.23	-0.25	-50.10	-6.39	-17.52	-8.03
48	-5.41	-4.45	-0.41	-1.10	-26.27	-3.32	-10.05	-4.41
72	-9.45	-5.88	-1.81	-0.56	-12.10	-6.35	-5.22	-2.66
96	-13.97	-5.29	-1.20	0.378	-4.95	-4.03	-1.95	-1.25
120	-12.36	-1.75	1.54	1.03	-1.38	-1.99	0.75	-0.47
144	-21.72	-7.62	-0.25	1.44	-4.54	-2.57	1.06	-0.08
168	-20.32	-10.56	-0.58	1.18	-2.60	-1.48	1.26	-0.28
192	-26.21	-9.59	-4.63	-0.18	-4.68	-1.12	0.37	-0.50
216	-26.58	-8.56	-7.03	-1.18	-5.10	-0.31	-0.35	-0.67
240	-26.51	-9.16	-4.99	0.73	-6.13	0.02	-0.42	-0.36
264	-26.60	-11.06	-5.06	0.86	-5.05	-0.47	0.16	-0.16
288	-28.24	-12.29	-5.34	1.37	-5.01	-0.67	0.23	0.03

4.3.8 Computational Complexity of the WTA methods

In terms of computational complexity, we consider data consisting of N sweeps with K sample points per sweep. Each WTA method will have to weight each sweep and sum the weighted sweeps which requires NK multiplications and NK additions. Since this is common to each WTA method it will be ignored. The computational complexity for each method is shown in Table 4.6. Appendix D provides more detail on the derivation of the computational complexity for each algorithm. In the past, the most expensive computation was multiplications and from this perspective we see that the least expensive algorithm in the V^{-1} method. The next most computationally expensive method is the ME method followed by the VNF and Gerull methods.

Table 4.6. Computational Complexity for each WTA method.

WTA method	Number of Multiplications	Number of Additions	Number of Divisions	Number of Square Roots	Number of Logs	Number of Inverses*
V^{-1}	2NK	3NK	4N	0	0	0
ME	KN^2+N	KN^2+N	N	0	0	1 (NxN)
Gerull	$K^2N/2$ + (11/7)NK	$K^2N/2$ + (18/7)KN	N(3 + K/14)	0	NK/7	0
VNF	$2N^2K+N$	$2N^2K+N$	2N	N^2	0	0

* This column contains the number of matrix inversions and the size of the matrix inverted in brackets.

4.3.9 WTA Algorithm Choice

For the 40 Hz SSABR case, the performance analysis of the WTA methods in *both* the Noisy and Quiet EEG Data cases showed that the V^{-1} method was always near the top (if not at the top) in terms of performance. In addition, PFFWC did not improve the performance of any WTA algorithm. Furthermore, Table 4.6 indicates that the V^{-1} method has the least amount of computational overhead. Thus, for the 40 Hz SSABR, the best WTA method to use for SSABR preprocessing is the V^{-1} method without PFFWC.

In terms of the 80 Hz SSABR, the choice is not as clear cut. For the Noisy EEG Data case, the V^{-1} method with PFFWC performs the best by about 1 dB over the ME method without PFFWC. However, in the Quiet EEG Data case, the ME method outperforms the other WTA methods in terms of SNR by about 0.5 dB early on (in the 30 to 80 sweeps averaged range), however after that all methods performed very similarly. The questions to consider here are whether more noisy recordings than quiet recordings will be encountered in practice and

also does the computational overhead of the ME method justify the 0.5 dB SNR performance improvement? Alternatively, the strategy could be to have two WTA methods for 80 Hz SSABR testing. For testing individuals who are quiet, the ME method could be used and when testing particularly noisy individuals such as infants, the V^{-1} method with PFFWC could be used.

4.3.10 Comparison of The Best WTA methods to STA with AFR

In this section, the performance of the best WTA method on the raw recorded data (i.e. no AFR) will be compared to the performance of STA with AFR. Only the RMSE and SNR results will be shown because the DRMSE and DSNR results added no extra information. For the 40 Hz SSABR, the V^{-1} method will be used in both the Quiet and Noisy EEG Data cases. For the 80 Hz SSABR, the V^{-1} method with PFFWC will be used for the Noisy EEG Data case and the ME method will be used for the Quiet EEG Data case.

Firstly, the extra amount of data needed for each subject to obtain the required 90 seconds of data for WTA analysis is shown in Table 4.7. This holds true for both the 40 and 80 Hz SSABR test cases. Table 4.7 indicates that for the Noisy EEG Data case, on average an extra 32.42 seconds of recording time was needed which represents an extra 36% of recording time. Meanwhile, for the Quiet EEG Data case, the subjects were well behaved and generated very few artifacts and thus only a very slight bit of extra recording time was needed.

Table 4.7. Extra recording time (in seconds) when using Artifact Rejection.

EEG Data	Subject					Extra Recording	Recording Time
Case	1	2	3	4	5	Time	% Increase
Noisy	22.7	86.5	30.8	4.9	17.2	32.42	36
Quiet	0	0	0	0.1	0.4	0.1	0.1

The RMSE and SNR results for the 40 Hz SSABR Noisy EEG Data case are shown in Figure 4.5. The plot shows that STA with AFR outperforms the V^{-1} method, however, as Table 4.7 indicated, STA with AFR required on average an extra 32.42 seconds (extra 36%) of data. Factoring this in, it is clear that the V^{-1} method is the better method to use in noisy recording conditions for 40 Hz SSABRs.

The RMSE and SNR results for the 40 Hz SSABR Quiet EEG Data case are shown in Figure 4.6. The results show that the V^{-1} method and STA with AFR perform very similarly with a slight edge to the V^{-1} method. This result is peculiar since when comparing V^{-1} and STA (with no AFR) for the Quiet EEG Data case, the SNR results were closer. In this case (Figure 4.6b) the difference is attributed to the fact that AFR was used in which case a few buffers of data were rejected for a few subjects. This had a cascade effect on the composition of the noise for the resulting sweeps which were formed for input to the STA algorithm.

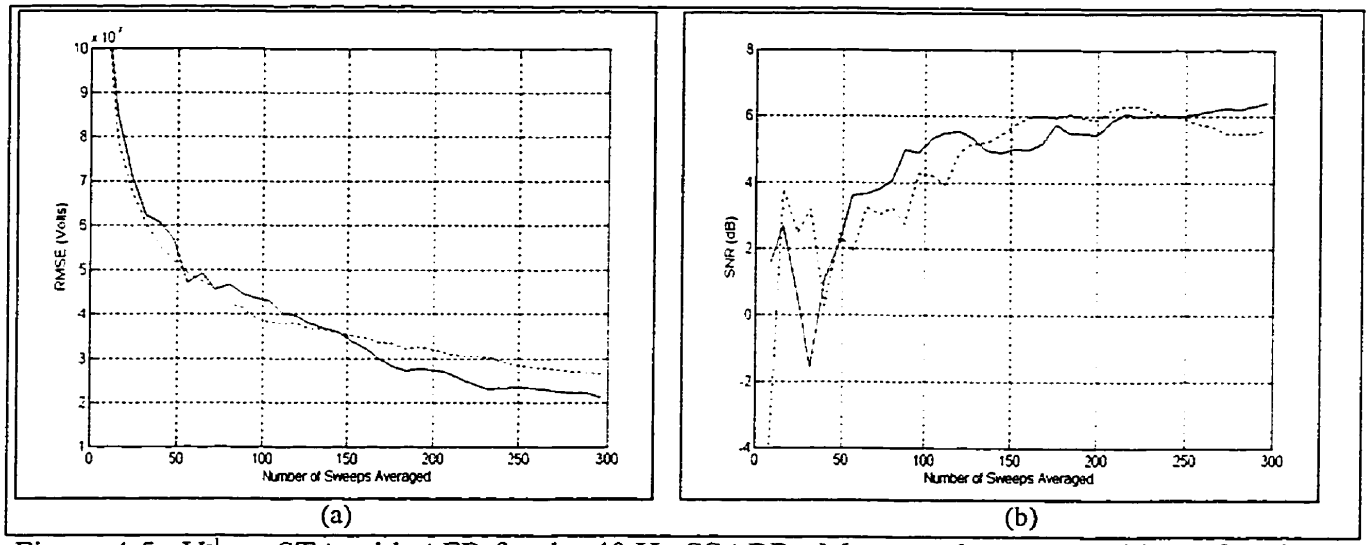

Figure 4.5. V^{-1} vs. STA with AFR for the 40 Hz SSABR. Mean results across subjects for Noisy EEG Data: (a) RMSE (b) SNR (V^{-1} '- -', STA with AFR '—').

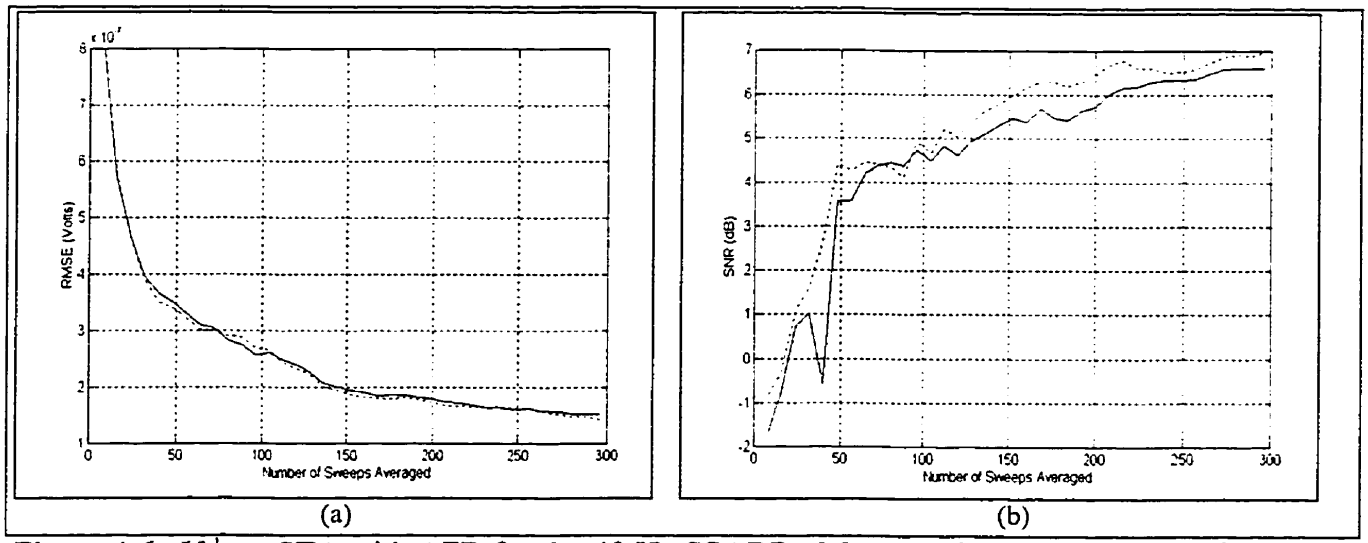

Figure 4.6. V^{-1} vs. STA with AFR for the 40 Hz SSABR. Mean results across subjects for Quiet EEG Data: (a) RMSE (b) SNR (V^{-1} '- -', STA with AFR '—').

The RMSE and SNR results for the 80 Hz SSABR Noisy EEG Data case are shown in Figure 4.7. The results show that the V^{-1} with PFFWC method outperforms the STA with AFR in both RMSE and SNR although this performance improvement diminishes by the end of the test. This is already on top of the reduction in recording time.

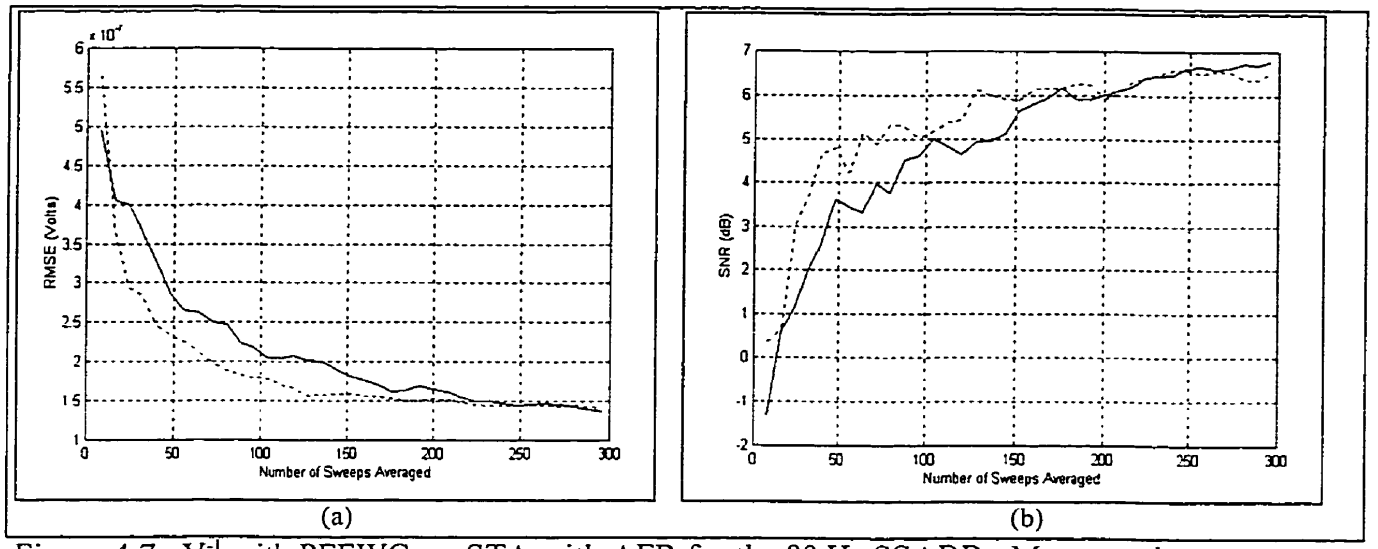

Figure 4.7. V^{-1} with PFFWC vs. STA with AFR for the 80 Hz SSABR. Mean results across subjects for Noisy EEG Data: (a) RMSE (b) SNR (V^{-1} '- -', STA with AFR '—').

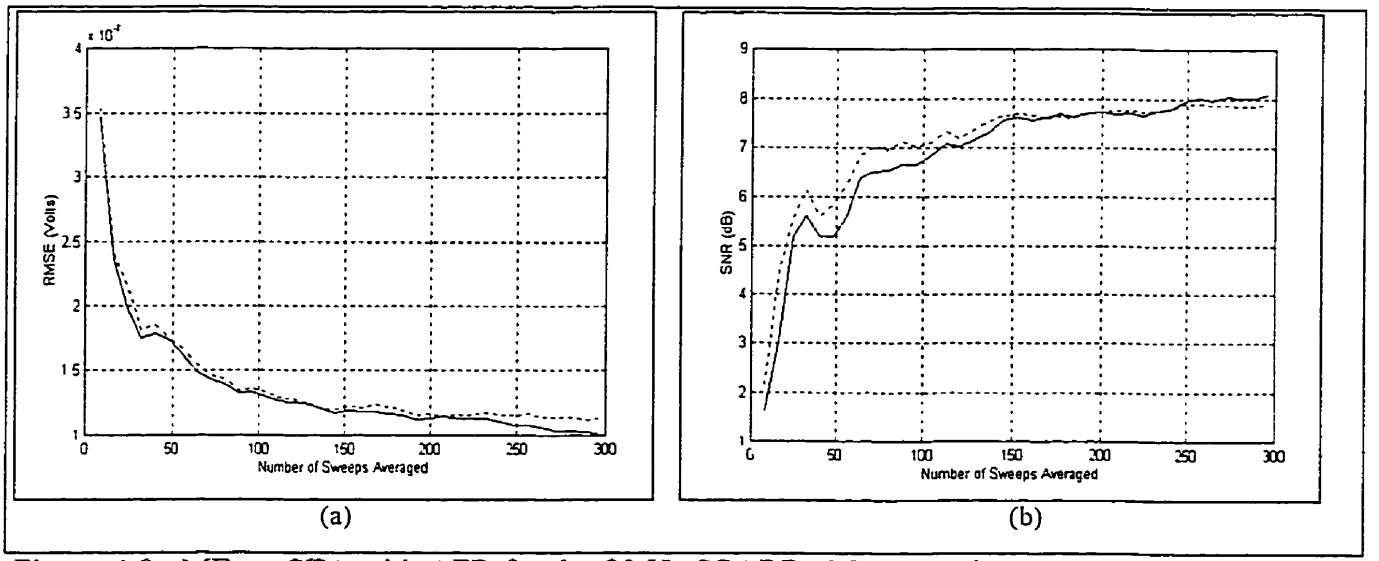

Figure 4.8. ME vs. STA with AFR for the 80 Hz SSABR. Mean results across subjects for Quiet EEG Data: (a) RMSE (b) SNR (ME '- -', STA with AFR '—').

The RMSE and SNR results for the 80 Hz SSABR Quiet EEG Data case are shown in Figure 4.8. The results show that in terms of RMSE, STA with AFR is only slightly better than the ME method, however in terms of SNR, the ME method performs slightly better than STA with AFR. Thus for both 80 Hz SSABR cases, the WTA methods perform better than STA with AFR in both Noisy and Quiet Recording conditions in addition to a time reduction of 36%.

4.4 WTA Results from the Literature

In the literature there have been both positive and negative findings when applying WTA to different types of EPs. The positive findings will be discussed first. Gasser et al. (1983) used a form of WTA (which was constructed to work well when the SNR was good) on transient VEPs. The sweeps were 313.5 ms long and the data was lowpass filtered with a cutoff frequency of 70 Hz and sampled at 408.5 Hz. Testing was done on 42 normal subjects (it is assumed they sat quietly during testing) in which 50 noise sweeps were obtained. There was 4 test cases in which a signal was added to 25, 50, 75 and 100% of the noise sweeps. WTA was found to correctly detect the signals when they occurred and to suppress impulsive noise. Davila et al. (1994) also successfully applied WTA to real ABRs (it is assumed the subjects sat quietly during testing) which resulted in an increase in SNR. The data was analog bandpass filtered from 10 to 3000 Hz, sampled at 10 kHz and digitally high-pass filtered with a cutoff frequency of 600 Hz.

Bataillou et al. (1991) used WTA on the ECG where an adaptive scheme was used to calculate weights based on the V^{-1} method. In their tests, data was sampled at 1 kHz and bandpass filtered from 1 to 300 Hz. Results showed that WTA was better than STA when the noise level varied across sweeps but for relatively constant noise levels WTA and STA gave similar results. A study by Bezerianos et al. (1995) using VNF on VEPs (sampled at 512 Hz and filtered from 1 to 100 Hz) showed that when similarity was detectable, WTA performance was good but when similarity wasn't detectable then WTA was restricted to artifact suppression only.

In work by Gerull et al. (1996), a rectangular signal was added to EEG data recorded under quiet and noisy conditions. This synthetic signal had a larger amplitude than normal EPs to inspect signal underestimation with a manageable number of sweeps. In their work, sweeps

were 10 ms long and were sampled at 10 kHz (later downsampled by a factor of 2). Using quiet data, they saw that weights based on the V^{-1} method underestimated the signal and had a much larger RMSE than STA. Using noisy EEG data, the V^{-1} method still gave a larger RMSE than that seen with STA. In both cases their proposed method (see section 4.2.3) performed better than STA. They stated that significant underestimation was not avoided by increasing the df for weight calculation but was avoided when both the signal was removed *and* there was an increase in the df for weight calculation.

Both positive and negative results were found by Lutkenhoner et al. (1985) for two types of EPs. Lutkenhoner found that WTA worked well for brainstem AEPs, however, whitening was important because the EPs they were interested in detecting had a frequency range of 0.25 to 1 kHz meanwhile the spectrum was dominated by EEG frequencies below 40 Hz. Alternatively for a certain class of cortical AEPs elicited by 500 ms tone bursts, Lutkenhoner found that ensembles of sweeps obtained during routine investigations were relatively stationary in power in the AEP frequency range (3 to 12 Hz). They stated that attempts to get the recordings to become more noisy by deliberate muscle activity had no effects and so WTA was not beneficial in this case.

4.5 Discussion and Conclusions

To conclude, we shall review the results of the 40 Hz SSABR test case followed by the 80 Hz SSABR test case. For the 40 Hz SSABR, results showed that WTA outperforms STA (without AFR) under noisy recording conditions which is not surprising, however, under quiet recording conditions there is no performance improvement. For the Noisy EEG Data case, the ME and V^{-1} method performed the best in terms of RMSE and SNR. Under quiet recording conditions, the V^{-1} method maintained its relative performance level, however the ME method suffered a bit. In both cases, the Gerull method resulted in signal amplitude overestimation. In addition, for both the Noisy and Quiet EEG Data cases, the use of PFFWC (from 35 to 55 Hz) did not improve WTA performance but in fact resulted in degradation with the ME method suffering the most. Furthermore, the V^{-1} method had the least computational overhead. Based on these findings, the best WTA method for 40 Hz SSABRs recorded under noisy and quiet conditions is the V^{-1} method.

For the 80 Hz SSABR test cases, WTA outperformed STA (without AFR) in the Noisy EEG Data case with the ME method performing slightly better than the V^{-1} method. For the Quiet EEG Data case, the WTA methods performed similarly to STA (without AFR), however, the ME method had slightly better performance with a 0.5 dB advantage in SNR in the 20 to 100 sweeps averaged range. When PFFWC (from 65 to 95 Hz) was used, the performance of the ME method suffered the most while the Gerull method was slightly affected. The VNF method performance was improved for the Noisy EEG Data case but unaffected for the Quiet EEG Data case. However, the V^{-1} method improved significantly for the Noisy EEG Data case and was only slightly affected for the Quiet EEG Data case. Consequently, the best WTA method for 80 Hz SSABRs recorded under noisy conditions was the V^{-1} method with PFFWC. For recordings under quiet conditions the ME technique was the best WTA method.

The best WTA methods were then compared to STA with AFR. First of all we notice that when using AFR, on average an extra 32 seconds (36%) of recording time was needed to collect 90 seconds of data for the Noisy EEG Data case. For the 40 Hz SSABR Noisy and Quiet EEG Data cases, the V^{-1} method was compared to STA with AFR and the result was that both methods performed similarly with a slight advantage to STA with AFR for the Noisy EEG Data case and a slight advantage to the V^{-1} method for the Quiet EEG Data case. In the 80 Hz SSABR Noisy EEG Data case, the V^{-1} method with PFFWC (65 to 95 Hz) was compared to STA with AFR. The V^{-1} method with PFFWC performed better in terms of both RMSE and SNR (at some points by a margin of 2 dB). For the 80 Hz SSABR Quiet EEG Data case, the ME method was compared to STA with AFR. In terms of RMSE, both methods performed similarly, however in terms of SNR, the ME method performed better by a margin of 1 to 2 dB early on and then 0.5 dB up to 150 sweeps averaged. Thus, in both the 40 Hz and 80 Hz SSABR cases, the best WTA methods performed just as good (in the 40 Hz case) and better (in the 80 Hz case) than STA with AFR. Furthermore, STA with AFR needed 36% more recording time, however, this extra recording time depends on how frequent the subject is noisy and the duration of each 'noisiness interval' (i.e. compare subjects 2 and 4 from Table 4.7). Thus, the amount of time reduction obtainable with WTA will be difficult to predict in general but nonetheless WTA does have an advantage over STA with AFR.

WTA does however have a disadvantage compared to STA with AFR because there is a need for short duration sweeps in order to make the method sensitive to physiological artifacts

which have short time durations. Thus, the final WTA result is also short in duration. Needless to say, this will have a negative impact in terms of the small number of data samples and the resulting decreased frequency resolution for spectral estimation using the periodogram. Alternatively, for STA with AFR, the buffer size can be chosen small to eliminate the short duration physiological artifacts (using AFR) and STA can then be done across groups of sweeps (concatenated end to end). Thus STA with AFR is somewhat more flexible than WTA.

In terms of nonstationarities in the recorded EEG, the performance of the WTA methods in all Quiet EEG Data cases (for both 40 and 80 Hz SSABRs) indicate that there are no major measurable nonstationarities that would negatively influence STA. Conversely and understandably, for the 40 and 80 Hz SSABR Noisy EEG Data cases, WTA was able to distinguish between the physiological artifacts and the regular ongoing noise and was able to improve on the performance of STA with AFR.

One may also argue that if subjects are awake while being tested then they should be able to control their behavior/state and thus reduce the amount of nonstationarity in their EEG and thus WTA should not be needed. While this maybe true for certain cases, we must also think of the ability of the different subject groups to control themselves (i.e. adults versus infants) and different states in which no one has no control over their behavior (i.e. sleep). In fact, for sleep, Sugimoto et al. (1978) showed that the degree of nonstationarity increases as segment length is increased in every stage of sleep (they found only 50% of 2.5 to 5 second sleep segments analyzed were stationary; see Appendix B). Thus WTA should definitely provide an advantage when testing individuals who are sleeping or who have no control over their behavior.

Another element of the WTA investigation was the use of different frequency ranges for weight calculation. Weights based on the range of 10 to 100 Hz were compared to weights calculated from the 35 to 55 Hz range for the 40 Hz SSABR and the 65 to 95 Hz range for the 80 Hz SSABR. This was done to see if there were nonstationarities outside of the signal frequency region which would adversely affect WTA or whether there were nonstationarities in the signal frequency region that were masked by the behavior of the rest of the spectrum of the recording. The results obtained in both the 40 Hz and 80 Hz SSABR test cases showed that in general PFFWC impaired the performance of the WTA methods. This was most likely due to the filtered range being too small which impaired the estimation of certain quantities in each

WTA method (e.g. variance in the V^{-1} method). Conversely for the VNF method, the smaller bandwidth should be more preferable since this test depended on the similarity/difference of the recorded sweeps, however, the performance of the VNF method did not really improve with PFFWC except for the 80 Hz SSABR test case.

Furthermore, in terms of SNR, PFFWC hindered the performance of the ME method the most, followed by the Gerull and V^{-1} methods. The ME performance with PFFWC suffered more in the 40 Hz SSABR case than the 80 Hz case. This was most likely due to the increased correlation between the sweeps due to the filtering (remember that a smaller bandpass was used for the 40 Hz SSABR). This increased correlation made the sweeps appear more similar which impaired the ability of the ME method to pick out the sweeps with better SNR and weight them accordingly. For the Gerull method the degradation was worse for the Noisy EEG Data cases and the 40 Hz versus the 80 Hz SSABR. This degradation was because the filtering was done after a non-ideal whitening of the spectrum. That is, in the ideal case, the spectrum would be flat after whitening and thus taking the variance estimate from a subsection of the spectrum for each sweep would still give a good idea of the relative variance of each sweep. However, for the 'real-life' situation, the spectrum after whitening is still a bit peaky and thus taking a smaller slice of the spectrum and using it for a variance estimate will lead to a less accurate picture of the relative variance of each sweep. To combat this effect in the 80 Hz SSABR, a larger bandwidth was taken which resulted in a reduced degradation effect when using PFFWC. In addition, the noise is more well behaved in the 80 Hz region.

In general, the performance of the V^{-1} method was also affected just as much as that of the Gerull method for the 40 Hz SSABR test cases when using PFFWC, however, for the 80 Hz SSABR Noisy EEG Data case there was a great improvement in performance. Firstly, the degradation seen in performance is due to a smaller number of frequencies contributing to the variance estimate. In addition, there was most likely a 'peaky' spectrum in the 40 Hz region used for the variance calculation (the idea here is the same as that given for the Gerull method). For the performance improvement in the 80 Hz SSABR Noisy EEG Data case, this was due to the fact that the noise spectrum was most likely more well behaved (i.e. smoother) than the noise in the 40 Hz region and the increased filter bandwidth provided a more accurate variance estimate (i.e. increased the degrees of freedom).

Finally, in all tests, the Gerull method overestimated signal amplitude because it had the highest RMSE results meanwhile maintaining SNR values which were comparable to those of the other WTA methods. The one case where RMSE was not as high was in the Noisy EEG Data case for the 80 Hz SSABR. It is known that the 40 Hz range is much noisier than the 80 Hz range. Thus there is something inherent in the Gerull weight calculation which suffers when the noise is increased. One explanation for this is the order of sweep summation and weight calculation. In the V^{-1} method, the sweeps are first scaled by the inverse of their variance before being summed but in the Gerull method summation between pairs of sweeps occurs before scaling by the inverse variance. However, in Gerull's work this overestimation was not seen for the transient signal used in his study. From our test results we can see that this overestimation decreases with increasing frequency (i.e. 80 Hz SSABR results versus the 40 Hz SSABR results). Thus looking at the frequency components of Gerull's transient signal we find that the majority of signal power occurs for frequencies above and including 100 Hz. This suggests that the overestimation effect diminishes when the noise power in the vicinity of the signal becomes smaller (recall that EEG noise follows a 1/f shape and thus noise in the 100 Hz range is much smaller than noise in the lower 0 to 40 Hz frequency range).

5.0 AR-BASED SSABR DETECTION

This chapter investigates the use of AR parametric modeling for SSABR detection. The strength of spectral estimation based on AR modeling is increased frequency resolution over the periodogram approach. This increased resolution will allow the use of shorter duration sweeps which will in turn allow for a greater amount of sweeps to be synchronously time averaged which should provide a greater amount of noise reduction. The amount of STA, however, will be limited by the amount of noise correlation between successive sweeps and on the limitations of the AR parameter estimators in terms of the minimum number of data samples needed for accurate parameter estimation. The increased resolution will also decrease the interference of the EEG noise without requiring longer recording durations or repeated measurements with short data records. All of these concepts should lead to improved SSABR detection.

This chapter begins with some background on Parametric Spectral Estimation with special emphasis on AR based methods. In particular, the Yule-Walker, Burg and Modified Covariance Methods will be discussed. Further, the gaussian nature of the EEG noise suggests the use of AR parameter estimation based on **H**igher **O**rder **S**tatistics (HOS) and so this topic will also be discussed. The issue of model order selection will then be examined in addition to some problems encountered in AR spectral estimation. The remainder of the chapter deals with the design of an AR-based SSABR detector for 40 Hz SSABRs. In particular, the AR-SSABR detector will be illustrated in block diagram format. Each block will then be discussed in terms of the parameters which characterize it. ROC analysis on real SSABR data will then be used to estimate values for each parameter. This chapter then concludes with a recap of the AR-SSABR detector and the parameters which should be used to yield the highest SSABR detection rate possible based on AR modeling.

5.1 Parametric Spectral Estimation

Parametric **S**pectral **E**stimation (PSE) is a modern spectral estimation technique that is most often chosen over the classical periodogram due to its enhanced resolution properties (Challis et al. 1991). However, the improved resolution depends on the SNR; usually for SNRs greater than 0 dB, PSE has better frequency resolution than the periodogram. In

addition, PSE does not suffer from the distortion produced by sidelobe leakage that occurs in the periodogram approach. Further, PSE yields reasonable spectral estimates for short data records, however, for low enough SNRs, PSE is often no better than periodogram processing (Kay et al., 1981).

The basic premise behind PSE is the choice of an appropriate time series model to represent the measured data for which a spectral estimate is desired. The parameters for the model are then estimated from the measured data and are inputted into the theoretical power spectral density expression for that model. Usually, the model that results in the fewest number of parameters should be chosen since this results in better estimation statistics (Marple, 1987). The advantage of modeling is that one can make a more realistic assumption about the measured data outside of the measurement interval other than to assume that it is zero (Kay et al., 1981).

The types of parametric models are the **A**uto**R**egressive **M**oving **A**verage (ARMA) model, the **M**oving **A**verage (MA) model and the **A**uto**R**egressive (AR) model. Each model assumes that an input white noise process is passed through a filter and the output is the observed data which has been measured. This concept is shown in Figure 5.1.

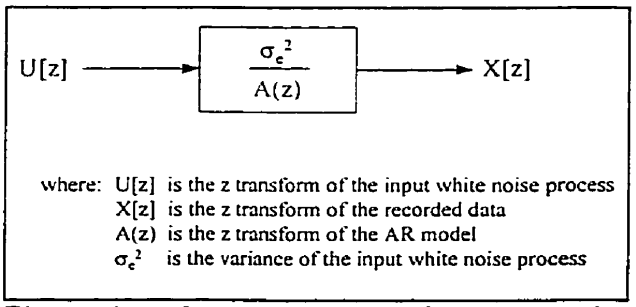

Figure 5.1. Conceptual Basis for Parametric Modeling.

The type of model chosen dictates the form of the filter. The most general model, the ARMA model, has the following linear-difference equation (Marple, 1987):

$$x[n] + \sum_{k=1}^{p} a[k]x[n-k] = \sum_{k=0}^{q} b[k]u[n-k] \qquad (5.1)$$

where x[n] is the measured data
u[n] is the input white noise process
a[k] are the AR parameters
b[k] are the MA parameters

which results in a filter with a transfer function given by:

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + \sum_{k=1}^{q} b[k] z^{-k}}{1 + \sum_{k=1}^{p} a[k] z^{-k}} \tag{5.2}$$

The MA model can be obtained by setting the parameter p in equation 5.1 to zero resulting in the following linear difference equation and filter transfer function:

$$x[n] = \sum_{k=0}^{q} b[k] u[n-k] \tag{5.3}$$

$$H(z) = B(z) = 1 + \sum_{k=1}^{q} b[k] z^{-k} \tag{5.4}$$

Similarly the AR model can be obtained by setting the parameter q in equation 5.1 to zero which results in the following linear difference equation and filter transfer function:

$$x[n] = -\sum_{k=1}^{p} a[k] x[n-k] + u[n] \tag{5.5}$$

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{k=1}^{p} a[k] z^{-k}} \tag{5.6}$$

The parametric spectral estimate is then calculated using the relation between the input and output of a linear system for stochastic signals. For an AR model this is (Marple, 1987):

$$P_{xx}(z) = P_{uu}(z) H(z) H^*(1/z^*) = P_{uu}(z) \cdot \frac{1}{A(z) A^*(1/z^*)} \tag{5.7}$$

where * denotes complex conjugation

Now using the relation $z = \exp(j2\pi fT)$ and scaling by the sampling interval T_s, the AR spectral estimate is (Marple, 1987):

$$P_{xx}(f) = \frac{\sigma_u^2 \cdot T_s}{|A(f)|^2} \tag{5.8}$$

where: $A(f) = 1 + \sum_{k=1}^{p} a[k] \exp(-j2\pi f k T_s)$

σ_u^2 is the variance of the input white noise process

The choice of model type depends on the characteristics of the spectrum of the data which is to be modeled. ARMA models are good for modeling spectra with sharp peaks and

valleys. MA models are good for modeling spectra with valleys but not peaks, meanwhile, AR models are good for modeling spectra with sharp peaks (i.e. sinusoids) but not valleys (Marple, 1987). Fortunately, all 3 models are interrelated by the Wold Decomposition Theorem so that if the wrong model is chosen then a satisfactory model may still be attained by choosing a high enough model order (Kay et al., 1981).

AR models have become the most popular parametric model because AR parameter estimation results in linear equations meanwhile MA and ARMA parameter estimation methods require the solution of nonlinear equations. Thus AR parameter estimation has much less computational overhead compared to the ARMA and MA parameter estimation techniques (Kay et al., 1981). More importantly, since AR based spectral estimates are good for modeling sinusoids, this parametric spectral estimator will be used for SSABR detection.

5.2 AR Parameter Estimation

As previously mentioned, AR models can describe processes with spectral peaks quite well and are thus often used to indicate the presence of sinusoidal components in measured data. However, AR modeling of sinusoidal spectra work well only as long as the SNR is adequate (Orfandis, 1985). Actually the frequency variance (i.e. the 'goodness') of an AR spectral estimate for a sinusoidal process is inversely proportional to both the data length and the square of the SNR (Kay et al., 1981). In addition, an important distinction between AR spectral estimates and the periodogram is that peaks in the AR spectral estimate are proportional to power squared whereas for the periodogram they are linearly related to power.

The high resolution of the AR spectral estimate is a result of the implied extrapolation of the **A**uto**c**orrelation **F**unction (ACF) when computing the spectrum. In fact, the lags for $|k| > p$ (for an AR(p) model) are obtained recursively by:

$$r_{xx}[m] = -\sum_{k=1}^{p} a[k] r_{xx}[m-k] \qquad (5.9)$$

where r_{xx} is the ACF

This can be continued to an infinite extrapolation of the ACF rather than windowing it to zero and thus AR spectra don't exhibit sidelobes due to windowing (Kay et al., 1981). Further, higher order AR models have better resolution potential than lower order models.

Much of the improved resolution is lost when noise is added to the AR time series because with low SNRs the AR model is less justifiable. This is because the addition of white noise results in an ARMA model (Kay, 1980) which contains the same poles as the AR model but there are now zeros whose location depends on the SNR. For low SNRs the zeros will locate themselves near the signal poles which will result in pole-zero cancellation thus yielding a smooth AR spectral estimate because the estimated signal poles will be near the origin[10]. The other problem for sinusoids in noise is that the peak location in the AR spectral estimate depends on the sinusoid's phase however this phase dependence decreases as the data length increases.

To combat the effect of noise on the AR spectral estimate there are 4 general techniques: (1) use an ARMA model (since it is the true model in this case), (2) filter the data to reduce the noise (this moves the estimated AR parameters closer to the true ones and greatly improves the AR estimate (i.e. more accurate frequency estimates at lower SNRs as shown by Kay (1978, 1984) and Paliwal (1986b))), (3) use a large order AR model (since a large order AR model approximates an ARMA model according to the Wold Decomposition Theorem) and (4) compensate the parameter estimates for the effect of noise (Kay et al., 1981).

The procedure for AR parameter estimation consists of defining the error between the modeled data and the actual measured data (Kay, 1987). The modeled data ($\hat{x}[n]$) is shown in Equation 5.10 (this is known as a forward prediction equation). The error ($e[n]$) between the modeled data and the actual data is defined in Equation 5.11. Next the error power (P_e) is obtained (equation 5.12) and is minimized with respect to the AR parameters (a[k]) which leads to the Yule-Walker equations (equation 5.13). Solution of equation 5.13 leads to the Yule-Walker AR parameter estimate.

$$\hat{x}[n] = -\sum_{k=1}^{p} a[k]x[n-k] \tag{5.10}$$

$$e[n] = x[n] - \hat{x}[n] \tag{5.11}$$

$$P_e = E[\,|e[n]|^2\,] \tag{5.12}$$

where E is the statistical expectation operator

[10] For a pure AR process the estimated pole positions would be at the correct pole locations and for white noise the estimated poles would locate themselves at the origin. If both processes are present then the estimated poles will locate themselves somewhere in between the true AR poles and the origin causing a spectral smoothing effect.

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[-1] & \cdots & r_{xx}[-(p-1)] \\ r_{xx}[1] & r_{xx}[0] & \cdots & r_{xx}[-(p-2)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[p-1] & r_{xx}[p-2] & \cdots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} \hat{a}[1] \\ \hat{a}[2] \\ \vdots \\ \hat{a}[p] \end{bmatrix} = -\begin{bmatrix} \hat{r}_{xx}[1] \\ \hat{r}_{xx}[2] \\ \vdots \\ \hat{r}_{xx}[p] \end{bmatrix} \tag{5.13}$$

The way that the error is defined and the parameters with which the error is minimized with respect to leads to different AR parameter estimation methods. In addition, preprocessing of the data (i.e. obtaining the HOS for the data) can lead to a different estimation method. The different ARPEs which will be investigated in this thesis are the Yule-Walker (Autocorrelation) method, the HOS-based Yule-Walker method using 4th order cumulants, the Burg method and the Modified Covariance Method. Each of these methods will now be briefly discussed.

5.2.1 The Yule-Walker (Autocorrelation) Method

The Yule-Walker equations (equation 5.13) detail the linear relationship between the AR parameters and the ACF. The solution to the YW equations can be had from the Levinson-Durbin recursion relation which provides an efficient solution that requires operations on the order of p^2 as opposed to p^3 required for Gaussian Elimination. In fact the Levinson-Durbin recursion relation can be used to generate successively higher order models until the modeling error is reduced to a desired value (Kay, 1981). Furthermore, the biased ACF estimate is used in the Levinson-Durbin algorithm to reduce the risk of ill-conditioning (Marple, 1980). Unfortunately, for short data records, the Yule-Walker approach produces poor resolution spectral estimates. This can be attributed to the data window associated with the YW approach (Kay et al., 1981). In later sections this ARPE will be called YW2.

5.2.2 The HOS-based Yule-Walker Method Using Fourth Order Cumulants

HOS is very useful because it can suppress additive Gaussian noise (white or colored) of unknown power spectrum. This is because for Gaussian data, all cumulants of order greater than two are theoretically zero (Nikias et al., 1993). In practice, this property does not hold for estimated cumulants, however, the estimated cumulant will have an additive noise term that decreases to zero as data size increases (Parat, 1994). Thus cumulants can lead to high SNR domains which will lead to more accurate parameter estimation. However, the biggest drawback to using HOS is larger computational overhead due to the requirement of longer data

lengths (than correlation based methods) which is needed to reduce the variance associated with estimating HOS from real data using sample-averaging techniques. The variance of the estimates is also affected by the SNR of the data (Mendel, 1991).

Cumulants can be calculated in terms of higher order moments. An n^{th} order moment of a random process X is given statistically as (Nikias et al., 1993):

$$m_n^x(\tau_1, \tau_2, \ldots, \tau_{n-1}) = E\{X(k)X(k+\tau_1)\ldots X(k+\tau_{n-1})\} \quad (5.14)$$

The 1^{st} order cumulant equals the 1^{st} order moment which is the mean. The 2^{nd} order cumulant is the covariance sequence. The 3^{rd} order cumulant is given by:

$$c_3^x(\tau_1,\tau_2) = m_3^x(\tau_1, \tau_2) - m_1^x\{m_2^x(\tau_1)+m_2^x(\tau_2)+m_2^x(\tau_1-\tau_2)\}+2(m_1^x)^3 \quad (5.15)$$

For a zero mean sequence, the 2^{nd} and 3^{rd} order cumulants equal the 2^{nd} and 3^{rd} order moments and the 4^{th} order cumulant simplifies to:

$$c_4^x(\tau_1,\tau_2,\tau_3) = m_4^x(\tau_1,\tau_2, \tau_3)-m_2^x(\tau_1)m_2^x(\tau_3-\tau_2) - m_2^x(\tau_2)m_2^x(\tau_3-\tau_1) - m_2^x(\tau_3)m_2^x(\tau_2-\tau_1) \quad (5.16)$$

In practice these quantities must be estimated in the same fashion as correlations and so cumulants are approximated by replacing the expectations by sample averages. Cumulants also have the desirable property that the cumulant of two statistically independent random processes equals the sum of the cumulants of the individual random processes which is not true for higher order moments (Mendel, 1991).

Higher order spectra (or polyspectra) can also be taken of the higher order moments and cumulants. The k^{th} order polyspectrum is the k-1^{th} dimensional discrete-time Fourier Transform of the k^{th} order cumulant (or moment). Moment spectra are useful in the analysis of deterministic signals whereas cumulant spectra are important in the analysis of stochastic signals. However, nonparametric polyspectral methods are subject to high estimation variance and low resolution (Mendel, 1991). Thus one can turn to parametric methods for polyspectral estimation. For random sinusoidal processes one would want to use the 4^{th} order cumulant because if the data contains sinusoids then any 1-D slice of the 4^{th} order cumulant retains all the pertinent sinusoidal information (Mendel, 1991) and is zero for a Gaussian process (i.e. noise).

HOS based AR parameter estimation can be obtained by starting from equation 5.13 written in a slightly different format:

$$\sum_n x[n]x[k+n] = -\sum_{i=1}^{p} a[k]\sum_n x[n]x[n+k-l], \qquad k = 1,2,\ldots,p \qquad (5.17)$$

Multiplying both sides with $x^3[n]$ leads to:

$$\sum_n x^3[n]x[k+n] = -\sum_{i=1}^{p} a[k]\sum_n x^3[n]x[n+k-l], \qquad k = 1,2,\ldots,p \qquad (5.18)$$

The terms $x^3[n]x[n + k]$ and $x^3[n]x[n + k - l]$ are recognized as one dimensional slices[11] of the 4th order cumulant and thus equation 5.18 is the Yule-Walker method using 4th order cumulants.

The Yule-walker method based on the ACF was prone to poor resolution and so the HOS-based Yule-walker method is most likely prone to poor resolution as well. However, the main advantage here is that gaussian noise will be greatly attenuated. In addition, as longer data lengths are used better parameter estimates should be obtained. This HOS-based method can also be augmented by pre-filtering the data which has been shown to have a detection advantage as long as the signal frequency region is known (Ioup et al., 1993). In later sections this ARPE will be called YW4.

5.2.3 The Burg Method

The Burg method is a constrained least squares estimation procedure using the sum of the forward and backward linear prediction error energies and minimizing them subject to the constraint that the AR parameters satisfy the Levinson recursion for all orders (Marple, 1980). The Burg method was very popular in the past due to its computational advantages, however, the Burg method introduces biases. For instance, the Burg method suffers from large frequency estimation errors when it is applied to short sinusoidal signals and spectral line splitting[12] for high AR model orders, high SNR, or for an odd number of quarter cycles of sinusoidal data with an initial phase of 45°.

The aforementioned problems may be solved by increasing the amount of data while using the original sampling rate (increasing the sampling rate to increase the data while keeping the observation interval unchanged is of no benefit) (Herring, 1980). Another approach is to apply weights to the estimation error. Two popular methods are the Hamming and optimum

[11] A one dimensional slice of an nth order cumulant is obtained by freezing n-2 of its n-1 indexes.
[12] Spectral line splitting means a single spectral line is seen as two or more spectral lines. Spurious peaks often accompany spectral line splitting. Spectral line splitting disappears when record length is increased.

parabolic weights which are applied to the forward and backward prediction error energies prior to minimization (Paliwal, 1985a). The Hamming weights are given by (Swingler, 1979):

$$W_m(k) = 0.54 - 0.46\cos\left(\frac{2\pi k}{N-m} + \frac{\pi}{N-m}\right) \qquad \text{for } k = 0,1,\ldots,N\text{-}1\text{-}m \qquad (5.19)$$

where m is the model order

The optimum parabolic weight function was achieved by minimizing the average variance of the frequency estimation error for a truncated real sinusoid measured over all phases in the range of 0 to π. The optimum parabolic weights are given by (Kaveh et al., 1983):

$$W_m(k) = \frac{6(k+1)(N-m-k+1)}{(N-m+1)(N-m+2)(N-m+3)} \text{ for } k = 0, \ldots, N\text{-}m \qquad (5.20)$$

Paliwal (1985a, 1985b, 1986a) studied the Hamming and optimum parabolic weights and found that the optimum parabolic weighted Burg method results in the best performance in terms of least frequency estimation bias for SNRs in the 25 to 30 dB range, however, the weighting methods result in performance degradation for actual AR signals. The deterioration however is least severe when using the optimum parabolic weighted Burg method. Thus, the optimum parabolic weighted Burg method will be used in this thesis and will be called Burg in later sections.

5.2.4 The Modified Covariance Method

The **M**odified **C**ovariance **M**ethod (MCM), also called the forward-backward Least Squares parameter estimator, is an unconstrained parameter estimator (the Levinson recursion constraint is removed). The foundation of this method stems from the fact that the errors from the forward and backward linear prediction estimates have the same statistical information and by combining them more error points are generated which results in improved AR parameter estimates. Thus, this method minimizes the energy of the sum of the forward and backward errors with respect to the AR parameters. This results in a performance improvement over the Burg method. In particular, researchers note that the MCM method has less frequency bias, reduced variance for frequency estimates and no spectral line splitting.

With the MCM method, if the model order is a small fraction of the number of data samples then the bias of the frequency estimates can be high. However, if the order is increased to half the number of data samples then a performance increase is obtained beyond which a

number of spurious spectral peaks occur. This has been attributed to the ill conditioning of the linear prediction equations when highly correlated signals like sinusoids are present in the data (SVD can be used to alleviate this problem (Kumaresan, 1982a)). In later sections this ARPE will be called MCM.

5.3 Relations between different ARPEs

The Burg and MCM methods work directly on the measured data to obtain better AR spectra than the YW2 method (Kay et al., 1981). In fact the MCM method has been noted as the best ARPE for estimating the spectra of data containing sinusoidal components when the SNR is fairly high. However, for long data records the YW2 method yields reasonable spectral estimates whereas for short data records, the YW2 method produces poor-resolution spectral estimates relative to the other ARPEs (Marple, 1987). In addition, the amount of phase dependence varies for different ARPEs with the MCM method being affected the least (Kay et al., 1981).

5.4 Model Order Selection

The best model order for the AR model is not known a priori and thus several model orders are usually tried and the best one chosen according to some error criterion. Choosing the correct model order is very important because if the process is truly an AR process of order p then using an order less than p will result in a smoothed spectral estimate meanwhile using too high a model order may result in spurious peaks. However, higher model orders provide increased resolution (Kay, 1988) and are needed for the detection of weak signals in noise (Swingler et al., 1988). The problem of model order selection becomes even more difficult in the presence of noise because an AR process in white noise has a theoretically infinite model order and so any choice of model order will introduce bias errors.

Most **M**odel **O**rder **E**stimators (MOEs) have a term which decreases with model order and a term which increases with model order (to penalize for increased parameter estimation). The MOE algorithm will thus produce a plot with a minimum which should occur at order p for an AR(p) process[13]. For data from a pure AR process, MOEs usually produce acceptable

[13] An AR(p) model denotes an AR model of order p.

results if the data record length is not extremely short but it has also been observed that for noise corrupted data, the model order chosen is usually not sufficient to resolve spectral details (Kay, 1988). There are many MOEs available but few guidelines on their use, however a good rule of thumb is that the model order shouldn't be lower than twice the number of spectral peaks and should be within the range of N/3 to N/2 for good spectral resolution with few spurious spectral peaks (Kay, 1988). The MOEs investigated in this thesis will now be discussed.

5.4.1 The Akaike Information Criterion

The **A**kaike **I**nformation **C**riterion (AIC) is a very popular MOE. When the data is assumed to have a Gaussian distribution the AIC is given by:

$$\mathrm{AIC(k)} = N \cdot \ln \hat{p}_k + 2k \; ; \tag{5.21}$$

where p_k is the estimation error
N is the number of data points

The AIC is based on the maximum likelihood estimate of the **p**robability **d**ensity **f**unction (pdf) of the signal and conceptually tries to measure the distance between an assumed pdf and the true pdf. This criterion has been developed under asymptotic conditions (i.e. many data samples). The AIC is popular although it has been noted as being an inconsistent estimator with a 15% probability of model order overestimation (Kay, 1988). Conversely, the AIC MOE may even underestimate the model order if the SNR is high (Marple, 1987). Further, Burshstein et al. (1985) found the AIC chooses low model orders for small data segments and high sampling frequencies and large model orders for large data segments. In addition, Schlindwein et al. (1990) found that for true AR processes when more samples were used (i.e. 64 samples were increased to 128 and 256) the AIC MOE tended to choose the correct model order.

5.4.2 The Modified Akaike Information Criterion

Other MOEs have been proposed to improve the AIC in order to fix the overestimation problem. One reformulation is the modified Akaike Information Criterion (BIC) which is calculated by (Priestly, 1981):

$$\mathrm{BIC(k)} = N \cdot \ln \hat{p}_k + k \cdot \log_{10} N + k \tag{5.22}$$

The basic difference between the BIC and AIC methods is the penalty (2^{nd}) term. Generally the minimization of the BIC function leads to lower model orders than the minimization of the AIC function. The BIC MOE is also a consistent estimator in that as the number of data samples is increased the BIC MOE tends to yield the correct model order (Holm, 1988).

5.4.3 The Pseudo-SNR Based Method

The Pseudo-SNR based MOE is not found in the literature and it was proposed by the author of this thesis (for an explanation of why it is named Pseudo-SNR please refer to section 5.7.2). The basic idea is that the model order yielding the model with the highest Pseudo-SNR at the SSABR frequency is chosen as the correct model order. The calculation of the Pseudo-SNR is outlined in section 5.7.2. In later sections, this method will be referred to as SNRp.

5.5 Some problems with AR Parameter Estimation

The idea of increasing the model order to increase accuracy in parameter estimation is only true for noiseless and moderately noisy data. For higher model orders in noisier situations, there can be considerable perturbation for parameter estimation. This will cause extraneous poles to fall close to the unit circle which results in spurious spikes in the spectral estimate. Silverstein (1991) stated that this numerical instability problem occurs because of the large spread in the eigenvalues of the autocorrelation matrix in equation 5.13. Some eigenvalues (due to noise) are small enough to cause arithmetic instabilities in the matrix inversion process. Thus AR parameter fluctuation is due to the large relative fluctuations of the smaller noise eigenvalues in the autocorrelation matrix.

5.6 Detection based on AR Spectral Estimation versus the Periodogram

In terms of detection using AR modeling versus the periodogram, there have been two conflicting reports in the literature. Dyson et al. (1981) compared the periodogram (with Hanning windowing and overlapped averaging) versus the Burg technique and found that the Burg method showed promise in reducing the observation time needed for sinusoidal detection for a useful SNR range for long observation times. The Burg technique used the Akaike Final Prediction Error method (very similar to the AIC MOE) for selecting the model order. The AR method consisted of segmenting the data into K non-overlapping unweighted records, obtaining

spectral estimates for each record and then averaging the spectral results. The simulations consisted of multiple tones with SNRs of -12 to 15 dB relative to a 1 Hz bandwidth. They found that the Burg method outperformed the periodogram technique down to an SNR of -9 dB. However, Hung et al. (1981) looked at the detection of complex sinusoids in additive noise using the Burg and periodogram methods and found the periodogram method consistently gave higher detection rates and more accurate frequency estimates than the Burg method.

5.7 AR-Based SSABR Detector Design

The block structure of the AR based SSABR detector is shown in Figure 5.2. The detector consists of a preprocessing block to improve the SSABR SNR and to robustify the detector against any recorded artifacts. The next step is to model the data with an AR model which is then used to obtain a spectral estimate. A detection statistic then operates on the spectral estimate to determine if there is a response. If an SSABR is detected then data collection is terminated. If an SSABR is not detected then more data can be collected unless a maximum recording time limit has been met in which case no SSABR is detected. Each block of the detector will now be discussed.

5.7.1 Data Preprocessing

To improve the accuracy of the ARPEs, the SSABR SNR must be increased. Thus, the ongoing EEG background noise must be reduced. In addition, the preprocessed data must be free of any artifacts which may be recorded. The analysis in Chapter 4 indicated that WTA can perform these two operations under both quiet and noisy recording conditions. Furthermore, the inverse variance technique maintained the performance of STA with AFR under both quiet and noisy recording conditions meanwhile requiring a significantly smaller amount of recording time under noisy recording conditions. Thus, the inverse variance technique will be used to combat physiological noise and improve SNR. Further, a small sweep time should be used to increase the number of sweeps averaged which will result in greater noise reduction (this will allow the V^{-1} method to be sensitive to short duration artifacts). A good starting point for the sweep time is 0.3 seconds (used in Chapter 4 and recommended in Appendix F). No spectral averaging is incorporated into the detector due to the observation made in section 3.5 that spectral averaging places a lower limit on the smallest sinusoid that is detectable. In

addition, for weight calculation, a notch filter must be used to remove the line noise (which was quite large in our recordings; 1 to 4 μV). As in Chapter 4, a first order Butterworth IIR notch filter, in the non-phase shift configuration (Antoniou, 1993), with a bandstop region of 58 to 62 Hz and a 40 dB notch was used. In addition, the first 0.4 seconds of recorded data are discarded due to filter startup transients and response latency.

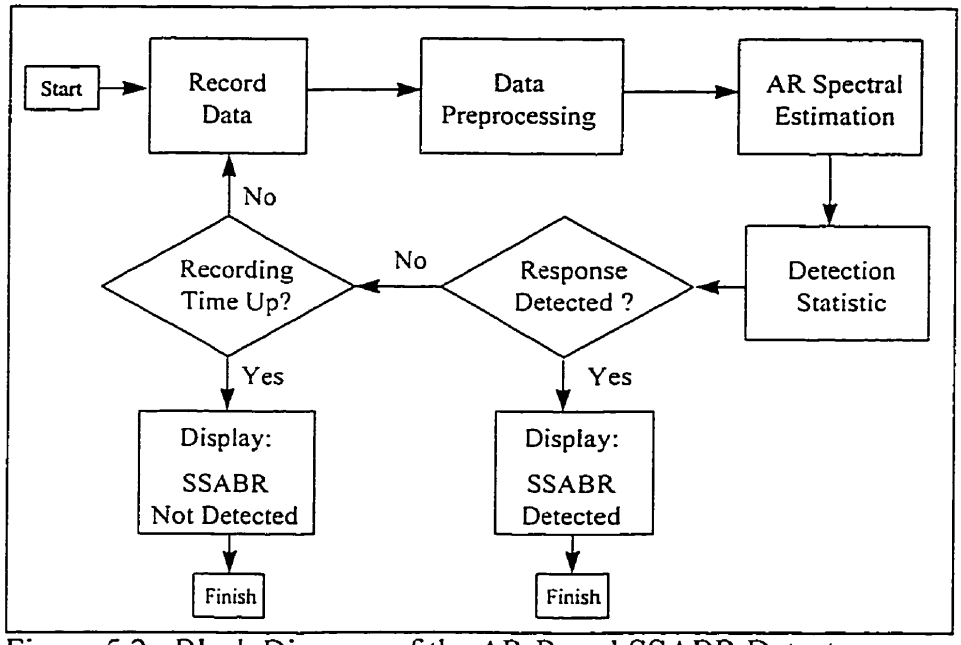

Figure 5.2. Block Diagram of the AR-Based SSABR Detector.

For further SNR enhancement, as noted in Section 5.2, bandpass filtering can be applied. This technique has already been used by Kay (1978, 1979, 1984), Paliwal (1986b) and Tang et al. (1993) to improve ARPE accuracy. In bandpass filtering, a small bandpass is desired for noise reduction, however it can not be too small because when noise is narrow bandpass filtered it will resemble a sinusoid whose frequency is at the center frequency of the passband and whose phase and amplitude vary slowly (Haykin, 1994). The bandpass filter used was an FIR filter with a filter order of 100. The FIR window design method was used with a Kaiser window ($\beta = 4$). The filtering method used was equivalent to taking the convolution of the filter impulse response and the recorded data and truncating a few data points at the beginning and end of the convolution result. This was beneficial for reducing the startup and ending transients. Please see Appendix G for an explanation on the parameters chosen and the bandpass filtering method. A reasonable starting point chosen for the passband of the bandpass

filter was 35 to 45 Hz for the 40 Hz SSABR. For other SSABRs, the bandpass filter bandwidth can be changed to coincide with the signal frequency region.

5.7.2 Detection Statistic

The detection statistic was based on a pseudo-SNR type measurement from the AR spectral estimate. The term pseudo is used because peaks in the AR spectral estimate are related to power squared, whereas in the periodogram the peaks are linearly related to power (Kay et al., 1981). Further pseudo-signal and pseudo-noise powers are obtained by averaging these peak values so it is not strictly proper to say that the averaged values are signal and noise powers, nonetheless, these values will indicate the relative signal and noise power.

The pseudo-signal power can be estimated from the bins in the signal region and the pseudo-noise power can be estimated using a technique commonly employed in radar processing called the split window technique (see Figure 5.3). The split window is comprised of two windows which are placed on either side of the signal region. The bins within the split window provide the pseudo noise power values which are averaged to obtain the pseudo-noise power estimate (Rohling, 1983). The ratio of the pseudo-signal power to the pseudo-noise power is then used to give an indication of whether or not a signal is in the SSABR recording. This is represented by the following equations:

$$SNR_p = 10\log\left(\frac{S_p}{N_p}\right) \tag{5.23}$$

$$N_p = \frac{1}{N_n}\sum_{i=1}^{N_n} N_i \tag{5.24}$$

$$S_p = \frac{1}{N_s}\sum_{i=1}^{N_s} S_i \tag{5.25}$$

where: SNR_p is the pseudo-SNR
N_p is the pseudo-Noise power
S_p is the pseudo-Signal power
N_n is the number of bins in the noise band
N_s is the number of bins in the signal band

In the remainder of chapter 5, a noise band of B Hz denotes a split window of size Q Hz such that Q + Signal Band (in Hz) = B Hz and the noise band is centered on the signal frequency. In addition, each bin in the spectral estimate is 0.25 Hz wide.

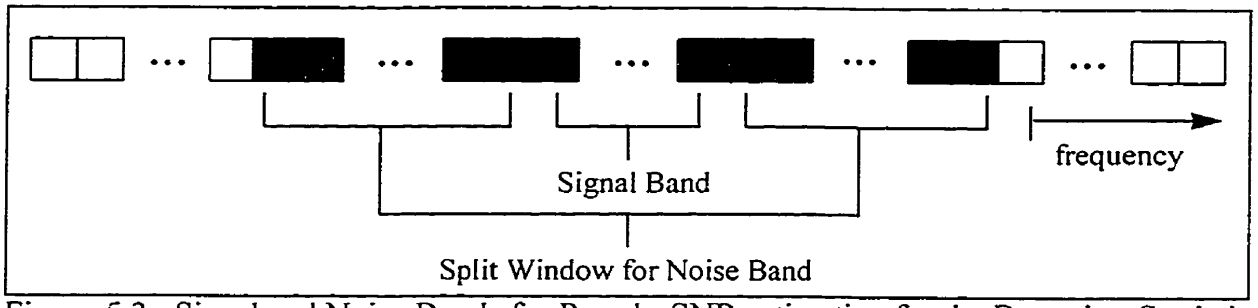

Figure 5.3. Signal and Noise Bands for Pseudo-SNR estimation for the Detection Statistic.

5.7.3 Detector Parameters: Questions

This section will raise a number of questions regarding the criteria used in the AR-based SSABR detector. First off which AR parameter estimator should be used? What should the sweep time be for weighted time averaging and is this compatible for AR parameter estimation (i.e. adequate amount of samples) meanwhile keeping in mind the WTA results from Chapter 4. Which MOE will result in the best detection performance? How large should the maximum model order estimated be? Can a constant model order be used instead of estimating a number of different AR models and then using an MOE? How large should the bandpass for the bandpass filter be? Finally, how large should the signal and noise bands be for the detection statistic?

5.7.4 Detector Parameters: Answers

In this section choices for each criteria mentioned in the previous section are made based on experimental data. To make these choices, real SSABR data was analyzed using ROC analysis (see Appendix E) to determine which parameter values resulted in the best detection performance. Statistical performance was quantified by the area under the ROC curve (RA) and the detection rate (also called Hit Rate(HR)) at a false alarm rate of 5% along with its 95% confidence interval. To obtain the experimental data, 3 male adult subjects aged 24 to 33 were tested. For each subject, two tests were performed to obtain one recording which contained signal plus noise (i.e. true positive data) and another recording which contained only noise (i.e. true negative data). Test #1, which provided true positive data, used a 70 dB SPL stimulus consisting of an AM waveform with f_c = 1kHz, f_m = 40 Hz and μ = 0.95. Test #2, which provided true negative data, used a 70 dB SPL stimulus consisting of a pure tone at 1 kHz which would not evoke an SSABR yet provided recording conditions similar to that of test #1.

For both recordings, the subject sat still in a sound proof booth and the stimulus was presented to the right ear. The duration of the recordings was 219.2 seconds. Please see Appendix C for a more thorough description of testing methods and instrumentation.

The following were the nominal parameters used in generating the ROC statistics (i.e. these parameters were used unless otherwise specified). The recorded data was partitioned into 4.8 second data segments (which resulted in 40 data trials per recording) and detection was performed on each segment. This resulted in 120 true positive and true negative data trials (40 trials per subject x 3 subjects). The time length of 4.8 seconds was chosen to keep in line with the goal of SSABR detection in a short amount of time. WTA was used with the sweep length set to 0.3 seconds which was small enough to provide many sweeps for averaging but was large enough to prevent noise in successive sweeps from being correlated (see Appendix F). Bandpass filtering (see Appendix G) was used with a passband from 35 to 45 Hz and an attenuation of 60 dB in the stopbands. For AR spectral estimation, the **M**aximum **M**odel **O**rder (MMO) was 20. For the SNRp MOE the number of signal bins was 3 and the noise bandwidth was 10 Hz centered at 40 Hz. For the Burg ARPE, optimum parabolic weighting was used.

The ROC results shown in each table are in the following format. For each table entry, the first line is the **ROC A**rea (RA) followed by the **S**tandard **E**rror (SE) in RA estimation in brackets. The second line is the detection rate at a false alarm rate of 5% (HR(0.05)) followed in brackets by the lower and upper bounds for the 95% confidence interval. All table entries are percentages from 0 to 100%.

The first criteria looked at was which MOE to use for each ARPE method using different signal bandwidths for the detection statistic. The ROC results for each ARPE are shown in Tables 5.1 to 5.4. From Table 5.1, we see that for the YW2 ARPE, in absolute terms, the AIC method had the highest RA although considering the SE there is an overlap in the performance of all MOEs. In terms of HR(0.05), the AIC and BIC MOEs are much better than the SNRp MOE. The AIC method also is slightly better than the BIC method. Thus the AIC MOE is chosen for the YW2 ARPE. For the YW4 ARPE, Table 5.2 indicates that the ROC results between all MOEs are very similar. The MOE is thus chosen arbitrarily to be the AIC method.

Table 5.1. ROC results for different MOEs and signal bin number for the YW2 ARPE.

MOE	Number of Signal Bins				
	1	3	7	11	21
AIC	71.82 (3.298) 22.5 (15.6, 31.2)	72.51 (3.258) 23.3 (16.3, 32.1)	73.99 (3.191) 24.2 (17, 33)	75.45 (3.12) 27.5 (20, 36.5)	76.82 (3.049) 29.2 (21.4, 38.3)
BIC	68.09 (3.43) 20 (13.5, 28.5)	69.35 (3.385) 20 (13.5, 28.5)	71.62 (3.296) 24.2 (17, 33)	73.69 (3.205) 23.3 (16.3, 32.1)	73.41 (3.217) 24.2 (17.3, 32.7)
SNRp	69.06 (3.396) 9.2 (5.1, 15.8)	70.09 (3.358) 11.7 (6.8, 14.1)	70.3 (3.35) 13.3 (8.1, 21)	70.27 (3.351) 15 (9.4, 22.9)	70.52 (3.341) 14.2 (8.7, 22)

Table 5.2. ROC results for different MOEs and signal bin number for the YW4 ARPE.

MOE	Number of Signal Bins				
	1	3	7	11	21
AIC	66.57 (3.48) 9.2 (4.9, 16.2)	66.88 (3.47) 9.2 (4.9, 16.2)	67.82 (3.44) 6.7 (3.2, 13.1)	69.25 (3.39) 8.3 (4.3, 15.2)	69.27 (3.388) 9.2 (4.9, 16.2)
BIC	67.8 (3.44) 8.3 (4.3, 15.1)	67.93 (3.44) 8.3 (4.3, 15.1)	67.78 (3.44) 12.5 (7.4, 20.1)	67.77 (3.44) 5.8 (2.6, 12.1)	68.45 (3.418) 4.2 (1.6, 9.9)
SNRp	69.49 (3.38) 10.8 (6.1, 18.1)	67.44 (3.452) 5.8 (2.6, 12.1)	67.35 (3.455) 6.7 (3.1, 13.1)	67.31 (3.457) 6.7 (3.1, 13.1)	66.94 (3.469) 6.6 (2.6, 12.1)

Table 5.3. ROC results for different MOEs and signal bin number for the Burg ARPE.

MOE	Number of Signal Bins				
	1	3	7	11	21
AIC	76.31 (3.074) 28.3 (20.7, 37.4)	78.17 (2.972) 33.3 (25.2, 42.6)	81.99 (2.732) 45.3 (34.4, 52.7)	81.87 (2.74) 38.3 (29.8, 47.7)	77.26 (3.023) 28.3 (20.7, 37.4)
BIC	72.02 (3.28) 27.5 (20, 36.5)	74.84 (3.15) 24.2 (17, 33)	80.03 (2.86) 22.5 (15.6, 31.2)	80.48 (2.832) 24.2 (17, 33)	76.72 (3.053) 25.8 (18.5, 34.8)
SNRp	57.08 (3.689) 9.2 (4.9, 16.2)	71.49 (3.302) 23.3 (16.3, 32.7)	71.9 (3.284) 19.2 (12.8, 27.6)	71.41 (3.31) 19.2 (12.8, 27.6)	71.61 (3.296) 20 (13.5, 28.5)

Table 5.4. ROC results for different MOEs and signal bin number for the MCM ARPE.

MOE	Number of Signal Bins				
	1	3	7	11	21
AIC	77.47 (3.012) 21.7 (14.9, 30.3)	78.44 (2.956) 21.7 (14.9, 30.3)	81.12 (2.79) 38.3 (29.8, 47.7)	81.53 (2.763) 38.3 (25.2, 42.6)	78.38 (2.96) 25.8 (18.5, 34.8)
BIC	71.58 (3.298) 15.8 (10.1, 23.9)	74.4 (3.171) 20 (13.5, 28.5)	79.39 (2.9) 35 (26.7, 44.3)	79.84 (2.872) 27.5 (20, 36.5)	78.55 (2.95) 23.3 (16.3, 32.1)
SNRp	57.91 (3.677) 10.8 (6.14, 18.1)	71.35 (3.307) 16.7 (10.7, 24.8)	69.91 (3.364) 19.2 (12.8, 27.6)	70.0 (3.36) 20.8 (14.2, 29.4)	69.64 (3.375) 20.8 (14.2, 29.4)

For the Burg ARPE, the AIC and BIC methods outperform the SNRp MOE. In addition, the AIC MOE has a slightly larger RA but considerably larger HR(0.05) than the BIC method. Thus the AIC method is chosen for the Burg ARPE. For the MCM method, the AIC and BIC MOEs outperform the SNRp method. The difference between the AIC and BIC methods is small with a slight advantage to the AIC method. The AIC method is thus chosen

arbitrarily over the BIC method. Furthermore, for all ARPEs, the SNRp MOE was repeated with 7 and 11 signal bins, however, ROC performance was still inferior to that seen for the AIC and BIC methods.

Table 5.5 shows the ROC results for each ARPE across different number of signal bins when the AIC is the MOE. From the table we can see that the MCM and Burg ARPEs had the best performance in terms of both RA and HR(0.05%) and the YW4 ARPE had the worst performance. Furthermore we see that the best number of signal bins to use could be either 7 or 11 for the Burg and MCM methods. In the future only the results for these number of signal bins will be shown. In addition, the increased performance for these larger signal bands (7 bins is 1.75 Hz wide and 11 bins is 2.75 Hz wide) indicates that there is some frequency estimation error on the part of the ARPEs due to observation noise and a small number of data samples.

The next parameter looked at was the **M**aximum **M**odel **O**rder (MMO) used by the ARPEs. The MMOs examined were 10, 20, 30, 40, 60, 80 and 100. The results for the Burg and MCM ARPEs are shown in Table 5.6 and 5.7 respectively. For the Burg ARPE there is no increase in ROC performance for MMOs of 40 or more. The ROC results for MMOs of 10 to 30 are similar however the MMO of 20 has a slight advantage in RA and a larger advantage for HR(0.05%). The MMO of 20 is thus chosen for the Burg ARPE.

Table 5.5. ROC results for the AIC MOE for all ARPEs and different number of signal bins.

ARPE	Number of Signal Bins				
	1	3	7	11	21
YW2	71.82 (3.298) 22.5 (15.6, 31.2)	72.51 (3.258) 23.3 (16.3, 32.1)	73.99 (3.191) 24.2 (17, 33)	75.43 (3.12) 27.5 (20, 36.5)	76.82 (3.047) 29.2 (21.4, 38.3)
YW4	66.57 (3.48) 9.2 (4.9, 16.2)	66.88 (3.47) 9.2 (4.9, 16.2)	67.82 (3.44) 6.7 (3.1, 13.1)	69.25 (3.39) 8.3 (4.3, 15.2)	69.27 (3.388) 9.2 (4.9, 16.2)
Burg	76.31 (3.074) 28.3(20.7, 37.4)	78.17 (2.972) 33.3 (25.2, 42.6)	81.99 (2.732) 45.3 (34.4, 52.7)	81.87 (2.74) 38.3 (29.8, 47.7)	77.26 (3.023) 28.3 (20.7, 37.4)
MCM	77.47 (3.012) 21.7 (14.9, 30.3)	78.44 (2.956) 21.7 (14.9, 30.3)	81.12 (2.79) 38.3 (29.8, 47.7)	81.58 (2.763) 33.3 (25.2, 42.6)	78.38 (2.96) 25.8 (18.5, 34.8)

For the MCM ARPE, similar performance is seen for the MMOs of 10 to 40 using either 7 or 11 bins for the signal band in the detection statistic. However, Table 5.7 indicates a slight advantage in RA for an MMO of 20 but a slight advantage in HR(0.05) for an MMO of 30. These differences are very small and considering computational overhead the choice of MMO is between 10 and 20. The MMO of 20 was chosen arbitrarily. In addition, for both the

Burg and MCM ARPEs we can see that performance decreases as MMO increases past 40. This is understandable because with an increase in model order there are more spurious peaks in the AR spectral estimate which will degrade detection performance.

Table 5.6. ROC results for the Burg ARPE for different MMOs and number of signal bins.

MMO	Number of Signal Bins	
	7	11
10	77.93 (2.985) 30.8 (22.9, 40)	80.06 (2.859) 20.8 (14.2, 29.4)
20	81.99 (2.722) 45.3 (34.4, 52.7)	81.87 (2.7[illegible]) 38.3 (29.8, 47.7)
30	79.02 (2.922) 30 (22.2, 39)	81.79 (2.745) 35 (26.7, 44.3)
40	73.7 (3.204) 20.8 (14.2, 29.4)	77.71 (2.998) 27.5 (20, 36.5)
60	74.15 (3.183) 20 (13.5, 28.5)	77.98 (2.983) 30 (22.2, 39.1)
80	74.18 (3.181) 20 (13.5, 28.5)	78 (2.981) 30 (22.2, 39.1)
100	74.08 (3.186) 20 (13.5, 28.5)	77.92 (2.986) 30 (22.2, 39.1)

Table 5.7. ROC results for the MCM ARPE for different MMOs and number of signal bins.

MMO	Number of Signal Bins	
	7	11
10	81.12 (2.79) 36.7 (28.2, 46)	80.72 (2.816) 31.7 (23.7, 40.9)
20	[illegible] [illegible]	81.53 (2.765) [illegible]
30	80.93 (2.803) 40.8 (32.1, 50.2)	81.35 (2.775) 35.8 (27.5, 45.1)
40	80.69 (2.818) 37.5 (29, 46.8)	81.02 (2.797) 35 (26.7, 44.3)
60	71.06 (3.319) 10 (5.5, 17.1)	70.9 (3.326) 6.7 (3.1, 13.1)
80	68.82 (3.405) 6.7 (3.2, 13.1)	69.27 (3.389) 6.7 (3.15, 13.1)
100	61.96 (3.603) 7.5 (3.7, 14.1)	60.99 (3.624) 5 (2.1, 11)

The next criterion which was investigated was the passband for the bandpass filter. The different bandwidths which were used were 35 to 45 Hz, 37 to 43 Hz, 38 to 42 Hz and 39 to 41 Hz. The noise bandwidths for the detection statistic for these different bandpass filter bandwidths were 10 Hz, 6 Hz, 4 Hz and 4 Hz. These noise bandwidths were chosen because they coincide with the bandpass filter passbands. The results are shown in Tables 5.8 and 5.9 for the Burg and MCM ARPEs respectively. Note that the AIC MOE was used and the MMO was 20. Results show considerable overlap in performance for the Burg ARPE across all bandpass filter bandwidths. However, the results do show a slight advantage in using the 37 to 43 Hz bandpass filter for the Burg ARPE. For the MCM ARPE, in terms of RA, the 37 to 43 Hz bandpass filter is preferable. In terms of HR(0.05) for 7 signal bins, the 35 to 45 Hz and 37 to 43 Hz bandpass filters have very similar performance and for 11 signal bins the 37 to 43 Hz and 38 to 42 Hz bandpass filters are very close in performance. Arbitrarily, the 37 to 43 Hz bandpass filter is chosen for the MCM ARPE. Thus for future analysis, the bandpass filter will have a passband from 37 to 43 Hz for all ARPEs.

Table 5.8. ROC results for different BPF bandwidths and number of signal bins for the Burg ARPE.

BPF (Hz)	Number of Signal Bins	
	7	11
35 to 45	81.99 (2.732) 43.3 (34.4, 52.7)	81.87 (2.74) 38.3 (29.8, 47.7)
37 to 43	82.84 (2.67) 48.3 (39.2, 57.6)	84.32 (2.564) 49.2 (40, 58.4)
38 to 42	80.45 (2.834) 45.8 (36.8, 55.1)	83.16 (2.65) 42.5 (33.7, 51.8)
39 to 41	80.51 (2.83) 40.8 (32.1, 50.2)	83.52 (2.623) 43.3 (34.4, 52.7)

Table 5.9. ROC results for different BPF bandwidths and number of signal bins for the MCM ARPE.

BPF (Hz)	Number of Signal Bins	
	7	11
35 to 45	81.12 (2.79) 38.3 (29.8, 47.7)	81.53 (2.763) 33.3 (25.2, 42.6)
37 to 43	83.78 (2.61) 42.5 (33.7, 51.8)	84.66 (2.532) 35 (26.7, 44.3)
38 to 42	76.75 (3.051) 29.2 (21.4, 38.3)	82.06 (2.727) 38.3 (29.8, 47.7)
39 to 41	74.58 (3.162) 20.8 (14.2, 29.4)	80.01 (2.861) 19.2 (12.8, 27.6)

The sweep time length was then varied to observe the effect on ROC performance. This was done to see the effect of the increased number of data samples on the Burg and MCM ARPEs as well as the AIC MOE. The performance of both the ARPEs and the AIC MOE should get better with an increase in the number of data samples. However, this increase in data samples comes at the expense of a decrease in SNR because with longer duration sweeps, there is a reduced number of sweeps being averaged (recall data segments of fixed time lengths of 4.8 seconds are being analyzed). The bandpass filter had a passband from 37 to 43 Hz, the MMO was 20 and the MOE was the AIC method. The results are shown in Table 5.10. The results for the YW2 ARPE are also shown because the performance of the YW2 method approaches that of the Burg and MCM ARPEs with an increase in the sweep time length. This effect has also been observed in the literature. For the YW2 method the best RA results were for sweep lengths of 0.4 and 0.6 seconds with 7, 11 or 21 signal bins for the detection statistic. In terms of HR(0.05), the performance was better for 0.6 second sweep lengths.

For the Burg ARPE, the best performance occurred for sweep lengths of 0.3 and 0.4 seconds and 7 or 11 signal bins. For the MCM ARPE, the highest ROC performance occurred for 7 and 11 signal bins for a 0.3 second sweep length and for 21 signal bins for a 0.6 second sweep length. This analysis was repeated with MMOs of 30 and 40 however, there was essentially a performance degradation. In conclusion, the ROC results for the Burg and MCM ARPEs for sweep lengths other than 0.3 seconds show no significant detection improvement and thus there is no advantage to using sweep length times larger than 0.3 seconds for the Burg and MCM ARPEs. Furthermore, the sweep length time of 0.3 seconds is convenient considering the sweep length chosen for the Weighted Time Averaging analysis of Chapter 4.

Table 5.10. ROC results for the YW2, Burg and MCM ARPEs for different sweeps lengths and number of signal bins.

SL*	ARPE	Number of Signal Bins			
		3	7	11	21
0.3	YW2	71.78 (3.289) 25.8 (18.5, 34.8)	75.41 (3.121) 32.5 (24.4, 41.7)	76.68 (3.055) 26.7 (19.2, 35.6)	76.78 (3.049) 25 (17.8, 33.8)
0.4	YW2	80.56 (2.826) 37.5 (29, 46.8)	82.24 (2.715) 37.5 (28.3, 46.8)	82.66 (2.685) 35 (27, 44.3)	82.69 (2.683) 35 (27, 44.3)
0.6	YW2	81.54 (2.762) 46.7 (37.6, 56)	82.18 (2.718) 49.2 (40, 58.4)	82.59 (2.69) 45.8 (36.8, 55.1)	83.57 (2.619) 45 (36, 54.3)
0.8	YW2	73.77 (3.201) 33.3 (25.2, 42.6)	75.39 (3.122) 32.5 (24.4, 41.7)	76.59 (3.06) 35.8 (27.5, 45.1)	79.42 (2.898) 41.7 (32.9, 51)
0.3	Burg	74.57 (3.163) 30 (22.2, 39.1)	82.84 (2.67) 48.5 (39.2, 57.6)	84.32 (2.564) 49.2 (40, 58.4)	78.58 (2.948) 28.3 (20.7, 37.4)
0.4	Burg	75.45 (3.119) 40 (31.3, 49.3)	80.24 (2.847) 42.5 (33.7, 51.8)	83.05 (2.657) 48.5 (39.2, 57.6)	81.04 (2.796) 40.8 (32.1, 50.2)
0.6	Burg	76.81 (3.048) 38.3 (30, 47.7)	80.26 (2.846) 37.5 (29, 46.8)	79.66 (2.883) 35.8 (27.5, 45.1)	81.42 (2.77) 40 (31.3, 49.3)
0.8	Burg	78.29 (2.965) 36.7 (28.2, 46)	78.85 (2.932) 40.8 (32.1, 50.2)	78.97 (2.925) 45 (36, 54.3)	77.7 (2.999) 38.3 (29.8, 47.7)
0.3	MCM	78.14 (2.974) 32.5 (24.4, 41.7)	83.75 (2.611) 42.5 (33.7, 51.8)	84.66 (2.537) 35 (26.7, 44.3)	79.81 (2.874) 27.5 (20, 36.5)
0.4	MCM	77.4 (3.015) 43.3 (34.4, 52.7)	79.03 (2.921) 29.2 (21.4, 38.3)	81.76 (2.748) 35 (26.7, 44.3)	81.42 (2.77) 35 (26.7, 44.3)
0.6	MCM	77.93 (2.986) 40 (31.3, 49.3)	80.97 (2.8) 39.2 (30.5, 48.5)	80.68 (2.819) 35.8 (27.5, 45.1)	82.22 (2.716) 40 (31.3, 49.3)
0.8	MCM	76.56 (3.061) 34.2 (25.9, 43.4)	77 (3.038) 39.2 (30.5, 48.5)	77.92 (2.986) 42.5 (33.7, 51.9)	80.02 (2.861) 42.5 (33.7, 51.8)

* SL stands for Sweep Length and the units are seconds.

The last question to be answered was whether there was any advantage in ROC performance when using a constant model order to estimate the AR parameters. This is in contrast to estimating AR models from order 1 to 20 and then using the AIC MOE to choose which AR model to use. This was done for all ARPEs with special emphasis on the YW4 ARPE since it is not strictly valid to use the AIC MOE to determine model order in this case (personal communication with Prof. Hatzinakos, 1998). The bandpass filter passband was from 37 to 43 Hz. The model order was varied from 2 to 30 in multiples of 2 and the sweep time length was either 0.3, 0.4, 0.6 or 0.8 seconds. The results in Table 5.11 highlight the important findings.

The ROC results for the YW4 ARPE are not listed because the performance was found to still be poor, however RA increased to the order of 77%. This was obtained with a 0.3 second sweep length and an AR order of 24. Increasing the sweep length to 0.4 seconds

resulted in very similar performance but increasing the sweep length to 0.6 or 0.8 seconds lead to worse performance.

Table 5.11. ROC results for bandpass filtering from 37 to 43 Hz with constant model order.

ARPE	ARO*	SL*	Number of Signal bins			
			3	7	11	21
YW2	14	0.4	82.98 (2.663) 53.3 (44.1, 62.4)	83.73 (2.608) 50 (40.8, 59.2)	84 (2.588) 46.7 (37.6, 56)	82.96 (2.664) 47.5 (38.4, 56.8)
YW2	16	0.4	83 (2.661) 40 (31.3, 49.3)	83.72 (2.609) 49.2 (40.2, 58.2)	83.88 (2.596) 45.8 (36.8, 55.1)	82.45 (2.7) 42.5 (33.9, 51.8)
YW2	18	0.4	83.29 (2.64) 45.8 (36.8, 55.1)	83.94 (2.592) 46.7 (37.6, 56)	84.03 (2.585) 47.5 (38.9, 56.8)	82.44 (2.701) 45 (36, 54.3)
YW2	20	0.4	83.4 (2.632) 45 (36.8, 54.3)	84.01 (2.587) 44.2 (35.2, 53.5)	84.17 (2.575) 41.7 (32.4, 50.8)	82.45 (2.7) 44.2 (35.2, 53.5)
YW2	14	0.6	80.55 (2.827) 50 (40.8, 59.2)	81.39 (2.772) 45 (36, 54.3)	81.9 (2.738) 42.5 (33.7, 51.8)	83.17 (2.649) 40.8 (32.1, 50.2)
YW2	16	0.6	81.32 (2.778) 53.3 (44.1, 62.4)	82.07 (2.727) 47.5 (38.4, 56.8)	82.61 (2.689) 38.3 (29.8, 47.0)	83.28 (2.614) 41.7 (32.9, 51)
YW2	18	0.6	82.08 (2.726) 52.5 (43.2, 61.6)	82.85 (2.672) 50 (40.8, 59.2)	83.22 (2.645) 43.3 (34.4, 52.7)	83.38 (2.637) 40 (31.3, 49.3)
YW2	20	0.6	81.55 (2.761) 50 (40.8, 59.2)	82.31 (2.71) 48.9 (39.2, 57.6)	82.82 (2.674) 45 (36, 54.3)	83.15 (2.65) 40 (31.3, 49.3)
Burg	14	0.3	75.55 (3.114) 26.7 (19.2, 35.6)	83.03 (2.659) 45.8 (36.8, 55.1)	84.16 (2.575) 45 (36, 54.3)	78.7 (2.941) 27.5 (20, 36.5)
Burg	16	0.3	75.73 (3.105) 25 (17.8, 33.9)	83.23 (2.644) 41.7 (32.9, 51)	84.58 (2.543) 43.3 (34.4, 52.7)	78.58 (2.948) 26.7 (19.2, 35.6)
Burg	18	0.3	74.76 (3.154) 26.7 (19.2, 35.6)	82.86 (2.671) 45.8 (36.8, 55.1)	84.25 (2.569) 44.2 (35.2, 53.5)	78.41 (2.958) 25.8 (18.5, 34.8)
MCM	14	0.3	80.08 (2.857) 40.8 (32.1, 50.2)	84.2 (2.572) 45 (36, 54.3)	84.5 (2.55) 42.5 (33.7, 51.8)	79.42 (2.898) 30 (22.2, 39.1)
MCM	16	0.3	80.2 (2.849) 45 (36, 54.3)	84.65 (2.538) 39.2 (30.5, 48.5)	84.95 (2.514) 37.5 (29, 46.8)	78.68 (2.942) 20 (13.5, 28.5)
MCM	18	0.3	79.22 (2.91) 34.2 (25.9, 43.4)	84.47 (2.552) 42.5 (33.7, 51.8)	84.25 (2.569) 37.5 (29, 46.8)	78.82 (2.934) 27.5 (20, 36.5)

* SL stands for sweep length with units of seconds.

* ARO stands for AR model order

The results show that with constant model orders in the range of 14 to 18, the ROC performance of the Burg and MCM methods can be maintained when the sweep length is 0.3 seconds (for longer sweep lengths there is similar performance). Thus, a constant model order of 14 can be used for both methods to reduce computational overhead. The same ROC performance holds for the YW2 method for model orders of 14 to 20 when the sweep length is 0.6 seconds meanwhile there is a slight advantage for a sweep length of 0.4 seconds. These findings are important because using constant model orders results in reduced computational overhead.

5.7.5 The AR-Based SSABR Detector (Putting it all Together)

This section recapitulates the findings of section 5.7. The parameters which lead to maximum detection for the AR-based SSABR Detector are a sweep length time of 0.3 seconds and a passband of 37 to 43 Hz for the bandpass filter. For the detection statistic, 11 bins can be used for the signal band (7 bins resulted in similar yet slightly smaller ROC values). Further, the noise bandwidth is 6 Hz centered on 40 Hz. The Burg and MCM ARPEs showed the best performance for a 0.3 second sweep length and thus either ARPE may be used. Furthermore, AR models can be generated up to an MMO of 20 and the AIC MOE can be used to choose the final AR model. Alternatively, a constant model order of 14 can be used. To decrease the amount of computational overhead, a constant model order of 14 will be used. These parameters are summarized in Figure 5.4.

Finally some comments on the HOS-based Yule-Walker (YW4) ARPE are warranted. Results showed that the YW4 method resulted in the poorest detection performance out of all of the ARPEs. This may have occurred for a number of reasons. Firstly, the number of data samples (in the WTA result) may not have been sufficient to reduce the estimation variance of the 4^{th} order cumulant. Increasing the sweep length would address this (we varied the samples from 147 (a 0.3 second sweep) to 392 (a 0.8 second sweep)) but this didn't improve detection performance. However, we must also remember that the increased sweep length resulted in a reduction in SNR due to a reduction in the number of sweeps averaged (i.e. for 0.3 second sweeps there are 16 sweeps averaged but for 0.8 second sweeps there are only 6 sweeps averaged). Secondly, an MOE for HOS-based AR parameter estimation has not really been devised yet and this makes choosing the correct model order more difficult. Finally the noise in the SSABR recordings may not have been truly gaussian. This is seen in the contrasting reports on the gaussianity of the EEG (see section 3.1).

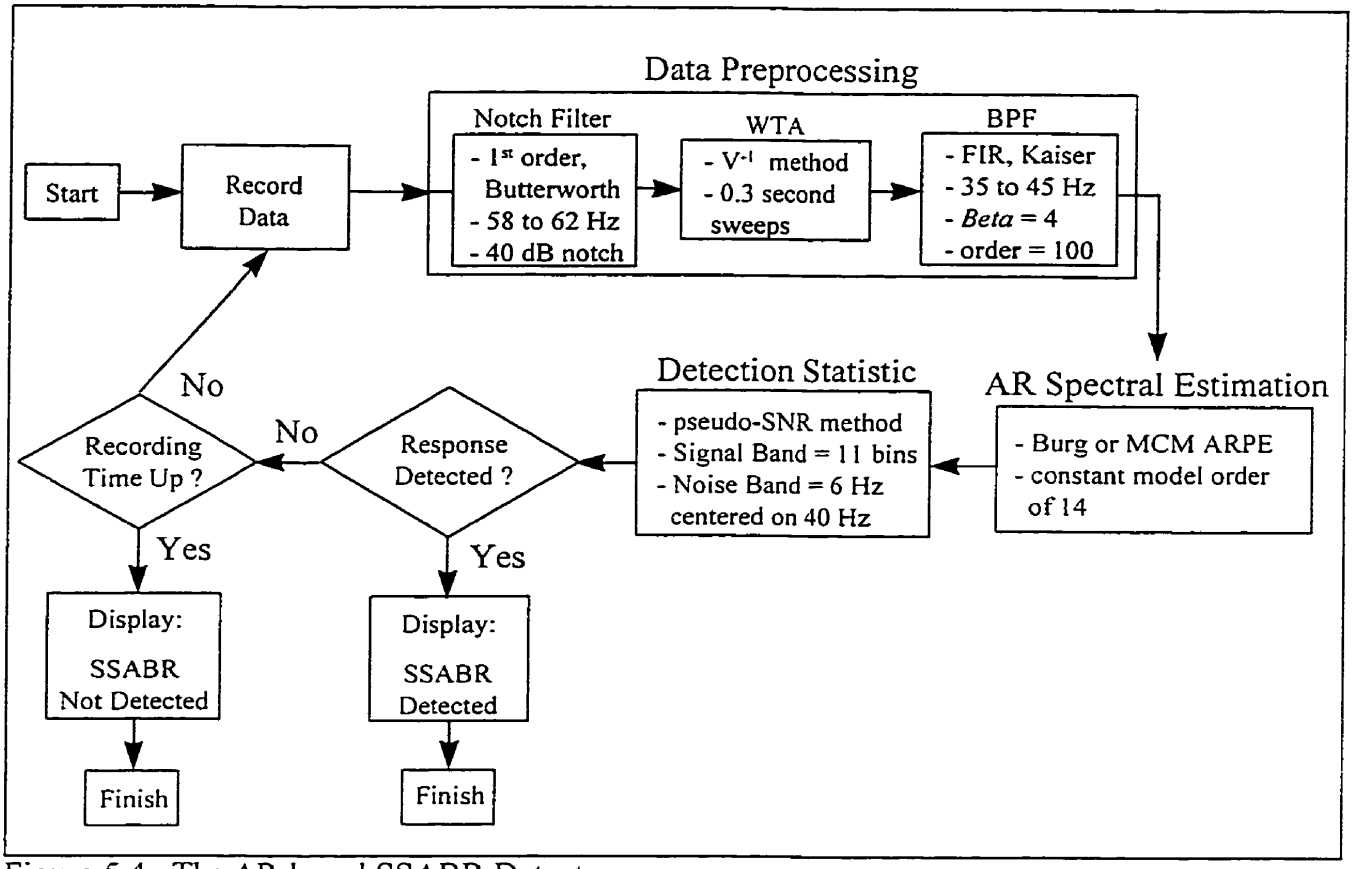

Figure 5.4. The AR-based SSABR Detector.

6.0 SSABR DETECTOR COMPARISON

In this chapter, the MSC, MSC-WA and AR-SSABR detectors will be compared on real 40 Hz SSABR data taken from human subjects under quiet and noisy recording conditions. Quiet recording conditions will most likely be the norm in clinical practice, however, data was also obtained under noisy recording conditions in order to determine the relative robustness of the different SSABR detectors. In addition, the stimulus intensity was varied to determine detector performance under different SSABR SNR conditions.

To begin with, the testing methods used and test cases performed will be considered in section 6.1. In section 6.2, the performance of the MSC and MSC-WA detectors will be examined using ROC analysis with particular emphasis on modifications made (WTA *within each* subaverage and filtering on each subaverage result) which should improve detector performance. Based on this analysis, choices will be made concerning which modifications for the MSC and MSC-WA detectors will be used in the comparison with the AR-SSABR detector. Section 6.3 will outline the implementation of each detector studied in the ROC comparison. In section 6.4, the detectors will be compared using ROC analysis. All of the ROC analysis in this chapter will be based on test trials of varying lengths to see how the performance of each detector varies with test time. In section, 6.5, the detectors will be compared in terms of computational complexity. To wrap up, section 6.6 will discuss the results obtained and make conclusions.

6.1 Data Acquisition: Methods, Subjects and Test Cases

The test group was 8 normal hearing adult male subjects with no history of auditory problems and hearing levels better than 20 dB HL at 1 kHz. For a description of the test setup and testing methods, please see Appendix C. In total, 8 different tests were conducted, however, there were only two test types: one test was done under quiet recording conditions (Quiet EEG test) and the other test was performed under noisy recording conditions (Noisy EEG test). In the Quiet EEG test, the subjects were instructed to remain very still to avoid adding any artifacts to the recording. In the Noisy EEG test, the subjects were instructed to introduce physiological noise of random time duration every two to five seconds.

Table 6.1 shows the different stimulus parameters and test types which were performed on each subject. Tests #1, 3, 5 and 7 were response tests (i.e. the stimulus would evoke an SSABR) to provide true positive data for ROC analysis. Tests #2, 4, 6 and 8 were no response tests to provide true negative data for ROC analysis. In these test cases, the stimulus was just a pure tone at 1 kHz which would not evoke an SSABR but would provide recording conditions similar to those for the 'response tests'. Noisy EEG tests were done only at the higher stimulus intensity because performing the test at the lower stimulus intensity would require much more data collection[14]. The stimulus intensity was also varied to determine the performance of each detector under different SSABR SNR conditions. The tests were performed with a 70 dB SPL stimulus (these are the Supra-threshold tests) and a 50 dB SPL stimulus (these are the Mid-threshold tests). In all test cases, the stimulus was applied to the subject's right ear and each recording was continuous and 219.2 seconds in duration. Adaptation should not play a role for this test duration (see section 2.5.4.3). In addition, the test subjects were given breaks of approximately 2 to 4 minutes between tests.

Table 6.1. Test type and Stimulus Parameters used for ROC Analysis.

Test Number	Test Type	f_c (Hz)	f_m (Hz)	μ	L_c (dB SPL)
1	Quiet EEG	1000	40	0.95	70
2	Quiet EEG	1000	0	0	70
3	Quiet EEG	1000	40	0.95	50
4	Quiet EEG	1000	0	0	50
5	Quiet EEG	1000	40	0.95	50
6	Quiet EEG	1000	0	0	50
7	Noisy EEG	1000	40	0.95	70
8	Noisy EEG	1000	0	0	70

6.2 MSC Algorithm Modification

This section will examine two modifications for the MSC and MSC-WA detectors: filtering on each subaverage result and WTA *within each* subaverage. In section 6.2.1, the basis for these modifications and their implementation will be outlined. In section 6.2.2, these modified MSC and MSC-WA methods will be compared to the unmodified MSC and MSC-WA methods under the test conditions described in section 6.1 using ROC analysis.

[14] Both the lower stimulus intensity and the large amount of physiological noise would make for very low SSABR SNR; thus longer data lengths would be needed for reasonable ROC analysis.

6.2.1 Proposed Modifications

As described previously in section 3.5, there are two modifications to the MSC methods which should be beneficial for SSABR detection for short data segments. Firstly, Dobie et al., (1995) used V^{-1} WTA on the MSC method in which each subaverage result was weighted with its inverse variance prior to MSC calculation. Dobie had noted that this resulted in slightly, however not statistically, improved detection over the MSC method. However, this weighting was applied after using STA *within each* subaverage. Thus any nonstationarities in the recording which were much shorter than 13 seconds in duration (the length of data used to create a subaverage) would probably go unnoticed in the weighting algorithm. Thus one improvement would be to use WTA *within each* subaverage instead of STA. Dobie also noted this but never investigated it. From the analysis of Chapter 4, the V^{-1} method would be a suitable WTA method, however, we should expect no real advantage under quiet recording conditions. The real advantage should come for data recorded under noisy recording conditions. Thus both the MSC and MSC-WA methods will be augmented by using V^{-1} WTA on sweeps *within each* subaverage. The sweep length in all subaverages is 0.3 seconds.

The other modification has to do with the limitations of the periodogram. Specifically, when trying to detect SSABRs in a short period of time then either each MSC subaverage will be short in duration or the number of averages within each subaverage will be small. To compound matters, the EEG has a huge amount of energy below 25 Hz. Thus for a short subaverage duration or for a small number of averages per subaverage, the EEG noise power will still be large and there will most likely be some spectral spreading from this noise in each subaverage FFT that will affect the calculation of the MSC value. To improve this situation, filtering can be done to each subaverage result before the FFT is done. The filter which was used was an FIR filter with a Kaiser window ($\beta = 4$). The filter order was 100 and the passband was from 35 to 45 Hz. A special filtering method was used (see Appendix G). Both the MSC and MSC-WA detectors will be augmented with this filtering technique.

6.2.2 Modified MSC methods: ROC Analysis

This section will discuss the ROC performance of the unmodified MSC and MSC-WA methods and the modified MSC and MSC-WA methods. In implementing the MSC detectors,

the literature states that the optimal number of subaverages for SSABR detection is between 8 and 16 and that the number of segments averaged per subaverage doesn't really affect detection performance as long as the same time length of data is used in calculating each subaverage[15] (Dobie et al., 1993). However, in Dobie's study, approximately 204 seconds were used to calculate the MSC value and in our study we are using much shorter time lengths (4.8 seconds at the least). To determine whether a different number of subaverages or different number of sweeps averaged for each subaverage are optimal for this shorter time length, ROC analysis was performed on 3 subjects (see Appendix I). The findings were that there was a slight performance enhancement when using 16 subaverages with no averaging in each subaverage (i.e. just 1 sweep in each subaverage) for a data length of 4.8 seconds. In addition, a sweep length of 0.3 seconds was used.

The ROC results for the MSC method are shown in Tables 6.2, 6.3 and 6.4 for the Supra-threshold Quiet, Mid-threshold Quiet and Supra-threshold Noisy EEG tests respectively. The ROC results for the MSC-WA method are shown in Tables 6.5, 6.6 and 6.7 for the Supra-threshold Quiet, Mid-threshold Quiet and Supra-threshold Noisy EEG tests respectively. For the Quiet EEG tests, **Artif**act **R**ejection (AFR) was used to reject any buffers (0.1 second duration) with data which had amplitudes exceeding ± 40 μV. To show the advantages of WTA and to increase the number of test cases for ROC analysis, AFR was not used for the Noisy EEG test.

For both the Quiet and Noisy EEG Supra-threshold tests, the data lengths for ROC analysis will be 4.8, 9.6 and 19.2 seconds. For the Mid-threshold test, the data lengths will be 14.4, 19.2 and 28.8 seconds (i.e. longer data lengths because of smaller SNR). In addition, the ROC entries in each table are of the following format. The first line of each table entry is the **R**OC **A**rea (RA) followed by the **S**tandard **E**rror (SE) in the estimation of the RA. The second line will be the Hit Rate at a False Alarm Rate of 5% denoted as HR(0.05) and the lower and upper 95% confidence bounds for this proportion follows in brackets.

The ROC results for the MSC detector will be discussed first. As Tables 6.2 and 6.3

[15] For instance if you have 100 sweeps per subaverage it doesn't matter whether you group 10 sweeps together (end to end) and average (10 averages per subaverage; each subaverage result has a length of 10 sweeps) or group 20 sweeps together and average (5 averages per subaverage; each subaverage result has a length of 20 sweeps) because both methods lead to similar detection performance.

indicate, the modification of filtering and/or WTA does not result in an increase in detector performance for the Supra- and Mid-threshold Quiet EEG tests. In fact, the ROC results of the unmodified MSC and MSC-WA methods are slightly larger than the modified MSC and MSC-WA employing both filtering and WTA. However, these differences are only statistically significant for the smallest time duration in each case (4.8 seconds for the Supra-threshold Quiet EEG test ($p < 3.0\%$, $\Delta RA=1.15\%$) and 14.4 seconds for the Mid-threshold Quiet EEG test ($p < 4.5\%$, $\Delta RA=1.85\%$)).

Table 6.2. ROC results for the MSC Detector with and without modifications for the Supra-threshold Quiet EEG test.

Filtering ?	WTA ?	Data Duration (seconds)		
		4.8	9.6	19.2
No	No	86(1.486) 54.4(48.8, 59.9)	93.26(1.468) 76.3(68.8, 82.4)	97.8(1.186) 97.5(90.5, 99.6)
No	Yes	86(1.486) 54.4(48.8, 59.9)	93.09(1.487) 77.5(70.1, 83.5)	97.79(1.188) 96.3(88.7, 99)
Yes	No	84.85(1.542) 47.2(41.6, 52.8)	92.54(1.545) 68.1(60.2, 75.1)	97.48(1.268) 92.5(83.8, 96.9)
Yes	Yes	84.85(1.542) 47.2(41.6, 52.8)	92.56(1.543) 70.6(62.8, 77.4)	97.45(1.276) 93.8(85.4, 97.7)

Table 6.3. ROC results for the MSC Detector with and without modifications for the Mid-threshold Quiet EEG test.

Filtering ?	WTA ?	Data Duration (seconds)		
		14.4	19.2	28.8
No	No	86.69(1.734) 56.3(49.5, 62.8)	89.05(1.868) 66.3(58.3, 73.4)	93.63(1.707) 76.8(67.7, 84)
No	Yes	86.3(1.758) 56.3(49.5, 62.8)	89.13(1.861) 63.1(55.1, 70.5)	92.81(1.815) 76.8(67.7, 84)
Yes	No	85.15(1.827) 49.1(42.4, 55.8)	88.22(1.935) 61.3(53.2, 68.7)	93.53(1.721) 77.7(68.9, 84.5)
Yes	Yes	84.84(1.844) 54.5(47.7, 61.1)	87.89(1.961) 63.1(55.1, 70.5)	92.74(1.823) 74.1(64.8, 81.7)

For the Noisy EEG test case we first note how poorly MSC performs without Artifact Rejection or WTA. Furthermore, we note the improvement of the MSC detector with WTA *in each* subaverage compared to the conventional MSC method (statistically significant for data duration = 9.6 and 19.2 seconds with $p < 0.2\%$ and $p < 0.06\%$ respectively). Furthermore, filtering does not improve detection at all. Thus considering the results for the Noisy and Quiet

EEG tests, the MSC detector should be used with WTA on sweeps *within each* subaverage to improve detection. In later sections this detector will be called MSC_{WTA}.

Table 6.4. ROC results for the MSC Detector with and without modifications for the Supra-threshold Noisy EEG test.

Filtering ?	WTA ?	Data Duration (seconds)		
		4.8	9.6	19.2
No	No	59.09(2.238) 15.3(11.7, 19.8)	63.77(3.081) 23.1(17, 30.6)	72.77(3.981) 26.3(17.4, 37.5)
No	Yes	59.09(2.238) 15.3(11.7, 19.8)	71.2(2.868) 34.4(27.2, 42.3)	84.73(3.105) 61.3(49.7, 71.7)
Yes	No	58.06(2.248) 15(11.4, 19.5)	62.15(3.115) 21.3(15.4, 28.5)	67.95(4.212) 33.8(23.8, 45.3)
Yes	Yes	58.06(2.248) 15(11.4, 19.5)	68.76(2.949) 36.9(29.5, 44.9)	83.92(3.181) 58.8(47.2, 69.4)

For the MSC-WA detector, we can make the same observations that were made for the MSC detector. For both Quiet EEG tests, the use of WTA and filtering do not really aid in detector performance. In fact the filtering modification results in a consistent degradation compared to the MSC-WA method which is statistically significant for data duration = 4.8 seconds ($p < 2.2\%$, $\Delta RA = 4.65\%$) for the Supra-threshold Quiet EEG test and data duration = 28.8 seconds ($p < 1.2\%$, $\Delta RA = 4.5\%$) for the Mid-threshold Quiet EEG test.

Table 6.5. ROC results for the MSC-WA Detector with and without modifications for the Supra-threshold Quiet EEG test.

Filtering ?	WTA ?	Data Duration (seconds)		
		4.8	9.6	19.2
No	No	86.53(1.459) 51.9(46.3, 57.4)	91.48(1.65) 71.9(64.2, 78.5)	97.52(1.26) 97.5(90.5, 99.6)
No	Yes	86.53(1.459) 51.9(46.3, 57.4)	92.66(1.532) 77.5(70.1, 83.5)	97.89(1.16) 98.8(92.3, 99.9)
Yes	No	81.88(1.675) 41.9(36.5, 47.5)	90.35(1.755) 61.3(53.2, 68.7)	96.12(1.576) 92.5(83.9, 96.9)
Yes	Yes	81.88(1.675) 41.9(36.5, 47.5)	90.9(1.705) 56.9(48.8, 64.6)	96.44(1.51) 91.3(82.3, 96.1)

For the Noisy EEG test, once again, the use of WTA *within each* subaverage results in a statistically significant improvement in detection performance over the MSC-WA method for data duration = 9.6 seconds ($p < 1.3\%$, $\Delta RA = 4.35\%$) and for data duration = 19.2 seconds ($p < 2.5\%$, $\Delta RA = 5.88\%$). Thus, filtering should not be used but WTA *within each* subaverage should definitely be used. In later sections, this detector will be called $MSC\text{-}WA_{WTA}$.

Table 6.6. ROC results for the MSC-WA Detector with and without modifications for the Mid-threshold Quiet EEG test.

Filtering ?	WTA ?	Data Duration (seconds)		
		14.4	19.2	28.8
No	No	83.22(1.934) 48.7(42, 55.4)	88.18(1.939) 58.8(50.7, 66.4)	93.3(1.751) 74.1(64.8, 81.7)
No	Yes	84.98(1.83) 56.7(50, 63.2)	87.57(1.986) 60.6(52.6, 68.1)	92.32(1.875) 72.3(63, 80.1)
Yes	No	84.21(1.88) 51.3(44.6, 58)	86.2(2.089) 57.5(49.5, 65.1)	91.15(2.012) 68.8(59.2, 77)
Yes	Yes	82.91(1.95) 52.2(45.5, 58.9)	84.87(2.182) 56.3(48.2, 64)	88.82(2.258) 67(57.4, 75.4)

Table 6.7. ROC results for the MSC-WA Detector with and without modifications for the Supra-threshold Noisy EEG test.

Filtering ?	WTA ?	Data Duration (seconds)		
		4.8	9.6	19.2
No	No	69.85(2.059) 27.8(23.1, 33.1)	73.43(2.784) 38.8(31.3, 46.8)	84.78(3.1) 60(48.5, 70.6)
No	Yes	69.85(2.059) 27.8(23.1, 33.1)	77.78(2.591) 44.4(36.6, 52.4)	90.66(2.448) 76(63.9, 83.7)
Yes	No	69.62(2.064) 25.6(21, 30.8)	71.53(2.856) 35.6(28.3, 43.6)	81.78(3.368) 43.8(32.9, 55.2)
Yes	Yes	69.62(2.064) 25.6(21, 30.8)	75.67(2.69) 33.1(26, 41)	84(3.174) 62.5(51, 72.8)

In conclusion we see that filtering did not improve detection performance for the MSC or MSC-WA detectors. Perhaps the filter bandwidth was not small enough (it was 10 Hz meanwhile the frequency resolution was 3.3 Hz). In retrospect a smaller bandwidth may have been used such as 38 to 42 Hz to further reduce noise contamination. The average of response power (16 subaverages) in the MSC calculation should aid in distinguishing signal from noise. The only disadvantage is that a smaller filter bandwidth would result in longer transients in the filter output which would harm signal detection. This effect remains to be determined. On the other hand, the modification of WTA *within each* subaverage greatly improved detection performance for the Noisy EEG test case for data durations of 9.6 and 19.2 seconds for the MSC and MSC-WA methods. In the other test cases, we saw comparable detection performance between the WTA modified and unmodified MSC and MSC-WA methods. Therefore, the WTA modified MSC and MSC-WA detectors (termed MSC_{WTA} and $MSC\text{-}WA_{WTA}$ respectively) will be used for future ROC analysis.

6.3 Implementation of SSABR Detectors

The detectors considered in this comparison are the MSC_{WTA}, $MSC\text{-}WA_{WTA}$, the AR-Based SSABR Detector and a detector based on the periodogram. All detectors had a preprocessing block which removed the line noise which was prominent in our recordings (fairly stable at 60 Hz with amplitudes from 1 to 4 μV across subjects). For this purpose, a first order Butterworth IIR notch filter, in the non-phase shift configuration (Antoniou, 1993), with a bandstop region of 58 to 62 Hz and a 40 dB notch was used.

In implementing the MSC detectors, as outlined in section 6.2.2, both the MSC_{WTA} and $MSC\text{-}WA_{WTA}$ detectors had 16 subaverages, 0.3 second sweeps and WTA (V^{-1} method) was done across sweeps *within each* subaverage. In addition, Artifact Rejection was used for both the MSC_{WTA} and $MSC\text{-}WA_{WTA}$ detectors for the Quiet EEG tests. The Artifact Rejection algorithm rejected buffers of data (0.1 seconds in duration) which had amplitudes exceeding ± 40 μV (Lins et al., 1996). AFR was not done for the Noisy EEG test because of a lack of data to perform a good statistical ROC analysis; furthermore an equal comparison (using the same amount of data) with the AR-based detectors was desired.

To implement the AR-based SSABR detectors we turn to Chapter 5 for the detector criteria. Basically Chapter 5 indicated that either the Burg or MCM ARPE can be used. We decided that both ARPEs would be used which resulted in two AR-based SSABR detectors: the SSABR-Burg and SSABR-MCM detectors which both used a constant model order of 14. In addition, the sweep length was 0.3 seconds, V^{-1} WTA was used on the recorded sweeps and the sweeps were filtered by an FIR filter (see Appendix G) designed with a Kaiser window (β = 4, filter order = 100) with a passband from 37 to 43 Hz. For the detection statistic, the signal band contained 11 bins (1 bin = 0.25 Hz) and the noise band was from 37 to 43 Hz (not including the signal band centered at 40 Hz). In addition, unlike the MSC detectors, the data was not passed through an AFR algorithm because of the use of WTA preprocessing. Please note that ROC results were also obtained using 7 bins for the signal band and the AIC MOE with an MMO of 20 (instead of using a constant model order of 14). These ROC results were quite similar (slightly less in value) to the ones shown in this section for the AR-based SSABR detectors and are shown in Appendix J.

A detector based on the periodogram was also implemented to see if AR modeling was indeed giving any performance advantage. The periodogram based detector was exactly like the AR-based detectors in that the V^{-1} WTA method was used with 0.3 second sweeps and the same filter was used. In addition, the same detection statistic was used with the signal band containing 7 bins (1 bin = 0.5 Hz) and the noise band occupying the 37 to 43 Hz region (excluding the signal band). These parameters were chosen for SSABR detection in the same manner used in Chapter 5 for the AR-based SSABR detectors. This detector will be called SSABR-PER.

6.4 SSABR Detector Comparison: ROC Analysis

ROC analysis (see Appendix E) was used for a comparison of the detection capabilities of the SSABR detectors. The performance of the different detectors was compared across different stimulus conditions (Supra-threshold and Mid-threshold), recording conditions (quiet and noisy) and different data time durations. The test results for the Supra-threshold Quiet EEG test case, Mid-threshold Quiet EEG test case and the Supra-threshold Noisy EEG test case are shown in Tables 6.8, 6.9 and 6.10 respectively.

The Supra-threshold Quiet EEG test results indicate that the MSC_{WTA} and $MSC\text{-}WA_{WTA}$ methods are the best detectors across the different data lengths in terms of RA and HR(0.05) except for the data duration of 19.2 seconds where the SSABR-Burg method is similar in performance. There was no statistical difference between the MSC_{WTA} and the $MSC\text{-}WA_{WTA}$ methods for all data durations for RA. The results for HR(0.05) are also very similar. For the AR based methods, the SSABR-Burg method was better than the SSABR-MCM method for all data durations for RA ($p < 0.3\%$, $p < 0.02\%$, $p < 0.4\%$, for data durations of 4.8, 9.6 and 19.2 seconds respectively). There is also a large discrepancy in HR(0.05) between these detectors. Furthermore, the performance advantage of the SSABR-Burg method over the SSABR-MCM method increased with data duration.

Comparing the SSABR-PER detector with the SSABR-Burg detector, the results are very similar and there is no statistical difference for RA across all data durations. However, the SSABR-Burg method does seem to have a larger HR(0.05) than the SSABR-PER detector across all data durations. Finally comparing the SSABR-Burg method to the $MSC\text{-}WA_{WTA}$ method, we see that the $MSC\text{-}WA_{WTA}$ method is statistically significantly better in RA for data

durations of 4.8 and 9.6 seconds but not 19.2 seconds ($p < 0.1\%$, $p < 2.5\%$, $p < 31\%$, for data durations of 4.8, 9.6, 19.2 seconds respectively). We also see a better HR(0.05) for the MSC-WA_{WTA} method than we do for the SSABR-Burg method. Thus either the MSC_{WTA} or MSC-WA_{WTA} detection methods should be used in this test situation.

Table 6.8. ROC results for each SSABR Detector versus data duration for the Supra-threshold Quiet EEG test.

SSABR Detector	Data Duration (seconds)		
	4.8	9.6	19.2
MSC_{WTA}	86(1.486) 54.4(48.8, 59.9)	93.09(1.487) 77.5(70.1, 83.5)	97.79(1.588) 96.3(88.7, 99)
MSC-WA_{WTA}	86.53(1.459) 51.9(46.3, 57.4)	92.66(1.532) 77.5(70.1, 83.5)	97.89(1.516) 98.8(92.3, 99.9)
SSABR-Burg	81.59(1.687) 43.8(38.3, 49.4)	88.96(1.875) 51.3(43.3, 59.2)	96.25(1.555) 90(80.8, 95.3)
SSABR-MCM	79.25(1.778) 28.4(23.6, 33.8)	84.63(2.197) 32.5(25.5, 40.4)	90.77(2.434) 52.5(41.1, 63.6)
SSABR-PER	81.54(1.689) 34.1(29, 39.6)	88.48(1.915) 47.5(39.6, 55.5)	95.23(1.675) 82.5(72.1, 89.7)

The Mid-threshold Quiet EEG test results indicate very similar results to those seen in the Supra-threshold Quiet EEG test. Once again both MSC detectors have similar detection performance however the MSC_{WTA} method seems to be performing slightly better in terms of RA and HR(0.05) although the performance increase for RA is not statistically significant for any data durations ($p < 6\%$, $p < 7.5\%$, $p < 54\%$, for data durations of 14.4, 19.2 and 28.8 seconds respectively). In addition, the SSABR-Burg method performs statistically significantly better than the SSABR-MCM method in terms of RA for all data durations ($p < 1\%$, $p < 0.7\%$, $p < 0.1\%$ for data durations of 14.4, 19.2, 28.8 seconds respectively). Also in terms of HR(0.05), we see an advantage for the SSABR-Burg method over the SSABR-MCM method.

Comparing the SSABR-PER and SSABR-Burg detectors, we see similar RA performance, however, the SSABR-PER detector has a much better HR(0.05) than the SSABR-Burg detector for data lengths of 14.4 and 28.8 seconds. Finally comparing the MSC_{WTA} and SSABR-Burg methods, the MSC_{WTA} method performs better for all data durations but none of the results were statistically significant. In terms of HR(0.05), the MSC_{WTA} method has better performance than the SSABR-Burg method for all data durations. Thus either the MSC_{WTA} or MSC-WA_{WTA} detection methods should be used in this test situation.

Table 6.9. ROC results for each SSABR Detector versus data duration for the Mid-threshold Quiet EEG test.

SSABR Detector	Data Duration (seconds)		
	14.4	19.2	28.8
MSC_{WTA}	86.3(1.758) 56.3(49.5, 62.8)	89.13(1.861) 69.1(65.1, 70.5)	92.81(1.815) 76.8(69.7, 84)
$MSC\text{-}WA_{WTA}$	84.98(1.85) 56.7(49.9, 63.2)	87.57(1.986) 60.6(52.6, 68.1)	92.32(1.875) 72.5(63, 80.1)
SSABR-Burg	81.78(2.007) 31.7(25.8, 38.3)	85.01(2.172) 50.6(42.7, 58.6)	88.21(2.317) 55.4(45.7, 64.6)
SSABR-MCM	79.38(2.12) 20.5(15.6, 26.5)	82.3(2.346) 28.8(22, 36.5)	84.38(2.649) 38.4(29.5, 48.1)
SSABR-PER	81.47(2.022) 42.4(35.9, 49.2)	84.69(2.191) 53.1(45.1, 61)	89.58(2.181) 70.5(61.1, 78.6)

The Supra-threshold Noisy EEG test provides some interesting results. First of all, no statistical testing on RA could be done for the 4.8 second data duration because RA < 70% and the method of Appendix E is not applicable in this case. Beginning with the MSC_{WTA} and $MSC\text{-}WA_{WTA}$ methods (recall no AFR was used), we see that the latter detector has statistically better RA for the 9.6 and 19.2 second data durations ($p < 1\%$ in both cases) This is also reflected in the HR(0.05) values. For the AR-based SSABR detectors, the difference in RA is not statistically significant for data durations of 9.6 seconds ($p < 30\%$) and 19.2 seconds ($p < 14\%$), although, in terms of HR(0.05), the SSABR-Burg method has a slight advantage over the SSABR-MCM method.

Table 6.10. ROC results for each SSABR Detector versus data duration for the Supra-threshold Noisy EEG test.

SSABR Detector	Data Duration (seconds)		
	4.8	9.6	19.2
MSC_{WTA}	59.09(2.238) 15.3(11.7, 19.8)	71.2(2.868) 34.4(27.2, 42.3)	84.73(3.105) 61.3(49.7, 71.7)
$MSC\text{-}WA_{WTA}$	69.85(2.059) 27.8(23.1, 33.1)	77.78(2.591) 44.4(36.6, 52.4)	90.66(2.448) 75(63.9, 83.7)
SSABR-Burg	68.24(2.094) 23.4(19, 28.5)	72.8(2.809) 28.1(21.5, 35.9)	83.45(3.224) 45(34, 56.5)
SSABR-MCM	68.49(2.089) 18.1(14.2, 22.9)	71.61(2.853) 25.6(19.2, 33.2)	81.02(3.43) 33.8(23.8, 45.3)
SSABR-PER	69.41(2.069) 20.5(16.1, 25.2)	73.12(2.796) 30(23.2, 37.8)	84.16(3.159) 63.8(52.2, 74)

Comparing the SSABR-PER and SSABR-Burg detectors, we see slightly better RA values for the SSABR-PER detector although these were not statistically significant for data

durations of 9.6 seconds ($p < 81\%$) and 19.2 seconds ($p < 65.3\%$). We also see a much better HR(0.05) value for the SSABR-PER detector for the 19.2 second data duration. Finally comparing the MSC-WA_{WTA} and SSABR-Burg detectors, we see statistically significant better RA values for the MSC-WA_{WTA} for the 9.6 second ($p < 5\%$) and 19.2 second ($p < 1.2\%$) data durations. This trend is also seen with the HR(0.05) values. Thus for Noisy EEG recordings, the MSC-WA_{WTA} method, results in the best detection performance.

6.5 Computational Complexity for the SSABR Detectors

The computational complexity of the different detectors is shown in Table 6.11. An explanation of these calculations is given in Appendix K. To put all detectors on an equal scale it is assumed that the data, upon which detection is performed, is N points long which can be segmented according to N = 16nM where 16n is the number of sweeps in the data and M is the number of points per sweep. For the MSC based detectors, the computational complexity is done for 16 subaverages.

The results indicate that both AR-based detectors have the same computational overhead and the SSABR-PER detector has much less computational overhead than the AR-based detectors. The highest complexity among the MSC methods belongs to the MSC-WA_{WTA} method which makes sense because all other MSC detectors are a subset of this detector. Comparing the AR-based detectors with the MSC-WA_{WTA} method is difficult but if we use the parameters from our ROC analysis (M = 149, n = 1, 2 or 4) then from Appendix K we see that the SSABR-PER method has the least amount of computational overhead followed by the SSABR-Burg method followed by the MSC-WA_{WTA} method.

Table 6.11. Computational Complexity for each SSABR Detector.

SSABR Detector	Number of Multiplications	Number of Additions	Number of Divisions	Number of log operations
MSC_{WTA}	$16(4nM+M) + 17M\log_2 M$	$16(5nM+M+n) + 17M\log_2 M$	64n	0
MSC-WA_{WTA}	$16(4nM+5M) + 17M\log_2 M$	$16(5nM+4M+n) + 17M\log_2 M$	64(n+1)	0
SSABR-Burg	$48nM+M^2$	$64nM+M^2+16n$	64n	1
SSABR-MCM	$48nM+M^2$	$64nM+M^2+16n$	64n	1
SSABR-PER	$48nM+M\log_2 M$	$64nM+M\log_2 M+16n$	64n	1

6.6 Discussion and Conclusions

In this chapter we have investigated various SSABR detectors via ROC analysis on real 40 Hz SSABR data recorded from 8 normal hearing adult subjects. The detectors investigated were the MSC and MSC-WA detectors as well as the modified MSC detectors: MSC_{WTA} and $MSC\text{-}WA_{WTA}$; two AR-based detectors: SSABR-Burg and SSABR-MCM and a periodogram based method: SSABR-PER. Firstly, the MSC and MSC-WA methods were modified to include filtering (35 to 45 Hz) on each subaverage result and WTA (V^{-1} method) on sweeps *within each* subaverage. ROC analysis showed that under quiet recording conditions, neither modification resulted in any detection improvement with the filtering method actually slightly degrading performance. However, under noisy recording conditions, the WTA modification significantly improved detection performance. Furthermore, the $MSC\text{-}WA_{WTA}$ method performed better than all of the other MSC methods (taking all test cases into consideration). Thus, for further analysis the WTA versions of the MSC and MSC-WA methods were used (MSC_{WTA}, $MSC\text{-}WA_{WTA}$).

ROC comparison between the SSABR-Burg and SSABR-MCM methods showed that the SSABR-Burg method always resulted in statistically significant better detection (using the RA and HR(0.05) measures) except for the Noisy recording case. However, as the SSABR SNR decreases the detection advantage of the SSABR-Burg method over the SSABR-MCM method diminishes. Furthermore, the SSABR-Burg and SSABR-PER detectors were compared and detection performance was similar across all data durations and recording situations (in some cases the SSABR-PER detector was much better in terms of HR(0.05)). Thus AR modeling did not really provide an advantage over the periodogram. This was most likely due to the small SSABR SNR after data preprocessing.

The last comparison made was the ROC performance between the $MSC\text{-}WA_{WTA}$ and SSABR-Burg method. Under the Mid-threshold quiet test, both methods were similar with a slight performance advantage going to the $MSC\text{-}WA_{WTA}$ method. Under the Supra-threshold quiet test, the $MSC\text{-}WA_{WTA}$ method was statistically significantly better for the 4.8 and 9.6 second data duration. Meanwhile, under the noisy recording condition, for the 9.6 and 19.2 second data durations, the $MSC\text{-}WA_{WTA}$ method resulted in better ROC performance. In fact, the most robust detector to noise was the $MSC\text{-}WA_{WTA}$ method.

It is also interesting to compare the performance of the detectors for the different stimulus levels and the different recording conditions (quiet versus noisy). Using the results of the MSC-WA_{WTA} detector we see similar detection for the Supra-threshold Quiet EEG test for a 4.8 second data duration and the Mid-threshold Quiet EEG test for a 14.4 second data duration. Further as the data duration is doubled for both recording conditions, we see similar increases in ROC Area. Thus, for the 50 dB SPL stimulus it takes about 3 times as much recording time (14.4/4.8 = 3) to get similar detection. We can also thus estimate that it will take approximately 60 seconds with the 50 dB SPL stimulus to get the high ROC areas and HR(0.05) values seen for the 70 dB SPL stimulus case using approximately 20 seconds.

Finally, the detectors were compared based on algorithmic complexity. Both AR-based methods had the same complexity which was more than that of the SSABR-PER detector which had the least amount of computational overhead. Furthermore, of all the MSC based detectors, the original MSC method without any modifications had the least computational overhead and the MSC-WA_{WTA} method had the most computational overhead which was larger than that of the AR-based detectors. Thus while the MSC-WA_{WTA} method is a little excessive in terms of computational overhead, it provides the best SSABR detection performance under all types of recording situations investigated in this work.

7.0 CONCLUSIONS AND FUTURE WORK

In section 7.1, the major contributions of this thesis will be summarized. Section 7.2 will discuss plans for future work which include other signal processing techniques which may be applied, changes to recording methodology and the analysis of different subject groups.

7.1 Summary of Research Contributions

In an attempt to improve SSABR detection in terms of reducing test time the following two objectives were found to be important: 1) identify an alternative method to Artifact Rejection which will make the detector robust against physiological artifacts while not requiring an excessive amount of recording time and 2) improve on the shortcomings of the periodogram based SSABR detectors and exploit any noise properties to achieve enhanced detection. To accomplish these objectives the following observations were originally made:

1) **W**eighted **T**ime **A**veraging (WTA) is more robust to the influence of physiological artifacts than the combination of **A**rtifact **R**ejection (AFR) and **S**ynchronous **T**ime **A**veraging (STA) and does not require an excessive amount of data elimination.
2) **A**utoregressive Modeling (AR) modeling can result in much better spectral estimates than periodogram based methods due to more realistic modeling assumptions about the data outside of the measurement interval. Further AR modeling based on **H**igher **O**rder **S**tatistics (HOS) can be used to attenuate gaussian noise (some researchers found the EEG noise to be gaussian).

Thus both methods were investigated to examine whether any performance advantages could be obtained over conventional SSABR signal processing methods.

For Weighted Time Averaging, four methods were investigated: the **M**inimum **E**nergy (ME) Technique, the Inverse Variance (V^{-1}) Technique, the Gerull method and a **V**ector **N**onlinear **F**iltering (VNF) method. These methods were studied using EEG noise recorded under both quiet and noisy conditions and simulated sinusoids (to simulate 40 and 80 Hz SSABRs). Using the measures of Root Mean Square Error, SNR and computational complexity the following observations were made:

1) WTA performs just as well as STA with AFR under quiet recording conditions for

both 40 and 80 Hz SSABRs. In addition, there was a slight advantage for the ME method in the 80 Hz SSABR case.

2) WTA performs just as well as STA with AFR under noisy recording conditions for both the 40 and 80 Hz SSABRs. In addition, there was a slight advantage for the V^{-1} method with PFFWC for the 80 Hz SSABR case.
3) Pre-filtering for weight calculation, in general, provided no performance improvement for the WTA methods (except for the V^{-1} method in the Noisy 80 Hz SSABR case).
4) For the 40 Hz SSABR recorded under noisy and quiet conditions, the V^{-1} method gave the best performance out of all the WTA methods.
5) For the 80 Hz SSABR recorded under noisy conditions, the V^{-1} method with pre-filtering for weight calculation gave the best performance meanwhile for quiet recording conditions, the ME method gave the best performance.
6) WTA resulted in a 36% savings in recording time compared to Artifact Rejection when applied to data obtained under noisy recording conditions.
7) The V^{-1} WTA method has the least amount of computational overhead among the different WTA algorithms.

Thus WTA, especially the V^{-1} method, is very beneficial as an SSABR preprocessing algorithm which will improve the SSABR SNR and robustify the detector from physiological artifacts meanwhile not throwing out an excessive amount of data.

The second part of this research consisted of designing an SSABR detector based on AR spectral estimation and comparing it to the current benchmarks in SSABR detection: the MSC and MSC-WA methods. For the design of the AR-based detectors, values for the different design parameters were estimated based on ROC analysis of real 40 Hz SSABRs. Basically, four different AR parameter estimators were investigated: the Yule-Walker (Autocorrelation) method, the Burg Method, the MCM method and the HOS-based Yule-Walker method using 4^{th} order cumulants. The HOS method was used because HOS can attenuate Gaussian noise; further there are reports of the EEG noise being nearly Gaussian. For preprocessing, the V^{-1} WTA method and pre-filtering with a 37 to 43 Hz bandpass filter was used. Investigation into the design of the AR-based SSABR detector resulted in the following observations:

1) The HOS-based AR parameter estimator always resulted in the worst detection performance. This could be due to an inadequate number of data samples, the lack of a model order estimation algorithm for HOS AR parameter estimation or the EEG may simply be non-Gaussian.
2) Based on a 0.3 second sweep length the Burg and MCM AR parameter estimators gave the best detection performance. It was also found that either the AIC model order estimator (with a maximum model order of 20) could be used for AR model selection or a constant model order of 14 can be used. Both techniques lead to very similar detection performance. For our work a constant model order of 14 was used.

For the ROC analysis, modifications were made to the MSC and MSC-WA detectors to improve performance. WTA was used *within each* subaverage (already proposed by Dobie, 1995a) and/or filtering was used on each subaverage result. Using ROC analysis for quiet recording conditions, the WTA modification did not really change detection performance and the filtering modification degraded performance slightly. For the noisy recording conditions, filtering still degraded performance, however, the WTA modification resulted in an improvement in detection. Thus the MSC and MSC-WA methods were used with the WTA modification in the ROC analysis with the other SSABR detectors. These modified MSC methods were called the MSC_{WTA} and $MSC\text{-}WA_{WTA}$ methods.

ROC analysis further showed that the MSC_{WTA} and the $MSC\text{-}WA_{WTA}$ detectors had similar detection performance for all test conditions except for the noisy EEG test in which the $MSC\text{-}WA_{WTA}$ method was statistically significantly better. For the AR-based detectors, the SSABR-Burg detector was statistically significantly better than the SSABR-MCM method in all test conditions except for the noisy EEG test. In addition, the SSABR-PER detector performed just as well as the SSABR-Burg detector for all test cases and data durations. This indicates that AR spectral estimation didn't really provide an advantage for SSABR detection. This was most likely due to inadequate SSABR SNR after data preprocessing. Finally, comparing the SSABR-Burg detector to the $MSC\text{-}WA_{WTA}$ detector, we note that the $MSC\text{-}WA_{WTA}$ detector had larger values of RA and HR(0.05) for practically all test conditions (stimulus level, recording type and data duration). Thus the $MSC\text{-}WA_{WTA}$ method is the best SSABR detector for data recorded under both quiet and noisy recording conditions.

7.2 Future Work and Directions

Future work can be done on three separate fronts: 1) repeat this analysis on neonates, 2) use alternate recording methods and 3) use alternate signal processing methods. Repeat analysis on neonates may be helpful to determine if the noise associated with their SSABRs has a different pdf which may warrant specialized signal processing techniques such as HOS or Locally Optimum Detection (discussed in section 7.2.2). The latter two methods will be discussed in more detail.

7.2.1 Alternate Recording Methods

Alternate recording methods comprise two different techniques. Firstly, different scalp electrode locations may be used where the noise properties are different. For instance, although McEwen et al. (1975) found similar gaussianity and stationarity test results for the EEG recorded from different scalp locations, they stated that the occipital EEG was more consistently gaussian and stationary than EEG data obtained from the frontal area. Thus one test could be to compare the detection performance of HOS-based detectors on data recorded from the frontal area versus the occipital area.

Another recording method would be to perform multi-channel SSABR recordings. This would involve recording the SSABR from different scalp sites simultaneously. The advantage here would be that more data can be obtained in such a way that the signal will be highly correlated between different scalp sites and the noise would not. The electrode sites can be determined by recording data with an array of electrodes and determining the MSC between each pair of electrodes for EEG noise and SSABRs to determine which electrode pairs have high signal correlation and low noise correlation. This idea was given by Challis et al. (1991).

7.2.2 Alternate Signal Processing Methods

There are numerous extensions to AR modeling in the detection of sinusoids. A few of those techniques which may show promise for SSABR detection are discussed here. The first technique consists of modeling the **A**utocorrelation **F**unction (ACF) of the measured data with an AR model instead of the measured data itself. This method was found by Gerhardt et al. (1986) to result in better AR parameter estimates in the case of additive white Gaussian noise. Their method included discarding the zeroth lag of the ACF which theoretically would contain

all of the noise power in the case of white noise. However, with this method only estimated ACF lags are available which will introduce estimation error. A possible modification to this method is to model a different segment of the ACF where there is greater SNR. This segment is determined according to the decorrelation time of the noise (Satorious et al., 1978).

The idea of modeling the ACF was taken one step further by a number of researchers who performed Successive Correlations (SC) on the ACF. With this method, the estimated ACF has its 0^{th} lag removed (i.e. white noise removal) and the last 20% of its lags thrown out because they have increased estimation variance. This sequence now has its ACF computed. This process is then repeated many times. McGinn et al. (1983) noted that SC results in reduced estimation bias because it is a pole-preserving, SNR improving method. The SNR improvement is signal dependent with long correlation-length signals (with respect to the sampling rate) showing the most improved performance. For sinusoids, McGinn et al. (1989) found that correlating up to 4 times proved beneficial. Further, using simulations, Castanie et al. (1987) found that SC was able to extract sinusoids from noise with the SNR as low as - 20 dB. The SNR improvement of SC depends on the initial SNR, the data length and the noise spectral peaks, however, the limitations of this method include the increasingly correlated nature of the noise and the smaller number of data points for each successive correlation. Further, when the algorithm fails it still gives a sine function.

Another method, discussed by Swingler et al. (1988), is the hybrid Fourier-Burg technique in which the Burg method is used to estimate the AR parameters of the data series of interest. The AR parameters are then used to extend the data sequence in the forward and reverse directions (Swingler recommended extrapolating N points in both the forward and reverse directions for an N point data series). The Fourier transform is then applied. Swingler noted that this method had greater frequency resolution compared to the Fourier technique and reduced sensitivity to model order and less frequency estimation variance compared to the Burg method. Swingler further noted that this method provided enhanced tonal detection in some cases.

Another idea is to use a different detection statistic. Polydoros et al. (1986) used ad hoc functions based on AR parameters estimated from the measured data or the ACF of the measured data. Polydoros created 5 AR detection statistics (no claim of optimality was made) which were constant false alarm rate detectors. They were generally not equivalent with each

other or with an AR spectral peak detector. With Monte Carlo simulations, Polydoros noted that using AR parameter estimation on the ACF of the measured data resulted in better detection performance (in terms of the lowest SNR value needed to achieve a certain detection rate) than using the FFT on the time samples directly.

In terms of improving the accuracy of the AR parameter estimators used in our work two options are available. First, to alleviate the numerical instability caused by observation noise (discussed in section 5.5) in the AR parameter estimation method, Kumaresan et al. (1982b) proposed the use of **S**ingular **V**alue **D**ecomposition (SVD) which is more robust to noise provided the SNR is adequate. In SVD, the principal 2M eigenvectors (where M is the number of sinusoids) are found. The eigenvalues for these eigenvectors are retained and the remaining eigenvalues (which are smaller noise eigenvalues) are set to zero. The problem with SVD in the case of SSABRs is that the signal eigenvalues are not much different than the noise eigenvalues because of the low SNR. This problem is further exasperated by narrow bandpass filtering the data. The solution here would be to look at the eigenvectors themselves to determine which ones have a significant amount of power in the signal frequency region however in low SNR cases these signal eigenvectors are also corrupted by noise (Hu et al., 1997).

The second option deals with HOS-based AR parameter estimation. A possible enhancement was given by Tichavsky et al. (1995) who recommended that the 4th order cumulant estimate be based on lags such that τ_1 is distant from τ_2 and τ_3 is far from τ_4 because the variance of the sample estimates is largest when $\tau_1=\tau_2$ and $\tau_3=\tau_4$. Furthermore, Tichavsky stated that the 4th order cumulant slice $C_{4y}^{(N)}(0,0,0,\tau)$ (which was used in the YW4 ARPE in our work) has the 2nd largest variance among all slices and should be avoided.

Another modification for AR modeling was proposed by Tang et al. (1993) and applied to steady state VEPs (basically a sinusoidal signal). Before AR spectral estimation, narrowband pre-filtering (with a passband of 2 Hz) around the sinusoid was used to increase the accuracy of the AR parameter estimates. Tang then estimated AR parameters using forward and backward prediction filters for many short data segments and averaged all of the results together to obtain the final AR model. Tang noted that this method provided substantially more accurate estimates of sinusoid amplitude in white noise and in human EEG. For real VEPs,

Tang noted a reduction in background noise by a factor of 5-6 dB over that obtained with the FFT alone. However, we must bear in mind that steady state VEPs have higher SNRs than SSABRs.

Finally, another technique which may be used is the concept of Locally Optimum Detection which is useful in detecting weak signals in non-Gaussian noise (Orr et al., 1990; Kuehls et al., 1989). In this method, a nonlinearity is used whose functional description depends on the probability distribution of the noise. The data is then passed through the nonlinearity to reduce the amount of noise power after which conventional signal processing techniques may be applied. Using simulated data, Orr and Kuehls showed some very dramatic improvements in spectral estimation when using Locally Optimum Detection instead of FFT processing on a non-Gaussian noise corrupted sinusoid. In applying this theory to our signal detection problem, we know that the EEG noise has a nearly gaussian pdf but that it has thicker tails (Gasser, 1986). Based on this, two pdfs which may be used to characterize the EEG are the generalized Gaussian distribution (Tesei et al., 1996) and the Cauchy distribution (Kassam, 1988). However, for this method to be successful either the EEG noise pdf must be similar across test subjects or we must be able to adaptively estimate the noise pdf for each test subject during testing (which will be difficult).

APPENDIX A: OCCURRENCE OF RETROCOCHLEAR LESIONS

An epidemiology study was done to determine the number of individuals with a retrocochlear based hearing disorder. The different subject groups which were investigated were neonates, infants, school children, and adults. Unfortunately, the literature did not give clear cut information on the number of cochlear versus retrocochlear cases of hearing loss.

A.1 Studies done on Neonates, Infants and School Children

A study by Hayes (1992), was done during the period of January 1985 to May 1991 on 1281 infants in a neonatal intensive care unit with click-evoked ABRs (air and bone conduction). 1162/1281 (90%) infants were normal and the others had losses as shown in Table A.1.

Table A.1 Degree and type of hearing loss in Hayes' study.

Type of Loss	sensorineural	conductive	brainstem pathway
mild	16	50	0
moderate	1	34	3
severe	3	4	4

Tieri et al. (1988), during the period of 1979 to 1986, examined 280 cases of unilateral sensorineural hearing loss in children aged 8 months to 12 years (211 children were 6 or older). They performed topodiagnosis on 29 cases by using the Metz test, the reflex decay test and ABRs and found that 27 cases had a cochlear hearing loss and that 2 cases had a retrocochlear hearing loss. They further noted that unilateral hearing loss in children is rare. Furthermore, Dr. Hyde (personal communication, 1996) said that approximately 25 out of 1000 infants with hearing loss have a retrocochlear based hearing loss and have normal otoacoustic emissions.

Parving (1983) looked at children aged 2 to 12 years old (median age was 8 years) born between January 1, 1970 and December 31, 1979 in Copenhagen, Denmark. The prevalence of permanent hearing loss in this region was noted at 1.4%. In this study 117 children were identified with a hearing loss; 106 (91%) of which had a sensorineural hearing loss. To evaluate the site of lesion in these sensorineural cases, impedance audiometry and cochlear microphonic testing was used. Results showed that 32 (30%) children had cochlear lesions, 21 (20%) had retrocochlear lesions and 52 (50%) had no distinction between cochlear or retrocochlear hearing involvement due to a lack of data.

Another study conducted by Arlsen et al. (1991) during 1973-1990 in Italy looked at 787 cases of children with sensorineural hearing loss. The study is not meant to be epidemiological, however, the rate of CNS involvement in these hearing loss cases was 31.9% (251/787). This however does not indicate how many cases involved only a retrocochlear based hearing loss.

A.2 Studies of Hearing Loss in Adult Populations

Sulkowski et al., (1982) looked at industrial workers during the period of 1976 to 1980 in Poland. The study looked at 697 cases with suspected occupational aetiology. Results showed that 210 (30%) of the cases had a cochlear noise induced hearing loss with 73 (34.8%) of those cases involving a retrocochlear impairment. Another study by Rizzo et al., (1991) looked at presbycusis (hearing loss due to aging) in 18 adult males. Rizzo used ABRs for site of lesion testing and found cochlear hearing loss in 9 males and retrocochlear in the other 9.

A.3 Cases of Sensorineural Hearing Loss With Retrocochlear Lesions

Wu et al. (1995) looked at 495 cases of sensorineural hearing loss investigated on suspicion of an acoustic pathway lesion. The subjects were 5 to 88 years old. Using MRI, 17% of cases had lesions affecting the retrocochlear auditory pathway. Another study (Hendrix et al., 1990), during the period of August 1985 to July 1988, looked at 225 consecutive cases of asymmetric sensorineural hearing loss. The average age of the subjects was 50.4 years (4 subjects were under the age of 20). With the use of ABR, CT and MRI tests, it was found that 194 cases (86%) had a cochlear site of lesion and 31 cases (14%) had retrocochlear lesions.

A.4 Conclusion and Recommendations

This appendix demonstrates that the prevalence of retrocochlear based hearing disorders in various subject groups is too large to ignore, however, the results of section A.3 may be inflated since the people in the study were referred to the researchers on the belief that their hearing loss was due to a retrocochlear lesion. Nonetheless, the data presented in this appendix substantiates the need for a more thorough audiometric test, one which also examines the retrocochlear components of the auditory system, such as the SSABR.

APPENDIX B: EEG AMPLITUDE PROBABILITY DISTRIBUTION

This appendix discusses the results of attempts in the literature to characterize the EEG amplitude probability distribution. In 1969, Elul measured the probability distribution of scalp EEG for one adult during an idle state and while performing mental arithmetic. The recorded EEG was filtered from 0.3 to 70 Hz and sampled at 200 Hz. The test used to check for gaussianity was the Chi-Square goodness of fit test. The EEG data segments were separated into 2 second segments to avoid nonstationarity. Based on a large sample, Elul found that in the idle state, the EEG had a gaussian distribution 66% of the time meanwhile during mental task performance the EEG was gaussian 32% of the time. From this data, Elul postulated that the gross EEG is gaussian when the generators are desynchronized but when the relation between the individual neurons is intensified, as in sleep or certain behavioral situations, then the EEG probability distribution is modified. Later, Elul et al. (1975) performed a study on children using the same test setup, recording procedures and analysis method as in his 1969 study. For young normal children, Elul saw a non-gaussian EEG distribution, however, with older children Elul noticed the EEG became more gaussian in nature (80% of trials were gaussian). However, the sampling rate value weakened the independence of data samples and this implies that the probability of the EEG being gaussian is greater than Elul found it to be (Siegel, 1981).

Some studies from Dumermuth in the late sixties and early seventies found that deviations from gaussianity occurred in a majority of twins in the waking state. Using the measures of skewness and kurtosis, Dumermuth also saw this effect in sleep EEG when analyzing 40 second segments with the EEG lowpass filtered at 40 Hz and sampled at 102.4 Hz (Dumermuth et al., 1975). McEwen et al. (1975) also conducted tests using the **K**olmogorov-**S**mirnov (KS) test on 30 subjects in which two 64 second trials were recorded before and during anesthesia. The EEG was bandpass filtered from 0.54 Hz to 30 Hz and different sampling rates of 64, 128 and 256 Hz were used to see the effect of sample dependence on the statistical tests. They concluded that it was best to use a sampling rate as little above the Nyquist rate as possible. They also noted that the percentage of gaussian and wide sense stationary (WSS) EEG segments had a strong dependence on segment length. The results were also strongly similar between the different EEG channels used. However, they noted that the occipital EEG activity appeared to be more consistently gaussian and stationary than the frontal

EEG activity. For 2 second segments, at least 90% of segments from all test derivations were stationary and gaussian and for 8 second segments this percentage dropped to about 65%.

Persson (1974) commented on previous statistical tests for the gaussianity of EEG data. He stated that for small sample sizes the KS test is more powerful than the Chi-Square test, however, for these tests the data must be stationary and there must be mutual independence between samples. Persson also showed that for the KS test, the sampling rate had to be 4 Hz in order to get independent samples. Further Persson showed that with sampling rates greater than 8 Hz there were many rejections for the EEG being gaussian because the samples were dependent. Due to the EEG autocorrelation shape, Persson recommended that a sampling rate less than 20 Hz (and sometimes much lower) was needed but this leads to weaker statistical tests because of smaller sample sizes.

Sugimoto et al. (1978) did a statistical analysis of the EEG of 2 sleeping adults. They used 5 second data segments and performed stationarity and normality tests based on non-parametric statistics (run test and slide test) and the parametric Fisher's K statistics (based on 3^{rd} and 4^{th} order moments). EEG data was bandpass filtered from 2 to 50 Hz and sampled at 100 and 500 Hz. The significance level for the statistical tests was 5%. They found that as segment length increased, the nonstationarity also increased in each stage of sleep. They also saw that 50% of 2.5 to 5 second segment lengths showed both stationarity and normality. They concluded that a decreased length of data implies an increased degree of stationarity and normality for the EEG in sleep.

In 1993, Galbraith conducted experiments to verify Elul's 1969 findings. The study used skewness and kurtosis to determine the gaussianity of the EEG. The data was bandpass filtered from 0.3 to 20 Hz and sampled at 200 Hz. They used a frontal-occipital bipolar derivation for EEG recording. They examined 40 digitized sweeps, 3.2 seconds in duration, and concluded that no difference in EEG probability distribution between idle states and states of mental arithmetic could be found. Meanwhile, Gasser (1986) used skewness and kurtosis on every 4^{th} data sample recorded to assess the normality of EEG noise. The data was lowpass filtered at 70 Hz and sampled at 408.5 Hz. The data was from pre- and post-stimulus intervals of length 313.5 ms in visual evoked potential studies. Skewness indicated a symmetric distribution with thick tails. Kurtosis also revealed slight deviations from normality.

APPENDIX C: EXPERIMENTAL SETUP AND INSTRUMENTATION

A diagram of the experimental setup is shown in Figure C.1. The LabView software package is used for stimulus generation and data acquisition. The stimulus is then converted to analog form by a Tahiti D/A converter and sent to a probe phone where the electrical stimulus is transduced to an acoustical stimulus. The subject, who is seated in an electromagnetically, shielded sound proof room, listens to this stimulus and his/her response is measured by skin electrodes. The response is then amplified, bandpass filtered and digitized by an A/D converter. The data is then post-processed using the Matlab software package.

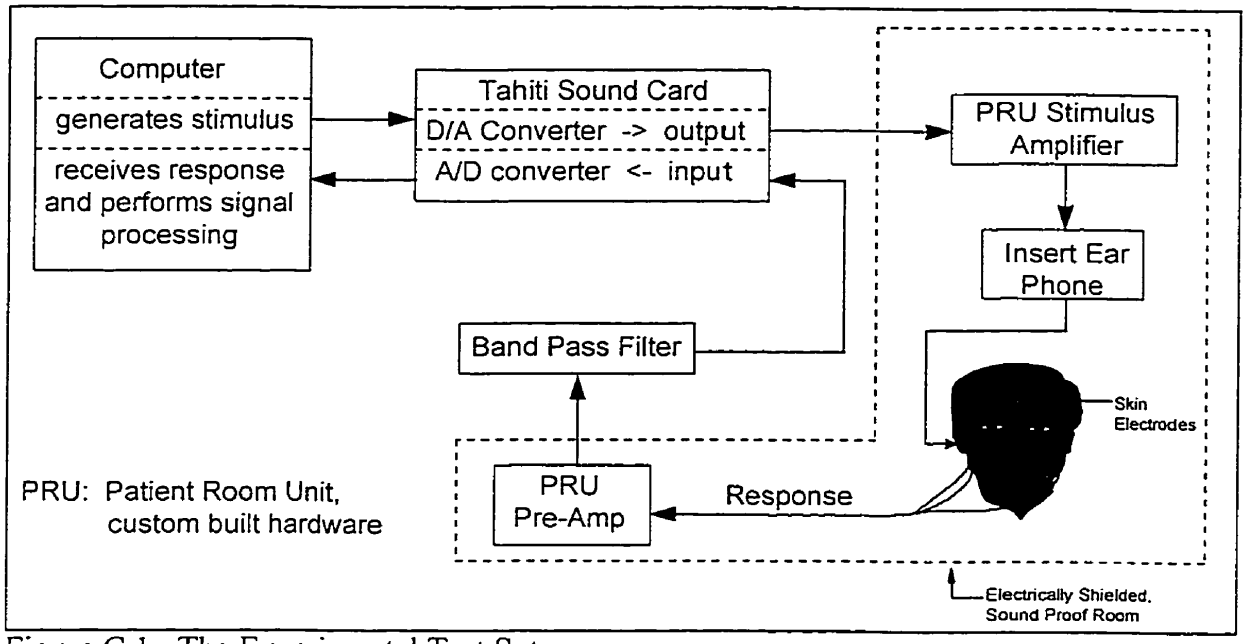

Figure C.1. The Experimental Test Setup.

The computer used was a 486, 100 MHz, IBM clone. Test recordings of four minutes in duration were performed with a sampling rate of 22,050 Hz which required 20 Megs of hard disk space. The A/D and D/A converters are contained in a Tahiti sound card which is made by Turtle Beach systems. This 16 bit sound card has 2 output channels and 2 input channels. The A/D and D/A rates can be either 11.025, 22.050 or 44.100 kHz. The data buffer size is also programmable with a maximum size of 16,384 data points when 2 data channels are being used. For our experiments, the A/D and D/A rates were both 22,050 Hz and the buffer size was 2205 points. Therefore each buffer contained 0.1 seconds of data. In postprocessing, the recorded data was downsampled so that the new sampling frequency was 490 Hz and the new

buffer size was 49 points. The buffer size allowed for an integer number of response cycles in each buffer to avoid spectral leakage (Kulik, 1995).

For the sound card, the total harmonic distortion is less than 0.005%, the intermodulation distortion is 0.01% and the frequency response is from DC to 19 kHz. The maximum voltage that the card can tolerate is 0.6 Volts RMS. Thus the minimum signal that can be resolved with this sound card is 2.6 nV_{pp} (where $_{pp}$ stands for peak to peak) which is adequate for SSABR measurement. The value is 2.6 nV_{pp} because the maximum sinusoidal amplitude that the sound card can receive is:

$$S_{MAX} = 0.6 \cdot (2)^{1/2} \cdot 2 \approx 1.697\ V_{pp} \qquad (C.1)$$

and so the minimum resolvable sinusoidal amplitude is:

$$S_{MIN} = 1.697/(2^{16}) \approx 25.9\ \mu V_{pp} \qquad (C.2)$$

and due to the gain of the differential amplifier (G = 10,000 V/V):

$$S_{MIN} = 25.9\ \mu V_{pp} / 10{,}000 \approx 2.6\ nV_{pp} \qquad (C.3)$$

The stimulus and response amplifiers are custom built hardware. The stimulus amplifier has a variable gain from -95.5 to 31.5 dB and the response amplifier is a differential amplifier with a constant gain of 80 dB (10,000 V/V). For the differential amplifier, the CMRR is 105 dB, the common mode impedance is 200 GΩ, the input impedance is typically 60 MΩ (the minimum specified value is 20 MΩ) and the input noise is 0.6 μV_{pp}. The bandpass filter was a Krohn-Hite (model 3700) filter. The filter rolloff was 24 dB/octave. The usual bandpass filter settings were from 10 to 100 Hz.

The acoustic probe phone used for stimulus delivery fits directly into the ear canal. A diagram of the acoustic probe is shown in Figure C.2. The acoustic probe was designed by Poul Madsen Medical Devices originally for Distortion Product OAE measurement, however, for our work only one speaker was used for AM wave presentation. The probe phone speakers were Knowles EF 1933 and the microphone (which was not used) was Knowles EM4068. The microphone and speakers both had a flat transfer function up to approximately 4 kHz.

To calibrate the acoustic stimulus used in testing, a Zwislocki coupler was used which simulates an average human ear. The stimulus was delivered to the Zwislocki coupler and measured with a microphone (at the other end of the coupler) which was situated in such a way that it resembles the eardrum. A Bruel and Kjaer microphone power supply (Type 2804) was

used in conjunction with a Bruel and Kjaer (Type 2231) Modular Precision Sound Level meter (BZ 7110) to measure the stimulus intensity delivered to the microphone. The measurements were taken with a Bruel and Kjaer octave band filter (Type 1625) to obtain more precise measurements of the stimulus intensity. The stimulus amplifiers on the PRU were then adjusted to get the correct stimulus intensity. The electrical stimulus (coming out of the Tahiti board) also had to be varied in amplitude to reduce the amount of acoustic distortion in the stimulus.

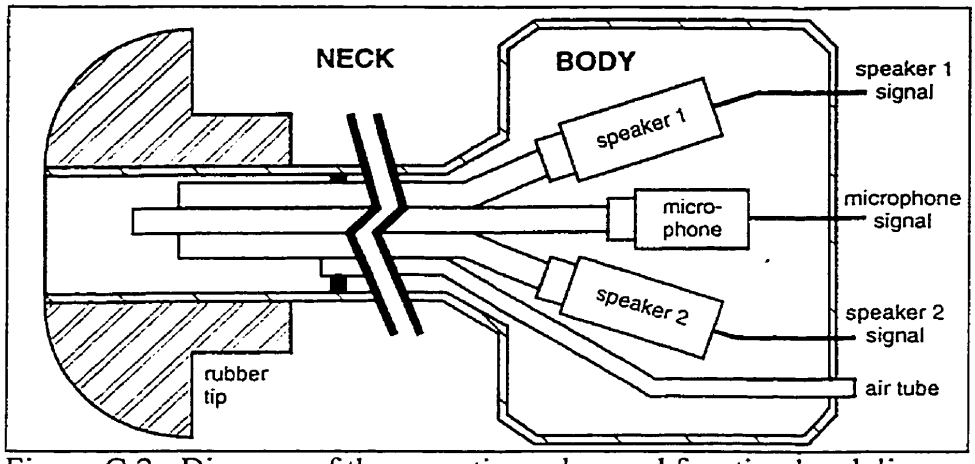

Figure C.2. Diagram of the acoustic probe used for stimulus delivery (adapted from Kulik, 1995).

The electrodes used were Nutab™, Silver Silver-Chloride resting ECG electrodes which have a low impedance for DC and low frequency potentials and exhibit less electrical noise than other electrodes (Neuman, 1995). The bottleneck in EP recording is the impedance at the skin-electrode junction which must be low (less than 5 kΩ). To achieve this, the skin was abraded and an electrode paste (Omni Prep) was also used to prepare the skin (adequate skin preparation can result in a skin-electrode impedance of 1 kΩ). To get an indication that the resulting skin-electrode impedance was acceptable, the DC offset between each pair of electrodes was measured. A low value (around 10 mV) which was approximately the same across all electrode pairs was acceptable (personal communication with Prof. Madsen, 1997). The active recording electrodes were placed on the subject's forehead and ipsilateral mastoid and the reference electrode was placed on the contralateral mastoid.

APPENDIX D: WTA ALGORITHM COMPUTATIONAL COMPLEXITY

For each WTA method, there will be N sweeps each containing K points. Each subtraction will be considered as 1 multiplication and 1 addition. For the V^{-1} method, the variance calculation will require N divisions and NK additions for the calculation of the mean and then 2NK additions, 2NK multiplications and N divisions for variance calculation. To compute the weights we need N additions and 2N divisions. Thus in total approximately 2NK multiplications, 3NK additions and 4N divisions are needed.

For the ME method, to compute the matrix FF^T requires KN^2 multiplications and KN^2 additions. Then there is the matrix inversion of FF^T which is NxN in size. Multiplication of $(FF^T)^{-1}$ by e requires N^2 multiplications and N^2 additions. The denominator requires N additions and N multiplications. The final step in the weight calculation involves N divisions. Thus in total there are approximately $KN^2 + N$ multiplications, $KN^2 + N$ additions, N divisions and the inversion of an NxN matrix.

For the Gerull method, there are NK additions and NK/2 multiplications to get the pair sums and pair differences. Whitening occurs N/2 times. Whitening involves the Yule-Walker method which requires K^2 additions and K^2 multiplications (Marple, 1987). The filtering involved in whitening has roughly (K-2p)p additions and (K-2p)p multiplications where p is the order of the whitening filter. The model order selection method involves p additions, p multiplications, 2p log operations and p divisions. Thus whitening involves $NK^2/2 + (N/2)(K-2p)p + Np/2$ multiplications, $NK^2/2 + N/2(K-2p)p + Np/2$ additions, pN log operations and Np/2 divisions. Assuming that p is K/7 gives approximately $NK^2/2 + KN/14$ multiplications, $NK^2/2 + KN/14$ additions, NK/7 log operations and NK/14 divisions. There are then N/2 variance estimates which results in NK multiplications, 3NK/2 additions and 2N divisions. To finally calculate the weights based on the inverse of these variance estimates there is N/2 additions and N divisions. Thus in total there is approximately $K^2N/2 + (11/7)NK$ multiplications, $K^2N/2 + (18/7)KN$ additions, NK/7 log operations and $N(3 + K/14)$ divisions.

For the VNF method, to calculate the N D_i requires 2N(N-1)K multiplications, 2N(N-1)K additions and N(N-1) square roots. The weight calculation from the D_i requires N multiplications, 2N divisions and N additions. Thus in total there are approximately $2N^2K + N$ multiplications, $2N^2K + N$ additions, 2N divisions and N^2 square roots.

APPENDIX E: RECEIVER OPERATOR CHARACTERISTIC (ROC) ANALYSIS

ROC analysis measures a detection algorithm's detection accuracy and can be used for performance comparison of diverse detection systems. This comparison is made in terms of comparing the probability of correctly identifying data which contains a signal and data which contains only noise versus the probability of misidentifying noise as a signal and a signal as noise. This discrimination is not done perfectly because values calculated based on noise only data and signal plus noise data may overlap (Swets, 1988). This concept is seen in Figure E.1 which shows the **p**robability **d**istribution **f**unctions (pdfs) of the energy level measured at a certain signal frequency (i.e. FFT bin) given observed data consisting of pure noise and data consisting of signal plus noise. The overlap of the pdfs gives rise to false detection outcomes.

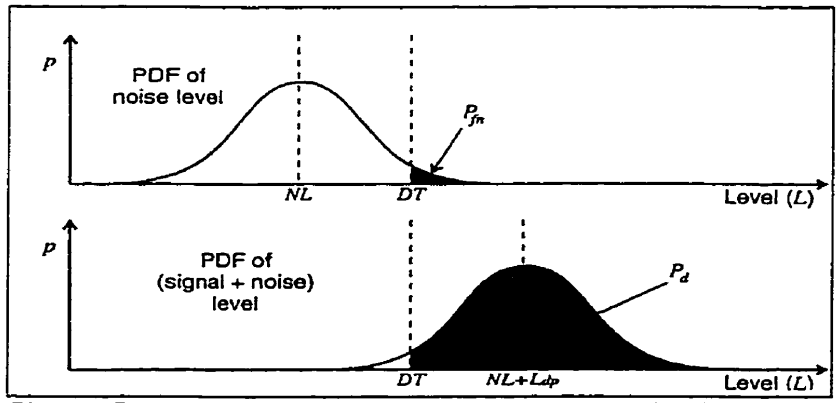

Figure E.1. PDFs of the energy level at a certain signal frequency given data consisting of pure noise (top curve) or signal plus noise (bottom curve) (adapted from Kulik, 1995).

ROC curves are generated by considering the possible detector outputs (signal detected or no signal detected) and the actual observed data (pure noise or signal plus noise). Pooling these events together results in the two-by-two contingency table shown in Table E.1. The two correct detection outcomes are hits (detecting a signal when a signal is present) and correct rejections (not detecting a signal when only noise is present). The two incorrect detection outcomes are misses (not detecting a signal when a signal is present) and false-alarms (detecting a signal when only noise is present). If the proportions of these outcomes are considered, then only two proportions (one from each row in Table E.1) are needed to quantify

detection algorithm performance[16]. Conventionally, the proportions of Hits (**Hit Rate**(HR)) and False Alarms (**False Alarm Rate** (FAR)) are used (Metz, 1986) to quantify detector performance.

Table E.1. Possible Detection Outcomes (adapted from Swets, 1988).

Data Type \ Detector Output	Signal Detected (D)	Signal not Detected (ND)
Signal Present (s)	Hit P(D|s)	Miss P(ND|s)
Signal not present (ns)	False Alarm P(D|ns)	Correct Rejection P(ND|ns)

P(D|s) + P(ND|s) = 1; P(D|ns) + P(ND|ns) = 1

The detector output itself is based on a decision criterion which is chosen depending on the prior probabilities of signals occurring and only pure noise occurring. If the signal has a high probability of occurring, and is easily discernible from noise, then the decision criterion can be lenient which will give a high HR. Conversely, a strict decision criterion is needed when the signal is not as likely to occur or the accompanying noise is very disruptive. This is to maintain a low FAR. Varying this decision criterion results in various pairs of HRs and FARs which are then plotted to obtain the ROC curve (see Figure E.2). On the ROC curve, movement from the lower left corner to the upper right corner represents a change from a strict decision threshold to a more lenient one which results in increases in both the HR and FAR. The major diagonal represents a detector which can not discern noise from signal in which case the curve is characterized by equal HRs and FARs for each value of the decision criterion.

To generate the ROC curve, values of HR and FAR are derived from data consisting of true negative cases (pure noise data) and true positive cases (data containing a signal plus noise) using the following equations:

$$HR = \frac{N_H}{N_P} \tag{E.1}$$

$$FAR = \frac{N_{fa}}{N_N} \tag{E.2}$$

where: N_P is the number of true positive test cases

[16] Whenever a signal is present, the detection algorithm can declare that a signal is detected or a signal is not detected and so the proportion of misses is the complement of the proportion of hits. Thus, the proportions in each row of Table E.1 add to one.

N_N is the number of true negative test cases
N_H is the number of Hits
N_{fa} is the number of False Alarms

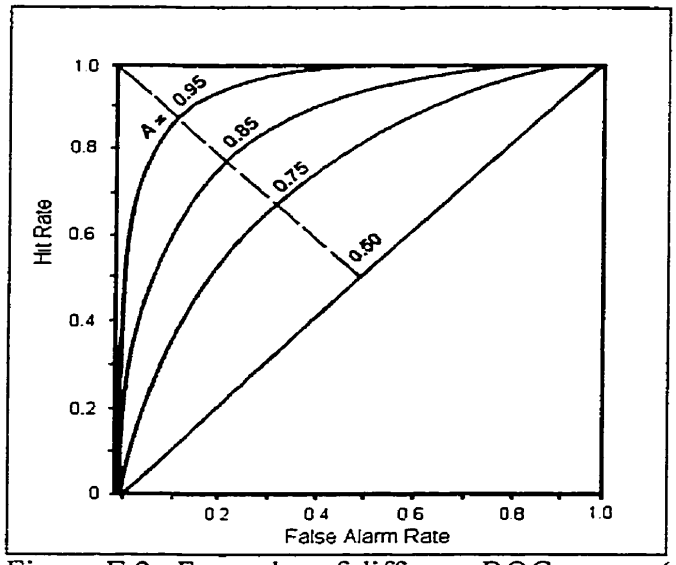

Figure E.2. Examples of different ROC curves (adapted from Swets, 1988).

For ROC data collection, care must be taken to avoid wrongly labeling the test cases as true positive data or true negative data. In addition, classifying the data as either positive or negative should be made independent of the detection algorithm. Finally, the data should fairly reflect the population of cases to which the detector will be applied to in practice.

E.1 ROC Performance Indices

One very popular ROC performance index is the area under the ROC curve (Swets, 1988). ROC Area (RA) summarizes detector performance across all decision criterion values. An ROC curve which is higher up and to the left indicates a higher discrimination capacity (because of larger HRs at each FAR value) which translates to a larger RA (Metz, 1986). RAs vary from 50% (random performance) to 100% (perfect detection). Another performance index is the HR(FAR) which is the HR at a particular FAR (usually the FAR is 1 or 5%). In this case it is highly desirable to have a detector with a high HR at the FAR of interest. In addition, this measure is valuable because ROC curves may cross and have the same RA yet be different curves with different HRs at all FARs except one point (Metz, 1989).

E.2 Statistical Analysis of ROC curves

The ROC performance indices are sound but require qualification because the test data that they are based on is of unsure quality[17]. Thus, when comparing the performance of two or more detectors, a statistical analysis on a sufficiently large data set must be done to ensure that measured differences are real and not due to sampling effects. Suggested data set sizes range from 50 (Hanley et al., 1983) to 100 (Metz, 1986). Larger data sets will reduce the variance of the ROC curve estimate (which is inversely proportional to data set size) and demonstrate more subtle differences in detection algorithm performance (Metz, 1978). The statistical analysis then consists of obtaining an empirical ROC curve, deciding on a parametric or nonparametric analysis technique, computing the performance index, estimating the ROC performance index variability due to case sampling and determining the statistical significance of the difference between performance indices of different detectors (Swets, 1988).

In practice, the parametric and non-parametric statistical analysis techniques result in very similar performance (Centor et al., 1985; Hajian-Tilaki et al., 1997) so choosing between them is a matter of convenience. For this thesis work the non-parametric method was chosen. Non-parametric ROC analysis involves estimating the RA by connecting successive ROC curve points with straight lines and using the trapezoidal rule. This approach is mathematically related to the Wilcoxon statistical test[18] and so the statistical properties of the Wilcoxon test are used to predict the statistical properties of the RA (Hanley et al., 1982). The rest of the section will describe the statistical tests employed in our ROC study.

The Standard Error (SE) for RA estimation (SE(RA)) was obtained with the Wilcoxon method which depends on two distribution specific quantities (Q_1 and Q_2) when the HR is greater than the FAR. The quantity Q_1 is the probability that two true positive data cases are given higher detector output values than a randomly chosen true negative data case. The quantity Q_2 is the probability that a randomly chosen true positive data case will be given a higher detector output value than two randomly chosen true negative data cases. The SE(RA) is then given by:

---

[17] There are a limited number of test cases and there is variability in reproducing ROC plots with different data sets.

[18] This is true for ROC curves created from continuous data or from rating data.

$$SE(RA) = \sqrt{\frac{RA(1-RA) + (N_P - 1)(Q_1 - RA^2) + (N_N - 1)(Q_2 - RA^2)}{N_P N_N}} \qquad (E.3)$$

The quantities Q_1 and Q_2 are complex functions of the underlying distributions for the detector outputs for the true positive and true negative data cases, however, equation E.3 is heavily determined by RA and only slightly influenced by the underlying distributions (Hanley et al., 1982). Hanley further noted that for RAs larger than 0.7, the negative exponential model yielded the most conservative SEs as well as convenient closed form expressions for Q_1 and Q_2:

$$Q_1 = \frac{RA}{2 - RA} \qquad (E.4)$$

$$Q_2 = \frac{2RA^2}{1 + RA} \qquad (E.5)$$

Therefore, the exponential model is used to give a conservative estimate of SE(RA).

For RA comparison, if the estimates of RA are roughly normally distributed, then one can determine the statistical significance of the difference RA_1 - RA_2 (for two detectors #1 and #2) by using a table of standard normal deviates (z) to find the probability that (Metz, 1989):

$$z > \frac{RA_1 - RA_2}{SE(RA_1 - RA_2)} \qquad (E.6)$$

This probability or p-level[19] is an estimate of the degree to which the observed RA difference is true (the higher the p level is the less we can believe that the observed relation is true). A p-level of 5% is 'border-line' statistically significant meanwhile p levels of 1% are statistically significant and p levels of 0.5 or 0.1% are highly statistically significant (Glantz, 1987).

The calculation of the denominator of equation E.6 depends on whether each ROC curve is estimated from different data cases (unpaired test) or from the same data cases (paired test). In our ROC study, each detector operates on the same data set so the paired test applies. In the paired test, ROC curve estimates (from different detectors) will tend to vary above and below their means together because they are applied to the same data cases. Thus, RA_1 - RA_2 will vary less than it would in the unpaired test case. The SE is then given by:

$$SE(RA1 - RA2) = \sqrt{SE^2(RA_1) + SE^2(RA_2) - 2rSE(RA_1)SE(RA_2)} \qquad (E.7)$$

where r is the correlation between RA_1 and RA_2

[19] In the vast majority of research undertakings, two-tailed tests are called for (Vida, 1993).

Now the denominator of E.6 is reduced (due to the correlation between RA_1 and RA_2) so that the statistical significance of the RA difference is increased which will increase the statistical power of a detector comparison study (Metz, 1989).

To calculate the correlation (r) between RA_1 and RA_2, two intermediate correlations must be calculated: the correlation (r_N) between detector output values for true negative data cases and the correlation (r_p) between detector output values for true positive data cases. The correlation can be calculated by the Pearson product-moment correlation method[20], which for two time series x and y is given by (Glantz, 1987):

$$r = \frac{\sum_{j=1}^{n} (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{j=1}^{n} (x_j - \bar{x})^2 \sum_{j=1}^{n} (y_j - \bar{y})^2}} \qquad \text{(E.9)}$$

The two correlation values, r_N and r_P, can then be converted to the correlation between RA_1 and RA_2 via a table provided by Hanley et al. (1983).

In terms of the other ROC performance index (HR(FAR)), an equation can be used to obtain the 95% confidence bounds of the proportion HR(FAR). The Matlab code for these upper and lower bounds is shown in Figure E.3 (Diamond, 1989). The FAR used in our ROC studies for the HR(FAR) ROC performance index is 5% (i.e. (HR(0.05)).

```
C = 1.96;       A = 0.025;
if (X == 0);    U = 1 - exp(log10(A)/N);     L = 0;
elseif (X == N);  U = 1;   L = exp(log10(A)/N);
else;   R = X+A-1/2;       L = (R+C*C/2-C*sqr(R*(1-R/N)+C*C/4))/(N+C*C);
        R = X-A+1/2;       U = (R+C*C/2+C*sqr(R*(1-R/N)+C*C/4))/(N+C*C);
end;
                    %where  U  is the Upper 95% confidence bound
                    %       L  is the Lower 95% confidence bound
                    %       X  is the numerator of the proportion
                    %       N  is the denominator of the proportion
```

Figure E.3. Matlab Code to get upper and lower 95% confidence levels for a proportion (adapted from Diamond (1989)).

[20] The Pearson product-moment correlation function determines to which extent two variables are linearly related to one another.

APPENDIX F: SWEEP LENGTH FOR SYNCHRONOUS TIME AVERAGING

In order to determine the appropriate length of each sweep used in synchronous time averaging, the Autocorrelation Function (ACF) was estimated for groups of data 4.8 seconds in duration. This was done after notch filtering the data using a first order Butterworth IIR notch filter, in the non-phase shift configuration (Antoniou, 1993), with a bandstop region of 58 to 62 Hz and a 40 dB notch. The time lag for which the noise was effectively not correlated was determined using a threshold level whereby the autocorrelation value at the zeroth lag was reduced by a factor of 10 dB. This was done on pure noise data recorded from 3 young males (24 to 33 years of age) who had no history of hearing impairment. For each subject, the data segments were taken 15 seconds apart. Figure F.1 shows an example of the noise ACF. Table F.1 contains the lag values at which negligible noise correlation was observed. The means and standard deviations of these lag values across subjects indicates that a sweep length of 0.3 seconds is sufficient for reducing the correlation between successive sweeps.

TABLE F.1. ACF Lag values at which the EEG noise is uncorrelated.

Data Segment Location (seconds)	Lag value at which noise is uncorrelated (seconds)		
	Subject 1	Subject 2	Subject 3
0.4 to 5.2	0.2265	0.195	0.1653
19.6 to 24.4	0.3327	0.1251	0.3853
38.8 to 43.6	0.1253	0.4353	0.1251
58 to 62.8	0.218	0.1041	0.2543
77.2 to 82	0.2833	0.0633	0.1796
Mean	0.2372	0.1846	0.2219
Std Dev	0.0779	0.1481	0.1026

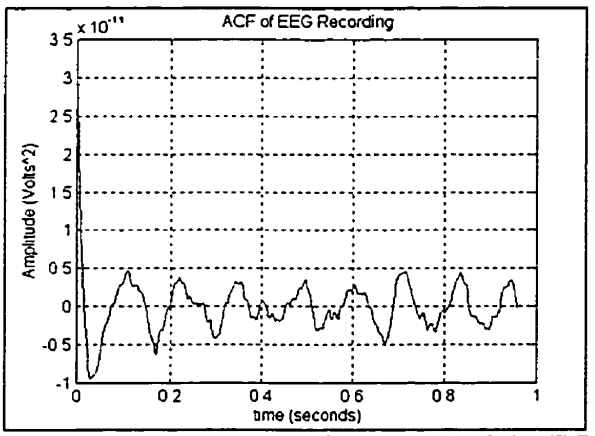

Figure F.1. Example of the ACF of the EEG noise.

APPENDIX G: BANDPASS FILTER PARAMETERS

In order to design the bandpass filter, estimates of typical SSABR spectrums after WTA were determined to yield the specifications for the filter's frequency domain characteristics. The no response data (used to avoid any biasing at the response frequency) was recorded from three healthy 24 to 25 year old male subjects under quiet recording conditions. Since the line noise was so large (2-4 μV; very stable at 60 Hz), a first order Butterworth IIR notch filter, in the non-phase shift configuration (Antoniou, 1993), with a bandstop region of 58 to 62 Hz and a 40 dB notch was used. WTA (inverse variance technique) was then performed on 4.8 second data blocks using 0.3 second sweeps. An FFT was done on this averaged data. To get a more accurate estimate groups of FFTs were averaged. Representative results from one subject are shown in Figure G.1 (no FFT averaging) and Figure G.2 (averaging two FFTs).

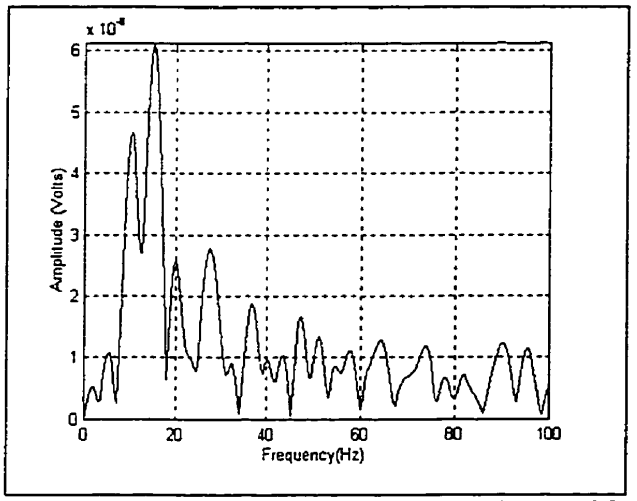

Figure G.1. Spectral Estimate based on 1 FFT.

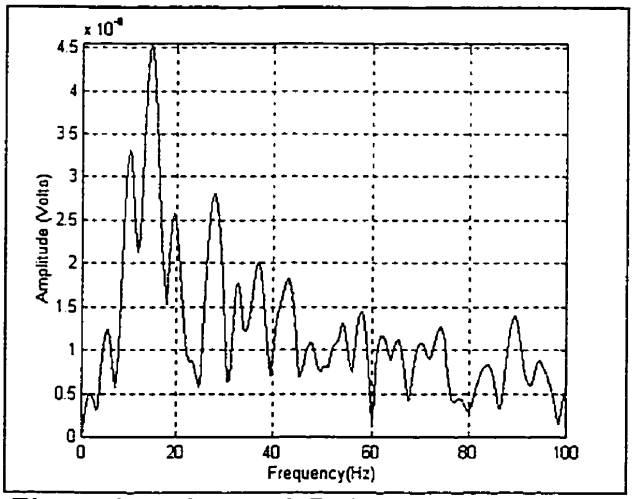

Figure G.2. Spectral Estimate based on spectral average of 2 FFTs.

To reduce the transients at the beginning and the end of the filter output and to allow for an arbitrary choice of filter length, an FIR filtering technique based from Chen et al. (1996) was used. It was equivalent to taking the convolution of the data sequence and the filter impulse response and then truncating the convolution result at both ends. The filter output was taken from (q+1)/2 to [(p+q-1) - (q+1)/2] where p is the length of the data sequence and q is the length of the filter impulse response.

The next step was choosing the filter passband. A passband of 10 Hz was chosen (for the 40 Hz SSABR the passband was 35 to 45 Hz) because this was narrow enough to remove as

much noise as possible but not so narrow that the noise would end up looking like a signal (as explained in section 5.7.1). The next step was to choose appropriate attenuation values for the filter and this was determined from data recorded from the 3 subjects (see Figures G.1 and G.2 for examples). The data showed that the filter needs to have an attenuation of at least 20 dB on the low frequency side (i.e. near 10 Hz) to make the energy in the 10 Hz region similar to the energy near the 40 Hz region. In order to get better AR spectral estimates and to reduce the biasing effects of nearby frequencies the filter attenuation was chosen to be at least 40 dB on either side of the passband.

The final step was designing the filter to meet these specifications. The Window FIR filter design method was chosen with a Kaiser window which provides a better tradeoff between mainlobe width and sidelobe amplitude (Oppenheim et al., 1989). A filter order of 100 and a β value of 4 met the design criteria specified above. The frequency response of the filter is shown in Figure G.3. The filter was also judged in terms of its rise time and percentage overshoot to a sinusoidal input with a frequency of 40 Hz and a peak amplitude of 300 nV. The rise time was defined as the time that the output took to reach 90% of its peak input value and the percentage overshoot was the peak of the output minus the peak of the input all divided by the peak of the input (Williams, 1988). The filter sinusoidal response is shown in Figure G.4. The rise time was 57 milliseconds and there was no overshoot.

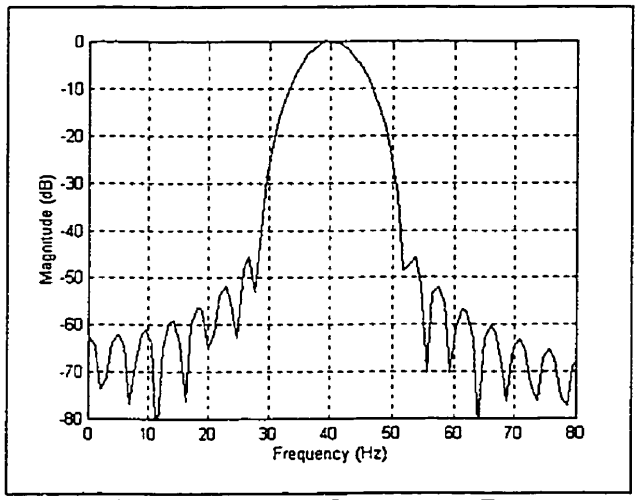

Figure G.3. FIR Filter Frequency Response.

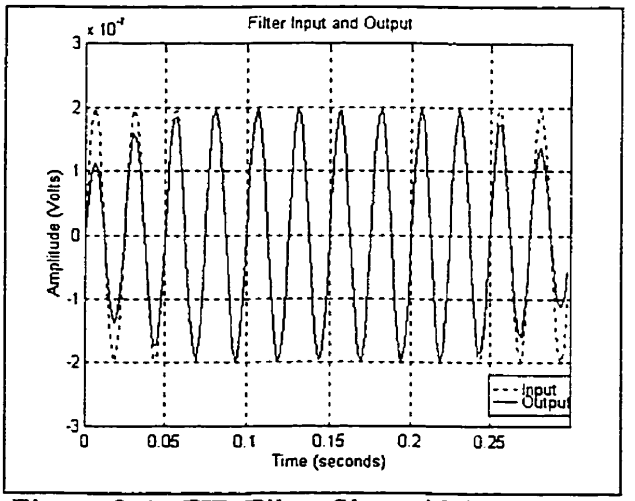

Figure G.4. FIR Filter Sinusoidal Response.

APPENDIX H: INSTRUMENTATION NOISE

The first type of instrumentation noise considered will be that seen in the stimulus. This noise is due to D/A conversion, stimulus amplifier noise and acoustic probe noise. To remove the digitization noise, the stimulus can be lowpass filtered at 10 kHz (this was provided by the frequency characteristics of the speaker in the probe phone). To determine stimulus quality, a syringe (0.3 cm length cavity) was used as an artificial ear. A recording was made of the stimulus by using the microphone in the probe phone. An FFT of the recording was then made and distortion was seen at the harmonics of the AM wave and at $f_c \pm af_m$ where a is an integer. These distortion peaks were less than 40 dB below the stimulus levels which is an acceptable limit (personal communication with Prof. Madsen, 1997). Also the amount of distortion was larger for the 70 dB SPL stimulus (Figure H.1) than the 50 dB SPL stimulus (Figure H.2).

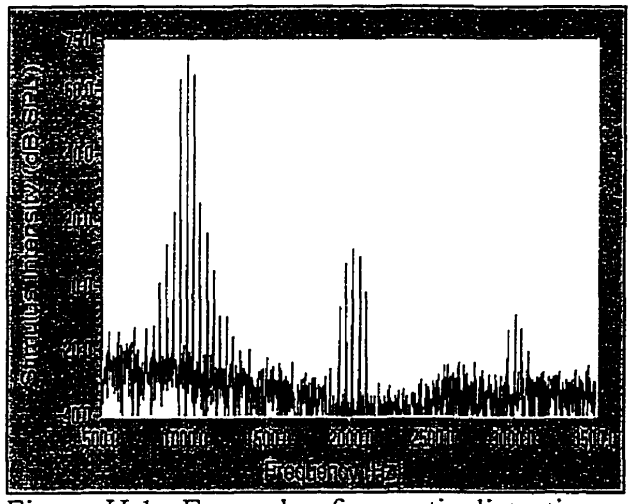

Figure H.1. Example of acoustic distortion in the 70 dB SPL SSABR stimulus.

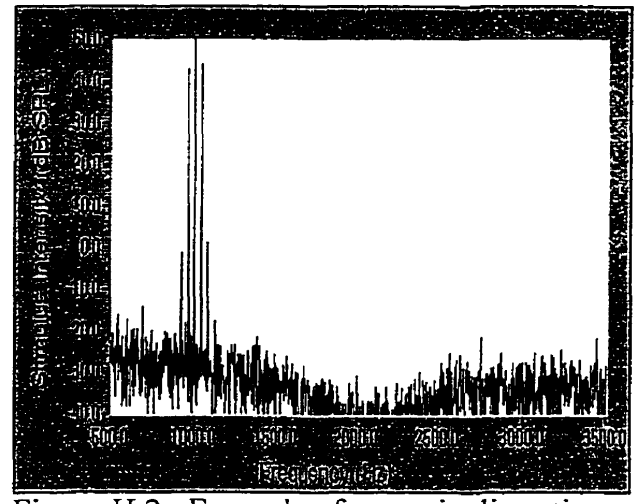

Figure H.2. Example of acoustic distortion in the 50 dB SPL SSABR stimulus.

The other instrumentation noise is from the equipment in the response path of the experimental setup. This noise will be added to the recording and is due to response amplifier noise, bandpass filter noise, electromagnetic noise and quantization noise. Bartoli et al. (1983) indicated that this noise could be considered as white noise with an RMS value of approximately 10 μV. For our recording equipment, an analytical examination of the quantization noise will be shown. Experimental measurements will also be shown to determine the total amount of instrumentation noise and to verify that no artificial SSABR was being generated by the instrumentation.

The quantization noise can be considered as uncorrelated, additive noise and is related to the number of bits (Nb) in the A/D converter and the maximum signal amplitude (V_{ref}) that the A/D converter can tolerate. With our equipment V_{ref} is 0.849 V_p, and the quantization error range is given by (Johns et al., 1994):

$$-1/2 \cdot V_{LSB} \leq V_x \leq 1/2 \cdot V_{LSB} \quad \text{(H.1)}$$

$$\text{where: } V_{LSB} = V_{ref}/2^{Nb} \quad \text{(H.2)}$$

If the input signal being quantized varies rapidly such that the quantization error is a uniformly distributed random variable then the RMS quantization noise voltage is given by:

$$V_{Q(rms)} = V_{LSB}/(12)^{1/2} \quad \text{(H.3)}$$

$$= (0.849/2^{16})/(12)^{1/2} \quad \text{(H.4)}$$

$$\approx 3.74\ \mu V \quad \text{(H.5)}$$

Thus, if the minimum SSABR amplitude we are trying to measure is 20 nV and the amplifier gain is 10,000 V/V then the SNR between the SSABR and the quantization noise is:

$$SNR = 20 \cdot \log_{10}(SSABR_{amp}/V_{Q(rms)}) \approx 34.6\ dB \quad \text{(H.6)}$$

Thus quantization noise does not pose much of a problem for SSABR recording.

The instrumentation noise was also analyzed by shorting the electrodes together and performing three recordings. Four 5 second segments, which were 20 seconds apart, were taken from each recording (after applying a 40 dB notch filter in the 58 to 62 Hz region) and further segmented into 1 second segments. The mean noise voltage and the variance of the noise voltages across the 1 second segments from the three recordings were calculated (see Table H.1). Results show the mean noise voltage is approximately 4 μV in the 10 to 100 Hz frequency range. This noise is much lower than the EEG noise. Thus, the EEG noise is the major bottleneck in the system. In addition, this noise resembled white noise and no contamination at the SSABR frequency (40 Hz) was seen.

Table H.1. Mean and Variance of Recording Instrumentation Noise Voltage for 1 second segments.

Recording	Mean Noise Voltage ($1x10^{-6}$ Volts)	Variance of Noise Voltage ($1x10^{-7}$ Volts)
1	3.85	2.36
2	4.17	3.18
3	4.02	2
Mean	4.01	2.51

APPENDIX I: MSC ALGORITHM PARAMETERS

This appendix discusses whether a different number of subaverages or a different number of sweeps synchronously time averaged in each subaverage should be used for the MSC method. This is a concern because our test cases are shorter in time duration (a minimum of 4.8 seconds) than those used in Dobie's study (time duration ≈ 200 seconds; Dobie et al., 1993). To this end, SSABR data was recorded from 3 adult subjects. The data was preprocessed using a notch filter (1st order Butterworth IIR notch filter in the non-phase shift configuration with a bandstop region from 58 to 62 Hz and a 40 dB notch) to remove line noise. The data was then passed through an Artifact Rejection Algorithm in which buffers that contained data exceeding ±40 μV in amplitude were thrown out. For ROC analysis, the true positive data was obtained with an SSABR stimulus which had f_c = 1 kHz, f_m = 40 Hz, μ = 0.95 and L_c = 70 dB SPL. For the true negative data, a 1 kHz, 70 dB SPL tone[21] was presented to the subject. In both cases the sound was applied to the subject's right ear (see Appendix C for the test setup). Recordings were 219.2 seconds long and were later partitioned into 4.8 second segments to yield 40 data trials per recording. This resulted in 120 data trials for ROC analysis (40 trials per subject x 3 subjects). The sweeps were 0.3 seconds long.

The number of subaverages (q) was either 8 or 16 and the number of sweeps synchronously time averaged per subaverage (naps) was either 1, 2, or 6 for q = 8 whereas it was either 1 or 3 for q = 16. The comparison was done using ROC analysis (see Appendix E). The results are shown in Table I.1. To understand the table, the 1st row for the MSC detector indicates that ROC Area is 83.4% with an SE of 2.632% and the HR at a FAR of 5% is 39.2% with 95% confidence bounds of 30.5% and 48.5%.

Table I.1. ROC results for MSC and MSC-WA detectors with varying subaverages (q) and number of sweeps synchronously time averaged per subaverage.

		Frequency	MSC Detector		MSC-WA Detector	
q	naps	Resolution (Hz)	ROC Area (SE)	HR(0.05)	ROC Area (SE)	HR(0.05)
8	6	10	83.4 (2.632)	39.2 (30.5, 48.5)	82.07 (2.726)	45.8 (36.8, 55.1)
8	2	3.33	83.4 (2.632)	39.2 (30.5, 48.5)	83.22 (2.645)	45 (36, 54.3)
8	1	1.67	83.4 (2.632)	39.2 (30.5, 48.5)	83.74 (2.607)	45 (36, 54.3)
16	3	10	84.85 (2.522)	57.5 (48.2, 66.4)	84.17 (2.575)	48.3 (39.2, 57.6)
16	1	3.33	84.85 (2.522)	57.5 (48.2, 66.4)	86.2 (2.414)	60 (50.7, 68.7)

[21] This won't evoke SSABRs but provides recording conditions similar to those used when obtaining SSABRs.

For the MSC detector, results show that all parameter choices lead to close ROC performance, however, there is a slight ROC Area increase (not statistically significant) when there are 16 subaverages and 1 sweep within each subaverage (so there is actually no averaging occurring in each subaverage for data durations of 4.8 seconds). In addition, this choice of parameters leads to the highest HR(0.05). Thus, 16 subaverages and 1 segment (length of 3 sweeps) per subaverage were chosen for the MSC parameters.

For the MSC-WA detector, the highest ROC Area also occurred for the detector with 16 subaverages and 1 sweep within each subaverage, however, this was only borderline statistically significant when compared to the other ROC Areas (p values ranged from 2.5% to 10.5%). Furthermore, this choice of parameters lead to the highest value for HR(0.05). Thus, 16 subaverages and 1 segment (length of 3 sweeps) per subaverage were chosen for the MSC-WA parameters.

APPENDIX J: ROC RESULTS FOR THE AR-BASED SSABR DETECTORS

This appendix shows the ROC results for the various formats of the AR-based detectors. The results of the Supra-threshold (70 dB SPL) Quiet EEG test is in Table J.1. The results of the Mid-threshold (50 dB SPL) Quiet EEG test is in Table J.2 and the results of the Supra-threshold (70 dB SPL) Noisy EEG test is in Table J.3. The column MOE Type indicates whether a constant model order of 14 was used (i.e. the entry will be CO) or whether the AIC MOE was used with an MMO of 20 (i.e. the entry will be AIC) for obtaining the AR parameters for the Burg and MCM ARPEs. The number of bins in the signal band is either 7 or 11 and will be indicated for each row of each table. Furthermore the table entries read as follows: the first line of each cell is the ROC Area followed by the Standard Error of the ROC Area estimate in brackets. The second line of each cell is the HR(0.05) followed by the 95% lower and upper confidence bounds in brackets. All table entries are percentages from 0 to 100%.

Table J.1. ROC results for each AR-SSABR Detector versus data duration for the Supra-threshold Quiet EEG test.

ARPE	MOE Type	Number of Signal Bins	Data Duration (seconds)		
			4.8	9.6	19.2
Burg	CO	7	78.44(1.808) 35.3(30.1, 40.9)	86.82(2.043) 57.5(49.4, 65.2)	94.16(1.936) 86.3(76.3, 92.6)
Burg	AIC	7	78.52(1.805) 34.7(29.5, 40.2)	86.24(2.086) 40.6(33, 48.7)	93.39(2.061) 80(69.3, 87.8)
Burg	CO	11	81.59(1.687) 43.8(38.3, 49.4)	88.96(1.875) 51.3(43.3, 59.2)	96.25(1.55) 90(80.7, 95.3)
Burg	AIC	11	81.09(1.707) 41.9(36.4, 47.5)	87.81(1.968) 49.4(41.4, 57.3)	94.48(1.882) 73.8(62.5, 82.6)
MCM	CO	7	77.69(1.834) 27.5(22.8, 32.8)	82.16(2.354) 30.6(23.8, 38.5)	88.4(2.723) 41.3(30.6, 52.8)
MCM	AIC	7	77.95(1.825) 25(20.4, 30.2)	82.1(2.358) 35(27.8, 43)	86.38(2.942) 41.3(30.6, 52.8)
MCM	CO	11	79.25(1.778) 28.4(23.6, 33.8)	84.63(2.197) 32.5(25.5, 40.4)	90.77(2.434) 52.5(41.1, 63.6)
MCM	AIC	11	79.48(1.77) 27.2(22.5, 32.5)	84.55(2.203) 36.9(29.5, 44.9)	88.08(2.759) 25(16.3, 36.1)

Table J.2. ROC results for each AR-SSABR Detector versus data duration for the Mid-threshold Quiet EEG test.

ARPE	MOE Type	Number of Signal bins	Data Duration (seconds)		
			14.4	19.2	28.8
Burg	CO	7	79.31(2.123) 27.7(22, 34.1)	82.95(2.306) 40(32.5, 48)	87.41(2.391) 66.7(57.4, 75.4)
Burg	AIC	7	78.46(2.16) 41.5(35.1, 48.3)	82.12(2.357) 41.9(34.2, 49.9)	85.68(2.543) 60.7(51, 69.7)
Burg	CO	11	81.78(2.007) 31.7(25.8, 38.3)	85.01(2.172) 50.6(42.7, 58.6)	88.21(2.317) 55.4(45.7, 64.6)
Burg	AIC	11	81.85(2.004) 42.4(35.9, 49.2)	84.69(2.194) 46.3(38.4, 54.3)	86.91(2.437) 50.9(41.4, 60.4)
MCM	CO	7	79.24(2.127) 21.4(16.4, 27.5)	80.78(2.434) 28.1(21.5, 35.9)	83.92(2.685) 36.6(27.9, 46.3)
MCM	AIC	7	78.49(2.159) 33(27, 39.7)	80.22(2.465) 33.8(26.6, 41.7)	81.76(2.844) 29.5(21.4, 38.9)
MCM	CO	11	79.38(2.12) 20.5(15.6, 26.5)	82.3(2.346) 28.8(22, 36.5)	84.38(2.649) 38.4(29.5, 48.1)
MCM	AIC	11	79.35(2.123) 23.7(18.4, 29.9)	81.95(2.366) 27.5(20.9, 35.2)	82.71(2776) 28.6(20.7, 38)

Table J.3. ROC results for each AR-SSABR Detector versus data duration for the Supra-threshold Noisy EEG test.

ARPE	MOE type	Number of Signal bins	Data Duration (seconds)		
			4.8	9.6	19.2
Burg	CO	7	66.22(2.135) 20.6(16.4, 25.6)	72.68(2.813) 33.1(26, 41)	81.79(3.367) 50(38.7, 61.3)
Burg	AIC	7	65.45(2.149) 20.3(16.1, 25.2)	72.66(2.814) 33.1(26, 41)	82.7(3.29) 33.8(23.8, 45.3)
Burg	CO	11	68.24(2.094) 23.4(19, 28.5)	72.8(2.809) 28.1(21.5, 35.9)	83.45(3.224) 45(34, 56.5)
Burg	AIC	11	67.89(2.102) 20(15.9, 24.9)	72.5(2.82) 20(14.3, 27.2)	83.39(3.229) 37.5(27.2, 49.1)
MCM	CO	7	67.68(2.106) 21.9(17.6, 26.9)	71.03(2.874) 28.8(22, 36.5)	79.05(3.581) 45(34, 56.5)
MCM	AIC	7	67.39(2.112) 19.7(15.6, 24.6)	70.18(2.903) 17.5(12.1, 24.5)	80.39(3.48) 40(29.4, 51.6)
MCM	CO	11	68.49(2.089) 18.1(14.2, 22.9)	71.61(2.853) 25.6(19.2, 33.2)	81.02(3.43) 33.8(23.8, 45.3)
MCM	AIC	11	68.32(2.093) 21.3(17, 26.2)	71.03(2.874) 16.9(11.6, 23.8)	81.35(3.403) 30(20.6, 41.4)

APPENDIX K: SSABR DETECTOR COMPUTATIONAL COMPLEXITY

This appendix shows the derivation of the computational complexity for each SSABR Detector. To put all detectors on an equal scale it is assumed that the data, upon which detection is performed, is N points long which can be segmented according to N = 16nM where 16n is the number of sweeps in the data and M is the number of points per sweep. Furthermore, each subtraction operation will be considered as 1 addition and 1 multiplication.

Starting with the SSABR-Burg, SSABR-MCM and SSABR-PER detectors we see that each of these detectors begin with the V^{-1} WTA method. From Appendix D we see that the variance calculation requires 16n divisions and (16n)M additions for the calculation of the mean and then 2(16n)M additions, 2(16n)M multiplications and 16n divisions for variance calculation. To compute the weights we need (16n) additions and 2(16n) divisions. To compute the weighted average requires weighting each sweep and summing the weighted sweeps which results in (16n)M multiplications and (16n)M additions. Thus in total approximately 64nM + 16n additions, 48nM multiplications and 64n divisions are needed for the V^{-1} method. On top of this both AR detectors require M^2 additions and M^2 multiplications for spectral estimation meanwhile the periodogram based detector requires $M\log_2 M$ additions and $M\log_2 M$ multiplications (Marple, 1987). For all 3 of these detectors, the detection statistic requires 24 additions, 1 multiplication, 3 divisions and 1 log operation. Thus, in total, the SSABR-Burg and SSABR-MCM methods require approximately $64nM + M^2 + 16n$ additions, $48nM + M^2$ multiplications, 64n divisions and 1 log operation. The SSABR-PER detector requires approximately $64nM + M\log_2 M + 16n$ additions, $48nM + M\log_2 M$ multiplications, 64n divisions and 1 log operation.

For the MSC family of detectors the complexity calculation is done for 16 subaverages. Starting with the original MSC detector, the STA in each subaverage results in nM additions and nM multiplications (for the weight which is always 1/n) and 1 division to obtain the constant weight (same for all subaverages). Further there is the averaging of the 16 subaverages which results in 16M additions, 16M multiplications and 1 division. In addition, there are 17 FFT operations each requiring $M\log_2 M$ multiplications and $M\log_2 M$ additions. There is then the averaging of the subaverage response powers (16 additions, 16 multiplications and 1 division) and the calculation of the MSC value which is 1 division. Thus in total there are approximately $16(nM+M) + 17M\log_2 M$ additions, $16(nM+M) + 17M\log_2 M$ multiplications

and 4 divisions for the MSC detector. To perform the MSC-WA algorithm requires all of the complexity just mentioned plus that needed to get the variance of each subaverage result and to weight each subaverage. Using the analysis for the V^{-1} WTA method in Appendix D, this V^{-1} *subaverage* weighting results in approximately 64M additions, 48M multiplications and 64 divisions. Thus in total for the MSC-WA algorithm we need $16(nM+5M) + 17M\log_2 M$ additions, $16(nM+4M) + 17M\log_2 M$ multiplications and 68 divisions.

Finally incorporating the V^{-1} WTA method *within* a subaverage requires $4nM + n$ additions, $3nM$ multiplications and $4n$ divisions. Generalizing this to 16 subaverages results in $64nM + 16n$ additions, $48nM$ multiplications and $64n$ divisions. Thus these values are added onto both the MSC and MSC-WA methods to get the computational complexity for the MSC_{WTA} and $MSC\text{-}WA_{WTA}$ methods. Thus for the MSC_{WTA} method we have $16(5nM+M+n) + 17M\log_2 M$ additions, $16(4nM+M) + 17M\log_2 M$ multiplications and $64n$ divisions. For the $MSC\text{-}WA_{WTA}$ method we have $16(5nM+4M+n) + 17M\log_2 M$ additions, $16(4nM+5M) + 17M\log_2 M$ multiplications and $64(n+1)$ divisions.

To help determine the relative complexity of the SSABR-PER, SSABR-Burg, MSC_{WTA} and $MSC\text{-}WA_{WTA}$ methods, numbers from our analysis our used. Using $M = 147$, $n = 1, 2, 4$ or 6 (recall $N = 16nM$) we get the results shown in Table K.1. We see the SSABR-PER method and $MSC\text{-}WA_{WTA}$ methods have the least and most computational overhead respectively.

Table K.1. Example of SSABR-Burg, MSC and $MSC\text{-}WA_{WTA}$ Complexity.

Detector	n	Multiplications	Additions	Divisions	Log Operations
SSABR-PER	1	28,114	30,482	64	1
SSABR-Burg	1	28,665	31,033	64	1
$MSC\text{-}WA_{WTA}$	1	39,161	39,177	132	0
MSC_{WTA}	1	29,752	32,121	64	0
SSABR-PER	2	15,170	19,906	128	1
SSABR-Burg	2	35,721	40,457	128	1
$MSC\text{-}WA_{WTA}$	2	48,569	50,953	196	0
MSC_{WTA}	2	39,160	43,880	128	0
SSABR-PER	4	29,824	38,754	192	1
SSABR-Burg	4	49,833	59,305	256	1
$MSC\text{-}WA_{WTA}$	4	67,385	74,505	324	0
MSC_{WTA}	4	57,977	67,400	192	0
SSABR-PER	6	43,594	57,602	256	1
SSABR-Burg	6	63,945	78,153	384	1
$MSC\text{-}WA_{WTA}$	6	86,201	98,057	452	0
MSC_{WTA}	6	76,792	57,602	256	0

APPENDIX L: SSABR SNR

The range of SSABR SNR values that can be expected in practice is estimated using real EEG background noise and simulated sinusoids. The EEG data was obtained from five male adults (23 to 33 years of age), using the experimental setup described in Appendix C. A sound (70 dB SPL tone at 1 kHz) was presented to the subjects which would not evoke an SSABR but would provide recording conditions similar to those used when evoking SSABRs. The data was then preprocessed by passing it through a 40 dB notch filter (1st order Butterworth filter in the non-phase shift configuration) with the notch in the 58 to 62 Hz range. This was done because the line noise was very prominent in the recordings (1 to 4 μV in amplitude; very steady at 60 Hz). Furthermore, to obtain the SSABR SNR in the frequency vicinity of the 40 Hz and 80 Hz SSABRs, the data was bandpass filtered from 35 to 45 Hz (for the 40 Hz SSABR) and from 75 to 85 Hz (for the 80 Hz SSABR). In both cases the filter was a 2nd order Butterworth filter in the non-phase shift configuration.

To derive the SSABR SNRs, signal power and noise power were calculated for 1 second data durations. This was done across data segments which were 5 seconds in duration (this segment would thus yield 5 values for SSABR SNR). This was also repeated on three other 5 second duration data segments. Further, all 5 second data durations were separated by 20 seconds. Thus in total, this analysis on each subject yielded 20 SSABR SNR values. Further for the 40 Hz SSABR, the two SSABR amplitude values used were 800 nVpeak and 50 nVpeak (in the time domain). For the 80 Hz SSABR, the two SSABR amplitude values used were 200 nVpeak and 50 nVpeak. These values were chosen to represent the range of SSABRs which would be encountered in practice. The results are shown in Table L.1.

Table L.1 Range of SSABR SNRs measured over 1 second intervals.

Subject	40 Hz SSABR		80 Hz SSABR	
	50 nVpeak	800 nVpeak	50 nVpeak	200 nVpeak
1	-32.2 ± 1.6	-8.1 ± 1.6	-29.9 ± 1.6	-17.9 ± 1.6
2	-27.6 ± 1.4	-3.6 ± 1.4	-27.2 ± 1.1	-15.2 ± 1.1
3	-26.8 ± 1.5	-2.8 ± 1.5	-24.9 ± 2.3	-12.9 ± 2.3
4	-30.3 ± 1.9	-6.2 ± 1.9	-27.8 ± 2.1	-15.8 ± 2.1
5	-33.7 ± 3	-9.6 ± 3	-27.3 ± 3	-15.3 ± 3.1
Mean (of the means)	-30.1	-6.1	-27.4	-15.4

Note: All table entries are in dB.

REFERENCES

Antoniou, A., Digital Filters: Analysis, Design and Applications, McGraw-Hill, Inc., Toronto, © 1993.

Aoyagi, M., Fuse, T., Suzuki, T., Kim, Y., Koike, Y., (1993a), 'An application of Phase Spectral Analysis to Amplitude-Modulation Following Response', Acta Otolaryngolica, Supplement 504, pp 82-88.

Aoyagi, M., Kiren, T., Kim, Y., Suzuki, Y., Fuse, T., Koike, Y., (1993b), 'Optimal modulation frequency for amplitude-modulation following response in young children during sleep', Hearing Research, 65(1-2), pp. 253-261.

Aoyagi, M., Kiren, T., Kim, Y., Suzuki, Y., Fuse, T., Koike, Y., (1993c), 'Frequency Specificity of Amplitude-Modulation-Following Response Detected by Phase Spectral Analysis', Audiology, 32(5), pp. 293-301.

Aoyagi, M., Kiren, T., Furuse H., Fuse, T., Suzuki, Y., Yokota, M., Koike, Y., (1994a), 'Pure-tone Threshold Prediction by the 80-Hz Amplitude-modulation Following Response', Acta Otolaryngolica, Supplement 511, pp. 7-14.

Aoyagi, M., Furuse, H., Yokota, M., Kiren, T., Suzuki, Y., Koike, Y., (1994b), 'Detectability of Amplitude-modulation Following Response at Different Carrier Frequencies', Acta Otolaryngolica, Supplement 511, pp. 23-27.

Aoyagi, M., Kiren, T., Furuse, H., Fuse, T., Suzuki, Y., Yokota, M., Koike, Y., (1994c), 'Effects of Aging on the Amplitude-modulation Following Response', Acta Otolaryngolica, Supplement 511, pp. 15-22.

Arslen, E., Trevisi, P., Genovese, E., Lupi, G., Prosser, S., (1991), 'Hearing Loss Etiology in a group of 996 children', Annals of the New York Academy of Sciences, 630, pp. 315-316.

Barlow, J.S., (1986), 'Artifact Processing (Rejection and Minimization) in EEG Data Processing', in Handbook of Electroencephalography and Clinical Neurophysiology: Clinical Applications of Computer Analysis and other Neurophysiological Signals, Eds. Lopes Da Silva, F.H., Storm Van Leeuwen, W., Remond, A., Elsevier, Amsterdam, pp. 15-62.

Barlow, J.S., The Electroencephalogram: Its patterns and Origins, MIT Press, © 1993.

Bartoli, F., Cerutti, S., (1983), 'An Optimal Filter for the Reduction of Noise Superimposed to the EEG Signal', Journal of Biomedical Engineering, 5(4), pp. 274-280.

Bataillou, E., Thierry, E., Rix, H., (1991), 'Weighted Averaging with Adaptive Weight Estimation', 1991 IEEE Conference on Computers in Cardiology, pp. 37-40.

Batra, R., Kuwada, S., Maher, V., (1986), 'The frequency following response to continuous tones in humans', Hearing Research, 21(2), pp. 167-177.

Bess, F.H., Humes, L.E., Audiology: The Fundamentals, Williams & Wilkins, Baltimore, © 1995.

Bezerianos, A., Laskaris, N., Fotopoulos, S., Papathanasopoulos, P., (1995), 'Data Dependent weighted averages for recording of evoked potential signals', Electroencephalography and Clinical Neurophysiology, 96(5), pp. 468-471.

Blanco, S., Garcia, H., Quiroga, R.Q., Romanelli, L., Rosso, O.A., (1995), 'Stationarity of the EEG Series', IEEE Engineering in Medicine and Biology, 14(4), pp. 395-399.

Burshstein, D., Weinstein, E., (1985), 'Some Relations Between the Various Criteria for Autoregressive Model Order Determination', IEEE Transactions on Acoustics, Speech, and Signal Processing, 33(4), pp. 1017-1019.

Campbell, F.W., Atkinson, J., Francis, M.R., Green, D.M., (1977), 'Estimation of Auditory Thresholds Using Evoked Potentials' in Auditory Evoked Potentials in Man: Psychopharmacology Correlates of Evoked Potentials, ed. Desmedt, J.E., Karger, Basel, pp. 68-78.

Castanie, F., Gasmi, F., (1987), 'Extraction of Sinusoids from Noise by Prony Modeling of Higher-Order Correlation Functions', in Digital Signal Processing - 87, eds. Cappellini, V., Constantinides, A.G., Elsevier Science Publishers, North-Holland, pp. 783-786.

Centor, R.M., Schwartz, J.S., (1985), 'An Evaluation of Methods for Estimating the Area Under the Receiver Operating Characteristic (ROC) Curve', Medical Decision Making, 5(2), pp. 149-156.

Challis, R.E., Kitney, R.I., (1991), 'Biomedical Signal processing (in four parts). Part 3: The power spectrum and coherence function', Medical & Biological Engineering & Computing, 29(3), pp. 225-241.

Chambers, R.D., Meyer, T.A., (1993), 'Reliability of Threshold Estimation in Hearing-Impaired Adults Using the AMFR', Journal of the American Academy of Audiology, 4(1), pp. 22-32.

Champlin, C.A., (1992), 'Method for detecting auditory steady-state potentials recorded from humans', Hearing Research, 58(1), pp. 63-69.

Chen, H., Van Huffel, S., Vandewalle, J., (1996), 'Bandpass prefiltering for exponential data fitting with known frequency region of interest', Signal Processing, 48(2), pp. 135-154.

Chiappa, K.H., (1990), 'Principle of Evoked Potentials', in Evoked Potentials in Clinical Medicine, ed. Chiappa, K.H., Raven Press, New York, pp. 1-35.

Childers, D.G., (1977), 'Evoked Responses: Electrogenesis, models, methodology and Wavefront Reconstruction and Tracking Analysis', Proceedings of the IEEE, 65(5), pp. 612-616.

Cohen, L.T., Rickards, F.W., Clark, G.M., (1991), 'A Comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans', Journal of the Acoustical Society of America, 90(5), pp. 2467-2478.

Davila, C.E., Azmoodeh, M., (1994), 'Weighted Averaging of Auditory Brainstem Responses', 1994 IEEE Engineering in Medicine and Biology: 16th Annual Conference, pp. 251-252.

Diamond, G.A., (1989), 'Limited Assurances', American Journal of Cardiology, 63(1), pp. 99-100.

Dobie, R.A., Wilson, M.J., (1989), 'Analysis of Auditory Evoked Potentials by Magnitude-Squared Coherence', Ear and Hearing, 10(1), pp. 2-13.

Dobie, R.A., Wilson, M.J., (1991), 'Optimal smoothing of coherence estimates', Electroencephalography and Clinical Neurophysiology, 80(3), pp. 194-200.

Dobie, R.A., (1993), 'Objective Response Detection', Ear and Hearing, 14(1), pp. 31-35.

Dobie, R.A., Wilson, M.J., (1993), 'Objective Response Detection in the Frequency Domain', Electroencephalography and Clinical Neurophysiology, 88(6), pp. 516-524.

Dobie, R.A., Wilson, M.J., (1994a), 'Phase Weighting: A method to improve objective detection of steady-state evoked potentials', Hearing Research, 79(1-2), pp. 94-98.

Dobie, R.A., Wilson, M.J., (1994b), 'Objective Detection of 40 Hz auditory evoked potentials: phase coherence vs. magnitude-squared coherence', Electroencephalography and Clinical Neurophysiology, 92(5), pp. 405-413.

Dobie, R.A., Wilson, M.J., (1995a), 'Comparison of Objective Threshold Estimation Procedures for 40 Hz Auditory Evoked Potentials', Ear and Hearing; 16(3), pp. 299-310.

Dobie, R.A., Wilson, M.J., (1995b), 'Objective versus human observer detection of 40 Hz auditory-evoked potentials', Journal of the Acoustical Society of America, 97(5), Pt. 1, pp. 3042-3050.

Dobie, R.A., Wilson, M.J., (1996), 'A comparison of t test, F test, and coherence methods of detecting steady-state auditory evoked potentials, distortion-product otoacoustic emissions, or other sinusoids', Journal of the Acoustical Society of America, 100(4), Pt. 1, pp. 2236-2246.

Dumermuth, G., Gasser, T., Lange, B., (1975), 'Aspects of EEG Analysis in the Frequency Domain', in Computerized EEG Analysis, eds. Dolce, G., Kunkel, H., Gustav Fischer Verlag-Stuttgart, pp. 429-457.

Dumermuth, G., (1977), 'Fundamentals of Spectral Analysis in Electroencephalography' in EEG Informatics: A Didactic Review of Methods and Applications of EEG Data Processing, ed. Remond, A., Elsevier, Amsterdam, pp. 37-55.

Dumermuth, G., Molinari, L., (1987), 'Spectral Analysis of EEG Background Activity', in Handbook of Electroencephalography and Clinical Neurophysiology: Methods of Analysis of Brain Electrical and Magnetic Signals, eds. Gevins, A.S., Remond, A., Elsevier Publishers, Amsterdam, pp. 85-130.

Durieux-Smith, A., Picton, T.W., Bernard, P., MacMurray, B., Goodman, J.T., (1991), 'Prognostic Validity of Brainstem Electric-response Audiometry (BERA) in infants of a neonatal intensive care unit (NICU)', Audiology, 30(5), pp. 249-265.

Dyson, T., Rao, S.S., (1981), 'Equal Observation Interval Comparison of Maximum Entropy and Weighted Overlapped Segment Averaging Spectrum Estimation Techniques', IEEE Transactions on Acoustics, Speech and Signal Processing, 29(4), pp. 919-922.

Elul, R., (1969), 'Gaussian Behavior of the Electroencephalogram: Changes during performance of a mental task', Science, 164(877), pp. 328-331.

Elul, R., Hanley, J., Simmons, J.Q., (1975), 'Non-Gaussian behavior of the EEG in Down's syndrome suggests decreased neuronal connections', Acta Neurologica Scandinavia, 51(1), pp. 21-28.

Fan, Z., Wang, T., (1991), 'A Weighted Averaging Method for Evoked Potential Based on the Minimum Energy Principle', 1991 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 13(1), pp. 411-412.

Fridman, J., Zapulla, R., Bergelson, M., Greenblatt, E., Malis, L., Morrell, F., Hoeppner, T., (1984) 'Application of Phase Spectral Analysis for Brain Stem Auditory Evoked Potential Detection in Normal Subjects and Patients with Posterior Fossa Tumors', Audiology, 23(1), pp. 99-113.

Fox, S. I., Human Physiology, Wm. C. Brown Publishers, Boston, © 1996.

Galbraith, G.C., Wong, E.H., (1993), 'Moment Analysis of EEG Amplitude Histograms and spectral analysis: Relative Classification of Several Behavioral Tasks', Perceptual and Motor Skills, 76(3), Part 1, pp. 859-866.

Gasmi, F., Castanie, F., (1988), 'Successive Correlation Based Prony Modeling very low SNR behavior Analysis', 1988 IEEE International Conference on Acoustics, Speech and Signal Processing, 4, pp. 2420-2423.

Gasser, T.H., Dumermuth, G., (1979), 'Non-Gaussianity and Non-Linearity in Electroencephalographic Time Series', in Stochastic Control Theory and Stochastic Differential Systems, eds. Kohlmann, M., Vogel, W., Springer-Verlag, Berlin, pp. 373-386.

Gasser, T., Mocks, J., Verleger, R., (1983), 'SELAVCO: A method to deal with Trial to Trial variability of Evoked Potentials', Electroencephalography and Clinical Neurophysiology, 55(6), pp. 717-723.

Gasser, T., Mocks, J., Kohler, W., (1986), 'Amplitude Probability Distribution of Noise for Flash-Evoked Potentials and Robust Response Estimates', IEEE Transactions on Biomedical Engineering, 33(6), pp. 579-583.

Gerhardt, L.A., Park, S., (1986), 'Estimating Multiple Sinusoids with Low Signal to Noise Ratios', 1986 IEEE International Communications Conference, 3, pp. 1636-1640.

Gerull, G., Graffunder, A., Wernicke, M., (1996), 'Averaging Evoked Potentials with an Improved Weighting Algorithm', Scandinavian Audiology, 25(1), pp. 21-27.

Gevins, A.S., (1987), 'Overview of Computer Analysis', in Handbook of Electroencephalography and Clinical Neurophysiology: Methods of Analysis of Brain Electrical and Magnetic Signals, eds. Gevins, A.S., Remond, A., Elsevier Publishers, Amsterdam, pp. 31-83.

Glantz, S.A., Primer of Biostatistics, McGraw-Hill Book Company, New York, © 1987.

Glasscock, M.E., Jackson, C.G., Josey, A.F., The ABR Handbook, Thieme medical publishers, New York, © 1987.

Griffiths, S.K., Chambers, R.D., (1991), 'The amplitude modulation-following response as an audiometric tool', Ear and Hearing, 12(4), pp. 235-241.

Guyton, A.C., Human Physiology and Mechanisms of Disease, W.B. Saunders Company, Philadelphia, © 1992, pp. 392-398.

Hajian-Tilaki, K.O., Hanley, J.A., Joseph, L., Collet, J.P., (1997), 'A Comparison of Parametric and Nonparametric Approaches to ROC Analysis of Quantitative Diagnostic Tests', Medical Decision Making, 17(1), pp. 94-102.

Hanley, J.A., McNeil, B.J., (1982), 'The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve', Radiology, 143(1), pp. 29-36.

Hanley, J.A., McNeil, B.J., (1983), 'A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases', Radiology, 148(3), pp. 839-843.

Hayes, D., (1992), 'Auditory Brainstem Response (ABR) in Infants: Screening and Diagnostic Applications' in Screening Children for Auditory Function, eds. Ross, F.H., Hall, J.V., Bill Wilkerson Centre Press, Nashville, Tennessee, pp. 127-143.

Haykin, S., Communication Systems, John Wiley & Sons Inc., New York, © 1994.

Hendrix, R.A., DeDio, R.M., Sclafani, A.P., (1990). 'The use of diagnostic testing in asymmetric sensorineural hearing loss', Otolaryngology - Head and Neck Surgery, 103(4), pp. 593-598.

Hernandez, J.L., Valdes, J.L., Biscay, R., Jimenez, J.C., Valdes, P., (1995), 'EEG predictability: adequacy of non-linear forecasting methods', International Journal of Bio-Medical Computing 38(3), pp. 197-206.

Herring, R.W., (1980), 'The Cause of Line Splitting in Burg Maximum-Entropy Spectral Analysis', IEEE Transactions on Acoustics, Speech, and Signal Processing, 28(6), pp. 692-701.

Holm, S., Omar, Z.I.A., (1986), 'A Comparison of Autoregressive Order-Determining Criteria', in Signal Processing III: Theories and Applications, ed. I.T Young, Elsevier Science Publishers, North-Holland, pp. 29-32.

Hu, B., Gosine, R.G., (1997), 'A New Eigenstructure Method for Sinusoidal Signal Retrieval in White Noise: Estimation and Pattern Recognition', IEEE Transactions on Signal Processing, 45(12), pp. 3073-3083.

Hung, E.K.L., Herring, R.W., (1981), 'Simulation Experiments to Compare the Signal Detection Properties of DFT and MEM Spectra', IEEE Transactions on Acoustics, Speech, and Signal Processing, 29(5), pp. 1084-1089.

Husar, P., Henning, G., (1997), 'Bispectrum Analysis of Visually Evoked Potentials', IEEE Engineering in Medicine and Biology, 16(1), pp. 57-63.

Ioup, G.E., Pflug, L.A., Ioup, J.W., Field, R.L., (1993), 'Pre-filtering for Higher Order Advantage', 1993 IEEE Signal Processing Workshop on Higher-Order Statistics, pp. 309-313.

Jerger, J., Chmiel, R., Frost, J.D., Coker, N., (1986), 'Effect of sleep on the Auditory Steady State Evoked Potential', Ear and Hearing, 7(4), pp. 240-245.

Jervis, B.W., Nichols, M.J., Johnson, T.E., Allen, E., Hudson, N.R., (1983), 'A Fundamental Investigation of the Composition of Auditory Evoked Potentials', IEEE Transactions on Biomedical Engineering, 30(1), pp. 43-49.

John, M.S., Lins, O.G., Boucher, B.L., Picton, T.W., (1997), 'Multiple Auditory Steady-State Responses (MASTER): Stimulus and Recording Parameters', submitted to Audiology.

Johnson, T.L., Wright, S.C., Segall, A., (1979), 'Filtering of muscle artifact from the electroencephalogram', IEEE Transactions on Biomedical Engineering, 26(10), pp. 556-563.

Kamp, A., Lopes Da Silva, F.H., (1982), 'Technological Basis of EEG Recording' in Electroencephalography: Basic Principles, Clinical Applications and Related Fields, eds. Lopes Da Silva, F.H., Niedermeyer, E., Urban and Schwarzenberg, Baltimore-Munich, pp. 27-39.

Kassam, S., Signal Detection in Non-Gaussian noise, Springer-Verlag, New York, © 1988.

Kaveh, M., Lippert, G.A., (1983), 'An Optimum Tapered Burg Algorithm for Linear Prediction and Spectral Analysis', IEEE Transactions on Acoustics, Speech and Signal Processing, 31(2), pp. 438-444.

Kay, S., (1978), 'Improvement of Autoregressive Spectral Estimates in the Presence of Noise', 1978 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 357-360.

Kay, S.M., (1979), 'The Effects of Noise on the Autoregressive Estimator', IEEE Transactions on Acoustics, Speech and Signal Processing, 27(5), pp. 478-485.

Kay, S.M., (1980), 'Noise Compensation for Autoregressive Spectral Estimates', IEEE Transactions on Acoustics, Speech and Signal Processing, 28(3), pp. 292-303.

Kay, S.M., Marple, L., (1981), 'Spectrum Analysis - A Modern Perspective', Proceedings of the IEEE, 69(11), pp. 1380-1419.

Kay, S.M., (1984), 'Accurate Frequency Estimation at Low Signal-to-Noise Ratio', IEEE Transactions on Acoustics, Speech, and Signal Processing, 32(3), pp. 540-547.

Kay, S.M., Modern Spectral Estimation: Theory and Application, Prentice-Hall , Englewood Cliffs, New Jersey, © 1988.

Kemp, D.T., (1978), 'Stimulated acoustic emissions from within the human auditory system', Journal of the Acoustical Society of America, 64(5), pp. 1386-1391.

Kemp, B., Lopes Da Silva, F.H., (1991), 'Model-based analysis of neurophysiological signals', in Digital Biosignal Processing, ed. Weitkunat, R., Elsevier, Amsterdam, pp. 129-156.

Kuehls, J.F., Geraniotis, E., (1989), 'A Technique for Detecting Unknown Weak Signals in Noise that is not Additive White Gaussian', 1989 IEEE Military Communications Conference, pp. 370-376.

Kulik, R., (1995), 'A Synchronous Averaging Instrument for the Measurement of Distortion Product Otoacoustic Emissions', M.A.Sc. thesis, Department of Electrical and Computer Engineering, University of Toronto.

Kumaresan, R., (1982a), 'Accurate Frequency Estimation Using an All-Pole Filter with Mostly Zero Coefficients', Proceedings of the IEEE, 70(8), pp. 873-875.

Kumaresan, R., Tufts, D.W., (1982b), 'Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in Noise', IEEE Transactions on Acoustics, Speech, and Signal Processing, 30(6), pp. 833-840.

Kuwada, S., Batra, R. and Maher V. L., (1986), 'Scalp Potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones', Hearing Research, 21(2), pp. 179-192.

Levi, E.C., Folsom, R.C., Dobi, R.A., (1993), 'Amplitude-modulation following response (AMFR): Effects of modulation rate, carrier frequency, age and state', Hearing Research, 68(1), pp. 42-52.

Linden, R.D., Campbell, K.B., Hamel, G., Picton, T.W., (1985), 'Human Auditory Steady State Evoked Potentials during Sleep', Ear and Hearing, 6(3), pp. 167-174.

Lins, O.G., Picton, T.W., (1995a), 'Auditory steady-state responses to multiple simultaneous stimuli', Electroencephalography and Clinical Neurophysiology, 96(5), pp. 420-432.

Lins, O.G., Picton, P.E., Picton, T.W., Champagne, S.C., Durieux-Smith, A., (1995b), 'Auditory steady-state responses to tones amplitude-modulated at 80-110 Hz', Journal of the Acoustical Society of America, 97(5), pp. 3051-3063.

Lins, O.G., Picton, T.W., Boucher, B.L., Durieux-Smith, A., Champagne, S.C., Moran, L.M., Perez-Abalo, M.C., Martin, V., Savio, G., (1996), 'Frequency-Specific Audiometry using Steady-State Responses', Ear and Hearing, 17(2), pp. 81-96.

Lopes Da Silva, F.H., (1982), 'Computer-assisted EEG Diagnosis: Pattern Recognition in EEG Analysis, Feature Extraction and Classification', in Electroencephalography: Basic Principles, Clinical Applications and Related Fields, eds. Lopes Da Silva, F.H., Niedermeyer, E., Urban and Schwarzenberg, Baltimore-Munich.

Lopes Da Silva, F.H., Mars, N.J.I., (1987), 'Parametric Methods in EEG Analysis', in Handbook of Electroencephalography and Clinical Neurophysiology: Methods of Analysis of Brain Electrical and Magnetic Signals, eds. Gevins, A.S., Remond, A., Elsevier Publishers, Amsterdam, pp. 243-260.

Lutkenhoner, B., Hoke, M., Pantev, C., (1985), 'Possibilities and Limitations of Weighted Averaging', Biological Cybernetics, 52(6), pp. 409-416.

Marple, L., (1980), 'A New Autoregressive Spectrum Analysis Algorithm', IEEE Transactions on Acoustics, Speech and Signal Processing, 28(4), pp. 441-454.

Marple, S.L., Digital Spectral Analysis: with applications, Prentice-Hall, Englewood Cliffs, New Jersey, © 1987.

Martin, F.N., Introduction to Audiology, Prentice Hall, Englewood Cliffs, New Jersey, © 1991.

Mast, J., Victor, J.D., (1991), 'Fluctuations of steady-state VEPs: interaction of driven evoked potentials and the EEG', Electroencephalography and Clinical Neurophysiology, 78(5), pp. 389-401.

Mathworks, Matlab for Windows, version 4.2.c.1, The Mathworks Inc., © 1994.

Mathworks, Matlab Signal Processing Toolbox, version 3.0b, The Mathworks Inc., © 1994.

Mathworks, Higher-Order Spectral Analysis Toolbox, version 2.0, The Mathworks Inc. and United Signals and Systems Inc., © 1995.

McEwen, J.A., Anderson, G.B., (1975), 'Modeling the Stationarity and Gaussianity of Spontaneous Electroencephalographic Activity', IEEE Transactions on Biomedical Engineering, 22(5), pp. 361-369.

McGillem, C.D., Aunon, J.I., (1987), 'Analysis of Event Related Potentials', in Handbook of Electroencephalography and Clinical Neurophysiology: Methods of Analysis of Brain Electrical and Magnetic Signals, eds. Gevins, A.S., Remond, A., Elsevier Publishers, Amsterdam, pp. 131-169.

McGinn, D., Johnson, D.H., (1983), 'Reduction of All-Pole Parameter Estimator Bias by Successive Autocorrelation', 1983 IEEE International Conference on Acoustics, Speech and Signal Processing, 3, pp. 1088-1091.

McGinn, D.P., Johnson, D.H., (1989), 'Estimation of All-Pole Model Parameters from Noise-Corrupted Sequences', IEEE Transactions on Acoustics, Speech and Signal Processing, 37(3), pp. 433-436.

Mendel, J.M., (1991), 'Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications', Proceedings of the IEEE, 79(3), pp. 277-305.

Metz, C.E., (1978), 'Basic Principles of ROC Analysis', Seminars in Nuclear Medicine, 8(4), pp. 283-298.

Metz, C.E., (1986), 'ROC Methodology in Radiologic Imaging', Investigative Radiology, 21(9), pp. 720-733.

Metz, C.E., (1989), 'Some practical Issues of Experimental Design and Data Analysis in Radiological ROC Studies', Investigative Radiology, 24(3), pp. 234-245.

Milford, C.A., Birchall, J.P., (1989), 'Steady-state auditory evoked potentials to amplitude-modulated tones in hearing-impaired subjects', British Journal of Audiology, 23(2), pp. 137-142.

Miskiel, E., Ozdamar, O., (1987), 'Real-time Weighted Averaging System for Evoked Potentials', IEEE 9th Annual Conference of the Engineering in Medicine and Biology Society, pp. 589-590.

Narasimhan, S.V., Dutt, D.N., (1985), 'Software Simulation of the EEG', Journal of Biomedical Engineering, 7(4), pp. 275-281.

National Institute of Health, (1993), 'Early Identification of Hearing Impairment in infants and young children', NIH Consensus Statement, 11(1), pp. 1-24.

National Instruments, Labview Graphical Programming for Instrumentation, National Instruments Corporation, © 1994.

Neuman, M.R., (1995), 'Biopotential Electrodes', in The Biomedical Engineering Handbook, ed. Bronzino, J.D., CRC and IEEE Press, pp. 745-757.

Niedermeyer, E., (1982), 'Maturation of the EEG: Development of Waking and Sleeping Patterns', in Electroencephalography: Basic Principles, Clinical Applications and Related Fields, eds. Lopes Da Silva, F.H., Niedermeyer, E., Urban and Schwarzenberg, Baltimore-Munich, pp. 107-130.

Newby, H.A., Popelka, G.A., Audiology, Prentice Hall, New Jersey, © 1992.

Nikias, C.L., Mendel, J.M., (1993), 'Signal Processing with Higher-Order Spectra', IEEE Signal Processing Magazine, 10(3), pp. 10-37.

O'Donnell, R.D., Berkhout, J., Adey, W.R., (1974), 'Contamination of Scalp EEG Spectrum During Contraction of Cranio-Facial Muscles', Electroencephalography and Clinical Neurophysiology, 37(2), pp. 145-151.

Oppenheim, A.V., Schafer, R.W., Discrete-Time Signal Processing, Prentice-Hall, Englewood Cliffs, New Jersey, © 1989.

Orfandis, S.J., Optimum Signal Processing: An Introduction, Macmillan Publishing Company New York, © 1985.

Orr, R.S., Buffalano, C., (1990), 'Locally Optimum Detection: A Linear Model that Explains How it Works', 1990 IEEE Military Communications Conference, 2, pp. 806-811.

Paliwal, K.K., (1985a), 'Comparative performance evaluation of weighted Burg methods for spectral estimation of AR signals', in MELECON '85, Volume II: Digital Signal Processing, eds. Luque, A., Figuerias Vidal, A.R., Cappellini, V., Elsevier Science Publishers, North Holland, pp. 49-53.

Paliwal, K.K., (1985b), 'Further Simulation Results on Tapered and Energy-Weighted Burg Methods', IEEE Transactions on Acoustics, Speech and Signal Processing, 33(6), pp. 1624-1626.

Paliwal, K.K., (1986a), 'Some Comments on the Tapered Burg Methods', Indian Journal of Technology, 24(11), pp. 677-681.

Paliwal, K.K., (1986b), 'Some Comments about the iterative filtering algorithm for spectral estimation of sinusoids', Signal Processing, 10(3), pp. 307-310.

Parat, B., Digital Spectral Processing of Random Signals: Theory and Methods, Prentice-Hall, Englewood Cliffs, New Jersey, © 1994.

Parving, A., (1983), 'Epidemiology of hearing loss and aetological diagnosis of hearing impairment in childhood', International Journal of Pediatric Otorhinolaryngology, 5(2), pp. 151-165.

Persson, J., (1974), 'Comments on estimations and tests of EEG amplitude distributions', Electroencephalography and Clinical Neurophysiology, 37(3), pp. 309-313.

Picton, T.W., Skinner, C.R., Champagne, S.C, Kellet, A.J.C., Maiste, A.C., (1987), 'Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone', Journal of the Acoustical Society of America, 82(1), pp. 165-178.

Pfurtscheller, G., Neuper, C., Kalcher, J., (1993), '40 Hz oscillations during motor behavior in man', Neuroscience Letters, 164(1-2), pp. 179-182.

Pfurtscheller, G., Flotzinger, D., Neuper, C., (1994), 'Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG', Electroencephalography and Clinical Neurophysiology, 90(6), pp. 456-460.

Polydoros, A., Nikias, C.L., (1986), 'Detection of Unknown-Frequency Sinusoids in Noise via Autoregressive Modeling', 1986 IEEE International Conference on Acoustics, Speech and Signal Processing, 4, pp. 2807-2810.

Priestly, M.B., Spectral Analysis and Time Series, Academic Press, New York, © 1981.

Probst, R., Lonsbury-Martin, B.L., Martin, G.K., (1991), 'A review of otoacoustic emissions', Journal of Acoustical Society of America, 89(5), pp. 2027-2067.

Rance, G., Rickards, F.W., Cohen, L.T., De Vidi, S., Clark, G.M., (1995), 'The Automated Prediction of Hearing Thresholds in Sleeping Subjects Using Auditory Steady-State Evoked Potentials', Ear and Hearing, 16(5), pp. 499-507.

Regan, D., Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine, Elsevier, New York, © 1989.

Rickards, F.W., Clark, G.M., (1984), 'Steady-State Evoked Potentials to Amplitude-Modulated Tones', in Evoked Potentials II, The Second International Evoked Potentials Symposium, eds. Nadar, R.H., Barber, C., Butterworth Publishers, Toronto, pp. 163-168.

Rickards, F.W., Tan, L.E., Cohen, L.T., Wilson, O.J., Drew, J.H., Clark, G.M., (1994), 'Auditory steady-state evoked potential in newborns', British Journal of Audiology, 28(6), pp. 327-337.

Rizzo, S.R., Gutnick, H.N., (1991), 'Cochlear versus Retrocochlear Presbyacusis: Clinical Correlates', Ear and Hearing, 12(1), pp 61-63.

Rohling, Hermann, (1983), 'Radar CFAR Thresholding in Clutter and Multiple Target Situations', IEEE Transactions on Aerospace and Electronic Systems, 19(4), pp. 608-621.

Ruchkin, D.S., (1988), 'Measurement of Event-Related Potentials: Signal Extraction', in Human Event Related Potentials, ed. Picton, T.W., Elsevier, Amsterdam.

Sauter, D., Tomczak, M., Richard, A., Amady, M.M., Cail, F., (1990), 'Artifacts Detection and Pre-Cleaning in Spectral EEG Analysis', Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 12(2), pp. 840-842.

Schauf, C.L., Moffett, D.F., Moffett, S.B., Human Physiology: foundations and frontiers, Mosby College Pub., St. Louis, © 1990.

Schlindwein, F.S., Evans, D.H., (1990), 'Selection of the Order of Autoregressive Models for Spectral Analysis of Doppler Ultrasound Signals', Ultrasound in Medicine and Biology, 16(1), pp. 81-91.

Siegel, A., (1981), 'Stochastic Aspects of the generation of the Electroencephalogram', Journal of Theoretical Biology, 92(3), pp. 317-339.

Silverstein, S.D., (1991), 'Linear Minimum Free Energy Estimation: A Computationally Efficient Noise Suppression Spectral Estimation Algorithm', IEEE Transactions on Signal Processing, 39(6), pp. 1348-1359.

Simpson, D.M., Infantosi, A.F.C., Carneiro, J.F., Peixoto, A.J., De Sousa Abrantes, L.M., (1995), 'On the Selection of Autoregressive Order for Electroencephalographic (EEG) Signals', 1995 Midwest Symposium on circuits and Systems, 2, pp. 1353-1356.

Stappells, D.R., Linden, D., Suffield, J.B., Hamel, G., Picton, T.W., (1986), 'Human Auditory Steady State Potentials', Ear and Hearing, 5(2), pp. 105-113.

Steinberg, H.W., Gasser, T., Franke, J., (1985), 'Fitting Autoregressive Models to EEG Time Series: An Empirical Comparison of Estimates of the Order', IEEE Transactions on Acoustics, Speech and Signal Processing, 31(1), pp. 143-150.

Sugimoto, H., Ishii, N., Iwata, A., Suzumura, N., (1978), 'On the stationarity and normality of the electroencephalographic data during sleep stages', Computer Programs in Biomedicine, 8(3-4), pp. 224-234.

Sulkowski, W., Kowalska, S., (1982), 'Hearing Disorders in industrial workers: clinical analysis of the cases observed in the audiological laboratory of the Lodz Institute of Occupational Medicine during the years 1976-1980', Medycyna Pracy, 33(1-3), pp. 65-79.

Swets, J.A., (1988), 'Measuring the Accuracy of Diagnostic Systems', Science, 240(4857), pp. 1285-1293.

Swingler, D.N., (1979), 'A Modified Burg Algorithm for Maximum Entropy Spectral Analysis', Proceedings of the IEEE, 67(9), pp. 1368-1369.

Swingler, D.N., Walker, R.S., (1988), 'Linear-Predictive Data Extrapolation for Narrow-Band Spectral Estimation', Proceedings of the IEEE, 76(9), pp. 1249-1251.

Tang, Y., T., Norcia, A.M., (1993), 'Improved Processing of the Steady-State Evoked Potential', Electroencephalography and Clinical Neurophysiology, 88(4), pp. 323-334.

Tesei, A., Regazzoni, C.S., (1996), 'Application to Locally Optimum Detection of a New Noise Model', 1996 IEEE International Conference on Acoustics, Speech and Signal Processing, 5, pp. 2467-2470.

Tichavsky, P., Swami, A., (1995), 'Statistical Characterization of Sample Fourth-Order Cumulants of a Noisy Complex Sinusoidal Process', IEEE Transactions on Signal Processing, 43(7), pp. 1620-1629.

Tieri, L., Masi, R., Ducci, M., Marsella, P., (1988) 'Unilateral Sensorineural Hearing Loss in Children', Scandanavian Audiology, Supplement 30, pp. 33-36.

Tseng, S.Y., Chen,. R.C., Chong, F.C., Kuo, T.S., (1995), 'Evaluation of parametric methods in EEG signal analysis', Medicine in Engineering and Physics, 17(1), pp. 71-78.

Tucci, D.L., Wilson, M.J., Dobie, R.A., (1990), 'Coherence Analysis of Scalp Responses to Amplitude-modulated Tones', Acta Otolaryngologica, 109(3-4), pp. 195-201.

Vaz, F, Oliveira, P.G., Principe, J.C., (1987), 'A study on the best order for autoregressive EEG modeling', International Journal of Biomedical Computing, 20(1-2), pp. 41-50.

Vida, S., (1993), 'A computer program for non-parametric receiver operating characteristic analysis', Computer Methods and Programs in Biomedicine, 40(2), pp. 95-101.

Weiss, M.S., (1989), 'Testing EEG Data for statistical Normality', IEEE Engineering in Medicine and Biology Society 11th annual International Conference, IEEE, pp. 704-705.

Williams, A.B., Taylor, F.J., Electronic Filter Design Handbook: LC, Active and Digital Filters, McGraw-Hill Publishing Company, Toronto, © 1988.

Wu, W., Thomas, K.A., (1995), 'MR Imaging of 495 consecutive cases with sensorineural hearing loss', Acta Radiologica, 36(6), pp. 603-609.

Zetterberg, L.H., Herolf, M., (1975), 'Spectral Analysis of EEG for On-Line Computer Implementation Based on an adaptive linear model', in Quantitative Analysis of the EEG: Methods and Applications, eds. Matejcek, M., Schenk, G.,K., Basal, Baltimore, pp. 461-474.

Zimmerman, R.L., (1994), 'Neurological Disorders and Examination', in, Handbook of Clinical Audiology, ed. Katz, J., Williams & Wilkins, Baltimore.

Zurek, P.M., (1992), 'Detectability of Transient and Sinusoidal Otoacoustic Emissions', Ear and Hearing, 13(5), pp. 307-310.