Time-frequency strategies for increasing high frequency oscillation detectability in intracerebral EEG
Nicolas Roehri, Jean-Marc Lina, John C. Mosher, Fabrice Bartolomei, Christian-George Bénar

To cite this version:
Nicolas Roehri, Jean-Marc Lina, John C. Mosher, Fabrice Bartolomei, Christian-George Bénar. Time-frequency strategies for increasing high frequency oscillation detectability in intracerebral EEG. IEEE Transactions on Biomedical Engineering, Institute of Electrical and Electronics Engineers, 2016, <10.1109/TBME.2016.2556425>. <hal-01316459>

HAL Id: hal-01316459
https://hal.archives-ouvertes.fr/hal-01316459
Submitted on 17 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Time-frequency strategies for increasing high frequency oscillation detectability in intracerebral EEG

Nicolas Roehri Student Member, IEEE, Jean-Marc Lina, John C. Mosher, Senior Member, IEEE
Fabrice Bartolomei, and Christian-George Bénar*, Member, IEEE

Abstract—Background: High Frequency Oscillations (HFOs) are considered to be highly representative of brain tissues capable of producing epileptic seizures. The visual review of HFOs on intracerebral electroencephalography is time-consuming and tedious, and it can be improved by time-frequency (TF) analysis. The main issue is that the signal is dominated by lower frequencies that mask the HFOs. Our aim was to flatten (i.e., whiten) the frequency spectrum to enhance the fast oscillations while preserving an optimal Signal to Noise Ratio (SNR). Method: We investigated 8 methods of data whitening based on either prewhitening or TF normalization in order to improve the detectability of HFOs. We detected all local maxima of the TF image above a range of thresholds in the HFO band. Results: We obtained the Precision and Recall curves at different SNR and for different HFO types and illustrate the added value of whitening both in the time-frequency plane and in time domain. Conclusion: The normalization strategies based on a baseline and on our proposed method (the “H0 z-score”) are more precise than the others. Significance: The H0 z-score provides an optimal framework for representing and detecting HFOs, independent of a baseline and a priori frequency bands.

Index Terms—high-frequency oscillation, epilepsy, wavelet transform, whitening, stereoelectroencephalography.

I. INTRODUCTION

STEREOELECTROENCEPHALOGRAPHIC (SEEG) recordings using clinical intracranial macroelectrodes are considered as a standard for identifying the epileptogenic zone (EZ), the part of the brain which has to be surgically removed for the patient to be seizure free. Brain activities of patients are recorded during one or two weeks, and the brain regions are then ranked according to their epileptogenicity, i.e., to the level of involvement in the initiation of epileptic discharges.

One of the most challenging aspects of these examinations is that the electrophysiological criteria are not clearly defined. A recently proposed marker consists in High-Frequency Oscillations (HFO, 80-500Hz) which have been shown to be an indicator of epileptogenicity [1]–[3]. HFOs can be divided into 3 bands: High-Gamma (HG, 80 – 150 Hz), Ripple (R, 150 – 250 Hz), and Fast-Ripple (FR, 250 – 500 Hz) bands. There is however neither formal consensus nor a tool that enable clinicians to detect and identify these events objectively. The visual review of HFOs is time-consuming, tedious, and hardly reproducible because of the short duration (about a hundred milliseconds or less) and because the frequency (f) spectrum is dominated by low frequencies (1/fα spectrum).

No automated detectors [4]–[7] have yet met a large consensus. Among possible strategies, time frequency (TF) analysis is an important tool for characterizing HFOs. In particular, it permits us to separate transients from actual oscillations, which have distinct signatures in the TF plane; however, the 1/fα spectrum impacts the TF maps. One obstacle that still remains is how to normalize the amplitudes across frequencies in order to capture equally well the activities in all the bands.

The purpose of this study is thus to propose a tool to objectively visualize and identify HFOs based on TF normalization or prewhitening. The underlying concept is to flatten, i.e., “whiten”, the spectrum in order to have a balance of power across frequencies. This flattening could however increase the noise and generate false detections. We therefore investigated and compared methods on different criteria and ranked them using the Area Under the Curve (AUC) of Precision and Recall (PR) curves. We applied these methods to simulated data with real human background (BKG) activity.

II. METHODS

A. Continuous Wavelet Transform

The Continuous Wavelet Transform (CWT) of a signal f is defined as

\[
T_f(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi \left( \frac{t-b}{a} \right) dt, \tag{1}
\]

with ψ the wavelet function, a the scaling factor, and b the shifting factor. We chose to normalize the wavelet in \(L^2\),

\[
\psi_{a,b}(t) = 1/\sqrt{a} \psi ((t-b)/a), \tag{2}
\]
in order for the estimate of the power spectrum – the wavelet spectral function [8] – to correspond to
\[ V_{B}^{L^{2}}(a) = \frac{1}{B} \int_{-B/2}^{+B/2} |T_{f}(b,a)|^2 \, db, \quad (3) \]
with \( B \) the duration of the time window, which needs to be large enough for standard ergodic arguments. In other words, the wavelet spectral function has to be the average over time of the TF image for each frequency. This is important to assure consistency between the prewhitening methods (that operate in the time domain) and the TF normalization methods. A normalization of the wavelet in \( L^{1} \) would lower even more the power of the low frequencies in the TF image and act as another filtering stage (more details in Appendix A). We therefore used the normalization in \( L^{2} \) defined in (2).

We used the analytic Derivative of Gaussian (DoG) wavelet (a specific case of the Morse wavelet) with \( n = 20 \). It is expressed in the frequency domain as
\[ \tilde{\psi}(f) = \left\{ \begin{array}{ll} f^n \exp(-f^2) & \text{for } f \geq 0 \\ 0 & \text{for } f < 0. \end{array} \right. \quad (4) \]

We used the analytic DoG wavelet for its good mathematical properties: it is null for negative frequencies and thus provides a better estimate of the phase [9], unlike the classical Morlet wavelet. Such analytic wavelets could provide a better strategy for analyzing HFOs [10]. To compare the different whitening strategies, we computed the TF image in log-scale with 3 octaves (Oct) and 12 voices (Voi) to cover the \( 68-512 \) Hz band (i.e. the band of interest for HFOs). For the figures, we increased the number of Oct to obtain a broader frequency range (below 68Hz) for a better overview of the data.

B. Whitening Strategies

1) Prewhitening in time-domain: The different prewhitening methods were applied in the time-domain. Let \( n \) be the time index.

The Diff method consists in a first-order backward differencing [5]. Let \( \hat{x} \) be the prewhitened signal
\[ \hat{x}[n] = x[n] - x[n-1]. \quad (5) \]
The motivation behind the diff method is that it suppresses the continuous component of the signal and lowers the low frequencies. The Discrete-time Fourier transform (DTFT) applied to a differentiated signal \( x \) gives
\[ \text{DTFT} \{x[n] - x[n-1]\} = (1 - e^{-j\omega}) X (e^{j\omega}), \quad (6) \]
Thus its power spectral density (PSD) is
\[ \left| (1 - e^{-j\omega}) X (e^{j\omega}) \right|^2 = \left| 1 - e^{-j\omega} \right|^2 \left| X (e^{j\omega}) \right|^2 \]
\[ = 2j \sin \left( \frac{\omega}{2} \right)^2 \left| X (e^{j\omega}) \right|^2 \]
\[ = 2 (1 - \cos (\omega)) \left| X (e^{j\omega}) \right|^2, \quad (7) \]
with \( X \) the DTFT of the signal \( x \), \( \omega = 2\pi f/f_s \) its normalized frequency, \( f \) its frequency and \( f_s \) the sampling frequency. The scalar \( (1 - \cos(2\pi f/f_s)) \) is monotonically increasing from \([0, f_s/4]\) in \([0,1]\), which indeed lowers the low frequencies compared to the high frequencies. The Diff prewhitening was applied on segments of interest (SOI), i.e. periods of BKG where an event of interest was added.

The autoregressive integrated moving average (ARIMA) [11] prewhitening computes the coefficients of a \( p^{th} \)-order AR model (e.g. Matlab’s LPC function) on the \( d^{th} \)-degree differentiated signal and filters the signal with the coefficient of the AR model. This method aims at finding the trend of the spectrum to flatten it afterward. Let \( a_i \) (with \( i \in \{1,2,\ldots,n\} \)) be the \( i \)th coefficient of a \( p^{th} \)-order AR model computed on \( x \). The DTFT of a signal \( x \) prewhitened with the AR prewhitening gives
\[ \text{DTFT} \{ (x * AR(p)) [n] \} = \left( 1 - \sum_{k=1}^{p} a_k e^{-jk\omega} \right) X (e^{j\omega}), \quad (8) \]
where “*” corresponds to the convolution. In other words, the AR prewhitening subtracts the autoregressive part of the spectrum from the spectrum. The ARIMA prewhitening is similar but the AR coefficients are computed on the \( d^{th} \)-degree differentiated signal, yielding
\[ \text{DTFT} \{ (x * ARIMA(p,d,0)) [n] \}
\[ = \left( 1 - \sum_{k=1}^{p} a_k e^{-jk\omega} \right) (1 - e^{-j\omega})^d X (e^{j\omega}). \quad (9) \]
It is noteworthy that the ARIMA(0,1,0) prewhitening is equivalent to the Diff method (6). As discussed in [11](Section 14.6), the differencing allows us to effectively manage the drifts and trends in the data that would otherwise overwhelm the frequencies. The fact that the last parameter is null means that there is no moving average (MA). We use ARIMA with the MA set to zero, because we want the “pre-emphasis” filter to be strictly FIR and not “smear” a transient too far in the data, which would happen if we allowed an MA component in the ARIMA. Several parameters were tested, and the set \((15,1,0)\) gave the best results. The ARIMA filter was applied on SOI and not on baselines. Increasing the order of the AR model would lead to over-fitting the spectrum and thus prewhiten also the signal of interest.

2) Time-Frequency Normalization: Normalization methods were applied to the TF image of the original signal. Let \( n \) and \( m \) be the time and frequency indices.

The Teager-Kaiser Operator Energy (TKEO) is mostly used to identify the instantaneous frequency and the amplitude of non-stationary signals [12], [13]. The TKEO applied to a continuous signal \( x(t) = A \sin (\Omega t + \phi_0) \) corresponds to
\[ \text{TKEO} \{ x(t) \} = \hat{x}^2(t) - x(t) \hat{x}(t) \]
\[ = A^2 \Omega^2 \cos^2 (\Omega t + \phi_0) \]
\[ + A^2 \Omega^2 \sin^2 (\Omega t + \phi_0) \]
\[ = A^2 \Omega^2, \quad (10) \]
where “*” and “’” describe the first and the second derivative respectively. The energy obtained depends on the amplitude and frequency of the wave. Its equivalence in discrete time is:
\[ \text{TKEO} \{ x[n] \} = x^2[n] - x[n-1]x[n+1]. \quad (11) \]
Note that the TKEO cannot be used as a prewhitening method, since its output is already an energy. Indeed TKEO normalization (used in [14]) operates on the complex coefficients of the TF image of the SOI over time and for each frequency taken separately,

\[
T_f^{\text{TKEO}}[n, m] = T_f[n, m]T_f[n, m] - \frac{1}{2}T_f[n - 1, m]T_f[n + 1, m] - \frac{1}{2}T_f[n - 1, m]T_f[n + 1, m],
\]

(12)

where “\( \tilde{\cdot} \)” denotes the complex conjugate.

The event-related spectral perturbation (ERSP) is a common way to normalize the TF maps in electrophysiological reviewing [15]. This method computes the mean \( \mu[m] \) of the square modulus of the TF coefficients \( |T_f[n, m]|^2 \) of a chosen baseline over time and for each frequency taken separately, and we apply this transformation

\[
\text{ERSP}_f[n, m] = |T_f[n, m]|^2 / \mu[m]. \quad (13)
\]

The baseline was taken in the same BKG but in a time-shifted window.

The z-score, another commonly used method, was applied to \( |T_f[n, m]|^2 \) using the mean \( \mu_{\text{type}}[m] \) and the STD \( \sigma_{\text{type}}[m] \) with type corresponding to either a baseline or the SOI itself, for each frequency taken separately. The baseline was selected the same way as for the ERSP. The two types of z-score will be further referred to as \( Z_{\text{baseline}} \) and \( Z_{\text{SOI}} \), and the TF transformation is

\[
Z_f^{\text{type}}[n, m] = \frac{|T_f[n, m]|^2 - \mu_{\text{type}}[m]}{\sigma_{\text{type}}[m]} \quad (14).
\]

In fact, \( \mu[m] \) is equivalent to the wavelet spectral function \( V_B \) defined in (3). ERSP and the \( Z_{\text{baseline}} \) whiten the data by either dividing the \( |T_f[n, m]|^2 \) coefficients by \( V_B \) or subtracting \( V_B \) from the coefficients, which would result in a flat spectrum in presence of BKG activity only. However, the results of \( Z_{\text{baseline}} \) and ERSP depend on the quality of the chosen baseline and its level of similarity with the BKG activity in the window of interest. For evoked potential, the baseline is usually taken in the window preceding the stimulus. It is to be noted that this does not guarantee that the BKG activity is similar before and after stimulus [16]. In HFO studies, finding a baseline for each channel can be difficult. This could be solved by the \( Z_{SOI} \) that uses the event time window as a reference, but this implies that the events occur rarely within the window. This is however a strong hypothesis which is usually not verified. We therefore propose a new method – “\( H_0 \) z-score” (\( Z_{H_0} \)) – which is built to be more robust to the SNR across frequency and does not require a baseline.

The first step of this method is to estimate the noise distribution in the complex plane at each scale. We thus made two hypotheses: 1) the global distribution \( H_0 \) per scale is made of a zero-mean Gaussian distributed noise \( H_0 \) and our signal of interest \( H_1 \); and 2) there is no correlation between the real and imaginary part in the center of the complex plan (the noise). This enables us to fit a Normal distribution on the center of the distribution of the real and imaginary part of the coefficients separately and calculate their mean and STD. Then the real and imaginary part are z-scored and the square modulus is taken to generate the TF image. The pseudo-code summarizing this technique is given below:

1: for \( a = 1^\text{st} \) scale to last scale do
2: \( T_f^{re} \leftarrow \Re \{T_f(:, a)\} \)
3: \( T_f^{im} \leftarrow \Im \{T_f(:, a)\} \)
4: calculate the first (\( Q_1 \)) and third (\( Q_3 \)) quartile and the interquartile range (IQR) of \( T_f^{re} \) and \( T_f^{im} \)
5: \( (\mu^{re}, \sigma^{re}) \leftarrow \text{Gaussian fit on } Q_1^{re} - 1.5\text{IQR}^{re} \leq T_f^{re} \leq Q_3^{re} + 1.5\text{IQR}^{re} \)
6: \( (\mu^{im}, \sigma^{im}) \leftarrow \text{Gaussian fit on } Q_1^{im} - 1.5\text{IQR}^{im} \leq T_f^{im} \leq Q_3^{im} + 1.5\text{IQR}^{im} \)
7: \( Z_{H_0,f}^{re} \leftarrow \frac{T_f^{re} - \mu^{re}}{\sigma^{re}} \)
8: \( Z_{H_0,f}^{im} \leftarrow \frac{T_f^{im} - \mu^{im}}{\sigma^{im}} \)
9: \( Z_{H_0,f}(\cdot, a) = Z_{H_0,f}^{re} + iZ_{H_0,f}^{im} \)
10: end for

One characteristic of the \( Z_{H_0} \) is that it whitens the TF by forcing the real and imaginary part of the coefficients to have broadly the same distribution across frequencies (Fig. 1).

The methods are illustrated in Fig. 2; their effect on the power spectrum is represented in Fig. 3.

Visualizing the signal in the time-domain is very important for clinicians, because they are more used to this type of representation than the TF map. One could argue that the normalization techniques do not permit to visualize the whitened signal in time, whereas the prewhitening methods do. In fact it is not the case for \( Z_{H_0} \), thanks to the properties of the analytic DoG wavelet. Given that \( f(\cdot) \) is a real-valued signal and the wavelet \( \psi \) is analytic, a reconstruction formula [17] is proportional to

\[
f(t) \propto \Re \int_{-\infty}^{+\infty} \int_0^{+\infty} T_f(b, a) \frac{1}{\sqrt{a}} \psi \left( \frac{t - b}{a} \right) \, db \, da / t^2. \quad (15)
\]

with \( \Re \) the real part of the complex number. Since \( \hat{\psi} \in L^1(\mathbb{R}_+, da/a) \), this could be simplified by using the linear
Fig. 2. Examples of the 5 types of simulated events in time and time-frequency domains. The events are represented in a 500 ms-windows. The columns correspond to an event type, the first line shows the time-series, and the other lines illustrate the result of each method in the TF map. The HFOs were generated with an SNR of 10 dB. Each colormap was normalized between the minimum and maximum value of the image. As expected, the HFOs are not visible in the raw TF representation but appear in the whitened signal. Note how the $Z_{SOI}$ alters the images, especially the spike and the HFO.

The analysis-reconstruction scheme [8] into

$$f(t) \propto \Re \int_{0}^{+\infty} T_f(t,a) \frac{da}{a\sqrt{a}}.$$  
$$\propto \int_{0}^{+\infty} \Re \{ T_f(t,a) \} \frac{da}{a\sqrt{a}}.$$  
$$\propto \int_{0}^{+\infty} \sigma(a) \Re \{ Z_{H_o,f}(t,a) \} \frac{da}{a\sqrt{a}}.$$  

We thus define the $Z_{H_o}$ whitened signal as

$$\hat{f}(t) = \int_{0}^{+\infty} \Re \{ Z_{H_o,f}(t,a) \} \frac{da}{a\sqrt{a}}.$$  

At the computational level, since we choose a log-scale representation and neglect the multiplicative factor, the whitened signal can be computed as

$$\hat{f}[n] \approx \sum_{m=1}^{\text{Oct} \times \text{VoI}} \frac{1}{a[m]} \Re \{ Z_{H_o,f}[n,m] \}.$$  

One should be aware that if the number of Oct is not high enough, the reconstructed signal will correspond to the original signal being band-pass filtered. Indeed the lower frequency captured by the TF decreases when the number of Oct increases. Moreover, the accuracy of the reconstruction increases with the number of voices per octave [17].

C. Simulated Data

In order to compare the different methods, we simulated five types of signals. Each segment measures 5 s and is composed of human BKG activity and a type of events. The different types are the following: BKG activity alone ($s_1$), BKG activity with an artifact ($s_2$), BKG activity with a simulated epileptic spike ($s_3$), BKG activity with a simulated HFO ($s_4$), and finally BKG activity with a simulated spike and simulated FR ($s_5$). The spike, the HFO and the artifact occur at 2.35 s, 2.5 s and 3.75 s respectively. Spikes and artifacts were chosen because they lead to false detection in common marking methods. Epileptic spikes were simulated using the spline function of MATLAB, which interpolates the curves between specific points taken from a real epileptic spike. The width of the spike randomly changed across trials. Its amplitude was set to be proportional to the standard deviation (STD) of the BKG. Three types of HFOs were produced, one for each HFO band with frequencies of 323 Hz, 181 Hz, and 114 Hz and a duration of six periods. To avoid edge effects, the spike and the FR were windowed beforehand and then added to the BKG. The
Fig. 3. The effect of the several methods on the spectrum are shown. 100 simulations of BKG with HFOs (SNR = 10 dB) were generated. The average estimate power spectrum $V_B$ of the BKG with HFO and the BKG are represented with solid and dotted line respectively. The dashed line corresponds to the standard deviation. The lines define the three simulated HFO types (FR: 323 Hz, R: 181 Hz, HG:114 Hz). For the representation, an offset of 1 was added to the $Z_{baseline}$ and $Z_{SOI}$ to have positive values. The raw spectrum has a “hockey-stick like” spectrum, meaning it follows two power laws $1/f^\alpha$, with $\alpha_1 \approx 3$ and $\alpha_2 \approx 1$ for the frequency range of $10^{-2}$ to $10^{-1}$ Hz and $250$ to $500$ Hz. The HFOs clearly stand out of the BKG around their respective frequency. All the whitening methods manage to flatten the spectrum but the Diff and TKEO do not succeed in removing the “hockey-stick like” spectrum and thus enhance the BKG high frequency.

Each $s_i$ event was processed for each prewhitening or normalization method. However, those which were prewhitened were not normalized in TF and conversely, as described in Fig. 4

D. Real Data with HFOs

The aforementioned techniques were applied to real HFOs to give an illustration of clinical applications. These real HFOs were marked using an automatic detector [5] and verified visually [20] and were recorded from the same patients as in the simulation part II-C.

E. Method Quantification

To capture how relevant the representations are, we want to quantify how the oscillations are separated from the BKG activity. It is known that oscillations are well localized in TF images. They appear as “islands” or “blobs” [21] whereas spikes and artifacts are extended across frequencies. While analyzing visually such representations, we pay attention to the local maxima rising above the noise level. The method proposed here is to quantify the detectability of HFO by constructing Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves. To compute these curves, all local maxima in the HFO band are detected and labeled as True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) for each value of the threshold. This threshold takes increasing values ranging from the minimum to the maximum values across all TF images for each method and each HFO type independently. For the three types of simulated HFOs, 30 events of each type were generated.

TPs are local maxima which are above the threshold and are our peaks of interest, i.e, local maxima of all signals

artifact was simply generated by increasing a single point by a certain level (impulse function). This level corresponds to five times the STD of the chosen BKG. Examples of the five events are shown in Fig. 2. The piece of human BKG activity was randomly selected from a collection of recordings (sampling frequency: 2048 Hz) which was previously labeled as BKG, i.e signal without one of these events, from several patients and several brain areas. These recordings were performed on patients undergoing pre-surgical evaluation of drug resistant epilepsy with SEEG during slow wave sleep where HFOs are usually studied [18]. For standardization, the collection of BKGs were normalized by dividing each signal by its own standard deviation and multiplied by the median STD of the real data collection. Unlike previous studies [6], [19], the STDs were not computed on raw data but on data which were digitally bandpass-filtered (4th-order Butterworth) in the HFO band. The SNR was also calculated on the filtered data on the time duration of the HFO. This approach is motivated by the fact that SEEG signals are dominated by the low frequencies which would have increased the SNR in a non-representative manner.
$s_4$ and $s_5$ which are above the threshold and are in the confidence zone. The confidence zone was set as the zone of the image where the blob of the HFO should theoretically appear. It is a rectangular zone centered on the simulated HFO and at the frequency of the oscillation with a time width of $2\sigma_t$ and a frequency width of $2\sigma_f$, where $\sigma_t$ corresponds to the theoretical width of the wavelet at the corresponding frequency. $\sigma_f$ was heuristically set to 30 Hz as being the accepted error on the frequency of the oscillation. FP are local maxima which are above the threshold, but are not our peaks of interest, i.e. local maxima of all signals $s_1$, $s_2$, and $s_3$, which are above the threshold, plus those of all $s_4$ and $s_5$ above the threshold, which are not in the confidence zone. TNS are local maxima which are not above the threshold and are not our peaks of interest, i.e. local maxima of all signals $s_1$, $s_2$, and $s_3$, which are above the threshold, plus those of all $s_4$ and $s_5$ under the threshold which are not in the confidence zone. FN are local maxima which are not above the threshold but are our peaks of interest, i.e. local maxima of all $s_4$ and $s_5$, which are not above the threshold but are in the confidence zone.

ROC and PR curves are obtained by calculating the True Positive Rate (TPR) or Recall, the False Positive Rate (FPR) and the Precision or Positive Predictive Value (PPV) as follows

$$TPR = TP/P = TP/(TP + FN),$$

$$FPR = FP/N = FP/(FP + TN),$$

$$PPV = TP/(TP + FP).$$

The ROC curve represents TPR as a function of FPR and the PR curve PPV as a function of TPR.

The Area Under the Curve (AUC) of the PR curves was used as a criterion to rank the different whitening strategies. This measure was repeated 30 times for each SNR with random human background activities. The different SNRs were chosen according to the range seen in real data. The STD of the oscillation was obtained by decomposing the signal using the Empirical Mode Decomposition (EMD) [22] and taking the mode corresponding to the HFO band. The obtained signal was checked on TF representation before and after decomposition to ensure the correct EMD filtering. The STD of the noise was computed on two pieces of filtered BKG before and after the HFO occurs with overall length of the oscillation. The SNR of the real data were found to lie between 0 and 17 dB with a median value of 9 dB.

### III. Result and Discussion

Fig. 3 represents the average spectrum of the signals $s_1$ and $s_2$ at SNR = 10 dB. In the first column, HFOs are dominated by the low frequencies in the non-whitened signal. This emphasized the importance of signal processing to visualize HFOs correctly. Moreover, the original spectrum has a “hockey-stick” like spectrum and is best fitted by two power laws $1/f^{\alpha_1}$, with $\alpha_1 \approx 3$ and $\alpha_2 \approx 1$ for the frequency range of $10 - 200$ Hz and $250 - 500$ Hz (the FR band), respectively. All the techniques managed to flatten the spectrum. The Diff and TKEO methods do not succeed in removing the “hockey-stick” like behavior and consequently amplify the high frequencies, whether they are due to FRs or BKG. This could lead to more FPs. Furthermore, the R band appears to be in the “elbow” of the raw spectrum which is turned into the global minimum of the Diff and TKEO spectra. In other words, a R will need a high enough amplitude to overcome the high frequency BKG. Visually, the ARIMA, ERSP and z-score methods exhibit better performance in flattening the spectrum. The average BKG spectrum is flat and the HFOs stand out of the BKG. The peaks of the HFOs on the $Z_{SOI}$ spectral is nevertheless smoother than on the other spectra which is probably due to the aforementioned drawback of this method. Because of the high number of negative events (N) compared to the number of positive (P) events the ROC curves are pushed to the left-hand side corresponding to the low FPR values and are thus not discriminative or even seems wrongly efficient. This imbalance is due to the intrinsic local maxima generated by the BKG activity. In contrast, the PR curves are not sensitive to this imbalance and thus highlight differences between the methods [23], [24]. In a clinical setting, it seems interesting to address the proportion of TP within all detections (PPV (21)) regardless of N, i.e. be more precise (high PPV (21)) than specific (high FPR (20)). PR curves are therefore preferred for further analysis.

Box and whisker plots of the AUC of the 8 methods applied to the simulated data are represented in Fig. 5 for
5 different SNRs. Raw TF representation which corresponds to the case without normalization nor prewhitening shows the worst performance overall, except for the ideal case with SNR = 20 dB. This is consistent with the aim of this study. Generally all the methods improve with increasing SNR. They exhibit poor performance for SNRs below 5 dB without having one method being significantly better than another. This is not the case for larger SNRs. The ERSP and z-score methods have better results for SNR = 5 dB for simulated FRs and Rs. As expected, the Z_SOI falls behind the other normalization methods for all the frequency bands from SNR = 10 dB and its performances decrease after SNR = 15 dB. The more the signal of interest weights in the distribution, the more the parameters for the z-score are overestimated. The ERSP, Z_{baseline} and Z_{Ho} perform the best across SNRs and frequency bands. The ARIMA method is the best prewhitening technique, but is still less efficient than the latter. The performance of the ERPS and Z_{baseline} may not be representative of the results which could be found in clinical settings because of the way the baseline was selected. In our study, the baseline was selected in the same BKG but in a time-shifted window. This means that the characteristics of the baseline were very similar to those of the BKG by construction. Z_{Ho} is thus better in principle since it does not require a baseline. Moreover the normalization and ARIMA methods are very sensitive to the duration of the chosen window. Using windows below 5 s (≈ 10000 samples) would lead to bias in the estimators and deteriorate the performance of the methods.

IV. APPLICATIONS

A. Objective identification of HFOs

The Z_{Ho} could improve the objective identification of HFOs since it is independent on a baseline and frequency bands and enables the HFOs to be easily spotted without increasing the BKG activity. In Fig. 6 four real events are represented in the TF maps using the aforementioned whitening processes. A spike with a FR, a gamma oscillation, a HG, and a FR enable the HFOs to be easily spotted without increasing the background or the signal. It is very clear in the first and second column respectively. As discussed above, the ARIMA and Z_SOI tend to increase high-frequencies whether they originate from the background or the signal. The Diff and TKEO methods seem to work fine when a single event occurs. In Fig. 3 we showed that these methods could not remove the distortion of the spike made by the Z_SOI. The Diff and TKEO methods seem to work fine when a single event occurs. In Fig. 3 we showed that these methods could not remove the distortion of the spike made by the Z_SOI.
Fig. 6. Examples of 4 real events in time and time-frequency domains. The events are represented in a 250 ms-window and are taken from 2 patients. The first to the last column corresponds to a spike with a FR, a gamma oscillation, a HG, and a FR respectively. The first line shows the original time-series, the second line displays the whitened reconstructed signals, and the other lines illustrate the result of each method in the TF map. Each color map was normalized between the minimum and maximum value of the image. As in Fig. 2, the HFOs are not visible in the raw TF representation but appear in the whitened signal and TF. However some slight differences can be noticed between the two baseline based methods and the \( Z_{H_0} \). The HG, FR and Gamma oscillations are visible, and most importantly the FR and the spike are distinguishable even while co-occurring. Fig. 7 illustrates in wider frames other real events which are clearly not visible in the normal TF representation and difficult to capture in the time-domain without stretching the signal but are evident in the \( Z_{H_0} \) image. This normalization could especially improve identification of HFOs which are not distinguishable in the time series. There is still a debate on HFOs occurring within spikes. There are some reports of HFOs not visible in spikes [1], but it is possible that part of these could arise from filtering artifacts [21]. Time frequency methods should in principle improve this situation [19], but may be “blinded” by the high energy contained in the spike [25]. Fig. 8 represents a simulated FR which occurs during the build up of the spike (dashed boxes) and is clearly not visible in the original signal, and only slightly distinguishable when we zoom in. The \( Z_{H_0} \) in the time or TF domain enables the researchers/clinicians to identify the HFOs without having to stretch or filter the signal to ensure the correctness of such “almost not visible HFO in spike”. One should be aware that even if the oscillation cannot be captured in the time-series, it does not necessarily mean that there is no oscillation. The \( Z_{H_0} \) whitened reconstruction enhances greatly the FR without denaturing the epileptic spike. This method is also convenient to study the signal without having to saturate the color map and thus avoid losing the global significance of the signal.
B. Application beyond SEEG

This study mainly focuses on SEEG because HFOs are mainly studied in SEEG recordings. Recently some groups have tried to study HFOs in scalp electroencephalography (EEG) [26], [27] and magnetoencephalography (MEG) [28]. Since the assumptions made by the $Z_{H_0}$ are not relative to SEEG, one could easily apply it to EEG or MEG data since they also have a $1/f^\alpha$ spectrum. A gamma wave (G, 25 – 80 Hz) marked in an MEG recording is shown in Fig. 9. The TF map is whitened as for SEEG data and the gamma is highlighted.

C. Framework for a new detector

Using our method should speed-up the visual marking of HFOs. Nevertheless, it would still be time-consuming and would need human resources. Interestingly, the method could be integrated in existing detectors such as [7]. These detectors are based on a two- or three-step algorithm. First a threshold is applied on the time-series according to its short-time energy [4] or short-time line length [5], but this step suffers high FP [5]; a second step was thus introduced to decrease this number. The second stage usually uses time-frequency images which could be strengthen by the $Z_{H_0}$. We however think that the signal should already be whitened at the first step because the short-time energy and short-time line length are greatly influenced by the $1/f^\alpha$ spectrum even if they are band-pass filtered between 80 – 500 Hz. In fact, the Diff method was used to prewhiten the signal in [5] to improve detections and decrease the number of missed HFOs (FN), but this method does not exhibit the best performance as we showed above. Moreover, the study was made on band-passed data between 0.1 – 100 Hz because of the low sampling frequency (200 Hz). We believe that FN would even be higher on data recorded with the current sampling frequency (2048 Hz) without whitening processing. Therefore to solve these issues, HFOs could be detected directly in the $Z_{H_0}$ TF map. Such a detector would select all local maxima above a threshold and label these maxima according to the shape of the “blob” into oscillation, spike or other. Using $Z_{H_0}$ normalization would actually have several interesting assets. By construction, the background activity can be modeled by a standard normal distribution ($\mu = 0$ and $\sigma = 1$) at each frequency. The local False Discovery Rate (IFDR) [29] is an empirical Bayes approach which is based on the similar hypothesis than the $Z_{H_0}$. In short, this method assumes that the observed data histogram $H_G$ is a mixture of $H_0$ (noise) and $H_1$ (signal of interest), and that $H_0$ is in the center of the histogram. It then defines a threshold such as

$$thr = x \quad | \quad IFDR(x) = H_0(x)/H_G(x) < Q,$$

(22)

with $thr$ the threshold and $Q$ the accepted level of IFDR. The IFDR would provide one single threshold for all frequencies. Thresholding methods used in previous detectors [4], [7], [30] calculated the mean and the standard deviation of skewed distribution. By combining the IFDR and $Z_{H_0}$ we could guarantee that the detected local maxima are not generated by the background. Another step would be needed to separate the oscillations from the spikes. Fortunately the $Z_{H_0}$ conserves the properties of the wavelet. An interesting property is that the frequency width of the wavelets is constant in log-scale. This implies that the oscillations which correspond to two Dirac distributions in the frequency domain have the same width whatever their frequency. A simple threshold on the frequency width could differentiate oscillations from spikes. The measured frequency width is the Full Width at Half Maximum (FWHM) of the island in the frequency axis. In the literature [1], [2], [4], a threshold on the duration is used to select only oscillations with at least 3-4 periods. To be
consistent this threshold should be different at each frequency. An approach could be to compare the time FWHM of the response of a Dirac in the TF map and the time FWHM of the selected island. This would provide a constant threshold across frequencies. It is noteworthy that this detector labels all oscillations without prior frequency bands. An example of the detection in the parameter plane is shown in Fig. 10. The black dots represent all the detected local maxima, the triangles correspond to the local maxima that were labeled as spikes and the circles designate the local maxima labeled as oscillations. Some local maxima in the oscillation or spike area were not labeled because they were too close in time from another local maxima. An illustration of the results of the detection is shown in Fig 11. We implemented a prototype of this detector in our open source software AnyWave [31], as well as a reviewer add-on that enables the user to review and edit the previous detections. We are currently testing them in a clinical setting.
HFOs at different frequencies for several SNRs and localizing increasing the BKG activity. This was made by simulated according to their capability to flatten the spectrum without used techniques, plus one method which we designed to over-

HFOs representation in TF maps. We compared 7 commonly given frequency ranges. These ranges can be manually set up. The name of the contacts in the mesiolateral axis.

TB: Temporo-basal, ': left hemisphere). The numbers indicate the indices (A: Amygdala, B: Hippocampus, H: Heschl’s Gyrus, OR: Orbitofrontal, the channels correspond to the name of the electrodes in bipolar montage

displayed in a bar graph giving the spike rates and the oscillation rates of local maxima that were labeled as spikes and oscillations respectively. The triangles, and the circles are the island relative to the local maxima. The black dots represent all the
data correspond to the same frequency. The frequency spread (FWHM) in time of the island relative to the local maxima compared to the

time of the impulse response at the same frequency. The frequency spread

effect of the whitening methods in lower frequency bands would like to point out that real HFOs could be not visible in time-domain because they are hidden by the slope of a sharp transients, which can be counter intuitive for neurophysiologists. The ZH0 provide an objective tool to identify HFOs since it does not over express high-frequency activity of the background and dissociate sharp transient from HFOs in time and TF domain. We have suggested that this technique could be applied on other types of electrophysiological recording with a power law spectrum. Future work will focus on the effect of the whitening methods in lower frequency bands and will test the detector on a large number of patients from multiple centers.

APPENDIX A

The estimate of the power spectrum VB in L2 corresponds to

\[ V_B^{L_2}(a) = \frac{1}{B} \int_{-B/2}^{+B/2} \left| T_f^L(b,a) \right|^2 db. \]  

In L1 it corresponds to

\[ V_B^{L_1}(a) = \frac{1}{B} \int_{-B/2}^{+B/2} \left| T_f^L(b,a) \right|^2 db. \]  

Note that \( V_B^{L_2}(a) = V_B^{L_1}(a) \). The main difference is that \( V_B^{L_2}(a) \) is simply the mean over time of the scalogram for the scale a, and \( V_B^{L_1}(a) \) is the mean over time of the spectrogram for the scale a, but multiplied by the scale a.

ACKNOWLEDGMENT

This study was financed by the ANR-13-TECS-0013. Research by JCM was supported in part by the National Institute of Neurological Disorders and Stroke under grant number R01NS089212 and National Institute of Biomedical Imaging and Bioengineering under grant number R01EB009048, both of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank Bruno Torrésani for his help on the analytic DoG wavelet implementation.

REFERENCES


V. Conclusion

This paper has examined several whitening methods for HFOs representation in TF maps. We compared 7 commonly used techniques, plus one method which we designed to overcome the drawbacks of other methods. We ranked the methods according to their capability to flatten the spectrum without increasing the BKG activity. This was made by simulated HFOs at different frequencies for several SNRs and localizing all local maxima in the TF image. The AUC of the PR curves was used as a performance criterion. We determined that the best methods were the ERSP, Zbaseline and ZH0. However, as the ERSP, Zbaseline require a baseline, the ZH0 outperforms them by its robustness to non-stationarity of background activities. This technique also permits to reconstruct the whitened signal, which is an obvious advantage for clinical application. We would like to point out that real HFOs could be not visible in time-domain because they are hidden by the

Fig. 10. Representation of the local maxima in the feature space. The time ratio in the x-axis corresponds to the Full Width at Half Maximum (FWHM) in time of the island relative to the local maxima compared to the width of the impulse response at the same frequency. The frequency spread in the y-axis is calculated in the TF image as the FWHM in frequency of the island relative to the local maxima. The black dots represent all the local maxima above the threshold. The triangles, and the circles are the local maxima that were labeled as spikes and oscillations respectively.

Fig. 11. Example of the detection rates. The results of the detections are displayed in a bar graph giving the spike rates and the oscillation rates of given frequency ranges. These ranges can be manually set up. The name of the channels correspond to the name of the electrodes in bipolar montage (A: Amygdala, B: Hippocampus, H: Heschl’s Gyrus, OR: Orbitofrontal, TB: Temporo-basal, ': left hemisphere). The numbers indicate the indices of the contacts in the mesiolateral axis.


