ABSTRACT

Title of Document: DA VINCI'S ENCEPHALOGRAM:
IN SEARCH OF SIGNIFICANT BRAIN SIGNALS

NAYEF AHMAR
Master of Science
2005

Directed By: Professor Jonathan Z. Simon
Department of Electrical and Computer Engineering

Magnetoencephalography is a noninvasive tool that measures the magnetic activity of the brain. Its high temporal resolution makes it useful for studying auditory and speech models. However, it suffers from poor signal to noise ratio caused by corruption from non-stationary external noise, biological artifacts, and non-auditory neural noise in the brain. We remove external noise from neural channels using a frequency domain block least mean square adaptive filter with the help of three reference sensors that measure environmental noise alone. Significance tests that build on F-statistics present ample evidence of the benefit of such de-noising by increasing the number of significant channels and reducing the variability of false positives. Finally, the least significant and noisiest channel is filtered and used to de-noise neural signals while minimizing interference with the auditory signal. We propose a method for finding such reference channels and assess performance through receiver operating characteristics and statistical significance.

DA VINCI'S ENCEPHALOGRAM:
IN SEARCH OF SIGNIFICANT BRAIN SIGNALS

By

Nayef Elian Ahmar

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
Of the requirements for the degree of
Master of Science
2005

Advisory Committee:
Professor Jonathan Simon, Chair
Professor Shihab Shamma
Professor David Poeppel

Dedication

To my Family for the love, sense of purpose, and joy they brought into my life.

♥ ☺ ♫

Acknowledgements

To the Computational Sensorimotor Systems Lab (CSSL), Neural Systems Lab (NSL), and Cognitive Neuroscience of Language Lab (CNL) family members, you touched every aspect of my life by taking part in this work.

Table of Contents

Introduction

"The poet ranks far below the painter in the representation of visible things, and far below the musician in that of invisible things."

Leonardo da Vinci

Magnetoencephalography (MEG) and electroencephalography (EEG) are promising techniques with excellent temporal resolution. Nevertheless, the problem of noise—environmental interference, biological artifact, and background brain activity—hinders their full potential.

Our focus was on magnetoencephalography. Whether we are looking at evoked responses, more general hypothesis testing, or medical diagnosis, the localization of the source of magnetic fields in the brain is crucial. This requires solving the inverse problem, which necessitates clean measurements of MEG signals. Any contaminating noise will give a wrong solution, hence a wrong source, hence a wrong diagnosis or model.

Our system is equipped with three orthogonal reference channels that record external noise. The noise is non-stationary, so we resort to an adaptive algorithm, more specifically a frequency domain block least mean square method (Fast LMS) to estimate such additive noise and exclude it from our signal. We apply such a filter to data collected from an auditory experiment using sinusoidal amplitude-modulated (SAM) stimuli, which are powerful tools in detecting auditory steady state responses (ASSR). ASSRs are responses whose frequency components remain constant in amplitude and phase over a long period.

As a first step in validating our de-noising algorithm, confidence tests are used to distinguish evoked auditory responses from background noise. This enables us to fit an accurate dipole to our signals by discarding neural channels that are not significant and that would contribute noise. The F-statistics that build on amplitude information in the Fourier domain outperformed all other tests in identifying auditory channels. However, the non-Gaussian nature of background noise challenges the accuracy of F-statistics in measuring false positives. In view of this, we changed our method so that instead of using the inverse F-distribution to find thresholds, we use averaged data to set and control the threshold that determines the number of false positives. Two validation mechanisms, significance test design and receiver operating characteristics (ROC), are used to validate the performance of Fast LMS.

Lacking a channel analogous to the environmental sensor that records only artifacts or neural noise generated in the brain, we resort to finding the best approximation from the 157 neural channels to use as a reference sensor. That is, we identify the most noise-prone channels, or better a function of these channels. Nevertheless, any operation must have minimal effect on our auditory response. Distance, significance, and correlation all play a major role in choosing a reference channel that will preserve auditory signal integrity.

Having developed two building blocks, a de-correlation algorithm and a significance test, we build a system that can address all aspects of noise in our signal. The system consists of three Fast LMS processes, each specific to one of the three different kinds of noise (external, artifacts, and background brain activity), while

significance tests and ROC ensure that each operation performed is indeed beneficial and does not interfere with our signal.

Figure 1.0 summarizes the proposed model.

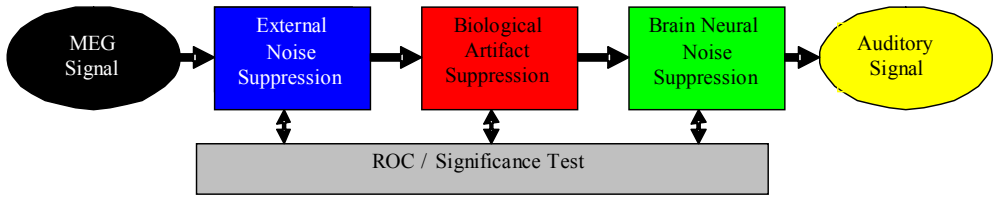

Figure 1.0: Flowchart of Noise suppression for recovering MEG Auditory Response. Four states, and three adaptive processes validated by ROC/ Significance test block.

Although our proposed model addresses MEG signals, it can be applied to other techniques, such as EEG. As for latency-motivated research that uses time information, a waveform significance test could be designed to replace the spectrum single bin significance test we used; the concept is the same.

Chapter 1: Preliminary Background and Experiment

"The noblest pleasure is the joy of understanding."

Leonardo da Vinci

What is Magnetoencephalography (MEG)?

MEG is a powerful functional tool for auditory experiments: it is non-invasive, fast (~1 ms), and spatially localizable (~5-10 mm). Compared to other imaging tools, such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and electroencephalography (EEG), MEG provides a different empirical base and results in observations reflecting the underlying neurophysiology, creating new ways to bridge the gap between imaging and cortical neurophysiology. Unlike PET and fMRI, we can draw conclusions from single-subject data, though group studies are preferable. MEG does not require subtraction between conditions (which assumes perhaps unjustified linearity), though such methods are still possible [9]. MEG complements EEG in many ways. Due to differences in how neurally generated magnetic and electric fields propagate, MEG is especially sensitive to neuronal activity in auditory areas, such as the supratemporal plane, making it a good tool for auditory research. MEG auditory responses lateralize strongly; EEG responses mix across cortical hemispheres and are strongest medially.

Neural Generation of Electromagnetic Fields

Excitatory postsynaptic potentials (EPSPs) are generated at the apical dendritic tree of a cortical pyramidal cell and motivate current that flows through the volume conductor from the non-excited membrane of the soma and basal dendrites to the apical dendritic tree sustaining the EPSPs. Some of the current travels within the dendritic trunk (primary current in blue Figure 1.1), while conservation of electric charges explains why the current loop is closed, with extracellular currents flowing through the volume conductor even at far distances (secondary currents in red) [5].

There are at least 10^{10} neurons in the human brain, equipped with 10^{14} interconnections, or synapses. When a signal is being processed, small currents flow in the neural system and produce a weak magnetic field. This is what the MEG system records. The MEG signal is derived from the net effect of ionic currents flowing in the dendrites of neurons during synaptic transmission and in the axons during action potentials (although net currents flowing in opposite directions down an axon from the point of action potential propagation give rise to magnetic fields that tend to cancel each other out). These net currents can be described as current dipoles with a position, orientation, and magnitude, but no spatial extent. According to the right-hand rule, a current dipole gives rise to a magnetic field that flows around the axis of its vector component. The magnetic field arising from the net current dipole of a single neuron is too weak to be directly detected. However, the combined fields from a region of about 50,000 active neurons can give rise to a net magnetic field that is measurable. Since current dipoles must have similar orientations in order to generate magnetic fields that reinforce each other, it is often the layer of pyramidal

cells in the cortex, which are generally perpendicular to its surface, that give rise to measurable magnetic fields. Furthermore, it is often groups of these neurons located in the sulci of the cortex with orientations parallel to the surface of the head that project measurable portions of their magnetic fields outside of the head.

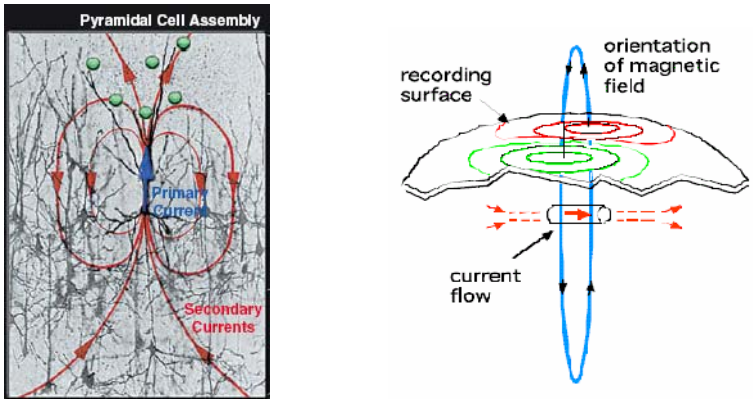

Figure 1.1: (Left) MEG is sensitive to the Primary Current (Blue) while EEG is sensitive to the volume current (Red) flowing in a pyramidal cell. [5]. (Right) Magnetic field (blue) caused by the primary current flow. On the head surface, MEG sensors record the strength and orientation of the field

The Problem of Low Signal to Noise Ratio (SNR)

A measured MEG signal is the combined result of the neural response to presented stimuli, additive background brain noise, biological artifacts, and external noise. Figure 1.2 lists many sources of additive external and biological noise with their relative strengths. For example, the cardiac magnetic field on the chest is 2 to 3 orders of magnitude larger than the fields outside the head generated by the brain [8]. As a result, the signal integrity of a given stimulus is weak. A robust de-noising algorithm is needed to improve signal to noise ratio.

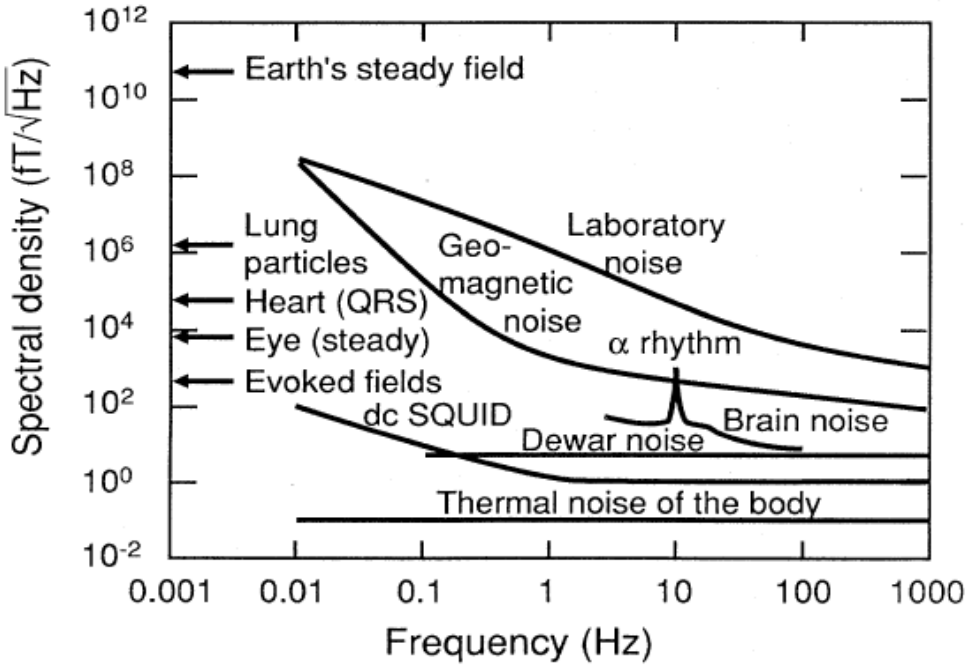

Figure 1.2: Peak amplitudes (arrows) and spectral densities of fields caused by typical biomagnetic and noise sources. Environmental noise is an order of magnitude higher than evoked fields signal. Artifacts and brain background noise are less powerful, but still strong [8].

Experiment

To explore the effects of noise and search for a de-noising algorithm, we designed an experiment that measures steady state response (SSR) using stationary sinusoidally amplitude modulated (SAM) stimuli. Auditory steady state response (ASSR) is a technique that predicts hearing sensitivity elicited with periodic modulated tones. It is frequency specific, hence suitable for our purpose. SAM stimuli, on the other hand are easy to track and analyze in the frequency domain for the simple spectrum they cover. We aim to test our algorithm using MEG data driven by SAM stimuli; however, the algorithm should not be limited to any MEG data.

Acoustic Stimuli

We present SAM sounds [24, 20] for two seconds 50 times each in a random order with inter-stimulus intervals uniformly distributed between 700 and 900 ms as described in [22]. A total of 20 stimuli were generated with five modulation frequencies (1.5 Hz, 3.5 Hz, 7.5 Hz 15.5 Hz and 31.5 Hz) and four distinctive carriers (pure tone; 1/3 octave pink noise; 1 octave noise and 5 octave noise all centered at 707 Hz). All stimuli were presented binaurally at a comfortable volume of approximately 70 dB SPL. Eight right-handed subjects, 5 female, were recruited.

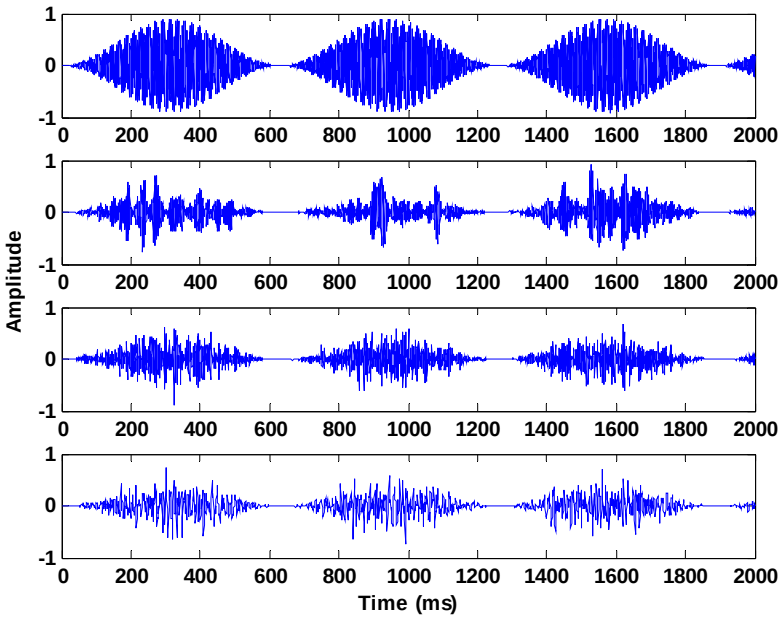

Figure 1.3: Waveforms for stimuli at a modulation frequency of 31.5 Hz with four distinctive carriers. From top to bottom: 1st Pure tone; 2nd, 3rd, and 4th pure tone plus pink noise at bandwidths of 1/3, 1, and 5 octave respectively. Narrow band nature of carrier is evident in the top figure. The random noise nature is more evident in the bottom one.

Recordings

The magnetic signals were recorded using a 160-channel, whole-head axial gradiometer system [12] housed in a magnetically shielded room. Its detection coils are arranged in a uniform array on a helmet-shaped surface on the bottom of the Dewar, with about 25 mm between the centers of two adjacent coils 15.5 mm in diameter. Sensors are configured as first-order axial gradiometers with a baseline of 50 mm; their field sensitivities are 5 fT/√Hz or higher at white noise region. Three of the 160 channels are magnetometers separated from the others and used as reference channels in noise filtering methods. The magnetic signals were bandpassed between 1 Hz and 200 Hz, notch filtered at 60 Hz, and sampled at the rate of 500 Hz.

MEG Signal

Responses to each stimulus were taken on each channel from 300 to 2300 ms post-stimulus (in order to guarantee steady state response [18]) and concatenated, resulting in 20 responses (of 2 ms resolution and 100 sec. duration) for each of the 157 channels. The fast Fourier transform (FFT) transformed each response. The result was 20 complex frequency responses (of 0.01 Hz resolution and 250 Hz extent) for each of the 157 channels. See Figure 1.4 for the magnitude of the FFT of the response of a single channel to the 31.5 Hz amplitude modulated sinusoid tone. The SSR peak at 31.5 Hz is stereotypically narrow with a width of 0.01 Hz. A valuable observation is that background responses became noisier with decreasing frequency challenging any proposed solution at such bands.

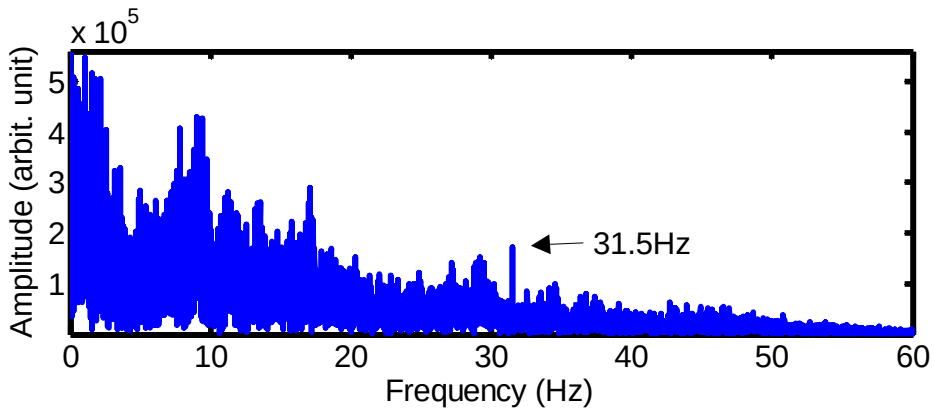

Figure1.4: Magnitude of the Fourier transform of single channel response to a pure tone sinusoidally amplitude modulated at 31.5 Hz. Note how the spectrum power decays as 1/frequency. The response at 31.5Hz is visually evident.

Chapter 2: External Noise Suppression

"Where the spirit does not work with the hand, there is no art."

Leonardo da Vinci

Introduction

Because the magnetic signals emitted by the brain are in the range of a few femto-Teslas (10^{-15}T), shielding from external magnetic signals, including the Earth's magnetic field (~$0.5x10^{-4}$T), is necessary. Shielding reduces noise by about 100dB; nevertheless, signal to noise ratio is still very low. As a result, suppressing external noise is crucial. There are some techniques, such as the Continuously Adjusted Least-Squares Method (CALM), that suppress such noise. These techniques are not powerful enough to clean our signal.

To remove such noise, which is typically non-stationary, we resort to adaptive filtering. Three reference channels, separated from the head, measure the noise alone, while 157 neuronal channels, arranged above the head's surface, record brain activity. The filter coefficients that linearly map the noise in the reference channels to the noise in the observed signal are calculated using the least mean square method (LMS) [27]. Then the estimated noise in the observed neuronal signal is subtracted. A fast version of LMS is adopted for speed [10].

Finally, we compare the Fast LMS method to the CALM, highlighting the weaknesses and strengths of each.

Continuously Adjusted Least Square Method (CALM)

CALM, used in the KIT-UMD MEG lab, reduces the non-periodical low frequency (<10 Hz) noise during MEG measurements [1]. The noise reduction procedure essentially eliminates any correlation that the MEG signal sensors have with any of the three reference magnetometers (set 25 cm apart from neural sensors) by removing the detected covariance from the neural MEG sensors. This is performed data point by data point, with computation extent of a moving window of a certain length. The spectrum for one neural signal and the three reference noise sensors are plotted in Figure 2.1. It is clear there is a high correlation between noise and raw signals, especially near 17, 25, and 180 Hz.

If $x_i(t)$ is the observed signal, $s_i(t)$ is the recovered signal, $n(t) = [n_1(t), n_2(t), n_3(t)]$ are the external noise references, and $w_i(t) = [w_{1i}(t), w_{2i}(t), w_{3i}(t)]$ is the set of filter weight coefficients vector, then all are related according to the equation:

$$s_i(t) = x_i(t) - w_i(t)n(t) \qquad (2.1)$$

The problem reduces to minimizing the power after subtracting the approximated noise components. We therefore find filter coefficients $w_i(t)$ that minimize L according to the equation:

$$L = \int_{t-T_1}^{t+T_2} \left[x_i(\tau) - w_i(\tau) n(\tau) \right]^2 d\tau \qquad \textbf{(2.2)}$$

where T1 and T2 are the lower and upper bounds for the sliding window used to calculate the weights.

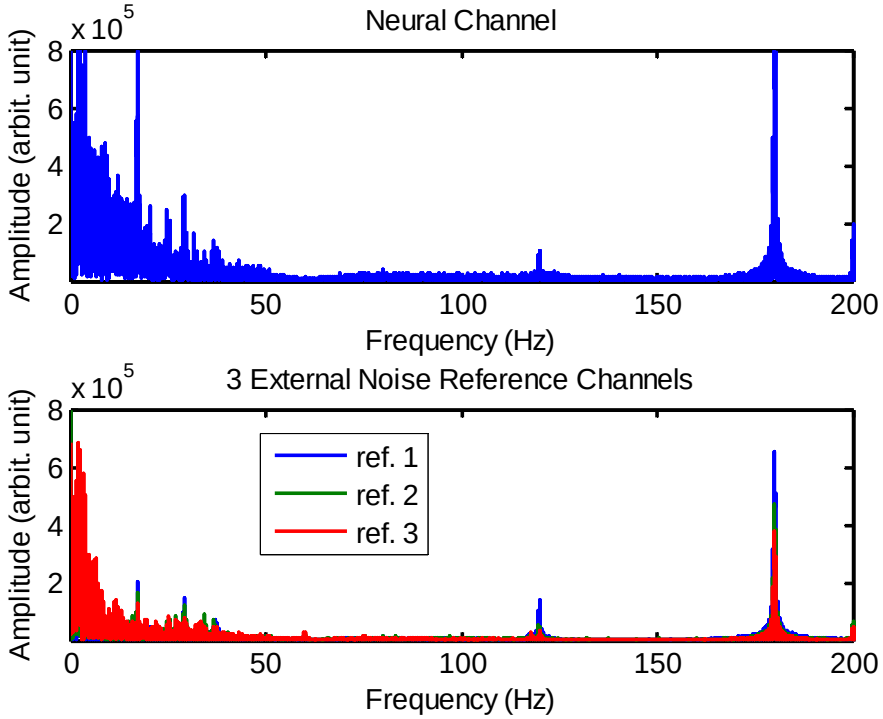

Figure 2.1: (Top) Magnitude of Fourier transform for raw data. (Bottom) Magnitudes of Fourier transform for the three reference noise channels are overlaid on top of each other. The reference and neural channels all have a similar envelope, whether it is wideband signal at low frequency or narrowband at 17, 25, 120, and 180Hz.

The CALM algorithm was not satisfactory for our purpose since it is not designed to extract narrowband noise that could have a severe effect on our signal. In addition, CALM is designed to deal with low frequency noise (it loses suppression capabilities for frequencies above 10Hz, Fig.2.2). Therefore, an alternative method to CALM needs to be addressed.

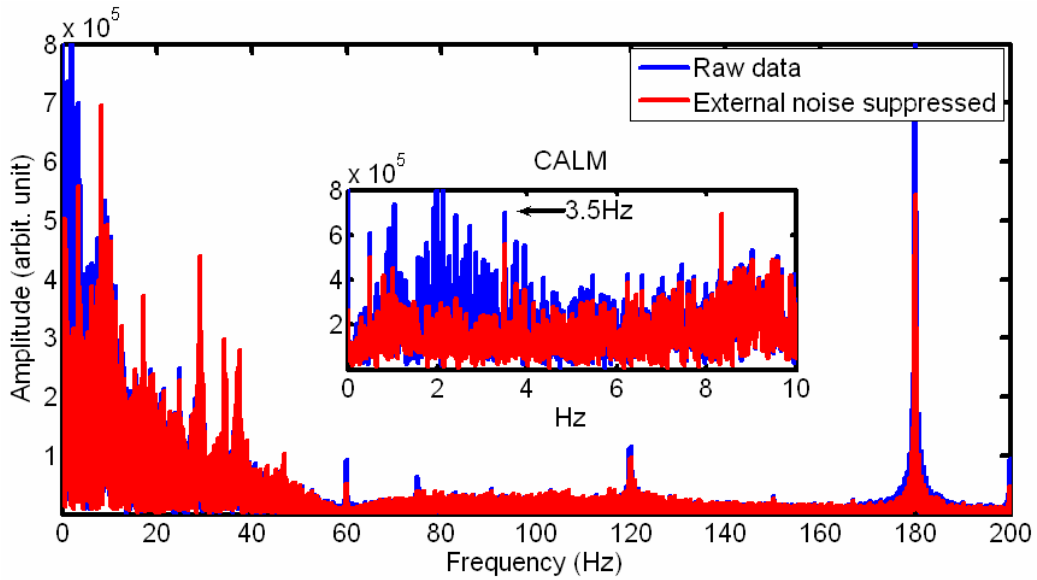

Figure 2.2: Magnitude of Fourier transform of raw signal (blue) and filtered with CALM (red) for one neural channel using the 3 external reference channels. Stimulus frequency is 3.5 Hz where the noise is most prominent, and where CALM is more effective. The response at that frequency is compromised, however. Less suppression is achieved at frequencies above 10Hz, including the high peak at 180Hz due to power line.

Adaptive Filtering

MEG brain signals are not stationary. This is expected, since the brain background is always changing even if we assume that the auditory cortex is responding to our stimuli in a stationary fashion. However, the noise is also non-stationary, as many of the magnetic sources are isolated events and are therefore variable in time and space (e.g., elevator, cars, a chair moved from one place to another). Simulation showed that, indeed, neither brain signals nor external noise is stationary, though the latter is less variable (experiment dependent).

Our target signal is corrupted with additive noise. We record the sum of the two and cannot tell them apart. However, we do have access to measured noise, which is a filtered version of source noise. The problem is to estimate filter coefficients that

transform noise in the reference channels into noise in the signal. If we can isolate the noise in the signal, we can subtract it from the measured signal and thereby recover the original signal.

A classic solution would be to use a Wiener filter (Figure 2.3) as a solution to our problem; however, the use of a linear shift-invariant Wiener filter will not be optimal because both neural signals and noise are non-stationary. Nevertheless, an adaptive Wiener filter that has filter coefficients that are allowed to vary as a function of time may provide effective noise cancellation in our non-stationary environment. It is worth mentioning that slicing the data in small frames that are almost stationary is not an accurate model, since we would be giving up resolution for stationarity, and hence we would lose the range of frequencies that are of interest to us.

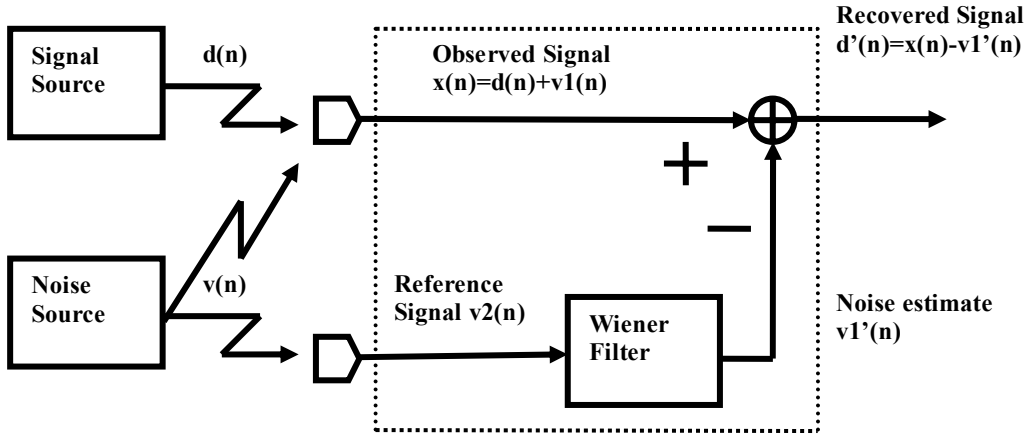

Figure 2.3: Wiener noise cancellation to recover unknown neural signal d(n) from a measurement of that signal x(n) after being corrupted with additive noise v1(n). A measurement of the filtered version of such noise v2(n) is used to estimate filter coefficients to recover and subtract the additive noise v1'(n) and subtract it from the observed signal. Lack of stationarity in the signal causes this algorithm to fail; as it has no feedback, it is not adaptive.

Adaptive noise canceling is an optimal filtering technique that is feasible every time we have a reference input available. What is special about this family of techniques are its adaptive capabilities, low output noise, and low signal distortion. It can also handle unknown and non-stationary inputs. Adaptive filtering is a stable operation that turns off when no improvements in SNR are achieved. It outperforms other conventional optimal filter configurations [27].

The adaptive filter should be able to take in a measured observed signal and a reference signal and be able to compute a minimized estimate of the original signal and an estimate of the noise. Two operators are derived:

$\hat{c}$: Correlate or filter noise channel to capture noise in observed signal

$\hat{d}$: De-correlate or subtract filtered noise captured in observed signal and recover original signal.

If X is one of the 157 neural channels, R is a reference channel, and T is a filter transformation

R $\hat{c}$ X = T(R) ∩X = recovered noise.

R $\hat{d}$ X = X- T(R)∩X = error = recovered signal.

Figure 2.4 illustrates the building block for the adaptive filter we are seeking to design.

Figure 2.4 Adaptive filter acts as a building block to de / correlate noise and observed signal x(n) by using one or multi-reference channels as inputs, and estimated signal d'(n) and noise v1'(n) as outputs.

Frequency Domain Block LMS (Fast LMS)

Background brain activity is always changing even if the area of interest responds to stimuli in a stationary fashion. External noise is also non-stationary since many of its sources are of random characteristics in space and time. We use an adaptive process that automatically adjusts the filter parameters to minimize estimation error.

Figure 2.5 shows the block diagram for Fast LMS. The design is a modification of Fast LMS described in [10] upgraded to handle multi-reference sensors. Instead of subtracting the filtered version of one reference channel, we subtract the scaled sum of the three filtered orthogonal references, each with its own filter coefficients to capture noise in observed signal. The scaling is done adaptively among all three references with equal probabilities. A validation of our algorithm follows in chapter 4.

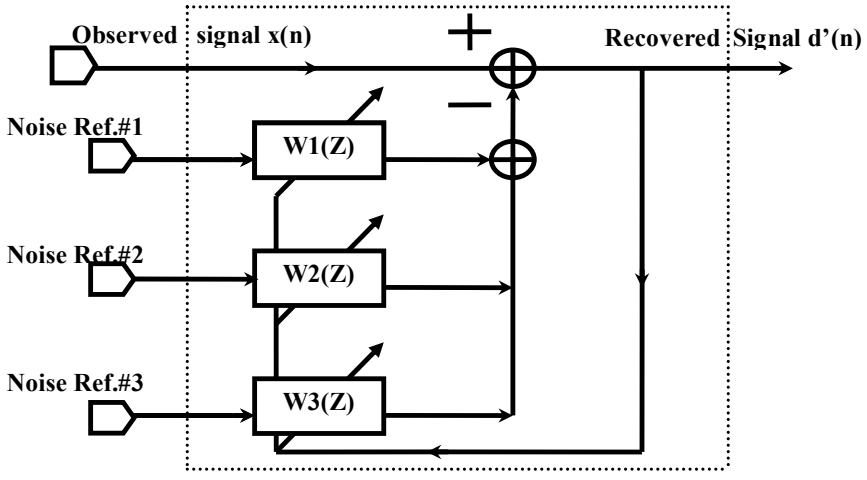

Figure 2.5: Three reference adaptive filter noise cancellation. Filter coefficients W are computed so the sum of the scaled version of the three observed reference noise channels captures as much of the observed signal as possible. An estimate of the noise is calculated and then subtracted to yield the recovered signal. Arrows are in both directions because the filter coefficients are adapting to the signals.

From Figure 2.6, we can see narrowband noise suppressed using Fast LMS. In the low frequency range, where noise is prominent and hard to suppress, our algorithm did a fair job in cleaning noise around our stimulus frequency, 3.5Hz. Fast LMS removed most noise at 180Hz (third harmonic of 60 Hz). Remember that a notch filter removed 60Hz power line noise after recording and prior to noise suppression. There is no neural interest in the 180Hz band; rather, it is a measure of performance for our de-noising algorithm.

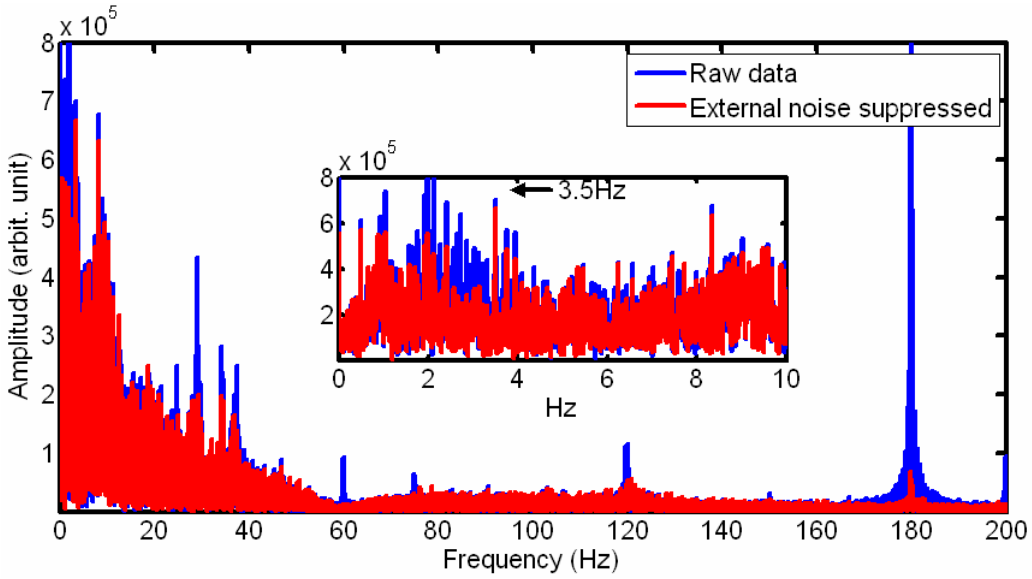

Figure 2.6: Magnitude of Fourier transform of raw signal (blue) vs. filtered signal (red) of one channel using Fast LMS,. Magnified spectrum (0-10Hz) shows the response at the stimulus frequency (3.5Hz) with suppressed noise in the vicinity. Note power line noise suppressed at 180Hz by 20dB. More narrow band signals at lower frequencies were suppressed too including 25Hz, 31Hz, and what is left after applying the notch filter at 60Hz. Most importantly, response at 3.5 Hz is barely compromised.

Adaptation coefficients we used in our design: Block size 128 (block size is equal to filter length), adaptation constant 0.01, forgetting factor 0.96 (the larger the

forgetting factor, the larger the memory). For less stationary signals, it is better to reduce filter size and/or forgetting factor.

Fast LMS vs. CALM

CALM is not efficient at removing multiple sources of narrow band noise, and it is almost passive at frequencies above 10Hz (see Figure 2.2). On the other hand, CALM is much faster than Fast LMS (the latter is slowed by the time-intensive discrete Fourier transform computations, and takes about the same time as signal acquisition). However, Fast LMS is a whole-spectrum de-noising algorithm and does an excellent job of narrowband noise suppression.

Comparative quantitative measures, summarized for both methods in Table 2.1, shows that Fast LMS better preserved the signal at the stimulus frequency. In addition, it removed 1.4 dB of noise for frequencies below 10Hz, and approximately 20 dB around 180Hz. On the other hand, CALM compromised stimulus frequency for more suppression at low frequency. It became less effective as frequency increased. This weakness is best seen at 180Hz, where very little noise is removed.

	Block-LMS Filter	**CALM Filter**
3.5Hz (Signal loss)	0.3 dB	1.8 dB
1-10Hz (Noise Suppression)	1.4 dB	2.0 dB
175-185Hz (Noise Suppression)	19.9 dB	4.3 dB

Table 2.1: Signal reduction comparison between Fast LMS and CALM filters. For a stimulus at 3.5Hz, Fast LMS preserved most of the auditory signal, reducing it by only 0.3dB, as opposed to 1.8dB for CALM. Also, LMS better removed most of power line noise at 180Hz. Calm, on the other hand, removed more noise, on average, below 10Hz, but at the expense of stimulus frequency reduction.

Several other methods were tried to suppress the noise: Least Mean Square (LMS), Recursive Least Square (RLS), QR-Decomposition Least Square Lattice (QRDLSL), and others. They are described in detail in [10]. Out of all these methods, Fast LMS prevails as the best technique in terms of both SNR improvements and speed. Although the algorithm exploits a block structure, the method is slower than other non-adaptive filtering methods because of discrete Fourier transform (DFT) computations. Nevertheless, SNR was improved with minimal invasiveness to our stimulus response signal.

Chapter 3: Detection and Significance tests

"Why does the eye see a thing more clearly in dreams than the imagination when awake?"

Leonardo da Vinci

Introduction

Neural and non-neural background noise impedes identification of the auditory stimulus evoked response. The problem of identifying those channels that have some auditory structure buried in noise is a detection problem. Knowing where to look temporally (M100 latency) or spectrally (SAM frequency) for a response adds more information, but does not answer the question of whether a channel is significant or not. This information is invaluable for accurately solving the inverse problem to determine the neural sources generating the measured magnetic field. Therefore, it is crucial to distinguish what is strong from what is significant.

We explore many existing confidence tests to assess and classify responses for signals corrupted by noise with our experimentally obtained MEG signals. This can be done either by measuring for consistency across different presentations (e.g., phase information) or by contrasting the signal strength at one frequency with the noise strength in neighboring bands (amplitude information).

Although various tests build on different techniques, they all tend to agree on what is significant and what is not, though some are more stringent than others. The F-test, as we shall see, proved to be the most powerful of all the tests we looked at.

Signal Detection

Our designed SAM stimulus is tuned at a very narrow band. To track any response,

it suffices to look for the same modulation frequency. If activity at that frequency is detected, it can only be credibly attributed to our stimulus if the activity is absent from neighboring bins. This is exactly how the F-test classifies signals as significant or not. A signal could be a true positive, in line with an auditory response, or it could be just background noise, a false positive.

Because of the non-stationarity of the signals in question, and hence the complexity for computing the posterior densities, we abandon Bayesian solutions. Our detection problem is nonparametric and data driven. Besides, the nature of our stimulus, narrowband in nature, makes it more feasible to deal with compared to wideband signals that require more complicated waveform detection framework.

The Neyman Pearson Criterion

The Neyman-Pearson criterion says that decision rules are constructed to have the maximum probability of detection (P_T) while not allowing the probability of false positives (P_F) to exceed a certain value α. We first set a false positive threshold not to be exceeded as explained below, then we maximize probability of detection for that particular false detection.

$$\textbf{Max } \{P_T \{\textbf{such that } P_F \leq \alpha\}\} \qquad (3.1)$$

F-Test for Hidden Periodicity

The F-test examines the signal to noise ratio for the signal at stimulus frequency compared to the background noise at neighboring frequencies [7, 18]. After taking the FFT of the concatenated 50 presentations, the average power of 120 frequency bins separated by 0.01Hz (60 below and 60 above the stimulus frequency) is measured,

denoting background noise. Total noise bandwidth is 1.2 Hz. The formula to compute this ratio is given by:

$$R_F = \frac{120\left|a_{sf}\right|^2}{\sum_{i=sf-60, i\neq sf}^{i=sf+60} \left|a_i\right|^2} \qquad (3.2)$$

The F-test ratio is computed for all 20 stimuli at each frequency. Four stimuli should yield a correct detection if there was a strong auditory response, while the other 16 can produce only false positives, since the experiment was designed such that there is exactly one frequency per stimulus (Table 3.1).

Average False Positives

MEG biological noise is strongly non-Gaussian. Experimental simulations failed when using the Gaussian assumption, irrespective of the central limit theorem (if one averages enough identical distributions, the resulting density converges to a Gaussian), again because of non-stationarity. This should not be surprising, since all the above tests statistically use Gaussian white noise as the null hypothesis, and typical MEG noise is non-Gaussian. As a result, a noisy signal often fails the marginal tests by exceeding the number of false positives allowed. For example, if we allow a rate of one percent of detections to be false, with any of the methods explored, it is more likely that the observed false positives are more than one percent due to the high structure of our background noise. As a result, we turned the problem around. Usually, we set a fixed false positive (α) probability that corresponds to a fixed theoretical *s* value (based on F-statistics). Instead, we computed an estimate of false positives at the frequency of interest by averaging a large number of

observations of false positives (α_{avg}) for responses at the other stimulus frequencies (where no auditory response we know of should exist). Accordingly, we tuned the α probability to achieve the desired α_{avg}. Table 3.1 shows how we can exploit such averaging.

Response	Stimuli	1.5Hz				3.5Hz				7.5Hz				15.5Hz				31.5Hz			
	Bandwidth	0	.3	1	5	0	.3	1	5	0	.3	1	5	0	.3	1	5	0	.3	1	5
1.5Hz		N	N	N	N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3.5Hz		1	1	1	1	N	N	N	N	1	1	1	1	1	1	1	1	1	1	1	1
7.5Hz		1	1	1	1	1	1	1	1	N	N	N	N	1	1	1	1	1	1	1	1
15.5Hz		1	1	1	1	1	1	1	1	1	1	1	1	N	N	N	N	1	1	1	1
31.5Hz		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	N	N	N	N

Table 3.1Ddistribution of correct detections vs. false detections. There are five stimuli modulation frequencies (first row) and five corresponding response frequencies (first column). Each frequency is presented with four different bandwidths (second row). Correct detection could be as high as 157 channels per stimuli (diagonal in green). False detection is set to be one on average per stimulus (off diagonal in red).

We sort F-scores for all false positives where we expect no auditory response (Table 3.1 off diagonal, 16 stimuli per frequency). We normalize scores and get *s* values in order to construct the cumulative distribution of false positives (Figure 3.1). We then find the normalized score *s* that meets an average α false positive (e.g. find *s* such that α = 1%). We do the inverse operation for true positives. First, we sort and normalize all true positives' F-statistics scores (Table 3.1 diagonal, per row for each frequency) and form the equivalent cumulative distribution. We then use the *s* value computed earlier as a threshold in order to find what true positives it corresponds to. Channels with scores above the threshold are labeled significant, while channels with scores below threshold are labeled as non-significant.

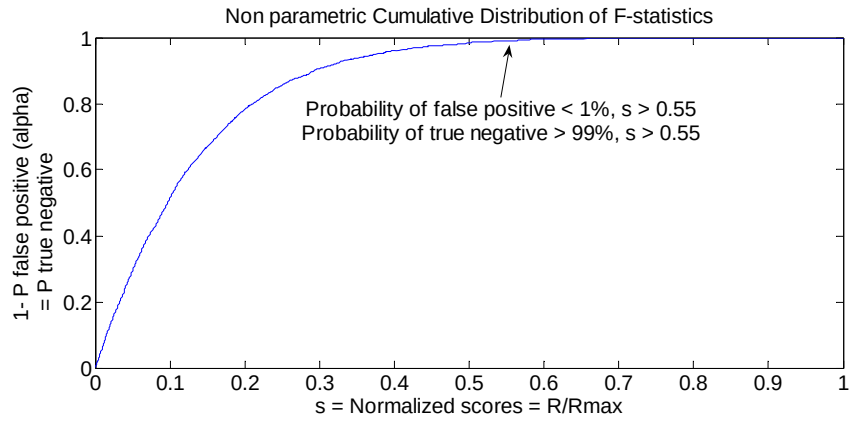

Figure: 3.1 Cumulative distribution of F- statistics for false positives. We normalize our F-test scores to get s values, and plot the corresponding probability of false positives (α) from the order statistics. E.g. There is <1% false positives for s > 0.55 of normalized F scores. Note s is not a probability, but a normalized F-score. False positives α is a probability.

Consistency across all significance tests

There are three kinds of significance tests: Those that carry phase information, such as Rayleigh's phase coherence test [19, appendix c1]; those that carry amplitude information, such as the F-test; and those that use both phase and amplitude, such as Hotelling's T^2 [11, appendix c3]. We explored all the aforementioned methods, in addition to a family of other significance tests, in search of the strongest auditory responses. This other family of tests includes but is not limited to: T-test, phase weighted test and phase coherence weighted test [18]; Multitaper DPSS [17, appendix c2]; permutation test [6]; and union and intersection joint tests of the former tests [2]. Although some of these tests are more stringent than others, all methods, within a small margin of difference, agreed on the results. Refer to Figure 3.2 for receiver operating characteristics (ROC) for the three most prominent methods tried.

In our preliminary analysis, Rayleigh's phase test, which carries phase information, complements F-test that builds on amplitude information. However, the improvement from combining the tests was not worth the computational complexity added. More importantly, as more subjects were analyzed, the F-test stood out as the best, least expensive test.

An improvement to the F-test was suggested by [18], where complex values were projected onto an expected phase, creating a t-test. For our MEG data, we used neighboring channels to compute the expected phase. This weighting method did not improve on the F-test, and typically had less power. This is consistent with noise contamination whose phase is spatially coherent.

Receiver Operating Characteristics (ROC)

We plotted receiver operating characteristics for the three most prominent tests that we explored. For each number of false positives, we counted the number of true positives that were detected. We averaged across six subjects and show results for all five modulation frequencies (Figure 3.2). Note that the F-test consistently outperforms other tests.

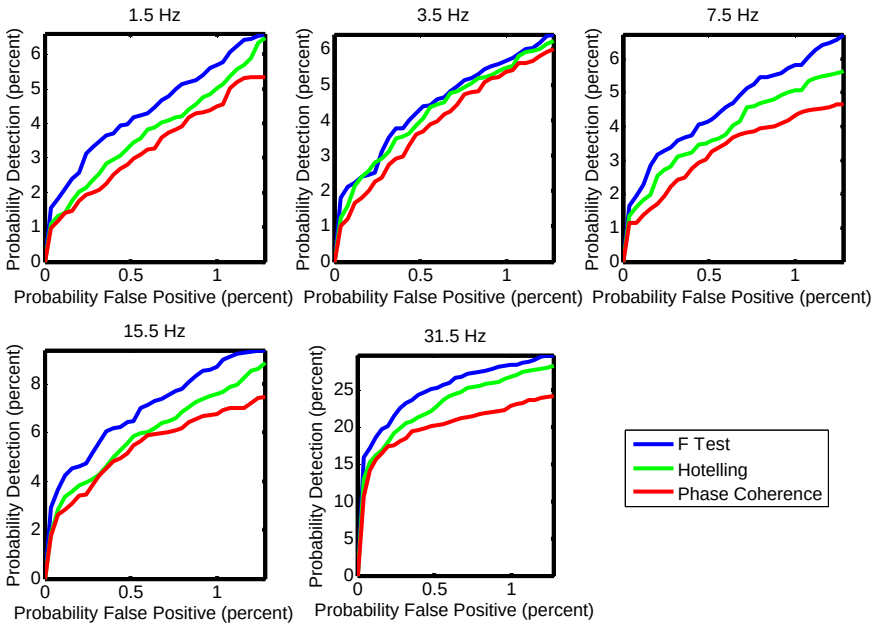

Figure 3.2: Receiver operating characteristics comparison for F, Hotelling's, and phase coherence significance tests averaged across six different subjects for signals after removing external noise. The F-test, on average, is better than all other tests; Hotelling's test ranks second, and phase coherence third across all frequencies tested. Low detection at frequencies at and below 15.5Hz is an indication of the high noise corruption.

Significance Head Map

We calculated significance tests for all channels then plotted the head map by showing amplitude and phase information for each channel. A channel is labeled as significant by drawing a circle around it.

Figure 3.3 (left) shows the complex field distribution at 3.5 Hz for a stimulus modulated at 3.5 Hz. Arrows represent the magnetic field response, at each of the 157 channels, as phasors: the length of the arrow denotes amplitude and the orientation denotes phase. Circles mark those channels identified as significant by the F-test ($\alpha < 1/157$). It is clear that many of the channels strong in magnitude are not significant (see especially the frontal channels in the right hemisphere). Figure 3.3

(right) shows the response at the same frequency (3.5 Hz) but from a stimulus whose modulation frequency was 1.5 Hz, and so only noise is expected. One significant channel is found, which is consistent with $\alpha < 1/157$ for 157 channels. Notice the apparent spatial coherence of the phase structure. It is neural but not stimulus driven.

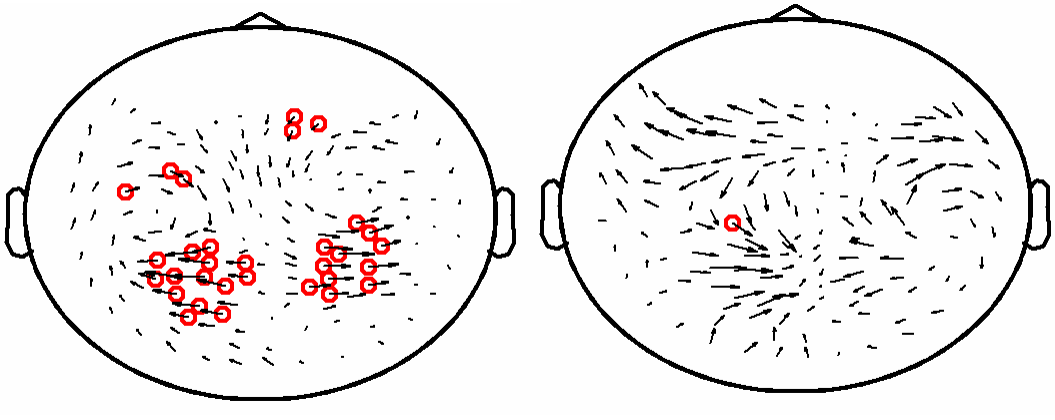

Figure 3.3: Head map plots for all 157 neural channels. Arrows represent amplitude (arrow's length) and phase (arrow's direction) of signals at the stimulus frequency. Red circles denote significant channels. (Left) Response at 3.5 Hz for a 3.5 Hz modulated stimulus. 31 significant channels. The structure of the dipole, distribution, and number of significant channels are evidence that this is indeed an auditory response. (Right) Response at 3.5 Hz for a 1.5 Hz modulated stimulus. The single circle is a false positive labeled as "significant". Note the strength and structure of the arrows; there is rich activity, but it is not auditory by experimental design. These head map plots were created after using Fast LMS.

Analogous maps for the 7.5 Hz (and 15.5 Hz) cases are shown in Figure 3.4. The number of significant channels may be higher in either hemisphere. In Figure 3.4 (right), there are two false positives. Recall that the test is designed so that there is, on average, one false positive for all responses in which there is no signal expected.

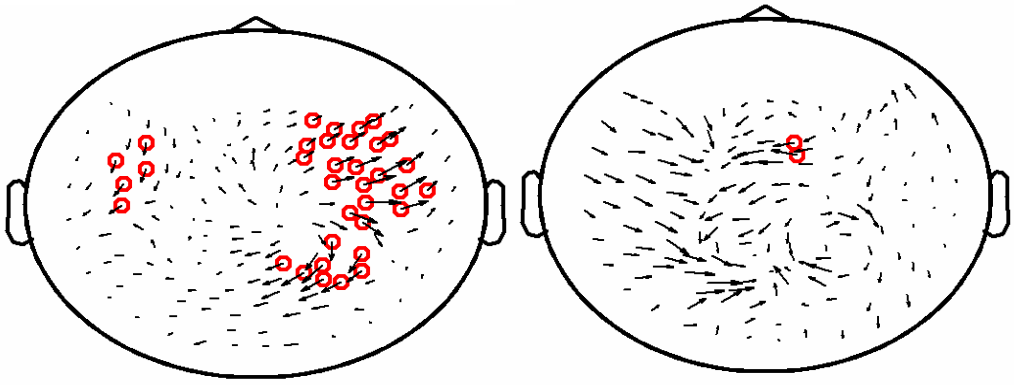

Figure 3.4: (Left) Response at 7.5 Hz for a 7.5 Hz modulated stimulus. 35 significant channels. The dipole on the left hemisphere is not formed; more noise suppression is needed. (Right) Response at 7.5 Hz for a 15.5 Hz modulated stimulus: two false positives. The constraint to allow only one false positive was on average across all 16 off diagonal stimuli from Table3.1.

To test how well the algorithm performed compared to other methods of selecting significant channels, we applied a permutation test to the magnetic distribution shown in Figure 3.4 (left). From the right hemisphere, 30 channels were chosen at random and labeled "significant", a dipole was fit to those channels, and its goodness of fit (GOF) was calculated. This process was repeated 1000 times in order to compute a cumulative distribution function of the GOF. The GOF of the dipole for the significant channels based on a joint test of F, and a phase coherence test, 84%, was not achieved in any of the 1000 permutations (i.e. $\alpha \leq 0.1\%$)

Summary

The F-test outperformed other tests in detecting significant channels for measuring MEG responses. Phase coherence tests performed well, but did not add valuable significance to the F-test with a reasonable cost of implementation. Expected-phase weighted tests fared more poorly, presumably because the expected phase used was

the local spatial average, which was typically coherent even when no signal was present. For this reason and because of similar properties of the noise, the null hypothesis of Gaussian noise, independent across channels, was not appropriate, leading us to rescale the probability distributions in order to match the measured false positive rate. For the purpose of fitting dipoles to auditory responses stimulated by SAM tones, using significant channels determined by the F-test yielded better goodness of fit.

Chapter 4: Evaluation of Fast LMS

"Experience does not err. Only your judgments err by expecting from her what is not in her power."

Leonardo da Vinci

Earlier we developed an adaptive algorithm to suppress external noise. Next, we need a performance measure to be able to validate the algorithm. Several techniques are explored. We first use the significance tests described in the previous chapter to show that Fast LMS indeed increases the number of true positives and reduces variability among false positives. Comparison of the raw data with the filtered data shows substantial improvement in many significant channels and a decrease in number and variability of false positives. Then we look at receiver operating characteristics, and we show that Fast LMS increases the probability of detection for fixed false positive.

Fast LMS and Significance

Fast LMS increases SNR on multiple dimensions. As illustrated in Figure 4.1 (left), we show the complex field distribution at 3.5 Hz for a stimulus modulated at 3.5 Hz. Arrows represent the magnetic field response, at each of 157 channels, as phasors: the length of the arrow denotes amplitude and the orientation denotes phase. Circles mark those channels identified as significant by the joint balanced test ($\alpha < 1/157$) [2]. Note that many of the channels, though strong in magnitude, are not found to be significant. Figure 4.1 (right), on the other hand, shows the response at the same frequency (3.5 Hz) after applying the noise suppression. The number of significant channels is increased, the structure of the background of the head map is established,

and most of the strong signals over the temporal lobes (where robust signals are expected in response to auditory stimuli) are de-noised.

Fast LMS also reduced the variability of false positives. In Figure 4.2, we illustrate the effects of Fast LMS on false positives by looking at responses at 15.5Hz for a 3.5Hz modulated stimulus. In the left plot, eight false positives are identified, and after applying the de-noising algorithm (Figure 4.2, right side) the number of false positives drops to two. Note that the test is designed so that there is, on average, one false positive for all responses in which there is no signal expected. Even so, the variance of the false positives among different responses per stimulus frequency is reduced after applying Fast LMS, providing more evidence of the value of the noise suppression.

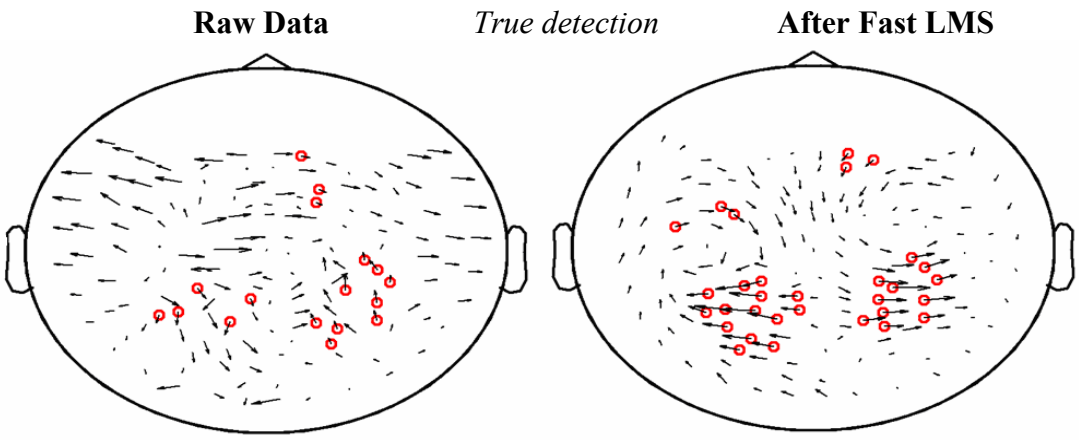

Figure 4.1Both stimulus and response are at 3.5Hz. (Left) Raw data. Although there are many significant channels, noise level hides the structure of the dipole. (Right) Filtered data using Fast LMS with the three external noise references. Note that the number of significant channels is increased, with dipoles both on the left and right hemispheres established; in addition the strong responses on the peripheral contour are suppressed.

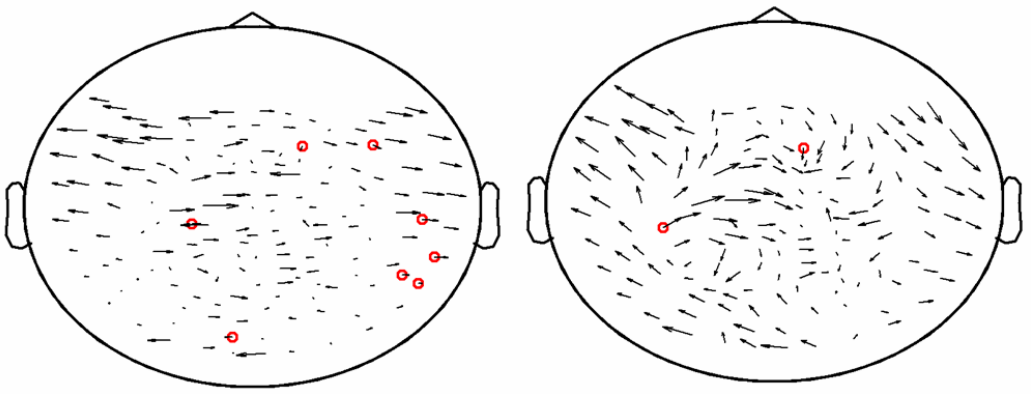

Figure 4.2 Stimulus presented at (3.5 Hz), response analyzed at 15.5Hz. (Left) Raw data. Note the many scattered false positives with random structures. (Right) Filtered data using Fast LMS. The number of false positives is reduced. Signal strengths are boosted after suppressing external noise but are comprised of artifacts and neural background noise. No auditory response is expected.

ROC curve

Receiver operating characteristics is a classical measure of performance in any detection problem. For any false detection probability (α value), there is a corresponding probability of detection. For any given false positive, we like to maximize our detection. We adopt a non-parametric approach in computing these probabilities, based on Table 3.1, which sets the threshold to find true positives for diagonal elements.

The ROC curve before and after applying the de-noising algorithm is plotted in Figure 4.3. ROC for 1.5Hz is almost linear; this is due to low number of detection with poor SNR. Averaging more subjects might smooth out some of the irregularities such as those found in stimulus 15.5Hz at (~α =0.5). For all frequencies, however,

there is an increase in probability of true detection for any particular α probability. Hence, Fast LMS increases the probability of detection.

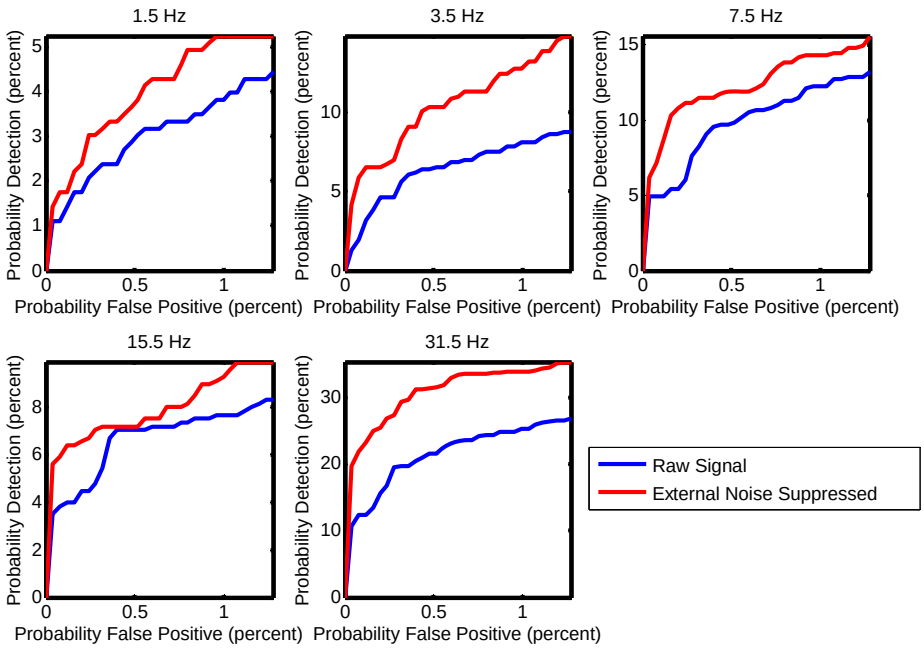

Figure4.3 Receiver Operating Characteristics curves comparison between raw and externally de-noised signals for one subject. Fast LMS improved significant detection for all frequencies. For example, a one percent false positive detection (α) threshold increases the number of significant channels identified from two(1.5Hz) to ten(31.5Hz) percent.

In addition, our quantitative measures (Table 2.1) showed that Fast LMS removed many of the narrowband noise that is classified as environmental, for example, the 180Hz power line source, while leaving our response at the stimulus frequency uncompromised.

In summary, Fast LMS improved SNR, increased the number of significant neuronal channels, and suppressed and regularized false positives.

Chapter 5: De-noising Biological Noise

"Life is pretty simple: You do some stuff. Most fails. Some works. You do more of what works. If it works big, others quickly copy it. Then you do something else. The trick is the doing something else."

Leonardo da Vinci

Having successfully suppressed external noise, we next look for ways to replicate the same process for biological noise. Now that we are equipped with a powerful adaptive algorithm, the problem is reduced to finding a reference channel, if it exists, that measures biological signals generated outside the brain, such as heartbeats, eye blinks, and other artifacts. It is even more valuable if this putative channel can also record neural brain activities that are not auditory. For our MEG-KIT system, there was no reference channel dedicated solely to this purpose; however, we have 157 channels, and the outermost (along the external contour) have a chance at capturing artifact noise. We conduct a brute force search for the channel that best measure such noise, apply the de-noising algorithm already developed and validate how acceptable the results are.

The remaining question is to what extent we can eliminate non-auditory brain-generated neural noise. We show that with some intelligent observation, we can reduce the non-relevant brain background noise with tolerable loss to our signals. Our approach should not be limited to auditory signal, it could generalize to other sensory modalities (e.g. visual, motor) as well.

Artifact Removal

Two widely used artifact suppression techniques for MEG are principal components analysis (PCA) and independent components analysis (ICA). The general idea is to transform the multi-channel dataset into another domain where the signal and noise are statistically independent or orthogonal (they contain no shared information). PCA can effectively remove artifacts from EEG signals; however, when the artifacts are an order of magnitude stronger than the signal, as in the case of MEG, PCA does not perform well. ICA, on the other hand, is sensitive enough to be able to pick out very small signals buried in noise. ICA tries to minimize the mutual information between mixed signals by maximizing the entropy between signals [14]. ICA is a powerful technique; however, it is still not applied effectively in MEG research.

Rather than solve the whole problem, we relax many of the constraints and try to identify a single noise channel rich in artifact that could be de-correlated from all other channels. The Fast LMS developed earlier is a de-correlation tool. We want to find sensors that are the most prone to noise, the least significant, and the most distant from our signals. We then subtract any correlation between our signals and the reference sensor channels. We are interested in maximum suppression of noise and minimum interference with our auditory signal without the need to separate different components.

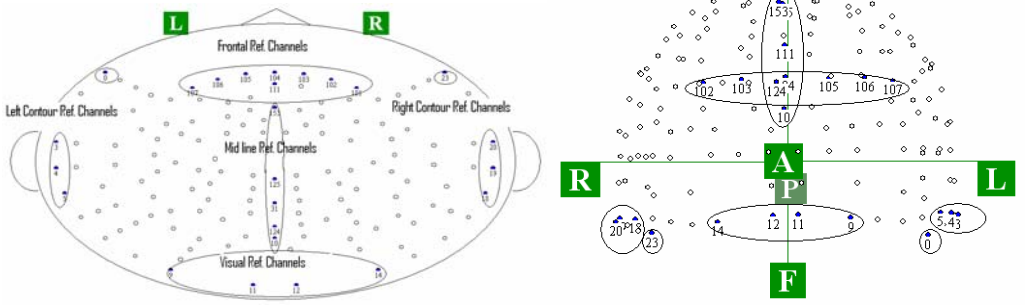

Figure5.1: (Left) 2D isofield contour map showing all 157 channels. Channels 0 and 23 are the top left and right on both hemispheres. We group channels on the peripheral contour and midline as potential reference channels for removing artifacts and background neural noise generated in the brain. They are furthest apart from neural brain activity and most susceptible to biological noise. (Right) 3D isofield contour map for same group of channels.

Our brute force search for the channels most prone to artifact noise led us to two candidates: channels 0 and 23 in Figure 5.1. These channels are furthest away from auditory signals and spatially best localized to contain primarily artifacts. Out of all the available neural sensors, these two channels, when de-correlated from other channels, yielded the highest SNR without hampering our SSR signal. ICA results confirm [29] that these two channels contributed most power to two major components of the heartbeat. Further validation would require looking at significance tests and ROC curves. Note that the nearest neighboring channels could yield similar results.

We demonstrate the effectiveness of our method through head map and spectrum plot examples.

On the left side of Figure 5.2, we have only six significant channels with no clear auditory dipole fit. After applying reference channels 0 and 23 using Fast LMS

(right), the number of significant channels increases, and there is the possibility of fitting a dipole on the left hemisphere. There is no noticeable change in the structure or flow of responses, indicating that the significance increases are due to noise removal that boosted SNR and hence lowered the threshold for true positives.

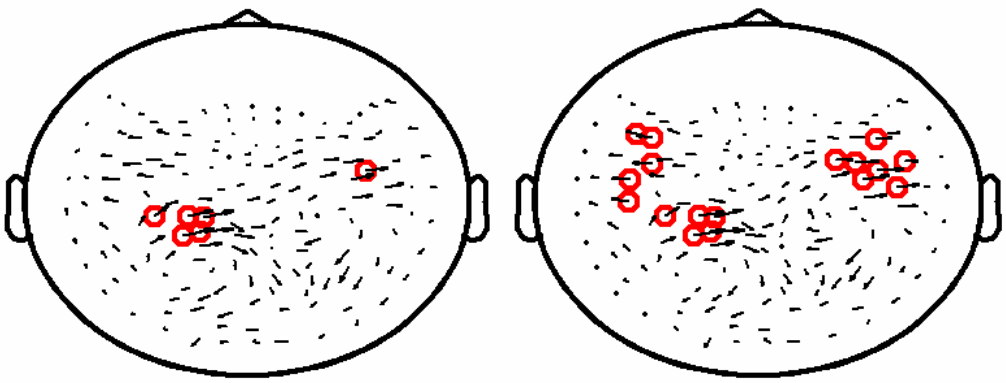

Figure 5.2 Head map plot for stimulus and response at 7.5Hz. (Left) After removing environmental noise. Note there are only six channels total with no dipole. (Right) After removing correlated noise with channels 0 and 23. Note how strong signals on the frontal side of the head are wiped out. In addition, number of significant channels increase, a dipole is forming on the left hemisphere.

A hierarchy of noise removal effects can be seen in Figure 5.3. The top figure shows the effect of removing external noise. This operation increased the number of significant channels and established ground for a dipole on the right hemisphere. However, in the left hemisphere, there is powerful peripheral noise impeding a dipole pattern. When we apply de-noising based on the reference channels 0 and 23 (left bottom), we can see that the number of significant channels decreased, but so did the noise on the two sides of the head, triggering the concern that these channels might be recording only noise. Note that the stimulus frequency is 1.5 Hz, rich in artifacts. It is an extremely noisy band.

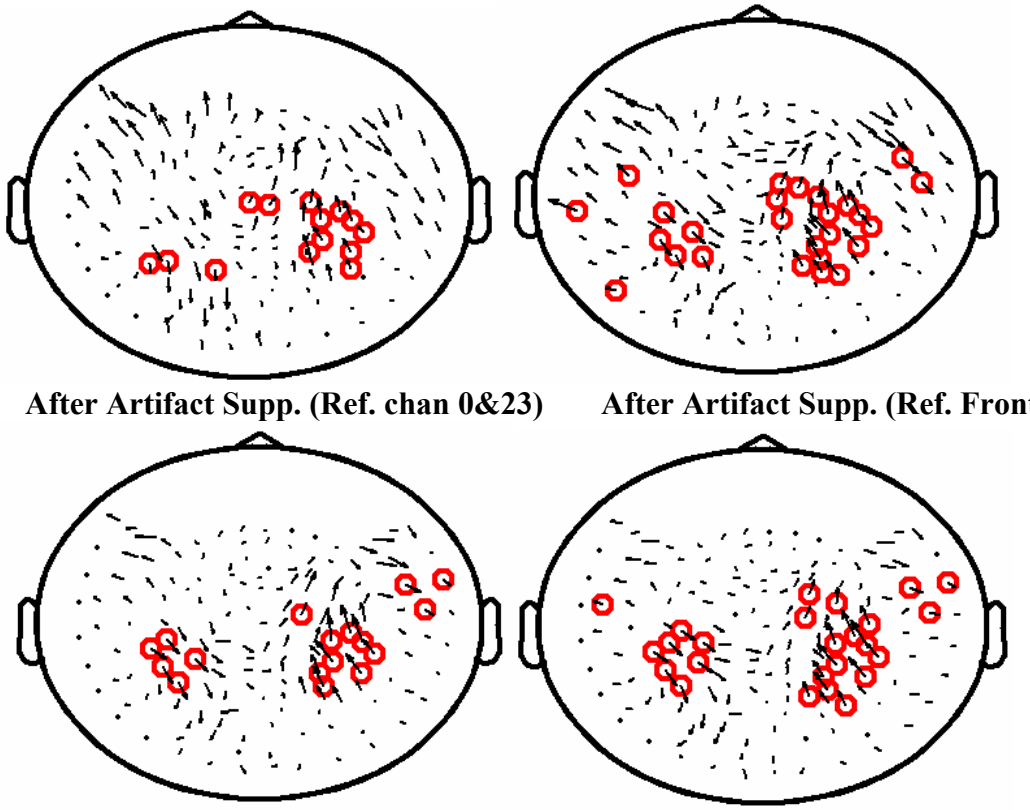

Figure 5.3 Head map plot for stimulus and response at 1.5Hz, an extremely noisy band. (Left top) Raw signal, 13 significant channels, poorly structured with no clear dipole to fit to. (Right top) Environmental noise is suppressed. The left hemisphere channels labeled significant lack the structure of a dipole; peripheral signals are strong. (Left bottom) After removing correlation with channels 0 and 23, which are rich in artifact noise, note that strong signals on the frontal side of the head are wiped out. In addition, what was labeled as significant on the left hemisphere is dropped out. On the other hand, the dipole on the right hemisphere is enhanced. (Right bottom) We remove correlation with frontal channels 101 to 107, rich in both artifact and neural background noise. The dipole on the left hemisphere is recovered, right hemisphere dipole is enhanced further.

To show the effect of such de-noising in the Fourier domain, see Figures 5.4 and 5.5. In the first figure, an already significant signal is minimally reduced while removing noise at 9, 17, and 31Hz in addition to other bands. In the other case

(Figure 5.5), the stimulus response signal was already buried in noise. The de-noising operation reduced power to all frequencies below 40 Hz. However, more noise was removed, and significant channels reemerged.

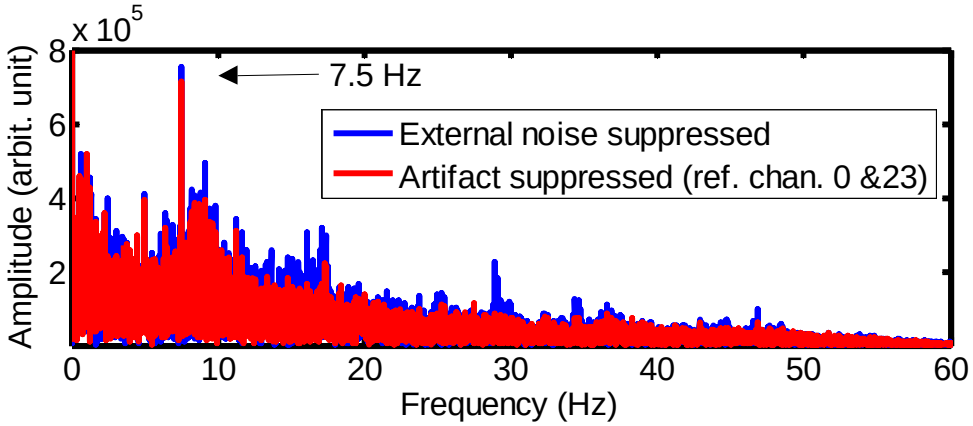

Figure 5.4: Artifact removal based on channels 0 and 23. Stimuli modulation frequency is 7.5 Hz. Peaks at 9, 17, and 31 Hz are suppressed. The signal is mildly affected. On the other hand, the 5Hz peak is not removed; further de-noising could yield better SNR.

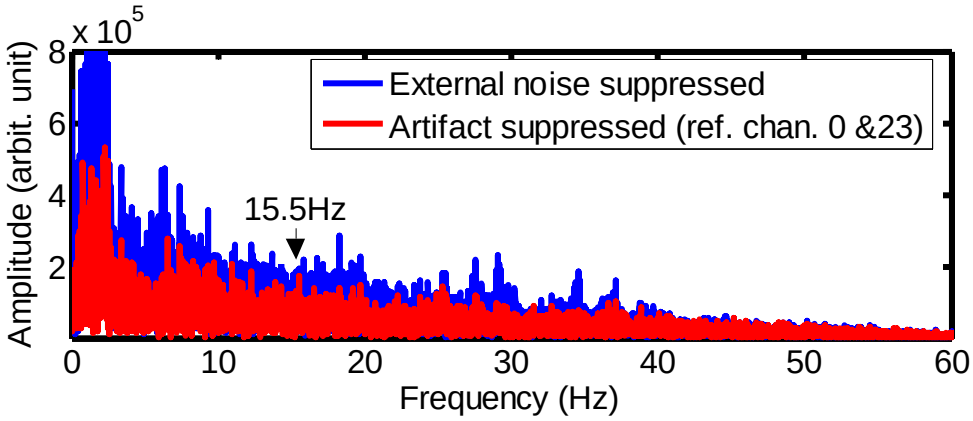

Figure 5.5: Removing correlation based on channels 0 and 23 for a stimulus modulated at 15.5Hz. The response was buried until we removed enough noise, then it emerged again. For most of the spectrum, the magnitude of the signal dropped, but unevenly, concentrated in narrow bands. In this case, a mixture of artifact and neural noise generated in the brain is suppressed.

The time series waveform in Figure 5.6 shows the effect of removing artifacts based on channels 0 and 23. We are plotting the difference between two neighboring channels to our reference channels, one on each side (before and after applying Fast LMS). The de-noising removed what looks like heartbeat in our signal and some other burst (at ~82sec.) that could be associated with eye movement.

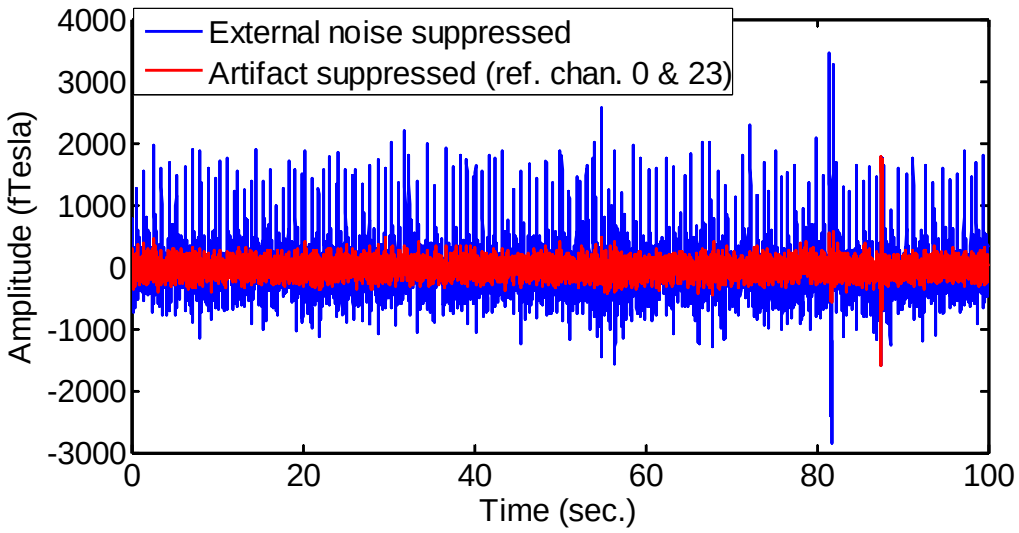

Figure5.6 Waveform of 50 concatenated responses at 31.5Hz for the difference between channels 1 and 22 before and after removing artifacts correlated with the two neighboring reference channels 0 and 23. Note the heartbeat and eye movement were filtered out. This operation was even consistently removing artifact from most channels on the peripheral contour map.

In some cases, valuable signal that exists in a potential reference channel could be filtered out from other channels. So far, we argue that the auditory contribution to this channel is very small. Nevertheless, we do not want to de-correlate such information from other neural channels. Indeed, it is possible to do the de-correlation with minimal or no loss to the signal, at the price of introducing a new filtering stage. First, we choose multi-reference channels that are corrupted with noise, such as channels 0

and 23. Instead of de-correlating the two, we look for common ground by correlating them and estimating the shared noise. This new constructed signal will be our new reference channel. It contains only noise information. It is very uncommon that an auditory signal from the same source will be significant in two distant channels with the same phase information. In other words, this new method will maximize noise presence in the new reference channel and suppress unidentified signals with different characteristics among multi-reference channels. This new reference channel is then cleared to be de-correlated from all other neural channels. In Figure 5.7, we plot the time waveform of the channel neighboring channel 0 (blue), then we contrast it with a filtered version (red) that is the result of de-correlating it with a reference channel constructed as we described earlier (reference channel is the filtered version of correlating channels 0 and 23). Note the suppression of heartbeat and other artifacts.

It is important to state that we cannot distinguish between biological artifact noise and neural brain background noise. Some channels, like 0 and 23, are more prone to artifacts because of their spatial location, but they could also be a source of brain noise. We separate de-noising artifact from suppressing brain noise to simplify the problem and divide it into 2 stages: In the first stage, we look for channels with mostly artifact, and in the second, we focus more on channels that could potentially be used to filter neural brain noise. The localized channels are used as references. It is evident that each stage could remove any biological noise as all 157 channels are in the magnetic field of such noise. Our goal is achieved when SNR improves regardless of the identity of the noise removed.

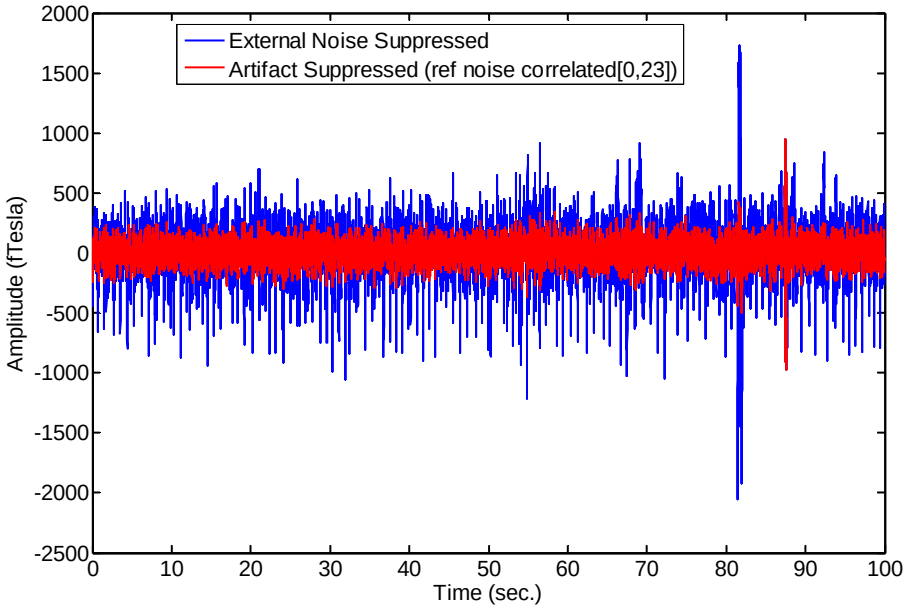

Figure 5.7 Waveform of 50 concatenated responses at 31.5Hz for channel 1 (neighbor of channel 0) before and after removing artifacts correlated with the reference channel constructed by capturing noise in channel 0 present in channel 23, i.e. the reference channel is correlated noise between channels 0 and 23. Note heartbeat and eye movement were filtered out. This operation is noninvasive to our signal because it is based on correlated noise between two channels on opposite sides of the hemisphere.

Finding a reference channel that captures noise and poses minimal threat to our signal is crucial for de-noising biological signals. We identify four criteria to find a reference channel in order to remove neural noise from other auditory channels:

1- Distance: channels should be apart from each other,

2- Significance: choose channels that are least significant, i.e., strong in noise, to subtract,

3- Correlation: The signal channels should have some correlation with the channel to de-noise,

4- Masking: Exclude those channels where we know that there might be auditory activity. For example, we look for reference channels in the visual cortex, frontal, left and right contour, and midline (Figure 5.1).

A case study: searching for reference channels

Here, we propose a method for finding noise reference channels. They are used to de-correlate from potential auditory neural channels. We plot in Figure 5.8 a correlation head map for a single channel. We identify the most significant channel with F-statistics (blue), and then we compute the correlation for this channel with all other channels.

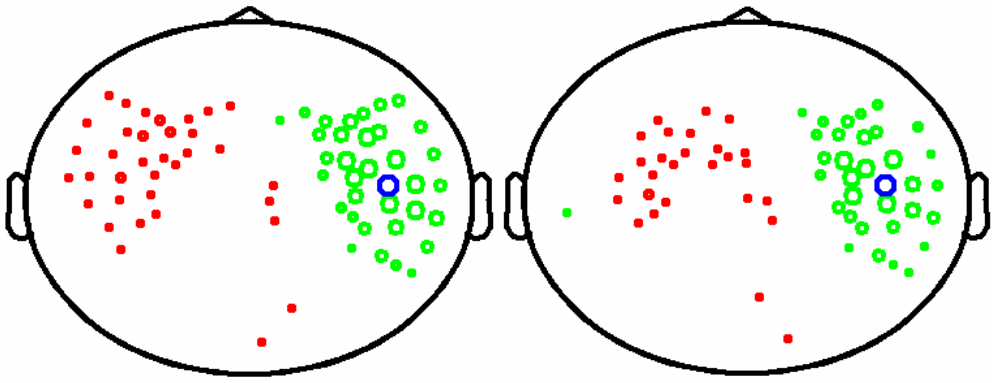

Figure 5.8 Correlation head map for channel 66 (most significant according to F statistics circled in blue) with all other channels. Positive (green) and negative (red) correlation with all other channels is computed. Larger circles correspond to higher correlation (only correlations above certain absolute threshold are plotted). Numbers indicate order statistics of correlation. One is most positively correlated, while 157 is most negatively correlated. (Left) Before applying artifact de-noising, with respect to channels 0 and 23. (Right) After applying artifact denoising. Many of the peripheral channels are de-correlated with our most significant channel. The cluster of channels that represent a dipole relocate to a more meaningful location. Note also the single outlier uncovered on the left hemisphere that is positively correlated with our channel on the right hemisphere. This could be due to artifacts.

Positive correlations are in green, and negative correlations are in red. After removing much of the peripheral noise with reference channels 0 and 23, the cluster of red

channels moves to the right, leaving an outlier close to the left ear positively correlated with blue channel.

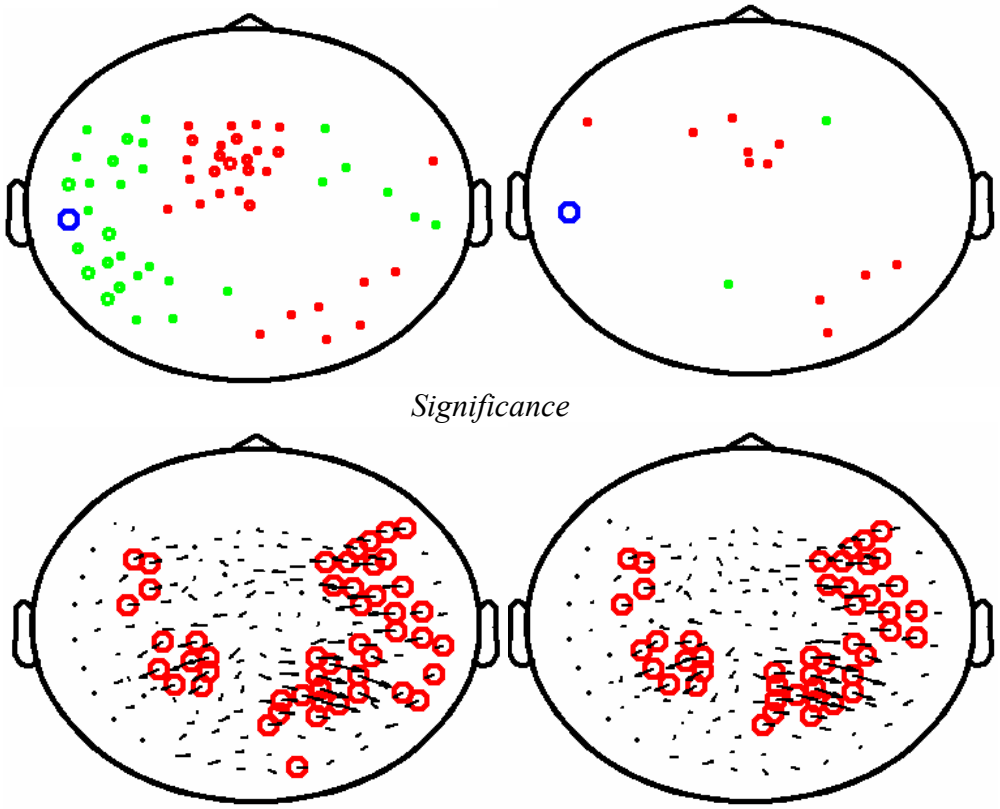

Figure5.9 (Top left) Correlation of channel 4 (blue) after de-noising based on channels 0 and 23. The correlation head map reveals that it could be rich in noise. (Top right) We suppress the correlation between channel 4 and all other channels. (Bottom left), Stimulus at 31.5Hz. The right hemisphere dipole is over fitted with noise prone signals labeled as significant. (Bottom right) After removing correlation with channel 4, the left hemisphere is enhanced, while the right hemisphere is more consistent after dropping outliers. Correlation threshold =0.2.

We pursue this channel further by plotting the correlation head map of this outlier with all other channels (Figure 5.9, top left) next to a significance head map (Figure 5.9, bottom left). The right hemisphere is overpopulated with significant channels that put constraints on a dipole fit due to the channels close to the right ear; their orientation is not in harmony with most other significant channels that are 180

degrees out of phase. In other words, we have a multi-dipole source. We are only interested in the dipole due to our stimulus. We label our outlier as a reference channel and de-correlate it from all other channels. Figure 5.9, bottom right, shows that the right hemisphere is cleaner for a dipole fit, strong noise is suppressed on the right contour, and the left hemisphere is slightly enhanced in significance. This disproves that the outlier was mostly measuring a neural signal.

Figure 5.10 is a time waveform plot for a channel neighboring our outlier. It appears that some artifacts not correlated with channels 0 and 23 are removed.

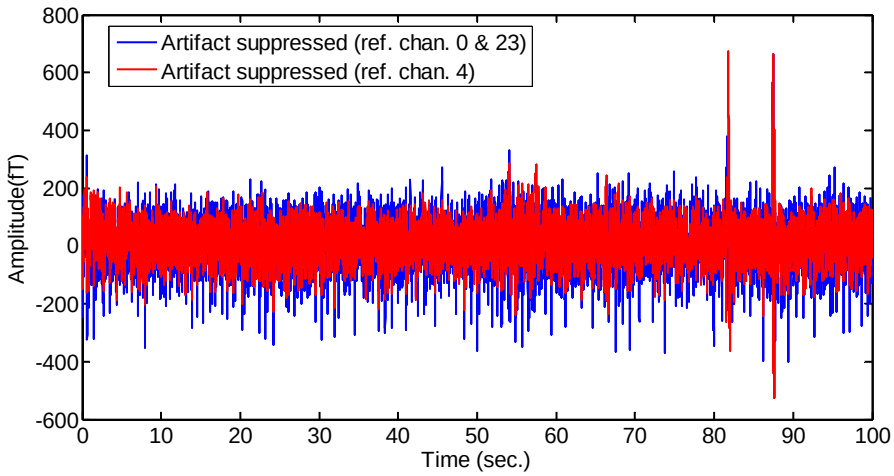

Figure 5.10 We remove correlation using channel 4 as a reference. The adjacent channel is plotted before and after noise suppression; it shows some artifact remaining from the previous de-noising process being removed.

Brain Background Noise Removal

If we are willing to sacrifice some of the strength of our auditory signal, it is possible to suppress much interference from other sources of non-auditory neural signal. We can compute a covariance matrix for all 157 channels and look at those with high correlation coefficients. We already know that two sources can be separated

only if the inter-source distance is at least of the same order of magnitude as the distance between sources and measurement locations [13]. Therefore, we can pick far distant channels. It is preferable to choose the most significant channels (those with the highest F-statistics), whether using order statistics for F-distribution based on SSR responses or using the most prominent M100 (in power) waveform channel.

Figure 5.11 shows the spectrum using frontal channels to de-noise our signal. Our auditory signal is uncompromised; noise removal was minimal in general, but of higher value at lower frequencies around 4 Hz.

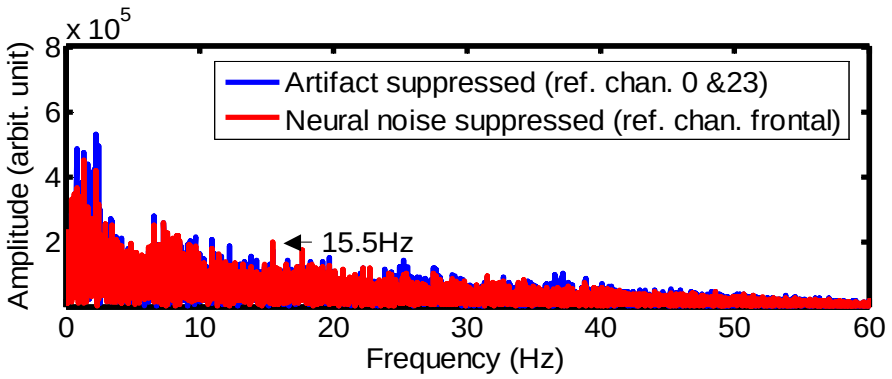

Figure 5.11: Removing correlation using frontal channels as references for biological noise for stimulus frequency at 15.5Hz. Minor improvements are recorded at low frequencies. Head map (not shown) revealed that this channel was not significant prior to de-correlation with frontal channels. This is the result of a drop in threshold level due to de-noising rather than an increase in signal power that remains the same (the channel is a neighbor of the channel in Figure 5.5; the latter was already significant before de-noising with frontal channels).

Regardless the four criteria, it is still possible that the reference channel we find has some correlation with the auditory signal, and so we risk subtracting some valuable signal. As in the case of artifact removal, we construct a reference channel that is an

estimate of the noise shared by multiple references rich in background noise. This will minimize signal loss.

Finally, a comparison of all de-noising stages using ROC curves (Figure 5.12) is consistent for all stimuli frequencies. Except for the 1.5Hz, which was very unstable for this particular subject, each stage increases in detection compared to the previous one. At high frequencies, the signal is already significant; hence, no valuable improvement is recorded.

Again, the results we present here are at an early stage of research; the tools and methods introduced in this chapter have great potential but need further refinement and investigation with more subjects.

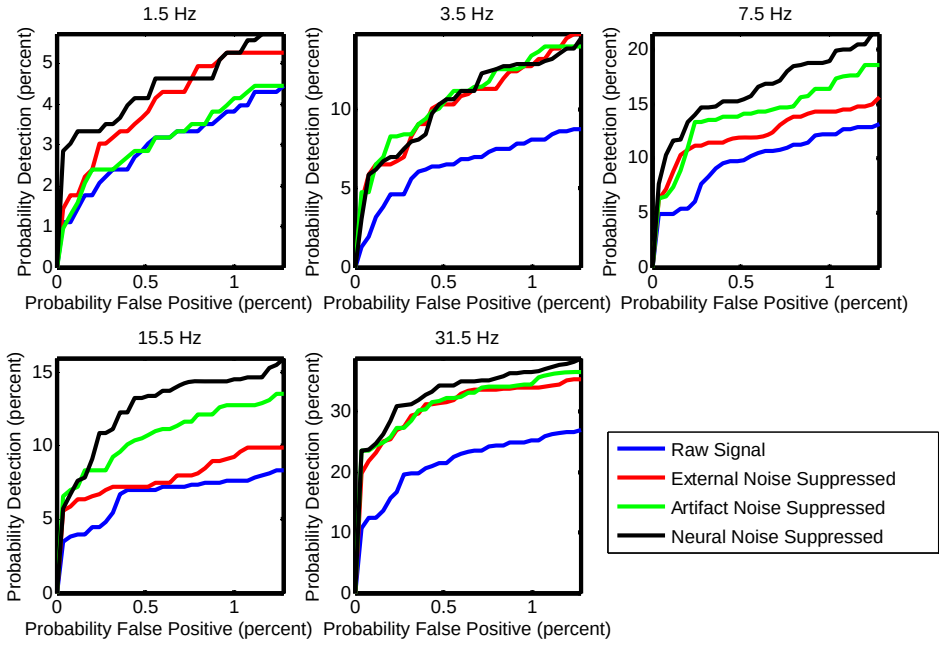

Figure 5.12 ROC curves comparison for all suppression stages for all modulation frequencies. In general, each de-noising process added some improvements to the detection of significance channels. The higher the frequency, the less improvement we get from suppressing biological noise. For stimuli at 1.5 Hz, it is very noisy and detection is very narrow, and hence one should avoid drawing any conclusion at that rate based on this single subject.

Conclusion

We have designed two powerful building blocks that form the foundation of a noise suppression model that addresses environmental noise, biological artifacts, and non-auditory brain activities. The first algorithm was a frequency domain adaptive filtering that exploits fast FFT and fast correlation to build an efficient de-noising tool. The second algorithm is a significance test that classifies every channel as a true positive or a false positive. As a result, we improved SNR by removing external noise as a first step. Various tests we conducted validated the universal application of the tools developed; however, these tests were spectral, not temporal.

Later, we expanded our de-noising to include artifacts and neural noise generated in the brain. Our results show that it is feasible to suppress such noise by de-correlating with respect to a noisy available or constructed reference channel. We constructed such a reference channel by estimating correlated noise in channels most susceptible to such noise. In the case of artifacts, because of the order of magnitude of the noise in comparison to our signal, our auditory signal is still intact, especially when using the constructed reference channel, although there is always a slight risk of compromise. In the case of background noise generated in the brain, the risk of compromise is higher because of the proximity of the auditory source and the source of noise; however, by respecting the 4 criteria described (significance, correlation, distance, and masking), the risk is reduced. Although we present de-noising modules as separate, there is crisscrossing between them. It is not necessary, though it is desirable, to separate different kinds of noise.

The purpose of the last chapter was to explore and introduce some potential techniques in suppressing biological noise. These novel, yet powerful tools still need to be investigated, especially in terms of replicating our findings in more subjects and turning the subjective solution into a general one.

In conclusion, by applying our de-noising algorithm, we improved SNR to a level where other techniques, such as ICA, are more capable of separating auditory components.

Appendices

A- Matlab code

A.1 Fast LMS

```
function filtinfo =
adapt_noise_supp_fix(srce,dest,blksize,num_blk_cte,forgetfact,adaptcte)
% filtinfo = adapt_noise_supp(srce, dest, blksize, num_blk_cte, forgetfact, adaptcte)
%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This file reads in from a sqd file: 157 neuronal channels, 3 reference channels,
% 32 trigger channels (and all other information). It filters the neuronal channels
% based on the 3 reference channels, then creates a new sqd file with the filtered
% neuronal data and everything else unmodified.
% The method is the Fast LMS adaptive filter.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
% srce : raw sqd file (if not specified, a sqd file is requested)
% dest : filtered sqd file (srce + '-filtered.sqd' is the default)
% blksize : size of LMS block in samples; typical: 64, 128(default), 256, 512
% num_blk_cte : number of blocks (weighted by sampling frequency in kHz) to filter
at
%       once (reduce if 'out of memory' errors occur); typical: 0.25, 0.5, 1(default), 2,
4
% forgetfact : forgetting factor,  0 < gamma < 1, (0.94 default)
% adaptcte : adaptation constant, 0 < alpha < 0.5 (0.01 default)
% filtinfo = filtering parameters and times
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Author: Nayef Ahmar aenayef@glue.umd.edu
%
% for the Computational Sensorimotor Systems Lab (CSSL) UMCP
% http://www.isr.umd.edu/Labs/CSSL/
% Version 0.91 Dec. 1 2005
%
% Latest version at
% http://www.isr.umd.edu/Labs/CSSL/simonlab/resources/resources.html
%
if nargin == 0,
   [fn, pn] = uigetfile('*.sqd','Select your SQD file source');
   if sum(fn==0)&&sum(pn==0),return,end
   srce = fullfile(pn,fn);
```

```
end
info = sqdread(srce,'info');%extract info
sf = info.SampleRate;%sampling frequency in HZ

if nargin < 6, adaptcte = 0.01; end;
if nargin < 5, forgetfact = 0.94; end;
if nargin < 4, num_blk_cte = 500/sf; end;
if nargin < 3, blksize = 2^ceil(log2(2*sf/8));end %128;
if nargin < 2, [pathstr,name,ext] = fileparts(srce); dest = fullfile(pathstr,[name '-
filtered' ext ]); end;

if exist(dest,'file'),delete(dest),end %delete any existing file because we use append
%command in sqdwrite
t0 = clock;
num_samples = info.SamplesAvailable;%total number of samples
tot_num_chan = info.ChannelCount;
%create a destination sqd file and write first sample
data = sqdread(srce,'Channels',[0:tot_num_chan-1],'Samples',[1 1]);
sqdwrite(srce,dest,data);

%Input parameters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
alpha = adaptcte;%0.01;%0.005;%adaptation constant
gamma = forgetfact;%0.94;%forgetting factor
M = blksize;%64;%Block size
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
cte = floor(1000*64/blksize);
num_blks = num_blk_cte*cte;%1000;%reduce if have out of memory problems
N = floor((num_samples-1)/M/num_blks);%number of blocks to pass to fast_lms,
should %be integer
blk_len = M*num_blks;

infoc.sf = sf;
infoc.blocksze = M;
infoc.forgetfact = gamma;
infoc.adaptcte = alpha;
infoc.multiblk = blk_len;
infoc.source = srce;
infoc.destin = dest;
%initialize P and W
P_chan = gamma*ones(2*M,3,157);
W_chan = zeros(2*M,3,157);
U = ones(M,3);%@for in between block transition update
h = waitbar(0,'Please wait...');

for blk = 1:N,
```

```
    blk_lb = (blk-1)*blk_len + 2;%start at 2 because we already copied first sample
    blk_ub = blk*blk_len + 1;
    data = sqdread(srce,'Channels',[0:tot_num_chan-1],'Samples',[blk_lb blk_ub]);
    data_chan = data;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %Warnings%@
    if max(abs(mean(data_chan(:,1:160)))) > 1000,%need to verify this threshold
        display('Warning: For better filtering, remove Dc before applying noise
suppression');
    end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    u = data_chan(:,158:160);
    for chan = 1:157,
        P = P_chan(:,:,chan);
        W = W_chan(:,:,chan);
        d = data_chan(:,chan);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        [E,p,w]=fastlms3ref(alpha,M,u,d,gamma,P,W,U);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        % Input arguments:
        % alpha =step size, dim 1x1
        % M =filter length, dim 1x1
        % u =input signal, dim Nx3
        % d =desired signal, dim Nx1
        % gamma =forgetting factor, dim 1x1
        % P =initial value, energy, dim 2Mx3
        % W =final filter vector from previous iteration, dim 2*Mx3 only last half is
used
        % U = previous input reference channel from previous multi-block, dim Mx3
        %
        % Output arguments:
        % e =estimation error, or filtered signal dim Nx1
        % P =output value, energy, dim 2Mx3
        % w =final filter vector, dim 2*Mx3 only last half is used
        data_chan(:,chan) = E;
        P_chan(:,:,chan) = p;
        W_chan(:,:,chan) = w;
    end %chan
    U = u(end-M+1:end,:);%@for in between block transition
    %sqdwrite goes here
    putdata(info,dest,'Action','Append','Data',data_chan);
    waitbar(blk/N,h);
end %blk
%append last fraction of a block if there is any
if blk_ub < num_samples,
```

```
    data = sqdread(srce,'Channels',[0:tot_num_chan-1],'Samples',[blk_ub+1
num_samples]);
    num_samp_fract = num_samples - blk_ub;
    num_app_zeros = blk_len - num_samp_fract;
    data_chan = [data; zeros(num_app_zeros,tot_num_chan)];
    u = data_chan(:,158:160);
    for chan = 1:157,
        P = P_chan(:,:,chan);
        W = W_chan(:,:,chan);
        d = data_chan(:,chan);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        [E,p,w]=fastlms3ref(alpha,M,u,d,gamma,P,W,U);
        data_chan_f(:,chan) = E(1:num_samp_fract,:);
    end %chan
    %sqdwrite goes here
    data_chan_f(:,158:160) = u(1:num_samp_fract,:);
    data_chan_f(:,161:tot_num_chan) = data(1:num_samp_fract,161:tot_num_chan);
    putdata(info,dest,'Action','Append','Data',data_chan_f);
end
close(h);
deltat = etime(clock,t0)
infoc.processtime = deltat;
filtinfo = infoc;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [E,p,w]=fastlms3ref(alpha,M,u,d,gamma,P,W,U);
% [e,p,w]=fastlms_3ref(alpha,M,u,d,gamma,P,W,U);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Input arguments:
% alpha =step size, dim 1x1
% M =filter length, dim 1x1
% u =input signal, dim Nx3
% d =desired signal, dim Nx1
% gamma =forgetting factor, dim 1x1
% P =initial value, energy, dim 2Mx3
% W =final filter vector from previous iteration, dim 2*Mx3
% U = previous input reference channel from previous multi-block, dim Mx3
%
% Output arguments:
% e =estimation error, or filtered signal dim Nx1
% P =output value, energy, dim 2Mx3
% w =final filter vector, dim 2*Mx3 only last half is used
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% initialization
N=length(u);
E=d;%Initially, the recovered signal is set to observed signal
zero1 = zeros(M,1);
```

```
zero3 = zeros(M,3);
twoM = 2*M;
% no.of blocks
Blocks=N/M;
ref = [1:3];
for k=0:Blocks-1
   if k>0,
      Uvec=fft([u((k-1)*M+1:(k+1)*M,ref)],twoM);
   else%in between multiblock
      Uvec = fft([U;u(1:M,:)]);%concatenate last block from previous multiblock, and
first block from current multiblock
   end
   yvec=real(ifft(Uvec(:,ref).*W(:,ref)));
   yvec2=yvec(M+1:twoM,ref);
   % block k; desired and error signal
   dvec=d(k*M+1:(k+1)*M);
   E(k*M+1:(k+1)*M,1)=dvec-sum(yvec2(:,ref),2);
   % FFT of estimation error
   Evec=fft([zero1;E(k*M+1:(k+1)*M)],twoM);
   % estimated power
   P=gamma*P(:,ref)+(1-gamma)*abs(Uvec(:,ref)).^2;
   % block k, inverse of power
   Dvec=1./P(:,ref);
   % estimated gradient
   phivec=ifft(Dvec(:,ref).*conj(Uvec(:,ref)).*[Evec Evec Evec],twoM);
   phivec=phivec(1:M,ref);
   % update of weights of filter coefficients
   W=W+alpha*fft([phivec;zero3],twoM);
end % no.of blocks
w=W;%save filter coefficient(fourier domain) to use in the next multiblock
p = P;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

A.2 F and Hotelling Significance test

```
function [sig_cell,sig_cell_tot] =
FH_sig_tests(srce,trigdirname,sig_test,p_indx,latency,stimlist);
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
%Significance test for F-test and Hotelling test
%The output is a cell of significant channels for all trigger files
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%srce : raw sqd file (if not specified, a sqd file is requested)
```

```
%trigdirname: path for trigger files
%sig_test: 1: F-test, 2: Hotelling test, 3: both F and Hotelling test
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if nargin == 0,
   [fn, pn] = uigetfile('*.sqd','Select SQD file source');
   if sum(fn==0)&&sum(pn==0),return,end
   srce = fullfile(pn,fn);
end
if nargin < 6, stimlist = [1.5 1.5 1.5 1.5 3.5 3.5 3.5 3.5 7.5 7.5 7.5 7.5 15.5 15.5 15.5
15.5 31.5 31.5 31.5 31.5];end ;
if nargin < 5, latency = 0.3;end ;%steady state response
if nargin < 4, p_indx = 1;end ;% number of false positives on average
if nargin < 3, sig_test = 1;end ;
if nargin < 2, trigdirname = uigetdir( 'Pick a Directory for trigger files'); end;
info = sqdread(srce,'info');%extract info
sf = info.SampleRate;%sampling frequency in HZ
t0 = clock;
num_stim = length(stimlist);
freqlist = unique(stimlist);
numfreq = length(freqlist);
num_false_pve = p_indx*16;
test_trig_sam_cube = zeros(2,num_stim,numfreq,157); %@initialize cube of all data
inc = 1;%matlab start at zero hence add one
pretrigger = 0;
posttrigger = 0;
stim_duration = 2;%stimulus duration in seconds
len = length(textread([trigdirname,'/trig_160.txt']));
epoc_samples = stim_duration*sf;
channel = 1:157;
fftw('planner','patient');%Choose fastest FFT
h = waitbar(0,'Please wait...');
for trigch = 160:179,%[160:175,178:181],
   triggerfile = [trigdirname,'/trig_',int2str(trigch),'.txt'];
   trigger = textread(triggerfile);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
   %Extract and concatenate sample paths of filtered data for all channels
   %maxtrig =  trigger(len)*sf + latency*sf + epoc_samples + posttrigger*sf ;
   %data_stim = zeros(1:maxtrig,157);
   data_stim = [];
   for j = 1:len,
      lowtrig = round(trigger(j)*sf - pretrigger*sf + latency*sf + 1) ;%trigger -
(pretrigger) + 1(matlab starts at 1 vs meg160 at 0)
      hightrig = round(lowtrig + pretrigger*sf + epoc_samples + posttrigger*sf -1);%
(pretrigger) + 1s recorded data +  (posttrigger)
      datain = getdata(info,'Channels',0:156,'Samples',[lowtrig hightrig]);
      data_stim = [data_stim; datain];
```

```
    end
    for sami = 1:numfreq,
        sam = freqlist(sami);%[1.5,3.5,7.5,15.5,31.5],
        amf = sam*stim_duration;
        am_freq = amf*len + inc;

        for chan = channel,
            data_ch = data_stim(:,chan);

            if sig_test ~= 2,
                fft_data_ch = fft(data_ch);
                fft_data_ch(1) = 0;
                fft_data_amf(chan) = fft_data_ch(am_freq,1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                %F test for hidden periodicity(Simplified)
                denom_f = sum(abs(fft_data_ch(am_freq-60:am_freq-
1)).^2)+sum(abs(fft_data_ch(am_freq+1:am_freq+60)).^2);
                num_f = 120*abs(fft_data_ch(am_freq))^2;
                R_f_test = num_f/denom_f;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                R_f_test_chlist(chan,:) = R_f_test;
            end
            if sig_test ~= 1,
                data_ch = data_ch - mean(data_ch);%
                %Hotelling
                for k = 1:len,
                    lb = (k-1)*epoc_samples + 1;
                    ub = k*epoc_samples;
                    datafft_1pr = fft(data_ch(lb:ub));
                    sam_complex_50(k) = datafft_1pr(amf+inc);
                end
                sam_50epoch = sam_complex_50;
                sam_50epoch = sam_50epoch(:);
                sam_realandimg(:,1) = real(sam_50epoch);
                sam_realandimg(:,2) = imag(sam_50epoch);
                [nr,nc] = size(sam_realandimg);
                mu0 = [0;0];
                sam_realandimg_mean=mean(sam_realandimg);
                dev = sam_realandimg-kron(ones(nr,1),sam_realandimg_mean);
                s=dev'*dev/(nr-1);
                sinv=inv(s);
                wrk = sam_realandimg_mean'-mu0;
                tst = nr*wrk'*sinv*wrk;
                R_hot = tst*(nr-nc)/(nc*(nr-1));
                R_hotelling_chlist(chan,:) = R_hot;
            end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        end%channel
        %To extract all possible values of all scenarios(#trig, #sam)
        trigi = trigch-159;%@
        if sig_test ~= 2, test_trig_sam_cube(1,trigi,sami,:)= R_f_test_chlist; end
%@%F-test
        if sig_test ~= 1, test_trig_sam_cube(2,trigi,sami,:)= R_hotelling_chlist; end
%@%Hotelling
    end%for sam
    trigch
    waitbar((trigch-159)/num_stim,h);
end%for trigch
sig_cell = cell(num_stim,2);
sig_cell_tot = cell(numfreq,num_stim,2);
for trig_in = 1:num_stim,%5:8,%1:20,
    for stimf_in = 1:numfreq,%2,%1:5,
        [freq_bnd,bound] = find(freqlist(stimf_in)~=stimlist);
        %F-Test
        if sig_test ~= 2,
            flist_fp = test_trig_sam_cube(1,bound,stimf_in,:);
            flist_fp = sort(flist_fp(:));
            f_thresh = flist_fp(length(flist_fp)-num_false_pve+1);
            flist_tp = test_trig_sam_cube(1,trig_in,stimf_in,:);
            [flist_sig,ind_F] = find(flist_tp>f_thresh);
            roc_f(num_false_pve,trig_in,stimf_in) = length(flist_sig);
            sig_cell_tot(stimf_in,trig_in,1) = {[ind_F]};
        end
        %Hotelling
        if sig_test ~= 1,
            hlist_fp = test_trig_sam_cube(2,bound,stimf_in,:);
            hlist_fp = sort(hlist_fp(:));
            h_thresh = hlist_fp(length(hlist_fp)-num_false_pve+1);
            hlist_tp = test_trig_sam_cube(2,trig_in,stimf_in,:);
            [hlist_sig,ind_H] = find(hlist_tp>h_thresh);
            roc_h(num_false_pve,trig_in,stimf_in) = length(hlist_sig);
            sig_cell_tot(stimf_in,trig_in,2) = {[ind_H]};
        end
    end
    [aa,stim_ind] = find(freqlist == stimlist(trig_in));
    sig_cell(trig_in,1) = sig_cell_tot(stim_ind,trig_in,1);
    sig_cell(trig_in,2) = sig_cell_tot(stim_ind,trig_in,2);
end
sig_cell
sig_cell_tot
deltat = etime(clock,t0)
close(h);
```

B- Table summarizing Fast Block LMS algorithm

Dimensions:
r=0,…,R ; reference channels, e.g. R = 3.
M; block size (e.g. 1024 samples)
i=0,…,2M-1

Initialization:

$\hat{W}_r(0) = zeros(2M, R)$; *Filter coefficients initialized to zero*

$P_{i,r}(0) = \delta_i$; *average signal power per Reference channel, initialized to small positive constant* δ .

Computation: For each block of *M* input samples:
Filtering:

$$U_r(k) = diag\left\{FFT[u_r(kM-M),...,u_r(kM-1),u_r(kM),...,u_r(kM+M-1)]^T\right\}$$

$y_r^T(k)$ = *last M elements of IFFT*$[U_r(k)\hat{W}_r(k)]$

Error estimation:

$$e(k) = d(k) - \sum_{r=1}^{R} y_r(k)$$

$$E(k) = FFT\begin{bmatrix} 0 \\ e(k) \end{bmatrix}$$

Signal-power estimation:

$$P_{i,r}(k) = \gamma P_{i,r}(k-1) + (1-\gamma)\left|U_{i,r}(k)\right|^2$$

$$D(k) = diag[P_0^{-1}(k), P_1^{-1}(k), ..., P_{2M-1}^{-1}(k)]$$

Tap-weight adaptation:

$\Phi_r(k)$ = first *M* elements of $IFFT[D(k)U_r^H(k)E(k)]$

$$\hat{W}_r(k+1) = \hat{W}_r(k) + \alpha FFT\begin{bmatrix} \Phi(k) \\ 0 \end{bmatrix}$$

FFT : Fast Fourier Transform, IFFT: Inverse Fourier Transform,
α *: adaptation constant <1/2*

Table b.1: Multi-reference Fast LMS adopted from [10] and modified for multiple references. The algorithm operates in the Fourier domain by slicing the spectrum in small block with slowly changing envelope. It takes advantage of the efficient implementation of convolution and correlation.

C- Significance Tests

C.1 Rayleigh's Phase Coherence Test

For each of the 2 seconds responses per stimulus (there are N=50 presentations), a DFT was performed and the phase at the stimulus frequency was measured. Then the projections onto the real and imaginary axes are summed separately for all presentations. The phase coherence, denoted *R*, ranges between zero and one where zero is uniformly random and 1 is most significant [4.10, 4.5, 4.12, and 4.2]. The phase coherence is formally:

$$R_P = \frac{1}{N}\sqrt{(\sum_{i=1}^{N}\cos\theta_i)^2 + (\sum_{i=1}^{N}\sin\theta_i)^2} \quad \text{(c.1)}$$

The significance of the result was assessed using approximation formula suggested by [6]:

$$P = \mathrm{e}^{-NR_P{}^2} \quad \text{(c.2)}$$

C.2 Multitaper DPSS

Based on Karhun loeve expansion, a multitaper method uses windows from the discrete prolate spheroid sequences (DPSS). It is also used to detect sinusoids embedded in noise based on their amplitude [17]. It is very similar to the F-test, but because of the nature of the DPSS windows, it averages over neighboring frequency bins. For our data, it has less power than the simpler F-test. This is consistent with our experimental design, which puts all the power of the signal into a single frequency bin, with no spectral splatter or frequency widening. In this design, smoothing in the

frequency domain serves little purpose and only allows additional noise into the signal's frequency bin.

C3. Hotelling's T^2 distribution

Hotelling's T^2 statistic is a generalization of Student's t statistic that is used in multivariate hypothesis testing. It is defined as follows:

For X_1 to X_n px1 column vector let $\overline{X}$ be the sample mean, and W the sample variance,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{(c.3)}$$

$$W = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)^T \qquad \text{(c.4)}$$

Let μ be some known px1 column vector that is a hypothesized value of a mean. To be compared with the sample mean. Then the Hotelling's T^2 statistic can be determined for any matrix of rank $\geq p$ and is defined as:

$$T^2 = n\left(\overline{X} - \mu\right)^T W^{-1}\left(\overline{X} - \mu\right) \qquad \text{(c.5)}$$

References

[1] Adachi Y., Shimogawara M., Higuchi M., Haruta Y., & Ochiai M., (2001). "Reduction of nonperiodical environmental magnetic noise in MEG measurement by continuously adjusted least squares method." IEEE Transactions on Applied Superconductivity, 11, 669–672.

[2] **Ahmar** N., Wang Y., Simon J., 'Significance tests for MEG response detection', International IEEE EMBS Conference on Neural Engineering 2005

[3] **Ahmar** N., Simon J.Z., Periodic Noise Suppression for MEG Signals UMCP-Tech2004

[4] **Ahmar** N., Simon J., 'MEG Adaptive noise suppression using fast-LMS', International IEEE EMBS Conference on Neural Engineering 2005

[5] Baillet S., Mosher J.C. and Leahy R. M., "Electromagnetic Brain Mapping", IEEE Signal Processing Magazine, Vol.18, No 6, pages: 14-30, 2001.

[6] Fisher NI. Statistical analysis of circular data, Cambridge, MA: Cambridge University Press, 1993.

[7] Fisher RA. Tests of significance in harmonic analysis. Proc R Soc Lond Ser A 1929;125:54-59.

[8] Hämäläinen, M., R. Hari, et al. (1993). "Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain." Rev Mod Physics 65(2): 413-497.

[9] Hari, R., Levanen, S., Raij, T, Timing of human cortical functions during cognition: role of MEG, Trends in cognitive neuroscience, vol. 4, no.12, December 2000.

[10] Haykin, S. S. (2002). Adaptive filter theory. Upper Saddle River, N.J., Prentice Hall.

[11] Hotelling H (1931). The generalization of student's ratio. Ann Math Statist

2:360-378.

[12] Kado, H., Higuchi, M., Shimogawara, M., Haruta, Y., Adachi, Y., Kawai, J., Ogata, H., & Uehara, G. (1999). Magnetoencephalogram systems developed at KIT. IEEE Transactions on Applied Superconductivity 9, 4057-62.

[13] Lutkenhoner B, Magnetoencephalography and its Achilles' heel. J Physiol Paris. 2003 Jul-Nov;97(4-6):641-58.

[14] Makeig S., Bell A.J., Jung T.-P., and Sejnowski T.-J.. Independent component analysis of electroencephalographic data. In Advances in Neural Information Processing Systems 8, pages 145-151. MIT Press, 1996.

[15] Manolakis D.G., Ingle V.K., and Kogon S. 2000. "Statistical and Adaptive Signal Processing: Spectral estimation, signal modeling, adaptive filtering, and array processing," McGraw Hill, New York., NY.

[16] Papoulis A., Probability, Random Variables, and Stochastic Processes, McGraw Hill, 2nd edition, 1984.

[17] Percival, D.B., and A.T. Walden. Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques. Cambridge: Cambridge University Press, 1993.

[18] Picton, T.W., Dimitrijevic, A., John, M.S. & Van Roon, P. 2001. The use of phase in the detection of auditory steady-state responses. Clin Neurophysiol, 112, 1698-1711.

[19] Rayleigh Lord. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. Philos Mag 1880;10:73-78.

[20] Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J Acoust Soc Am 108:679-691.

[21] Ross B, Pantev C.Auditory steady-state responses reveal amplitude modulation

gap detection thresholds.J Acoust Soc Am. 2004 May;115(5 Pt 1):2193-206.

[22] Simon J.Z., Wang Y., Poeppel D., Xiang J., Ahmar N. (2004), "MEG Steady State Responses To Auditory Stimuli Of Varying Complexity," Society for Neuroscience abstracts.

[23] Vigário R., Jousmäki V., Hämäläinen M., Hari R., and Oja E.. Independent component analysis for identification of artifacts in magnetoencephalographic recordings. In Advances in Neural Information Processing Systems 10, pages 229-235. MIT Press, 1998.

[24] Wang Y., **Ahmar** N., Xiang J., Ma L., Simon J. Complex Valued Equivalent-Current Dipoles Fitting for MEG Responses', International IEEE EMBS Conference on Neural Engineering 2005.

[25] Wang Y., **Ahmar** N., Xiang J., Poeppel D., Simon J., "Auditory Steady State Responses to Broadband Noise in Human Auditory Cortex", submitted to ARO 2005 MidWinter meeting.

[26] Wang Y., **Ahmar** N., Xiang J., Poeppel D., Simon J., "MEG Steady State Response to Broadband Sounds", BIOMAG 2004.

[27] Widrow B., "Adaptive Noise Cancelling: Principles and Applications," Proceedings of the IEEE, vol. 63, pp. 1692--1716, Dec. 1975.

[28] Widrow B. and Stearns S. D., "Adaptive Signal Processing," Prentice Hall, Englewood Cliffs, 1985.

[29] Xiang J, Wang Y, Simon JZ (2005) MEG Responses to Speech and Stimuli With Speechlike Modulations. In: International IEEE EMBS Conference on Neural Engineering 2005.