

2018

Automated High Frequency Oscillation Detection and Seizure Onset Zone Estimation Using TCRES

Michael Tamayo
University of Rhode Island, mickytamayo@gmail.com

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Terms of Use

All rights reserved under copyright.

Recommended Citation

Tamayo, Michael, "Automated High Frequency Oscillation Detection and Seizure Onset Zone Estimation Using TCRES" (2018). *Open Access Master's Theses*. Paper 1234.
<https://digitalcommons.uri.edu/theses/1234>

This Thesis is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

AUTOMATED HIGH FREQUENCY OSCILLATION DETECTION AND
SEIZURE ONSET ZONE ESTIMATION USING TCRES

BY

MICHAEL TAMAYO

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL ENGINEERING

UNIVERSITY OF RHODE ISLAND

2018

MASTER OF SCIENCE THESIS
OF
MICHAEL TAMAYO

APPROVED:

Thesis Committee:

Major Professor Walter Besio

Steven Kay

Gavino Puggioni

Yalda Shahriari

Nasser H. Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2018

ABSTRACT

High frequency oscillations (HFOs) have become a predominant topic in neurology research in recent years as they are considered markers of epileptic activity, potential indicators of seizure onset zone (SOZ), and possibly hint at an oncoming seizure in epileptic patients. These oscillatory signals are frequently seen in intracranial electroencephalography (iEEG) but have proven difficult to record from conventional scalp EEG electrodes. However, with the advent of tripolar concentric ring electrodes (TCREs) it is becoming easier to record HFOs without the need of a surgical procedure. While several different methods have been proposed for detection of HFOs in iEEG signals, no single method has been identified as a best detector and it is unknown how any of these methods will perform on the activity recorded from TCREs (tEEG). In this study, a novel detection design is derived and evaluated on real and simulated tEEG and compared to two of the more popular HFO detection routines designed for iEEG. Subsequently, an estimate of SOZ based strictly on HFO event occurrence is made on a few patients and compared to the markings of a clinician.

ACKNOWLEDGMENTS

I would like to thank several people who helped me with my thesis.

Thank you to all of the professors I have had in my many years at the University of Rhode Island. I would especially like to recognize my advisor, Dr. Walter Besio, and professor, Dr. Steven Kay, both of whom have an incredible ability to simplistically describe the most complex topics.

A big thank you to my family. My wife, April, encouraged me to continue at times I was ready to quit and allowed me to isolate myself at times I was willing to work. My sons, Elias and Henry, were born in the last year of my Master's work and made the effort greater but also far more enjoyable.

Thanks also goes to my supervisor Nitin Shali and the Research & Development department at Natus Medical for providing the funding and support for this research.

Thank you to the rest of my friends and family who helped proofread and provide guidance.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	viii
CHAPTER	
1 Introduction	1
1.1 Neurophysiological Signal Analysis	3
1.1.1 High Frequency Oscillations	6
1.1.2 Epileptic Spikes	7
1.2 HFO Detection	8
1.3 Statistical Signal Processing	10
1.3.1 Autoregressive Processes	11
1.3.2 Generalized Likelihood Ratio Test	13
1.3.3 Canonical Autoregressive Decomposition	14
List of References	15
2 Methodology	19
2.1 Detector Overview	19
2.2 Anomaly Detector	22
2.2.1 Overview	22

	Page
2.2.2 Derivation	23
2.2.3 Explanation	26
2.2.4 Expected Performance	28
2.2.5 Implementation	29
2.3 Sinusoid Detector	30
2.3.1 Overview	30
2.3.2 Derivation	31
2.3.3 Explanation	36
2.3.4 Expected Performance	38
2.3.5 Implementation	38
2.4 tEEG Simulator	40
2.5 Comparison Detectors	46
2.6 SOZ Estimation	46
List of References	48
3 Findings	50
3.1 Detector Performance	50
3.1.1 Simulated Data	50
3.1.2 Real Data	53
3.1.3 Timing Comparison	55
3.2 SOZ Estimation	56
List of References	57
4 Conclusion	62
List of References	66

BIBLIOGRAPHY 68

LIST OF FIGURES

Figure		Page
1	Detection Routine Block Diagram	21
2	Anomaly Detector Block Diagram	27
3	GLRT Theoretical Asymptotic Performance	29
4	Sinusoid Detector Block Diagram	37
5	10 seconds of Simulated TVAR tEEG	43
6	Simulated HFO and Spike Events	44
7	Simulated Alpha tEEG with Events	45
8	Simulated Gamma tEEG with Events	45
9	Detector Stage 1 Estimated ROC	51
10	Detector Stage 1 Estimated Probability of Detection	51
11	Detector Stage 2 Estimated ROC against H_0 background tEEG noise	54
12	Detector Stage 2 Estimated ROC against H_0 simulated spike . .	54
13	Detector Stage 2 Estimated Probability of Error	55

LIST OF TABLES

Table		Page
1	Detection Count on Real Data	55
2	Algorithm Run Time	56
3	Patient 1 SOZ Results, 400 Hz study	57
4	Patient 2 SOZ Results, 1600 Hz study	58
5	Patient 3 SOZ Results, 200 Hz study	59
6	Patient 4 SOZ Results, 200 Hz study	60
7	Patient 5 SOZ Results, 1600 Hz study	61

CHAPTER 1

Introduction

Epilepsy is a brain disorder characterized by recurrent seizures and affects an estimated 1% of the world's population [1]. Nerve cell communication in the brain is made up of electrical activity balanced by excitatory and inhibitory signals—diseased brain tissue can cause an imbalance of these neurotransmitter signals leading to a higher likelihood of seizures [2, 3]. Antiepileptic drugs are often prescribed to treat this imbalance but are only effective in controlling symptoms for 70% of patients [1, 2]. Those suffering from partial seizures—epileptic activity restricted to a certain area of the brain—can often be helped with a surgical procedure that removes the part of the cortex responsible for seizure propagation [2]. However, it can be difficult to find the location of such diseased brain tissue [4, 5] and it is not possible to determine if a resection of the epileptic zone will be successful in preventing future seizures before the operation has been performed [4].

Electroencephalography (EEG) is a monitoring of the brain's electrical activity from the scalp surface and is a routine practice for epileptic patients—it is often the first step in epilepsy diagnosis [2, 3]. While routine surface EEG (recording from electrodes placed directly on the scalp) is typically concerned with frequencies below 70 Hz [3], modern EEG research commonly uses intracranial or depth electrodes in order to record higher frequency signals that were previously difficult or impossible to find [3, 6, 7]. A craniotomy, open brain surgery, is required to place the intracranial electrodes. The advantage is better localization and less environmental noise than scalp recordings, yet determining the SOZ is still a challenge [4]. In epileptic cases where a lesionectomy is chosen as a remedy, multiple

surgical procedures consisting of iEEG recordings and functional brain mapping with a cortical stimulator are often needed [3, 4]. iEEG is associated with a risk to patients and only obtained when required [8]. Therefore, the ability to accurately detect neurophysiological signals of interest and localize areas responsible for seizure propagation without requiring additional complex procedures has the potential to revolutionize how advanced neurological studies are performed.

TCREs are a newer technology capable of recording signals with better spatial resolution than standard EEG by generating a Laplacian montage at a single scalp recording site [9, 10]. The appeal is the ability to investigate highly localized brain activity without requiring a surgical procedure. This technology applies to the study at hand as tripolar electroencephalography (tEEG) recordings have been used to find HFOs preceding seizure onset in epileptic patients [11, 12]. Pairing an effective HFO detection routine with TCREs has the potential to reduce the number of open-brain procedures necessary when intervening with high-risk epileptic patients.

The tradeoff is a higher background noise in tEEG than iEEG, making automatic signal detection more difficult. Therefore, using detectors that have been tried with success on iEEG data may not be a fruitful endeavor as iEEG has more physiological data contributing to the overall signal and less influence due to environmental noise. This brings us to the purpose of the study: to develop a detection method designed for the signal characteristics of tEEG and to compare to detectors in the literature that have had success evaluating the HFO data type in iEEG studies.

It should be noted that HFOs have been found in recordings from both healthy and epileptic brains, yet they are more commonly seen in patients with epilepsy and their frequency of occurrence appears to be correlated with ictal state [13, 14, 15, 5].

Recent studies have used HFOs recorded from iEEG to try and determine the SOZ [14, 15]. It is known that resection of brain areas with ripples correlates with a good postsurgical outcome when compared to those areas that do not generate oscillatory activity and several studies have shown the correlation between fast ripples and SOZ, though it appears most likely it is the rate of occurrence or channels with an abnormally high number of HFOs that indicate epileptic zones and not the individual characteristics of the events themselves [15, 16, 8, 17]. Detecting HFOs when recording using standard scalp electrodes has proven difficult but HFOs can occasionally be seen [5]. However, the typical amplitude of ripples detected from the scalp is very low and only a small proportion of HFOs has a broad enough spatial extent to be recorded from a standard cup electrode [8]. As previously discussed, the advantage of tEEG is better source localization than EEG and a less invasive routine than intracranial, making high-quality recordings of neural activity, such as HFOs, easier to obtain. Therefore, it should be theoretically possible to not only design an automated detector that can locate HFOs without requiring a surgical procedure to do so, but to subsequently use an event detection count and provide an initial estimate of which area of brain tissue contributes to seizure propagation.

1.1 Neurophysiological Signal Analysis

Neurophysiological signal processing imposes new complexities over linear time-invariant systems upon which the majority of signal processing techniques have been designed. The signal recorded from any electrode designed to measure electrical properties of the cortex is a summation of activity from many neurons, each of which can be considered an independent random process [3]. Some research suggests EEG can be treated as a nonlinear deterministic process by describing the firing of individual neurons; however, the electrical signals from the brain are often

considered stochastic in nature and it is generally understood that if the signals recorded are the result of a deterministic system, then that system is so complex that it is best modeled as a random process [2]. In order to accurately reconstruct neurological signals we turn to a parametric approach.

There are two types of noise to consider when working with human data: environmental and physiological. With surface electrode recording, standard cup or TCRE, the environmental noise could be due to other equipment in the vicinity, 60 Hz line noise, sensor movement (gradually increasing electrode impedance), etc [2, 3] for which most of the parameters cannot be known. This noise can vary based on the type of electromagnetic coupling, the area of the pickup point, and the patient's proximity to the source [2, 18]; and it may exist for the duration of the recording or appear intermittently, either gradually coming and going or appearing in short irregular bursts. Physiological data, on the other hand, describes all the data intentionally recorded by the electrodes that are not part of the signal being observed at the time. As an example, a single study may consist of slow delta waves during sleep patterns and broadband spiking activity corresponding to seizures when the patient is awake. Depending on the nature of the study and what the clinician intends to observe, one of these signal types could be considered data of interest whereas the other would be physiological noise. Since the parameters of this latter type of noise are generally known, digital filters are frequently used to reject the physiological background not currently under observation when possible.

TCREs record further from the neuronal source and closer to obstacles in the patient environment when compared with intracranial contacts for which the majority of HFO detectors are designed. Naturally, the HFO amplitude range is highly dependent on the proximity to the HFO generators [8]. This complicates the design of an HFO detection routine for tEEG as there is apt to be more environmental

noise and a lesser magnitude of the physiological data under observation.

The EEG signal is nonstationary long-term due to its slowly time-varying signal properties; abrupt state changes; and transient events, such as spikes, that appear superimposed on the background signal [2, 19, 20]. Neurological data has a $1/f$, or pink, spectrum but due to the lack of stationarity, estimating the noise covariance and applying a pre-whitening filter for the length of the recording is not feasible [21]. Fortunately, as much as 90% of EEG less than 4 seconds in duration can be modeled as a Gaussian wide-sense stationary (WSS) process [21]. It is important to ensure any signal processing techniques that rely on a stationary signal, such as frequency spectrum analyses, must be adapted to consider short-term stationarity.

It has been shown that short, locally stationary EEG signals can be described with an autoregressive (AR) model as an all-pole filter is able to describe the spectral parameters of several different EEG rhythms [2]. The AR model also has the advantage of no spectral leakage offering a better frequency resolution when compared to nonparametric methods [20]. Fourier analysis is limited for EEG since it is designed for linear, stationary systems [16], and requires a tradeoff between short record length and accuracy in the frequency domain [22]. The AR model on the other hand, offers the advantage of relatively few samples for accurate frequency analysis and computational efficiency [22] due to the analytically derived AR spectrum having an infinite frequency resolution with a less restrictive dependence on the length of analyzed data [20]. It is also good for the analysis of signals with sharp spectral features at the cost of poor spectral estimation when inappropriate model orders are used [20].

It is easy to understand the suitability of an all pole model to neurological data as the EEG signal is described by dominant frequencies, such as sharply defined

peaks and rhythmic bands describing different brain states, opposed to spectral notches (absence of power at certain frequency bands) which are better described by a moving average (MA) model [22].

Typical, non-adaptive signal analysis models require that the signal is stationary which we know neurophysiological data is not. Performing a spectral analysis on a data record exhibiting a non-stationarity will result in a severely biased estimate [23]. Thus, it is necessary to either use an adaptive model or divide the signal into short data segments that can be assumed wide-sense stationary [22]. In this study, non-stationarities will be found in the data records followed by signal analysis on short windows located at the events of interest believed to exhibit an anomaly.

1.1.1 High Frequency Oscillations

The HFO is a signal found in brain wave activity, typically described by duration of less than 125 ms and frequency in the 80-500 Hz range [8, 5]; however, some studies have found HFOs at frequencies well above 500 Hz [6, 24]. They can be further categorized into ripples: frequencies between 80 and 200 Hz, and fast ripples: frequencies greater than 200 Hz [25, 5]. HFOs have a minimum of four oscillations and an interval of greater than 25 ms between events [14, 8]. Although HFOs can be found by reviewing recorded patient data, performing a manual search is time consuming and automated analyses are preferred when adequate [26].

It should be noted due to the pink noise spectrum of neurological data, the amplitude of ripples is higher than that of fast ripples for the same signal-to-noise (SNR) ratio [27]. Similarly, the overall duration is dependent on the frequency of the signal. HFOs must have a minimum of four oscillations [8], and past studies attempting to use strictly a signal duration to decide between HFOs or spikes have failed since the duration of fast ripples will be shorter than that of ripples

[27, 28]. Therefore, it is important when using a duration threshold to base it on the frequency of the event [27].

In setting thresholds for event detectors it is important to understand that HFOs occur in an on-off intermittent pattern [16] and rates are not consistent across channels or patients [29]. The pathological significance of HFOs is they reflect fields of synchronized action potentials within small neuron clusters which are responsible for seizure generation [16]. Therefore the electrodes contributing the most HFOs in a study should be indicative of the SOZ.

1.1.2 Epileptic Spikes

Epileptic spikes are a non-oscillatory signal characterized by an abrupt change in voltage polarity of less than 70 ms duration [3, 8] which can incidentally be selected by HFO detection methods due to the spurious oscillations they exhibit following the use of a digital filter [25, 30]. It is known that filtering spikes or artifacts could produce sustained oscillations at high amplitudes due to Gibbs' phenomenon [8, 27], and that sharp transients with high amplitudes will contribute more to generating spurious oscillations [28]. However, oscillations resulting from filtering will have a low correlation with the raw signal in the frequency band of interest [8] and result in relatively few oscillations [28]. While epileptic spikes and sharp waves are a highly specific marker of epilepsy, they do not carry the same physiological significance as HFOs since they are not markers of the SOZ; therefore, it is important to distinguish between the two when designing a detection method [8, 5]. Fortunately, spikes are broadband in the frequency domain whereas HFOs are sinusoidal and exhibit a narrow frequency band, making it theoretically possible to classify the two with an automated method.

1.2 HFO Detection

The short-term Fourier transform (STFT) has been used on TCRE data, visualized with spectrograms, to find areas where high-frequency narrow-band activity is present [12] as this may be a spectral indicator of HFOs. However, there is need for examining other techniques for several reasons. First, there is no way to determine from the spectrogram alone if the high-frequency, narrow-band activity represents a HFO or is due to some other physiological or environmental noise. Second, this is a manual technique for making follow-up review easier and not a detection routine on its own. Knowing that the statistical properties of EEG signals, and in the case of this problem the underlying noise, varies over time is reason to believe a better estimator exists than the STFT.

It is important to design a detector based on valid assumptions about the data. There is no shortage of ad hoc approaches in the literature, such as basing on the eyeballing of a frequency spectrum so that it looks similar to what other events have shown [31, 28]. While some of these methods will work, the result is likely to be a biased detector and performance demonstrated on limited data subsets. A majority of the detectors available today are limited in clinical practice due to their large number of false detections [27]. Additionally, many detectors are strictly compared against the markings as done by a clinician, which does not show how each algorithm works on different data types and can be a fairly subjective analysis [27]. Similarly, many others are compared only during slow-wave sleep when datasets are relatively clean and therefore is not a valid performance evaluation of the detector [29]. It is important to evaluate each detection routine on controlled computer simulated data, designed to be realistic to the actual environment, in order to verify the efficacy of the algorithm and gain a meaningful performance metric [18].

It is also unknown if the methods that work well for iEEG or EEG will have similar results for tEEG. As previously discussed, while information conveyed is similar, the signals being analyzed are not identical and statistical properties may differ. Of the methods that have been previously evaluated for HFO detection no single detector has been established as a best method [30, 26]; therefore a comparison of detectors is worthwhile.

A majority of the successful approaches consist of a two-stage approach: separate events of interest from the background and perform a subsequent test to decide if the event of interest consists of an HFO. Two such methods will be compared to the novel detection routine established here. The first, which will be called the Staba approach, was developed by R. J. Staba et al. in [6] and is most often used as a baseline method for comparison in the literature since it was one of the first publications to offer automated HFO detection with good results. The second, which will be called the Birot method, was developed by G. Birot et al. in [32] and has gained traction recently for its great performance on capturing fast ripples from iEEG.

The Staba routine is a filtering of the signal followed by a short-term energy detector for marking events of interest (EOIs) and the setting of subsequent amplitude and duration thresholds for HFO classification. The first stage of the detector uses a band-pass filter for the HFO frequency range (80-500 Hz) and computes a three ms sliding window of the RMS signal. Successive RMS values > 5 standard deviation above the mean RMS signal with a minimum of six ms duration are flagged as EOIs. The second stage uses the rectified signal and classifies an EOI as an HFO if a minimum of six peaks > 3 standard deviation above the mean amplitude are present.

The first stage of the Birot routine uses a Hanning window to find the local

energy of the high-pass filtered signal which is compared to a threshold for stage 2 follow-up. The real contribution of this work is in the second stage which uses a ratio of the power in the periodogram (for full-band data) for high and low-frequency bands for each EOI. The intent is to reject spikes which are broadband in nature and therefore will result in a lower test statistic than HFOs, which are described by high frequencies only. Results are compared in [32] using both the Fourier and wavelet transforms. Best results are achieved using a low-band of 32-64 Hz and a high-band of 256-512 Hz, although the comparison between different frequency band sets tested does not vary greatly.

Both methods use a different implementation of what is essentially an energy detector as the first stage, for which there are known drawbacks with neurophysiological data. When detectors rely on thresholds calculated using a standard deviation or percentile, the event count will be based on the distribution of actual events; i.e., it will cause a high number of false alarms on inactive channels and missed events on active ones [27]. The energy detector is known to fail for HFOs in active channels [8].

Though the Staba method is generally accepted as the baseline detector to which the majority of new routines are compared, it is designed for use with short, relatively noiseless datasets and is prone to detecting artifacts as HFOs [29].

Each method the author believes to be a fair implementation and baseline comparison.

1.3 Statistical Signal Processing

There are multiple statistical signal processing approaches used in this thesis. A high level overview of some of them is provided here.

1.3.1 Autoregressive Processes

It has been previously mentioned that neurological data can be well-described by an AR process. This is because the data is defined by its strength of power in certain frequency bands, opposed to its lack thereof which would be better described by a MA process in order to capture spectral notches [22]. The AR process is

$$w[n] = u[n] - \sum_{j=1}^p a[j]w[n-j] \quad (1)$$

for model order p , where the data $w[n]$ is modeled as the output of a causal all-pole discrete filter excited at the input by white Gaussian noise (WGN). It is a simple model in which the current value is generated by a weighted sum of previous values plus a noise term [33]. The input to the filter, $u[n]$, is termed the excitation noise and is necessary to ensure the output describes a WSS random process [34]. By using this model to describe EEG data, the pink noise assumption becomes inherent to the filter design and passing generated WGN as input results in simulated data characteristic of EEG.

A single sample of WGN with mean μ takes the form [34]

$$p_x(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(x - \mu)^2\right] \quad (2)$$

Making the approximate probability density function (PDF) of an AR process excited by WGN for record length N

$$p_x(x) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} \left(x[n] + \sum_{j=1}^p a[j]x[n-j]\right)^2\right] \quad (3)$$

A more intuitive way to think of the AR model is a coloring of the white noise input to allow the modeling of a wide variety of power spectral densities (PSD).

The theoretical PSD for data which fits an AR model can be described by

$$\hat{P}_{xx}(f) = \frac{\hat{\sigma}_u^2}{|1 + \sum_{m=1}^p \hat{a}[m] \exp(-j2\pi fm)|^2} \quad (4)$$

where the “hat” denotes the parameters as estimated from the data, or

$$P_{AR}(f) = \frac{\sigma^2}{|A(f)|^2} \quad (5)$$

In order to solve for the AR parameters from a data set, the following AR term must be minimized

$$J(a) = \sum_{n=p}^{N-1} \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)^2 \quad (6)$$

The partial derivative with respect to the k th AR parameter is

$$\frac{\delta J(a)}{\delta a[k]} = 2 \sum_{n=p}^{N-1} \left(x[n] + \sum_{j=1}^p \hat{a}[j]x[n-j] \right) x[n-k] = 0; \quad k = 1, 2, \dots, p \quad (7)$$

Resulting in

$$\sum_{j=1}^p \hat{a}[j] \sum_{n=p}^{N-1} x[n-j]x[n-k] = - \sum_{n=p}^{N-1} x[n]x[n-k]; \quad k = 1, 2, \dots, p \quad (8)$$

Using matrix notation, the MLE of the set of AR parameters can be expressed by

$$\begin{bmatrix} c_{xx}[1, 1] & c_{xx}[1, 2] & \dots & c_{xx}[1, p] \\ c_{xx}[2, 1] & c_{xx}[2, 2] & \dots & c_{xx}[2, p] \\ \vdots & \vdots & \ddots & \vdots \\ c_{xx}[p, 1] & c_{xx}[p, 2] & \dots & c_{xx}[p, p] \end{bmatrix} \begin{bmatrix} \hat{a}[1] \\ \hat{a}[2] \\ \vdots \\ \hat{a}[p] \end{bmatrix} = - \begin{bmatrix} c_{xx}[1, 0] \\ c_{xx}[2, 0] \\ \vdots \\ c_{xx}[p, 0] \end{bmatrix} \quad (9)$$

where

$$c_{xx}[j, k] = \frac{1}{N-p} \sum_{n=p}^{N-1} x[n-j]x[n-k] \quad (10)$$

The AR parameters can now be solved with a Cholesky decomposition. This result is an estimate of the optimal forward predictor, often termed the covariance method [33].

It is possible to instead base the estimate on the optimal forward and backward predictors of the AR process using

$$c_{xx}[j, k] = \frac{1}{2(N-p)} \left(\sum_{n=p}^{N-1} x[n-j]x[n-k] + \sum_{n=0}^{N-1-p} x[n+j]x[n+k] \right) \quad (11)$$

which is termed the modified covariance method [33].

The derivation above is one of multiple methods commonly used to estimate AR parameters from a dataset. The entirety of AR parameter estimation in this thesis utilizes the modified covariance method as it is a common choice in mixed-spectrum problems and generally results in a lesser shifting of spectral estimate peaks from their true frequency locations due to additive noise [35, 33].

1.3.2 Generalized Likelihood Ratio Test

The generalized likelihood ratio test (GLRT) is a statistical hypothesis test considered best suited for problems where the signal PDF is not entirely known, i.e., may be known with the exception of certain parameters. The GLRT replaces the unknown parameters by their maximum likelihood estimates (MLEs). The advantage of using the MLE is it produces an efficient parameter estimate if one exists [34]. The neurological data being analyzed in this thesis consists of several signal and noise parameters that are quite variable in practice; with the GLRT, no prior knowledge of the unknown signal parameters is necessary. Additionally, the GLRT can be shown to be the uniformly most powerful test asymptotically when compared to other invariant detectors [36]. While no optimality can be claimed with using the GLRT, it works well in practice.

The GLRT decides a signal is present based on the following test statistic

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_1, H_1)}{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_0, H_0)} > \gamma \quad (12)$$

where H_1 is the case when a signal is present, H_0 is the null hypothesis, and $\hat{\boldsymbol{\theta}}_i$ represents the MLE of the unknown parameters. The benefit of using the GLRT

for detection problems where information about the unknown parameters is desired should be clear: the first step in calculating the test statistic is to find the maximum likelihood estimate.

1.3.3 Canonical Autoregressive Decomposition

The canonical autoregressive decomposition (CARD) is a technique for estimating unknown sinusoidal parameters when obscured by an unknown AR process. The intent of utilizing this approach for HFO detection should now be clear: neurophysiological data can be described by an AR background and we are searching for an unknown brief sinusoidal burst. CARD offers a simplistic way of addressing mixed spectral problems as a sum of an AR process and sinusoids so that both signal components can be estimated simultaneously. The intuitive, yet rudimentary, explanation of the method is a prewhitening of data followed by a periodogram peak search. A full derivation is outside the scope of this document but can be found in [35] where it is shown to require a numerical maximization with respect to only the sinusoidal frequencies.

For implementation here, the CARD method uses a fixed sinusoidal model order of $s = 1$ and is run on full-band data to reject spurious oscillations due to the filtering of spikes or sharp signal artifacts. It is believed this is an adequate approach as CARD should accurately estimate a pink pre-whitener for the AR background and find the sinusoidal burst representing an HFO.

A background has been provided on signals of interest and techniques that may aid us in event detection from tEEG. The following chapters provide a description and performance evaluation of a novel HFO detection routine and subsequent SOZ estimator based on the assumptions provided above.

List of References

- [1] D. J. DiLorenzo and J. D. Bronzino, *Neuroengineering*. Boca Raton, FL: Taylor and Francis Group, 2008.
- [2] L. Sornmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. London, UK: Elsevier Academic Press, 2005.
- [3] K. E. Misulus and T. C. Head, *Essentials of Clinical Neurophysiology*. Burlington, MA: Butterworth Heinemann, 2003.
- [4] F. Rosenow and H. Luders, “Presurgical evaluation of epilepsy,” *Brain*, vol. 124(9), pp. 1683–1700, 2001.
- [5] J. Jacobs, R. Staba, E. Asano, H. Otsubo, J. Y. Wu, M. Zijlmans, I. Mohamed, P. Kahane, F. Dubeau, V. Navarro, and J. Gotman, “High-frequency oscillations (HFOs) in clinical epilepsy,” *Progress in Neurobiology*, vol. 98(3), pp. 302–315, 2012.
- [6] R. J. Staba, C. L. Wilson, A. Bragin, I. Fried, and J. E. Jr, “Quantitative analysis of high-frequency oscillations recorded in human epileptic hippocampus and entorhinal cortex,” *Journal of Neurophysiology*, vol. 88(4), pp. 1743–1752, 2002.
- [7] A. B. Gardner, G. A. Worrell, E. Marsh, D. Dlugos, and B. Litt, “Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings,” *Clinical Neurophysiology*, vol. 118(5), pp. 1134–1143, 2007.
- [8] G. Worrell, K. Jerbi, K. Kobayashi, J. Lina, R. Zelman, and M. L. V. Quyen, “Recording and analysis techniques for high-frequency oscillations,” *Progress in Neurobiology*, vol. 98(3), pp. 265–278, 2012.
- [9] W. G. Besio, K. Koka, R. Aakula, and W. Dai, “Tri-polar concentric ring electrode development for laplacian electroencephalography,” *IEEE Transactions on Biomedical Engineering*, vol. 53(5), pp. 926–933, 2006.
- [10] K. Koka and W. G. Besio, “Improvement of spatial selectivity and decrease of mutual information of tri-polar concentric ring electrodes,” *Journal of Neuroscience Methods*, vol. 165(2), pp. 216–222, 2007.
- [11] O. Makeyev, X. Liu, L. Wang, Z. Zhu, A. Taveras, D. Troiano, A. V. Medvedev, and W. G. Besio, “Feasibility of recording high frequency oscillations with tripolar concentric ring electrodes during pentylenetetrazole-induced seizures in rats,” *34th Annual International Conference of the IEEE EMBS*, pp. 4599–4602, 2012.

- [12] W. G. Besio, I. E. Martinez-Juarez, O. Makeyev, J. N. Gaitanis, A. S. Blum, R. S. Fisher, and A. V. Medvedev, “High-frequency oscillations recorded on the scalp of patients with epilepsy using tripolar concentric ring electrodes,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 2, 2014.
- [13] A. Pearce, D. Wulsin, J. A. Blanco, A. Krieger, B. Litt, and W. C. Stacey, “Temporal changes of neocortical high frequency oscillations in epilepsy,” *Journal of Neurophysiology*, vol. 110(5), pp. 1167–1179, 2013.
- [14] U. Malinowska, G. K. Bergey, J. Harezlak, and C. C. Jouny, “Identification of seizure onset zone and preictal state based on characteristics of high frequency oscillations,” *Clin Neurophysiol*, vol. 126(8), pp. 1505–1513, 2015.
- [15] S. Burnos, B. Frauscher, R. Zelman, C. Haegelen, J. Sarnthein, and J. Gotman, “The morphology of high frequency oscillations (HFO) does not improve delineating the epileptogenic zone,” *Clinical Neurophysiology*, vol. 127(4), pp. 2140–2148, 2016.
- [16] L. Huang, X. Ni, W. Ditto, M. Spano, P. Carney, and Y. C. Lai, “Detecting and characterizing high-frequency oscillations in epilepsy: a case study of big data analysis,” *R. Soc. open sci.*, vol. 4, 2017.
- [17] M. Dumpelmann, J. Jacobs, and A. Schulze-Bonhage, “Temporal and spatial characteristics of high frequency oscillations as a new biomarker in epilepsy,” *Epilepsia*, vol. 56(2), pp. 197–206, 2015.
- [18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*. Westford, MA: Prentice Hall, 2013.
- [19] H. Ombao, J. Raz, R. V. Sachs, and W. Guo, “The SLEX model of a non-stationary random process,” *Ann. Inst. Statist. Math.*, vol. 54(1), pp. 171–200, 2002.
- [20] A. Fahoum and A. Fraihat, “Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains,” *ISRN Neuroscience*, 2014.
- [21] I. Ghaleb, C. Davila, and R. Srebro, “Prewhitening of background brain activity via autoregressive modeling,” *IEEE*, pp. 242–245, 1997.
- [22] J. Pardey, S. Roberts, and L. Tarassenko, “A review of parametric modelling techniques for EEG analysis.”
- [23] S. Kay, “A new nonstationarity detector,” *IEEE Transactions on Signal Processing*, vol. 56(4), pp. 1440–1451, 2008.

- [24] K. Kobayashi, T. Agari, M. Oka, H. Yoshinaga, I. Date, Y. Ohtsuka, and J. Gotman, “Detection of seizure-associated high-frequency oscillations above 500 hz,” *Epilepsy Research*, vol. 88, pp. 139–144, 2010.
- [25] C. G. Benar, L. Chauviere, F. Bartolomei, and F. Wendling, “Pitfalls of high-pass filtering for detecting epileptic oscillations: a technical note on “false” ripples,” *Clinical Neurophysiology*, vol. 121(3), pp. 301–310, 2010.
- [26] N. Roehri, J. M. Lina, J. C. Mosher, F. Bartolomei, and C. G. Benar, “Time-frequency strategies for increasing high-frequency oscillation detectability in intracerebral EEG,” *IEEE Transactions on Biomedical Engineering*, vol. 63(12), 2016.
- [27] N. Roerhi, F. Pizzo, F. Bartolomei, F. Wendling, and C. G. Benar, “What are the assets and weaknesses of HFO detectors? a benchmark framework based on realistic simulations,” *PLoS ONE*, vol. 12(4), 2017.
- [28] M. Amiri, J. M. Lina, F. Pizzo, and J. Gotman, “High frequency oscillations and spikes: Separating real HFOs from false oscillations,” *Clinical Neurophysiology*, vol. 127, pp. 187–196, 2016.
- [29] S. V. Gliske, Z. T. Irwin, K. A. Davis, K. Sahaya, C. Chestek, and W. C. Stacey, “Universal automated high frequency oscillation detector for real-time, long term EEG,” *Clinical Neurophysiology*, 2015.
- [30] N. Jmail, M. Gavaret, F. Wendling, A. Kachouri, G. Hamadi, J.-M. Badier, and C. Benar, “A comparison of methods for separation of transient and oscillatory signals in EEG,” *Journal of Neuroscience Methods*, vol. 199(2), pp. 273–289, 2011.
- [31] S. Burnos, P. Hilfiker, O. Surucu, F. Scholkmann, N. Krayenbuhl, T. Grunwald, and J. Sarnthein, “Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis,” *PLoS ONE*, vol. 9(4), 2014.
- [32] G. Birot, A. Kachenoura, L. Albera, C. Benar, and F. Wendling, “Automatic detection of fast ripples,” *Journal of Neuroscience Methods*, vol. 213(2), pp. 236–249, 2013.
- [33] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1988.
- [34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993.
- [35] V. Nagesha and S. Kay, “Spectral analysis based on the canonical autoregressive decomposition,” *IEEE Transactions on Signal Processing*, vol. 44(7), pp. 1719–1733, 1996.

- [36] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, 1998.

CHAPTER 2

Methodology

This chapter starts with a high-level concept on the HFO detection routine used and follows with an algorithm description for each of the two detector stages derived as part of this routine. Then an overview on the tEEG simulator design created for performance evaluation is discussed, how the detection algorithms are to be compared, and how the SOZ is to be estimated based on the results.

2.1 Detector Overview

Similar to the majority of the detection routines in the literature, the detection approach generated here starts by filtering out the unwanted background, locating events of interest, and deciding based on signal parameters if an HFO is present. Novel approaches based on the GLRT are used for locating the events of interest and detecting the signal. The detection routine is run on each tEEG channel individually.

The HFO signal can be thought of as containing a transient component and a rhythmic component [1]. A two-stage detector is therefore designed to first locate transients via an anomaly detector and second find oscillations via a sinusoid detector.

The tEEG signal consisting of an HFO is modeled by a sinusoid embedded in AR noise.

To locate the events of interest, an anomaly detector is derived using the GLRT in order to find stationarity changes described by a change in AR filter coefficients or excitation white noise variance. The derivation is based on finding a stationarity change in a single window of tEEG data and is implemented by a sliding window (recursively calculated for each sample shift) so that all non-

stationarities are located in the entire tEEG data record. The use of a sliding window estimator in order to find a transition point is preferred when interested in an abrupt change in data and not any gradual changes in the background [2].

The second stage decides if the events of interest found from stage 1 contain HFOs. This is called the sinusoid detector and is again derived using the GLRT; the result is essentially a normalized replica correlator using pre-whitened estimates. A required input to this detector is an estimate of the sinusoidal frequency and AR process filter and noise parameters. Two suggestions for estimating these parameters are offered: CARD parameter estimates based on the full-band signal and use of the AR spectrum as already estimated from stage 1, the latter of which requires a subsequent period count to distinguish between HFO and filtered artifact.

Corresponding block diagram for the overall detection routine is shown in figure 1.

Before stepping into the algorithm descriptions, it is worth discussing why this is believed to be a valid approach. In the introduction an overview was provided on neurological signal analysis and the signals of interest. We have essentially built a list of prior assumptions about our data, each of which needs to be tackled by the designed detector. EEG signals are slowly time-varying and non-stationary long term; the anomaly detection stage is designed to reject slowly time-varying signals due to AR estimates being similar between consecutive windows. This same stage will reject any gradual increases in noise throughout the recording due to the noise estimates being slowly time-varying. Any abrupt changes in environmental noise or signal artifacts will be detected by the anomaly detector, but subsequently rejected by the sinusoid detector in stage 2. The data record can be assumed stationary in small segments—this allows the use of a window size much less than a second and

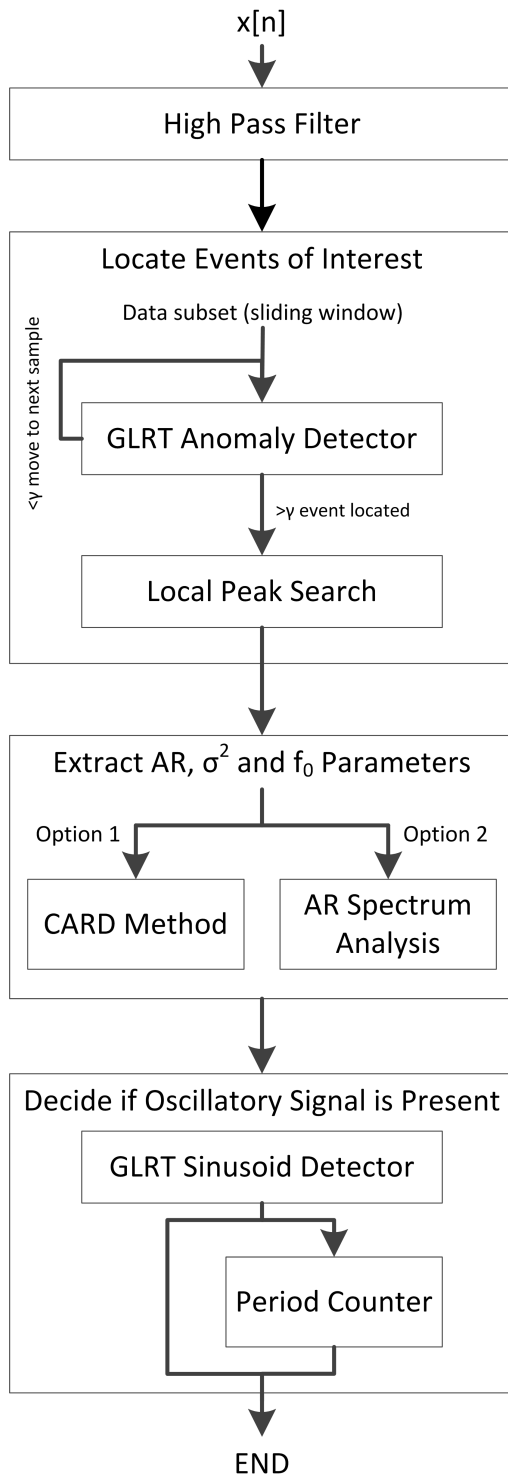


Figure 1. Detection Routine Block Diagram

subsequent spectrum analysis to be performed since the window can be assumed WSS. EEG has a $1/f$, or pink, noise spectrum requiring any signal analysis to use a pre-whitener—this is inherent to the AR filter that is estimated from the data.

Similar arguments can be made for the signals to be detected. The HFO is a transient event occurring above 80 Hz and can be separated from the background; therefore, an 80 Hz high pass filter is applied to the data first to isolate high frequencies, the anomaly detection stage finds transient events, and the subsequent stage is based on the frequency peak. The signal is also less than 125 ms in duration, and the detector window can be adjusted based on the sample rate to run on a quarter-second of data at a time (allowing for a max-duration HFO to be captured in half of the window). Because the frequency of event occurrence varies between channels, the test statistic threshold is constant for all channels and there is not an influence due to, say, an energy detector threshold that is based on the mean of the channel. In order to reject epileptic spikes, which are broadband events and can resemble oscillations due to high-pass filtering, two implementations of the sinusoid detector are employed. The first, to count the number of detected oscillations from the high-pass band based on estimated signal duration and frequency and use this as a subsequent test statistic as HFOs have a minimum of 4 oscillations whereas filtered spikes show fewer [3, 4]. The second, to instead perform sinusoid detection and AR estimation on the full-band signal using the CARD method.

2.2 Anomaly Detector

2.2.1 Overview

The goal of the first stage is to detect a stationarity change in the PSD of a locally wide sense stationary random process. The PSD is modeled as an AR

process of model order p with

$$P_{xx}(f) = \frac{\sigma_u^2}{|1 + a[1]e^{-j2\pi f} + \dots + a[p]e^{-j2\pi fp}|^2} \quad (13)$$

where $a[j]$ represents the AR parameters and σ_u^2 represents the noise variance. It is known that the signal of interest is a transient signal, separable from the background, so it is desired to only detect significant changes in the AR filter coefficients or driving white noise variance and reject slowly time-varying signals.

For the use of HFO detection, it is known that the signal of interest is above 80 Hz. Therefore, a digital 80 Hz high-pass filter is applied before anomaly detection is performed.

2.2.2 Derivation

For a single window length M , the data is broken into \mathbf{x}_0 , the full window length; \mathbf{x}_1 , the first half of the data; and \mathbf{x}_2 , the second half of the data. Similarly, the data is described by AR parameters \mathbf{a}_0 , \mathbf{a}_1 , and \mathbf{a}_2 each of model order p and excitation white noise variances σ_0^2 , σ_1^2 , and σ_2^2 . This can be shown as

$$\begin{aligned} \mathbf{x}_0 : \quad & \mathbf{a}_0, \sigma_0^2 \quad n = 0, 1, \dots, M - 1 \\ \mathbf{x}_1 : \quad & \mathbf{a}_1, \sigma_1^2 \quad n = 0, 1, \dots, \frac{M}{2} - 1 \\ \mathbf{x}_2 : \quad & \mathbf{a}_2, \sigma_2^2 \quad n = \frac{M}{2}, \frac{M}{2} + 1, \dots, M - 1 \end{aligned} \quad (14)$$

where the **bold** parameters represent column or row-vectors of data.

The hypothesis test is

$$\begin{aligned} H_0 : \quad & \mathbf{a}_0 = \mathbf{a}_1 = \mathbf{a}_2 \quad \text{AND} \quad \sigma_0^2 = \sigma_1^2 = \sigma_2^2 \\ H_1 : \quad & \mathbf{a}_0 \neq \mathbf{a}_1 \neq \mathbf{a}_2 \quad \text{OR} \quad \sigma_0^2 \neq \sigma_1^2 \neq \sigma_2^2 \end{aligned} \quad (15)$$

The GLRT decides H_1 if

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}_1; \hat{\mathbf{a}}_1, \hat{\sigma}_1^2, H_1)p(\mathbf{x}_2; \hat{\mathbf{a}}_2, \hat{\sigma}_2^2, H_1)}{p(\mathbf{x}_0; \hat{\mathbf{a}}_0, \hat{\sigma}_0^2, H_0)} > \gamma \quad (16)$$

The PDF of an AR process excited by WGN of length M can be written as

$$p(\mathbf{x}_i; \hat{\mathbf{a}}_i, \hat{\sigma}_i^2, H_i) = \frac{1}{(2\pi\sigma_i^2)^{M/2}} \exp \left[-\frac{1}{2\sigma_i^2} \sum_{n=0}^{M-1} (x[n] + a_i[1]x[n-1] + \dots + a_i[p]x[n-p])^2 \right] \quad (17)$$

or equivalently

$$p(\mathbf{x}_i; \hat{\mathbf{a}}_i, \hat{\sigma}_i^2, H_i) = \frac{1}{(2\pi\sigma_i^2)^{M/2}} \exp \left[-\frac{1}{2\sigma_i^2} \sum_{n=0}^{M-1} \left(x[n] + \sum_{j=1}^p a_i[j]x[n-j] \right)^2 \right] \quad (18)$$

The test statistic can now be written as

$$\frac{\frac{1}{(2\pi\sigma_1^2)^{M/4}} \exp \left[-\frac{1}{2\sigma_1^2} \sum_{n=0}^{\frac{M}{2}-1} (x[n] + \sum_{j=1}^p a_1[j]x[n-j])^2 \right]}{\frac{1}{(2\pi\sigma_2^2)^{M/4}} \exp \left[-\frac{1}{2\sigma_2^2} \sum_{n=\frac{M}{2}}^{M-1} (x[n] + \sum_{j=1}^p a_2[j]x[n-j])^2 \right]} \cdot \frac{1}{(2\pi\sigma_0^2)^{M/2}} \exp \left[-\frac{1}{2\sigma_0^2} \sum_{n=0}^{M-1} (x[n] + \sum_{j=1}^p a_0[j]x[n-j])^2 \right] \quad (19)$$

To solve the MLE of σ_1^2 we maximize

$$p(\mathbf{x}_1; \hat{\mathbf{a}}_1, \hat{\sigma}_1^2, H_1) = \frac{1}{(2\pi\sigma_1^2)^{M/4}} \exp \left[-\frac{1}{2\sigma_1^2} \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=1}^p a_1[j]x[n-j] \right)^2 \right] \quad (20)$$

$$\ln p(\mathbf{x}_1; \hat{\mathbf{a}}_1, \hat{\sigma}_1^2, H_1) = -\frac{M}{4} \ln 2\pi\sigma_1^2 - \frac{1}{2\sigma_1^2} \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=1}^p a_1[j]x[n-j] \right)^2 \quad (21)$$

$$\frac{\delta \ln p(\mathbf{x}_1; \hat{\mathbf{a}}_1, \hat{\sigma}_1^2, H_1)}{\delta \sigma_1^2} = -\frac{M}{4\sigma_1^2} + \frac{1}{2(\sigma_1^2)^2} \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=1}^p a_1[j]x[n-j] \right)^2 \quad (22)$$

Setting equal to zero and solving results in

$$\hat{\sigma}_1^2 = \frac{2}{M} \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=1}^p a_1[j]x[n-j] \right)^2 \quad (23)$$

Equivalently,

$$\hat{\sigma}_2^2 = \frac{2}{M} \sum_{n=\frac{M}{2}}^{M-1} \left(x[n] + \sum_{j=1}^p a_2[j]x[n-j] \right)^2 \quad (24)$$

$$\hat{\sigma}_0^2 = \frac{1}{M} \sum_{n=0}^{M-1} \left(x[n] + \sum_{j=1}^p a_0[j]x[n-j] \right)^2 \quad (25)$$

To find the MLE of the AR parameter sets we must minimize

$$J(a) = \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=1}^p a_1[j]x[n-j] \right)^2 \quad (26)$$

$$\frac{\delta J(a)}{\delta a_1[k]} = 2 \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=1}^p a_1[j]x[n-j] \right) x[n-j] = 0 \quad (27)$$

This is solved using $k=1$ as an example

$$2 \sum_{n=0}^{\frac{M}{2}-1} a_1[1]x[n-1]x[n-1] = -2 \sum_{n=0}^{\frac{M}{2}-1} \left(x[n] + \sum_{j=2}^p a_1[j]x[n-j] \right) x[n-1] \quad (28)$$

$$\hat{a}_1[1] = \frac{\sum_{n=0}^{\frac{M}{2}-1} (x[n] + \sum_{j=2}^p a_1[j]x[n-j])x[n-1]}{\sum_{n=0}^{\frac{M}{2}-1} x^2[n-1]} \quad (29)$$

It then becomes clear that the AR parameter estimates are

$$\hat{a}_1[k] = \frac{\sum_{n=0}^{\frac{M}{2}-1} (x[n] + \sum_{j=1}^p a_1[j]x[n-j])x[n-k]}{\sum_{n=0}^{\frac{M}{2}-1} x^2[n-k]} \quad \text{for } j \neq k \quad (30)$$

$$\hat{a}_2[k] = \frac{\sum_{n=\frac{M}{2}}^{M-1} (x[n] + \sum_{j=1}^p a_2[j]x[n-j])x[n-k]}{\sum_{n=\frac{M}{2}}^{M-1} x^2[n-k]} \quad \text{for } j \neq k \quad (31)$$

$$\hat{a}_0[k] = \frac{\sum_{n=0}^{M-1} (x[n] + \sum_{j=1}^p a_0[j]x[n-j])x[n-k]}{\sum_{n=0}^{M-1} x^2[n-k]} \quad \text{for } j \neq k \quad (32)$$

When implementing, we will use the modified covariance method to solve for the AR parameters.

When plugging in the estimates for σ^2 in the test statistic, the summations in the exponent will cancel. This is shown below for the null hypothesis.

$$p(\mathbf{x}_0; \hat{\mathbf{a}}_0, \hat{\sigma}_0^2, H_0) = \frac{1}{(2\pi\hat{\sigma}_0^2)^{M/2}} \exp \left[-\frac{M \sum_{n=0}^{M-1} (x[n] + \sum_{j=1}^p \hat{a}_0[j]x[n-j])^2}{2 \sum_{n=0}^{M-1} (x[n] + \sum_{j=1}^p \hat{a}_0[j]x[n-j])^2} \right] \quad (33)$$

The resulting test statistic is

$$L_G(\mathbf{x}) = \frac{\frac{1}{(2\pi\hat{\sigma}_1^2)^{M/4}} e^{-M/4} \frac{1}{(2\pi\hat{\sigma}_2^2)^{M/4}} e^{-M/4}}{\frac{1}{(2\pi\hat{\sigma}_0^2)^{M/2}} e^{-M/2}} \quad (34)$$

$$L_G(\mathbf{x}) = \frac{(\hat{\sigma}_0^2)^{M/2}}{(\hat{\sigma}_1^2)^{M/4} (\hat{\sigma}_2^2)^{M/4}} \quad (35)$$

Or equivalently

$$L_G(\mathbf{x}) = \left(\frac{(\hat{\sigma}_0^2)^2}{\hat{\sigma}_1^2 \hat{\sigma}_2^2} \right)^{M/4} > \gamma' \quad (36)$$

2.2.3 Explanation

Three sets of AR filter coefficients and σ^2 values are estimated: for the full window length, the first half of the window, and the second half of the window. The σ^2 values can be thought of as having the AR filter coefficient estimates and the input data embedded in their estimates. The resulting value from the hypothesis test then becomes a relationship between the estimated noise variance for the entire window and the estimates for its two halves. The test statistic will only produce positive values. Significant signal changes will result in a high valued test statistic whereas a stationary signal will produce values closer to zero.

The white excitation noise can be thought of as a fitting order for an AR process. The addition of a sinusoid increases the model order estimate; however it should also cause the AR coefficient estimates to change when estimating assuming

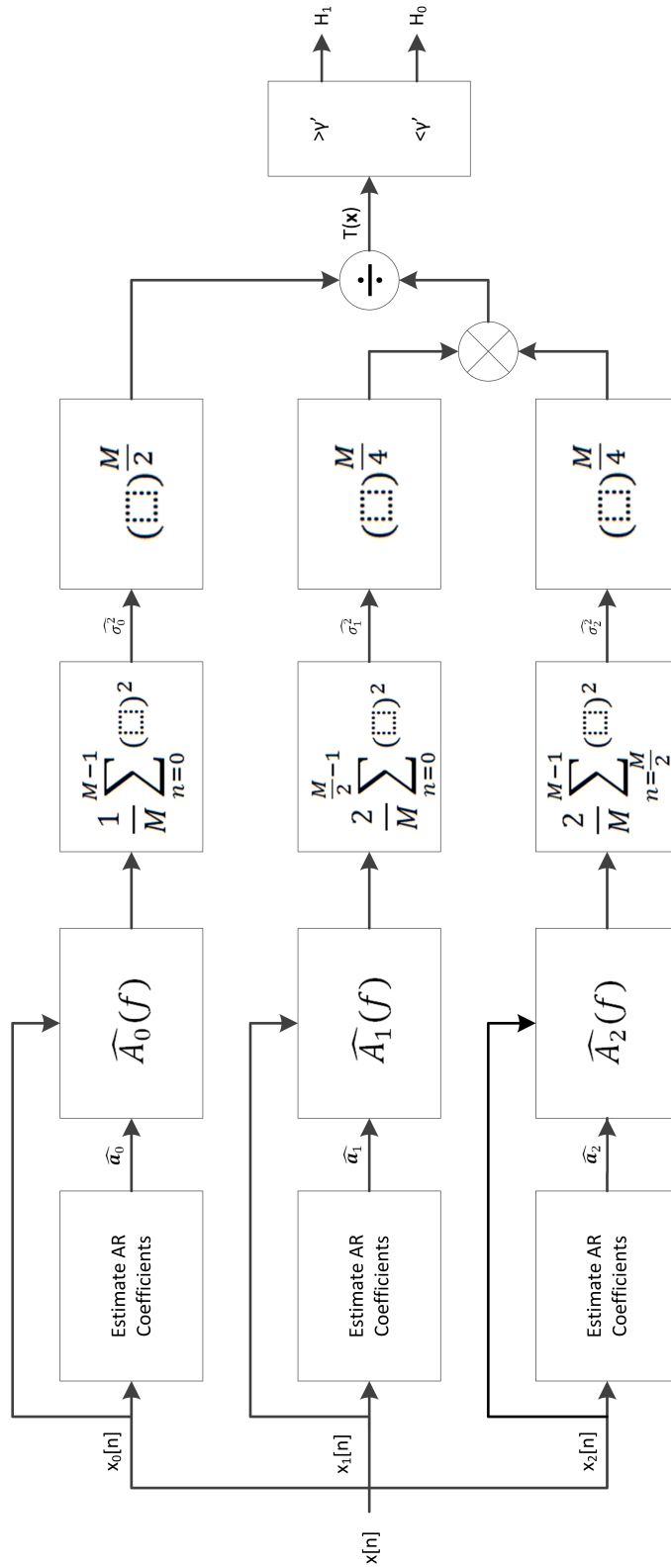


Figure 2. Anomaly Detector Block Diagram

a fixed model order. Therefore, the anomaly detector can detect transient events when implemented with a fixed p for the duration of the signal.

By choosing the threshold correctly, the test statistic should identify transient signals but not those that may be described by a time-varying autoregressive (TVAR) model or an environment with variant noise over time. It is expected that the output of this stage, when run on tEEG data, will be a set of events consisting of HFOs; epileptic spikes; and instances of sudden increases in environmental noise, such as electromagnetic interference or a disconnected electrode.

2.2.4 Expected Performance

The asymptotic performance of the GLRT statistic is given by [5]

$$2 \ln L_G(\mathbf{x}) \stackrel{a}{\sim} \begin{cases} X_r^2 & \text{under } H_0 \\ X_r'^2(\lambda) & \text{under } H_1 \end{cases} \quad (37)$$

Where X_r^2 and $X_r'^2$ are the central and non-central Chi-squared distributions, respectively. In certain uses of the GLRT these can be computed to get an estimate of the expected performance of the detector; however, the result under H_1 can take on a large variety of signals, so it is not feasible to compute the Chi-squared non-centrality parameter for this problem. It is possible to compare H_0 to the expected asymptotic results by performing a Monte Carlo simulation of the detector on tEEG simulated data (computed in chapter 3). The issue of performance estimation is further complicated in that the Chi-squared distribution assumes independent, identically distributed (IID) trials [5] of which the data for this problem does not fit—the background noise is known to be slowly time-varying and trials are highly correlated since the detector recursively computes the test statistic for single-sample shifted windows. The probability of false alarm (P_{FA}) for these cases is plotted against the threshold in figure 3 with the central Chi-squared distribu-

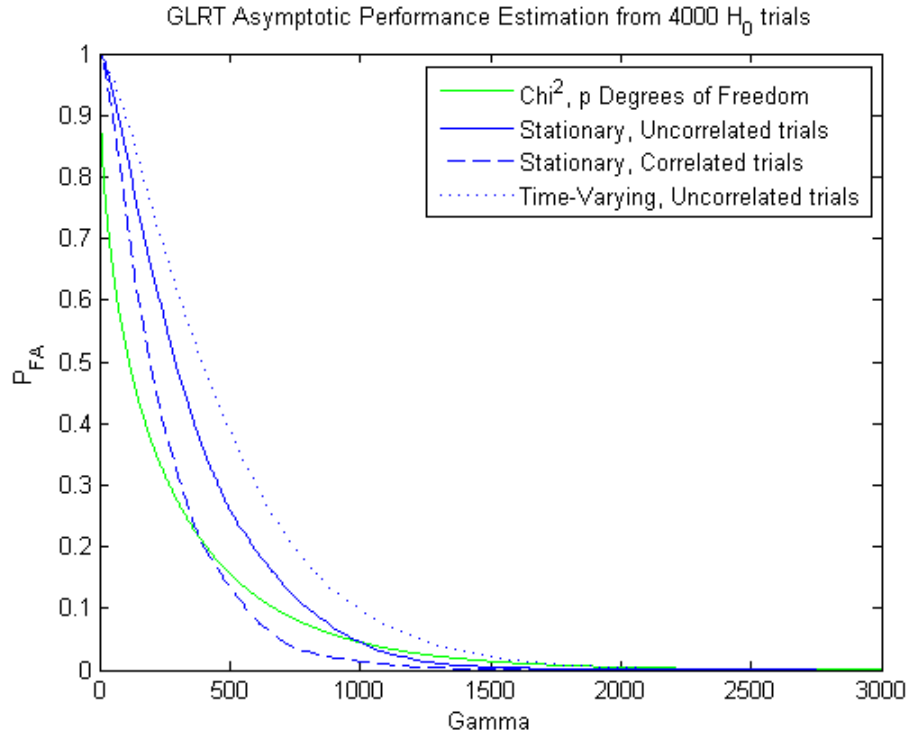


Figure 3. GLRT Theoretical Asymptotic Performance

tion with p degrees of freedom. The performance is clearly difficult to estimate and a performance evaluation must be run on simulated tEEG data—this is shown in figure 9 in the results section.

2.2.5 Implementation

The detector derived above is based on detecting a stationarity change in a single window of M samples. M ideally should be chosen to be equal to two times the length of a stationary segment so that the stationary signal is captured in half the window length. To find stationarity changes in a data record of N samples, the test statistic must be calculated for each sample delay n_0 , from 0 to $N - M - 1$. This can be thought of as a sliding window of length M : recursively compute the test statistic and shift window by one sample. When the test statistic exceeds the threshold, sample location $n_0 + M/2$ is considered a stationarity change.

A window duration of 250 milliseconds is used for the detection of HFOs in this study in order to ensure the full HFO signal can be captured in half the window length.

In addition to M , only the AR model order (p) and the threshold (γ) need to be defined for implementation. γ is based on a Monte Carlo simulation of data and corresponds to a P_{FA} of 0.05. It should be noted that the threshold will vary based on window size and sample rate. EEG data has successfully been modeled with an AR process of model order 6 previously [6]; however, filtering the data at 80 Hz will remove the strongest contributing frequency components from the signal. For implementation in this study, a model order of 4 is used for anomaly detection on the filtered data.

Implementing this routine programmatically is straightforward. The modified covariance method is used to obtain the estimates of the AR parameters. These estimates are then plugged into equations (23) - (25) and (36) to get the excitation white noise variance estimates and resulting test statistic.

It is established that a threshold crossing signifies the area of a stationarity change. Since the detector is implemented by recursively computing the test statistic for single-sample shifted windows, the results between windows will be highly correlated. In other words, each transient event should be marked by several threshold crossings. To avoid double-counting events, a local peak search with a buffer of $M/2$ samples is performed on the test statistic results. The event window is determined to be the location of the local peak $\pm M/2$.

2.3 Sinusoid Detector

2.3.1 Overview

The input signals of the second stage are discrete data subsets of fixed duration assumed to contain an HFO, epileptic spike, or noise characterized by an

abrupt change in signal stationarity. The goal is to decide if an oscillatory signal characteristic of an HFO is present: a sinusoid of frequency above 80 Hz with a minimum of 4 cycles. The signal of interest may or may not take up the entire duration of the window. Because HFOs are transient events separable from the background activity, we model the signal to be detected as a sinusoid added to AR noise of model order p , or

$$x[n] = \begin{cases} w[n] & n = 0, 1, \dots, n_0 - 1 \\ A \cos(2\pi f_0(n - n_0)) + w[n] & n = n_0, n_0 + 1, \dots, n_0 + M - 1 \\ w[n] & n = n_0 + M, n_0 + M + 1, \dots, N - 1 \end{cases} \quad (38)$$

where f_0 is the signal frequency, A is its amplitude, n_0 the signal onset, M the signal duration, and

$$w[n] = u[n] - \sum_{j=1}^p a[j]w[n-j] \quad (39)$$

is the AR noise model of order p , all of which are unknown.

The previous detection stage uses the modified covariance method to estimate the AR parameters and the MLE to find σ^2 . However, if the stationarity change includes a sinusoidal signal, the estimates for the current window may be inaccurate since the AR model order, p , should be greater than originally assumed. It is also unknown if a frequency analysis of the window to find the peak is adequate to estimate f_0 due to the interference of the AR background on the frequency spectrum. Therefore, other methods may need to be utilized to obtain estimates of these parameters.

2.3.2 Derivation

To start, we will assume the sinusoidal frequency and parameters of the AR filter are already known and derive the rest of the problem. The collection and use of these parameters will be discussed in the Implementation section for this

detector.

Our hypothesis test is

$$\begin{aligned}
 H_0 : x[n] &= w[n] & n &= 0, 1, \dots, N-1 \\
 H_1 : x[n] &= \begin{cases} w[n] & n = 0, 1, \dots, n_0 - 1; \\ & n_0 + M, n_0 + M + 1, \dots, N - 1 \\ As[n - n_0] + w[n] & n = n_0, n_0 + 1, \dots, n_0 + M - 1 \end{cases} & (40)
 \end{aligned}$$

where

$$s[n - n_0] = \cos(2\pi f_0(n - n_0)) \quad (41)$$

The noise is the result of an autoregressive process excited by WGN with excitation noise variance σ^2 . It is easily shown that the whitening filter for an AR process of model order p is $x[n] + \sum_{j=1}^p a[j]x[n - j]$. By pre-whitening we get for the noise only case

$$x[n] + \sum_{j=1}^p a[j]x[n - j] = w[n] + \sum_{j=1}^p a[j]w[n - j] \quad (42)$$

$$x[n] = u[n] - \sum_{j=1}^p a[j]x[n - j] \quad (43)$$

And when the signal is present under H_1

$$\begin{aligned}
 x[n] + \sum_{j=1}^p a[j]x[n - j] &= \\
 A \left(s[n - n_0] + \sum_{j=1}^p a[j]s[n - n_0 - j] \right) &+ w[n] + \sum_{j=1}^p a[j]w[n - j] & (44)
 \end{aligned}$$

$$x[n] = As[n - n_0] + A \sum_{j=1}^p a[j]s[n - n_0 - j] + u[n] - \sum_{j=1}^p a[j]x[n - j] \quad (45)$$

The GLRT decides H_1 if

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{A}, \hat{n}_0, \hat{M}, H_1)}{p(\mathbf{x}; H_0)} > \gamma \quad (46)$$

The approximate distribution under the null hypothesis is

$$p(\mathbf{x}; H_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} \prod_{n=0}^{N-1} \exp \left[-\frac{1}{2\sigma^2} \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)^2 \right] \quad (47)$$

And the approximate H_1 distribution is

$$\begin{aligned} p(\mathbf{x}; \hat{A}, \hat{n}_0, \hat{M}, H_1) &= \frac{1}{(2\pi\sigma^2)^{N/2}} \prod_{n=0}^{n_0-1} \exp \left[-\frac{1}{2\sigma^2} \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)^2 \right] \\ &\cdot \prod_{n=n_0}^{n_0+M-1} \exp \left[-\frac{1}{2\sigma^2} \left(x[n] + \sum_{j=1}^p a[j]x[n-j] - As[n-n_0] \right. \right. \\ &\quad \left. \left. - A \sum_{j=1}^p a[j]s[n-n_0-j] \right)^2 \right] \\ &\cdot \prod_{n=n_0+M}^{N-1} \exp \left[-\frac{1}{2\sigma^2} \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)^2 \right] \end{aligned} \quad (48)$$

We can expand

$$\begin{aligned} &\left(x[n] + \sum_{j=1}^p a[j]x[n-j] - As[n-n_0] - A \sum_{j=1}^p a[j]s[n-n_0-j] \right)^2 = \\ &\left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)^2 - 2Ax[n]s[n-n_0] - 2Ax[n] \sum_{j=1}^p a[j]s[n-n_0-j] \\ &- 2As[n-n_0] \sum_{j=1}^p a[j]x[n-j] - 2A \sum_{j=1}^p a^2[j]x[n-j]s[n-n_0-j] \\ &+ A^2s^2[n-n_0] + 2A^2s[n-n_0] \sum_{j=1}^p a[j]s[n-n_0-j] \\ &+ A^2 \sum_{j=1}^p a^2[j]s^2[n-n_0-j] \end{aligned} \quad (49)$$

so that

$$p(\mathbf{x}; \hat{A}, \hat{n}_0, \hat{M}, H_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} \prod_{n=n_0}^{n_0+M-1} \exp \left[-\frac{1}{2\sigma^2} J(A) \right] \cdot \prod_{n=0}^{N-1} \exp \left[-\frac{1}{2\sigma^2} \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)^2 \right] \quad (50)$$

where

$$\begin{aligned} J(A) = & A^2 s^2[n - n_0] + 2A^2 s[n - n_0] \sum_{j=1}^p a[j]s[n - n_0 - j] \\ & + A^2 \sum_{j=1}^p a^2[j]s^2[n - n_0 - j] - 2Ax[n]s[n - n_0] \\ & - 2Ax[n] \sum_{j=1}^p a[j]s[n - n_0 - j] - 2As[n - n_0] \sum_{j=1}^p a[j]x[n - j] \\ & - 2A \sum_{j=1}^p a^2[j]x[n - j]s[n - n_0 - j] \end{aligned} \quad (51)$$

This allows the likelihood function to reduce to

$$L_G(\mathbf{x}) = \prod_{n=n_0}^{n_0+M-1} \exp \left[-\frac{1}{2\sigma^2} J(A) \right] > \gamma \quad (52)$$

and taking the natural log results in

$$\ln L_G(\mathbf{x}) = -\frac{1}{2\sigma^2} \sum_{n=n_0}^{n_0+M-1} J(A) > \ln \gamma \quad (53)$$

To find the MLE of A , we maximize $J(A)$

$$\begin{aligned} \frac{\delta J(A)}{\delta A} = & \sum_{n=n_0}^{n_0+M-1} 2As^2[n - n_0] + \sum_{n=n_0}^{n_0+M-1} 4As[n - n_0] \sum_{j=1}^p a[j]s[n - n_0 - j] \\ & + \sum_{n=n_0}^{n_0+M-1} 2A \sum_{j=1}^p a^2[j]s^2[n - n_0 - j] - \sum_{n=n_0}^{n_0+M-1} 2x[n]s[n - n_0] \\ & - \sum_{n=n_0}^{n_0+M-1} 2x[n] \sum_{j=1}^p a[j]s[n - n_0 - j] - \sum_{n=n_0}^{n_0+M-1} 2s[n - n_0] \sum_{j=1}^p a[j]x[n - j] \\ & - \sum_{n=n_0}^{n_0+M-1} 2 \sum_{j=1}^p a^2[j]x[n - j]s[n - n_0 - j] = 0 \end{aligned} \quad (54)$$

$$\hat{A} = \frac{\sum_{n=n_0}^{n_0+M-1} \left(x[n]s[n-n_0] + x[n] \sum_{j=1}^p a[j]s[n-n_0-j] + s[n-n_0] \sum_{j=1}^p a[j]x[n-j] + \sum_{j=1}^p a^2[j]x[n-j]s[n-n_0-j] \right)}{\sum_{n=n_0}^{n_0+M-1} \left(s^2[n-n_0] + 2s[n-n_0] \sum_{j=1}^p a[j]s[n-n_0-j] + \sum_{j=1}^p a^2[j]s^2[n-n_0-j] \right)} \quad (55)$$

$$\hat{A} = \frac{\sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right)}{\sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right)^2} \quad (56)$$

The test statistic can now be re-written as

$$\begin{aligned} & -\frac{1}{2\sigma^2} \left[\hat{A}^2 \sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right)^2 \right. \\ & \quad \left. - 2\hat{A} \sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right) \right] \\ & > \ln \gamma \end{aligned} \quad (57)$$

Which allows us to make a substitution to get

$$\begin{aligned} & -\frac{1}{2\sigma^2} \left[\hat{A} \sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right) \right. \\ & \quad \left. - 2\hat{A} \sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right) \right] \end{aligned} \quad (58)$$

Or equivalently

$$\frac{\hat{A}}{\sigma^2} \sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j] \right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j] \right) > 2 \ln \gamma \quad (59)$$

$$\begin{aligned}
L_G(\mathbf{x}) &= \\
&\frac{\left(\sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j]\right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j]\right)\right)^2}{\sigma^2 \sum_{n=n_0}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j]\right)^2} \\
&> \gamma'
\end{aligned} \tag{60}$$

Which must be maximized with respect to the unknown onset time, n_0 , and duration, M . This can be done by computing the test statistic for each n_0, M pair.

It should also be noted that summations should start at $n_0 + p$ to prevent out-of-bounds limitations. The effect should be negligible for $M \gg p$.

The test statistic is then

$$\max_{n_0, M} \frac{\left(\sum_{n=n_0+p}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j]\right) \left(x[n] + \sum_{j=1}^p a[j]x[n-j]\right)\right)^2}{\sigma^2 \sum_{n=n_0+p}^{n_0+M-1} \left(s[n-n_0] + \sum_{j=1}^p a[j]s[n-n_0-j]\right)^2} \tag{61}$$

which is compared to γ' .

2.3.3 Explanation

The resulting detector is a signal pre-whitener followed by a normalized replica correlator using the pre-whitened signal as the signal replica.

The test statistic is calculated for all possible n_0 (between 0 and $N - M$) and M (between pre-defined limits m_1 and m_2) combinations and the maximum value is compared to a threshold. If the test statistic exceeds the threshold then it is decided that a signal is present and the $\max_{n_0, M}$ location represents the signal onset and duration length in samples.

It is possible to include phase as an unknown parameter in the GLRT derivation with the assumption that the signal amplitude is positive. However, the purpose of this application is to detect >80 Hz signals and the worst-case for an incorrect phase estimate would cause the signal onset estimation to be off by a

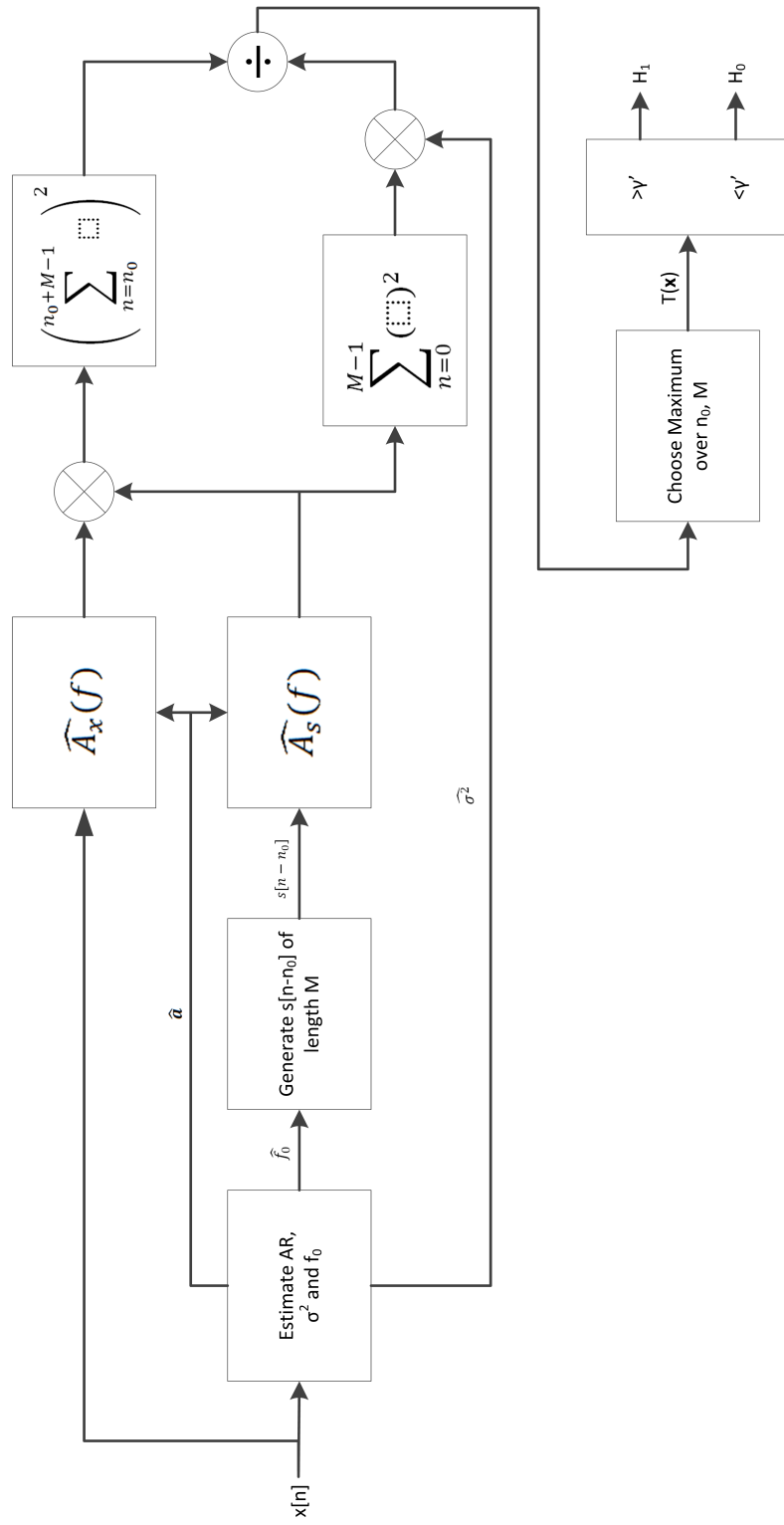


Figure 4. Sinusoid Detector Block Diagram

half-cycle at 80 Hz, or approximately 6 milliseconds. This is considered negligible and therefore the phase is assumed equal to zero.

2.3.4 Expected Performance

The asymptotic performance discussion is similar to that of the anomaly detector as this is another sample-shifting detector based on the GLRT. Again the windows being compared are highly correlated, in this case for each (n_0, M) pair, so IID cannot be assumed. A P_{FA} must be chosen, with threshold based on the result from a Monte Carlo simulation. Performance of this detector can be found in figures 11 and 12 in chapter 3.

2.3.5 Implementation

The derivation assumes the AR parameters and sinusoidal frequency are already known. Two suggestions are offered for the implementation of this second stage. The first is based on CARD as it should allow the estimation of the AR parameters and sinusoidal frequency of the full-band data so that the filtering of spikes does not cause false alarms. The second is to use AR and frequency estimates from the filtered data.

The CARD method is utilized to get an estimate of the AR model order, AR filter coefficients, white excitation noise variance, and sinusoidal signal frequency. In utilizing this method, it is assumed that a single sinusoid is present. CARD estimates use the full-band data subset in order to avoid detecting spurious oscillations due the filtering of sharp transients. This approach should be able to resolve the HFO frequency peak and AR background from the unfiltered data, and the same unfiltered data must be passed to the sinusoid detector as the AR estimates from CARD act as a pre-whitener for the full-band data. An added test for this stage is to first confirm CARD returns a frequency >80 Hz; otherwise it can be

assumed no HFO is present.

An alternate option is to base the AR pre-whitener on the previous window of filtered background tEEG data. i.e., if the event of interest concerns samples on the range $[0 : M]$, then the modified covariance method and noise estimates on the range $[-M : 0]$ are used. The assumption here is that the background is slowly time-varying and therefore estimates less than half a second away are still useful as a pre-whitener. Subsequently, the frequency peak from the event of interest is used as the f_0 value since for such a short window length, any HFO frequency should be dominant in the high-pass band. The results from this study happen to show that these are valid assumptions. Implementation of this approach does not require any additional estimation since the AR and therefore frequency spectrum parameters have already been calculated from the first detection stage. For this approach, high-pass band data must be provided as input to the sinusoid detector. In order to ensure the detected oscillations are from an HFO and not a spike, the number of sinusoidal cycles are calculated using the sampling frequency (F_s), estimated HFO frequency (f_0), and now available MLE of signal duration (M), using

$$cycles = \frac{f_0 \cdot M}{F_s} \quad (62)$$

Both implementations have their own way of deciding if an event is an HFO or a spike. The sinusoid detector in both implementations should limit false alarms due to other noise as the test statistic will be low if no oscillations are present.

In computing the test statistic it can be seen that $\sum_{n=n_0+p}^{n_0+M-1} (s[n - n_0] + \sum_{j=1}^p a[j]s[n - n_0 - j])$ is dependent on the signal duration only and not the onset location n_0 . For a quicker implementation this can be replaced by $\sum_{n=p}^{M-1} (s[n] + \sum_{j=1}^p a[j]s[n - j])$. The corresponding pseudo-code is shown in Algorithm 1.

```

signal  $\leftarrow$   $\cos(2 \cdot \pi \cdot f_0 \cdot n)$ ;
foreach duration value  $M$  do
     $s \leftarrow \text{signal}(0 : M - 1)$ ;
    denominator  $\leftarrow \sigma^2 \sum_{n=p}^{M-1} \left( s[n] + \sum_{j=1}^p a[j]s[n-j] \right)^2$ ;
    foreach onset value  $n_0$  do
         $x \leftarrow \text{inputData}(n_0 : n_0 + M - 1)$ ;
        numerator  $\leftarrow$ 
             $\left( \sum_{n=p}^{M-1} \left( s[n] + \sum_{j=1}^p a[j]s[n-j] \right) \left( x[n] + \sum_{j=1}^p a[j]x[n-j] \right) \right)^2$ ;
         $T_x(M, n_0) \leftarrow \text{numerator} / \text{denominator}$ ;
    end
end
if  $\max(T_x) > \gamma$  then
    signal present ( $H_1$ );
     $\hat{n}_0 = \max T_x(:, :)$ ;
     $\hat{M} = \max T_x(:, :)$ ;
else
    no signal present ( $H_0$ );
end

```

Algorithm 1: Sinusoid Detector Pseudocode

2.4 tEEG Simulator

It is difficult to assess performance of a detector based on observed data alone. For a stationary process it is often possible to use averaging should the underlying process be ergodic [7, 2]. Clearly, this is not the case with neurophysiological data. In order to avoid developing a biased detector based on a limited subset of data, detector performance should be evaluated on realistic simulations [4, 2]. The following describes the tEEG simulator design created for the evaluation of HFO detection routines.

Multiple sets of AR parameters, and corresponding white noise variances, have been extracted from real, full-band tEEG and saved to file using the modified covariance method. Data subsets corresponding to different tEEG rhythms were hand-selected in order to allow for a varying background, i.e., such that delta /

theta (<7 Hz), alpha / beta (8-30 Hz), gamma (>30 Hz) and seizure (broadband spiking) frequency components contribute to the overall signal. The estimated model order was used for each data subset resulting in model order ranges from $p = 10$ to $p = 15$. The range in model orders is preferred for the simulator to ensure detector performance is not biased by selecting a model order for the detector that is identical to each set of filter coefficients used to generate the data. To simplify implementation, each AR parameter set has additional filter coefficients added with value equal to zero so that a fixed model order of $p = 15$ can be used. By observing the AR model it should be clear that this does not affect the actual filter model order but merely impacts array size programmatically.

In order to simulate tEEG data based on a single AR parameter set, WGN is generated and filtered based on the AR process formula used in equation (1). Unfortunately, once a set of AR parameters is chosen and used to generate data, the output of the simulator is stationary which is not true to the nature of neurological signals. The purpose of retaining multiple sets of AR parameters is to have the ability to cycle through different types of tEEG background data before adding in physiological events (HFOs and epileptic spikes). However, simulating from an AR parameter set for a duration of time and immediately changing to a different filter parameter set would be equivalent to a neuron cluster shutting down abruptly and another one gaining full effect at the same time which is not a realistic scenario [7]. In order to overcome these impracticalities, the simulator randomly selects a subsequent AR parameter set and linearly spaces the filter coefficients over a period of one to two seconds, such that each filtered sample is effectively described by a slightly different AR filter. At no time does the simulator dwell on a single set of AR parameters and the result is a truly time-varying background for realistic performance estimation of the developed detection routines.

An example of simulated TVAR tEEG is shown in figure 5.

HFOs and spike events can be described as appearing superimposed on background EEG. The simulator design first generates the background data and the physiological events are subsequently added to the TVAR data. The simulator is able to randomize (between set limits) onset times and types of events, which are then outputted with the data so the accuracy of the simulator can be tested without needing to manually find and mark events.

The HFO signal is simply a Gaussian-modulated sinusoidal pulse in which the duration, frequency, and amplitude can be modified. Each parameter is randomized on a uniform distribution between defined limits representing the physiological characteristics of an HFO to ensure the detectors are not biased for certain event features.

The spike signal is designed as an exponential growth function concatenated with a negated exponential decay function. The simulator uses a coin flip to decide if the simulated waveform is positive or negative before adding to the tEEG background since the polarity in practice is dependent on the electrode's position relative to the neural source and the type of polarization shift in the corresponding neuron cluster [8]. Although spike polarity typically only varies on a channel-to-channel basis, this allows the testing of detection routines without assuming a certain type of deflection.

Example simulations of an HFO and a spike signal that get added to the TVAR background are shown in figure 6.

It is important to have the ability to modify the SNR of the tEEG simulator so that an evaluation of detection routines against different qualities of recordings can be performed. In this case we consider the signal an HFO or a spike (as spikes are expected to be detected in the first detection stage) and noise the TVAR tEEG

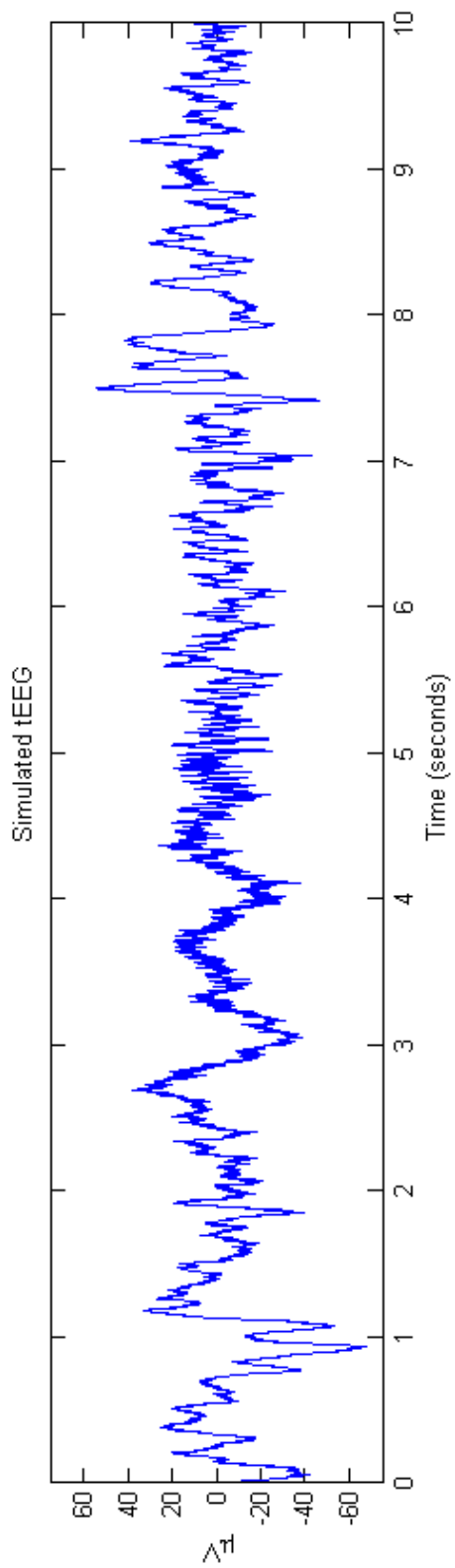


Figure 5. 10 seconds of Simulated TVAR tEEG

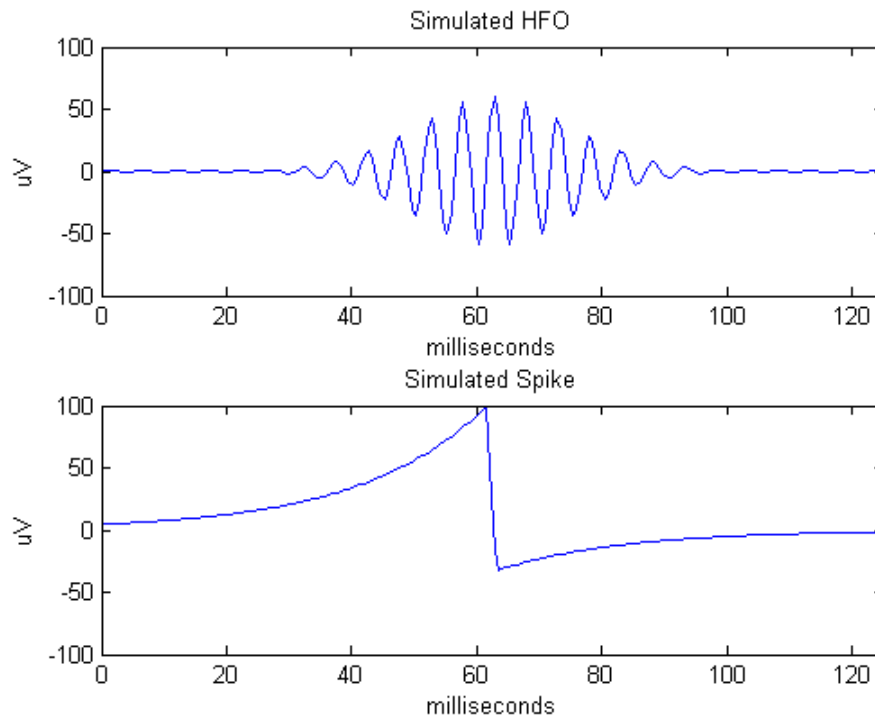


Figure 6. Simulated HFO and Spike Events

background. For this, we define a new measure termed the Event-to-Noise Ratio (EvNR). EvNR is an event amplitude multiplier based on the most recent half-second of tEEG data. With a unity EvNR, the spike max value will be equal to the amplitude of the recent full-band signal; for an HFO, the peak-to-peak max of the sinusoidal pulse is equal to the amplitude of the recent signal high-pass filtered at 80 Hz. An EvNR equal to one appears realistic for these events as recorded from TCREs, though this is a subjective measure. It is believed that as EvNR approaches two the events are unrealistically large but may be closer to iEEG in which more physiological data contributes than environmental noise. Naturally, as EvNR approaches zero the signal becomes indistinguishable.

An example of simulations of different backgrounds with an HFO and spike added are shown in figures 7 and 8. In both, an HFO is added at the half-second mark and a spike at the 1.5-second mark, each with unity EvNR.

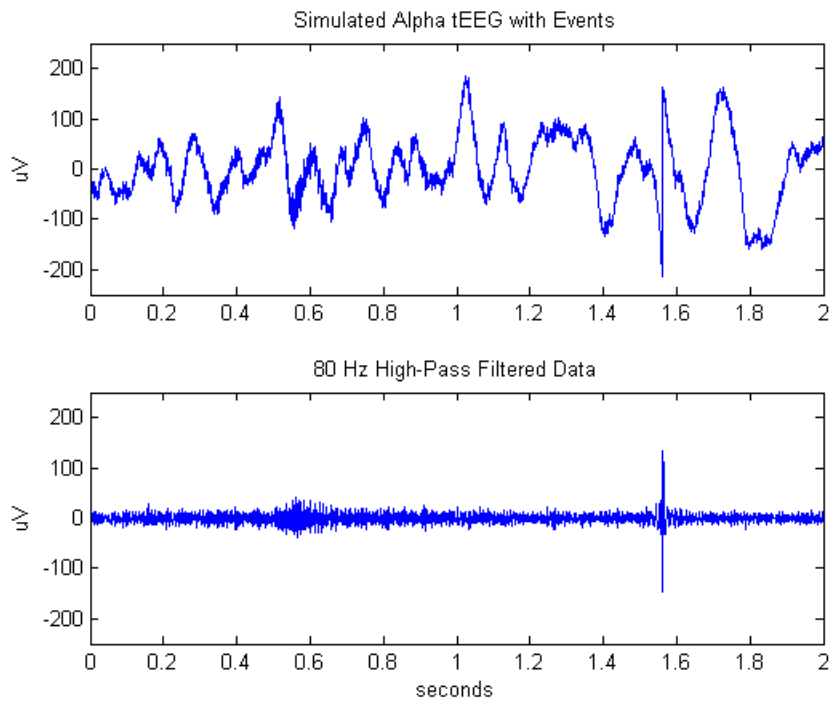


Figure 7. Simulated Alpha tEEG with Events

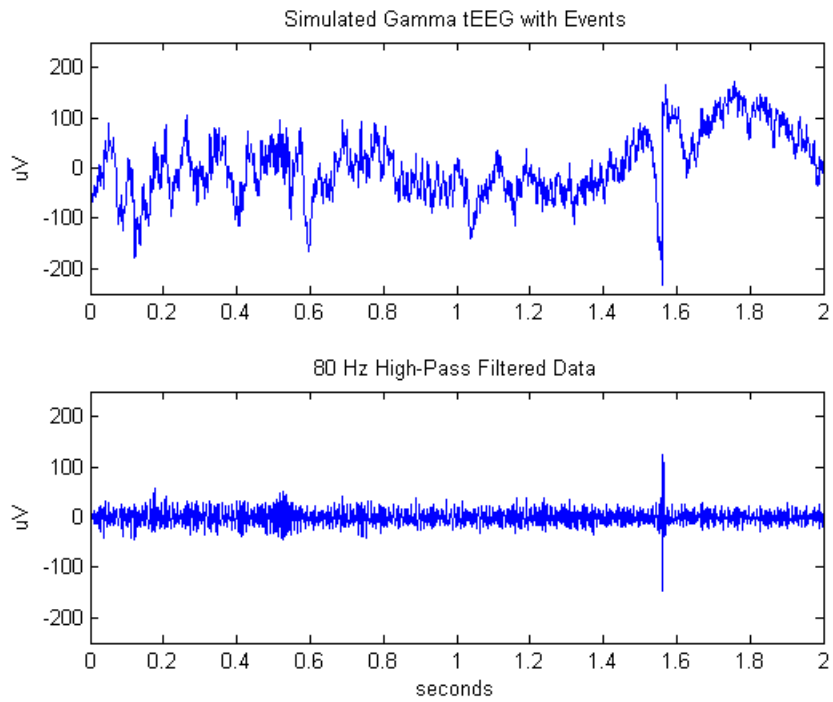


Figure 8. Simulated Gamma tEEG with Events

2.5 Comparison Detectors

The Staba and Birot detection algorithms are discussed in the introduction. Both routines have been implemented in MATLAB for comparison to the detection routine derived above. The Staba routine remains unmodified from its original publication in [9]. The Birot routine has been modified slightly from [10] to work with HFOs in both the ripple and fast ripple band for use on tEEG data (where the sample rate is not identical to that used in the original study). The frequency range is therefore expanded for the Birot routine to ratio the power of frequencies above and below 80 Hz. This should be considered a fair adaptation of the detection routine as the original assumptions of the detector remain unmodified.

All detection routines are evaluated on simulated tEEG in order to obtain performance metrics. Receiver operating characteristic (ROC) curves are provided for the first and second stage of each detection routine to show detection performance against a fixed EvNR. In order to show how the performance varies with SNR, a first stage comparison of the probability of detection (P_D) to EvNR is made. Similarly, a stage two probability of error (P_e) vs. EvNR comparison is done to see how each detector classifies simulated HFOs against simulated spikes.

An assessment of each routine is then made on a 10 minute subset of real tEEG data, for which actual HFOs were manually identified, in order to get a more realistic detection count; and a comparison of time to run each algorithm is provided for discussion on clinical applications of each routine.

2.6 SOZ Estimation

It is believed that the count of HFO events on a channel and not the individual HFO parameters can be an indicator of SOZ [11, 12, 13]. In [14] an epileptogenicity index is created based on which channels a certain ratio of high-to-low frequency activity shows up immediately preceding a seizure. Since we know the HFO is

a higher frequency signal, basing an SOZ estimate on the count of HFO events during the pre-ictal stage is reasonable.

The difficulty with SOZ estimation based on the count or frequency of detected HFO events is there will always be a channel with the most events and one with the least. This alone is not enough to quantify which channels are indicative of the SOZ since HFOs will be detected in patients who do not suffer from seizures at all. Therefore, it is important to select only those channels that contain a statistically significant number of events as compared with the rest of the channels. Due to the risk factor with making incorrect estimations, no decision should be made over a wrong decision.

In [12] a measure was done to see if there was statistical significance of HFO channels to the SOZ using a comparison of the difference between detections inside and outside the SOZ to the sum of the same detection counts. This has been adapted in an attempt to find channels contributing to the SOZ by the following measure

$$\phi_{CH} = \frac{D_{CH} - D_{AVG}}{D_{CH} + D_{AVG}} \quad (63)$$

which is computed for each channel, where D_{CH} is the number of detections on the current channel and D_{AVG} is the average of the number of detections on all other channels. ϕ_{CH} is the result, between ± 1 . Values closer to +1 have a higher count of detected HFOs and therefore should be more likely to be in the SOZ.

This is to be tested after the HFO detector has been run on 5 tEEG studies for which a clinician has already identified the SOZ. Results will also be compared with output of the first stage detector alone to see if a statistical significance is available for all events of interest, opposed to only those that have been identified as HFOs by the sinusoid detector

List of References

- [1] S. Chaibi, T. Lajnef, Z. Sakka, M. Samet, and A. Kachouri, “A reliable approach to distinguish between transients with and without HFOs using TQWT and MCA,” *Journal of Neuroscience Methods*, vol. 232, pp. 36–46, 2014.
- [2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*. Westford, MA: Prentice Hall, 2013.
- [3] M. Amiri, J. M. Lina, F. Pizzo, and J. Gotman, “High frequency oscillations and spikes: Separating real HFOs from false oscillations,” *Clinical Neurophysiology*, vol. 127, pp. 187–196, 2016.
- [4] N. Roerhi, F. Pizzo, F. Bartolomei, F. Wendling, and C. G. Benar, “What are the assets and weaknesses of HFO detectors? a benchmark framework based on realistic simulations,” *PLoS ONE*, vol. 12(4), 2017.
- [5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, 1998.
- [6] L. Sornmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. London, UK: Elsevier Academic Press, 2005.
- [7] J. P. Kaipio and P. A. Karjalainen, “Simulation of nonstationary EEG,” *Biological Cybernetics*, vol. 76(5), pp. 349–356, 1997.
- [8] K. E. Misulus and T. C. Head, *Essentials of Clinical Neurophysiology*. Burlington, MA: Butterworth Heinemann, 2003.
- [9] R. J. Staba, C. L. Wilson, A. Bragin, I. Fried, and J. E. Jr, “Quantitative analysis of high-frequency oscillations recorded in human epileptic hippocampus and entorhinal cortex,” *Journal of Neurophysiology*, vol. 88(4), pp. 1743–1752, 2002.
- [10] G. Birot, A. Kachenoura, L. Albera, C. Benar, and F. Wendling, “Automatic detection of fast ripples,” *Journal of Neuroscience Methods*, vol. 213(2), pp. 236–249, 2013.
- [11] M. Dumpelmann, J. Jacobs, and A. Schulze-Bonhage, “Temporal and spatial characteristics of high frequency oscillations as a new biomarker in epilepsy,” *Epilepsia*, vol. 56(2), pp. 197–206, 2015.
- [12] S. V. Gliske, Z. T. Irwin, K. A. Davis, K. Sahaya, C. Chestek, and W. C. Stacey, “Universal automated high frequency oscillation detector for real-time, long term EEG,” *Clinical Neurophysiology*, 2015.
- [13] G. Worrell, K. Jerbi, K. Kobayashi, J. Lina, R. Zelmann, and M. L. V. Quyen, “Recording and analysis techniques for high-frequency oscillations,” *Progress in Neurobiology*, vol. 98(3), pp. 265–278, 2012.

- [14] F. Bartolomei, P. Chauvel, and F. Wendling, “Epileptogenicity of brain structures in human temporal lobe epilepsy: A quantified study from intracerebral EEG,” *Brain*, vol. 131, pp. 1818–1830, 2008.

CHAPTER 3

Findings

Performance metrics are provided based on the tEEG simulator design. The detector and SOZ estimator are also evaluated based on real tEEG data subsets. A discussion on the data shown here is provided in chapter 4.

3.1 Detector Performance

3.1.1 Simulated Data

What follows are Monte Carlo simulation results on generated tEEG to determine performance metrics for the two detectors. All simulations are based on sampling rate of 1600 Hz. For the evaluation of the detection method derived as part of this thesis, a window size M equal to 400 samples (1/4-second) and AR model order of 4 is used.

Figure 9 shows a ROC curve for the first detection stage as estimated from a Monte Carlo simulation. In order to generate, 1000 one-second simulations of TVAR tEEG data are generated under H_0 and a simulated HFO is added to each simulation with frequency and duration randomized using EvNR equal to one to represent H_1 . The test statistic result for stage 1 of each detector is calculated and the detection threshold is varied in order to get an estimate of P_D versus P_{FA} at each point. The “Coin Flip” trace is intended to show the level at which performance could be improved with a random selection. The Anomaly detector based on the GLRT outperforms both comparison detectors for simulated tEEG data.

Figure 10 is the estimated P_D versus EvNR for a P_{FA} of 0.05. The purpose is to show how the performance is affected when the signal is outweighed by noise. In order to generate, 10,000 seconds of TVAR tEEG data are generated as H_0

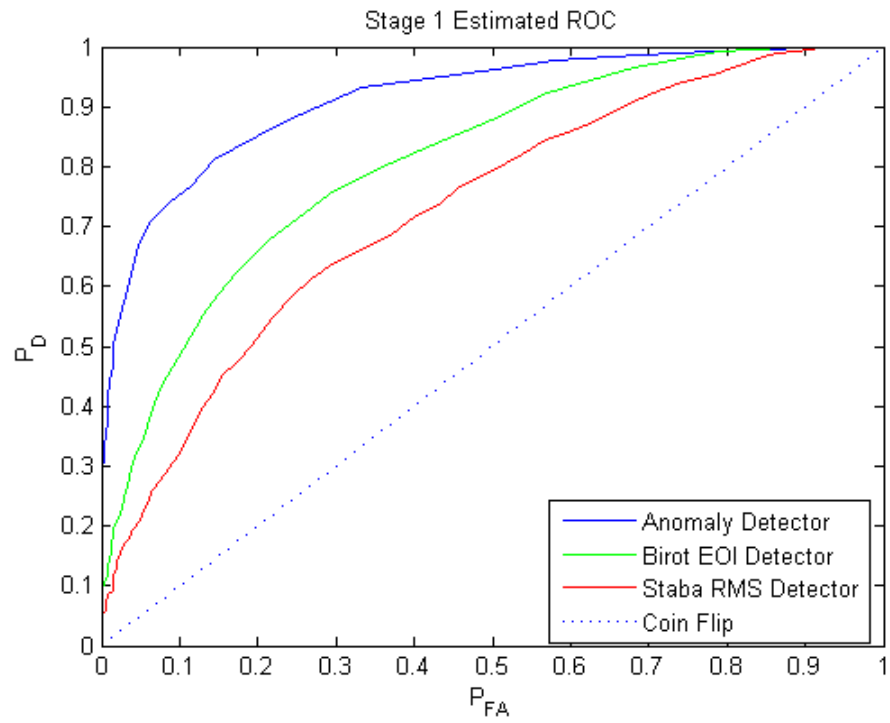


Figure 9. Detector Stage 1 Estimated ROC

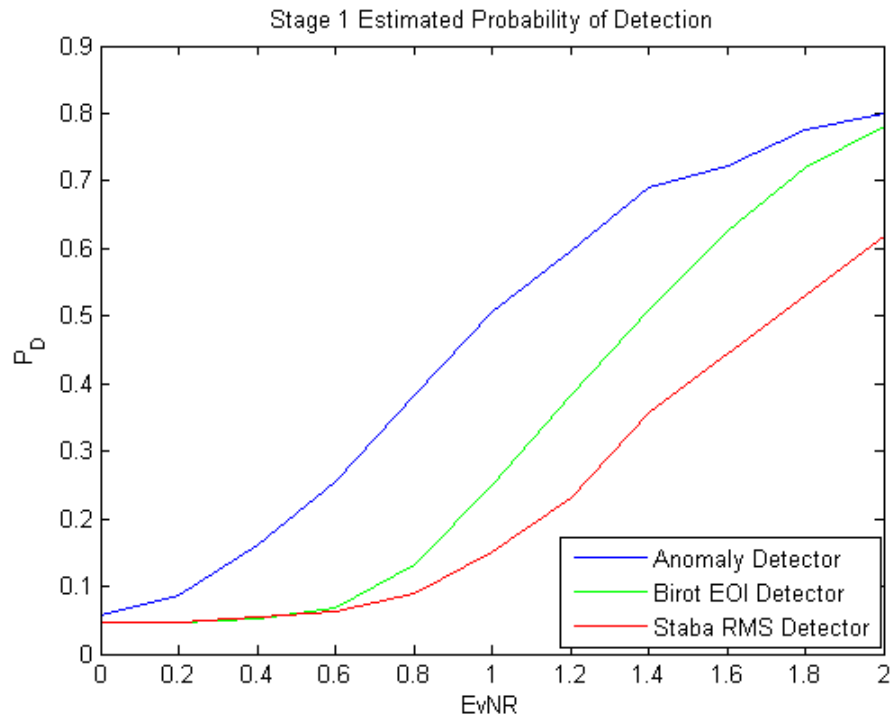


Figure 10. Detector Stage 1 Estimated Probability of Detection

and the test statistic is calculated for each detector. The corresponding detection thresholds are then calculated based on a P_{FA} of 0.05. HFOs are subsequently added to 1000 of the 10,000 seconds of background with frequency and duration randomized and the first detection stage of each routine re-evaluates the data. The latter step is performed for all EvNRs from zero (no discernible signal) to two.

The results show that the GLRT anomaly detector is superior for simulated tEEG data at all signal-to-noise ratios.

Two ROCs are shown for the second stage in order to estimate how each detection method performs against different type of data. In both cases, 1000 one-second simulations of TVAR tEEG data are generated and under H_1 HFOs are added to each simulation with frequency and duration randomized and EvNR equal to one. The null hypothesis in figure 11 is based on the background TVAR tEEG noise only whereas in figure 12 it is based on the same background with a simulated epileptic spike signal added with EvNR equal to one.

As described in chapter 2, the sinusoid detector is implemented in two different ways: the first with estimation from the CARD method and the second via the modified covariance method (MCM). The results of the MCM approach are also fed to a Period Counter to determine if there are enough oscillations for the signal to be considered an HFO. Therefore, both the MCM and Period Counter thresholds must contribute to the performance evaluation for the latter implementation. As can be seen from the results in figure 11, the sinusoid detector as implemented via the CARD method is best at distinguishing HFOs from tEEG background. Ideally, these results are less useful than those in figure 12 since the first detection stage should reject cases of background noise.

When the null hypothesis is instead a simulated spike signal, the period counter is clearly the best performer followed closely by the second Birot detec-

tion stage [see figure 12]. The sinusoid detection based on MCM does poorly as the magnitude of the test statistic for a spike signal resulting in spurious oscillations is greater than that of an HFO. However, figure 11 shows that the MCM sinusoid detector does well against signals without oscillations and spikes are ultimately rejected since it is paired with the period counter. The CARD and Staba implementations are clearly inconsistent.

The intent of the final performance metric in figure 13 is to show how each routine does at properly classifying an EOI, assumed to contain either an HFO or a spike signal. Five-thousand seconds of TVAR tEEG background data are generated. The first iteration adds a spike at each second and has the second-stage of each routine decide if an HFO is present. Then, the same background data is used but a simulated HFO is added at every second and again each routine decides if an HFO is present. The resulting P_e is the number of incorrect classifications over 10,000. This is done for each EvNR from zero to two with P_{FA} fixed at 0.05.

For this example, the period counter and sine detector using MCM results are combined since the two work in series. The results of the P_e curve shows that the sine detector paired with a count of the number of periods is better than the comparison methods at all SNRs. The sine detector based on CARD implementation with full-band data is unable to adequately distinguish between HFO and spike signals.

3.1.2 Real Data

A ten-minute subset of real tEEG data from a 1600 Hz study provided by CRE Medical has been annotated for HFO events by a manual review in MATLAB. This study consists of 23 tEEG channels and the ten-minute segment contains 1535 HFOs. Each detection routine has been run on the data subset with results of the analysis is shown in table 1.

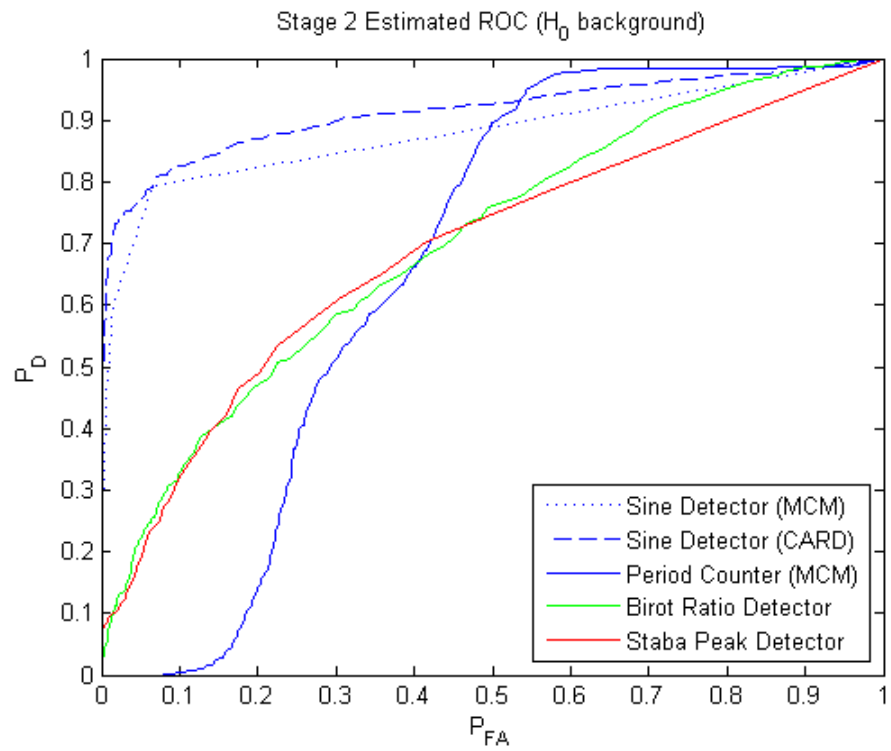


Figure 11. Detector Stage 2 Estimated ROC against H_0 background tEEG noise

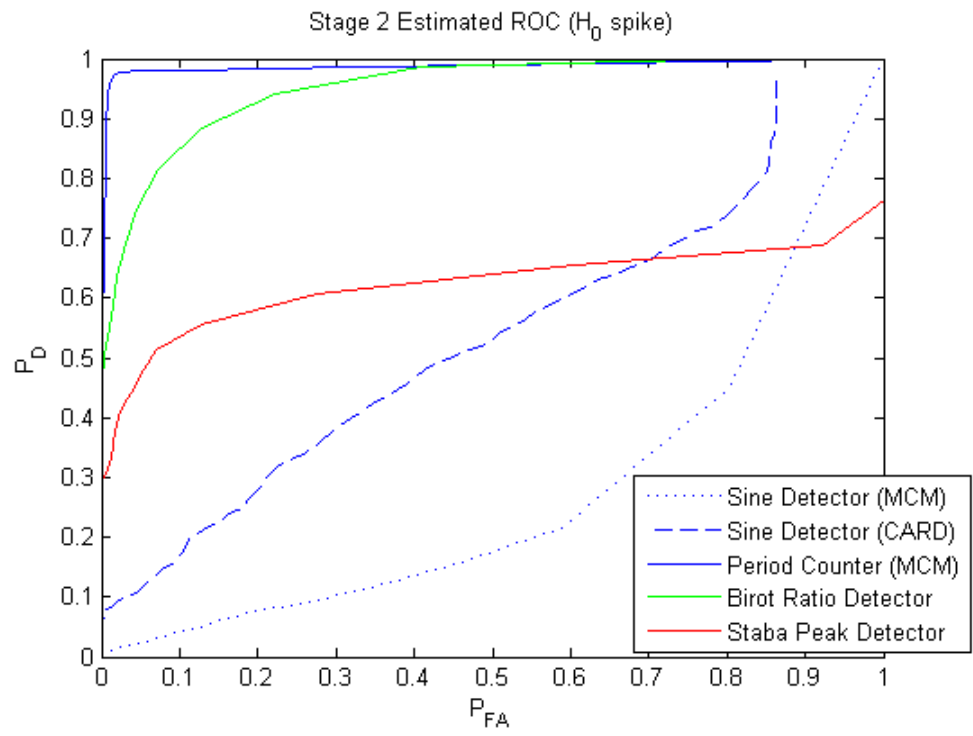


Figure 12. Detector Stage 2 Estimated ROC against H_0 simulated spike

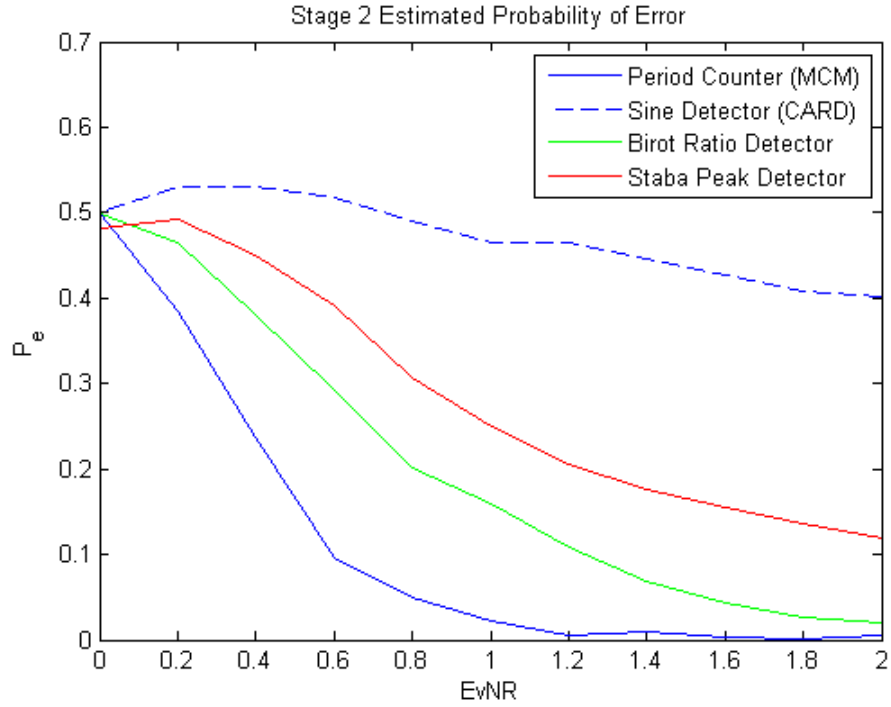


Figure 13. Detector Stage 2 Estimated Probability of Error

HFO Detector	Correct Detections	Missed Events	False Alarms
GLRT(MCM)	1349	186	1233
GLRT(CARD)	1242	293	1650
Birot	661	874	77
Staba	3	1532	50

Table 1. Detection Count on Real Data

3.1.3 Timing Comparison

A timing comparison was performed for the three detection routines. Results will vary based on computer hardware (processor speed, available RAM, etc.), the parameters of the study, and algorithm implementation; however, the intent is to show how the three routines compare for runtime. Timing test is based on running each detection routine in MATLAB on one minute of tEEG at 1600 Hz with two simulated HFOs and three spikes. Results are shown in table 2.

It is clear from these results that the improved performance of the GLRT

HFO Detector	Approximate Run Time (seconds)
GLRT(MCM)	307.984
GLRT(CARD)	317.878
Biro	1.462
Staba	0.659

Table 2. Algorithm Run Time

detection routine comes at the cost of processing speed.

3.2 SOZ Estimation

Five studies have been provided by CRE Medical for which a clinician has already identified the SOZ based on seizure data, previously used in [1]. HFO detection has been performed on all studies without any confirmation of accuracy of events—the intent is to see if the results from the detection routine alone can determine SOZ. Recordings were performed using TCRES connected to a Grass AuraLTM64 amplifier with recording at 200, 400, or 1600 Hz. Data is exported from Grass TWin software into MATLAB for applying the detection routine.

Previously described ϕ_{CH} values are calculated based on the results of HFO detection for each channel. This is done twice: first based on the results of the first stage of the detector only, then based on the results of the entire HFO detection routine. This is noted in the results tables as Stage 1 and Stage 2, respectively. The channels in **bold** font are those identified by the clinician. SOZ estimation results are based on HFO detection of the pre-ictal state, or the data collected leading up to the start of the first seizure in the study. The length of the tEEG segment considered varies by patient, up to half an hour in length, due to the availability of tEEG data preceding the first seizure.

Results are shown in tables 3 through 7 for the 5 studies. No threshold has been set for deciding if a channel contributes to the SOZ. Instead, it is assumed that the greater the ϕ_{CH} value, the more likely the channel is to contribute to the

Channel	Stage 1	Stage 2
F7	-0.0151	0.5660
F3	0.0389	0.4555
T3	-0.0078	0.4521
P4	0.0077	-0.0980
T6	0.0414	-0.1634
Fz	0.0211	-0.1750
C4	-0.0006	-0.1987
T4	0.0154	-0.2359
T5	0.0109	-0.2618
C3	0.0438	-0.4071
P3	0.1207	-0.5276
F8	0.1069	-0.5654
F4	-0.6350	-1.0000

Table 3. Patient 1 SOZ Results, 400 Hz study

SOZ. As can be seen from the data, results are varied and this technique alone cannot determine the SOZ as identified by the clinician.

List of References

- [1] W. G. Besio, I. E. Martinez-Juarez, O. Makeyev, J. N. Gaitanis, A. S. Blum, R. S. Fisher, and A. V. Medvedev, "High-frequency oscillations recorded on the scalp of patients with epilepsy using tripolar concentric ring electrodes," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 2, 2014.

Channel	Stage 1	Stage 2
Fp1	0.0961	0.8136
P3	0.1331	0.4559
T3	0.0395	0.1713
P4	0.0074	0.0618
F4	-0.0658	-0.1567
O1	-0.0945	-0.2045
F7	-0.0379	-0.2335
C4	-0.0510	-0.2639
Pz	0.3282	-0.3287
F3	-0.0573	-0.3287
Fp2	-0.0539	-0.3373
T6	-0.0623	-0.3459
Cz	-0.0718	-0.3907
Fz	-0.0554	-0.4094
C3	-0.0206	-0.4383
T4	-0.0623	-0.4094
T5	-0.0437	-0.4684
F8	-0.0976	-0.5321
O2	-0.0398	-0.5892

Table 4. Patient 2 SOZ Results, 1600 Hz study

Channel	Stage 1	Stage 2
O1	-0.0144	0.7002
C3	0.0831	0.3912
F8	-0.0067	0.3425
P4	0.0335	0.0168
T4	-0.0109	0.0025
P3	-0.0032	-0.0347
O2	-0.0050	-0.0746
F3	-0.0156	-0.0829
Cz	-0.0073	-0.0913
F7	-0.0032	-0.1445
T3	-0.0174	-0.1925
T6	-0.0055	-0.2444
Fz	-0.0168	-0.2663
F4	-0.0061	-0.2775
Fp2	-0.0044	-0.3243
C4	-0.0085	-0.3489
T5	-0.0132	-0.3615
Oz	-0.0150	-0.3615
Pz	0.0565	-0.5642
Fp1	-0.0332	-0.5808

Table 5. Patient 3 SOZ Results, 200 Hz study

Channel	Stage 1	Stage 2
P3	0.0116	0.6898
O2	-0.0040	0.5934
Fp2	0.0047	0.5265
O1	-0.0538	0.5265
Pz	0.0379	0.2946
F4	0.0042	-0.2425
T4	-0.0017	-0.3286
T5	-0.0058	-0.3286
C3	0.0927	-0.4264
Fz	-0.0070	-0.4264
Oz	-0.0005	-0.4264
T6	-0.0005	-0.4264
Fp1	-0.0443	-0.5385
Cz	0.0007	-0.5385
P4	0.0498	-0.5385
C4	0.0053	-0.6681
T3	-0.0171	-0.8199
F3	-0.0512	-1.0000
F7	-0.0386	-1.0000
F8	-0.0028	-1.0000

Table 6. Patient 4 SOZ Results, 200 Hz study

Channel	Stage 1	Stage 2
Cz	-0.1960	0.5827
O1	0.0055	0.2239
C4	-0.0125	0.2151
P4	0.0221	0.2061
F8	0.0328	0.1192
O2	0.0286	0.0617
Fp1	0.0058	0.0463
P3	0.0191	0.0140
C3	0.0118	-0.0029
T3	-0.0878	-0.0837
Fz	0.0304	-0.1202
T6	0.0164	-0.1546
F4	0.0293	-0.1909
F3	0.0266	-0.2196
F7	0.0225	-0.3025
Fp2	-0.0537	-0.3080
T4	0.0401	-0.4548
T5	0.0043	-0.5327
Pz	0.0092	-0.7327

Table 7. Patient 5 SOZ Results, 1600 Hz study

CHAPTER 4

Conclusion

Finding a superior HFO detector is a popular topic in neurophysiological research and determining a detector that works with the noise characteristics of surface-recorded TCRE data is a mostly unexplored field. The current approach that has been used for HFO detection in TCRE data, the spectrogram, is really an “eyeball” indicator to find high-power narrow-band activity before attempting a manual search. The detector presented here offers an automated approach to HFO detection and demonstrates that it is still possible to find a detector that performs well in less-than-ideal noise conditions.

With the Staba HFO detector, it is difficult to discern an HFO signal from an epileptic spike. Other studies [1, 2] have used the length of the detected signal to no avail since higher-frequency oscillations tend to be shorter signals. This is why it is important to weigh the length of the signal against its frequency in order to get an estimate of the number of oscillations. Spurious sinusoidal waveforms due to the high pass filtering of spike signals results in a lower number of oscillations per signal. The approach derived here combats the problem by defining two thresholds in the second detection stage. A threshold on the GLRT test statistic properly rejects signals with no discernible oscillations. A second threshold on the number of sinusoidal periods rejects signals with spurious oscillations due to filtering of epileptic spikes.

The voltage recorded at each contact site can vary based on a high number of factors, including electrode impedance. Therefore, the threshold of energy detectors are typically based on some average of the measured voltage for each channel. The issue with using an energy detector for HFO detection is the threshold will be

influenced by channels with a lot of events vs. those with none. A large number of EOIs will drive up the mean and cause more events to be missed. A channel with no such events will still cause detections. This effect worsens with surface recordings over iEEG due to the higher, and more frequently sporadic, environmental noise. The GLRT implementation on the other hand, compares consecutive windows to detect a change in the composition of a signal, regardless of event count. As shown, the GLRT generally performs well in practice.

The difficulty with automatic HFO detection using TCRES is there is a higher environmental noise when compared with intracranial recordings for which the HFO detectors in the literature are designed. The Staba algorithm was used as a baseline detector as it is compared to the most often and was the first automated method to demonstrate some success. The Birot algorithm has gained traction recently and is arguably the best method for detecting fast ripples from iEEG. It has been shown that the implementation of the GLRT presented here outperforms both of these methods asymptotically for detection of HFOs from surface recordings.

A significant finding from using the tEEG simulator at low EvNR values is that the sinusoid GLRT detector is able to locate oscillations that are not as obviously discernible from a manual review. This has the potential to pick up on more real events that would otherwise be missed by most detectors and some clinicians.

It should be noted that the comparison detectors were developed for iEEG whereas the detector presented here is derived with surface recordings in mind. Since recording HFOs from surface potentials is a fairly new discovery, there are not many routines in the literature to compare to. This is why the author selected the detectors he thought would best adapt to tEEG data. As HFO recording from scalp electrodes gains more traction, finding the best HFO detector will become a

more popular pursuit. To level the playing field with the comparisons presented here, a next step should be to use this same detector on real and simulated iEEG to further evaluate performance and determine if the GLRT-based algorithms offer any advantage over other detectors.

The tradeoff with the detector presented is run-time. Both stages operate by comparing single-sample shifted windows which can take a long time to process. Run time can be reduced in stage 1 by using less of an overlap between windows and in stage 2 by defining upper and lower limits on the duration of the signal, both at the cost of a lower probability of detection. However, even with these adaptations the processing performance will not rival the detectors presented as comparisons. In practice, EEG recordings can run anywhere from several hours to several days in duration. For this detector to be clinically tractable, the clinician would not need immediate results and its implementation would have to be clever enough to run a post-study batch analysis on a server or several PCs operating on a single channel at a time. Faster PCs are available today than the one used in this study and more RAM and higher processing speeds continue to be offered. Implementation for the comparison here is also performed in MATLAB; implementing in a language that is able to take advantage of multiple threads and better memory allocation has the potential to improve processing speeds. Although this detection routine is slow to run today, the algorithms presented will only become more practical as technology advances.

An argument could be made against the performance evaluation that the simulator design is based on a TVAR process which is an assumption of the derived anomaly detector. However, this implementation is based on the nature of neurophysiological data and the simulator was designed to be as accurate as possible. High accuracy is expected if the observed signal is in agreement with the model

[3]. The MATLAB simulator is a good representation of time-varying tEEG signals that contain spontaneous epileptic spikes and HFOs. In practice, there will be more types of neurophysiological and environmental noise that are difficult to account for in a simulation. It should also be understood that detectors evaluated on computer simulated data will almost always out-perform an equivalent in-field study; however, simulation is necessary for an algorithm implementation to define a meaningful performance metric and compare to competitor algorithms [4]. The results show that the detection routine derived as part of this thesis have good performance on both real and simulated tEEG data.

It was found that the CARD estimation is an unnecessary step in the majority of the simulations. High-pass filtering reduces, but does not eliminate, the background EEG described by the AR model. When an HFO signal is present and the window size is kept short the filtered signal window chosen by the anomaly detector should be mostly dominated by the sinusoidal frequency. Therefore using strictly a frequency peak estimate and basing the pre-whitener on the AR filter described by preceding data works well when compared to the CARD method which assumes the signal is overshadowed by AR noise. The conclusion here is that this approach is the best detector to use when overall run-time is not a concern.

The GLRT-based detection routines performed well on the small subset of real tEEG data at the cost of a high number of false alarms. In both implementations it was found that the active HFO channels contributed more to the false alarm count than those channels with no events. The Birot routine contributed a fewer number of false alarms but consequently missed more than half of the actual HFOs. The Staba method is clearly unfit for HFO detection using scalp recordings.

Unfortunately, the SOZ estimation method offered here did not work to positively identify electrode contacts corresponding to the brain regions responsible

for propagating seizures in the patients studied. Clinician-identified SOZ channels can be fairly subjective making it a difficult item to test against; however it is important that any SOZ estimation method be well validated since incorrectly identifying the SOZ can have dire consequences in practice [5]. The latter goal of determining SOZ from the anomaly detector only is now understood to be flawed—epileptic spike signals appear at a higher rate than HFOs in all examined studies and therefore the resulting phi estimate will always be low.

A typical 10-20 EEG montage has 21 electrodes whereas intracranial can often use hundreds of patient contacts at a time [3]. Therefore, even if it is possible to get a rough estimate of the SOZ from a tEEG surface recording it is not going to be as accurate, nor as localized, as an intracranial recording. However, it can be a first step to determining the SOZ and offer the clinician an idea of the general area to look before an open-brain procedure is necessary to perform a functional brain mapping or lesionectomy. The assumption that HFOs can be used to identify the SOZ remains a common topic in neurology research and the efforts of doing so using tEEG data should not be abandoned. It is likely that SOZ estimation from HFO detection will only work for epileptic patients with well-localized epileptic zones and not those suffering from generalized onset seizures. The study in [6] suggests channels in which high frequency activity shows up earlier preceding seizure onset are a positive indicator of SOZ. Therefore, next efforts should be made with TCRE HFO detection to not only base the SOZ estimate on event occurrence in the preictal state, but to add a greater weight to positively-identified HFOs the sooner they occur in an epileptic study.

List of References

- [1] N. Roerhi, F. Pizzo, F. Bartolomei, F. Wendling, and C. G. Benar, “What are the assets and weaknesses of HFO detectors? a benchmark framework based on realistic simulations,” *PLoS ONE*, vol. 12(4), 2017.

- [2] M. Amiri, J. M. Lina, F. Pizzo, and J. Gotman, “High frequency oscillations and spikes: Separating real HFOs from false oscillations,” *Clinical Neurophysiology*, vol. 127, pp. 187–196, 2016.
- [3] L. Sornmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. London, UK: Elsevier Academic Press, 2005.
- [4] S. M. Kay, *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*. Westford, MA: Prentice Hall, 2013.
- [5] F. Rosenow and H. Luders, “Presurgical evaluation of epilepsy,” *Brain*, vol. 124(9), pp. 1683–1700, 2001.
- [6] F. Bartolomei, P. Chauvel, and F. Wendling, “Epileptogenicity of brain structures in human temporal lobe epilepsy: A quantified study from intracerebral EEG,” *Brain*, vol. 131, pp. 1818–1830, 2008.

BIBLIOGRAPHY

- Amiri, M., Lina, J. M., Pizzo, F., and Gotman, J., “High frequency oscillations and spikes: Separating real HFOs from false oscillations,” *Clinical Neurophysiology*, vol. 127, pp. 187–196, 2016.
- Bartolomei, F., Chauvel, P., and Wendling, F., “Epileptogenicity of brain structures in human temporal lobe epilepsy: A quantified study from intracerebral EEG,” *Brain*, vol. 131, pp. 1818–1830, 2008.
- Benar, C. G., Chauviere, L., Bartolomei, F., and Wendling, F., “Pitfalls of high-pass filtering for detecting epileptic oscillations: a technical note on “false” ripples,” *Clinical Neurophysiology*, vol. 121(3), pp. 301–310, 2010.
- Besio, W. G., Koka, K., Aakula, R., and Dai, W., “Tri-polar concentric ring electrode development for laplacian electroencephalography,” *IEEE Transactions on Biomedical Engineering*, vol. 53(5), pp. 926–933, 2006.
- Besio, W. G., Martinez-Juarez, I. E., Makeyev, O., Gaitanis, J. N., Blum, A. S., Fisher, R. S., and Medvedev, A. V., “High-frequency oscillations recorded on the scalp of patients with epilepsy using tripolar concentric ring electrodes,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 2, 2014.
- Biot, G., Kachenoura, A., Albera, L., Benar, C., and Wendling, F., “Automatic detection of fast ripples,” *Journal of Neuroscience Methods*, vol. 213(2), pp. 236–249, 2013.
- Burnos, S., Frauscher, B., Zelmann, R., Haegelen, C., Sarnthein, J., and Gotman, J., “The morphology of high frequency oscillations (HFO) does not improve delineating the epileptogenic zone,” *Clinical Neurophysiology*, vol. 127(4), pp. 2140–2148, 2016.
- Burnos, S., Hilfiker, P., Surucu, O., Scholkmann, F., Krayenbuhl, N., Grunwald, T., and Sarnthein, J., “Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis,” *PLoS ONE*, vol. 9(4), 2014.
- Chaibi, S., Lajnef, T., Sakka, Z., Samet, M., and Kachouri, A., “A reliable approach to distinguish between transients with and without HFOs using TQWT and MCA,” *Journal of Neuroscience Methods*, vol. 232, pp. 36–46, 2014.
- DiLorenzo, D. J. and Bronzino, J. D., *Neuroengineering*. Boca Raton, FL: Taylor and Francis Group, 2008.

- Dumpelmann, M., Jacobs, J., and Schulze-Bonhage, A., “Temporal and spatial characteristics of high frequency oscillations as a new biomarker in epilepsy,” *Epilepsia*, vol. 56(2), pp. 197–206, 2015.
- Fahoum, A. and Fraihat, A., “Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains,” *ISRN Neuroscience*, 2014.
- Gardner, A. B., Worrell, G. A., Marsh, E., Dlugos, D., and Litt, B., “Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings,” *Clinical Neurophysiology*, vol. 118(5), pp. 1134–1143, 2007.
- Ghaleb, I., Davila, C., and Srebro, R., “Prewhitening of background brain activity via autoregressive modeling,” *IEEE*, pp. 242–245, 1997.
- Gliske, S. V., Irwin, Z. T., Davis, K. A., Sahaya, K., Chestek, C., and Stacey, W. C., “Universal automated high frequency oscillation detector for real-time, long term EEG,” *Clinical Neurophysiology*, 2015.
- Huang, L., Ni, X., Ditto, W., Spano, M., Carney, P., and Lai, Y. C., “Detecting and characterizing high-frequency oscillations in epilepsy: a case study of big data analysis,” *R. Soc. open sci.*, vol. 4, 2017.
- Jacobs, J., Staba, R., Asano, E., Otsubo, H., Wu, J. Y., Zijlmans, M., Mohamed, I., Kahane, P., Dubeau, F., Navarro, V., and Gotman, J., “High-frequency oscillations (HFOs) in clinical epilepsy,” *Progress in Neurobiology*, vol. 98(3), pp. 302–315, 2012.
- Jmail, N., Gavaret, M., Wendling, F., Kachouri, A., Hamadi, G., Badier, J.-M., and Benar, C., “A comparison of methods for separation of transient and oscillatory signals in EEG,” *Journal of Neuroscience Methods*, vol. 199(2), pp. 273–289, 2011.
- Kaipio, J. P. and Karjalainen, P. A., “Simulation of nonstationary EEG,” *Biological Cybernetics*, vol. 76(5), pp. 349–356, 1997.
- Kay, S., “A new nonstationarity detector,” *IEEE Transactions on Signal Processing*, vol. 56(4), pp. 1440–1451, 2008.
- Kay, S. M., *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1988.
- Kay, S. M., *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993.
- Kay, S. M., *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, 1998.

- Kay, S. M., *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*. Westford, MA: Prentice Hall, 2013.
- Kobayashi, K., Agari, T., Oka, M., Yoshinaga, H., Date, I., Ohtsuka, Y., and Gotman, J., “Detection of seizure-associated high-frequency oscillations above 500 hz,” *Epilepsy Research*, vol. 88, pp. 139–144, 2010.
- Koka, K. and Besio, W. G., “Improvement of spatial selectivity and decrease of mutual information of tri-polar concentric ring electrodes,” *Journal of Neuroscience Methods*, vol. 165(2), pp. 216–222, 2007.
- Makeyev, O., Liu, X., Wang, L., Zhu, Z., Taveras, A., Troiano, D., Medvedev, A. V., and Besio, W. G., “Feasibility of recording high frequency oscillations with tripolar concentric ring electrodes during pentylenetetrazole-induced seizures in rats,” *34th Annual International Conference of the IEEE EMBS*, pp. 4599–4602, 2012.
- Malinowska, U., Bergey, G. K., Harezlak, J., and Jouny, C. C., “Identification of seizure onset zone and preictal state based on characteristics of high frequency oscillations,” *Clin Neurophysiol*, vol. 126(8), pp. 1505–1513, 2015.
- Misulus, K. E. and Head, T. C., *Essentials of Clinical Neurophysiology*. Burlington, MA: Butterworth Heinemann, 2003.
- Nagesha, V. and Kay, S., “Spectral analysis based on the canonical autoregressive decomposition,” *IEEE Transactions on Signal Processing*, vol. 44(7), pp. 1719–1733, 1996.
- Ombao, H., Raz, J., Sachs, R. V., and Guo, W., “The SLEX model of a non-stationary random process,” *Ann. Inst. Statist. Math.*, vol. 54(1), pp. 171–200, 2002.
- Pardey, J., Roberts, S., and Tarassenko, L., “A review of parametric modelling techniques for EEG analysis.”
- Pearce, A., Wulsin, D., Blanco, J. A., Krieger, A., Litt, B., and Stacey, W. C., “Temporal changes of neocortical high frequency oscillations in epilepsy,” *Journal of Neurophysiology*, vol. 110(5), pp. 1167–1179, 2013.
- Roehri, N., Lina, J. M., Mosher, J. C., Bartolomei, F., and Benar, C. G., “Time-frequency strategies for increasing high-frequency oscillation detectability in intracerebral EEG,” *IEEE Transactions on Biomedical Engineering*, vol. 63(12), 2016.
- Roehri, N., Pizzo, F., Bartolomei, F., Wendling, F., and Benar, C. G., “What are the assets and weaknesses of HFO detectors? a benchmark framework based on realistic simulations,” *PLoS ONE*, vol. 12(4), 2017.

- Rosenow, F. and Luders, H., “Presurgical evaluation of epilepsy,” *Brain*, vol. 124(9), pp. 1683–1700, 2001.
- Sornmo, L. and Laguna, P., *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. London, UK: Elsevier Academic Press, 2005.
- Staba, R. J., Wilson, C. L., Bragin, A., Fried, I., and Jr, J. E., “Quantitative analysis of high-frequency oscillations recorded in human epileptic hippocampus and entorhinal cortex,” *Journal of Neurophysiology*, vol. 88(4), pp. 1743–1752, 2002.
- Worrell, G., Jerbi, K., Kobayashi, K., Lina, J., Zelmann, R., and Quyen, M. L. V., “Recording and analysis techniques for high-frequency oscillations,” *Progress in Neurobiology*, vol. 98(3), pp. 265–278, 2012.