Epileptiform Detection Algorithm (V8.6)

Authors: Michael Chang, Christopher Lucasius, Liam Long and Taufik A. Valiante

Table of Contents

Summary of Detection Algorithm

The Epileptiform Detection Algorithm (herein referred to as the "algorithm") operates on the fundamental basis that epileptiform events are comprised of spikes, either individual spikes (i.e., interictal spikes) or a collection of spikes (i.e. interictal events, interictal bursts, or Ictal events, a.k.a. seizure-like events). Spikes that are near each other (i.e. within 10 seconds) are considered as part of the same event. The detected events are then classified using unsupervised machine learning (k-means clustering, where k = 2) on two engineered features that can discern the shape of epileptiform events from brain slice recordings, as either an ictal event or interictal event. The algorithm's unsupervised classifier is highly accurate (>94%) because a multitude of fail-safes are built on top of the k-means clustering algorithm (i.e. removal of artifacts, and implementation of a floor-threshold, hard-cored threshold and Michael's customized thresholds). Most importantly, the algorithm is designed to be fast, require no training (with labelled data), and user-friendly for researchers with no computer programming experience.

Abstract

The Epileptiform Detection Algorithm (herein referred to as the "algorithm") can detect and classify epileptiform events (and artifacts) from the seizure models presented in Chang et al., 2018b, *JoVE*. The algorithm works in three stages: Detection → Feature Extraction → Classification. First, the algorithm isolates the baseline activity of the local field potential (LFP) recording and locate all the epileptiform

spikes. Epileptiform spikes near each other (i.e., 10 s) are grouped as a single “event”. A 'crawler' function then locates the exact onset and offset of the event, according to the methods described in Chang et al., 2018a, *Neurobio Dis*. Second, the algorithm extract features from the detected events that are specially engineered to capture the unique spatial and temporal dynamics of events from *in vitro* brain slices. Third, the algorithm uses unsupervised machine learning, k-means clustering (where k = 2), on these engineered features to classify detected events either as an ictal event or interictal event. Many fail-safe mechanisms (i.e. floor thresholds, hard-coded thresholds, etc.) are built on top of the k-means clustering algorithm to ensure high accuracy in a wide range of situations. The algorithm is fast, requires no training, and correctly classifies events and artifacts >90% of the time, which is on par with humans.

Acknowledgements

This work is the cumulative effort over 3 years by Christopher Lucasius (initial frame work for code and plotting figures to create .pptx presentation), Fu-der (Fred) Chen (signal processing and strategies to isolate epileptiform events), Liam Long (Advice and guidance on coding in Matlab, Boolean logic), Thomas Lordello (advanced tips in Matlab), Kramay Patel (helped to code spectral energy and template matching with convolution), David Groppe (understanding k-means clustering was the best way to proceed with this algorithm), Uilki (discussions on signal processing and strategies to isolate baseline), Vitaly Topekha (studying the population of interictal spikes and artifacts), Dr. Taufik A. Valiante (for his professional insights into ictal events and motivation to build the IIE classifier), and Michael Chang (manually studying ictal events over 4 years for his PhD and writing the detection algorithm based on the cumulative knowledge and combining together everyone’s ideas). We would also like to thank the research assistants: Shadini Dematagoda, Alina Rizvi, and Barret Kaplan that helped to annotate and organize the ictal events and LFP recordings.

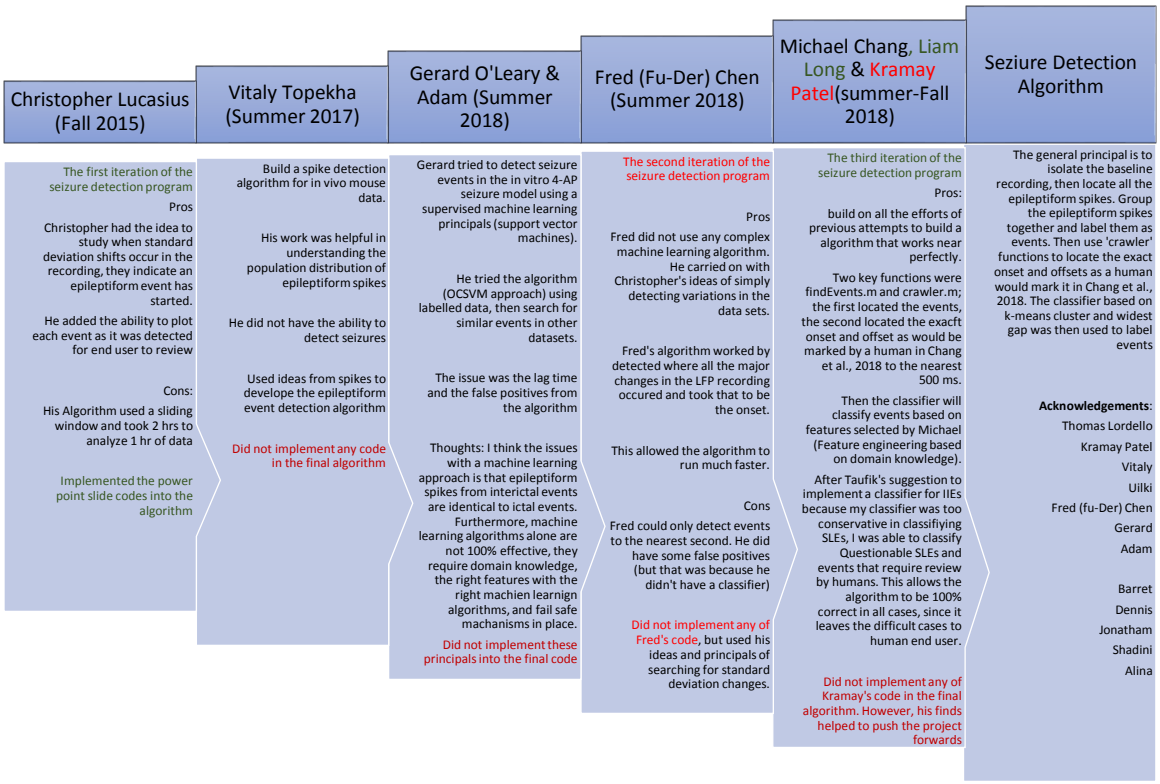


Figure 1. Development timeline of the Epileptiform Detection Algorithm

Prelude:

- The algorithm was developed based on the labelled dataset from Chang et al., 2018a.
- Different events from 4-AP seizure models described in Chang et al., 2018b were labelled as ictal event, interictal event, and artifacts.
- Labelling was supervised by a team of clinically practicing medical doctors specializing in epilepsy, as well as, PhD candidates studying epilepsy models from the Toronto Western Hospital

Introduction

The strategy upon which the algorithm was designed allows it to be fast, light-weight, and require no prior training. It correctly classifies events and artifacts >90% of the time, which is on par with human experts (Gaspard et al., 2014). The algorithm is user-friendly for life science students and researchers with minimal computer programming experience. Furthermore, the algorithm is comprised of three separate modules for each stage of the algorithm (Spike Detection → Feature Extraction → Classification). This feature allows more advanced users with programming experience the opportunity to modify modules for different seizure models. The code to the algorithm can be found on the Valiante Lab's GitHub Repository and a video tutorial is also available.

Stage 1: Spike Detection

The detection algorithm is based on the standard techniques of filtering the signal and locating peaks in in the signal. The algorithm detects spikes using the *findpeaks* function. Epileptiform events in neural recordings are visually observed as spikes (deviations from the baseline) in the time series. Spikes detected in the time series are then assumed to be epileptiform spikes. The detection algorithm dynamically sets a threshold to detect spikes based on a hyperparameter determined by the end user; the threshold is based on the baseline voltage activity, where epileptiform spike threshold = average baseline activity + sigma of baseline activity*hyperparameter). The optimal hyperparameter is 3.9 for in vitro data is, and 10 for in vivo data (or noisy data) based on the recordings from (Chang et al., 2018a). Future iterations of the algorithm will perform a hyperparameter grid search (trial and error of a wide range of hyperparameters) to find the best optimal hyperparameter if the recommended one fails to capture ictal event[1].

Based on the literature, epileptiform events need to be at least 10 seconds apart (reference), thus spikes that are separated by more than 10 seconds are from separate epileptiform events. The first spike being the end of the preceding event, and the second spike being the start of the subsequent event. This is the fundamental basis of the detection algorithm. Once the first and last spikes of all the epileptiform events are detected, a second functions *Crawler.m*, finds the exact onset and offset point based on when the signal's power (feature) increases or decreases, respectively. The crawler function can be modified to detect the onset and offset of different events from other types of seizure models or clinical recordings. The strength of the detection algorithm is its speed to analyze datasets and its ability to

[1]The algorithm can determine how well captured ictal events from the recording based on the intensity ratio of what it deems 'ictal events'; if the intensity ratio is <0.3 it is likely not a single isolated ictal event, and will repeat the process with another hyperparameter or perform hyperparameter grid searching to determine which hyperparameter will result in detected ictal events with the highest intensity ratio.

dynamically set the thresholds for detecting events from many different *in vitro* recordings of the 4-AP seizure model.

Stage 2: Feature Extraction

Domain knowledge of ictal events from the *in vitro* 4-AP model (Chang et al., 2018a, *Neurobio Dis*) was used to engineer new features that were effective in capturing the spatial-temporal aspects of ictal events. In total, the algorithm uses four features to classify detected events: "Frequency", "Intensity", Duration, and peak-to-peak amplitude.

Engineered Features (averaged value of feature)

The key features of ictal events that distinguishes it from interictal events is that their population spikes (deflections in the LFP) occur at a higher frequency (>1 Hz) and higher intensity (>15 mV^2/s). Thus, we engineered two features to capture those aspects of the ictal events:

1) **Frequency**, of population spikes in the LFP recording (herein referred to as "frequency")
2) **Intensity**, defined as $\text{total power}/\text{duration}$, where total power = $\sum \text{amplitude}^2$

These two engineered features work effectively with k-means clustering (where K = 2) to distinguish epileptiform events as either an putative ictal event or interictal event.

Standard Features (total value of feature)

Interictal spikes and artifacts also have a high frequency and high intensity. For this reason, they can also be misclassified as an ictal event. Duration is used as a third feature to find any interictal spikes that may be misclassified as an ictal event.

3) **Duration**, length of the epileptiform event

All events classified as a putative ictal event are only classified as an ictal event if their duration is longer than the threshold. Threshold is set in the following manner:

> Duration Threshold is either Michael's Customized Threshold or Algo determined threshold, depending on whichever one is lower. If both are below the floor threshold, the floor threshold is used. The floor threshold is determined based on the literature and findings from Chang et al., 2018a

- **Michael's customized Threshold** = duration of the average putative ictal event – 2*sigma of putative ictal events
  or
  Michael's customized Threshold = Sigma, if average ictal event < 2*sigma of putative ictal event
- **Floor threshold** = 10 seconds
- **Algo determined threshold** = k-means clustering determine threshold + widest gap

Artifacts can also be filtered out based on their peak-to-peak amplitude, which is significantly (i.e., 15x) larger than physiological spiking activity in the LFP recording. If any spikes labelled as an outlier in amplitude (because it's way too large), it's labelled as an artifact.

4) **Peak-to-peak amplitude**, the maximum amplitude – the minimum amplitude of an event

The classifier uses these four features described above to classify the epileptiform events detected (which are just a collection of spikes that make an event) as either an ictal event, interictal event, or artifact. In the future, artifacts can also be filtered out based on their narrow width; **if they are less than 10 ms, they are likely non-physiological** (this rule will be implemented in future versions).

Engineered Features (per second values during the event)

The key features about epileptiform events that distinguish SLEs from IIEs, and IIEs from IISs are

- **Tonic phase**, the frequency (population spike rate) at each second through the event
    - SLEs must have a tonic phase, or tonic firing, which is high frequency (>1 Hz) population activity that is consistently maintained for at least 2 seconds. The clonic phase begins when the population firing activity decreases and is interspersed with no population activity (i.e 0 Hz).

- **Intensity Ratio**, the power (amplitude^2) at each second through the event.
    - The power content of an event can be described by analyzing the total period of time the event is at high power, a feature I called the 'intensity ratio' (time at high power/total time). These features can be used to create a classification system to classify interictal events as questionable SLEs (QSLEs) or IISs.

Stage 4: Classifier

The classification algorithm is based on k-means clustering, where k = 2 (the ictal events and interictal events). K-means clustering is performed on two specially engineered features: average rate of population spikes (herein referred to as "frequency") and average the amplitude of voltage activity squared (herein referred to as "intensity"). These specially engineered features are designed to work effectively with the k-means clustering (k=2) algorithm in classifying epileptiform events detected from *in vitro* 4-AP treated brain slices, as either interictal events or ictal events. Where the events with a larger average value (for whatever feature) classified as the ictal event.

The threshold resulting from k-means clustering algorithm is the mid-point between the IIE with the largest value and the SLE with the lowest value. It is important to note that the results of k-means clustering are inconsistent if data is seeded randomly. To address this, the k-means clustering algorithm is repeated 25x, and the lowest threshold value is used.

The following fail-safe mechanisms are in place

- The widest gap between data points for a feature

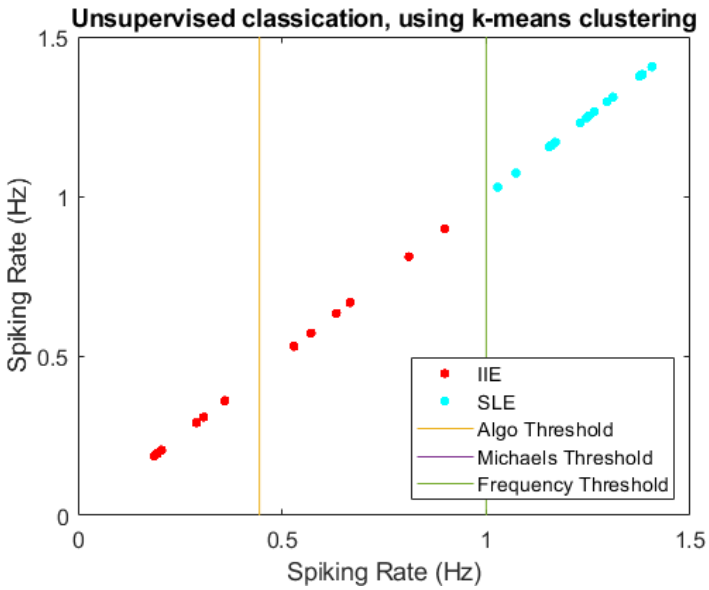


Figure 2. The algo-determined threshold (yellow) is based on k-means clustering but will move to a lower threshold if there is a wider gap that exist between data points for a feature (i.e. spiking rate).

- Michael's customized threshold
    - Based on empirical experience from Chang et al., 2018a
    - Usually some algorithm based on population statistics of event (i.e. average – sigma)
    - Used if it is a lower threshold than the algo determined threshold
- More liberal Threshold
    - Between the algorithm-determined threshold and Michael-determined threshold, the lower value is used as the threshold to classify SLEs, unless it is below a floor value (the minimum value).

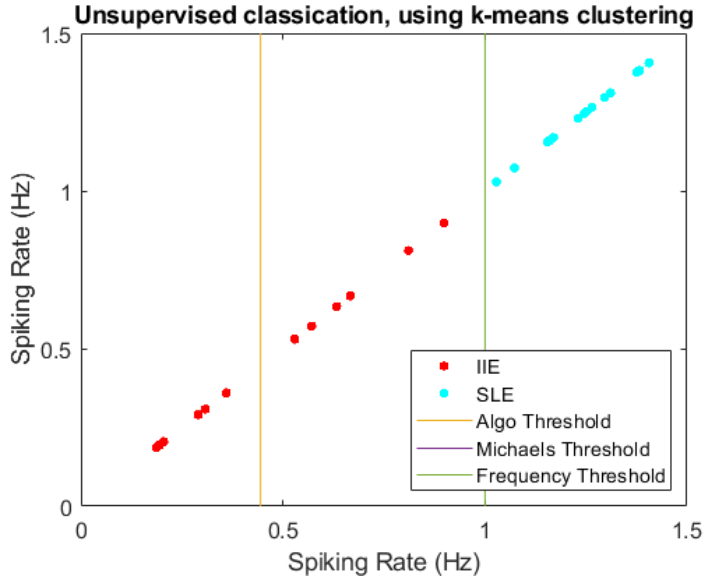
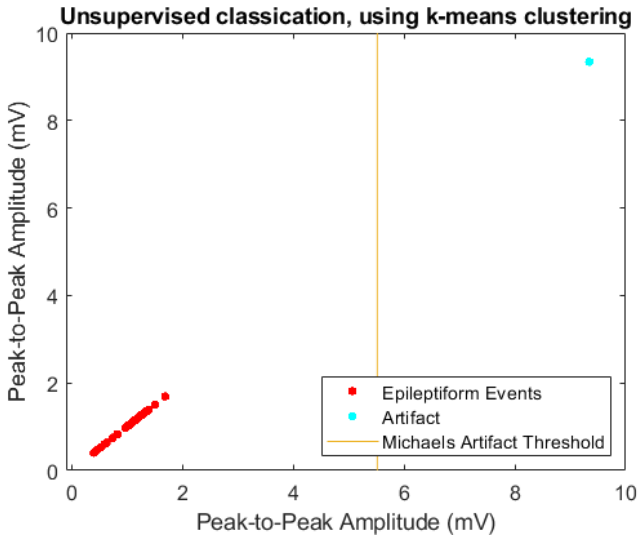


Figure 3. Since the algo determined threshold for frequency (yellow line) is below the floor threshold for frequency (1 hz) the floor frequency is used

- Artifact Removal
    - Based on the feature: peak-to-peak amplitude
    - Threshold for artifact is determined as
    - The widest gap between features, if there is an existing feature that is greater than 4x the population's sigma



Test Script (Demonstration Purposes)

- Github Link
- Video Tutorial to use algorithm
- Run the code
- Explain the input and output from the program

Results:

Training Algorithm

There are seven .abf files from Chang et al., 2018; these files were labelled by human makers under the supervision of a neurologist. The algorithm was developed to analyze these files. The algorithm detected 100% of the ictal events that were detected by human markers (140 events from these 7 .abf files), plus an additional 4 ictal events that were missed by human markers. The algorithm made 1 false positive detection (a preictal event was detected as an ictal event). Importantly, there were no false negatives, and no human-detected ictal events were missed. The algorithm's final score in analyzing this "training" dataset was +3 and considered on par with humans in detecting epileptiform events. The ictal events' onset and offset detected by the algorithm differed from human markings on average by 0.43 seconds and 10.62 seconds, respectively.

Table 1. Statistics for difference between Algorithm detection vs Human detection. A score of +1 awarded if algorithm detects an ictal event humans missed, and -1 if algorithm missed an ictal event marked by human. SLE detected is by algorithm, SLE total is how many ictal events in the recording (confirmed by human). Onset (Diff) and Offset (Diff) is the average (absolute) difference between onset detected by algorithm and by human.

FILENAME	SCORE	ALGO DETECTED	HUMAN DETECTED	ONSET (DIFF), AVG	OFFSET (DIFF), AVG
13226009	0	16	15	0.04 (+.02)	8.53 (-.05)

13227004	1	12(+7)	18	0.01 (-0.02)	9.05
13725002	2	22(+2)	19	0.02(+.01)	0.10 (+.01)
13725005	0	33	33	0.02(+.01)	0.23(+0.01)
14609000	0	15(+6)	14	0.63(-0.03)	2.89(+0.02)
16125003	0	23	23	2.26	4.58(-0.08)
16201017	0	10	10	0.02(+0.01)	48.97(+0.11)
TOTAL	**+3**	**131(+15)**	**140**	*0.43 (avg)*	*10.62 (avg)*

Note: After reviewing of the results of the algorithm's detection points, I would argue the algorithm is more faithful and accurate than human markers. Human markers were biased in the selection of onset positions by a few ms to demonstrate light-triggered correlation. Furthermore, detecting the exact location of ictal event offset is extremely difficult and subjective.

Ictal Event Characteristics

Table 2. Statistics for the features of all the ictal events (7 abf files) from chang et al., 2018. Frequency is the average spike rate per minute; intensity is the average power per minute; amplitude is the max peak-to-peak amplitude value.

FEATURE	**FLOOR**	**CEILING**	**AVERAGE**
DURATION	12.84	194.00	67.41
FREQUENCY	1.03	4.00	2.35
INTENSITY	16.21	76.28	32.43
AMPLITUDE	0.58	2.14	1.22

Table 3. The intensity ratio statistics the light-triggered ictal events from each file from 7 .abf files from Chang et al., 2018.

FILE	**13226009**	**13227004**	**13725002**	**13725005**	**14609000**	**16125003**	**16201017**	AVG
MEAN	0.56	0.50	0.82	0.88	0.77	0.66	0.57	0.68
STDEV	0.08	0.14	0.11	0.10	0.12	0.12	0.02	0.10
MAX	0.70	0.79	0.95	1.00	0.92	0.91	0.61	1.00
MIN	0.46	0.27	0.43	0.62	0.56	0.40	0.53	0.27
COUNT	16	17	23	33	15	23	10	

Note: the intensity ratio is the amount of time event has high intensity divided by the total duration of event. Intensity is considered high if it's at least 1/3 the max intensity value during the event.

From studying 137 ictal events (a majority of which are light-triggered) from 7 recordings, it was found that the average intensity ratio (when all values were grouped together) was 0.7±0.2 mW^2/s. The floor intensity for ictal events was 0.4 mW^2/s, however, one ictal event (out of the 137 analyzed) was observed to have a intensity ratio as low as 0.27 mW^2/s. The average intensity ratio for each recording was 0.68 mW^2/s.

Table 4. The average ictal event is characterized by spiking rates that change over the course of 3 main phases: preictal, tonic-like, and clonic-like; averages of light-triggered ictals events from 7 recording.

	PREICTAL FREQ, AVG	**TONIC FREQ, AVG**	**CLONIC FREQ, AVG**	**TONIC FREQ, MIN**

AVERAGE	2.54884523	3.341822624	1.450735803	1.526649118
STDEV	0.815611207	0.441700854	0.407066319	0.616959118
MAX	5.5	4.536585366	4	4
MIN	0.6	1.9	0.779220779	1

Note: The average ictal event (75% are light-triggered) from 7 recordings. Each slice's average is then averaged to report the values in this table.

The different phases of an ictal event are characterized in relation to the tonic-like phase where high frequency activity occurs during the ictal event. The tonic-like phase is considered to be the region of the ictal event that is characterized by high frequency spiking activity. High frequency activity is considered to be when the spiking rate per second is 1/3 the maximum spiking rate per second during the event (or at least 1 Hz, floor threshold). The tonic phase begins when there are two consecutive seconds of high frequency activity and considered over when there are two consecutive seconds of low frequency activity (or the spiking rate drops to 0 Hz for 1 second–because of the results in Table 4). The Preictal phase is the region between the sentinel spike and beginning to the tonic phase. The post-ictal phase is the region after the tonic phase.

From studying 137 ictal events (a majority of which are light-triggered) from 7 recordings, it was found the minimum floor for the tonic-like phase of ictal event was 1 Hz, so this was set as the floor threshold for the tonic-like phase in the algorithm (Table 4). Furthermore, the greatest variability in spiking rate was observed during the preictal phase (Table 4), as expected. the highest frequency activity occurs during the tonic-like phase, on average, and expected because the algorithm was designed to classify the region with the highest, continuous, frequency activity as the tonic-like phase.

These results (statistical analysis) was used to develop the *classifer_in_vitro.m*, with hard-coded thresholds and the floor value threshold for *classifierDyanmic.m*. The floor values for Duration is 10 secs, Frequency is 1 Hz, intensity is 10 mW^2/s, and ceiling for amplitude is 4.9 mV. The floor value for the tonic-like phase was 1 Hz (Table 4).

Testing Algorithm

Detection Algorithm V5.3 was tested on six different .abf files from Chang et al., 2018; these files were recordings of activity from *in vitro* 4-AP cortical seizure model, wild-type C57/Bl6 mice. These observed spontaneous epileptiform events were labelled by human makers under the supervision of a neurologist. The algorithm detected 93% of the ictal events that were detected by human markers (106/114 events from 6 .abf files), plus an additional 12 ictal events that were missed by human markers. Unfortunately, the algorithm made 6 false positive detections (preictal events and IIEs with large amplitudes were mislabelled as ictal events). This algorithm final score in this "testing" dataset was -2 and considered to be on par with humans in detecting epileptiform events (Gaspard et al., 2014). The spontaneous ictal events' onset and offset detected by the algorithm differed from human markings on average by 0.83 seconds and 1.60 seconds, respectively.

Table 5. Statistics on the results of the algorithm analyzing spontaneous activity from *in vitro* 4-AP cortical seizure model from WT C57/Bl6 mice

FILENAME	SCORE	ALGO DETECTED	HUMAN DETECTED	ONSET (DIFF), AVG	OFFSET (DIFF), AVG

13802004	-3	30	33	0.27	0.32
13802006	-2	20	19	0.10	0.82
13803004	2	20	20	1.46	0.57
13805000	2	8	9	2.79	0.27
14528001	-2	14	16	0.32	5.36
14528003	1	15	17	0.06	2.29
TOTAL	**-2**	**107**	**114**	*0.83 (avg)*	*1.60 (avg)*

Table 6. Statistics on the intensity ratio of spontaneous ictal events from WT C57/Bl6 mice, 6 recordings in 4-AP in vitro cortical slices

	13802004	13802006	13803004	13805000	14528001	14528003
AVERAGE	0.62	0.51	0.55	0.63	0.71	0.24
STDEV	0.12	0.16	0.19	0.13	0.20	0.07
MAX	0.82	0.83	0.81	0.91	0.91	0.48
MIN	0.37	0.10	0.17	0.50	0.38	0.15
COUNT	30.00	40.00	21.00	11.00	14.00	17.00

Note: recording 14528003 had more post-ictal spiking than average, as a result, the interictal spike connected two (or more) ictal events together (Figure 4. The result is detected events with low intensity ratios (avg. 0.24) because there is a lot of silence (interictal period) in between the two ictal events. So, I plan to build into the detection algorithm a While-loop, to repeat the analysis at a higher Sigma for threshold detection, i.e., 10, if the avg. intensity ratio detected is below 0.3; and issue a warning for human review if it is below 0.5.

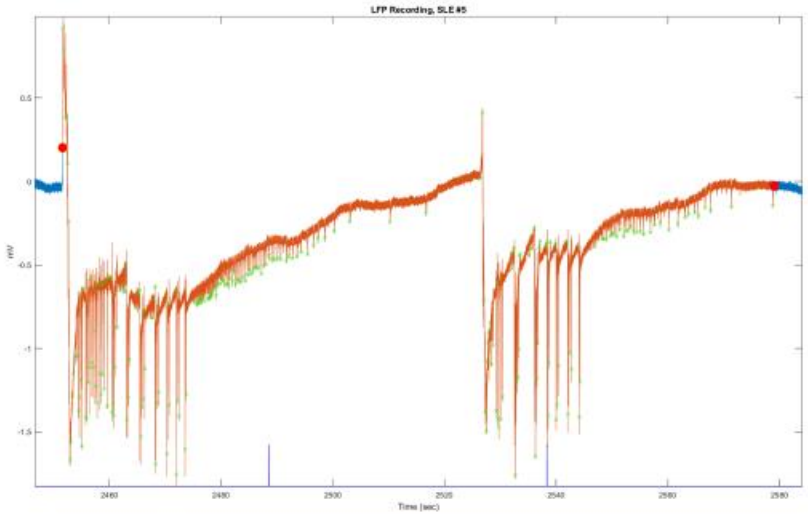


Figure 4. This is the result if the sigma threshold for detection is too low, it picks up all the noise (low-level spikes) and it results in two ictal events getting connected. As a result, these ictal events have a much lower intensity ratio.

Discussion

The detection algorithm V5.3 is on par with human raters, as it detects ictal events within the range of what is considered *perfect* (90–100% agreement) with human encephalographers (Gaspard et al., 2014). It is very difficult to study EEGs (traces) without an accompanying video for electrical events that are associated with a seizure (behaviour). From developing the detection algorithm and testing it, I've come to understand that some of the variability that exist between human markers is when the features of epileptiform events approach the "hard" threshold of what is an ictal event (Figure 5). Based on studying all the light-triggered ictal events, it was observed that ictal events have a frequency (spike rate) of 2.35 Hz on average, and the lowest frequency ever observed from 140 human labelled ictal events (from 7 recordings) was 1.03 Hz (Table 2). Thus, the detection algorithm's classifier system set a floor value for frequency as 1 Hz for ictal events. Thus, in nature, if there was an epileptiform event in the EEG recording with a frequency at 0.99 Hz, it would be classified as a interictal event, not sufficient in frequency feature to be classified as a ictal event (seizure). This example can be observed in Figure 5, the last event (far right) is an epileptiform event with a frequency below 1 Hz (0.66 hz), however, it is difficult for the human eye to visually discern the lower spike rate in the 5th epileptiform event (@6363s), unless you zoom in close, manually count the spikes and divide by duration of event) (Figure 6). Most human encephalographer will not analyze 1 hour long recordings in this manner. Thus, they rely on their intuition, and that is often where mistakes occur, leading to the interrater reliability when it comes to marking EEG recordings (reference). Furthermore, human intuition is difficult to built into a computer, a human can see that the 5th event looks like the preceding ictal events, and will mark it as such. A detection algorithm can not be programmed to be intuitive, it function with hard-coded rules to follow when classifying epileptiform events.

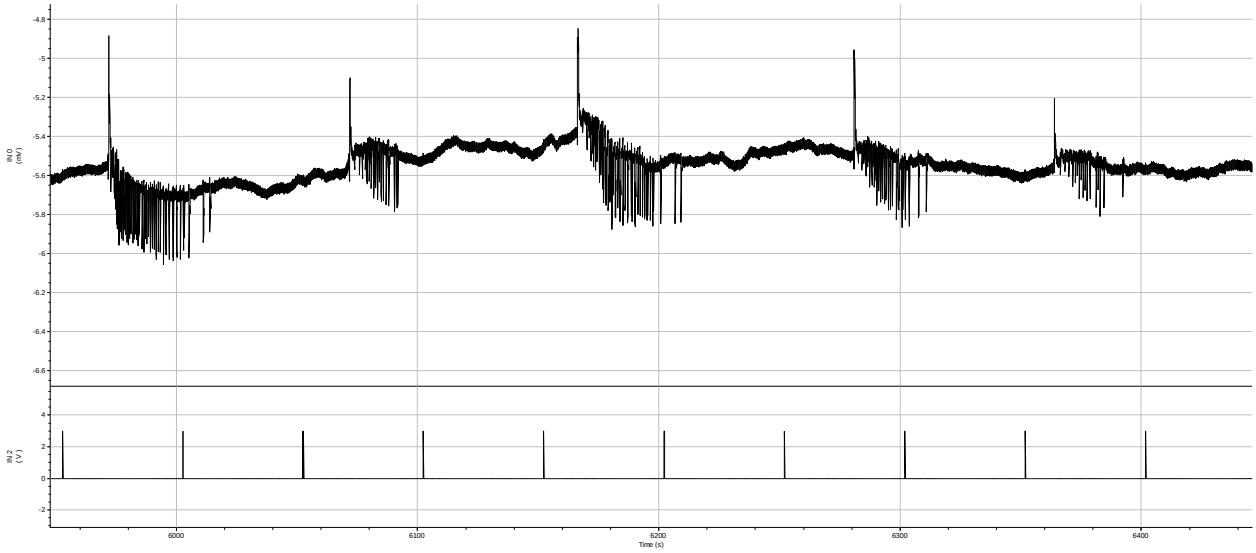


Figure 5. LFP recording of epileptiform events; the first 4 are ictal events and the last one is a IIE, 13802004.abf, 5500-6500s

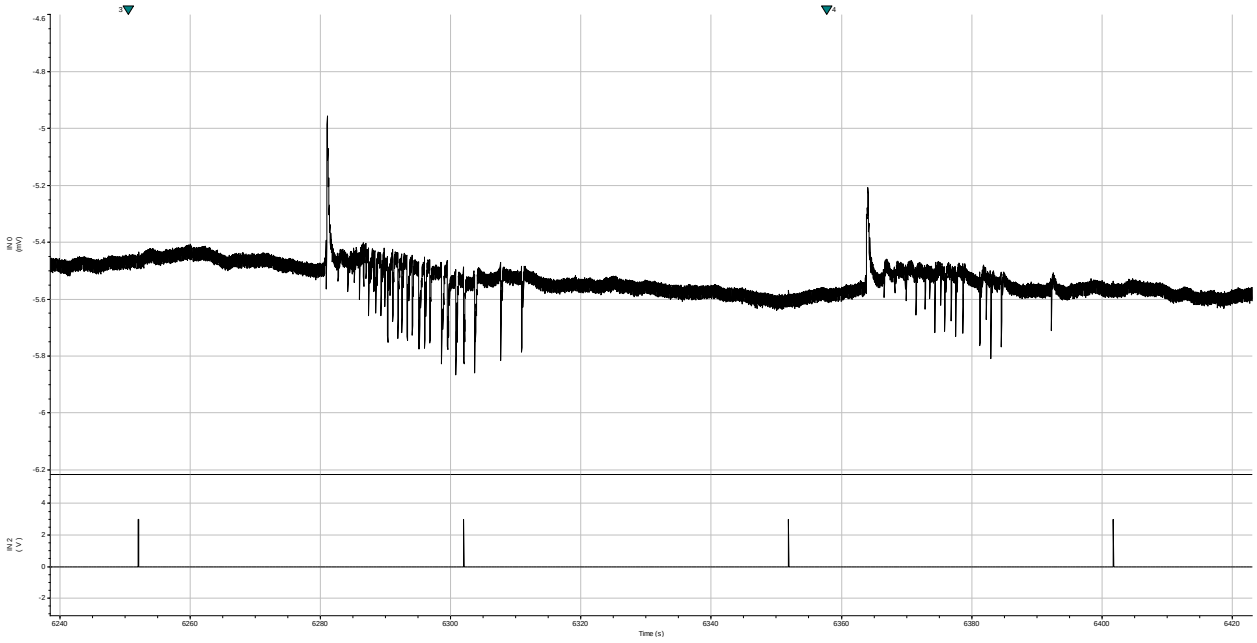


Figure 6. zoom-in on epileptiform events #4 and#5 from figure above (13802004.abf) @ 6200 - 6400 s

This is the strength of the detection algorithm in detecting interictal events. This detection algorithm was designed for the purposes of testing the size vs frequency distribution of interictal events (hypothesis is that they follow a power-law distribution).

Second, I learned that the algorithm works perfectly for the light-triggered ictal events, because these events are elicited on a regular interval, and as such, the events are incredibly uniform and easy to detect and classify. However, spontaneous ictal events are much more variable, because they can happen at any random moment, and the size of the ictal event will also vary accordingly because the trigger (for the ictal event) is not consistent. When I compared the population of ictal events from the training dataset (light-triggered ictal events) and ictal events from the testing dataset (spontaneous ictal events), they were from different populations and had a medium effect size (Cohen's d = 0.5). This was apparent, as the detection algorithm's classification system scored lower on the testing dataset (-2) than the training dataset (+3), therefore the population of ictal events must be difficult.

Third, the spontaneous ictal events in wild-type C57/Bl6 mice can result in two different population of epileptiform events (Figure 7). As a human marker, I labelled the smaller ictal events, as interictal events, however, they pass the thresholds in place by the classification system to be considered an ictal event by the detection algorithm. Post Analysis: I made this mistaken as a human marker because I didn't measure the absolute features of the ictal event, just relatively, how they compare to the rest of the population; since the orange labelled ictal events (Figure 7) were smaller, I labelled them as interictal (bursting) events. However, on the absolute scale, they qualify to be considered as ictal events. In fact, the smaller ictal events (orange labelled, Figure 7) are much more standard sized ictal events observed, and the larger ictal events (blue labelled, Figure 7) are actually super sized, unusually large ictal events.

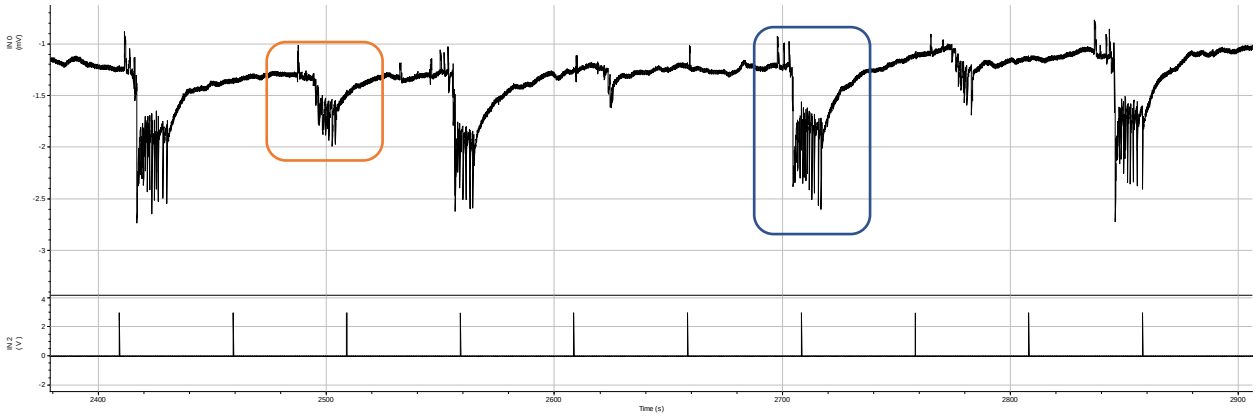


Figure 7. two disinct populations of ictal events from 13802005.abf (LFP recording of WTC57/Bl6, *in vitro* 4-AP cortical seizure model).

In an alternate scenario, there are two populations of epileptiform events in the next recording (of a different slice from the same mice). However, in this case, the two events appear to be much more similar in size, however, the smaller green labelled events (Figure 8) are below the threshold to be considered ictal events, both absolutely and relatively. Furthermore, the orange labelled ictal events are similar in magnitude and feature values to the orange labelled ictal events in Figure 7, which were initially accidently labelled as interictal events. In summary, human markers are very subjective. Furthermore, there are different sizes and types of ictal events, would be nice to classify them (what is it about scientists that make us just love to classify things).

You got to draw the line somewhere to distinguish interictal events from ictal events (transition: you need to have strict threshold that can segment ictal events from interictal events). That's why I built this classification algorithm based on k-means clustering. It works quite well because it's all relative, however, there are floor values to prevent interictal events from being accidently labelled as ictal events. The thresholds determined by the algorithm are dynamic and change between different recordings. The only constant that remains is the floor values of minimum values acceptable to be considered an ictal event (based on data from Chang et al., 2018a).

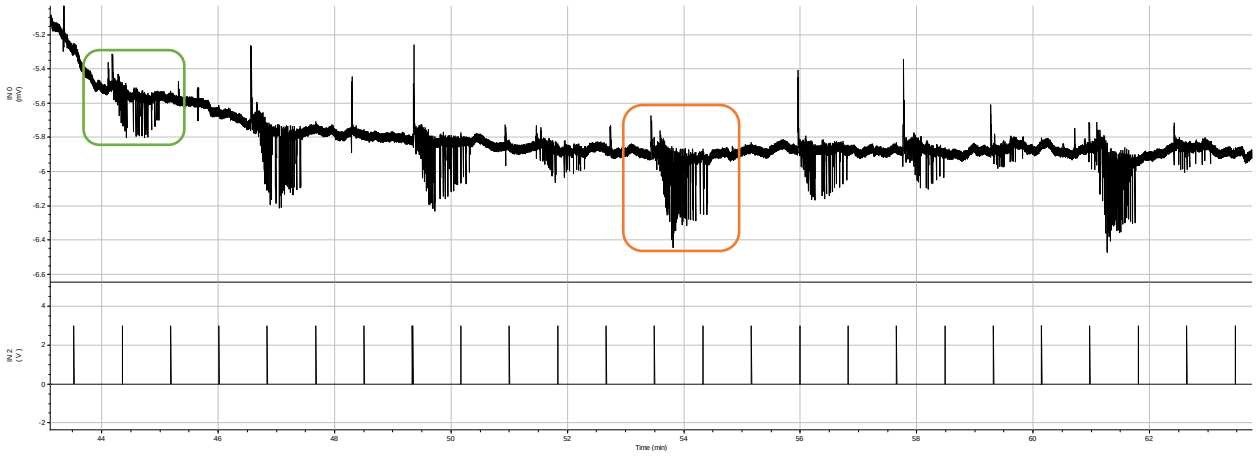


Figure 8. Two distinct population of epileptiform events from 13802004.abf (LFP recording of WTC57/Bl6, *in vitro* 4-AP cortical seizure model). Green labelled events are interictal bursts that are below the threshold to be considered a SLE, while orange labelled events are standard sized ictal events.

Methods:

Function Descriptions

Detection.m

The detection algorithm script is based largely on *findpeaks*, to detect spikes in the time series.

statistics.m

Calculates the population statistics for the time series provided. This function will provide the average, sigma, and quartile ranges for the time series.

findArtifact.m

This function is based on Matlab's *findpeaks* function to detect the peaks of spikes that occur in the time series. The function uses *statistics* to find the population statistics and quartile values. The end user can specify the threshold for spikes to be considered as an artifact. By default, spikes are labelled an artifact if it is 70 standard deviations above the average value within the time series. The putative artifacts must be smaller than 10.5 ms in width to prevent large epileptiform (or physiological) spikes from being mislabelled. A small window (±0.6 s) is made around the artifact spike, and within this window, all the spiking activity that is greater than five-times the third quartile value is detected to be a part of the artifact. The first spike and last spike are the start and end of the artifact event. These hard coded 1.2 s window is enough to detect the entire artifact because artifacts tend to be very short in duration (10 ms).

findEvents.m

This function is also based on *findpeaks* to locate the onset and offset of epileptiform events. It locates all the spikes in the time series, assuming them to be epileptiform spiking activity. The main logic is that distinct epileptiform events are separated by at least 10 s (based on rules #1 and #2). Thus, the function assumes spikes that are separated by more than 10 sec belong to distinct events. Based on this logic, the preceding spike in the gap represents the end of one event and the subsequent spike represents the start of another event. This function also relies on the *findArtifact.m* function to locate artifacts in the time series and distinguish them from epileptiform (or physiological) spiking activity. This prevents the algorithms from assigning an artifact to be the start or end of an epileptiform event.

If no specification is provided for epileptiform spike characteristics, the default threshold for epileptiform spikes is 20-times the 1st quartile value and that they are separated by 0.1 seconds. This allows the algorithm to perform faster by reducing the number of spikes detect. (Detection of onset and offset accuracy is not important here because only peaks of spikes are being detecfted, *SLECrawler.m* function is later used to detect the exact onset and offset point.

SLECrawler.m

This function analyzes the power feature of the absolute (value) dataset to determine the exact onset and offset points of SLEs. This function will go to the approximate onset and offset times of epileptiform events and scan 0.5 seconds before the event and 0.5 seconds after the event to search for the exact onset and offset point. The function applies Chang et al., 2018's arbitrary rules to locate the onset and offset points of SLEs, such as ensuring the last spike of a SLE is not due to a light pulse or an artifact. The function determines the last spike is due to an artifact if it occurs within 1.5 seconds of a light pulse; if any of the area between the putative SLE offset and the crawler's detected offset occur within 1.5 of a

light pulse, it's determined to be caused by light. This function can also determine if an event is light triggered (occurs within 100 ms of a light pulse).

The hard-coded values are that onset is the first point when power in the time series reaches 1/3 the value of the power of the first major epileptiform spike (detected by *findEvents.m*). The offset is the last point when the power of the timeseries returns of the half the baseline's mean value.

findThresholdSLE.m

The threshold calculation is very complex, uses a combination of methods that include k-means clustering and widest gap; it takes the widest gap in the feature set that is at or below the gap created by k-means clustering. The population with the smaller values is labelled as IIEs and the populations with higher values is labelled as SLEs. Note: k-means clustering was performed 25x to find the lowest threshold possible.

classifier_dynamic.m

This function has been updated to be truly dynamic and adaptable. It can use Michael's empirical thresholds used to separate a population of event based on some feature related to peak-to-peak amplitude, spike rate, intensity, and duration. There are floor values in place, which are implemented if the algorithm calculate a threshold below the floor value or if instructed in through the function's input.

This function separates a population of events, into two either SLEs or IIEs, based on one given feature. It calculates the threshold to separate the population using two different algorithms, called "Michael's" (based on populations statistics of ictal events from Chang et al., 2018) and "algo's" (which uses the *findThresholdSLE.m function)*. Michael's algorithm is basic population statistics to find the threshold (i.e. 2 standard deviations below average value, or the standard deviation, whichever value is higher). The algo's algorithm is based on k-means clustering and finding the widest gap in the feature set below the threshold found by k-means clustering (see *findThresholdSLE.m* above). The lower (more liberal) threshold, between Michael's and algo's calculation, is used as the threshold to classify the population of events as SLEs or IIEs. It is important to mention that there is a floor value (lowest minimum threshold) that is allowed for a event to be considered an SLE. This prevents a large collection of IIEs from being considered as a SLE. The classifier will label events as SLEs or IIEs based on three features: frequency, intensity, and duration. If the event cross the SLE threshold for all three features, it will be classified as a SLE. Lastly, the peak-to-peak amplitude feature is used to remove any events that have contain an artifact from being analyzed and contaminating the dataset.

Classifier_in_vitro.m

This function is like the *classifier_dynamic.m* function, but it uses only the floor values as the threshold. This classifier contains hard-coded threshold values specific for ictal event from the *in vitro* 4-AP cortical seizure models. It is used when there are less than 6 events that need to be classified, because the *classifier_dynamic.m* is ineffective in a small population of events.

Detection Algorithm

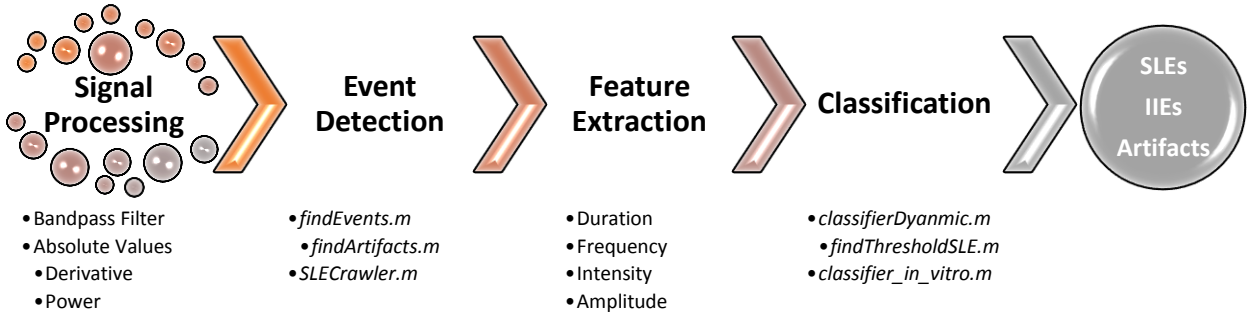


Figure 9. General overview of the key steps in the Detection Algorithm

0) Signal Processing

The raw signal (.abf file) is converted into a time series and processed in three sequential steps. First, the data is bandpass filtered, 1-100 hz (**filtered dataset**), Second, the filtered data is converted into absolute values (**absolute dataset**). Third, the derivative of this signal is calculated (**derivative dataset**).

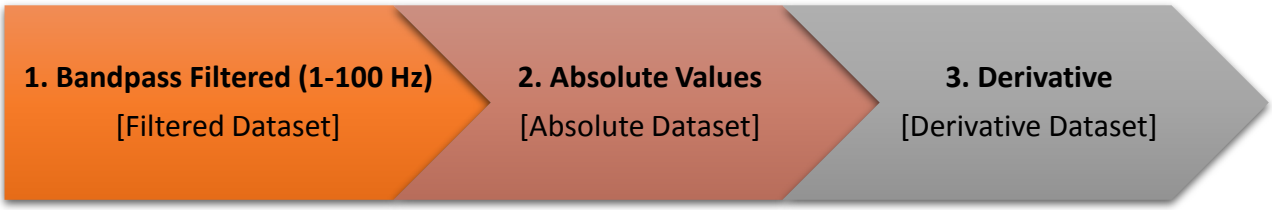


Figure 10. The sequential steps in which the raw LFP signal (time series) is processed for analysis

1) Event Detection

The epileptiform event detection algorithm works in three stages, using two functions *findEvents.m* (includes *findArtifacts.m*) and *SLECrawler.m*. First, it quickly approximates the location of events and artifacts in the derivative time series. Second, it removes these events from the absolute dataset to isolate the baseline signal and calculate new thresholds to detect epileptiform events in the absolute dataset. Third, the exact onset and offset times of SLEs are detected according to rules used by Chang et al., 2018. This detection algorithm is very quick and effective because of the nature of how events are detected as a collection of spiking activity in the time series. Furthermore, the third stage (*SLECrawler.m* function) can be easily modified for detecting events from other seizure models.

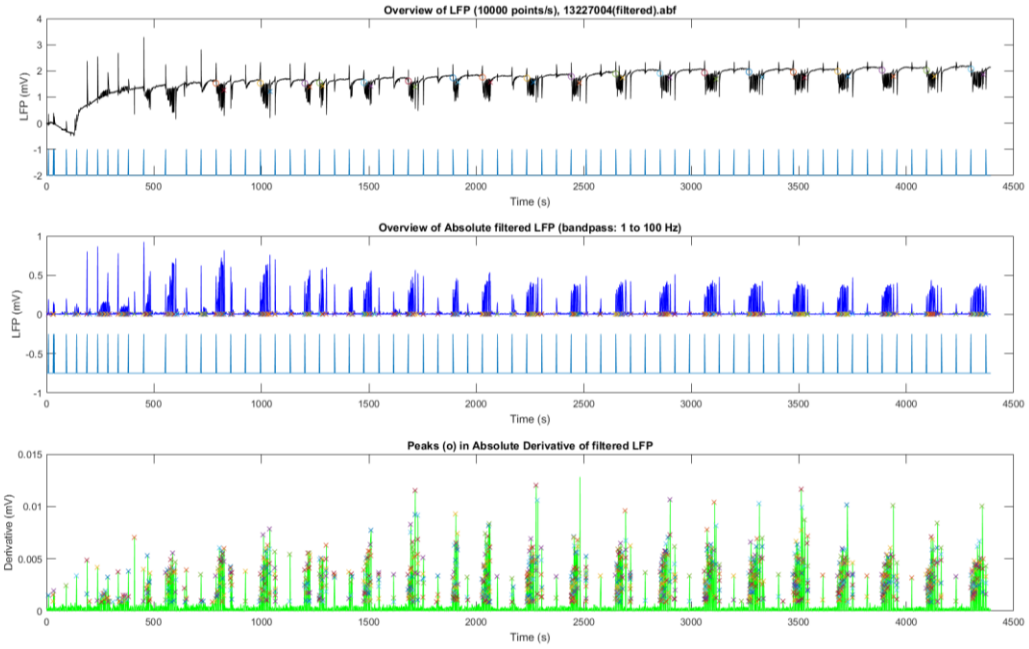


Figure 11. Unfiltered LFP dataset (top panel), absolute dataset (middle panel), and derivative dataset (bottom panel).

Stage 1: Derivative

The derivative dataset is first analyzed using the *findEvents.m* function to location all the major deviations in the signal, assumed to be epileptiform spikes. If these detected spikes are more than 10 seconds apart (Rules #1 and #2), then they are assumed to belong to separate epileptiform events. The preceding spike in the gap is assumed to be the end of the preceding event, and the subsequent spike is assumed to be the start of the subsequent event. The *findEvents.m* function detects epileptiform spikes with amplitudes beyond a specified threshold. By default, this algorithm dynamically sets the threshold to detect spikes as 20 times the first quartile value. The algorithm also detects artifacts, which are spikes that cross a threshold set as 70 standard deviations beyond the average time series value. In summary, the purpose of stage 1 is to quickly approximate the location of epileptiform events and artifacts that occur in the signal.

Stage 2: Absolute values of bandpass filtered LFP

Once these putative epileptiform events and artifacts are detected, they are removed from the absolute time series data. The remaining time series is considered an approximation of the baseline signal. The detection algorithm then calculates the averages and sigma of the approximate baseline and uses the variables to set the threshold to detect epileptiform events in the absolute dataset. The thresholds for epileptiform spikes and artifacts can be adjusted by user input (#1 and #2), however, the default recommended threshold for epileptiform spikes is (3.9*sigma) + average value, while the artifact is (70*sigma) + average value. Note: a noisier dataset may require a higher multiple of sigma. In summary, stage 2 is a repeat of stage 1, but with specified thresholds for epileptiform spikes as a multiple of baseline conditions. This strategy is very effective at detecting all the first and last spike of all epileptiform events in the *in vitro* 4-AP seizure model. If less than two epileptiform spikes are detected,

there is a feature build into the script to inform the end user to review the raw data (to make sure there are events) and consider using a lower multiple of sigma to detect the events.

Stage 3: Crawlers

To locate the exact onset and offset times of SLEs (ictal events) as Chang et al., 2018 would mark them, we use the *SLECrawler.m* function to analyze the power features of the absolute dataset. The *SLECrawler.m* function has built in arbitrary rules used by Chang et al., 2018, such as making sure an epileptiform event's last spike is not light-triggered event or an artifact. Furthermore, it can determine if an event was triggered by an applied stimulus or spontaneous. Most importantly, it locates the SLE onset to be the point where power in the signal reaches at least 1/3 prominence of the max amplitude (first spike detected) and the offset to be the last point where power in the signal returns to half the mean value of the baseline. This *SLEcrawler.m* function is specifically designed for epileptiform events from the *in vitro* 4-AP cortical seizure model. This function would need to be modified and tailored for onset and offset detection of SLEs from other seizure models or clinical recordings. In summary, stage 3 is to locate the exact onset and offset points for the epileptiform event according to the arbitrary rules used by Chang et al., 2018.

2) Feature Extraction

Built into the *detection.m* script.

- Three primary (absolute) features are extracted from the neural recording,
- One (absolute) feature is derived from the primary feature
- Two (normalized[2]) features are derived from the primary and secondary features.
    - These two features are specially engineered to capture the spatial and temporal aspect of epileptiform events from *in vitro* brain slices.

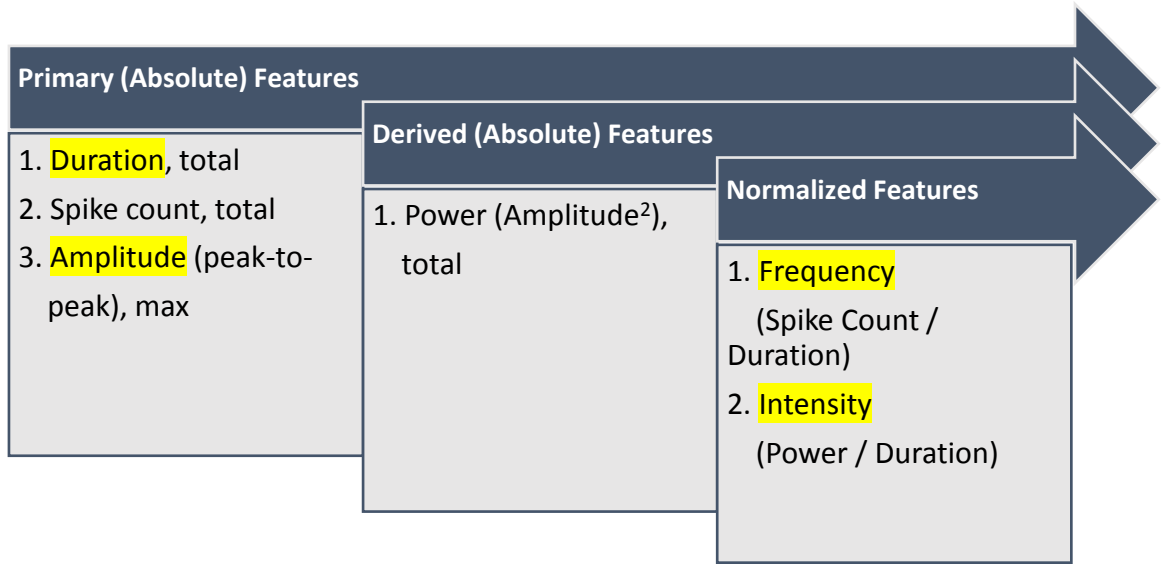


Figure 12. Features extracted from the filtered (LFP) dataset

[2] Normalization in this case is dividing by the duration of the event.

These four main features (highlighted) are used to classify epileptiform events, two absolute features and two normalized features (Figure 12). The duration, frequency*, and intensity* of events are used to classify whether events are ictal events (herein referred to as “SLEs”). The peak-to-peak amplitude is used to detect if epileptiform events contain artifacts (outliers), are subsequently removed from the dataset analyzed. *These features were specially engineered to be able to characterize the spatial and temporal aspects of epileptiform events.

3) Classification

There are two different variations of the classification system, *classifer_dynamic.m* (uses the *findThresholdSLE.m* function) and *classifier_in_vitro.m* (uses hard-coded thresholds based on statistical population data). The classification system classifies epileptiform events as SLEs using the two normalized features, frequency and intensity. SLEs have high values for these features that distinguish them from general bursting activity, noise, and IIEs. Notably, short interictal spikes (IISs) will also register high for these normalized features. Thus, the absolute duration is an important third feature to help filter out IISs and isolate only the SLEs. Lastly, peak-to-peak amplitude is the ideal feature to detect artifacts because artifact spikes tend to be very high in amplitude, typically 3-10 x higher than physiological spikes. If an event crosses the threshold to be considered an artifact (outlier), it is removed from the data for further analysis. The remaining events are labelled as a SLE if it surpasses the threshold for all the remaining 3 features, duration, frequency, and intensity. Otherwise, events are classified as interictal events (sub-seizure events) that occur during the interictal period. The two variations of the classification system differ in how they determine which threshold value to use.

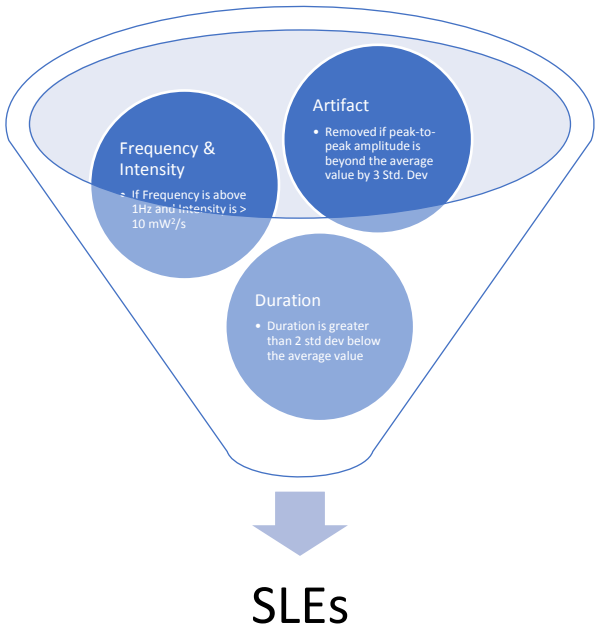


Figure 13. The classification system is essentially a filtering system for features contained by SLEs, to isolate such events.

i. Dynamic Classifier, SLEs

The first algorithm uses either Michael's empirical algorithm, or the *findThresholdSLE.m* function (written by Michael), whichever generates a lower (more liberal) threshold, to separate SLEs from IIEs,. However, it cannot go below the floor value (the minimum value acceptable for an SLE), a priori set by studying the statistics of the population data. The thresholds are dynamic and will differ between different datasets as it will adapt to each specific recording and segment the population events into SLEs and IIEs, like how a human would. The original motivation for this detection algorithm was to detect these interictal events and classify them as such in an objective manner. This is what inspired the development of the dynamic classification algorithm (*classifier_dynamic.m*). This algorithm classifies epileptiform events effectively classifies events like humans, consequently it is also subject to similar inconsistencies observed from human raters (reference). As a result, the performance of the dynamic classifier will depend on how the length of the recordings and the events that are present. It may classify the same epileptiform event differently, if only a smaller portion of the recording is analyzed by the algorithm. This is because the threshold that classifies a population of events into two populations (SLE and IIE) is dynamic and adapts to the recording that is under analysis.

The dynamic classification system works in three stages. First, it will detect if there are any outliers, or events that contain a potential artifact. It does this by using the peak-to-peak feature; artifacts tend to be very high in amplitude, much greater than biological activity. Artifacts (outliers) are events with a peak-to-peak amplitude that is greater than the population's average value by 3.1 standard deviations and removed from the subsequent analysis. The second stage uses the two normalized features, frequency and intensity, to segment the populations of epileptiform events into either IIEs or SLEs. If both features characterize the events as a SLE, it is labelled as a putative SLE. The third stage calculates the average duration and the standard deviation of the duration for putative SLEs. Then, it will use Michael's empirical algorithm to calculate the duration threshold as either the sigma, or 2*sigma below the average value, whichever value is higher (this is important because datasets with high sigma, will result in a negative value if you use Michael's second method). The *findThresholdSLE.m* function is also used to identify a potential threshold value. Between the Algo determined threshold and Michael's determined threshold, the lower (more liberal) value is used, with the lowest accepted threshold (the floor) set at 3 second. This third stage, is what I did as a person, if most ictal events were 50 s longer, then the ictal event that was around 10 s was excluded from my population of ictal events I analyzed.

ii. Hard-coded Classifier, SLEs

In time series with less than 6 events detected, the findThresholdSLE.m function does not work, and the dynamic classifier is ineffective. In these cases (if no SLEs are detected), the detection algorithm will repeat the classifier process using a hard-coded classification algorithm (specifically designed for events from *in vitro* 4-AP cortical seizure model) and classify events in a fixed manner as a computer would. The thresholds used to classify events are the floor values used in the dynamic classifier. The algorithm will classify events as a SLE if it satisfies the floor threshold for all three thresholds. Events are labelled as an IIE, otherwise. Similarly, if only 1 event is observed in the dataset (i.e., if end user wants to analyze the feature of a single SLE observed in the neural recording using the detection algorithm), the detection algorithm will switch over to use the hard-coded algorithm to classify the single event.

iii. Dynamic Classifier, epileptiform events

The epileptiform events that were detected are run through the dynamic classifier a second time, however at a potentially lower floor threshold for frequency (at Taufik's suggestion). Artifacts were not detected among the epileptiform events, because they should have already been filtered out from the first run through the dynamic classifier for SLEs.

iv. Hard-coded Classifier, epileptiform events

The floor frequency for questionable SLEs was the algo determined threshold for separating SLEs from IIEs, in case it was the case that the algo determined the threshold should be 0.95 Hz, but my floor at 1 Hz, resulted in some events with a frequency at 0.99 Hz being mislabelled as an IIEs. However, if the algo determined threshold went below 0.60 Hz, the floor threshold of 0.60 Hz was used. the absolute floor threshold for frequency was set at 0.60 Hz for the following reasons:

- There was a interictal events that Taufik said is likely a SLE, however it's frequency was below the floor threshold that Michael set at 1 Hz (I set it at 1Hz, because the lowest frequency observed from all the SLEs that I hand marked was 1.03 Hz; Taufik said I had implied bias in my classifier for SLEs that I recognized).
- The questionable SLE that Taufik found had a frequency around 0.66 Hz
- A ictal event should have at least 1 spike every 2 seconds minimum, or else its just bursting activity.

If the algo determined threshold was above 1 Hz, the standard floor threshold of 1 Hz was used.

Feature Extraction, per second values –a stage within the *detectionInVitro4AP.m* function

Spike Rate

The number of spikes that occur each second during the epileptiform event. I.e., Spike Rate of 1 means only 1 spike occurring in that 1 second period.

Tonic Phase

If there is consecutive high frequency firing for more (i.e., >1 s) that is by definition a tonic period. The threshold for high frequency firing is set as 1/3 the maximum frequency of the epileptiform event, with a floor set at 1 Hz. The tonic phase ends once the frequency drops to low frequency consecutively (i.e. >1s) or drops to 0 Hz at any time. Once you determine the location of the tonic phase, you can locate all the other phases of the ictal event.

Clonic Phase

The clonic phase begins when (and if) the tonic phase ends

Preictal Phase

The preictal phase is between the sentinel spike (start) and tonic phase

Intensity Ratio

The intensity ratio is the amount of time the epileptiform event is at high intensity divided by the total duration of the epileptiform event. Generally, SLEs have a higher intensity ratio than IIEs, because they are much larger in amplitude and spike frequency. Furthermore, the intensity ratio for IIEs is much

higher than a collection of IISs, since spikes are just concentrated bursts of energy in the signal fixed to a small period of time, where as IIEs have much more spike and wave activity that is higher in energy through the event (no silence/non-activity in the middle)

Intensity Ratio, SLEs

The threshold for high intensity when calculating the intensity ratio for SLEs is the epileptiform event's maximum intensity/10.

Intensity Ratio, IIEs

The threshold for high intensity when calculating the intensity ratio for IIEs is the average intensity (typically 0.002 mV^2/s) of the entire bandpass filtered [1 Hz, 100 Hz] LFP filtered time series. | used in Feature extraction stage

Note

The reason for calculating the intensity ratios differently for SLE classification and IIE classification is because there is a different level of sensitivity that needs to account for when you scale down from a SLE to an IIS. i.e. the power components of an IIS are much lower, so greater sensitivity is required to distinguish between when intensity is high vs low, and the average intensity of the entire recording is generally very low (0.02 mV^2/s). This makes great sense as well because if you're above the average intensity, it's fair to consider it as high, if you're below the average intensity, it's fair to consider it as low. An IIS can be considered to be a part of the baseline recording.

However, during a SLE, the intensity is generally very high, thus a new threshold is required to account for this. So, we use a threshold that is dependent on each event that is analyzed (1/10 of the max intensity observed is the threshold).

Rounding the intensity ratios

The reported intensity ratio is rounded to the first decimal place. However, all calculations to determine the threshold values are based on the true values (with no rounding). Then the classification is based on comparing the rounded intensity ratios (to first decimal place) to the unrounded threshold values. This helps to deal with any issues related to values close to the threshold value being missed.

Interictal Event Classifier – not a function, a stage within the *detectionInVitro4AP.m* function

The interictal event classifier, uses the additional information about the second by second features of individual events to further classify them into:

0: Unclassified (Review Event)

1: SLE

1.5: Questionable SLE

2: IIE

3: IIS

4: Artifact

Tonic Phase

- If an event is lacking a tonic phase, it cannot be considered a SLEs.
- If an IIE has a tonic phase, it is likely not a collection of IISs (which tend to be dispersed)

Intensity Ratio

- The intensity ratio threshold required to be considered a SLE is dynamic
    - The lowest intensity ratio observed among all the confirmed SLEs/IIEs
    - The algo-determined threshold (k-means clustering the feature and finding the widest gap below the cluster)
    - Floor set at 0.3 for SLEs, 0.2 for IIEs

In summary, the dynamic classifier calculates a new threshold for each feature based on the dataset being analyzed. The hard-coded classifier will use the same threshold for every file. An important feature in the dynamic classifier is the minimum value (floor) of the features accepted as a threshold to classify SLEs. For Frequency, it is 1 Hz (justify with literature) and intensity is 10 mV^2/s (based on a study on the population of ictal events from Chang et al. 2018a). Setting a floor value is important to make sure the larger IIEs are not classified as SLEs in a population composed of only IIEs. Furthermore, the duration feature is important to make sure short IIEs (i.e. IISs) are not mistaken as SLEs because they have the same average frequency and intensity values. Lastly, peak-to-peak amplitude is an important feature to guard against events contaminated with artifacts.

Hard-coded Values

detectEvents.m

```
    minPeakHeight = Q(1)*20;    %spike amplitude threshold = 20x(3rd quartile)
    minPeakDistance = 0.1 * frequency;     %spikes seperated by 0.1 seconds
    minArtifactHeight = mean(LFP) + (70*std(LFP));
    minArtifactDistance = 0.6 * frequency;
```

findArtifacts.m

```
    average = mean(LFP);                %Average
    sigma = std(LFP);               %Standard Deviation
    minPeakHeight = average+(sigma*70);    %artifact amplitude >120x 3rd
quartile
    minPeakDistance = 6000;     %artifact spikes seperated by .6 seconds
%% Finding real Artifacts, width <10 ms
locs_artifact = locs_artifact_potential(width<105);
```

Future Works

Questions for Liam/Fred

i. Logical indexing in SLECrawler.m without using the find function.

ii. How can I set up matlab, so enduser can specify what kind of file it wants to analyze, then have preset, hardcoded values (parameters) set in place (i.e. multiple of sigma) for each file type. Another example is light-triggered inputs, or none

- Classifier for other seizure models
- Remove the Michael's threshold algorithm and just use the floor value in its place
    - Michael's threshold algorithm is useful, in some cases of intensity. Keep Michael's algorithm in place
    - However, a floor value is important to protect against a population of IIEs being mislabelled as SLEs
    - I did drop the floor value for threshold of intensity to 5 $mV^2/2$
        - Because of file 13805000.abf; it had one event that looked (to me) very much like a SLE, but it's intensity was extremely low (had lower amplitude spikes) that were around 5.6 mV^2/s
- Add a rule to classify all detected spikes with a width < 10 ms as an artifact
    - Follow-up: This only works for large amplitude spikes.
    - Some biological spikes and lower amplitude spikes, and preictal (sentinel) spikes also have a very narrow width <10 ms, and as low as 2.6 ms.
- Play with sigma to see which is best for seizure detection (10x works well, but is it much better than 3.9x, as recommended?)
- Use a While loop to test many different multiples of sigma to see which results are best. – this would be better than a human; because it's like analyzing it many times and picking the best outcome (like the k-means clustering algorithm with widest gap, you wrote)
- Future iterations of algorithm will perform hyperparameter grid searching (trial and error of a wide range of hyperparameters).

4-AP induced epileptiform-like events

Seizure-like Events (SLEs)

SLEs, properly known as ictal-like events, are electrographic signals in rodent brain tissue treated with 4-AP [100 uM] that are that are reminiscent of the electrographic signals observed in the EEG-recordings of patients while they are experience a seizure episode. A seizure is comprised of the following characteristics based on the detection algorithm and supported by literature [provide references):

- Duration is at least 10 s long
    - They are longer than interictal events, on average
    - at least 3 seconds long in some physiology papers from the 90's,
    - Generally accepted value today is >10 s
- Intensity is at least 5 mV^2/s
    - They are greater than interictal events, on average
    - The lowest observed value was 16 mV^2/s (Table 2)
- Spike rate per second (frequency) is at least 1 Hz
    - They spike rate is greater than interictal events, on average
- Tonic phase is present and at least >3s
    - Tonic phase is high frequency spiking activity
    - The algorithm defines the spiking rate as high if it's at least 1/3 of the max spiking rate/s during the epileptiform event (with a minimum floor set at 1 Hz)
- Intensity ratio is high

    - The intensity ratio is the total amount of time in high intensity divided by total time of event
    - The algorithm defined the threshold between onset and offset as the epileptiform event's max intensity divided by 10
- Ictal event Onset is defined as the start of the preictal spike (a.k.a. sentinel spike) (Figure 14)
    - The algorithm defined the peak power of the preictal spike
    - Then it locates the point when power is below the threshold before the spike
    - The threshold is set as 5% of the spike's peak power
- Ictal event offset is defined as the last spike of the epileptiform event
    - The algorithm defines the offset as the point where power of the signal decreases to half the average power of the last spike.
- Light-trigger; ictal events were considered to be light triggered if the sentinel spike occurred within 130 ms after the light pulse was turned on.
- Examples of SLEs detected by the algorithm
    - 13226009(interictalEvents).pptx

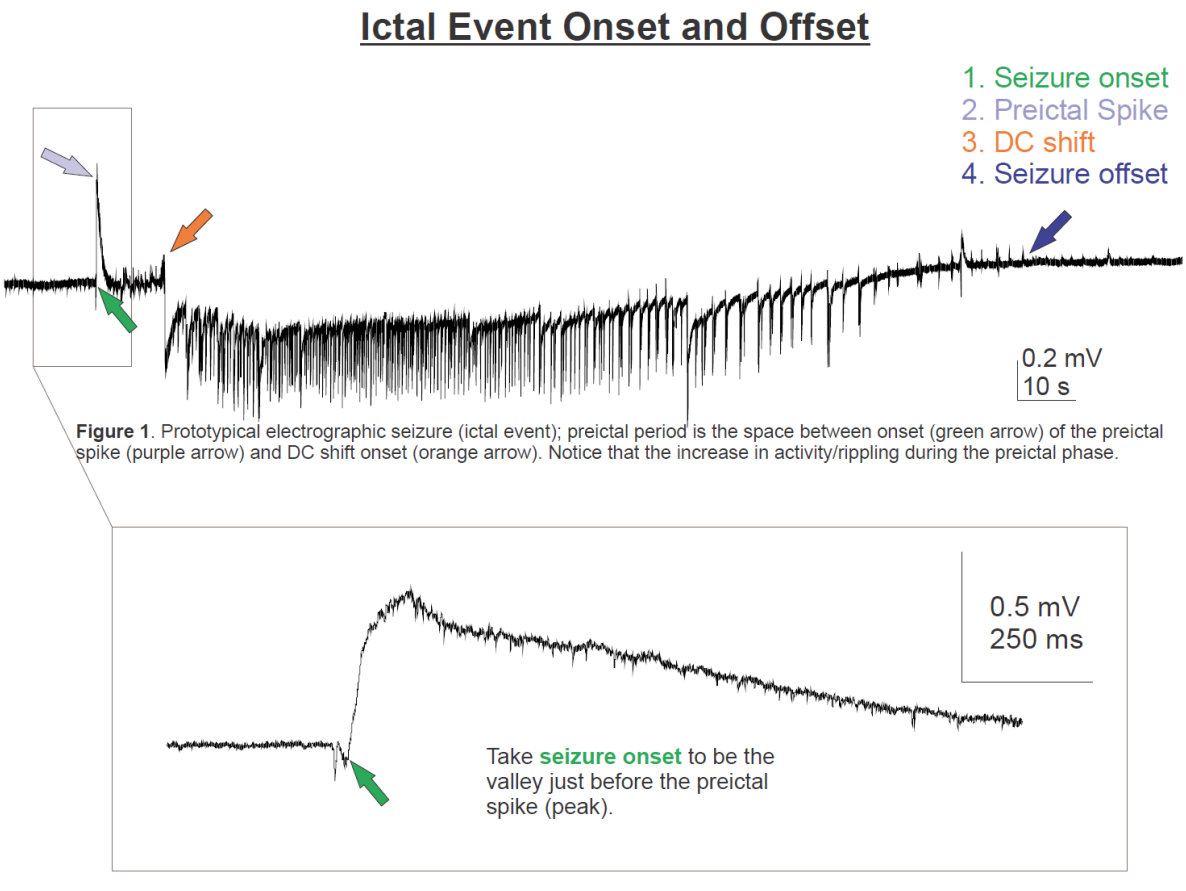


Figure 1. Prototypical electrographic seizure (ictal event); preictal period is the space between onset (green arrow) of the preictal spike (purple arrow) and DC shift onset (orange arrow). Notice that the increase in activity/rippling during the preictal phase.

Figure 2. Zoom-In on the preictal spike to show the exact point of the seizure onset.

Figure 14. Figure used to teach students how to define a ictal event, when training them to mark ictal events from .abf files

Interictal-like events (IIEs)

Interictal spikes are all the epileptiform events that occur between two seizure events. In other words, they are all the epileptiform events that are not a seizure. According to that definition, there are a great variety of IIEs, and they all look very different. There are subcategories of IIEs:

i. Interictal spikes (IISs)
   - A single deflection in the LFP recording

ii. Spike and wave events
   - A single deflection in the LFP recording followed by small oscillations

iii. Interictal events (IIEs)
   - A general interictal event will appear to be a 'runt' SLEs. Almost like a failed SLE because it's lacking all the ionic factors.
   - They are a SLEs lacking on of the necessary features or thresholds to be considered a SLE
   - The algorithm will define as IIEs as
      i. Duration below threshold to be SLE
      ii. Intensity below threshold to be SLE
      iii. Frequency below threshold to be SLE
      iv. Lacking tonic phase
      v. Lacking the intensity ratio to be a SLE
      vi. 
      13226009(interictalEvents).pptx

iv. Bursting activity
   - Oscillation activity spread a part

Supplementary Information

Overview of the Epileptiform Detection Algorithm

The Epileptiform Detection Algorithm detects epileptiform events from LFP recordings (the first channel by default) and classifies them as an ictal event, questionable ictal event (that requires human verification), interictal events, and major artifacts. The Epileptiform Detection Algorithm operates on the fundamental basis that epileptiform events are comprised of spikes, either individual spikes (i.e., IIS) or a collection of spikes (i.e. IIEs, bursts, or SLEs).

The detection algorithm detects the peak of spikes in the LFP recording, where spikes are defined as deflection above the threshold (known as the hyperparameter). The hyperparameter is optimally defined as baseline's average amplitude + 3.9 x sigma. In future iterations, the algorithm will have the option to search for the ideal threshold using hyperparameter grid searching.

Using domain knowledge of the ictal events from the 4-AP, we performed feature engineering and found that four features are optimal at defining epileptiform events: frequency of population spikes, intensity (amplitude^2/duration), duration of event and peak-to-peak amplitude of event.

The classification algorithm then classifies the detected epileptiform events based on a modified version of k-means clustering, where k = 2 (ictal events and interictal events) for each of the four features, separately. If a detected event is grouped into the cluster with larger values for all of the follow three features (frequency, intensity, and duration) it is classified as an ictal event, otherwise it is a interictal event. If a detected event is grouped into the larger group for peak-to-peak amplitude it is classified as an artifact.

The modified version of k-means clustering, where k = 2 simply takes the widest gap below the detected threshold by k-means clustering. Widest gap is when the distance between events based on some feature [provide figure as demonstration of concept].

The Epileptiform Detection Algorithm is fast, light-weight, and requires no training. It can analyze 1 hour of data in 1 minute; it achieves this by limiting intense (long) calculations to limited periods of the time series (such as when locating exact point of seizure onset). The detection algorithm has a modular design, so only the *crawler.m* function (which locates exact onset and offset) needs to be modified to detect specific events from other seizure models or clinical recordings.

There are three main additions to V7.0 (from 5.0) is i) the ability to classify interictal events and questionable SLEs, ii) modified the crawler function so can be customized to detect different events, such as spikes, and iii) print out the metadata of the .abf files on the excel files to keep track of abf recording data. The detection algorithm (V 5.0+) can accurately detect and classify electrographic events from the *in vitro* 4-AP seizure model such as SLEs, IIEs, or artifacts. Detection algorithm is 100% and 93% consist with Chang et al., 2018's markings of light-triggered ictal events (n = 140) and spontaneous ictal events (n = 114), respectively.

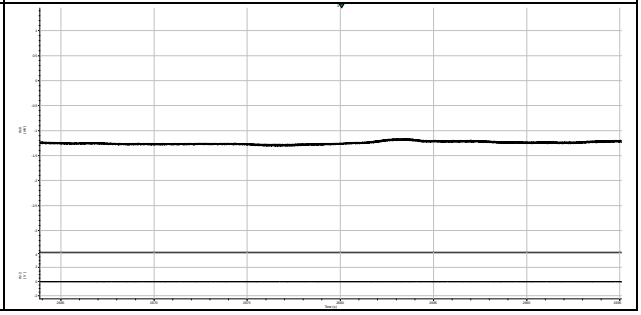
Hardcoded Values

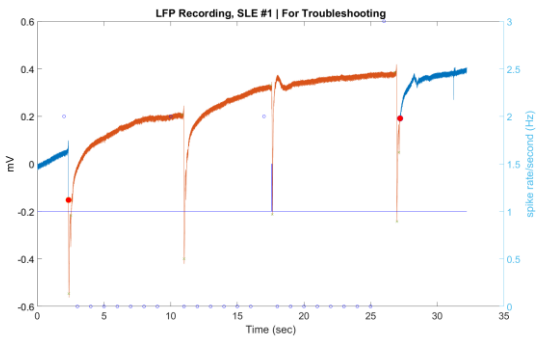
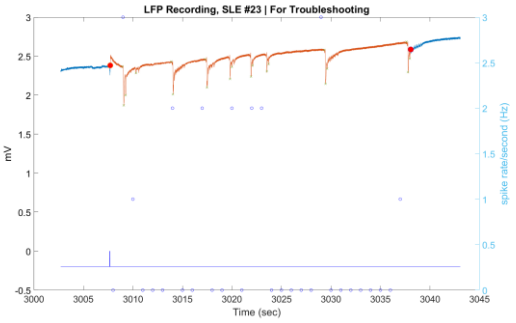
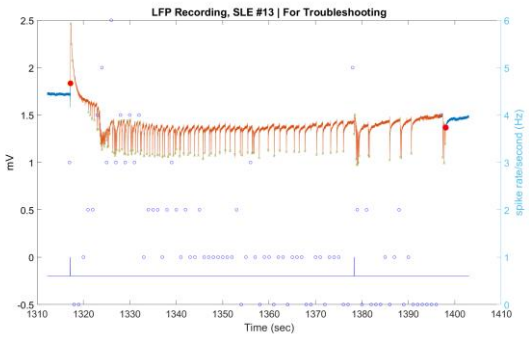
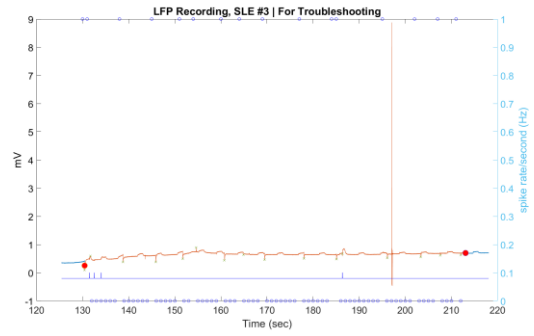
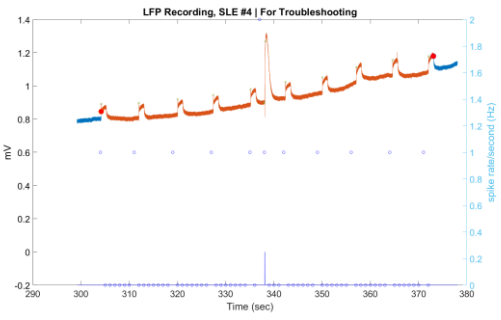
Detection Algorithm, Stage 1: Derivative

```
minPeakHeight = Q(1)*20;   %spike amplitude >20x 3rd quartile
minPeakDistance = 0.1 * frequency;    %spikes seperated by 0.1 seconds
minArtifactHeight = mean(LFP) + (70*std(LFP));
minArtifactDistance = 0.6 * frequency;
```

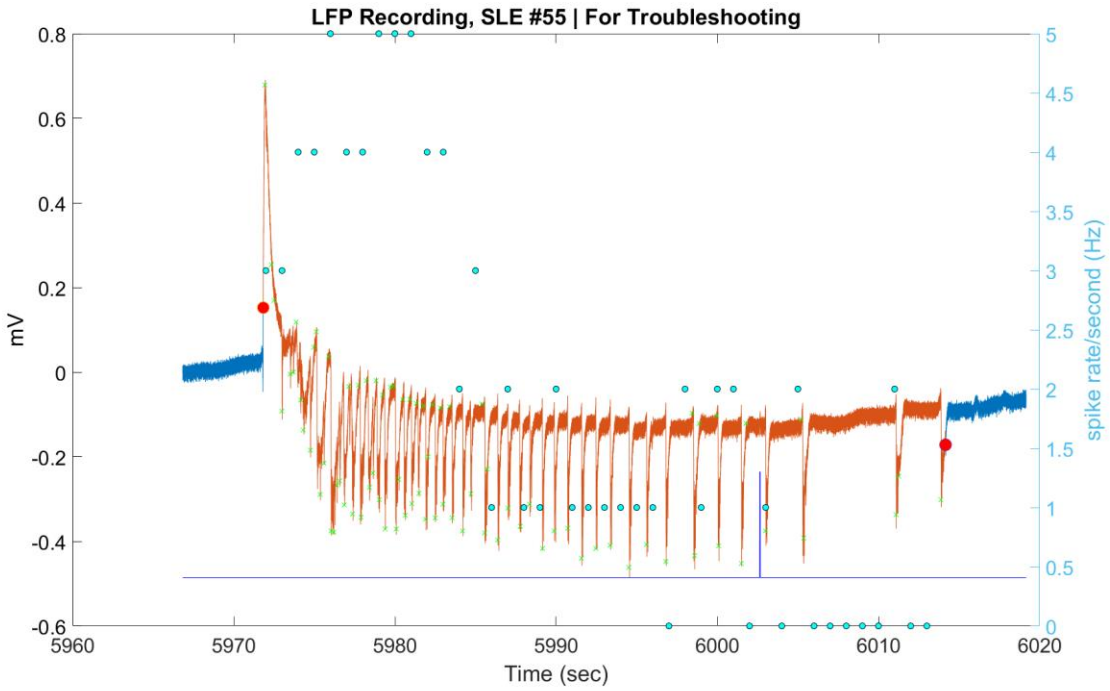
Six possible events observed in in vitro *4-AP seizure models*

Table 7. Commonly observed events in *in vitro* LFP recordings, 13226009.abf

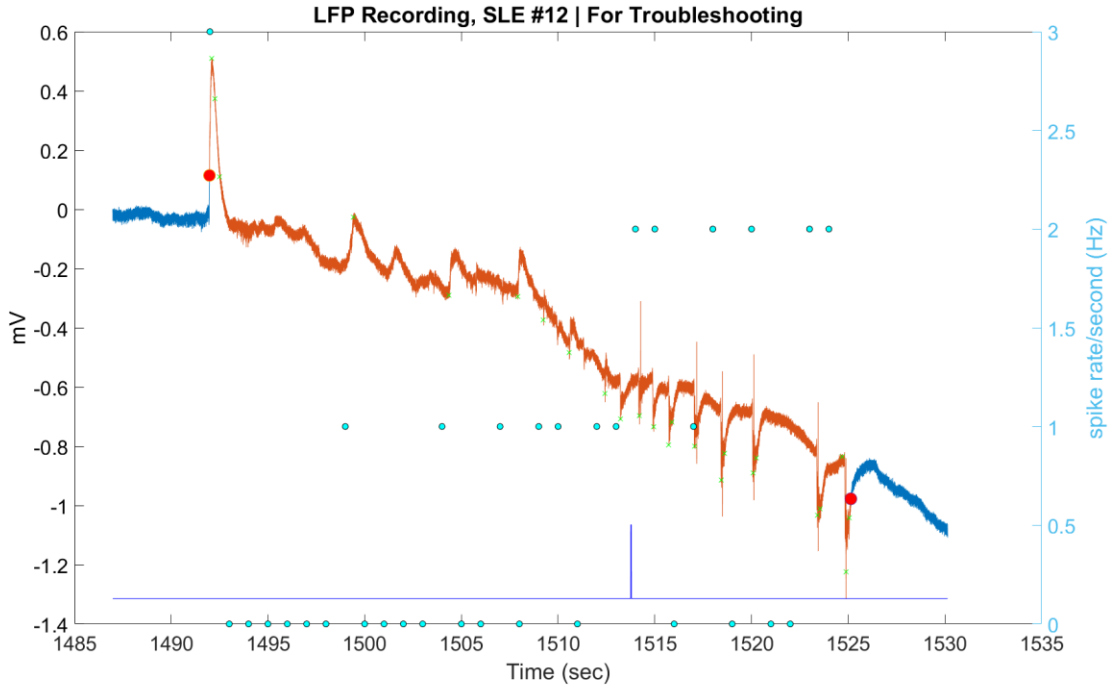
Epileptiform Event	Description	Image
Baseline	Flat line	

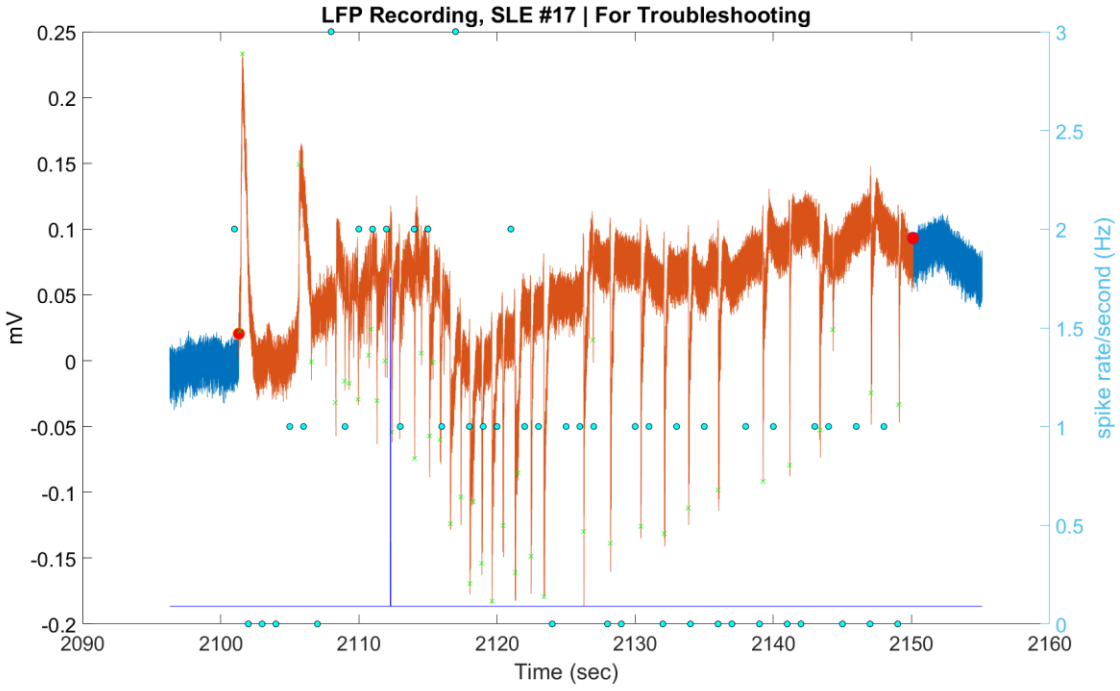
Interictal Spike(s)	Sharp spike, narrow width ~ 100-250 ms, with high amplitude that has a physiological maximum amplitude (~2 mV)	
Interictal Event	A collection of spikes, sparse and spread out (low frequency)	
Ictal Event	A high frequency oscillation with spikes of high amplitude (intense)	
Artifact	A high amplitude spike that is narrow in width >10 ms, and can have unlimited amplitude (saturate the recording limits)	
Noise	Repetitive oscillations that are shaped obtuse.	

13802004.abf SLEs detected by both algorithm and humans



13802004.abf SLEs detected by humans, but missed by algorithm





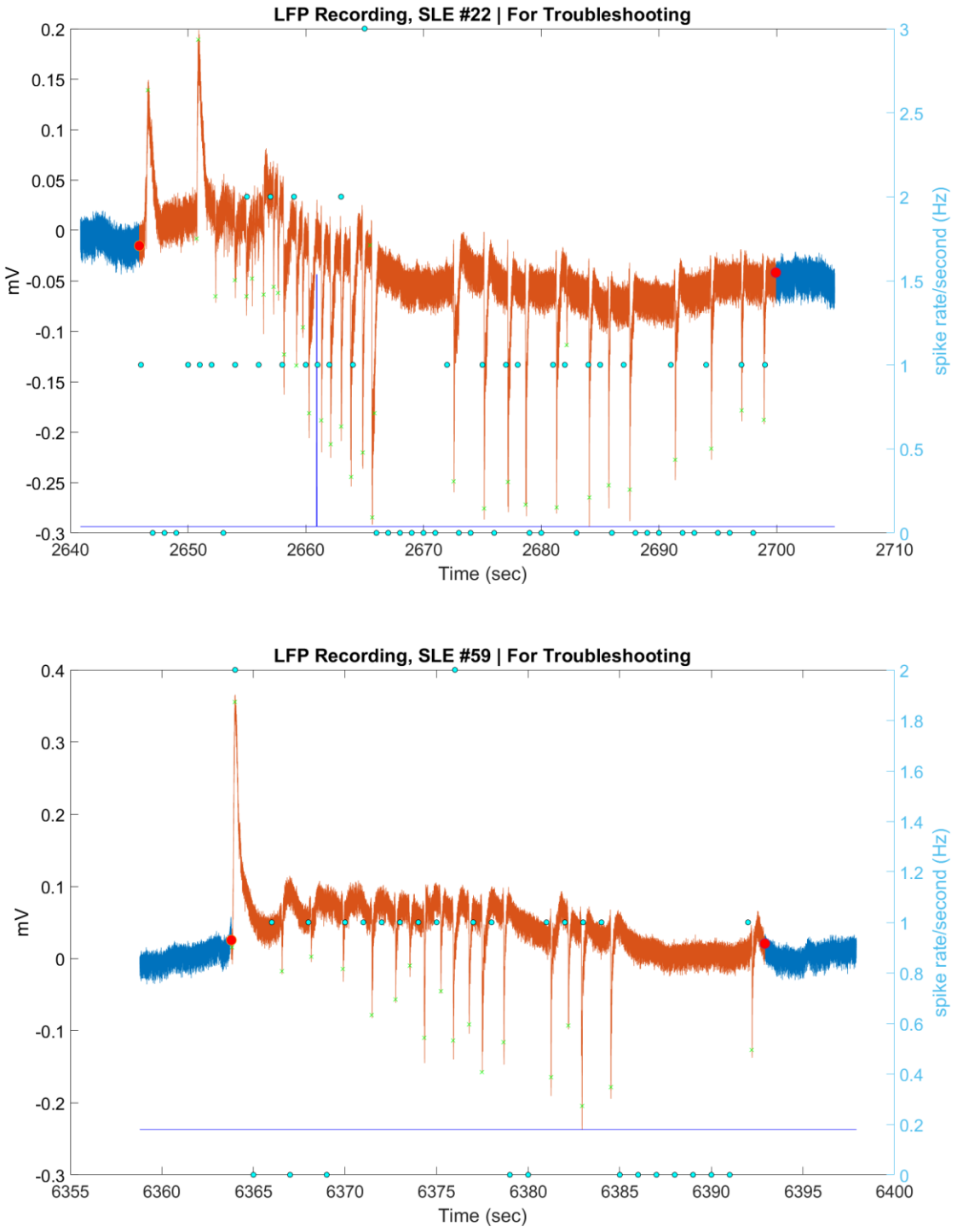
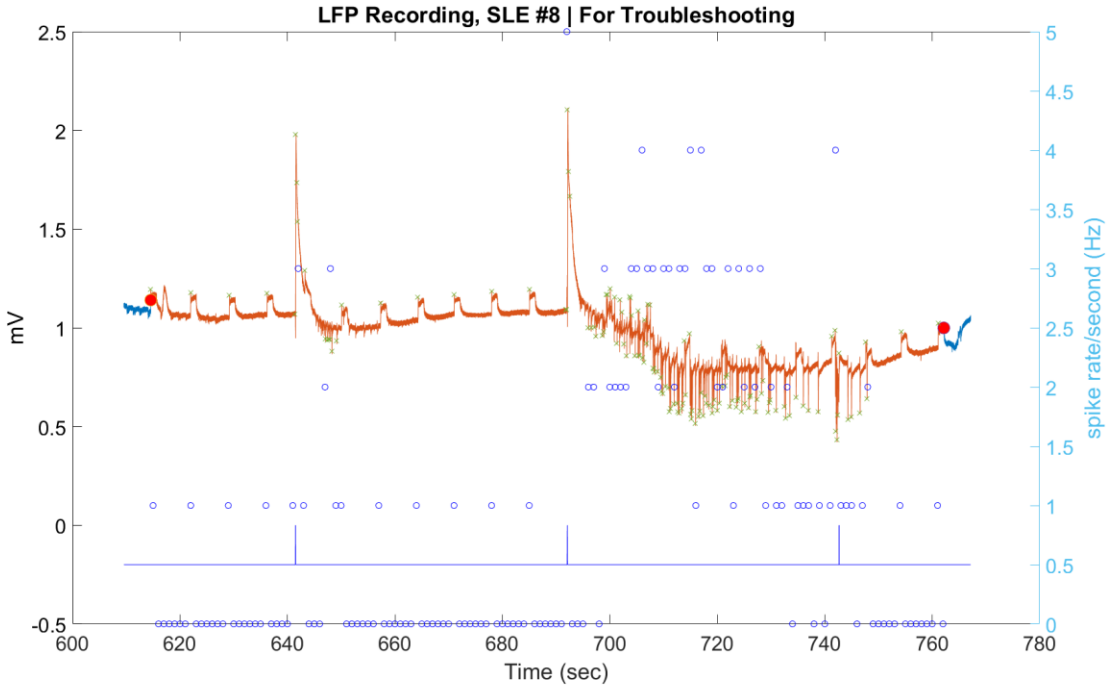


Figure 15. SLE detected by human markers, but not considered to be an ictal event by algorithm.

File 13226009.abf, Detection algorithm can miss ictal events embedded in noise



Excel File output

Column

A. The onset time (seconds) for the event
B. The offset time (seconds) for the event
C. Duration (seconds) for the event
D. Spike Rate (Hz); the average number of spikes detected per second, during the event
E. Intensity (mV^2/s); the average power per second, during the event
F. Peak-to-peak amplitude (mV); the maximum amplitude – minimum amplitude, observed during the event
G. Classification; arbitrary number system (enumeration) designed by Michael Chang
    - 0: unclassified event, these events typically have artifacts or are contaminate by low-level noise. Require human intuition to remove from analysis or keep.
    - 1: Seizure-like Event (SLE)
    - 1.5: Questionable SLE (smaller seizure, may require human intuition to determine if SLE
    - 2: Interictal Event (IIE)
    - 2.5: Questionable IIE, not yet sure if it's a IIE or a collection of spikes (IISs)
    - 3: Interictal spike (IIS), or a collection of IIS
        - Regardless, events that are classified as such (3) will be returned to the recording to be detected again using a wavelet transform
    - 4: Artifact: these are events that have been contaminated by noise
H. Light-triggered; indicate if event was triggered by a external stimulus (1), or if it was spontaneously initiated (0)
I. Results of the classification system for Spike Rate;

- 1, if event was above threshold
- 0, if event was below threshold.
- Number in the top column indicates the threshold, determined by the algorithm.

J. Results of the classification system for Intensity;
- 1, if event was above threshold
- 0, if event was below threshold.
- Number in the top column indicates the threshold, determined by the algorithm.

K. Results of the classification system for Duration;
- 1, if event was above threshold
- 0, if event was below threshold.
- Number in the top column indicates the threshold, determined by the algorithm.

L. Results of the detection algorithm for Artifacts (based on peak-to-peak amplitude);
- 1, if event was above threshold
- 0, if event was below threshold.
- Number in the top column indicates the threshold, determined by the algorithm.

M. Tonic Phase; the tonic phase was defined as 2 continuous seconds of frequency that is at least 1/3 maximum frequency during the event (floor set at 1 Hz).
- 1, if tonic phase is present
- 0, if tonic phase is absent

N. Average Preictal Frequency; number is only valid if tonic phase is present (Column M = 1)

O. Average Tonic Frequency; number is only valid if tonic phase is present (Column M = 1)

P. Average Clonic Frequency; number is only valid if tonic phase is present (Column M = 1)

Q. Minimum Frequency observed during the tonic phase, if tonic phase is present (Column M = 1)

R. Intensity Ratio, time period of high intensity divided by time period of low intensity, during the epiletiform event

S. Minimum intensity ratio minimum to be considered a SLE
- 1, if event was above threshold
- 0, if event was below threshold.
- Number in the top column indicates the threshold, determined by the algorithm.

T. Minimum intensity ratio minimum to be considered a IIE
- 1, if event was above threshold
- 0, if event was below threshold.
- Number in the top column indicates the threshold, determined by the algorithm.

U. Start Time (Tonic Phase); number is only valid if tonic phase is present (Column M = 1)

V. End Time (Tonic Phase); number is only valid if tonic phase is present (Column M = 1)

W. Label, determined by a human (Michael)
- 0: unclassified event, these events typically have artifacts or are contaminate by low-level noise. Require human intuition to remove from analysis or keep.
- 1: Seizure-like Event (SLE)
- 1.5: Questionable SLE (smaller seizure, may require human intuition to determine if SLE
- 2: Interictal Event (IIE)
- 2.5: Questionable IIE, not yet sure if it's a IIE or a collection of spikes (IISs)
- 3: Interictal spike (IIS), or a collection of IIS

        i. Regardless, events that are classified as such (3) will be returned to the recording to be detected again using a wavelet transform
    - 4: Artifact: these are events that have been contaminated by noise
X. Michael's comments about the event based on the .pptx file

References

Chang M, Dian JA, Dufour S, Wang L, Chameh HM, Ramani M, Zhang L, Carlen PL, Womelsdorf T, Valiante TA (2018a) Brief activation of GABAergic interneurons initiates the transition to ictal events through post-inhibitory rebound excitation. Neurobiol Dis 109:102-116.

Gaspard N, Hirsch LJ, LaRoche SM, Hahn CD, Westover MB (2014) Interrater agreement for Critical Care EEG Terminology. Epilepsia 55:1366-1373.