Characterization of the role of high-frequency oscillations in epileptic networks

Nicolas Roehri

Nicolas Roehri. Characterization of the role of high-frequency oscillations in epileptic networks. Neuroscience. Aix Marseille Université, 2018. English. NNT : 2018AIXM0001 . tel-02063483

HAL Id: tel-02063483
https://hal.science/tel-02063483
Submitted on 11 Mar 2019

AIX-MARSEILLE UNIVERSITÉ
INSTITUT DE NEUROSCIENCE DES SYSTEMES, INS, INSERM
ÉCOLE DOCTORALE : EDSVS (ED62)

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline: Sciences de la vie et de la santé
Spécialité: Neuroscience

Nicolas Roehri

Caractérisation du rôle des oscillations à haute fréquence dans les réseaux épileptiques

Characterization of the role of high-frequency oscillations in epileptic networks

Soutenue le 16/01/2018 devant le jury composé de:

Pr. Jean Gotman	MNI, Montréal, Canada	Rapporteur
Dr Michel Le Van Quyen	ICM, Paris, France	Rapporteur
Dr Maeike Zijlmans	UMC, Utrecht, Pays-Bas	Examinatrice
Pr. Philippe Kahane	GIN, Grenoble, France	Examinateur
Dr Christian-George Bénar	INS, Marseille, France	Directeur de thèse
Pr. Fabrice Bartolomei	INS, Marseille, France	Directeur de thèse

Numéro national de thèse/suffixe local: 2018AIXM0001/001ED62

Résumé

Touchant plus de 50 millions de personnes dans le monde, l'épilepsie est un problème majeur de santé publique. Un tiers des patients souffrent d'épilepsie pharmaco-résistante. Une chirurgie visant à enlever la région cérébrale à l'origine des crises – la zone épileptogène – est considérée comme l'option de référence pour rendre libre de crises ces patients. Les cliniciens cherchent à localiser cette zone lors d'un examen pré-chirurgical consistant à enregistrer l'activité électrique des régions présumées pathologiques grâce à des électrodes implantées dans le cerveau. Alors que l'enregistrement des crises (la période critique) est le moment privilégié pour délimiter cette zone, le taux d'échec chirurgical non négligeable a poussé la recherche d'autres marqueurs. Certains ont choisi de mieux quantifier la phase critique alors que d'autres ont préféré sonder la période intercritique (entre les crises). Pour cette dernière, les transitoires pointus appelés pointes épileptiques en sont le marqueur standard. Ce domaine a connu un essor après la découverte de petites rafales d'oscillations rapides : les oscillations à haute fréquence (HFOs).

Une HFO est une brève oscillation entre 80-500 Hz qui dure au moins 4 périodes. Les HFOs sont divisées en deux catégories, les ripples 80-250 Hz – connues pour être impliquées dans des processus physiologiques – et les fast-ripples 250-500 Hz qui ont initialement été trouvées en conditions épileptiques.

Par leur caractère très bref, le marquage visuel de ces petites oscillations est fastidieux et chronophage. Il faut plusieurs heures pour marquer dix canaux de 10 min. Il semble impératif de trouver un moyen de détecter automatiquement ces oscillations pour étudier les HFOs sur des cohortes de patients. Aucun détecteur automatique existant ne fait cependant l'unanimité.

Cette thèse a pour but de développer un nouveau moyen de visualiser les HFOs pour mieux les détecter automatiquement. Puis, une stratégie est mise en place pour caractériser et valider des détecteurs. Enfin, le nouveau détecteur est appliqué à une cohorte de patients pour déterminer la fiabilité des HFOs et des pointes dans la prédiction de la zone épileptogène.

Nous avons conçu une normalisation robuste du plan temps-fréquence qui permet de mieux saisir les HFOs tout en préservant un rapport signal sur bruit. Dans ce plan temps-échelle normalisé, les éléments se détachant du bruit ont été caractérisés et classifiés comme des oscillations ou des pointes. Ce détecteur a été nommé Delphos.

Comme le marquage visuel est chronophage et moyennement fiable, nous avons mis au point un banc de test pour la validation des détecteurs de HFOs basés sur des simulations réalistes. Ce banc de test permet de contrôler plusieurs paramètres de la survenue des HFOs et des pointes. Delphos a montré une sensibilité plus élevée que les quatre détecteurs standards et ses détections étaient presque toujours correctes (médiane : 100%).

Finalement, Delphos a été appliqué aux données de 30 patients avec des épilepsies variées. Nous avons comparé le taux par minute des HFOs, pointes et différentes variantes à un estimateur de la zone épileptogène. Le résultat principal était l'absence d'évidence montrant que les HFOs (ou ses sous-groupes) étaient meilleurs que les pointes pour prédire la zone épileptogène. L'un des marqueurs combinés, le taux croisé, i.e. la moyenne géométrique des taux de HFOs et de pointes, était meilleur que tous les autres marqueurs.

Dans cette thèse, nous avons développé des outils qui pourraient transférer les HFOs dans le domaine clinique tout en gardant l'information des pointes. La question concernant la séparation des ripples physiologiques et pathologiques reste ouverte. Une piste que l'on propose est l'analyse des HFOs et pointes en réseau de couches multiples plutôt que par regroupement de caractéristiques ou de co-occurrence entre les ripples et d'autres évènements physiologiques (e.g. fuseaux ou ondes lentes).

Abstract

Epilepsy is a major health problem as it affects 50 million people worldwide. One third of the patients are resistant to medication. Surgical removal of the brain areas generating the seizure – the epileptogenic zone – is considered as the standard option for these patients to be seizure free. Clinicians attempt to localize this zone during the presurgical evaluation. This investigation involves implanting depth electrodes into the brain to directly record putative pathological brain regions. While the ictal phase (i.e. during a seizure) is the preferred period to delineate this zone, the non-negligible rate of surgical failure has led to seek other electrophysiological criteria. Better quantification of the ictal stage was the path chosen by some teams whereas others have suggested to further examine the interictal period (the period between seizures). For the latter, sharp transients called epileptic spikes are the standard marker of the interictal period. This field however boomed after the discovery of small bursts of fast oscillations: the high-frequency oscillations (HFOs).

An HFO is a brief oscillation between 80-500 Hz lasting at least 4 periods. HFOs can be defined as ripples (80-250 Hz) – known to be implicated in physiological processes – and fast-ripples (250-512 Hz), which were initially found in epileptic conditions only.

Due to their short-lasting nature, visually marking the occurrence of these small oscillations is tedious and time-consuming as it takes several hours to mark ten channels of 10 min recording. Automatically detecting these oscillations seems an imperative stage to study HFOs on cohorts of patients. There is however no general agreement on existing detectors.

This thesis aims firstly to develop a new way of representing HFOs and to use this novel representation as a base for detecting HFOs automatically. A second objective is to design a strategy to properly characterize and validate automated detectors. Finally, it intends to characterize, in a cohort of patients, the reliability of HFOs and spikes as predictors of the epileptogenic zone using the validated detector.

A robust normalization of the time-frequency image was designed to better capture the activity of HFOs while preserving the spikes and an optimal signal-to-background-activity ratio for every frequency. In this normalized scalogram, events standing out from the background were characterized and classified as oscillations or spikes. This detector was named Delphos.

As visual marking is time-consuming and leads to low inter-rater agreement, we developed a benchmark to validate HFO detectors based on realistic simulations. This simulation permits to control the timing of the HFOs and spikes, their co-occurrences, their relative amplitudes, and their quantity. Delphos had higher sensitivity than the four standard detectors and had a constant high precision (median: 100%), i.e. the detections were almost always correct.

We finally applied Delphos on 30 patients with various refractory epilepsies and compared the rate per minute of the HFOs, spikes and different variations to a quantified estimator of the epileptogenic zone. The main result was that there was no statistical evidence showing that HFOs (or any of their subgroups) were better biomarkers than spikes in predicting the epileptogenic zone. One of the combined markers, the cross rate, i.e. the geometric

mean of the spikes and HFO rates, was statistically better than all the other markers.

In this thesis, we developed tools that could be translated to the clinical environment and could push further the application of HFOs while conserving the information given by the epileptic spikes. The question regarding the separation of the physiological and pathological ripples remains. As a future perspective, we propose multi-layer network analysis of HFOs and spikes as an alternative to feature clustering or co-occurrence of ripples and other physiological events (e.g. spindles or slow waves).

Acknowledgements

First of all, I am very grateful to my two supervisors Dr Christian-George Bénar and Pr. Fabrice Bartolomei for having accepted me as their PhD student. I also thank them for trusting me and alloying me to investigate my own ideas. I really appreciated the time they managed to find to answer my questions which often turned into hour-long methodological/philosophical discussions.

I would like to thank Pr. Jean Gotman, Dr Michel Le Van Quyen, Dr Maeike Zijlmans and Pr. Philippe Kahane for having accepted to be part of my thesis committee.

I also express my gratitude to all members of the INS institute and especially to director Dr Viktor Jirsa.

My thanks go to all members of the ANR TechSan FORCE project, and especially the members I interacted the most with in the LTSI team in Rennes: Dr Fabrice Wendling, Dr Anca Nica, Dr Isabelle Merlet, Dr Amar Kachenoura and Dr Nisrine Jrad; in the GIN team in Grenoble: Pr. Philippe Kahane, Dr Olivier David and Dr Manik Bhattacharjee and in Micromed: Mr Jean-Luc Magnon and Mr Cristiano Rizzo.

I further thank my co-workers from the PhD student office, Dr Francesca Pizzo, Dr Stanislas Lagarde, Mr Samuel Medina, and Dr Isabelle Lambert (who was unfortunately not enough present), with whom I shared interesting scientific discussions but also laughs and relaxing moments.

My thanks go also to the rest of the MEG team, Dr Jean-Michel Badier, Dr Bernard Giusiano, Dr Sophie Chen and Mr Bruno Colombet, with whom I could learn about topic non-related to my project and share our coffee break together.

I am very thankful to the medical doctors of the Clinical Neurophysiology Unit for letting me assist to there staff, and especially Pr. Agnes Trébuchon and Dr Aileen McGonigal for our exchange on various epilepsy-related topics, and Romain Carron for allowing my presence during one of his SEEG implantation.

I also thank Pr. Jean-Marc Lina and Pr. Bruno Torrésani who helped me a lot with the time-frequency analysis and especially the wavelet transform, and John C. Mosher for helping me with my first publication.

I would like to address a special thank to my parents who let me choose my own path and let me study what I wanted, a chance they did not have. Finally and foremost, thank you Clara for sharing my life and for being my haven, my shelter. These three years would have been extremely more difficult without you; I hope I can be as reassuring and supportive for you as you were for me.

Abbreviations

Am	amygdala
ANR	French National Research Agency
AR	autoregressive
ARIMA	autoregressive integrated moving average
AUC	Area Under The Curve
BKG	background
CI	confidence interval
CT	computed tomography
CWT	continuous wavelet transform
Delphos	Detector of ElectroPhysiological Oscillations and Spikes
DoG	derivative of Gaussian
DTFT	discrete-time Fourier transform
DWT	discrete wavelet transform
DYS	dysplasia
ECoG	Electrocorticography
EEG	Electroencephalography
EI	Epileptogenicity Index
EIZ	Epileptogenicity Index Zone
EMD	Empirical Mode Decomposition
ER	Energy Ratio
ERSP	event-related spectral perturbation
EZ	Epileptogenic Zone
FDR	False Discovery Rate
FLE	Frontal Lobe Epilepsy
FN	false negative
FP	false positive
FPR	False Positive Rate
FR	fast ripple
FWHM	Full Width at Half Maximum
GUI	Graphical User Interface
$\mathbf{H_0}$	null hypothesis
HFO	High frequency oscillation
HG	high-gamma
HIL	Hilbert detector
Hip	hippocampus
HS	hippocampus sclerosis
iEEG	intracranial EEG
ILAE	International League Against Epilepsy
IQR	interquartile range
IZ	irritative Zone
Les	lesions
lFDR	local False Discovery Rate
LTLE	Lateral Temporal Lobe Epilepsy
LVFA	low-voltage fast activity
MA	moving average
MCD	malformation of cortical development
MLTLE	Mesial-Lateral Temporal Lobe Epilepsy
MRI	Magnetic Resonance Imaging
MTLE	Medial Temporal Lobe Epilepsy
N	negative instance
NDT	neurodevelopmental tumors
NEIZ	Non Epileptogenicity Index Zone
$\mathcal{N}$	Normal distribution
NREM	non-rapid eye movement
Oct	octave
P	positive instance
pAUC	partial Area Under the Curve
PET	Positron Emission Tomography
PPV	Positive Predictive Value
PR	Precision and Recal

Prec	precision
PSD	power spectral density
R	ripple
REM	rapid eye movement
R-FR	Rs co-occurring with FRs
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
SCHF	semi-continuous high-frequency
SD	standard deviation
SEEG	Stereo-electroencephalography
Sens	sensitivity
SLL	Short Line Length
SNR	Signal-to-Noise Ratio
SOI	signal of interest
SOZ	seizure onset zone
Spe	specificity
Spk	interictal spike
Spk × HFO	spike and HFO cross rate
Spk-FR	Spks co-occurring with FRs
Spk-HFO	Spks co-occurring with HFOs
Spk-R	Spks co-occurring with Rs
Spk-R-FR	Spks co-occurring with Rs and FRs
STE	Short Time Energy
STFT	short-time Fourier transform
$\mathbf{T_N}$	test negative outcome
$\mathbf{T_P}$	test positive outcome
TF	time-frequency
$\mathbf{Thr_{Opt.\ i}}$	optimal threshold i
TKEO	Teager-Kaiser Energy Operator
TLE	Temporal Lobe Epilepsy
TN	true negative
TP	true positive
TPR	True Positive Rate
$\mathcal{U}$	Uniform distribution
V-EEG	Video-electroencephalography
Voi	voice
WE	wavelet entropy
$\mathbf{Z_{H_0}}$	H_0 z-score

Contents

List of Figures

List of Tables

List of Publications

[1] Nicolas Roehri, Jean-Marc Lina, John C Mosher, Fabrice Bartolomei, and Christian-George Bénar. Time-frequency strategies for increasing high frequency oscillation detectability in intracerebral. *IEEE Transactions on Biomedical Engineering*, 63(12):2595–2606, 2016.

[2] Nicolas Roehri, Francesca Pizzo, Fabrice Bartolomei, Fabrice Wendling, and Christian-George Bénar. What are the assets and weaknesses of HFO detectors? A benchmark framework based on realistic simulations. *PLoS ONE*, 12(4):e0174702, apr 2017.

[3] Nicolas Roehri, Francesca Pizzo, Stanislas Lagarde, Anca Nica, Aileen McGonigal, Fabrice Bartolomei, and Christian-George Bénar. High-frequency oscillations are not better biomarkers of epileptogenic tissues than spikes (Accepted). *Annals of Neurology*, 2017.

[4] Nicolas Roehri, Christian-George Bénar, and Fabrice Bartolomei. Method for detecting elements of interest in electrophysiological signals and detector (Patent), 2017.

[5] Francesca Pizzo, Nicolas Roehri, Hélène Catenoix, Samuel Medina, Aileen McGonigal, Bernard Giusiano, Romain Carron, Didier Scavarda, Karine Ostrowsky, Anne Lepine, Sébastien Boulogne, Julia Scholly, Edouard Hirsch, Sylvain Rheims, Christian-George Bénar, and Fabrice Bartolomei. Epileptogenic networks in nodular heterotopia: A stereoelectroencephalography study. *Epilepsia*, pages 1–12, 2017.

[6] Isabelle Lambert, Nicolas Roehri, Bernard Giusiano, Romain Carron, Fabrice Wendling, Christian-George Bénar, and Fabrice Bartolomei. Brain Regions and Epileptogenicity Influence Epileptic Interictal Spikes Production and Propagation during NREM Sleep in comparison with Wakefulness. *Epilepsia*, in press, 2017.

[7] Stanislas Lagarde, Nicolas Roehri, Isabelle Lambert, Agnes Trebuchon, Aileen McGonigal, Romain Carron, Didier Scavarda, Mathieu Milh Francesca Pizzo, Bruno Colombet, Bernard Giusiano, Samuel Medina Villalon, Maxime Guye, Christian-G Bénar, and Fabrice Bartolomei. Interictal SEEG Functional Connectivity in Refractory Focal Epilepsies using SEEG: the model of the malformations of cortical development (under review).

[8] Kostiantyn Maksymenko, Bernard Giusiano, Nicolas Roehri, Christian-G. Bénar, and Jean-Michel Badier. Strategies for statistical thresholding of source localization maps inmagnetoencephalography and estimating source extent. *Journal of Neuroscience Methods*, 2017.

[9] Samuel Medina, Rodrigo Paz, Nicolas Roehri, Romain Carron, Andrea Brovelli, Anne-Sophie Dubarry, Agnès Trébuchon, Olivier Coulon, Fabrice Bartolomei, and Christian-George Bénar. EpiTools: A software suite for presurgical brain mapping in epilepsy (under review). 2017.

I. General Introduction

I.1. Context

With 1% of the worldwide population being affected and an annual total cost in Europe of 15.5 billion euro [Pugliatti et al., 2007], epilepsy is a major health problem. One third of the epileptic patients have drug-resistant epilepsy. For these patients, resective surgery is considered as the only treatment capable, under certain conditions, of suppressing seizures. The main conditions are that the "site of the beginning and of the primary organization" of the epileptic discharge, i.e. the epileptogenic zone, has to be 1) focal, 2) clearly identified and 3) can be safely removed [Bancaud et al., 1965, Kahane et al., 2006]. Although often effective, epilepsy surgery still fails in a substantial percentage (more than 35% [Jobst et al., 2015]) of patients despite extensive and costly investigations. These investigations involve at least two hospitalizations of several weeks in order to record seizures and test the potential post-operative deficits. There is thus a legitimate need of improving the identification of the brain tissues that generates seizures.

Intracranial EEG recordings are considered as the "gold standard" for the identification of the cortical area to be removed [Schuele, 2016, Cardinale et al., 2016, Iida and Otsubo, 2017]. Such recordings are only performed when data obtained non-invasively remain insufficiently concordant (discordant, inconclusive or impacting functional areas). Invasive electrodes are placed either at the surface of the brain (electrocorticography) or deep inside the brain (stereo-electroencephalography). The main difficulty clinicians are facing is that the electrophysiological criteria for delineating the epileptogenic zone are not precisely defined nor understood. On the one hand, the ictal period – the period during a seizure – is the privileged moment to delineate the epileptogenic zone based on the region involved at seizure onset. The main drawback is that it necessitates the recording of spontaneous seizures. On the other hand, the interictal period – the period between seizures, which is the most recorded period – has been extensively studied as a source of potential biomarkers of epileptogenicity.

Interictal spikes (Spk), which are sharp and transient events, are the standard marker of this period. Recently however, high-frequency oscillations (HFOs, 80-500 Hz) [Staba et al., 2002, Urrestarazu et al., 2007] have ousted spikes from the research spotlight. In studies using intracerebral EEG recordings, these brief and small oscillations are considered to be more focal and specific to the epileptogenic zone than spikes [Jacobs et al., 2008, Jacobs et al., 2009, Haegelen et al., 2013, van't Klooster et al., 2015]. The visual review of HFOs is however time-consuming, tedious, and hardly reproducible [Menendez de la Prida et al., 2015, Spring et al., 2017]. Moreover, no automated detectors [Staba et al., 2002, Gardner et al., 2007, Zelmann et al., 2010, Birot et al., 2013, Burnos et al., 2014] have yet met a large consensus. Furthermore, the impact of physiological HFOs on the identification of epileptogenic tissues is not clearly determined. Some teams have thus proposed to distinguish physiological from pathological HFOs by studying their relationship with spikes [Jacobs et al., 2008, Wang et al., 2013], with slow waves [von Ellenrieder et al., 2016a] or with spindles [Bruder et al., 2016]. Although promising, none of these approaches has so far succeeded in fully separating them.

The aim of the thesis is thus to:

1) improve the visualization of HFOs to offer an objective tool for both the visual and automatic detection (Chapter II Improving HFO visualization),

2) characterize and validate the detection procedure (Chapter III Detection Validation), and

3) quantify the reliability of each interictal marker in predicting the epileptogenic zone (Chapter IV Clinical Application).

To properly understand each step of this work, epilepsy and the presurgical evaluation will be firstly introduced. The focus will then be on the ictal and interictal recordings of the stereo-EEG. Later, the characteristic of the HFOs will be described. We will then travel in time and frequency space to explore the different means of visualizing non-stationary time-series and finally conclude this introduction with a description of the performance testing of binary classifiers and its potential pitfalls.

I.2. Epilepsy and presurgical evaluation

I.2.1. Epilepsy

Defining epilepsy is not an easy task. Even the International League Against Epilepsy (ILAE) acknowledges that little common agreement exists on the definition of the terms *seizure* and *epilepsy* [Fisher et al., 2005]. The ILAE defines an epileptic seizure as “a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain”. A person is said to have epilepsy if he or she has had “at least two unprovoked (or reflex) seizures occurring more than 24 h apart” [Fisher et al., 2014]. Despite 50 million people worldwide having epilepsy (WHO, 2017), this disease is poorly known by the general public. In fact, epilepsy has routinely been referred to as a family of disorders rather than a disease because of its wide spectrum of symptoms and conditions [Fisher et al., 2014]. This diversity blurs, for both the global opinion and patients, what epilepsy really is.

The first report of epilepsy dates back to 2000 B.C. and originates from Babylon [Weber, 2005]. In 400 B.C., Hippocrates wrote that the *Sacred Disease*, as it was referred to by the Greek, was nowise more divine nor more sacred than other diseases, already pointing out the role of the brain in causing epileptic seizures [Adams, 1868]. In a recent survey by the Fédération Française pour la Recherche sur l’Épilepsie (2016), only 29% of the respondents knew that epilepsy is a brain disease; 50% thought it was related to the nerves and 18% declared having no idea. Surprisingly, 9% still believed it has supernatural causes. Most people picture epileptic seizures as sudden muscle contractions and violent shaking of limbs. This seizure – called tonico-clonic or Grand Mal – is one of the many possible seizure types. Indeed, seizures can be divided into three groups according to their onset: focal, generalized or unknown [Fisher et al., 2017]. A seizure is qualified as generalized when both hemispheres are rapidly engaged (e.g. tonico-clonic or absence), whereas it is designated as focal when it originates within a limited network. The latter seizure type can manifest itself as repeated movements such as pedaling or jerking hand movements (for seizures with motor onset), or as a feeling of fear or déjà vu vu (for seizures with non-motor onset). The unknown category serves as a placeholder when there is not enough argument to assign the seizure to one of the two previous categories. Each category can be further subdivided as described in Fig I.1. This classification is very recent and does not elicit unanimity among clinicians. “This classification potentially offers a number of advantages, but it leaves some ambiguities and needs to be rigorously tested before entering routine clinical practice” [Beghi, 2017]. Once epilepsy is diagnosed, medication can be used to help patients.

Several classes of anti-epileptic drugs exist to relieve epileptic patients from seizures. However 30% of the patients have drug-resistant epilepsy. For these patients, epilepsy surgery may be considered. This surgery aims at resecting or disconnecting a part of the brain in order for these patients to be seizure free. Beforehand, clinicians have to determine and localize the zone to be resected during the presurgical evaluation.

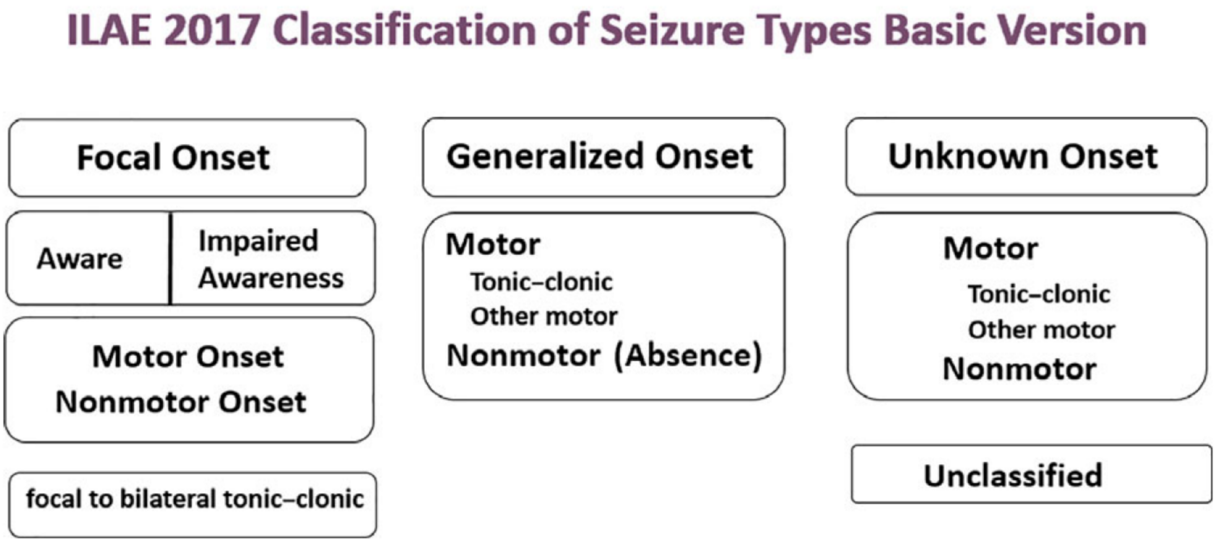


Figure I.1. – ILAE 2017 Classification of Seizure Type (taken from [Fisher et al., 2017])

I.2.2. Presurgical evaluation

During presurgical evaluation, clinicians face the difficult task of defining the epileptic zone (EZ) and its overlap with the eloquent/functional zone. The EZ is defined as “the minimum amount of cortex that must be resected (inactivated or completely disconnected) to produce seizure freedom” by the North Americans [Lüders et al., 2006]. The French and Italians define it as “the site of the beginning of the epileptic seizures and of their primary organization” [Bancaud et al., 1965, Kahane et al., 2006]. The eloquent zone is the cortical area necessary for neurological or cognitive processes (motor, sensory, visual, language cortex). This evaluation can be divided into two phases: the non-invasive and invasive phases.

During the first phase, the patient is monitored several days by video-electroencephalography (V-EEG). This procedure captures both the ictal period, i.e. the seizures, and the interictal period, i.e. the period between seizures. The EEG electrodes placed on the scalp of the patient record the electrical activity of the brain. This measure gives clinicians a rough clue on the area, i.e. a brain lobe or hemisphere, which generates the seizures or the interictal epileptiform activities. The video is mainly used to retrace the change and evolution of the patient’s behavior during the seizure (semiology). By comparing synchronously the electrical information and semiology, clinicians mentally visualize the putative area generating the seizure and its propagation.

Patients also get a magnetic resonance imaging (MRI) scan, which images the anatomy of a patient’s brain. This may be used to identify several brain abnormalities such as hippocampus sclerosis, atrophy or lesions. This first phase may further include a magnetoencephalography (MEG) recording and positron emission tomography (PET). The former records the magnetic activity of the brain similarly to the EEG except that it measures the magnetic field. The latter images the consumption of glucose in the brain and can reveal hypometabolism (lower glucose consumption) possibly caused by epileptogenic tissues.

PET scans may disclose a lesion not visible in MRI. If all information gathered by these examinations points toward the same area and spares the eloquent cortex, the patient can be operated. However, if ambiguous or discordant data are obtained, the patient is further invasively investigated.

There are two types of invasive EEG recording: electrocorticography (ECoG) and stereo-EEG (SEEG). These techniques are considered to be invasive because they imply a brain surgery to place the electrode prior to the investigation. For the ECoG, the surgeon operates a craniotomy and places a sheet with grid-patterned electrodes (grids) on the exposed surface of the cortex. For the SEEG, the surgeon drills several holes in the skull and inserts needle-like multi-lead electrodes deep inside the brain. Both techniques have the advantage of bypassing the skull and thus record the neuronal activity with a better fineness and accuracy. The placement of the electrodes is based on all available non-invasive data for a given patient and aims at targeting structures which could possibly be involved in seizure onset and/or propagation. The number of electrodes and the selection of sampled regions in each patient is necessarily restricted for both safety and technological issues. ECoG is said to be better to map functional areas because of its better coverage of the cortex [Iida and Otsubo, 2017], whereas SEEG simultaneously records from multiple sites including mesial and lateral regions.

Once the clinicians have identified the areas generating the seizure, resection can occur. In the next section, we focus on the signal recorded by SEEG, its origin and its use for the ictal and interictal periods.

I.3. Ictal and Interictal recording in Stereo-EEG

In these sections, only focal seizures are discussed since patients with generalized seizures are not implanted.

I.3.1. Origin of the signal

Several means of recording the neuronal activity exist and each have different spatial and temporal resolutions. When a neuron fires, it produces an electrical current. This current is due to ion flows in the cellular membrane. The main neuronal potentials are the fast action potentials and the post-synaptic potentials. Each of these potentials contributes to the local field potential but affects the recording differently according to the chosen technique. For instance, scalp-EEG cannot measure single neuron activity. It records the activity of large populations of synchronized cells and is preferentially sensitive to post-synaptic potentials because they last longer than action potentials (their summation yield stronger potential) [Buzsáki et al., 2012]. Moreover, the EEG signal is also influenced by the electrical conductive properties of the tissues between the neuronal sources and the scalp electrode, mainly the skull and the scalp.

To bypass their effects, Jasper and Penfield pionnered electrocorticography (ECoG) in the early 50s [Lüders et al., 2006]. This invasive technique records the electrical activity of the brain with a sheet with grid patterned electrodes (grids) on the exposed surface of the cortex. This method allows measuring with better fineness the neuronal populations. It however does not permit to record deep structures such as the amygdala and hippocampus known to be involved in various forms of epilepsy.

It is to solve this issue that depth electrodes were implanted to record deep structures. Jean Bancaud and Jean Talairach developed the stereo-electroencephalography (SEEG) in the late 50s [Bancaud et al., 1965, Cardinale et al., 2016] which is a precise implantation procedure to record both the cortex and deep structures using depth electrodes. By implanting several needle-like electrodes, they were the first to record the onset of a seizure directly from the human cortex [Cardinale et al., 2016]. Notably, the prefix ‘stereo’ does not refer to the stereotaxic frame used for placing the electrode but to its etymology: solid or volumetric and by extension three-dimensional [Cardinale et al., 2016]. While SEEG seems more invasive and dangerous than ECoG because the electrodes pass through the cortex, SEEG is associated with better safety and less major complications [Cardinale et al., 2013, Iida and Otsubo, 2017].

Both techniques were thought to be sensitive only to post-synaptic potentials like the scalp-EEG but synchronous action potentials from numerous neurons may contribute to high-frequency components [Buzsaki et al., 1992, Bragin et al., 2002, Buzsáki et al., 2012, Demont-Guignard et al., 2012].

As electrodes measure differences in electric potentials and not directly voltage (or electric tension), the traces of all these techniques can be differently visualized according to the reference channel used. The referential montage corresponds to the difference between the potential measured at a given contact and a common reference. For SEEG, this common reference can be a contact in the white matter or an external reference, such as linked ears. A variation of this montage is the average common montage. In that case, the reference used is an artificial reference obtained by calculating the average recorded potential across all channels. This is a good estimation of potential at infinity for the scalp-EEG but less intuitive for ECoG and SEEG. Finally, the montage preferred in clinical setting for SEEG is the bipolar montage. The reference is not common anymore but is the adjacent contact within the electrode. Here onwards, any traces shown are displayed in bipolar montage.

Each of the above techniques is used in chronic long-term monitoring and records ictal and interictal signals.

I.3.2. Interictal period

The interictal stage, i.e. the period between seizures, is the most recorded period. During this period, the brain undergoes various physiological states and transition phases with many different rhythmic and arrhythmic activities. Resting state, i.e. the state when the brain is at rest, doing nothing in particular, is a period which was intensively studied in functional MRI [Achard and Bullmore, 2007, Rubinov and Sporns, 2011, Achard et al., 2012, Besson et al., 2017]. Finding interictal elements which could accurately predict the EZ would decrease the hospitalization time (thus the cost) and above all improve patients' well-being since seizure would not need to be recorded anymore.

One aspect these brain states have in common is the complex structure of their background activity: the $1/f^{\alpha}$ power spectrum[1], often simply refer to as the $1/f$ power spectrum (Fig I.2). This means that the background activity has arrhythmic activity of high-amplitude low frequency components and low-amplitude high frequency components. The steeper the slope (α), the longer and stronger the correlation in time [He, 2014]. This behavior can be attributed to the low-pass frequency filtering property of dendrites or network mechanisms (more time is needed to recruit more neurons generating slow but ample signal) [Buzsáki et al., 2012]. It is however a misnomer to call the background activity a strict $1/f$-spectrum since its shape is more complex and has different slopes (Fig I.2).

Epileptiform discharges or epileptic spikes are present mainly in epileptic patients and already visible in scalp-EEG and can be measured during the interictal period. For these reasons, it is the standard marker of this period. These activities are sharp transients and are called epileptic spikes when lasting less than 70 ms and sharp waves otherwise. We will not make the difference and refer to both of them as spikes. Examples are given in Fig I.2. The cortical areas involved in producing these spikes are called the "irritative zone" [Kahane et al., 2006, McGonigal et al., 2007]. This zone usually overlaps with the EZ but is often more extended. Removing all tissues generating spikes do not necessary lead to seizure freedom but neither does leaving too much spike generating areas [Stefan

---

[1] The definition and calculation of the power spectrum are described in I.5.1 Fourier transform

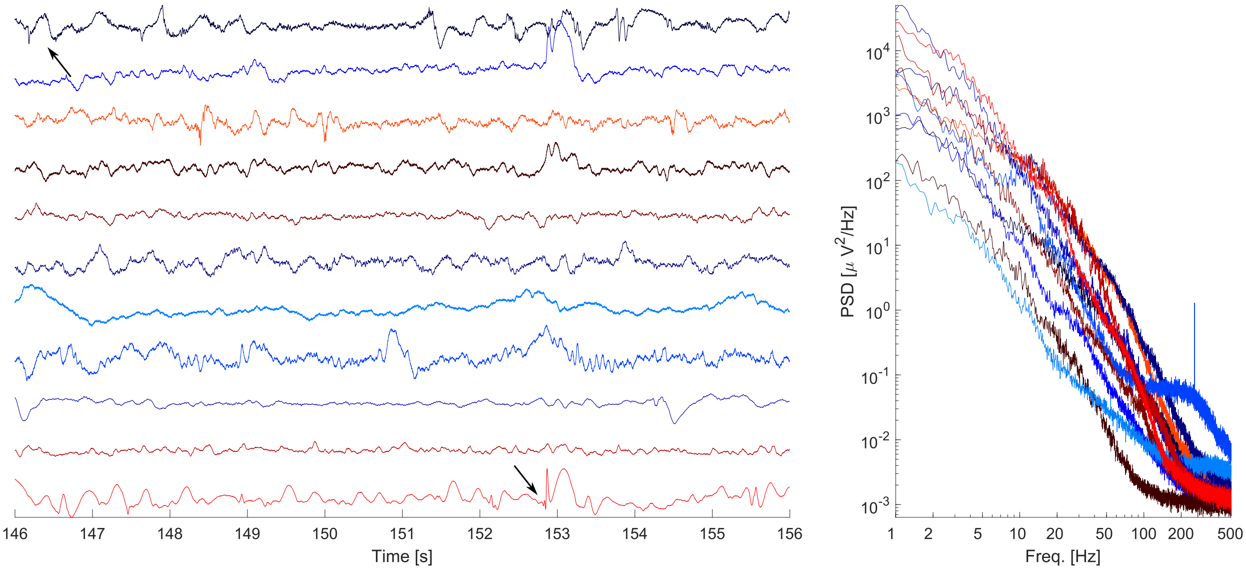


Figure I.2. – Illustration of interictal signal. The left panel shows interictal traces with some epileptic spikes (black arrow). The right panel displays the power spectrum of the respective traces.

et al., 2008, Jacobs et al., 2010]. Recently however, a new marker has concentrated the attention of this field, the high-frequency osxcillations (HFOs). This marker is extensively described in the next section (I.4).

One state of particular interest for analyzing interictal epileptic activities is sleep. One obvious reason is that the patient while asleep moves less resulting in less artifacts. Secondly, it is known that during sleep – especially during non-rapid eye movement (NREM) sleep – epileptic activities are facilitated [Rossi et al., 1984, Sammaritano et al., 1991]. This is true for both the spikes and HFOs [Sammaritano et al., 1991, Bagshaw et al., 2009, Frauscher et al., 2016]. More activities does not necessarily mean that they have better predictive values. In fact, rapid eye movement (REM) sleep exhibits less epileptic activities than NREM sleep but these activities are said to be more specific [Sammaritano et al., 1991, Frauscher et al., 2016]. Most HFOs studies have investigated the NREM sleep, which is why this stage was considered for this thesis.

I.3.3. Ictal period

The “area of cortex that initiates clinical seizures” is named the seizure-onset zone (SOZ) [Lüders et al., 2006]. Seizure recordings may be divided according to their onset patterns and the presence of preceding epileptiform discharges (Fig I.3). The most common pattern is the low-voltage fast activity (LVFA, Fig I.3 A-D). In general, focal LVFA is associated with favorable outcome whether preceded by pre-ictal spiking or not [Lee et al., 2000, Jiménez-Jiménez et al., 2015, Lagarde et al., 2016]. Other types of onsets are not necessary associated with poor outcome. The main factor is whether the onset is focal or diffuse/extended, correlating with good or bad outcome, respectively. It is to better understand how to handle extended onsets that computer-based method were designed to better quantify them [Bartolomei et al., 2008, David et al., 2011, Gnatkovsky et al., 2011, Andrzejak et al., 2015].

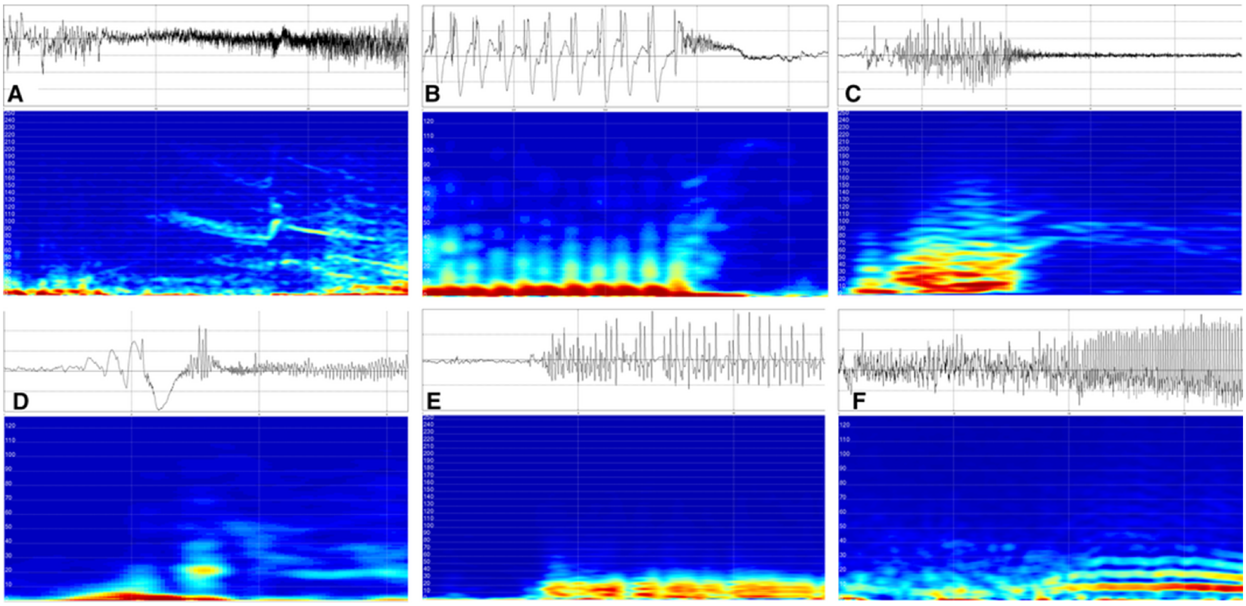


Figure I.3. – Different seizure onset patterns. (A) Low-voltage fast activity (LVFA). (B) Preictal spiking with rhythmic spikes of low frequency followed by LVFA. (C) Burst of polyspikes of high frequency and amplitude followed by LVFA. (D) Slow wave or baseline shift followed by LVFA. (E) Rhythmic spikes or spike-waves, at low frequency and with high amplitude. (F) Theta/alpha sharp activity with progressively increasing amplitude (taken from [Lagarde et al., 2016]).

Most of these methods aim at translating the visual expertise in a semi-automated algorithm and thus have similar cores based on quantifying the frequency content and the delay of involvement of each structure. This paragraph mainly describes the Epileptogenicity Index [Bartolomei et al., 2008] as it is the method used for this thesis.

The first step is to extract a feature reflecting the LVFA. The feature used here is the energy ratio (ER) between the high frequency and the low frequency calculated in a sliding window. Indeed, the LVFA is characterized by a decrease in low frequency and an increase of high frequency activities yielding an augmentation of the ER (Fig I.4 A). The ER is defined as:

$$\mathrm{ER}[n] = \frac{E_\beta[n] + E_\gamma[n]}{E_\theta[n] + E_\alpha[n]} \tag{I.1}$$

with the energy calculated in the Fourier domain[2] and the corresponding frequency bands: $\theta \in [3.5, 7.4)$ Hz, $\alpha \in [7.4, 12.4)$ Hz, $\beta \in [12.4, 24)$ Hz and $\gamma \in [24, 97)$ Hz. As ictal discharges last several seconds, a threshold based uniquely on the energy is not sufficient. The Page-Hinkley algorithm compares the chosen metric at a given time to its mean value

---

[2] The Fourier transform is defined in I.5.1 Fourier transform

calculated until the same time point (U_N, Eq. I.2, Fig I.4 B,C).

$$\mathrm{U}_N = \sum_{n=1}^{N} \mathrm{ER}[n] - \mathrm{ER}_n - \nu \tag{I.2}$$

$$= \mathrm{U}_{N-1} + \mathrm{ER}[N] - \mathrm{ER}_N - \nu, \tag{I.3}$$

$$\mathrm{ER}_n = \frac{1}{n} \sum_{k=1}^{n} \mathrm{ER}[k]. \tag{I.4}$$

The change has to be strong and last enough ($\mathrm{U}_N - \mathrm{u}_n > \lambda$) to be labeled as a significant change. u_N is the local minimum of the set of U_n:

$$\mathrm{u}_N = \underset{n \in [1,N]}{min} (\mathrm{U}_n)\,. \tag{I.5}$$

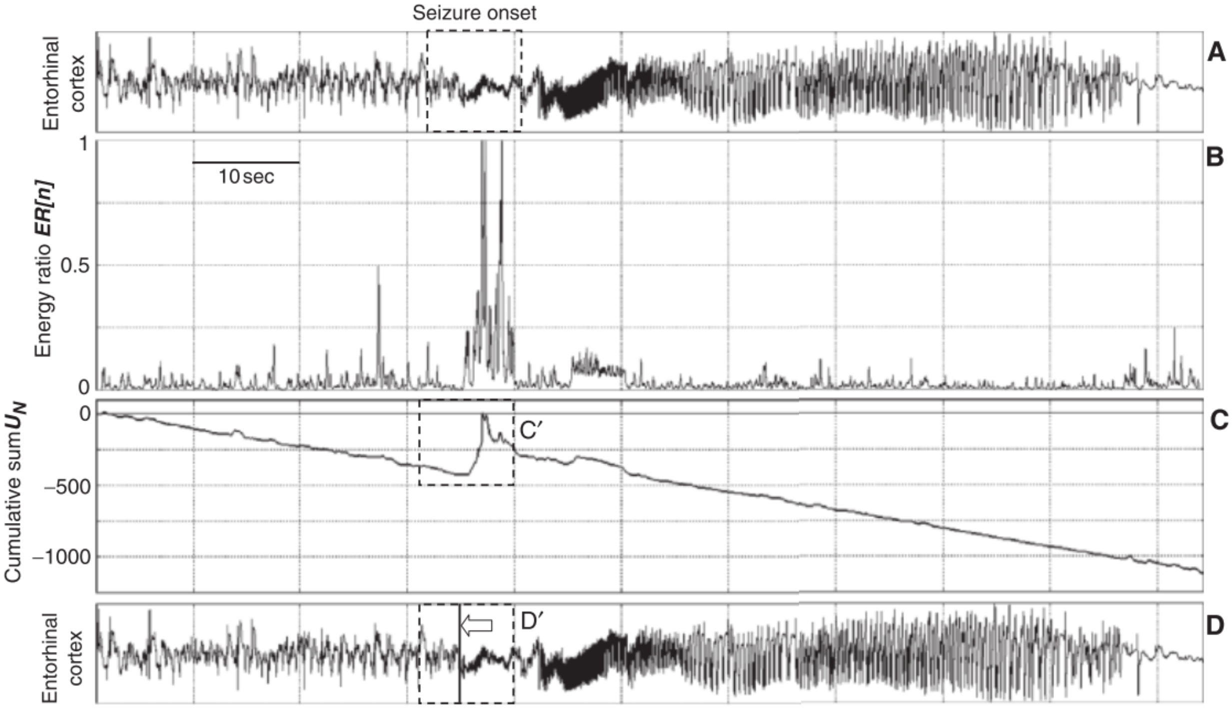


Figure I.4. – Features utilized by the Epileptogenicity Index. **A** recording of a fast discharge in the enthorinal cortex. **B** Plot of the energy ratio over time. **C** Decreasing U_N curve over time. Note the increase during the fast discharge (C'). **D** Trace of the fast discharge seen in **A** with the automated detected seizure onset (D')(adapted from [Bartolomei et al., 2008])

Importantly, the bias ν has to be high enough for the U_N curve to be decreasing (during the stationary period). When the U_N curve passes the threshold, the algorithm sets the onset of the change (N_i) as the last u_N. This is a better estimation of the onset since it corresponds to the moment when the U_N curve starts to increase, i.e. when a change occurs. By tuning the bias ν and threshold λ the clinician can fit the onset calculated by the algorithm according to his or her visual expertise. Finally, the Epileptogenicity Index (EI) takes into account both the energy ratio of the fast discharge and the delay of involvement according to the onset of the seizure. It is defined as:

$$\mathrm{EI}_i = \frac{1}{N_i - N_0 + \tau} \sum_{n=N_i}^{N_i+T} \mathrm{ER}[n], \quad \tau > 0, \tag{I.6}$$

with i a channel, N_0 the first detected onset, N_i the detected onset of channel i, τ the decay of the time-weighting and T the time of integrating of the fast discharge. The EI values are usually divided by the maximum EI of the studied seizure to obtain measures between 0 and 1. Regions with a EI value above 0.2-0.3 are labeled as epileptogenic [Bartolomei et al., 2008, Marchi et al., 2016, Pizzo et al., 2017].

The EI algorithm was shown to be relevant in quantifying the epileptogenicity of brain structures [Aubert et al., 2009, Bartolomei et al., 2017b]. For the purpose of this thesis I developed an EI plugin (Appendix B.3 Epileptogenicity Index) for the home-made open-source software AnyWave [Colombet et al., 2015].

I.4. High-frequency oscillations in epilepsy

I.4.1. What is an HFO?

High-frequency oscillations are small-amplitude oscillations between 80-500 Hz lasting more than 4 periods (Fig I.5). In humans, HFOs were first recorded using micro-electrodes measuring the hippocampus and enthorinal cortex [Bragin et al., 1999a, Bragin et al., 1999b]. In 1999, Bragin and colleagues found ripples (R, 250-500 Hz) – already described in animal studies [Buzsaki et al., 1992, Chrobak and Buzsáki, 1996] – but also HFOs between 250-500 Hz they named fast-ripples (FR). The most stimulating fact about the FRs was that there were found only under epileptic conditions. The major leap forward happened when Gotman's group showed in 2007 that both interictal Rs and FRs could be recorded using macroelectrodes [Urrestarazu et al., 2007]. This opened the way for investigating the putative clinical application of HFOs as new biomarkers of epileptogenicity. It is however unclear whether HFOs captured with microelectrodes are the same as the those detected with macroelectrodes.

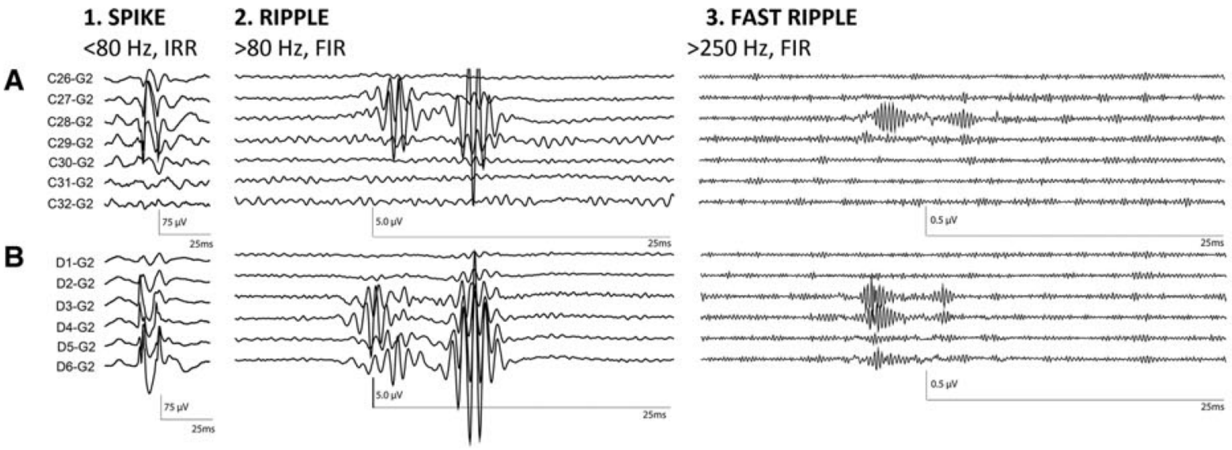


Figure I.5. – Example of HFOs taken from [Zijlmans et al., 2017]

The rise of HFOs as biomarkers of epilepsy followed several steps. Firstly, the rate of HFOs recorded with clinical electrodes was shown to be higher inside the SOZ than outside [Jacobs et al., 2008, Crépon et al., 2010], linking HFOs to the SOZ. Several teams then revealed a correlation between the surgical removal of HFOs and especially FRs generating tissues and post-surgical outcome [Wu et al., 2010, Jacobs et al., 2010, Haegelen et al., 2013, Akiyama et al., 2011]. Jacobs and colleagues [Jacobs et al., 2008] compared HFOs to the standard interictal biomarker, the epileptic spikes. They showed that adding information of HFOs improved the prediction of the SOZ. The main contrast between HFOs and spikes is that medication reduction results directly in an increase in the rate of HFO occurrence whereas it does not increase the spike rate [Zijlmans et al., 2009]. Therefore, HFOs may be a better marker of disease activity than spikes. The impact of physiological HFOs in delineating the EZ is however poorly understood and could be a potential drawback.

I.4.2. Visual Marking

Historically, the presence of HFOs was assessed through visual analysis. Gotman's team designed a procedure to reliably and reproducibly mark HFOs. The computer screen has to be split vertically with an 80 Hz high-pass filter on one side and a 250 Hz high-pass filter on the other side. Only events containing at least four consecutive oscillations are regarded as HFOs. Visual marking is however highly time consuming, taking around 10 h to visually mark HFOs in a 10-channel 10 min recording [Zelmann et al., 2012]. Using this protocol, they managed to have an inter-rater agreement of 0.7, i.e. there was a high agreement between the reviewers. However, several other studies showed lower agreement [Menendez de la Prida et al., 2015, Spring et al., 2017]. Spring and colleagues [Spring et al., 2017] reported an agreement of only 0.4 for HFOs alone and 0.57 for HFOs within spikes. Studies with visual marking of the HFOs are reported as the gold standard but such studies could be biased not only by the data available at any particular epilepsy center, but also by the tendencies of the selected reviewers [Spring et al., 2017]. Moreover, the filtering of sharp transients such as epileptic spikes can result in artefactual oscillatory activity which resembles HFOs (the so-called "false-ripple" [Bénar et al., 2010] illustrated in Fig I.6). To overcome these drawbacks, several automated detectors were designed.

I.4.3. Reviews on automated detectors

Most automated detectors usually extract features from band-pass filtered signals. Filtering in the HFO band allows increasing the signal to noise ratio, i.e how well the HFO stands out of the noise, by removing the high-amplitude low frequency components. From these signals, different features can be extracted to distinguish oscillations from background activity. These features can be energy based [Staba et al., 2002, Gardner et al., 2007, Zelmann et al., 2012, Birot et al., 2013, Jrad et al., 2016] or amplitude based [Crépon et al., 2010, Burnos et al., 2014]. Intuitively, these metrics should increase in presence of HFOs and remain relatively low during background activity. The difficult task is to estimate a threshold which optimally and robustly separates HFOs from the background in the feature space. Some have assumed that the feature was normally distributed and has set the threshold equal to Thr $= \mu + a\sigma$ with μ the mean, σ the standard deviation and a a positive integer [Staba et al., 2002, Crépon et al., 2010, Burnos et al., 2014]. Others have taken a non-parametric approach and have set the threshold as the p^{th}-percentile of the histogram of the feature. The choice of the feature as well as the thresholding strategy greatly impacts the performance of the detector and should thus be chosen carefully. Indeed, some features could increase in presence not only of HFOs but also of the high frequency components of spikes or even of non-stationarity changes in the background activity. The thresholding strategies affect the performance in the same manner. Additional criteria can be applied on the candidate HFOs to increase the performance, such as the minimum number of cycles [Staba et al., 2002] or the minimum duration [Gardner et al., 2007]

The first generation of HFO detectors suffered from high number of erroneous detections [Staba et al., 2002, Gardner et al., 2007, Crépon et al., 2010]. To handle this issue,

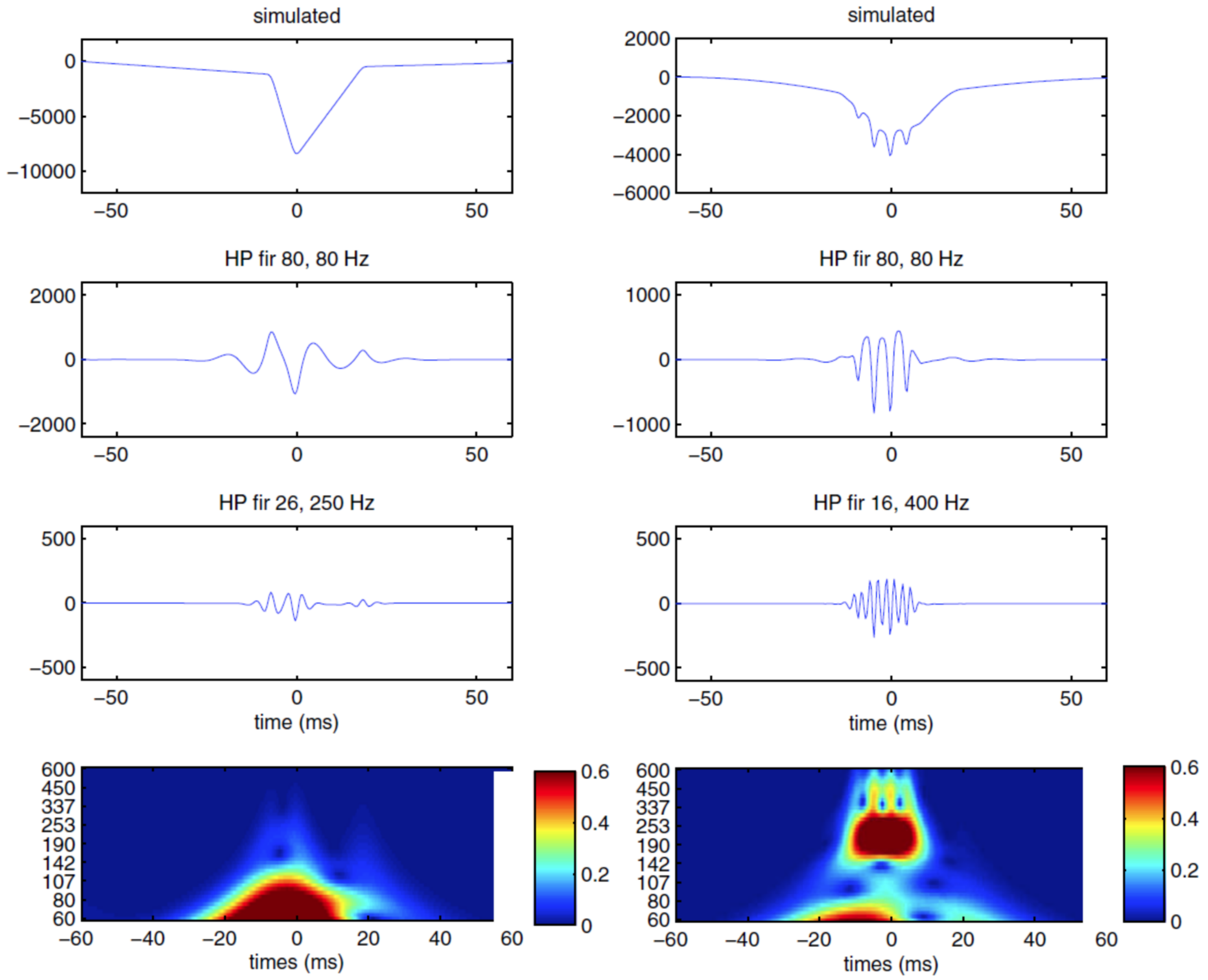


Figure I.6. – Illustration of false-ripples in time- and time-frequency domain (taken from [Bénar et al., 2010])

some teams have chosen to add other stages to previously designed detectors or to improve the thresholding method through baseline detections. For the former, Burnos and colleagues have added a time-frequency analysis [Burnos et al., 2014] to discard oscillatory activity produced by filter ringing. To be labeled as HFO, a detection has to be clearly separated from spikes in the time-frequency image [Bénar et al., 2010] (Fig I.6). For the latter, Gotman's team has thought of estimating the background level not on the whole data but on baseline sections [Zelmann et al., 2012]. This was the first attempt, to our knowledge, to handle the fact that the background level estimation based on the whole channel could be biased since the histogram of the feature has two modes, a main mode produced by the background activity and a second mode produced by all non stationary elements (e.g. spikes and HFOs). By finding a proper baseline, one can directly estimate the background level since the histogram only reflects the background activity. This was done preliminary using a baseline detector based on wavelet entropy. Fig I.7 summarizes the usual pipeline of HFO detections.

These two approaches outperform previous detectors. These techniques handle how-

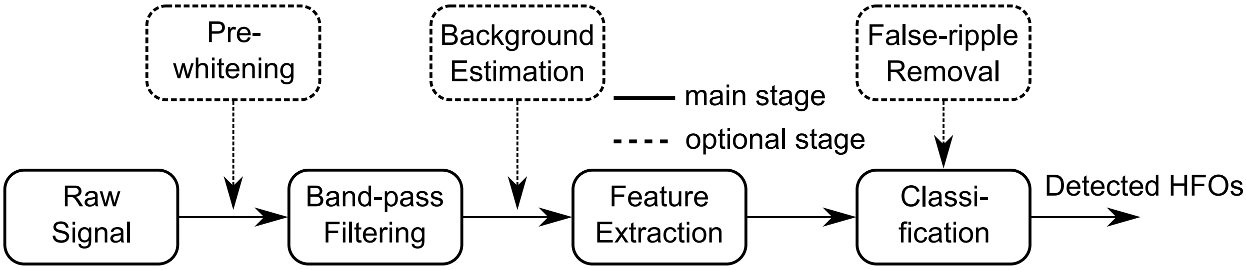


Figure I.7. – General pipeline of HFO detectors

ever different issues – threshold estimation and rejecting false detections – it thus seems crucial to combine them to reach optimal performance. Moreover, only one detector has thought of counteracting the effect of the $1/f$-spectrum to increase the sensitivity to upper frequencies of the HFO bands [Gardner et al., 2007]. These are the three aspects we discuss in chapter II.

I.5. Time-frequency analysis

I.5.1. Fourier transform

The Fourier transform is the first tool one thinks about when attempting to bridge time and frequency. Intuitively, the Fourier transform simply measures the amount of oscillations at frequency f there is in a signal; it represents any integrable or finite energy signal as a sum of sinusoidal waves e^{ift}. This translates mathematically into:

$$\tilde{s}(f) = \int_{-\infty}^{+\infty} s(t)e^{-ift}dt, \tag{I.7}$$

with s the signal of interest belonging to the space $L^1(\mathbb{R})$ of integrable functions or $L^2(\mathbb{R})$ of finite energy function. To represent the constituent frequencies of s, one computes the power spectrum density (PSD, or simply the power spectrum) as:

$$\Sigma_s(f) = |\tilde{s}(f)|^2 . \tag{I.8}$$

The Fourier transform provides answers for most questions concerning stationary signals. However, as soon as one is interested in representing transient phenomena, such as interictal spikes and HFOs, the Fourier transform is not adequate. Let us use two toy examples to illustrate the need of going beyond the Fourier transform. Let s_1 and s_2 be two signals defined as:

$$s_1(t) = \frac{1}{4}w_D\left(t - \frac{D}{2}\right) \times \sum_{i=1}^{4} \sin(2\pi f_i t), \tag{I.9}$$

and

$$s_2(t) = \sum_{i=1}^{4} w_{D/4}\left(t - (2i-1)\frac{D}{8}\right) \sin(2\pi f_i t), \tag{I.10}$$

with D the duration of the signal and w_D the Hann window function centered on zero and of length D. Fig I.8 plots the two signals for a duration of $D = 1$ s and $f_i \in$ {87, 142, 278, 382} Hz along with their power spectrum. One can see that both power spectra are similar. Indeed, s_1 and s_2, by construction, carry the same frequency information (f_1, f_2, f_3, f_4). The temporal information, i.e. when the difference waves occurs, is however different. Every oscillations of frequency f_i occurs at the same time in s_1 whereas the oscillations in s_2 arrive one after the other.

This is the main limitation of the Fourier transform. One can however solve this issue by adding a little something to the actual computation of the Fourier transform. Let us see how time meets frequency!

I.5.2. Short-time Fourier transform

The Fourier transform is applied to the whole signal to provide its frequency content. Intuitively, one could apply the Fourier transform on overlapping sections of the same duration rather than on the whole signal. In this manner, the power spectrum could be represented at each time step, and thus recover the temporal information. Computing the Fourier

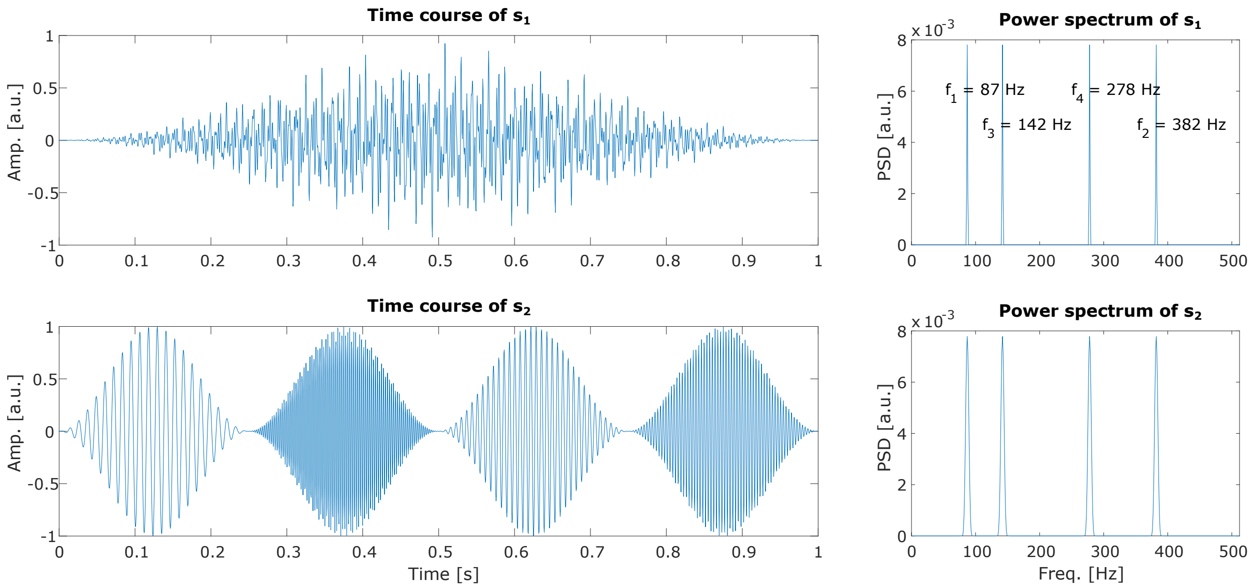


Figure I.8. – Application of the Fourier transform on two toy examples. The two signals have the same frequency content yielding the same power spectrum. The signals in the time domain are however completely different.

transform on sections is equivalent to transform sliding windowed versions of the signal. This transform is called the short-time Fourier transform (STFT). The mathematical expression of the STFT is given by:

$$S_s(\tau, f) = \int_{-\infty}^{+\infty} s(t) w_D(t-\tau) e^{-ift} dt, \tag{I.11}$$

with S_s the STFT of the signal s, w_D a window function centered on zero and of duration D and τ the time shift. Any type of window function could be considered but to avoid edge effects one should not use the rectangle window and prefer a Hann or Tukey window. For the rest of the paragraph, the Hann window will be used. Noteworthy, the Gabor transform is a STFT with a Gaussian window. The product of the STFT depends not only on the frequency but also on the time shift. Therefore, instead of plotting a power spectrum, one draws an image whose x- and y-axis correspond to the time and the frequency, respectively, and the colors code for the power at a given time and frequency. This image is named a spectrogram, and is obtained by calculating the squared modulus of S_s:

$$Sp_s(\tau, f) = |S_s(\tau, f)|^2 . \tag{I.12}$$

Along with the type of window function, one can change the length of this window. This will control the time-frequency resolution. To understand that, one should refer the Heisenberg uncertainty principle. It states that:

$$\Delta_t \Delta_f \geq c, \tag{I.13}$$

with Δ_t the uncertainty about the time, Δ_f the uncertainty about the frequency and c a constant [Carmona et al., 1998]. In other words, one cannot know, with the best precision, both the occurrence time and the frequency of an oscillation. One can know with a very good precision the frequency of an infinite sine wave but it occurring time is impossible to mark. Conversely, the time instant of a wave bursting few periods is easy to mark but its frequency is uncertain because it does not last.

The results of the STFT of s_1 and s_2 for two different window durations $D \in \{31.3, 125\}$ ms are pictured in Fig I.9. The STFT with the high frequency resolution ($D = 125$ ms) effectively pictures s_1 with four lines at the corresponding frequency. Moreover, s_2 is also correctly imaged since the "islands" corresponding to each wave appear at their correct occurring time. The STFT with the high temporal resolution ($D = 31.3$ ms) also manages to represent s_2, even though the "blobs" are wider in frequency. It however has troubles representing s_1: the two lowest blobs collide and generate interferences.

Could one say that with a well chosen window size the STFT is suitable to image interictal activities? The question is: do these activities last about the same amount of time whatever their frequency or do they burst for a similar number of oscillations at different frequencies? The STFT seems more adapted for the ictal activities where the fast discharge (in the beta or gamma band) lasts several seconds. However, interictal oscillations are bursts of several periods. Therefore their duration, i.e. their time support, highly depends on their frequency. The HFOs last a certain amount of periods, about 3 to 6 periods and are not well modeled by s_2. Let us introduce a third signal:

$$\begin{aligned} s_3(t) = \sum_{i=1}^{4} \Bigg[& w_{\frac{n}{f_i}}\left(t-(2i-1)\frac{D}{16}\right)\sin(2\pi f_i t) + \\ & w_{\frac{n}{f_i}}\left(t-(2i-1)\frac{D}{16}+\frac{D}{2}\right)\sin(2\pi f_i t)\Bigg], \end{aligned} \tag{I.14}$$

where n is the number of periods of the oscillation at frequency f_i. We again used the same four frequencies but this time these waves oscillate only during a fixed number n of periods. Therefore their duration $\frac{n}{f_i}$ depends on their frequency. Now, the previously adequate STFT ($D = 125$ ms) barely preserves the higher frequencies (f_2 and f_4). Indeed, the analyzing window is longer than the duration of the oscillation . On the other hand, the other STFT manages to image these high frequencies. Its representation of the lower frequencies is however suboptimal. Indeed the oscillations at f_1 and f_3 last longer; therefore their frequency spread is narrower than the one represented (Heisenberg uncertainty principle I.13). This is one of the limitation of the STFT. The STFT has a constant analyzing frame in both time and frequency because the window size is constant [Carmona et al., 1998, Mallat, 2008]. This means that it is not possible to properly capture oscillations of about the same amount of periods at both low and high frequencies. To represent HFOs, the analyzing frame should change according to the frequency. Let us meet the wavelet!

I.5.3. Continuous wavelet transform

In order to analyze HFOs, it is necessary to use an analyzing frame with different time supports. Instead of representing the signal as a sum of sine waves like the Fourier transform, the wavelet transform decomposes the signal over dilated and translated versions of

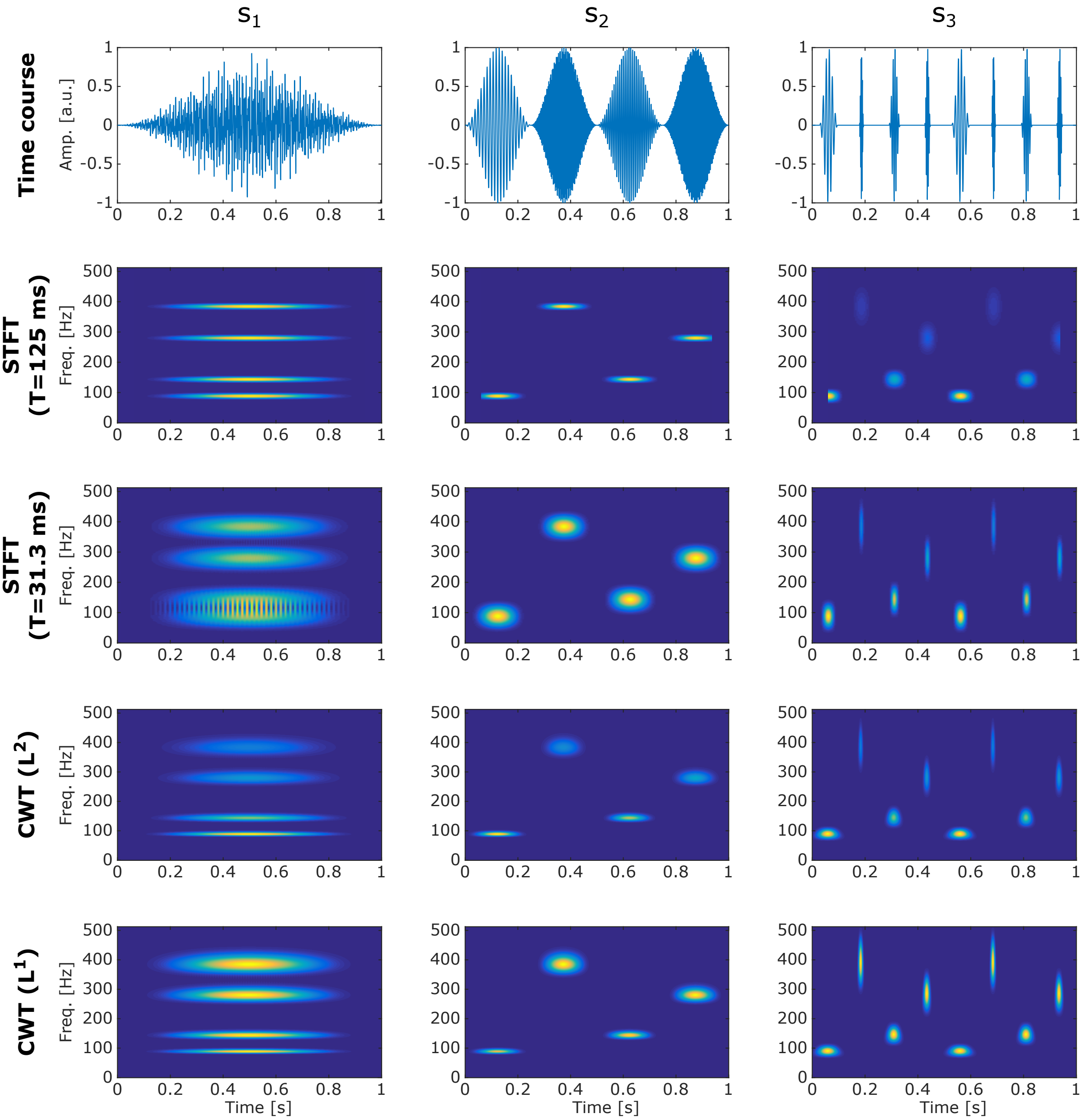


Figure I.9. – Application of the short-time Fourier and continuous wavelet transforms on three toy examples.

the same atom. This atom is called the mother wavelet and is a function $\psi \in L^2(\mathbb{R})$ with a zero average:

$$\int_{-\infty}^{+\infty} \psi\left(t\right) dt = 0. \tag{I.15}$$

The mother wavelet is normalized such as $\|\psi\| = 1$. The dictionary of wavelets is obtained by scaling, i.e. squeezing or stretching, ψ by a factor $a \in \mathbb{R}^+$ and translating it by $\tau \in \mathbb{R}$:

$$\psi_{a,\tau}\left(t\right) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right). \tag{I.16}$$

Note also that $\|\psi_{a,\tau}\| = 1$. Fig I.10 illustrates two complex dilated wavelets. Observe the opposite behavior in the time- and frequency-domain. The dilated wavelet at 35 Hz is larger in the time-domain but narrower in the frequency-domain than the wavelet at 58 Hz. This recalls the Heisenberg uncertainty principle seen in Eq. I.13.

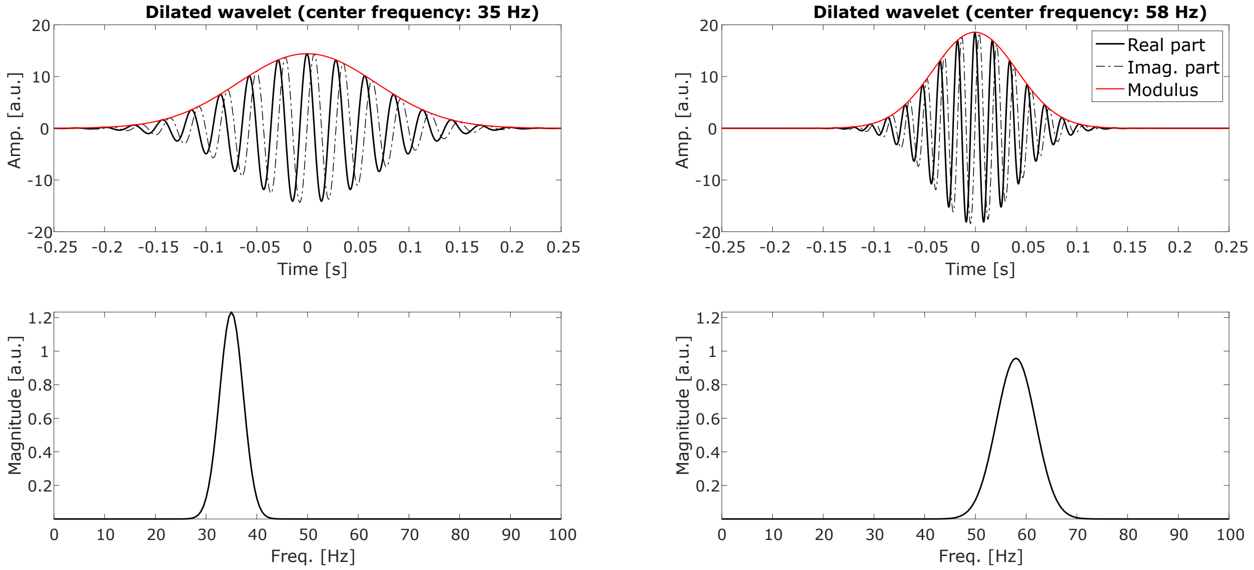


Figure I.10. – Example of two wavelets dilated by two different scaling factors. Note the dilatation in the time domain yielding a compression in the frequency domain, and inversely.

The continuous wavelet transform (CWT) of a signal $s \in L^2(\mathbb{R})$ is thus defined as

$$T_s^{L^2}(\tau, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} s(t)\overline{\psi}\left(\frac{t-\tau}{a}\right) dt, \tag{I.17}$$

with $\overline{\psi}$ the complex conjugate of ψ. Note that the high frequencies are captured by the smallest wavelets (which aim at the details) and are obtained with a small scaling factor, whereas the low frequencies are represented by the large wavelets (which aim at coarse elements) and are obtained with a high scaling factor. The CWT can be rewritten as a convolution product to highlight its similarity with linear filtering [Carmona et al., 1998, Mallat, 2008]:

$$T_s^{L^2}(\tau, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} s(t)\overline{\psi}\left(\frac{t-\tau}{a}\right) dt = \int_{-\infty}^{+\infty} s(t)\overline{\psi}_c\left(\tau - t\right) dt, \tag{I.18}$$

with

$$\overline{\psi}_c(t) = \frac{1}{\sqrt{a}}\overline{\psi}\left(\frac{-t}{a}\right) dt. \tag{I.19}$$

Finally one can express the CWT in the Fourier space as:

$$T_s^{L^2}(\tau, a) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sqrt{a}\tilde{s}(f)\overline{\tilde{\psi}}\left(af\right) df. \tag{I.20}$$

This has two main consequences. $\tilde{\psi}$ is the transfer function of a band-pass filter $\left(\text{because } \tilde{\psi}(0) = \int_{-\infty}^{+\infty} \psi(t)dt = 0\right)$ and the CWT is equivalent to applying dilated band-pass filters.

Moreover I.20 permits to use the FFT algorithm which is computationally more efficient than the convolution one. This is the algorithm which will be further used.

There also exists another type of normalization in $L^1(\mathbb{R})$, rarely used in the neuroscience field:

$$T_s^{L^1}(\tau, a) = \frac{1}{a}\int_{-\infty}^{+\infty} s(t)\overline{\psi}\left(\frac{t-\tau}{a}\right) dt. \tag{I.21}$$

Very useful for representing oscillations [Carmona et al., 1998], this normalization is however not adapted for separating HFOs from background activity. This will be discussed afterward. This normalization is in fact proportional to the S-transform [Ventosa et al., 2008] used in one automated detector [Burnos et al., 2014, Burnos et al., 2016, Fedele et al., 2016].

Real wavelets, i.e. whose value are real numbers, have important applications for detecting transients and analyzing fractal signals [Mallat, 2008]. However, in order to retrieve the phase and amplitude at different frequencies of signals, analytic wavelets have to be used. A signal $s_a \in L^2(\mathbb{R})$ is said to be analytic if its Fourier transform is zero for negative frequencies:

$$\tilde{s}_a(f) = 0 \quad \text{if} \quad f < 0. \tag{I.22}$$

Such analytic signal is necessarily complex but is fully described by its real part $s = \Re[s_a]$ [Carmona et al., 1998, Mallat, 2008] since:

$$\tilde{s}(f) = \frac{\tilde{s}_a(f) + \overline{\tilde{s}}_a(-f)}{2}. \tag{I.23}$$

In the following paragraphs, only analytic wavelets are used.

Similarly to the spectrogram, one can picture the fluctuation in time of a signal s by:

$$Sc_s(\tau, a) = |T_s(\tau, a)|^2. \tag{I.24}$$

This image is called a scalogram. For the sake of comparison with the STFT, we used the complex Morlet wavelet, which can easily be applied at the same frequencies of the STFT. Its expression in the Fourier domain is:

$$\tilde{\psi}(f) = \left(4\pi\sigma^2\right)^{\frac{1}{4}} e^{-\frac{\sigma^2}{2}(f-f_0)^2}, \tag{I.25}$$

with σ the width of the Gaussian and f_0 the center frequency. In fact, the Morlet wavelet is only approximately analytic [Lilly and Olhede, 2009] because the Gaussian envelope tends to zero towards infinity and is thus not exactly null at negative frequencies. With a high enough oscillation parameter ξ, it can however be approximated as analytic. In the following chapter, the analytic Derivative Of Gaussian (DoG) wavelet is used because of its better properties and true analytic behavior [Lilly and Olhede, 2009]. Its expression in frequency domain is:

$$\tilde{\psi}(f) = \begin{cases} f^n \exp(-f^2) & \text{for } f \geq 0 \\ 0 & \text{for } f < 0, \end{cases} \tag{I.26}$$

with n the number of vanishing moment. Notice that already for $f = 0$ the DoG wavelet equals 0 ($\psi(0) = 0$). This wavelet is thus analytic and not approximately analytic as the complex Morlet wavelet. Moreover, since by definition the analyzing frame of the wavelet changes across frequencies, it is redundant to calculate them for frequencies linearly spaced. From chapter II onwards, the time-frequency plots are shown in logarithmic scale for the frequency axis. The main consequences of this scaling are that wavelets have the same width along the frequency axis independently of the frequency and that time-frequency images have fewer lines (improving the computational cost).

The CWT(L^2) was applied to the three previous toy signals (Fig I.9). One can see that it retrieves the frequency and temporal contents of the three signals. Note the change in width from low frequency components to high frequency components. The signal s_3 is optimally represented with the lowest components being wide in time and narrow in frequency and the highest frequency being narrow in time but wide in frequency. One could be surprised that even with the optimal analyzing window, the peak of the island corresponding to the highest frequency is lower than the one of the lowest frequency even though the amplitudes in the time domain are the same. This is again due to the Heisenberg uncertainty I.13. In fact, the Parseval theorem states that:

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 \, dt = \int_{-\infty}^{+\infty} \left|\tilde{\psi}(f)\right|^2 dt, \tag{I.27}$$

i.e. the energy of a signal is independent of its domain of representation; its energy in the time domain equals its energy in the frequency domain. Intuitively, if the energy is unchanged in time and frequency domain but the time support and frequency support change, the amplitude or power has to change. For the high frequency f_2, the amplitude is 1, the time support is short but the frequency support is large therefore the maximum power has to be lower.

One can see that this natural consequence is solved by the second normalization (CWT(L^1)). This normalization seems perfect to capture oscillations. One thing which has not been illustrated yet is the representation of noisy signals.

Let us see how these normalizations handle white noise. To all the s_i signals, white noise ($\epsilon \sim \mathcal{N}(0, \sigma^2 = 0.4^2)$) was added. One can calculate the noise level at each scale/frequency. It is given by [Carmona et al., 1998]:

$$\sigma^2_{L^2}(a) = \mathbf{E}\left\{\left|T_\epsilon^{L^2}(\tau, a)\right|^2\right\} = \sigma^2 \left\|\psi\right\|, \tag{I.28}$$

$$\sigma^2_{L^1}(a) = \mathbf{E}\left\{\left|T_\epsilon^{L^1}(\tau, a)\right|^2\right\} = \frac{\sigma^2}{a} \left\|\psi\right\|. \tag{I.29}$$

$\mathbf{E}$ is the expectation value. For the normalization in L^2, the noise level $\sigma^2_{L^2}$ is constant across scales and proportional to the noise level in the time domain σ. The noise level in the L^1-normalized CWT decreases when the scale increases, i.e. it increases with the frequency. As illustrated in Fig I.11, the peak power relative to the oscillation at f_2 is high but so is the noise. Because we aim at finding a strategy to better represent and potentially detect the HFOs, time-frequency images should have constant noise level across

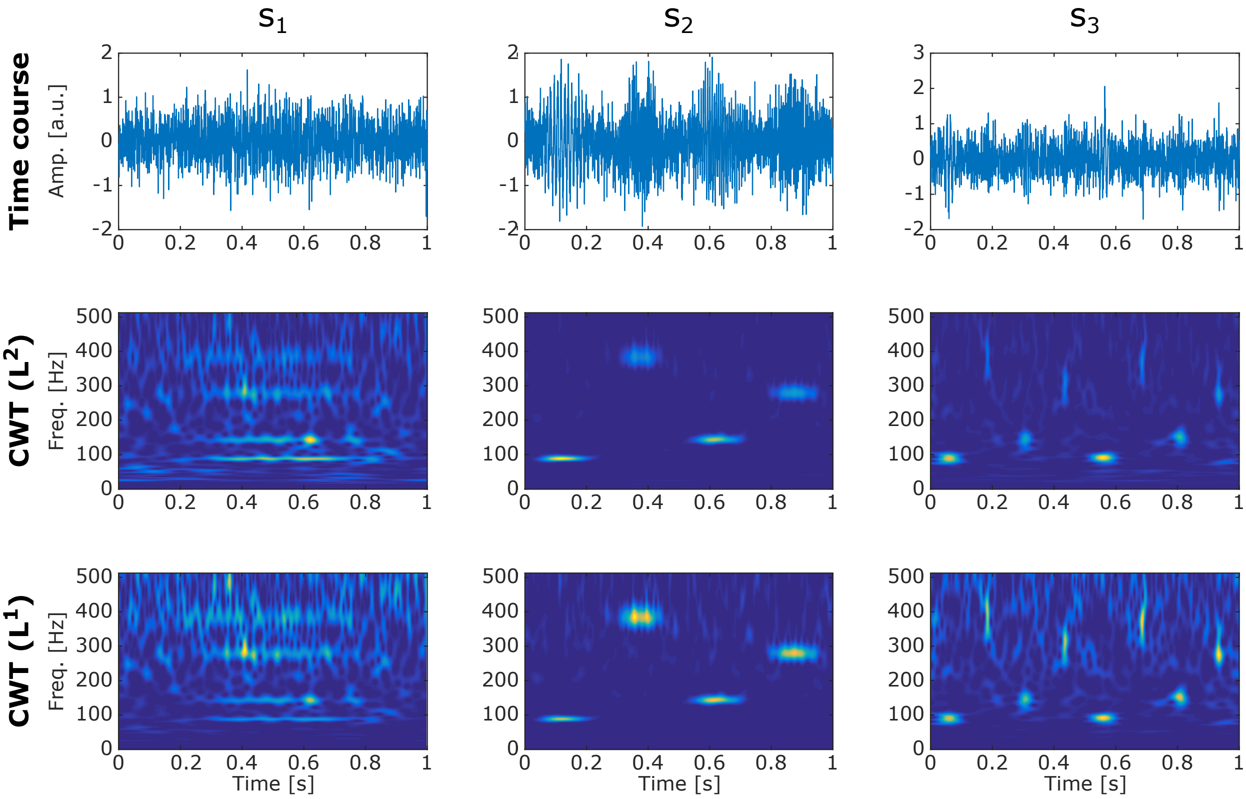


Figure I.11. – Application of the continuous wavelet transforms on three noisy toy examples.

scales. Any attempt to adjust the noise level would correct the L^1-normalization anyhow. We will therefore use the L^2-normalization.

Once a strategy is found, its performance in separating the HFOs from the background activity should be measured. Since there are two classes (the HFOs and the background activity), one could evaluate the strategy in the binary classification paradigm.

I.6. Binary Classification

The value of a diagnostic/classification test lies in its ability to detect the elements of interest (its sensitivity), e.g. to detect patients with a disease, interictal spikes, high-frequency oscillations, epileptogenic zone, and to exclude the other elements (its specificity), e.g. to exclude patients without disease, background activity, non-epileptogenic areas [McNeil et al., 1975]. Importantly, this presumes that the true/real status of the elements has been established by other means than the current test, i.e. identified by a gold standard [Linnet, 1988]. Let us start with tests with binary outcomes.

I.6.1. Tests with binary outcomes

A test with binary outcomes simply outputs whether an element of a set belongs to one group or another. Table I.1 shows a confusion matrix which resumes the results of a test. True positives (TP) are the positive elements that are predicted as positive. False positives (FP) are the negative elements that are predicted as positive. True negatives (TN) are the negative elements that are predicted as negative. False negatives (FN) are the positive elements that are predicted as negative.

		True conditions	
		P	N
Test outcome	T_P	TP	FP
	T_N	FN	TN

T_P: the test positive outcome
T_N: the test negative outcome

P: the positive instances/conditions
N: the negative instances/conditions
TP: the true positive
FP: the false positive
TN: the true negative
FN: the false negative

Table I.1. – Confusion matrix of a binary test

From these four fundamental measures, one can derive several metrics, such as the sensitivity (Sens), specificity (Spe), false positive rate (FPR) and precision (Prec):

$$\mathrm{Sens} = \frac{\mathrm{TP}}{\mathrm{P}} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{I.30}$$

$$\mathrm{Spe} = \frac{\mathrm{TN}}{\mathrm{N}} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}, \tag{I.31}$$

$$\mathrm{FPR} = 1 - \mathrm{Spe} = \frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}, \tag{I.32}$$

$$\mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{T_P}} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}. \tag{I.33}$$

Sensitivity is the proportion of the positive instances which are labeled as positive, e.g. the proportion of patients with the tested disease revealed by the test or the proportion of HFOs detected by the algorithm.
Specificity is the proportion of the negative instances which are labeled as negative, e.g. the proportion of individuals without the tested disease excluded by the test or the proportion of non-epileptogenic tissues classified as not epileptogenic.
False positive rate is the proportion of the negative instances which are labeled as positive,

e.g. the proportion of individuals without the tested disease revealed by the test or the proportion of non-epileptogenic tissues classified as epileptogenic.
Finally, precision is the proportion of correct positive outcome, e.g. the proportion of individuals revealed by the test which have the tested disease or the proportion of detected HFOs which are true HFOs.

These measures are the most commonly used to characterize biomarkers and detectors [Staba et al., 2002, Jacobs et al., 2008, Blanco et al., 2010, van't Klooster et al., 2011, Zelmann et al., 2012, Birot et al., 2013, Burnos et al., 2014, Nissen et al., 2016a, Jrad et al., 2016, Fedele et al., 2016]. One important aspect of these measures is that they are always paired. A test cannot be characterized by only one metric. Imagine a test being positive for every individual. Its sensitivity is one but its specificity is zero. Assessing the performance of the test only by its sensitivity does not estimate the error of misclassifying the negative instances (FP), and conversely. One has to use at least three components of the confusion matrix. The main pairs are *sensitivity and specificity*, and *precision and recall* (recall being a synonym of sensitivity). The best test has a sensitivity and specificity of one which is equivalent to having a sensitivity and precision of one.

Before applying a test, one should always verify that each of the four sub-classes of the confusion matrix exists. This may play an important role in the design of the performance testing. For instance, one may want to diagnose a given disease in a group of individuals. The TPs are the individuals with a disease revealed by the test. The FNs are the individuals with a disease excluded by the test. The FP are the individuals without a disease revealed by the test. The TN are the individuals without a disease excluded by the test. In this example each sub-class is perfectly defined. Now, one wants to quantify the performance of an HFO detector. Here, the TPs are the correctly detected HFOs. The FNs are the missed HFOs. The FPs are the wrongly detected HFOs. And the TNs are... What are they? Every single sample which does not belong to an HFOs? (That is a lot of points!) Every section in between HFOs? (These are long sections!) Does that mean that it is not possible to quantify HFO detectors? Of course not. Zelmann and colleagues [Zelmann et al., 2012] cleverly proposed to discretize their dataset into sections of HFOs and baselines. A TN is thus a section of baseline. Another possibility is to avoid handling TNs. Instead of calculating the specificity, one could calculate the precision which does not take into account the TNs. This is the solution we used in chapter II and III. The precision has properties that could possibly be more interesting in assessing performance of HFOs detectors. To illustrate them, let us continue with tests with outcomes of continuous values.

I.6.2. Tests with continuous value outcomes

Tests with outcomes of continuous values output values of a certain range. To classify the elements of the dataset into two groups, one applies a threshold to these values and labels elements with a value below this threshold as predicted negatives and elements with a value above this threshold as predicted positives. After this stage one falls back into the previous binary outcome case and the aforementioned metrics can be calculated. Since this threshold is not trivial, one has to study the behavior of the test for different thresholds. One can represent Sens as a function of FPR for each threshold. This curve

is named *Receiver Operating Characteristic* (ROC) curve. Similarly, the *Precision and Recall* (PR) curve draws Prec as a function of Sens for different thresholds. From these curves, the optimal thresholds can be derived as well as a global measure of performance.
The *Area Under the Curve* (AUC), which literally calculates the area below the ROC or PR curves, gives the global performance of a classifier. The best classifier has an AUC of one since a threshold which perfectly separates the dataset into the two real groups exists. The closer to one the AUC is, the better the classifier/test is.

As for any optimization procedure, the optimal threshold is a trade-off between the acceptable loss and required gain of the test. Two optimization procedures are illustrated below. As discussed above, the best test has a sensitivity and specificity of one. Because the ROC curve uses the FPR on the x-axis, this point is the upper left corner ($P = (0, 1)$). One optimal threshold for a given curve can be the one which gives the point closest to the upper left corner. In other words, it is the one which minimizes the distance between the curves and the upper corner:

$$\begin{aligned}\text{Thr}_{\text{Opt. 1}} = \underset{\text{Thr}\in\mathbb{R}}{argmin}(d(P,M)) \Leftrightarrow \text{Thr}_{\text{Opt. 1}} &= \underset{\text{Thr}\in\mathbb{R}}{argmin}\left(\sqrt{(1-\text{Sens})^2 + (0-\text{FPR})^2}\right) \\ \Leftrightarrow \text{Thr}_{\text{Opt. 1}} &= \underset{\text{Thr}\in\mathbb{R}}{argmin}\left(\sqrt{\left(\frac{\text{FN}}{\text{P}}\right)^2 + \left(\frac{\text{FP}}{\text{N}}\right)^2}\right), \end{aligned} \tag{I.34}$$

with Thr the threshold, $d(.,.)$ the euclidean distance, P the point at the upper left corner and M a point on the ROC curve. By rewriting the geometrically driven expression, one can interpret this optimal threshold as the threshold which minimizes the error rates. This optimal threshold is the one usually used [Zelmann et al., 2012].

A second technique aims at finding the optimal threshold by minimizing a cost function [Linnet, 1988]:

$$\Delta L = \frac{1}{\text{P}+\text{N}}\left[\Delta L_{\text{FP}}\,\text{FP} + \Delta L_{\text{FN}}\,\text{FN}\right], \tag{I.35}$$

with ΔL the cost function. ΔL_{FP} and ΔL_{FN} are defined as:

$$\Delta L_{\text{FP}} = L_{\text{FP}} - L_{\text{TN}}, \tag{I.36}$$

$$\Delta L_{\text{FN}} = L_{\text{FN}} - L_{\text{TP}}, \tag{I.37}$$

with L_{FP} the cost of misclassifying a negative, L_{TN} the cost of correctly classifying a negative, and conversely for L_{FN} and L_{TP}. The optimal threshold for this cost function is obtained when the slope of the ROC curve equals [Linnet, 1988]:

$$m = \left.\frac{d\,\text{Sens}}{d\,\text{FPR}}\right|_{\text{Thr}=\text{Thr}_{\text{Opt. 2}}} = \frac{\Delta L_{\text{FP}}}{\Delta L_{\text{FN}}}\cdot\frac{N}{P}. \tag{I.38}$$

Properly estimating the slope of ROC curve may be difficult for non-parametric curves (staircase-like curves). This optimal threshold can be obtained by minimizing the following expression [Zweig and Campbell, 1993]:

$$\text{Thr}_{\text{Opt. 2}} = \underset{\text{Thr}\in\mathbb{R}}{argmin}\left(\text{FPR} - \frac{\text{Sens}}{m}\right). \tag{I.39}$$

For the following examples, the cost of the FP and FN are the same $\left(\Delta L_{\text{FP}} = \Delta L_{\text{FN}}\right.$ and $\left.m = \frac{\text{N}}{\text{P}}\right)$.

Let us now study four cases:

- case 1 - a random classifier or feature set whose values for the Ns and Ps are randomly drawn from the same Gaussian distribution $(\lambda_{\text{P}} \sim \mathcal{N}(6.5, 2)$ and $\lambda_{\text{N}} \sim \mathcal{N}(6.5, 2))$ and a dataset with the same amount of P and N;
- case 2 - a classifier/features whose values for the Ns and Ps are randomly drawn from two overlapping Gaussian distributions $(\lambda_{\text{P}} \sim \mathcal{N}(8, 2); \lambda_{\text{N}} \sim \mathcal{N}(6.5, 2))$ and a dataset with the same amount of P and N;
- case 3 - a classifier or feature set whose values for the Ns and Ps are randomly drawn from two overlapping Gaussian distributions $(\lambda_{\text{P}} \sim \mathcal{N}(8, 2); \lambda_{\text{N}} \sim \mathcal{N}(6.5, 2))$ but a dataset with ratio of $\frac{\text{P}}{\text{N}} = \frac{1}{10}$;
- case 4 - a classifier or feature set whose values are randomly drawn from two overlapping Gaussian distributions for forty percent of Ns and all the Ps $(\lambda_{\text{P}} \sim \mathcal{N}(8, 2);$ $\lambda_{\text{N}_1} \sim \mathcal{N}(6.5, 2))$, the sixty remaining percent of the Ns are drawn from a uniform distribution $(\lambda_{\text{N}_0} \sim \mathcal{U}(0, 1))$ and a dataset with ratio of $\frac{\text{P}}{\text{N}} = \frac{1}{10}$.

The histograms as well as the ROC and PR curves of each case are plotted in Fig I.12. Each case was realized 250 times to calculate the average and the standard deviation of the sensitivity and precision at a given specificity and sensitivity, respectively (vertical average [Fawcett, 2006]).

For case 1, the average curves of the ROC and PR spaces are straight lines. The area under both curves equals 0.5. Both optimal points are at $(0.5, 0.5)$ in both spaces. Notice that the variance of the precision increases when the sensitivity decreases. Intuitively, the sensitivity as well as the number of detections decrease when the threshold increases; the precision is therefore calculated on fewer elements, increasing its variance.

For case 2, the area under both curves equals 0.77 and both optimal points are at $(0.3, 0.7)$ in ROC space and at $(0.7, 0.7)$ in PR space. The classifier has a higher AUC and a better operating point than the random classifier. For cases 1 and 2, the PR and ROC curves are pretty similar. For cases 3 and 4 it is a different matter.

The classifier of case 2 was used for case 3 but the amount of N and P instances is different $\left(\frac{\text{P}}{\text{N}} = \frac{1}{10}\right)$. This example illustrates the main difference between the ROC and PR curves, and the two optimal thresholds. ROC curves as well as optimal thresholds based on the left corner distance are insensitive to the class skew, i.e. the ratio of Ps and Ns [Fawcett, 2006]. The ROC curves as well as the optimal operating point 1 are similar

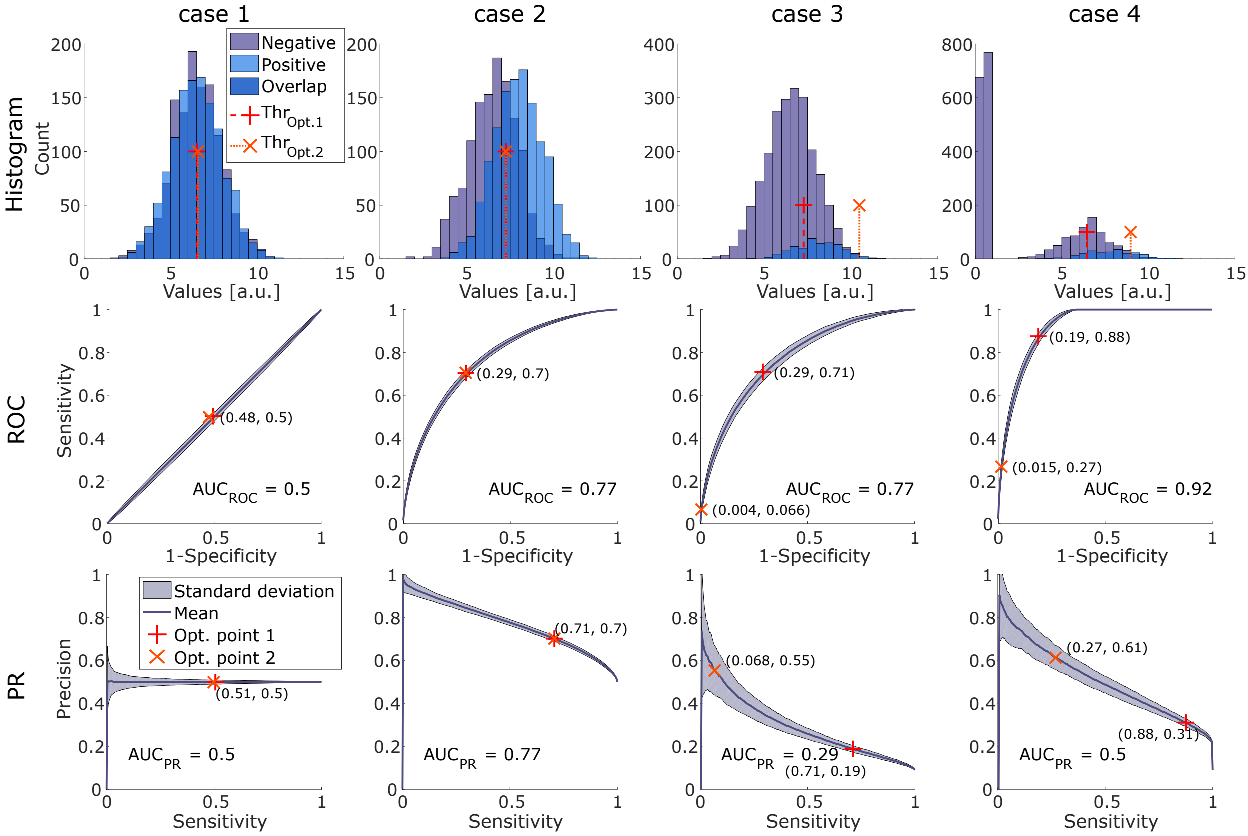


Figure I.12. – Illustration of the behavior of the ROC and PR curves for different classifiers and samplings

in case 2 and 3. From case 2 to 3, the sensitivity and specificity of this operating point are the same but its precision changes drastically. The precision drops from 0.70 to 0.20. In other words, in case 2 three out of ten predicted positives were wrongly classified but in case 3 it is seven out of ten predicted positives. How can an operating point, which minimizes the error rate and thus have a specificity of 70%, have such a low precision?

It may not seem obvious but there is a matter of misunderstanding with the specificity metric. Indeed, there is a dissonance between the meaning of the word and the meaning of the metric. The *word specificity* signifies the quality of belonging uniquely to a particular subject, whereas the *metric specificity* is the capacity to exclude elements of non-interest. In other words, *the specificity is the sensitivity of the negative instances*! When, in [Jacobs et al., 2008], it is written that the high specificity allows "identifying channels that are almost surely in the SOZ", it is not entirely true. Indeed, high specificity is obtained by lowering the false positive but it is obtained by lowering the false positive *compared to the negative instances*! It is the adverb "surely" which is not necessarily correct. High specificity means that the test properly rejected most of the negative instances. It does not assess how certain it is that a predicted positive is an actual positive (this is the precision). The histogram for case 3 shows that the first threshold minimizes the error rate but tolerates a lot of FPs compared to the TPs. The second threshold however, which minimizes the cost function, removes a lot of FPs but less TPs remain. The question is not which curves, from the ROC or the PR, is better but what the actual goal of the classifier is.

For instance, a first-stage HFO detector should be optimized to minimize the error rate because the second-stage handles the high FP. It seems unsuitable to optimize a first-stage HFO detector according to the ratio of Ps and Ns ($\mathrm{Thr}_{\mathrm{Opt.}\ 2}$). This would most likely reduce the sensitivity which would again be reduced by the second-stage. On the other hand, the precision of HFO detectors, as a whole (several stage detector or one block detector), should be optimized because the number of negative instances is usually higher $\left(\frac{7994}{51061} \approx 0.16\right.$ in [Zelmann et al., 2012]$\left.\right)$. It seems inefficient to have a detector with 99% of sensitivity and 99% of specificity but only have 20% of precision, i.e. which is wrong 80% of the time, simply because the ratio was $\frac{\mathrm{P}}{\mathrm{N}} = 0.25\%$...

The sensitivity of the PR curves to the class skew begs the question of the AUC of the random case. Is it always 0.5 like the AUC of ROC curves? Naturally, the probability of correctly classifying a P is, for the random case, the probability of finding a P $\left(\frac{\mathrm{P}}{\mathrm{P+N}}\right)$ and this probability is constant whatever the threshold. Therefore the random AUC depends on the class skew of the dataset and equals $\frac{\mathrm{P}}{\mathrm{P+N}}$. For case 1, it equals $1/2$ and for case 3, it equals $1/11$.

The final case is similar to case 3 but with the Ns having a bi-modal distribution. Notice that the ratio is the same as case 3; only the distribution of the negatives changes. In that case, the ROC and PR curves improved and the AUCs increased. Previously, the ROC curves were said to be insensitive to the class skew [Fawcett, 2006]. This claim has to be nuanced. It may be the case for classifiers trained in machine learning but in neurophysiology, Ns may behave less predictively. Let us assume two different datasets used to tests HFO detectors. One dataset contains a lot of sections with HFO-like activities (the negatives), e.g. spikes producing ringing artefacts, oscillations at lower frequencies which somehow affects the detection metric, and less HFO sections (the positives), e.g. section with HFOs alone or riding spikes. This is similar to case 3. The second dataset has the same number of Ns and Ps but whose Ns also have some baseline segments. These baseline sections may have little to no influence on the detection metric. The second dataset corresponds to case 4. The issue here is that the $\mathrm{AUC}_{\mathrm{ROC}}$ of a detector is high simply because it was tested on the second dataset. This highlights the need of comparing detectors on the same dataset (III).

Let us now assume two SEEG implantations with the same number of electrodes. The first one records the EZ and IZ. Allegedly, the spike and HFO rates are higher in the EZ than in the IZ and one can draw a ROC curve to predict the EZ based on these rates (case 3). The second implantation, however, records not only the EZ and IZ but also some physiological areas. If the detector is well tuned, the rates should be null in the physiological area (unless it records physiological HFO generators). The second implantation is similar to case 4. The performance of the features in predicting the EZ is completely different. This again highlights the need of testing features on the same dataset and building datasets that reflect reality as much as possible.

With the previous examples, comparisons of features or classifiers was partially introduced. Let us now focus on how to compare two features/classifiers and the relative metrics used.

I.6.3. Comparing classifiers or features

To compare classifiers, one has to extract characteristics from the ROC or PR curves. The AUC is generally used to analyze the curves. In the following paragraph, examples are shown to illustrate to what extent the AUC is a reliable metric for comparing two classifiers. Let us study four cases:

case 1 - one classifier whose outputs follow: $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,1)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,1)$, a second classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,2)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,2)$ and a dataset with the same amount of Ps and Ns;

case 2 - one classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,1)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,1)$, a second classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,2)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,2)$ and a dataset with a ratio of $\frac{\mathrm{P}}{\mathrm{N}} = \frac{1}{10}$;

case 3 - one classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,1)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,1)$, a second classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,2)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,1)$ and a dataset with the same amount of Ps and Ns;

case 4 - one classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,1)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}(6.5,1)$, a second classifier with $\lambda_{\mathrm{P}} \sim \mathcal{N}(8,2)$ and $\lambda_{\mathrm{N}} \sim \mathcal{N}\left(6.5,\frac{1}{4}\right)$ and a dataset with a ratio of $\frac{\mathrm{P}}{\mathrm{N}} = \frac{1}{10}$.

The histograms as well as the ROC and PR curves of each case are plotted in Fig I.13. Each case was again realized 250 times to calculate the average and the standard deviation of the sensitivity and precision at a given specificity and sensitivity, respectively (vertical average [Fawcett, 2006]).

Case 1 illustrates two classifiers with different levels of overlap between the Ns and the Ps. The overlap is larger for the second classifier. It seems logical that its performance is weakened by this greater overlap. Indeed, its relative curve is always below the curves of the first classifier, in both ROC and PR spaces. Consequently, the AUC of the first classifier is higher than the one of the second. For this case, the first classifier is the best.

Case 2 allows studying the impact of the class skew. The overlaps of the two classifiers are the same as in case 1; only the fraction of Ps is reduced. From the previous section, it is expected that the ROC curves and therefore the $\mathrm{AUC}_{\mathrm{ROC}}$ remain the same but that the PR curves as well as the $\mathrm{AUC}_{\mathrm{PR}}$ have to drop. No surprise here. Interestingly, despite the drop, the first classifier still outperforms the second. In fact, for a fixed number of positive and negative examples, one curve dominates a second curve in ROC space *if and only if* the first dominates the second in PR space [Davis and Goadrich, 2006]. In other words, if a curve is always above, i.e. dominates, a second curve in ROC space, then it is always above in the PR space, and conversely. Things get a bit more complicated when the curves cross.

Case 3 is similar to case 1 with the difference that the Ns of the second classifiers are less spread out. Intuitively, this makes the sensitivity drop rapidly since all the Ns

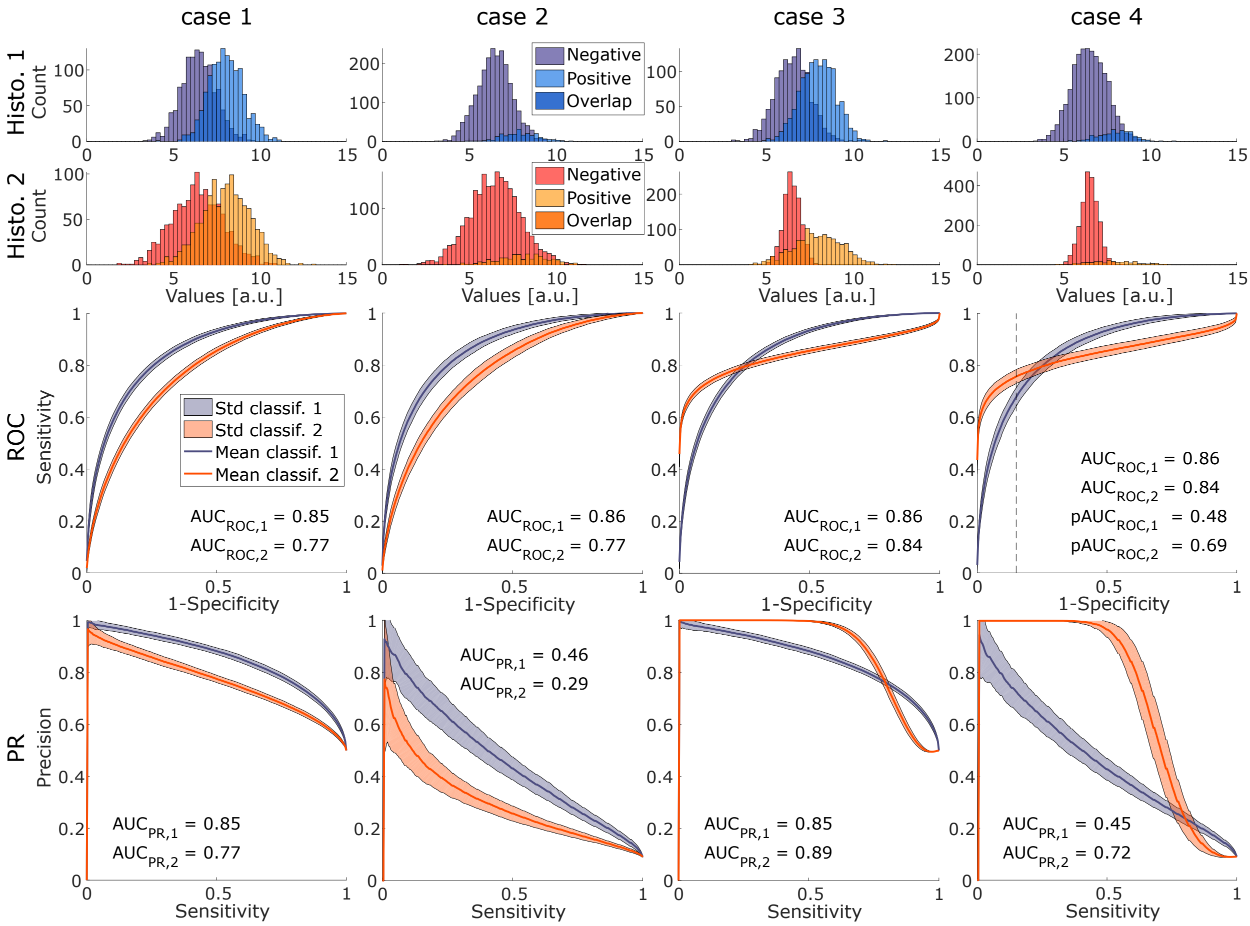


Figure I.13. – Illustration of methods to rank different classifiers

overlap with the Ps. After a certain value of the threshold, however, there are few Ns and still lots of Ps. The sensitivity thus stays relatively high for high specificity. In both the ROC and PR spaces, no curve dominates. Regrettably, the AUCs are contradictory ($\mathrm{AUC}_{\mathrm{ROC},1} > \mathrm{AUC}_{\mathrm{ROC},2}$ but $\mathrm{AUC}_{\mathrm{PR},1} < \mathrm{AUC}_{\mathrm{PR},2}$). Case 4 illustrates this issue even more. This time $\mathrm{AUC}_{\mathrm{PR},2}$ is much higher than $\mathrm{AUC}_{\mathrm{PR},1}$ due to the change in ratio of Ps and Ns. This is the main limitation of the AUC; it cannot discriminate classifiers whose curves cross each other without any other criterion. Here again, the final users has to decide what is more important: limit the FP or limit the error rate?

A solution proposed to solve this issue is to compute the AUC not for the whole ROC space but for a section of interest (partial AUC) [Ma et al., 2015]. For instance, if the aim is to minimize the FP with an acceptable decrease in sensitivity, one can calculate the AUC only for high specificity. The partial AUCs were calculated for specificity above $\mathrm{Spe}_{\mathrm{min}} = 0.85$. Since the area is not calculated for the whole range of specificity, the partial AUC cannot reach 1. The value is thus normalized by the maximum value possible to range from 0 to 1 as follows:

$$\mathrm{pAUC}_{\mathrm{ROC}} = \frac{\mathrm{AUC}_{\mathrm{ROC}}^{\mathrm{Spe} \geq \mathrm{Spe}_{\mathrm{min}}}}{1 - \mathrm{Spe}_{\mathrm{min}}}. \tag{I.40}$$

Notice that the partial AUC for a random classifier now equals:

$$\mathrm{pAUC}_{\mathrm{ROC},\epsilon} = \frac{\frac{1}{2}\cdot(1-\mathrm{Spe}_{\mathrm{min}})^2}{1-\mathrm{Spe}_{\mathrm{min}}} = \frac{1-\mathrm{Spe}_{\mathrm{min}}}{2} = 0.075. \tag{I.41}$$

One could possibly calculate a partial AUC for the PR curves but that supposes to limit the sensitivity, which is a counter-intuitive task.

Now the partial AUC of the second classifier is higher than the one of the first classifier $\left(\mathrm{pAUC}_{\mathrm{ROC},2} > \mathrm{pAUC}_{\mathrm{ROC},1}\right)$. One can say that for high specificity values, the second classifier outperforms the first classifier. This is the method we use in chapter IV when comparing the different biomarkers (e.g. HFOs, FRs, spikes). This compromise seems adequate since we wish to determine which biomarker better describes the EZ while preserving the functional areas, i.e. which biomarker has globally the highest sensitivity for high specificity.

II. Improving HFO visualization

Roehri N. *et al.*, IEEE Transaction on Biomedical Engineering 2016

Preamble

Following the work on false-ripples [Bénar et al., 2010], filtering is prone to ringing artifacts and time-frequency (TF) analysis seems adequate to separate ripples riding spikes from spurious oscillations. TF analysis has however other drawbacks that need to be handled. As illustrated in Fig II.1, high amplitude spikes and the $1/f$ spectrum have a huge impact on this representation. Because of their high power, they raise the upper limit of the color map and hinder the visibility of the HFOs. In an automatic procedure, this would not be a problem but since we aim at finding an optimal representation for both visual and automated detection, this must be addressed.

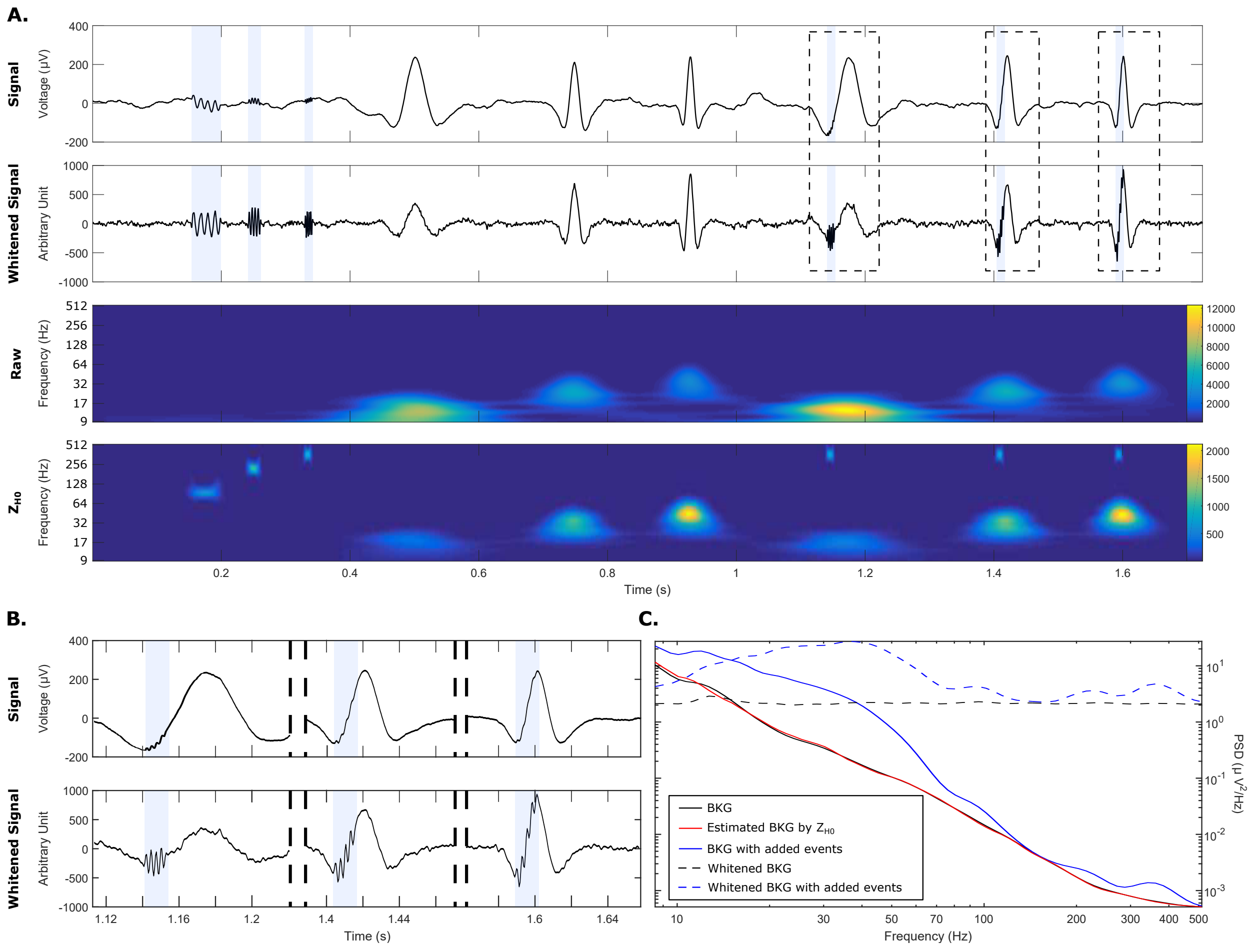


Figure II.1. – Illustration of the purpose of whitening

This chapter describes the solution that we have chosen, i.e. normalizing the TF image based on the background activity. Please note that normalizing the TF plot on the background activity is similar to transforming the background activity (with its $1/f$ spectrum) into a white noise (a flat spectrum). The last row of Fig II.1 shows how this normalization highlights the HFOs while preserving the entire frequency content of the signal (which would be lost or drastically decreased with filtering)[1].

The cartoon below (Fig II.2) describes how whitening can help automatic detection as well. One single threshold is needed to separate the emerging activities from the background activity. Indeed, after normalization, the background is flat, i.e. is a white noise. A single threshold controls the amount of false detections produced by the background across frequency and this threshold has the same meaning at each frequency.

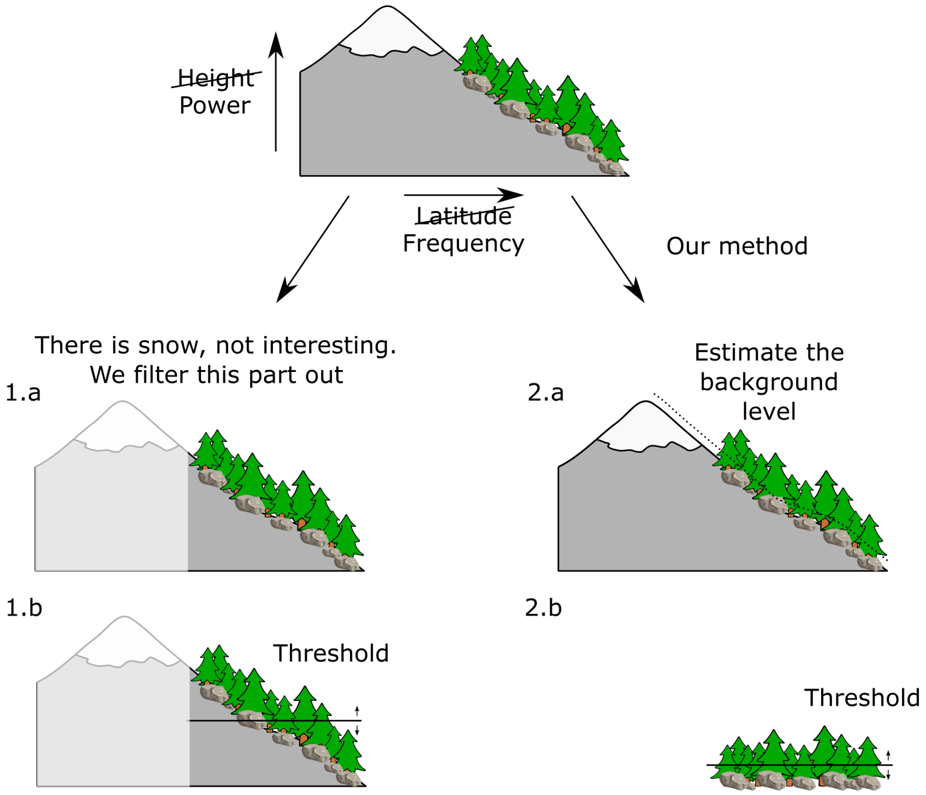


Figure II.2. – Whitening in cartoon. 1. illustrate the method used in the literature, i.e. 1.a filtering low frequencies and 1.b applying a unique threshold. 2. illustrate out method using a whitening stage, i.e. 2.a estimating the background activity and 1.b applying a unique threshold after compensating the background activity.

---

[1] Importantly, one could change the scale of the colormap or cut out lower frequency of the TF plane to enhance a certain frequency range. For the former, changing the scale to highlight the HFOs would yield saturation of the lower frequency components and therefore a loss of information. For instance, changing the scale to better visualize the FR could saturate the ripple band and obscure their visibility. Moreover, the choice of a scale is subjective because the energy values depends on the signal properties whereas, in a normalized and whitened TF, the energy values have the same statistical meaning across frequency. For the latter, events can leak into the HFO band and look like HFOs, e.g. an oscillation at the upper gamma band or high frequency components of a spike.

Time-frequency strategies for increasing high frequency oscillation detectability in intracerebral EEG

Nicolas Roehri[1], Jean-Marc Lina[2], John C. Mosher[3], Fabrice Bartolomei[1,4] and
Christian-George Bénar[1]

[1] Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France
[2] Department of Electrical Engineering, École de technologie supérieure 1100, Notre-Dame Street West Montréal, Québec, Canada H3C 1K3
[3] Epilepsy Center, Cleveland Clinic Neurological Institute, Cleveland, Ohio, 44195
[4] APHM, Timone hospital, Clinical Neurophysiology, Marseille, France

***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 8methods 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 "H_0 z-score") are more precise than the others. *Significance*: The H_0 z-score provides an optimal framework for representing and detecting HFOs, independent of a baseline and a priori frequency bands.**

II.1. 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, 80-500 Hz) which have been shown to be an indicator of epileptogenicity [Urrestarazu et al., 2007, Jacobs et al., 2008, Jacobs et al., 2009]. 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 ($1/f^{\alpha}$ spectrum). No automated detectors [Staba et al., 2002, Gardner et al., 2007, Birot et al., 2013, Burnos et al., 2014] 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 plan; however, the $1/f^{\alpha}$ 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.2. Methods

II.2.1. Continuous Wavelet Transform

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

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

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

$$\psi_{a,b}(t) = 1/\sqrt{a}\psi((t-b)/a), \tag{II.2}$$

in order for the estimate of the power spectrum – the wavelet spectral function [Carmona et al., 1998] – to correspond to

$$V_B^{L^2}(a) = \frac{1}{B}\int_{-B/2}^{+B/2} |T_f(b,a)|^2 \, db, \tag{II.3}$$

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 II.A). We therefore used the normalization in L^2 defined in (II.2).

We utilized an analytic Derivative of Gaussian (DoG) wavelet (a specific case of the Morse wavelet) with $n = 20$. It is expressed in the frequency domain as

$$\tilde{\psi}(f) = \begin{cases} f^n \exp(-f^2) & \text{for } f \geq 0 \\ 0 & \text{for } f < 0. \end{cases} \tag{II.4}$$

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 [Lilly and Olhede, 2009], unlike the classical Morlet wavelet. Such analytic wavelets could provide a better strategy for analyzing HFOs [Worrell et al., 2012].

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 68 Hz) for a better overview of the data.

II.2.2. Whitening Strategies

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 [Gardner et al., 2007]. Let $\tilde{x}$ be the prewhitened signal

$$\tilde{x}[n] = x[n] - x[n-1]. \tag{II.5}$$

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

$$\mathrm{DTFT}\left\{x[n]-x[n-1]\right\}=\left(1-e^{-j\omega}\right)X\left(e^{j\omega}\right). \tag{II.6}$$

Thus its power spectral density (PSD) is

$$\begin{aligned}\left|\left(1-e^{-j\omega}\right)X\left(e^{j\omega}\right)\right|^2&=\left|1-e^{-j\omega}\right|^2\left|X\left(e^{j\omega}\right)\right|^2\\&=\left|2j\sin\left(\frac{\omega}{2}\right)\right|^2\left|X\left(e^{j\omega}\right)\right|^2\\&=2\left(1-\cos\left(\omega\right)\right)\left|X\left(e^{j\omega}\right)\right|^2,\end{aligned} \tag{II.7}$$

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) [Ljung, 1987] 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

$$\mathrm{DTFT}\left\{\left(x*AR(p)\right)[n]\right\}=\left(1-\sum_{k=1}^{p}a_k e^{-kj\omega}\right)X\left(e^{j\omega}\right), \tag{II.8}$$

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

$$\mathrm{DTFT}\left\{\left(x*ARIMA(p,d,0)\right)[n]\right\}=\left(1-\sum_{k=1}^{p}a_k e^{-kj\omega}\right)\left(1-e^{-j\omega}\right)^d X\left(e^{j\omega}\right). \tag{II.9}$$

It is noteworthy that the ARIMA(0,1,0) prewhitening is equivalent to the Diff method (II.6). As discussed in [Ljung, 1987](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.

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 Energy Operator (TKEO) is mostly used to identify the instantaneous frequency and the amplitude of non-stationary signals [Kamwa et al., 2011, Kamath, 2013]. The TKEO applied to a continuous signal $x(t) = A\sin(\Omega t + \phi_0)$ corresponds to

$$\begin{aligned} \mathrm{TKEO}\{x(t)\} &= \dot{x}^2(t) - x(t)\ddot{x}(t) \\ &= A^2\Omega^2\cos^2(\Omega t + \phi_0) \\ &\qquad + A^2\Omega^2\sin^2(\Omega t + \phi_0) \\ &= A^2\Omega^2, \end{aligned} \tag{II.10}$$

where " $\dot{}$ " and " $\ddot{}$ " 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:

$$\mathrm{TKEO}\{x[n]\} = x^2[n] - x[n-1]x[n+1]. \tag{II.11}$$

Note that the TKEO cannot be used as a prewhitening method, since its output is already an energy. Indeed TKEO normalization (used in [Zelmann et al., 2014]) operates on the complex coefficients of the TF image *of the SOI* over time and for each frequency taken separately,

$$\begin{aligned} T_f^{\mathrm{TKEO}}[n,m] = T_f[n,m]&\overline{T_f[n,m]} \\ &- \frac{1}{2}T_f[n-1,m]\overline{T_f[n+1,m]} \\ &- \frac{1}{2}\overline{T_f[n-1,m]}T_f[n+1,m], \end{aligned} \tag{II.12}$$

where " $\overline{\cdot}$ " denotes the complex conjugate.

The event-related spectral perturbation (ERSP) is a common way to normalize the TF maps in electrophysiological reviewing [Grandchamp and Delorme, 2011]. 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

$$ERSP_f[n,m] = |T_f[n,m]|^2 / \mu[m]. \tag{II.13}$$

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 $\mu_{type}[m]$ and the SD $\sigma_{type}[m]$ 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 $\mathrm{Z}_{\mathrm{baseline}}$ and $\mathrm{Z}_{\mathrm{SOI}}$, and the TF transformation is

$$Z_f^{type}[n,m] = \frac{|T_f[n,m]|^2 - \mu_{type}[m]}{\sigma_{type}[m]}. \tag{II.14}$$

In fact, $\mu[m]$ is equivalent to the wavelet spectral function V_B defined in (II.3). ERSP and the $\mathrm{Z_{baseline}}$ 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 $\mathrm{Z_{baseline}}$ 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 [Krieg et al., 2011]. In HFO studies, finding a baseline for each channel can be difficult. This could be solved by the $\mathrm{Z_{SOI}}$ 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 – "$\mathrm{H_0}$ z-score" ($\mathrm{Z_{H_0}}$) – 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_G per scale is made of a zero-mean Gaussian distributed noise $\mathrm{H_0}$ and our signal of interest $\mathrm{H_1}$; 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 SD. 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:

```
for $a = 1^{st}$ scale to last scale do
    $T_f^{re} \leftarrow \Re\{T_f(:,a)\}$
    $T_f^{im} \leftarrow \Im\{T_f(:,a)\}$
    calculate the first ($\mathrm{Q_1}$) and third ($\mathrm{Q_3}$) quartile and the interquartile range (IQR) of $T_f^{re}$ and $T_f^{im}$
    $(\mu^{re},\sigma^{re}) \leftarrow$ Gaussian fit on $\mathrm{Q}_1^{re} - 1.5\mathrm{IQR}^{re} \leq T_f^{re} \leq \mathrm{Q}_3^{re} + 1.5\mathrm{IQR}^{re}$
    $(\mu^{im},\sigma^{im}) \leftarrow$ Gaussian fit on $\mathrm{Q}_1^{im} - 1.5\mathrm{IQR}^{im} \leq T_f^{im} \leq \mathrm{Q}_3^{im} + 1.5\mathrm{IQR}^{im}$
    $Z_{H_0,f}^{re} \leftarrow \left[T_f^{re} - \mu^{re}\right]/\sigma^{re}$
    $Z_{H_0,f}^{im} \leftarrow \left[T_f^{im} - \mu^{im}\right]/\sigma^{im}$
    $Z_{H_0,f}(:,a) = Z_{H_0,f}^{re} + iZ_{H_0,f}^{im}$
end for
```

One characteristic of the $\mathrm{Z_{H_0}}$ 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 II.3).

The methods are illustrated in Fig II.4; their effect on the power spectrum is represented in Fig II.5.

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 $\mathrm{Z_{H_0}}$, thanks to the properties of the analytic DoG wavelet. Given that $f(.)$ is a real-valued signal and the wavelet ψ is analytic, a reconstruction formula [Mallat, 2008] is proportional to

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

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 analysis-reconstruction scheme* [Carmona et al., 1998] into

$$\begin{aligned} f(t) &\propto \Re \int_0^{+\infty} T_f(t,a) \frac{da}{a\sqrt{a}}, \\ &\propto \int_0^{+\infty} \Re\left\{T_f(t,a)\right\} \frac{da}{a\sqrt{a}}, \\ &\propto \int_0^{+\infty} \sigma(a) \Re\left\{Z_{H_0,f}(t,a)\right\} \frac{da}{a\sqrt{a}}. \end{aligned} \tag{II.16}$$

We thus define the Z_{H_0} whitened signal as

$$\tilde{f}(t) = \int_0^{+\infty} \Re\left\{Z_{H_0,f}(t,a)\right\} \frac{da}{a\sqrt{a}}. \tag{II.17}$$

At the computational level, since we choose a log-scale representation and neglect the

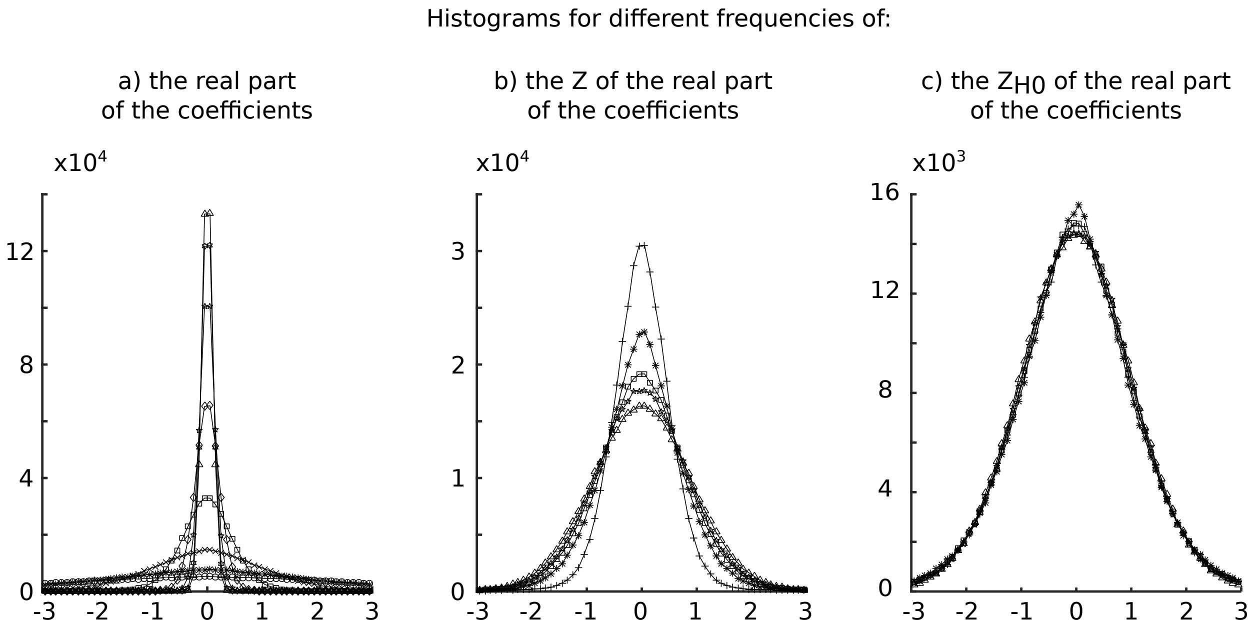


Figure II.3. – Distributions of the real part of the TF coefficients at different scales for different normalizations. a) shows distributions without normalization where large differences in width can be observed across frequencies. b) represents the distributions z-scored with μ and σ estimated on the whole distribution. The variability is reduced but there is still a difference in width. c) displays the distributions normalized with Z_{H_0}. In that case distributions are similar across frequencies. The same behavior is observed with the imaginary part (data not shown).

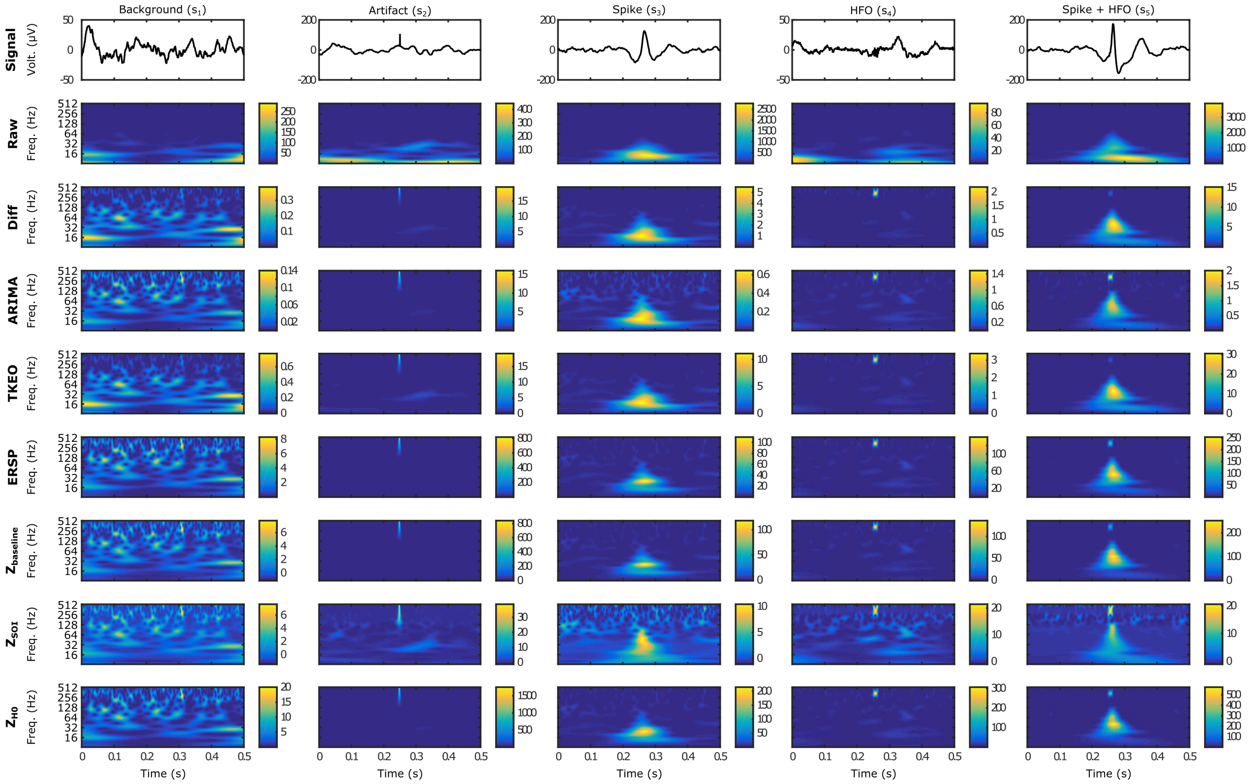


Figure II.4. – Examples of the 5 types of simulated events in time and time-frequency domains. The events are represented in a 500 ms-windows. The columns correspond to an event type, the first line shows the time-series, and the other lines illustrate the result of each method in the TF map. The HFOs 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.

multiplicative factor, the whitened signal can be computed as

$$\tilde{f}[n] \approx \sum_{m=1}^{\text{Oct}\times\text{Voi}} \frac{1}{\sqrt{a[m]}} \Re\left\{Z_{H_0,f}[n,m]\right\}. \tag{II.18}$$

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 [Mallat, 2008].

II.2.3. 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

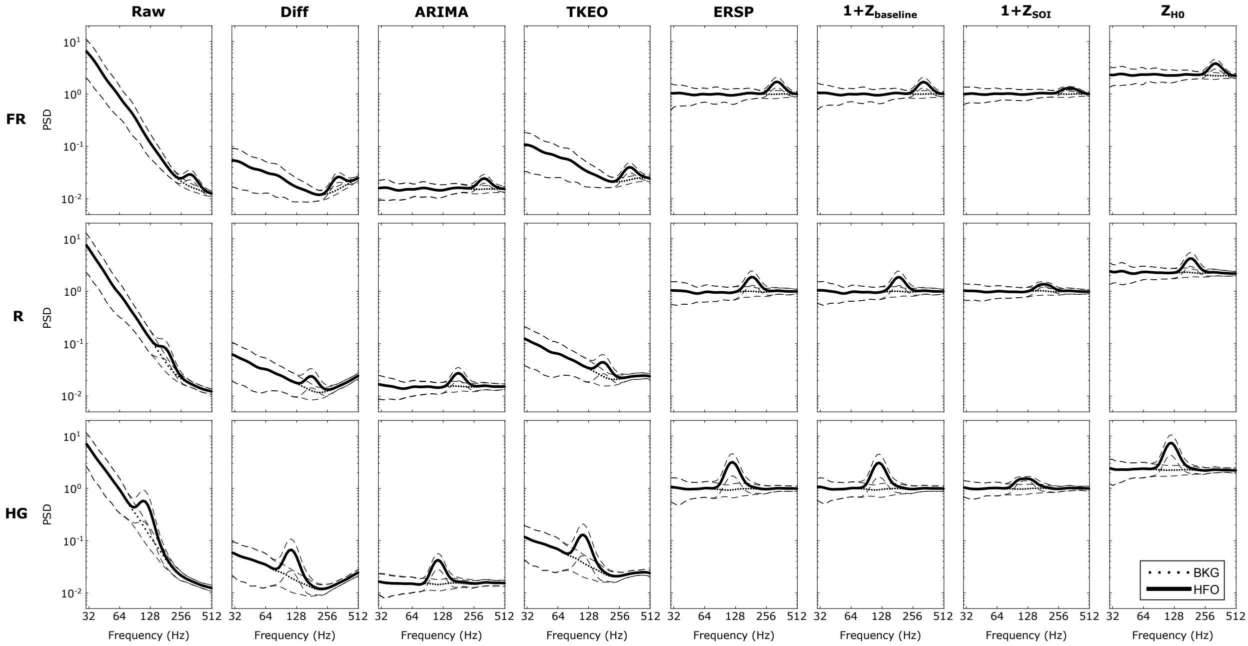


Figure II.5. – The effect of the several methods on the spectrum. 100 simulations of BKG with HFOs (SNR = 10 dB) were generated. The average estimate power spectrum V_B of the BKG with HFO and the BKG are represented with solid and dotted line respectively. The dashed line corresponds to the standard deviation. The first column shows the spectrum of the non-whitened signal (Raw) and the other columns correspond to the different methods. 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 $1/f^{\alpha_i}$, with $\alpha_1 \approx 3$ and $\alpha_2 \approx 1$ for the frequency range of 10-200 Hz and 250-500 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.

(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 (SD) 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 artifact was simply generated by increasing a single point by a certain level (impulse function). This level corresponds to five times the SD of the chosen BKG. Examples of the five events are shown in Fig II.4.

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 [Bagshaw et al., 2009]. For standardization, the collection of BKGs were normalized by dividing each signal by its own standard deviation and multiplied by the median SD of the real data collection. Unlike previous studies [Jmail et al., 2011, Birot et al., 2013], the SDs 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.

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 II.6

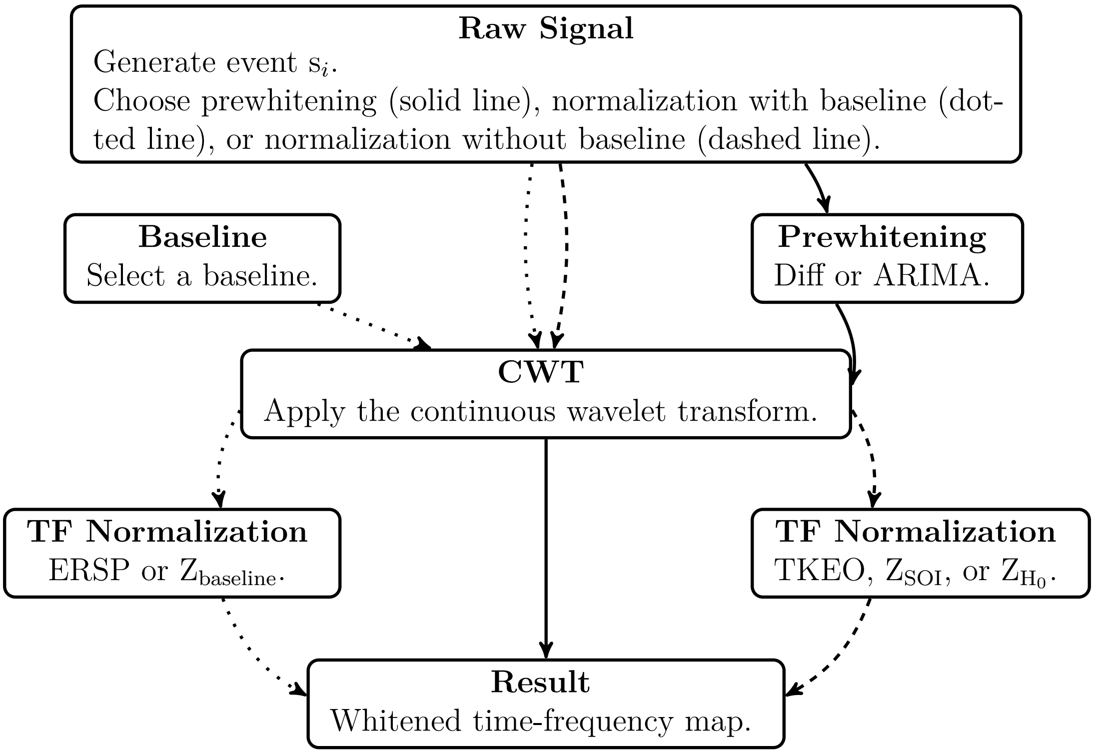


Figure II.6. – Schematic representation of the pipeline. The boxes correspond to a stage and the arrows to a path. In the first step, we generate an event s_i. This event is either processed through the prewhitening pipeline (solid line), the normalization pipeline with baseline (dotted line), or the normalization pipeline without baseline (dashed line). The prewhitening is done in the time-domain before the CWT, and the normalization is done after the CWT. This results in a whitened TF map.

II.2.4. 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 [Gardner et al., 2007] and verified visually [Zelmann et al., 2009] and were recorded from the same patients as in the simulation part II.2.3.

II.2.5. 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" [Bénar et al., 2010] 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 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 σ_t corresponds to the theoretical width of the wavelet at the corresponding frequency. σ_f was heuristically set to 30 Hz as being the accepted error on the frequency of the oscillation. FPs 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 not above the threshold, plus those of all s_4 and s_5 under the threshold which are not in the confidence zone. FNs 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

$$TPR = TP/P = TP/(TP + FN), \tag{II.19}$$

$$FPR = FP/N = FP/(FP + TN), \tag{II.20}$$

$$PPV = TP/(TP + FP). \tag{II.21}$$

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 SD of the oscillation was obtained by decomposing the signal using the Empirical Mode Decomposition (EMD) [Huang et al., 1998] 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 SD 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.

II.3. Results and Discussion

Fig II.5 represents the average spectrum of the signals s_1 and s_4 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 $1/f^{\alpha_i}$, 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 [Davis and Goadrich, 2006, Fawcett, 2006]. In a clinical setting, it seems interesting to address the proportion of TP within all detections (PPV (II.21)) regardless of N, i.e. be more precise (high PPV (II.21)) than specific (high FPR (II.20)). PR curves are therefore preferred for further analysis.

Box and whisker plots of the AUC of the 8methods applied to the simulated data are represented in Fig II.7 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 one. This is not the case for larger SNRs. The ERSP and z-score methods have

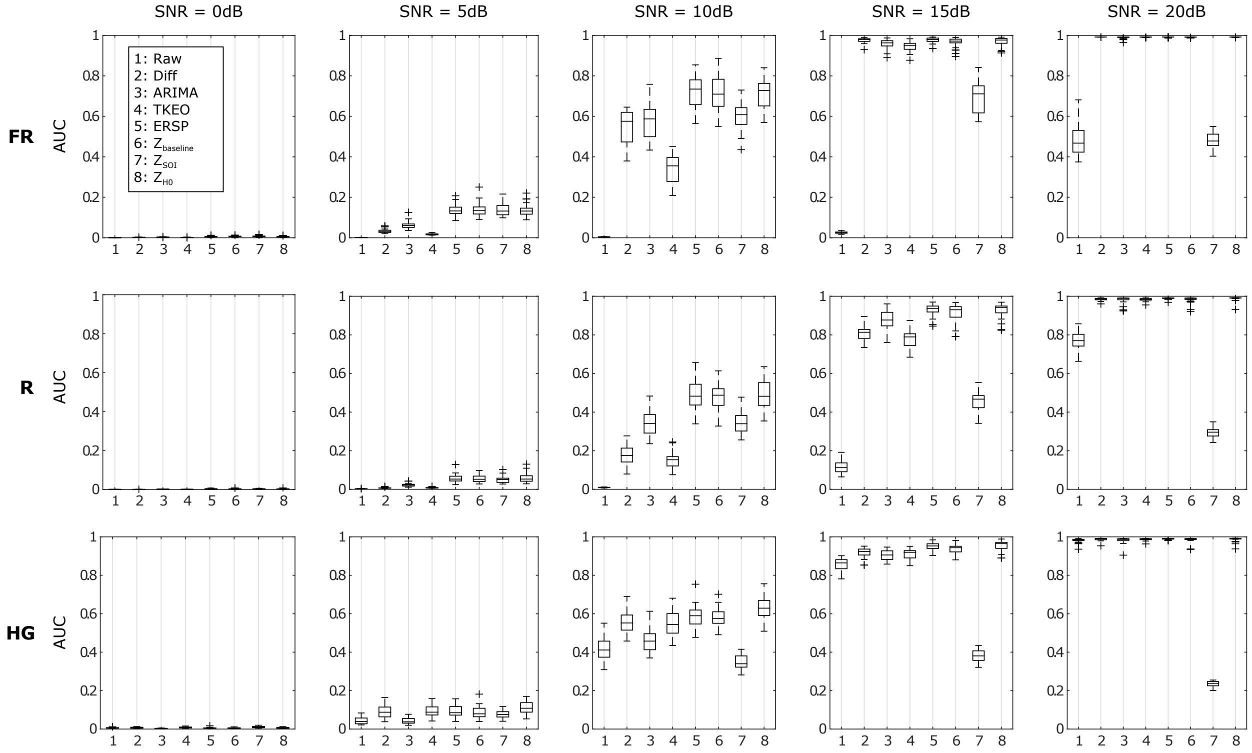


Figure II.7. – Box and whisker plot of the AUC of PR curves of the original signal and the 8techniques for several SNRs. Generally, all methods show better results when the SNR increases except the Z_{SOI}. All methods exhibit poor performances for SNRs below 5dB without having one method being significantly better than another one. In most cases, the ERSP, the $Z_{baseline}$ and Z_{H_0} methods exhibit the best results. It is noteworthy that Z_{H_0} is the only method which does not require to define a baseline.

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_{H_0} 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_{H_0} 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 ($\approx$ 10000 samples) would lead to bias in the estimators and deteriorate the performance of the methods.

II.4. Applications

II.4.1. Objective identification of HFOs

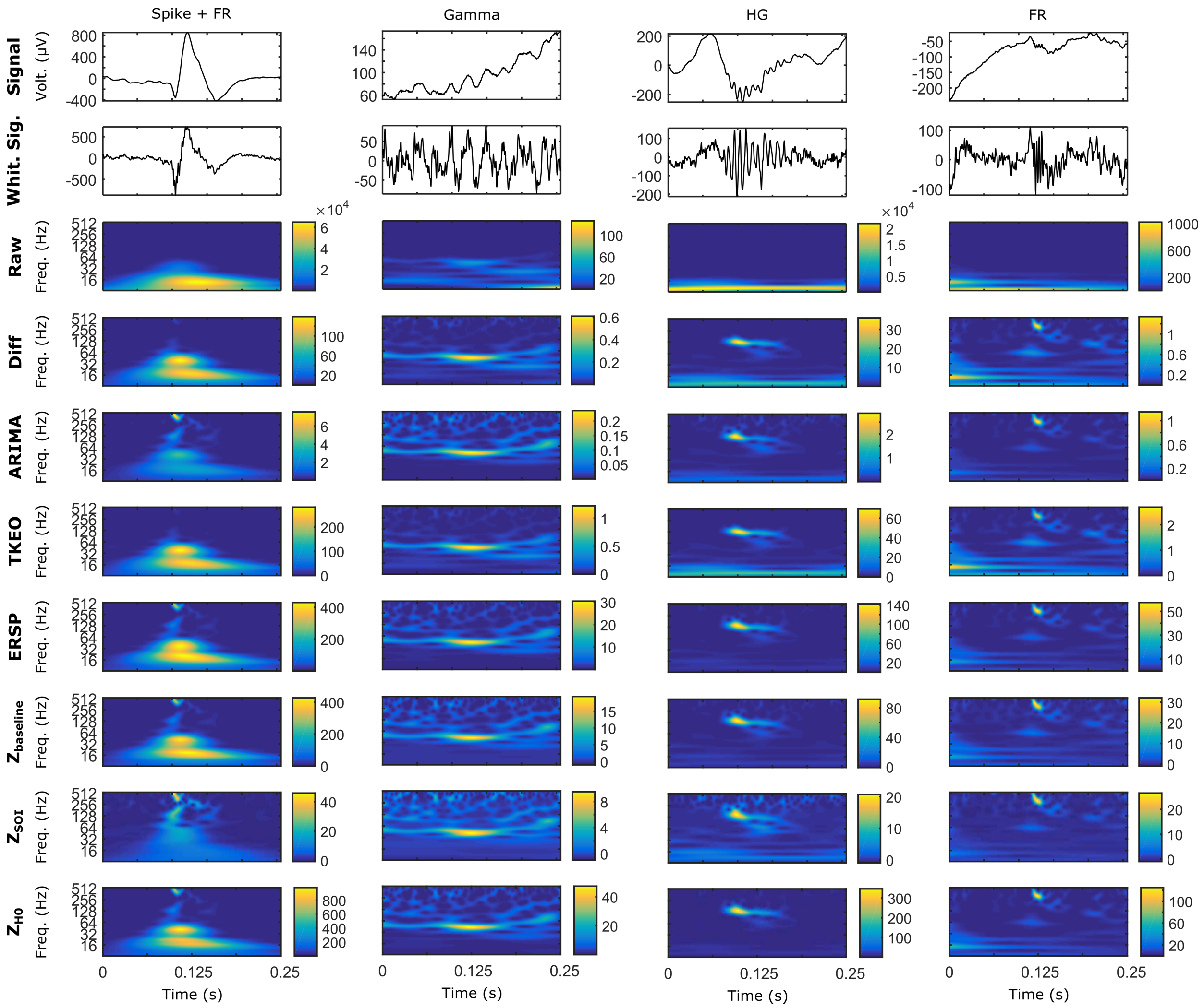


Figure II.8. – Examples of 4 *real* events in time and time-frequency domains. The events are represented in a 250 ms-windows and are taken from 2 patients. The first to the last column corresponds to a spike with a FR, a gamma oscillation, a HG, and a FR respectively. The first line shows the original 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 II.4, the HFOs are not visible in the raw TF representation but appear in the whitened signal and TF.

The Z_{H_0} 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 II.8 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 are depicted in the first, second, third and fourth column respectively. As discussed above, the ARIMA and Z_{SOI} tend to increase high-frequencies whether they originate from

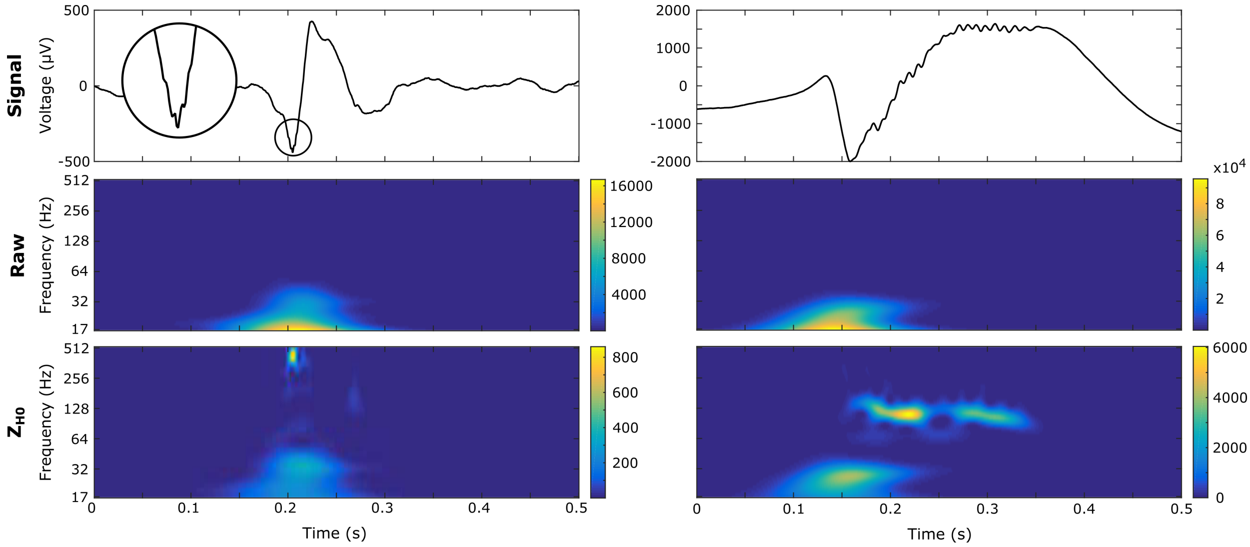


Figure II.9. – Examples of real FR and HG are illustrated in a wider frame. The FR and HG are represented in the time-domain and in the Raw and Z_{H_0} TF image in the left and right panel respectively. The round inset is a stretched version of the FR. Both HFOs are completely hidden in the Raw TF map but clearly visible in the normalized version.

the background or the signal. It is very clear in the first and second column. The spike is dulled and the FR very energetic but the high frequency background is visible above the HG for these two techniques whereas it is not for the others. Moreover, note 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 II.5 we showed that these methods could not remove the elbow of the spectrum. Indeed in the first column the FR appears very dull compared to the spike for both methods. The baseline based methods (ERSP and $Z_{baseline}$) and Z_{H_0} manage to capture all the frequency contents of the examples. 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 II.9 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 [Urrestarazu et al., 2007], but it is possible that part of these could arise from filtering artifacts [Bénar et al., 2010]. Time frequency methods should in principle improve this situation [Jmail et al., 2011], but may be “blinded” by the high energy contained in the spike [Amiri et al., 2015].

Fig II.10 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 searchers/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 oscilla-

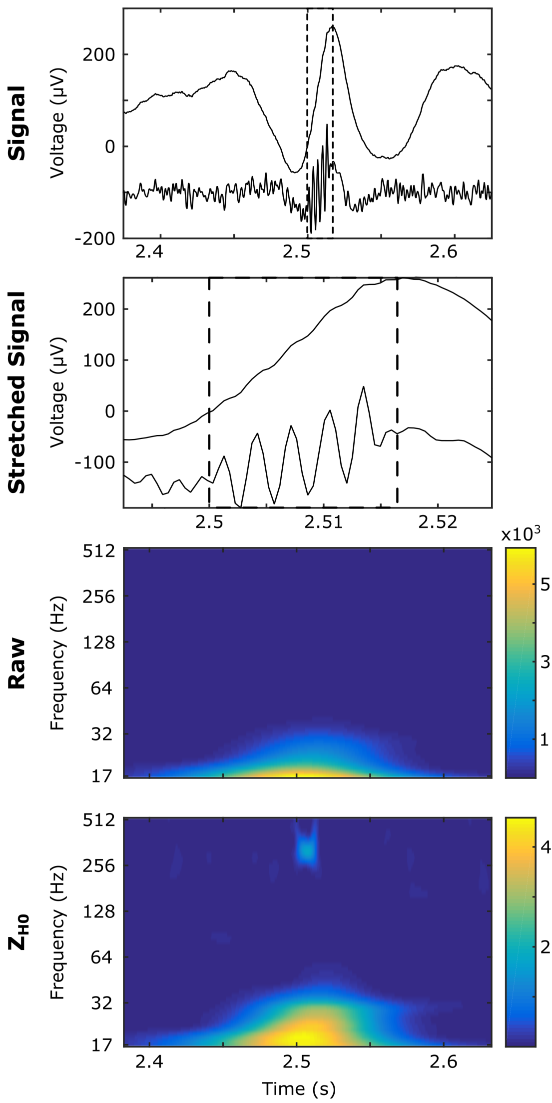


Figure II.10. – An example of a *simulated* Fast-Ripple (FR) occurring during a sharp transient. The box with dashed lines delineate the FR. The first line corresponds to the time course of the raw (above) and the Z_{H_0} whitened reconstructed signal (below.) The second line is a stretched version of the first line. The third and fourth lines are the raw and Z_{H_0} TF maps relative to the first line respectively. This HFO is hardly distinguishable in the time and TF domain on the original data. It is however evident in the Z_{H_0} frame and whitened reconstructed signal. Note that the spike is still present in the whitened reconstructed signal.

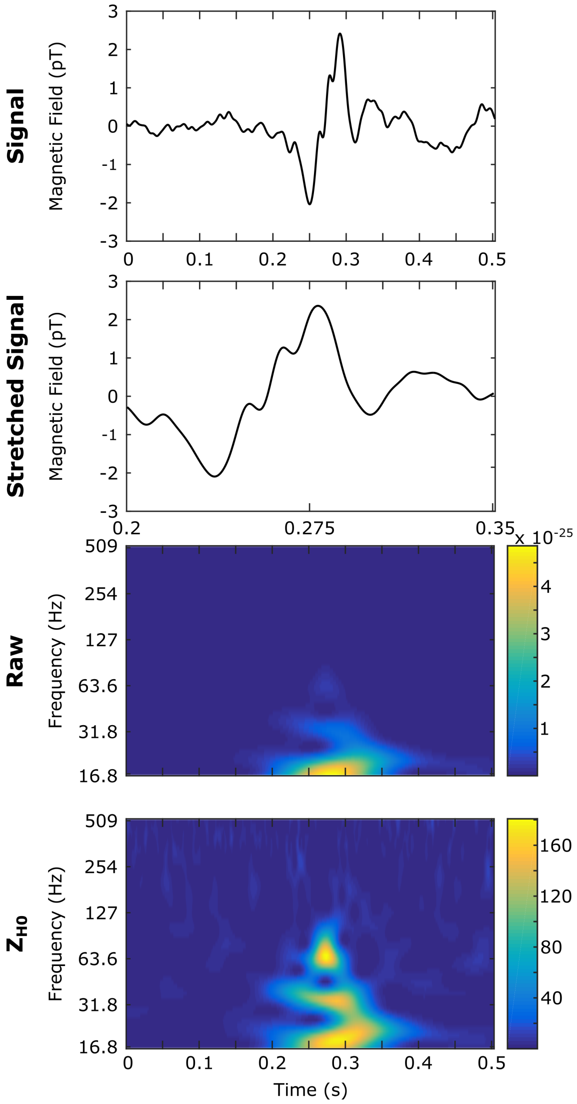


Figure II.11. – An example of a real gamma oscillation (γ) from MEG data is illustrated. The γ is represented above in the time-domain and in the Raw and Z_{H_0} TF image. The oscillation is enhanced in the Z_{H_0} image.

tion 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.

II.4.2. 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) [Chaitanya et al., 2015, Kobayashi et al., 2010] and magnetoencephalography (MEG) [van Klink et al., 2015]. 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 (γ, 25-80 Hz) marked in an MEG recording is shown in Fig II.11. The TF map is whitened as for SEEG data and the gamma is highlighted.

II.4.3. 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 [Burnos et al., 2014]. 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 [Staba et al., 2002] or short-time line length [Gardner et al., 2007], but this step suffers high FP [Gardner et al., 2007]; 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 [Gardner et al., 2007] 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 (lFDR) [Efron, 2005] 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

$$thr = x \quad | \quad \text{lFDR}(x) = H_0(x)/H_G(x) < Q, \tag{II.22}$$

with thr the threshold and Q the accepted level of lFDR. The lFDR would provide one single threshold for all frequencies. Thresholding methods used in previous detectors [Staba et al., 2002, Khalilov et al., 2005, Burnos et al., 2014] calculated the mean and the standard deviation of skewed distribution. By combining the lFDR 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 [Urrestarazu et al., 2007, Jacobs et al., 2008, Staba et al., 2002], 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 II.12. 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 II.13. We implemented a prototype of this detector in our open source software AnyWave [Colombet et al., 2015], 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.

II.5. Conclusion

This paper has examined several whitening methods for HFO representations 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, Z_{baseline} and Z_{H_0}. However, as the ERSP, Z_{baseline} requires a baseline, the Z_{H_0} 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 applications.

We would like to point out that real HFOs could be not visible in time-domain because they are hidden by the slope of sharp transients, which can be counter intuitive for neurophysiologists. The Z_{H_0} provides an objective tool to identify HFOs since it does not over express high-frequency activity of the background and dissociates sharp transient from HFOs in time and TF domain. We have suggested that this technique could be applied on other types of electrophysiological recordings with a power law spectrum. Future work

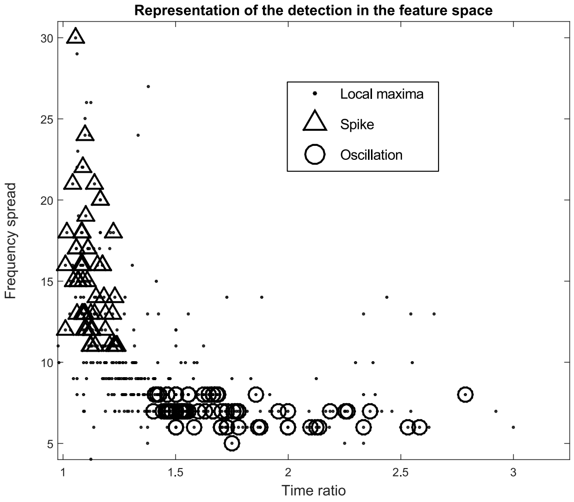


Figure II.12. – Representation of the local maxima in the feature space. The time ratio in the x-axis corresponds to the Full Width at Half Maximum (FWHM) in time of the island relative to the local maxima compared to the width of the impulse response at the same frequency. The frequency spread in the 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.

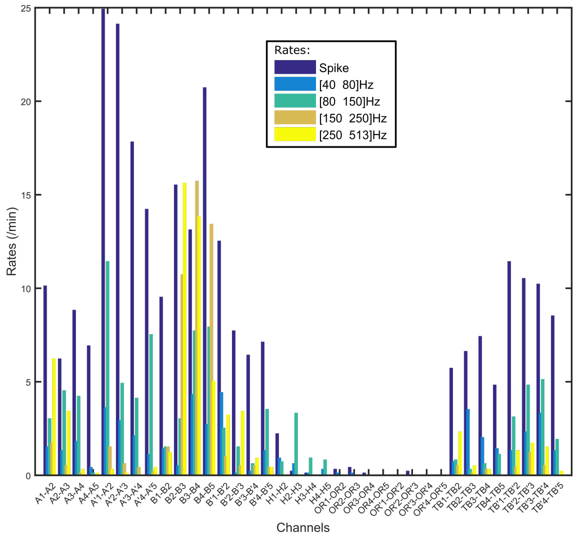


Figure II.13. – Example of the detection rates. The results of the detections are displayed in a bar graph giving the spike rates and the oscillation rates of given frequency ranges. These ranges can be manually set up. The name of the channels correspond to the name of the electrodes in bipolar montage (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.

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.

II.A. Appendix

The estimate of the power spectrum V_B in L^2 corresponds to

$$V_B^{L^2}(a) = \frac{1}{B} \int_{-B/2}^{+B/2} \left| T_f^{L^2}(b,a) \right|^2 db. \tag{II.23}$$

In L^1 it corresponds to

$$V_B^{L^1}(a) = \frac{a}{B} \int_{-B/2}^{+B/2} \left| T_f^{L^1}(b,a) \right|^2 db. \tag{II.24}$$

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.

Closing remarks

In this article we have designed and validated a new method (Z_{H_0}-normalization) for normalizing the TF image. Using this normalization, we have created an automatic procedure which marks each island in the TF space and classifies them as oscillations and spikes based on wavelet properties. This detector was subsequently named *Delphos: Detector of ElectroPhysiological Oscillations and Spikes*.

In the Appendix, we describe the two AnyWave plugins developped during this thesis. One is named *Delphos* (Appendix B.1.1 Delphos) and corresponds to the detector. The second is called *Delphos Viewer* (Appendix B.1.2 Delphos Viewer) and permits to visualize the Z_{H_0}-transformed data and to review previous detections.

Noteworthy, the thresholding method described above is not used anymore because of its computational cost in long datasets. It was preferred to manually set a threshold based on the standard normal distribution (which is the one of the background after normalization). This threshold controls the amount of FP.

Other properties of this normalization are discussed in V Discussion and future perspectives (V.2.1 Z_{H_0}-suppression and V.2.2 Ridge-Riding Denoising).

III. Detection Validation

Roehri N. *et al.*, Plos One 2017

Preamble

After having developed this new detector, Delphos had to be characterized and validated. The usual way of validating detectors is to compare the detections with markings of one or several experts [Zelmann et al., 2012, Spring et al., 2017]. One can then calculate the performance of the detector using the binary classification metric described in I.6 Binary Classification. One could wonder if having such a static measure is sufficient. Indeed, to apply a detector, shouldn't we be able to know how the detectors handle ripples and fast-ripples, low amplitudes HFOs, various background activities or numerous spikes? These questions remain open because of the fixed content of real data. In this chapter, we describe a benchmark based on realistic simulations which aims at comprehensively characterizing HFO detectors across several parameters. This benchmark is then used to compare Delphos to four standard detectors available in the Ripplelab Toolbox [Navarrete et al., 2016].

What are the Assets and Weaknesses of HFO detectors? A Benchmark Framework based on Realistic Simulations

Nicolas Roehri[1], Francesca Pizzo[1,2], Fabrice Bartolomei[1,2], Fabrice Wendling[3] and Christian-George Bénar[1]

[1] Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France
[2] APHM, Timone hospital, Clinical Neurophysiology, Marseille, France
[3] INSERM U1099, Université de Rennes 1, LTSI, Rennes, France

High-frequency oscillations (HFO) have been suggested as biomarkers of epileptic tissues. While visual marking of these short and small oscillations is tedious and time-consuming, automatic HFO detectors have not yet met a large consensus. Even though detectors have been shown to perform well when validated against visual marking, the large number of false detections due to their lack of robustness hinder their clinical application. In this study, we developed a validation framework based on realistic and controlled simulations to quantify precisely the assets and weaknesses of current detectors. We constructed a dictionary of synthesized elements – HFOs and epileptic spikes – from different patients and brain areas by extracting these elements from the original data using discrete wavelet transform coefficients. These elements were then added to their corresponding simulated background activity (preserving patient- and region- specific spectra). We tested five existing detectors against this benchmark. Compared to other studies confronting detectors, we did not only rank them according their performance but we investigated the reasons leading to these results. Our simulation sets, thanks to their realism and their variability, enabled us to highlight unreported issues of current detectors: (1) the lack of robust estimation of the background activity, (2) the underestimated impact of the $1/f$ spectrum of physiological data, and (3) the inadequate criteria defining an HFO. We believe that our benchmark framework could be a valuable tool to translate HFOs into clinical environment.

III.1. Introduction

High-frequency oscillations (HFOs) are putative markers of epileptogenicity [Urrestarazu et al., 2007, Jacobs et al., 2012]. The visual review of HFOs on intracerebral electroencephalography is time-consuming and tedious, and suffers from poor inter-reviewer reliability [Worrell et al., 2012]. There is thus a crucial need for automatic detection in order to translate HFOs into clinical practice. Many different HFO detectors have been designed [Staba et al., 2002, Gardner et al., 2007, Crépon et al., 2010, Zelmann et al., 2012, Birot et al., 2013, Burnos et al., 2014, Jrad et al., 2016] but none has yet met a large consensus. The validation of the detectors is thus an important stage. Such validation has to show how well the elements of interest are detected and how robust each algorithm is to variations of signals, e.g. the ongoing background (BKG) activity or the numbers of events. Most HFO detectors were validated using real data [Staba et al., 2002, Gardner et al., 2007, Zelmann et al., 2012, Burnos et al., 2014], based on visual marking, using marking by an expert as ground truth. The validation process typically consists in three steps [Staba et al., 2002, Gardner et al., 2007, Zelmann et al., 2012, Burnos et al., 2014]: firstly, expert reviewers mark events of interest, i.e. ripple (R, 80-250 Hz) and fast-ripple (FR, 250-500 Hz), on a recording; then the detector is run over the same data and finally the detections are compared to the reviewers' marking and its performance is measured.

This strategy however suffers from several issues. The main drawback is that there is no gold standard: is an HFO marked by the detector but not by the reviewers necessarily wrong? It could be that the detector extracts a feature that is not evident for the reviewers. Without gold standard, this situation will always be considered as an erroneous detection. Furthermore, the performance highly depends on the content of the tested signal. On the one hand, the performance is measured as a global agreement over two classes (R and FR) between the detections and the reviewers' marking. Since there are more Rs than FRs [Bagshaw et al., 2009] and that Rs are probably easier to detect than FRs due to their higher amplitude and longer duration [Staba et al., 2002] and the $1/f$ spectrum, the performance measure does not reflect the actual efficiency of the detector on the R and FR class separately. All the FRs could be missed but the performance of the detector could still be high because of this imbalance. Since the FRs are a putative better marker of epileptic tissues [Urrestarazu et al., 2007, van't Klooster et al., 2011, Fedele et al., 2016] this may be a critical issue. On the other hand, robustness cannot be studied in real data since the content is fixed. For instance, aforementioned detectors rely on thresholds which are calculated using measures such as standard deviation or percentile. These measures depend on the distribution of the data and therefore on the content of the recordings (number of spikes and number of HFOs). Moreover, if detectors were tested on recordings with high amplitude HFOs, the results cannot be extrapolated to cases where the amplitude of the HFOs is closer to the amplitude of the BKG.

Realistic simulations may be a solution to these drawbacks. The balance between the classes can be controlled and the gold standard is obviously available. We propose a method that aims at building a dictionary of representative extracted events (Spike, R, FR) from different brain areas and integrating them in a controlled manner in a simulated BKG from the same area to determine the performance and robustness of the detectors.

In a second step, we show how to use the results of this benchmark to diagnose currently-available detectors, in term of assets and weaknesses.

III.2. Methods

III.2.1. Clinical Database

Simulated data were created by using recordings of non-REM slow wave sleep of drug resistant epileptic patients undergoing pre-surgical examination with stereoelectroencephalography (SEEG) using macro-electrodes of diameter 0.8 mm. The same type of electrodes was used in all patients. The frequency sampling was 2048 Hz with an anti-aliasing filter set at one third of the sampling frequency (688 Hz). We only selected contacts which were inside the epileptogenic zone defined by expert neurophysiologists (FP and FB). Selected channels had to exhibit interictal spikes, ripples and fast-ripples on bipolar montage. Explored brain areas included the mesio-temporal, lateral temporal, frontomesial and frontolateral regions. We did not simulate posterior regions because they are under-explored and little is known about HFOs in these areas.

III.2.2. Simulation

Extraction of events

To build the dictionary of representative events, we only used visual marking and no automatic detection to avoid favoring one detector. We visually marked the events of interest on a bipolar montage according to [Zelmann et al., 2009]. We extracted them from the signal using the following method. We developed a Graphical User Interface (GUI) which displays in three different panels simultaneously the raw signal with the extracted signal, the normalized time-frequency (TF) obtained using Continuous Wavelet Transform (CWT) and, finally the TF relative to the discrete wavelet transform (DWT) (Wavelet Toolbox of MATLAB). By selecting the relevant coefficients of the DWT TF, one can reconstruct parts of the signal and thus extract elements without the surrounding BKG. If elements were superimposed, i.e. an HFO was superimposed on a spike or a FR on a R, we reconstructed the R, FR and spike separately by selecting the appropriate DWT coefficients. We skipped the elements which were overlapping both in time and frequency [Jmail et al., 2011]. Two examples of extracted events are given in Fig III.1. To avoid introducing unwanted oscillations, the user ensured both in normalized TF and in the original and synthesized band-pass filtered signals the correct extraction of the elements[1].

---

[1] Only obvious HFOs were extracted (about a hundred). They had to have more than four oscillations in the filtered signals. Moreover, the synthesized HFOs were generally longer than four oscillations since no background was present in the extracted versions. The extracted oscillations were discussed between NR and FP to ensure their correct extractions and that each of them passed the different criteria of the tested detectors. Moreover, each extracted ripples and fast-ripples were detected by every detectors when simulated with very high SNR without spikes. The aim was not to mark every possible HFOs of a given regions but to have some real variability such as amplitude and frequency modulations.

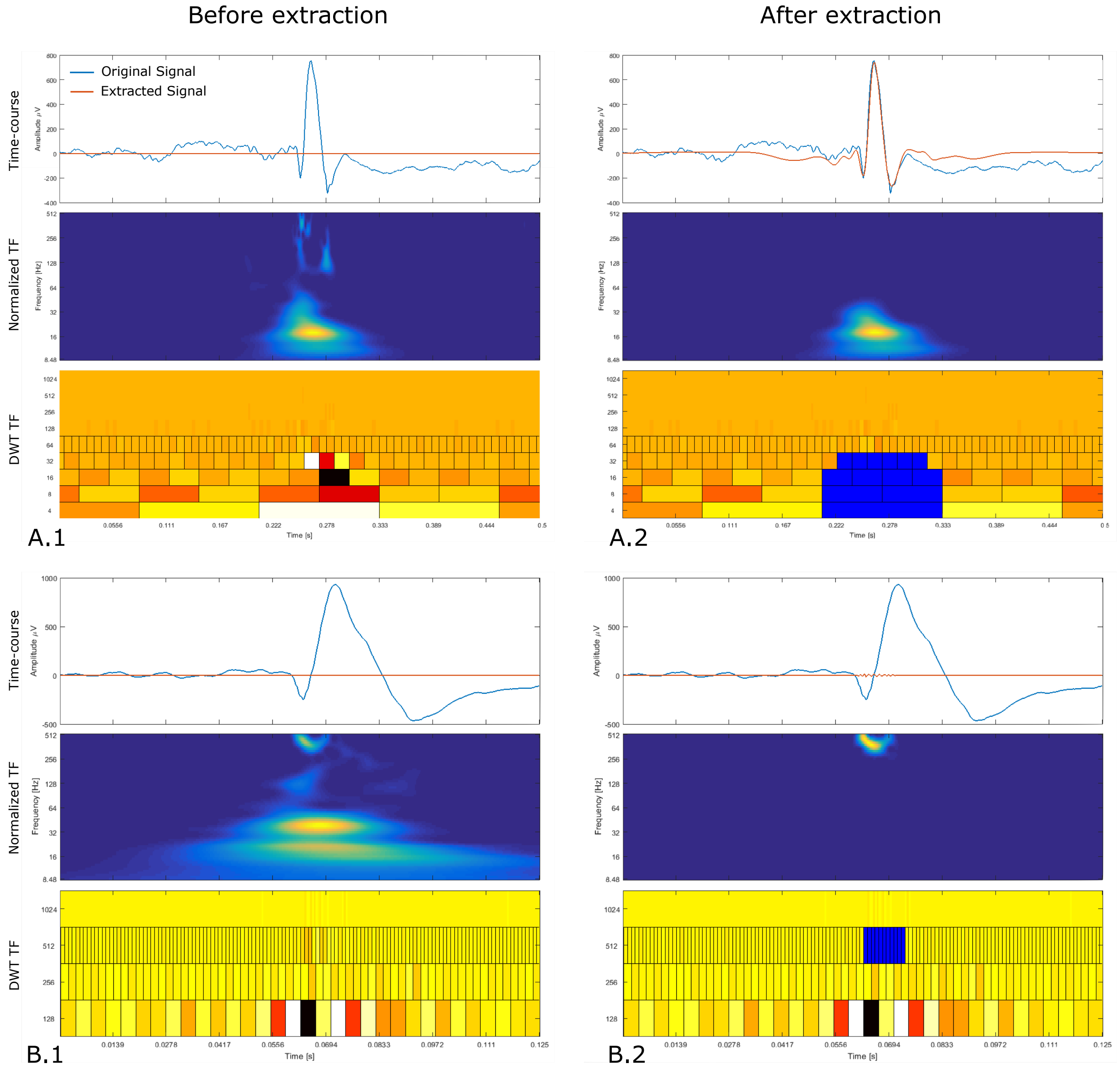


Figure III.1. – Each panel is composed, from top to bottom, of the original signal (blue) with the synthesize signal (orange), the normalized continuous TF image and the discrete TF image. Panels A and B display the process of extracting an epileptic spike and a fast-ripple riding a spike, respectively, at the beginning (.1) and the end (.2) of the process. We also provide an animated version of these two panels here : `http://meg.univ-amu.fr/wiki/AnyWave:Plugin_Simul`. User selects the coefficients of interest in the discrete TF image by left-clicking the appropriate tile. Once a tile is selected, it turns blue and the inverse transform is applied to the corresponding coefficients in order to reconstruct the signal. At the beginning of the process, the continuous TF represents the TF of the original data and during the process it shows the TF of the extracted signal.

This paradigm allows us to control the Signal to Noise Ratio (SNR) of each event separately according to their own frequency band[2]. This also guaranty the stability of the

---

[2] The SNR is calculated in the ripple and fast-ripple band for the ripples and fast-ripples, respectively.

BKG of the simulation which would not have been the case if we had cut and added real elements which would have ineluctably carried a part of human BKG.

HFOs could have been simulated by tapering sine waves of specific frequencies as it was done in [Navarrete et al., 2016, Roehri et al., 2016] but we wished to keep the original shape of the HFOs as it was recorded to build a dictionary relative to the structure. This also allows us to use frequency and amplitude modulations found in real HFOs which makes the simulation more realistic. For the same reasons, we used this technique to extract spikes.

Background Simulation

The ongoing background activity (BKG) of the areas where the HFOs and spikes were extracted were simulated. This step is useful to verify the robustness of the algorithm to variability in BKG activity across brain structures. We marked several baseline sections of the recording, i.e. containing only BKG activity. Since errors may occur, reviewers were asked to mark several pieces of baseline of several seconds for each channel of interest. The coefficients of an autoregressive (AR) model[3] were estimated for each section (e.g. Matlab's LPC function) and averaged over the pieces of the same channel to have an estimate of the BKG for each channel. To generate the simulated BKG activity, we filtered a Gaussian white noise with the averaged AR coefficients. Every BKG is generated using a new realization of the Gaussian white noise. Fig III.2 represents a schematic view of the BKG simulation.

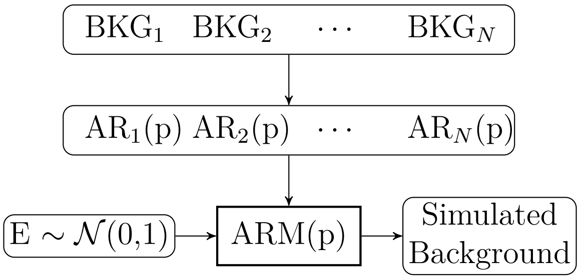


Figure III.2. – Background simulation pipeline. BKG segments are taken from the given BKG sections, the AR coefficients are estimated for each BKG section and averaged. The simulated BKG are obtained by filtering a Gaussian white noise with the average AR coefficients.

Simulation Generator

Once the events and BKGs of the different areas are extracted, they can be added together in a controlled manner by the following procedure. This generator aims at simulating

---

[3] As we wanted to reproduce the whole spectrum of the baseline sections, we purposely over-fitted them by using 200 coefficients. There is thus no reason why the high frequency components of the spectrum were underestimated by the AR model. Moreover, there is no issue with over-fitting since we do not want to generalize what we are estimating contrary to usual model fitting.

seven classes of events corresponding to: 1. Spike (Spk), 2. Spk co-occurring with a R (Spk-R), 3. Spk co-occurring with a FR (Spk-FR), 4. Spk co-occurring with a R and a FR (Spk-R-FR), 5. R , 6. FR and, 7. R co-occurring with a FR (R-FR).

Using the same GUI, the user sets the rates of each class, the values of SNRs and the number of realization per simulated channel and per SNR. The event class are randomly drawn according to the input rates for each timing. Finally, the events are selected randomly in the dictionary of the channel of the patient according to the selected class. Spks and HFOs are added using different processes. First of all, the SNR of the HFOs is fixed for each realization of the BKG whereas the SNR of the Spk varies randomly. Each inserted Spk is first stretched or compressed by a pseudo-random factor k, scaled to fit the randomly selected SNR and inserted in the BKG. By doing so, we can create spikes which exhibit a high-frequency component and reproduce the overlap in time and frequency of HFOs superimposed on a spike which we could not extract properly as explained in section Extraction of events. Moreover we simulated the post-spike silencing [Kobayashi et al., 2009] by reducing the amplitude of the BKG during the Spk. This generates non-stationarity in the BKG and challenges the estimation of features used for the threshold. Each inserted HFO is first scaled to fit the fixed selected SNR in its relative frequency band and added to the BKG. Globally the timing of the events follows a random Poisson process. Locally, we added a timing jitter to the HFO when superimposed to HFO or spike. An example of the insertion of the elements is given in Fig III.3. Fig III.4A illustrates real channels facing their simulated versions, Fig III.4B displays the location of the simulated channels in a 3D brain mesh and Fig III.5 shows examples of simulated HFOs. We generated a set of simulations with a rate of 3 /min for each event class. It contains 30 realizations of each channel for SNRs set to 0-5-10-15 dB

III.2.3. Performance Evaluation

Once the simulations have been generated, detectors can be run. We defined a small time window of 100 ms centered on each inserted HFO as the confidence interval (CI). CIs containing detections were considered as True Positives (TP), those without detections were defined as False Negatives (FN) and, detections falling outside CIs were labeled as False Positives (FP) . We use the precision (Prec) and sensitivity (Sens) criteria as well as the F-measure which combines Prec and Sens to characterize the performance over the SNRs. We avoid using the specificity because of the issue of the imbalance in the distribution of positive and negative instances [Fawcett, 2006] and because the True Negative are not clearly defined. Sens and Prec are defined as follows

$$Sens = TPR = TP/P = TP/(TP + FN), \tag{III.1}$$

$$Prec = PPV = TP/(TP + FP). \tag{III.2}$$

We also computed the sensitivity per class to determine whether the quality of detection was class dependent.

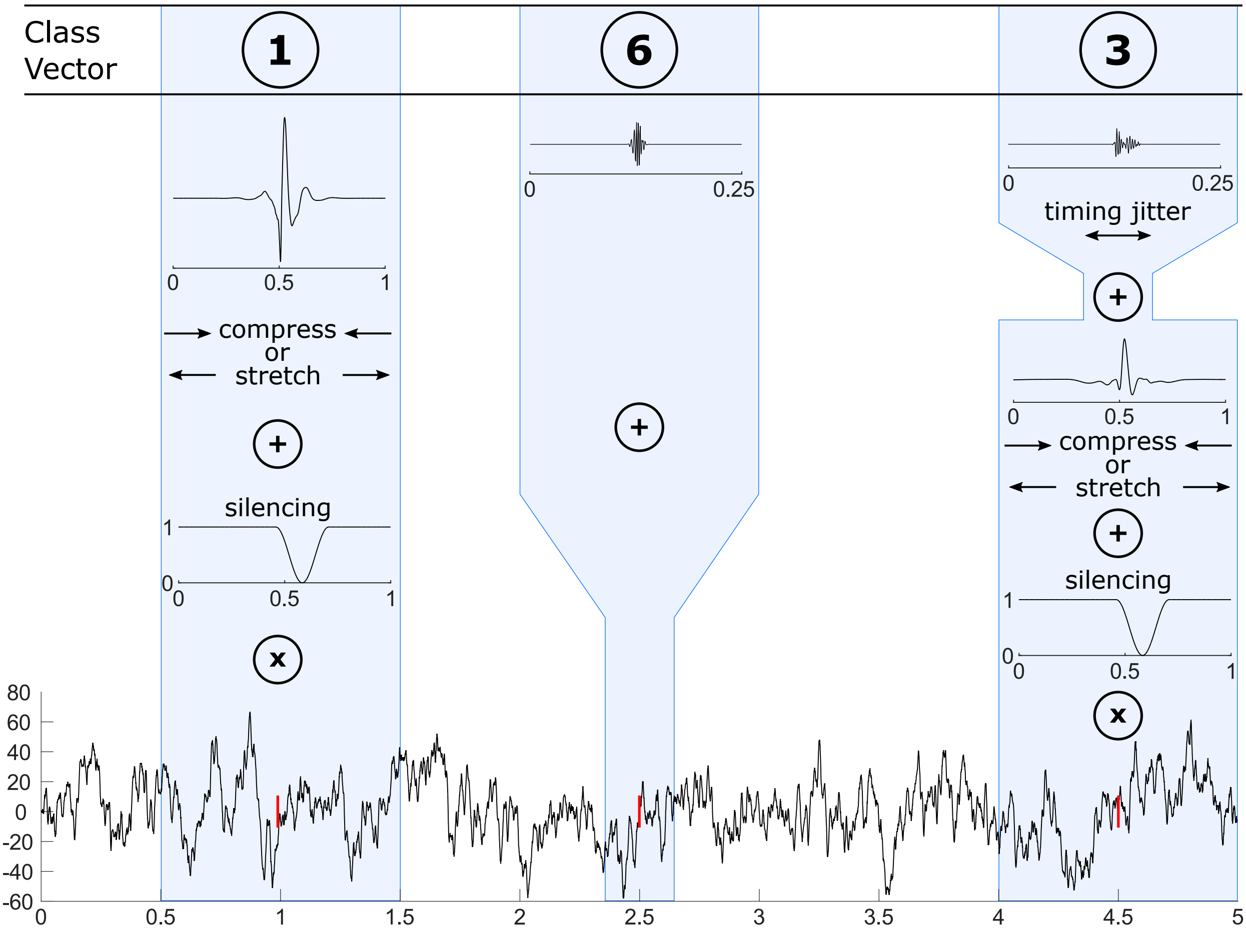


Figure III.3. – Event insertion pipeline. The class vector contains a set of randomly drawn numbers corresponding to the event classes for each timing (red line). At 1, 2.5 and 4.5 s, events of the first (Spk), third (FR) and sixth (Spk-FR) classes are added to the BKG. Each Spk is stretched or compressed and scaled to fit a random SNR (ranging from 0-15 dB). The BKG is multiplied by a notch to reproduce post-spike silencing. HFOs are scaled to fit the chosen SNR and have a timing jitter when riding a Spk.

III.2.4. HFO Detectors

Five automatic detectors were tested; four of them come from the RIPPLELAB Toolbox [Navarrete et al., 2016] which we integrated in our open source software AnyWave [Colombet et al., 2015]. The fifth detector is our detector called Delphos (Detector of ElectroPhysiological Oscillations and Spikes) described in [Roehri et al., 2016]. The next paragraph will briefly summarize the different detection methods.

The Short Time Energy detector (STE) [Staba et al., 2002] is based on the moving average of the root mean square amplitude of the filtered signal. The segments above five times the standard deviation (SD) plus the mean energy lasting more than 6 ms are considered as putative HFOs. These events are kept if containing more than 6 peaks greater than 3 times the SD above the mean value of the rectified band-pass signal.

The Short Line Length detector (SLL) [Gardner et al., 2007] calculates the line length energy of a sliding window applied to a first-order backward differencing and band-pass filtered signal. A detection is retained if its amplitude is greater than the 97.5th percentile

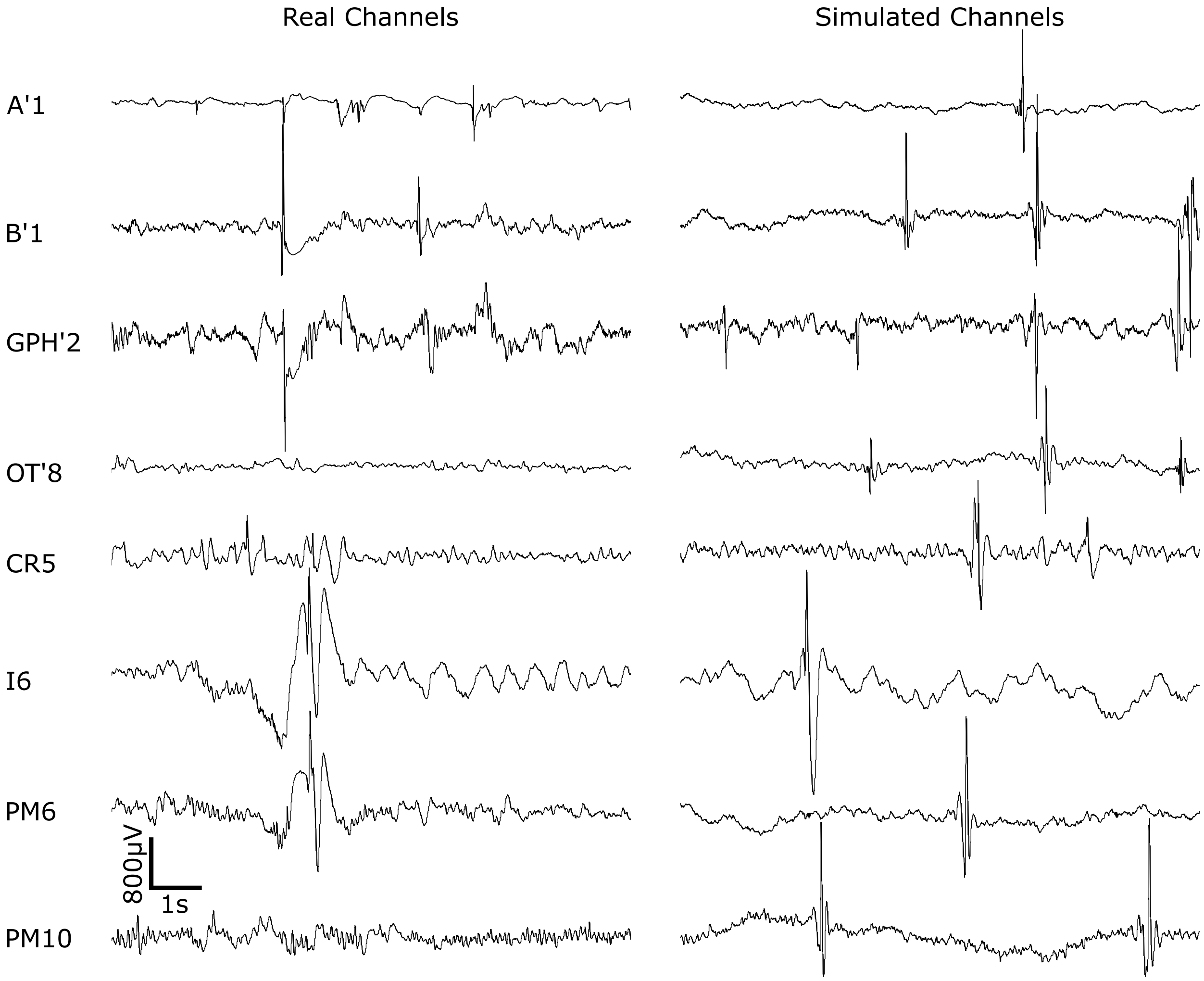


Figure III.4. – Example of real data and their corresponding simulations with their 3D localization. Each color corresponds to a different patient.

of the empirical cumulative distribution function and longer than 12 ms.

The Hilbert detector (HIL) [Crépon et al., 2010] computes the envelope of the filtered signal using the Hilbert Transform. The local maxima exceeding 5 SD of the envelope of the whole signal with a minimal time length of 10 ms are labeled as HFOs.

The MNI detector (MNI) [Zelmann et al., 2012] is a two-stage algorithm. Firstly, baseline segments are detected using the wavelet entropy of the auto-correlation function of the band-pass filtered signal. If there is enough baseline segments, the detection procedure is similar to the STE detector except that the threshold is set to the 99.9999 percentile of the empirical cumulative distribution function of the baseline segments. Otherwise, the STE detector is run with a threshold optimized by iteratively removing previously detected HFOs. We modified the algorithm of RIPPLELAB concerning this detector to be closer to the original one described in [Chander, 2007][4].

---

[4] In the RIPPLELAB toolbox, a section was labeled as baseline when its mean wavelet entropy was above the chosen threshold whereas in [Chander, 2007] and also in [Zelmann et al., 2012] the criterion was on the minimum wavelet entropy. We modified the code in that sense.

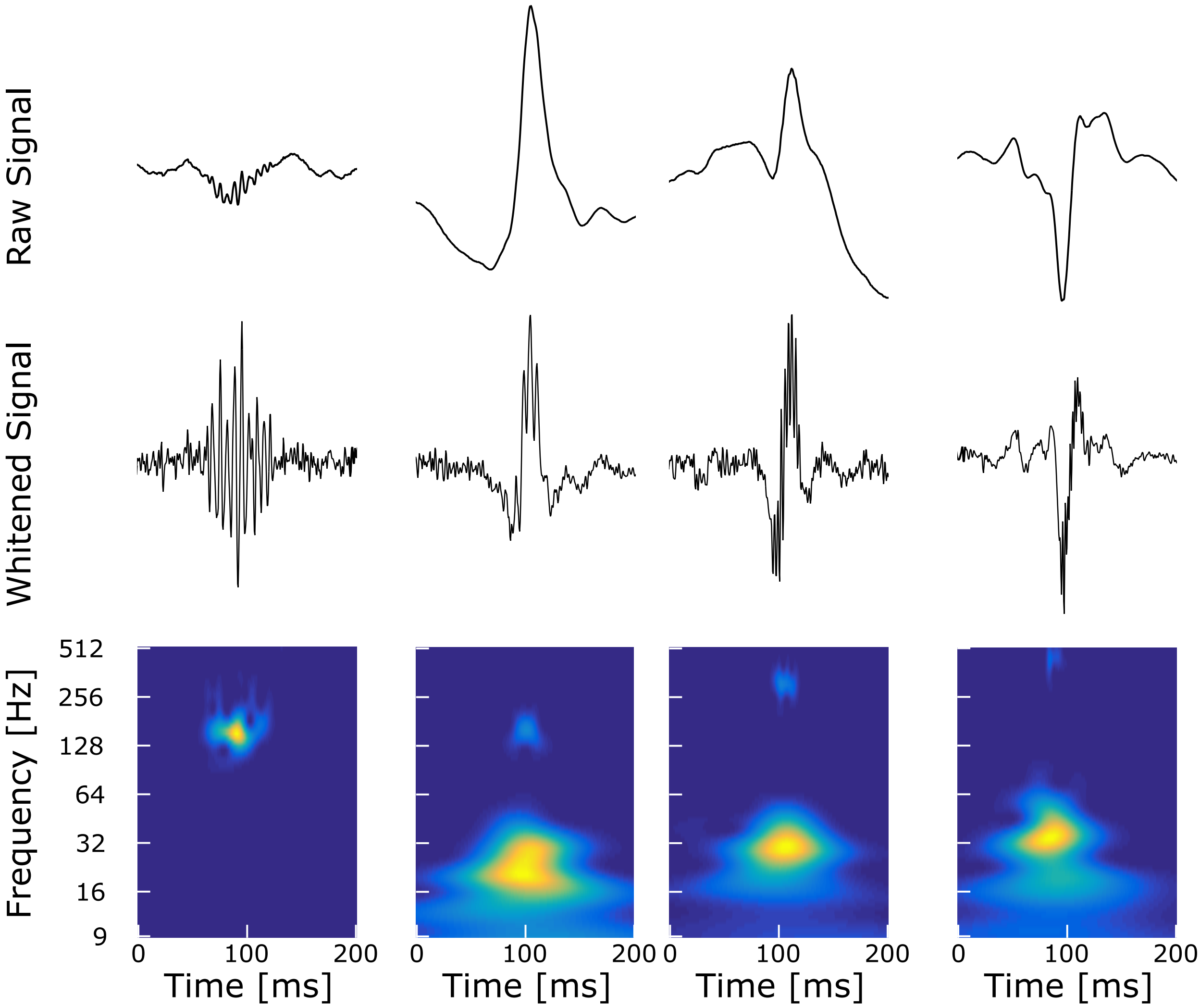


Figure III.5. – Example of simulated HFOs at 15 dB (R, Spk-R, Spk-FR, Spk-FR). Each column is composed, from top to bottom, of the raw signal, the whitened reconstructed signal and Z_{H_0} normalized TF image [Roehri et al., 2016]. Note that in the raw signal the HFOs above are hidden in the slope of the spikes. There is however no doubt about their existence since there were manually added. Also observe that the whitened signal unveils these HFOs while preserving the shape of the spikes.

All the parameters of the detectors correspond to the default setting of RIPPLELAB which corresponds to the one used in the respective articles. The signals were bandpass filtered between 80-500 Hz for all detectors as in the original publications, except for Delphos which does not rely on filtered signals.

Delphos detects oscillations and spikes in the Z_{H_0} TF [Roehri et al., 2016] representation by analyzing the time width and frequency spread of peaks above a threshold. The values in the Z_{H_0} TF correspond to an normalized energy. We set the threshold to 30 in the Z_{H_0} TF. Detections are classified as oscillation, if their frequency spread is similar to the one of the wavelet and their time width is greater than the one of a Dirac impulse, or as spike if their frequency spread is greater than the one of the wavelet and their time width is similar to the one of a Dirac impulse.

III.3. Results and Discussion

Fig III.4 illustrates the realism of our simulation. Note the similarity between the simulated BKG and its original form and its variability. Moreover the shape of the spikes and the HFOs are more complex and diverse than previous simulations [Bénar et al., 2010, Roehri et al., 2016, Navarrete et al., 2016], which is more realistic. Finally the structure of the simulation, i.e. the statistical similarity of the BKG corresponding to each channel, the control of the SNR and the control of the rates of every classes enables us to test the robustness of the detection.

According to the design of the simulation, one can make some assumptions on the results. First of all, the sensitivity should increase with increasing SNR. For 0 dB, the amplitude of the HFOs is similar to that of the BKG; consequently the sensitivity should be close to zero. Moreover, the variability of the precision should be greater than for other SNRs because of the low sensitivity. For 15 dB however, the amplitude of the HFOs is a lot higher than the one of the BKG; consequently the sensitivity should approach one. This is consistent with our results (Fig III.6 and Fig III.11), where the sensitivity of all detectors increases with the SNR. However, some detectors (MNI, HIL, STE) do not reach 1 in sensitivity. This observation is partly explained by the fluctuation of the threshold of the STE, HIL, SLL detectors (Fig III.7).

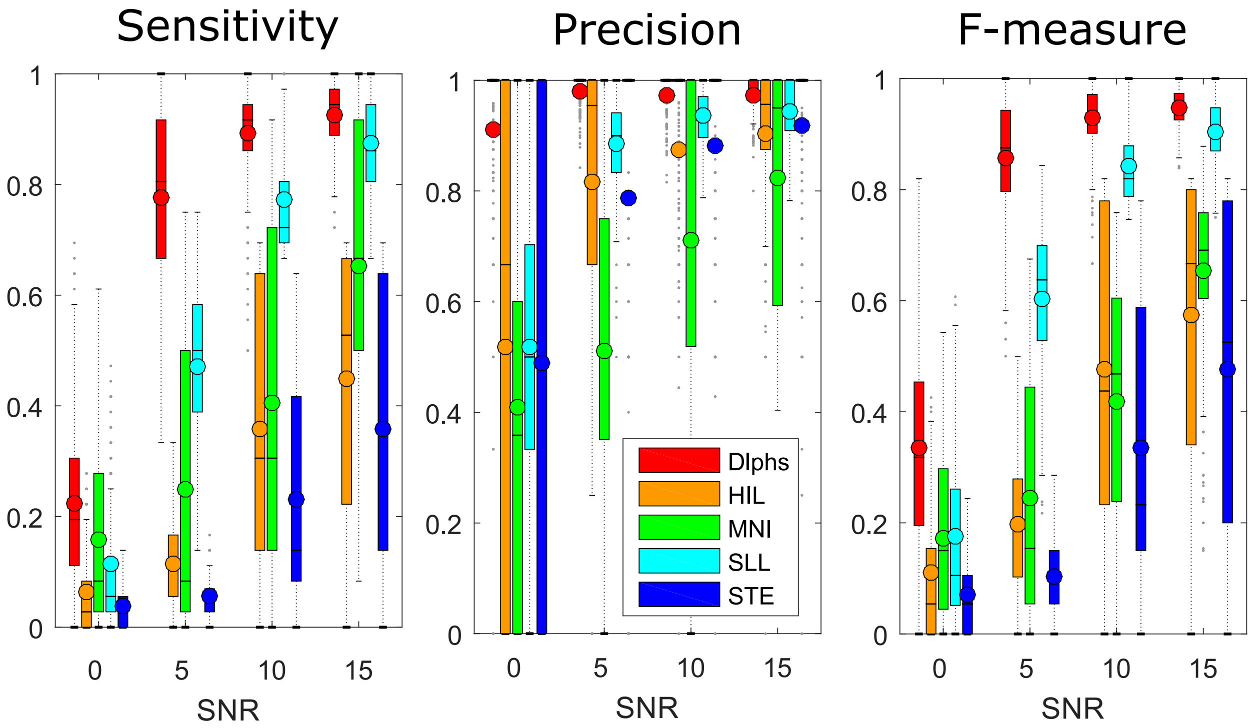


Figure III.6. – Boxplots of the sensitivity, precision and F-measure of the detectors ran over this benchmark for different SNRs. The circles correspond to the means and the dots represents the outliers. The sensitivity of all detectors increases with the SNR. Interestingly, the precision also increases with the SNR except for Delphos whose precision stays at 1. The F-measure combines the sensitivity and precision.

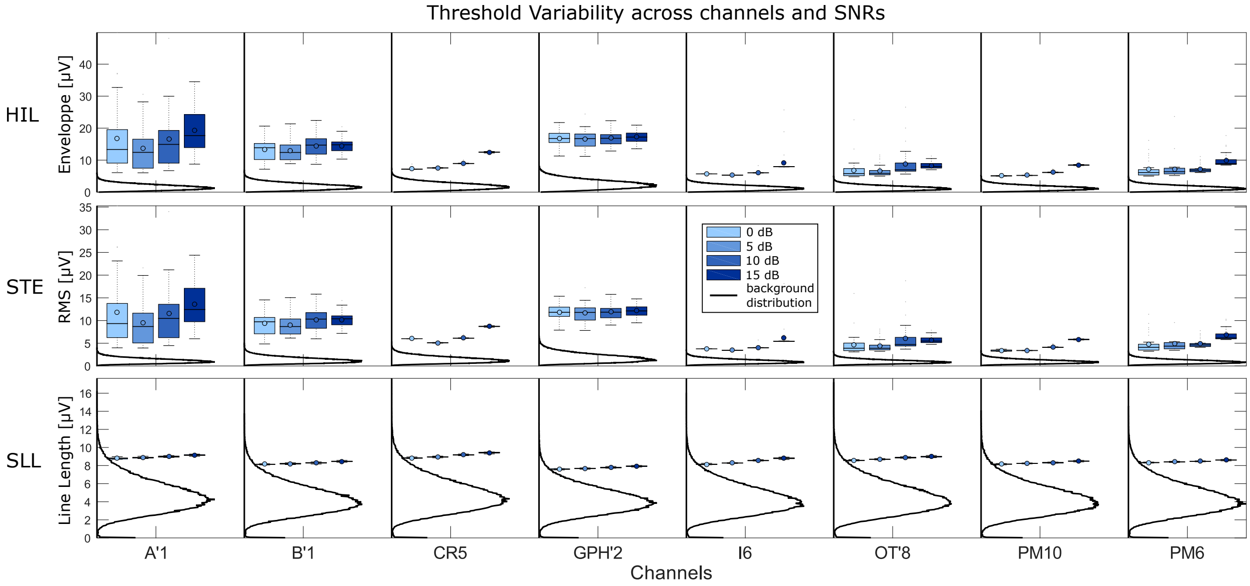


Figure III.7. – Boxplots of thresholds for the different metric (RMS, SLL and hilbert enveloppe) used by the one-stage detectors (STE, SLL and HIL) compared to the histogram of the metric applied to the simulated background.

Robust estimation of the background activity

In Fig III.7, we see that all thresholds are increasing with the SNR, especially for the HIL and STE detectors. This is consistent with the fact that they both rely on SD to calculate their threshold which leads to a biased estimation of the BKG boundary since the events weight more in the histogram for higher SNRs. The SLL nevertheless uses a percentile threshold which is more robust to outliers and thus more stable but still slightly increasing with the SNR. Note that this can be studied only because the simulated BKGs are statistically identical for each channel since it was generated using the same AR coefficients and white noise with the same properties. In other words, different realizations of a BKG from the same simulated channel yield the same histogram. Consequently the increase of the thresholds is only due to changes in amplitudes of the HFOs. Moreover, the STE and HIL thresholds are high compared to the tail of the BKG distributions; the threshold could be lowered to increase the sensitivity without losing in precision. Interestingly, the variability of the HIL and STE thresholds is the largest for all SNRs in A'1, B'1 and GPH'2, which are the channels recording from the amygdala, hippocampus cephalis and hippocampus caudalis, respectively. In these channels, we found the sharpest spikes, which created large outliers in the distributions. This shows again how the SD can be biased by outliers and how the percentile method of the SLL is more stable. Nevertheless, the percentile will always give a value in the range of the data; this means that if there is only BKG activity, the threshold will still select portion of the BKG as putative HFOs. These issues highlight the need of estimating the threshold based on the histogram of the BKG – the null hypothesis H_0 – and not on the histogram of the mixture (BKG plus events; H_0+H_1). This is what the MNI detector and Delphos aim at solving by either finding baseline/BKG sections or estimating robustly the BKG activity, respectively.

While Delphos uses the Z_{H_0} TF normalization to estimate the BKG activity, the MNI detector finds baseline/BKG segments based on wavelet entropy (WE). In theory, both techniques are similar since the threshold is calculated on either a robust estimate of the BKG histogram derived from the total histogram (Delphos) or on the histogram of the detected BKG segments (MNI detector). Delphos estimates the BKG activity at each frequency by fitting a Gaussian distribution on the histogram of the real coefficients of the wavelet transform within the Tukey's range (for more details please refer to [Roehri et al., 2016]). One can estimate the power of the BKG by squaring the obtained SD and also normalize the TF plan by z-scoring each line with the same SD. In Fig III.8, we can see that each histogram follows a standard Normal distribution – this shows that the estimation stage was effective – and that the spectra of the estimated BKGs computed on the simulated channels thoroughly follow the spectrum of the original BKG signal independently of the whole frequency content. The threshold was set to 30 which correspond to $thr = \sqrt{30}\sigma = \sqrt{30}$ and is thus stable since the estimation is robust.

On the contrary, the WE calculated on the auto-correlation function of the filtered signal depends on the power spectrum density (PSD). In Eq (III.14) of Appendix A, we show that applying the wavelet transform to the auto-correlation function is equivalent to applying it to the PSD of the filtered signal in the frequency domain. In other words, the values of the WE changes with the power of the spectrum whether there is rhythmicity, i.e. oscillations, or stronger broad-band activity.

The histogram of the minimum WE of all the sections (Fig III.9) demonstrates a shift for BKGs with higher power in the 80-500 Hz band. This shift is interpreted as channels with semi-continuous high-frequency (SCHF) activities and the MNI detector switches into the no baseline mode. Interestingly, this shift occurs in the B'1, GPH'2 and CR5 simulated channels which correspond to the synthesized hippocampus cephalis, hippocampus caudalis, and the cingulate region. These regions are prone to generating SCHF activities [Mari et al., 2012, Melani et al., 2013] and this illustrates again the realism of our simulation. The idea underlying the concept of SCHF activities is that, in those channels, it is harder to visually identify HFOs due to this active BKG. Therefore, to increase the sensitivity, the threshold has to be optimized to be closer to the BKG. Our simulation reproduces this context for the three aforementioned channels at 0 dB. In this context, the MNI detector exhibit the best performance in sensitivity compared to other detectors but have the worst precision (Fig III.10). The main issue is that the MNI remains in this "no baseline mode" even when the HFOs are distinguishable from the BKG (SNR>0 dB). It still performs well in term of sensitivity, even though it is caught up by Delphos and the SLL detector, but has the worst precision among the detectors. This is of clinical importance since R occurring in flat BKG seems to correlate more with seizure freedom than R occurring in oscillatory BKG [Kerber et al., 2014]. Despite this low threshold, Delphos and the SLL detectors manage to have an equal or better sensitivity generally but especially in those channels compared to the MNI detector. This introduces the second result of this benchmark; namely that detectors do not have the same sensitivity across the classes.

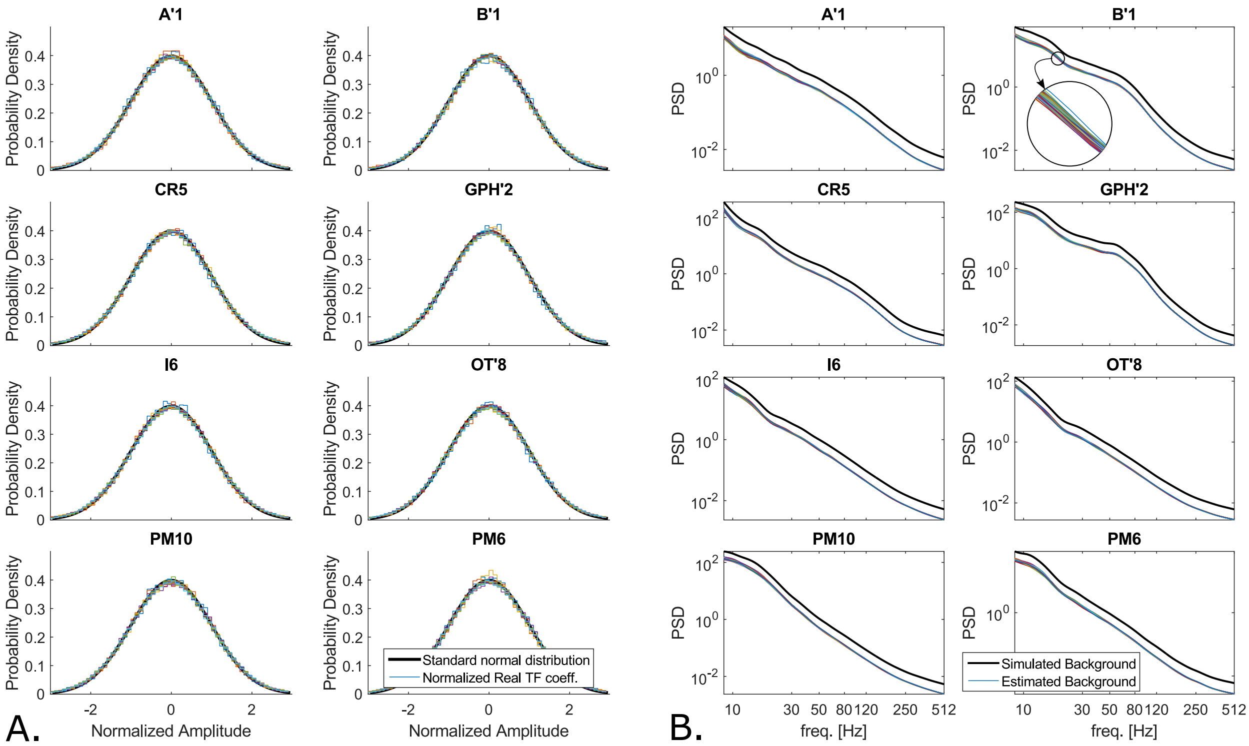


Figure III.8. – Estimation of background activity and spectra of different backgrounds calculated by Delphos on the simulated channels. A. shows in each panel the probability density function of a standard normal distribution ($\mu = 0$ and $\sigma = 1$) in solid black line and the histogram of the normalized real coefficients of the TF at different frequencies calculated for one realization at 15 dB of each simulated channels. The quality of the estimation of the BKG activity can be assessed by comparing the closeness of these histograms to a standard Normal distribution. B. represents the estimated BKG spectra (color coded) retrieved by Delphos from all the realizations of the simulated channels and the actual spectra of the simulated BKG of these channels (black line). We added an offset to the simulated BKG spectra to avoid overlapping of the estimated and the actual spectra. In B'1 panel, we added an insert showing the small variability of the method.

III.3.1. Accounting for the power spectrum

The MNI, HIL and STE detectors have a better sensitivity to the classes containing R (Spk-R, Spk-R-FR, R, R-FR) than to classes with FR only (Spk-FR and FR) (the second to the fourth panel of Fig III.10). The SLL detector is the opposite; it has high sensitivity to FR classes (Spk-FR, Spk-R-FR, FR, R-FR) but low sensitivity to classes with R only (Spk-R, R) (last panel of Fig III.10). Delphos seems to have either an equal or better sensitivity per class except for the Spk+R which is somehow lower than the other classes (first panel of Fig III.10). The first statement is explained by the $1/f$ structure of the physiological data. In our simulation, the SNR of the HFOs is constant for each realization and set according to the level of the BKG in the frequency band of the added HFOs, i.e. in the R or FR band. The amplitude of a R is therefore higher than that of a FR for the same SNR; the faster the oscillation, the smaller its amplitude [Demont-Guignard et al., 2012, Fink et al., 2015]. Determining a threshold in the 80-500 Hz band is thus suboptimal. Lowering the threshold to increase the sensitivity in the FR band would increase the number of false detections due to the higher BKG in the R band. A solution would be to run the detector

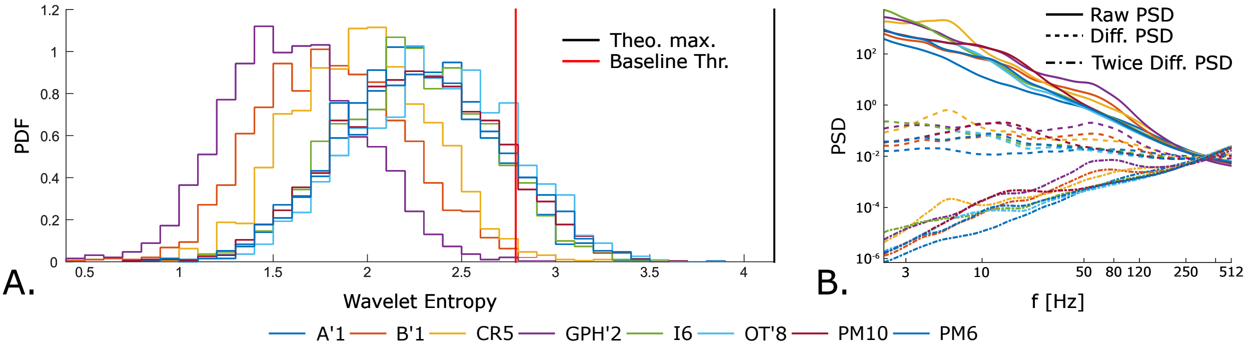


Figure III.9. – Wavelet Entropy (WE) of the simulated background of each channel and their differentiated power spectrum density. A. histograms of the minimum WE calculated on the autocorrelation function of the filtered simulated BKG (without inserted events) of each simulated channels. The theoretical maximum and the threshold value are represented in black and red lines respectively. B. Power spectrum densities (PSD) of the simulated BKG activity and the effect of differencing. Colors refer to the same channels in A. and B. Note that the histograms which have the fewest number of sections above the threshold values (A.) correspond to B'1(orange), CR5(yellow) and GPH'2(purple) which exhibit the highest power in the 80-500 Hz band (B.). Also observe that (i) the differentiated spectra (B., dashed line) is flatter than the raw spectra but not completely flat and that (ii) the 2nd-order differentiated spectra (B., dashed and pointed line) inverse the slope and over-express the high-frequencies.

on the R and FR band separately as was done in [Fedele et al., 2016, Burnos et al., 2016]. The slope of the spectrum still remains in each band but it undoubtedly improves the results for a longer computing time.

A solution is to whiten the spectrum, i.e. to flatten the spectrum in order to balance the power across frequencies as done by the SLL detector and by Delphos. However the latter performs better for each class compared to the SLL. This is due to the quality of the whitening stage. Delphos uses the Z_{H_0} TF whitening as discussed above in section Robust estimation of the background activity. In [Roehri et al., 2016], we showed that it was as efficient as normalizing using a baseline. The main advantage of the Z_{H_0} normalization is that it adapts to the shape of the BKG (Fig III.8) whereas the pre-whitening stage of the SLL is fixed. Moreover, we showed in [Roehri et al., 2016] that the pre-whitening by differencing was not optimal but would not lead to over-express the higher frequencies. In theory the SLL detector should detect the R as well as the FR since its threshold is robust and a pre-whitening stage is applied, In fact, in Eq III.6 of Appendix A, we derive that the line length metric used to detect the HFOs is equivalent to a second differencing stage, which over-expresses the higher frequencies (Fig III.9 B.). This explains the very high sensitivity in the FR band and the poor sensitivity in the R band. Furthermore, the differencing method is dependent on the sampling frequency (Eq III.8, Appendix A), i.e. the pre-whitening stage will not have the same effect if applied to data recorded at 1024 Hz or 2048 Hz. This is an important aspect to consider when discussing clinical results.

Globally, every detector has issues with detecting R superimposed on Spk when they are overlapping both in time and frequency. The main difficulty for detectors using time-series is to separate actual Rs riding on Spks from filter ringing [Bénar et al., 2010]. In [Amiri

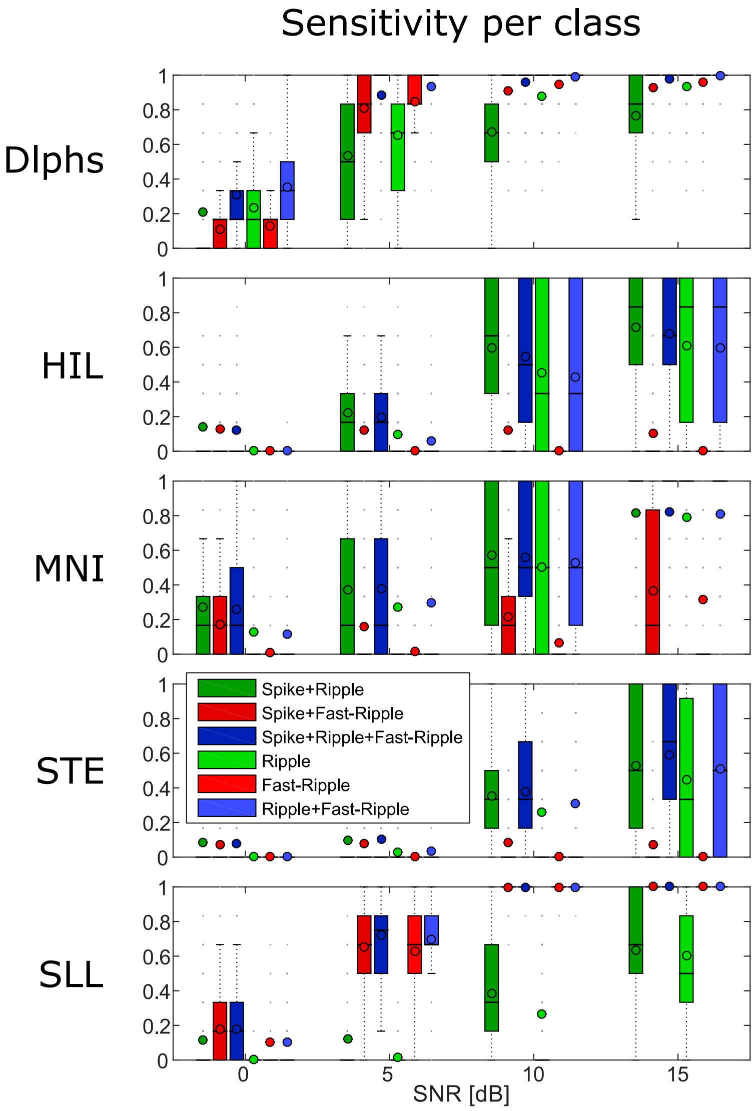


Figure III.10. – Boxplots of the sensitivity per class for each detector. Globally, the sensitivity per class increases with the SNR. All detectors have issues detecting ripple riding spike. The HIL,MNI and STE detectors are more sensitive to the ripple classes thant to the fast-ripple classes. The SLL has a better sensitivity in the FR band than in the R band. Delphos has a comparable sensitivity per class at each SNR.

et al., 2015] a solution was proposed based on feature extraction and unsupervised clustering but this method is not yet implemented in current detectors. In this latter article, the authors also highlight a potential problem of TF analyzes. Because of the structure of the analysis, the Spk could hide the HFOs. In fact, Delphos which detects HFOs in TF plane suffers from this masking but still outperform the others. This is probably due to the fact that masking only occurs for a very specific value of the phase and thus generate a single peak in the TF image instead of two separate. A possible future step would be to subtract the spike waveforms before running the detectors [Chaibi et al., 2014, Jmail et al., 2016].

Moreover, even though Delphos seems to give good results, it has some drawbacks. For instance, it misses HFOs with strong frequency modulation (chirp). This type of HFOs create a "comma-like blob" – similar to the FR shown in Fig III.1 – which makes the calculation of the time width difficult. A solution may be to calculate the width using the ridge to have a better estimate. This would probably increase the computing time greatly but the significant increase in sensitivity has yet to be proven. One could also add features to describe the "blob" such as its surface or volume but it would be difficult to compare it to a theoretical model.

III.3.2. Clear definition of HFOs

In previous paragraphs, we focused on the sensitivity (number of TP); we will now concentrate on the precision (impact of the FP on the detection). One key prerequisite of a detector is that its precision should ideally be constant across the SNRs[5] and should not have more FP at lower SNR than and at higher SNR and inversely. This involves two factors, a robust threshold – we discussed that in section Robust estimation of the background activity but also a strict definition of HFOs [Menendez de la Prida et al., 2015].

Most detectors defined them as a sustained increase in energy in the HFO band (SLL, HIL, MNI, STE). The STE detector adds the criterion of the number of cycle. It was however shown in [Bénar et al., 2010] that filtering artefact could produce sustained oscillations of high amplitude. The aforementioned criteria lack in distinguishing these false ripple from real ones and thus decrease their precision (Fig III.6 and Fig III.10). Also the duration threshold is an inadequate criterion. For the STE, it was fixed at 6 ms which corresponds to the duration of 3 periods of a 500 Hz oscillations. A 2-cycle oscillation at 250 Hz has a longer duration than 6 ms but do not fall under the definition of an HFO. To be efficient, this duration threshold should be calculated and applied after determining the frequency of the oscillation in order to fit the 3 or 4 cycles criterion.

Naturally, one can think of the TF analysis for quantifying this duration at a specific frequency. Delphos uses this technique to validate the cycle criterion. Moreover it takes advantage of a practical property of the log-scale wavelet TF image, namely that oscillations have the same frequency width whatever their frequency. The definition of an HFO in the TF image is an increase in the TF representation which is wide enough in time but has a limited spread in frequency. This definition enables Delphos to have a high and stable precision for all SNR. Two-stage detectors such as [Burnos et al., 2014] also use a TF analysis in the second stage to increase the precision on previously detected HFOs. This however only increases the precision but does not solve the issues, previously discussed in sections Robust estimation of the background activity and Accounting for the power spectrum, concerning the sensitivity of the first stage detector.

---

[5] Indeed, if the number of FP is constant across SNR, the precision increases with the SNR since the sensitivity and thus the number of TP also increase. This, however, means that, at SNR$=0$ dB, the threshold does not differentiate properly background from genuine HFOs or that the detector does not distinguish HFOs from high frequency-component of spikes. In other words, we believe that at low SNR it is better to detect nothing than to wrongly point toward an area where there is no real HFOs. This is especially true when the detector has difficulties distinguishing false-ripples due to spikes since spikes are usually found in a more extended area than the HFO and epileptogenic zone.

III.3.3. Clinical implication

To push forward HFOs into a clinical context, one should know the exact features and limitations of the detectors. Imagine a study which aims at mapping the rates of the Rs and FRs in various brain structures using a detector with a better sensitivity in the R band than in the FR band. Quantifying its performance using only visual marking would not have revealed this bias and thus the clinical conclusion could have been that structure *A* never exhibits FRs even thought it does. The lack of characterization could lead to contradictory results. Team *1* uses detector *A* while Team *2* uses detector *B* which have a better sensitivity in FR than in R. After running their detector on a large cohort of patients, Team *1* would conclude that Rs are betters marker of the epileptic zone than FRs whereas Team *2* would say the opposite. Both are right if you consider only statistical significance; but both are wrong because they did not take into account the detector biases.

Höller and collaborators [Höller et al., 2015] proposed to validate automated detectors against outcome instead of visual detection. In a scientific perspective, it seems to be an ambiguous idea. Assume that there is an HFO detector that succeeds in predicting the outcome. We would know which part of the brain to resect but we would not know why. We would have found the perfect outcome detector but not the perfect HFO detector. This detector could have detected “false” HFO [Bénar et al., 2010] resulting from very sharp spikes or oscillations contained in the semi-continuous high-frequency activity [Melani et al., 2013]. All these features could indeed be a manifestation of epileptogenesis but in the perspective of treating the patients with less invasive method, we would not know what to target... However, combining sensitive and precise detectors of all types of epileptic features to generate a meta-detector to predict the outcome could be an interesting goal. This way we could understand how the different features are related to epileptogenicity.

Concerning the separation between physiological and pathological HFOs, we did not address this delicate question because in our simulation we only used HFOs from epileptogenic regions which are likely to be pathological. We however believe that all HFOs have to be detected and, only in a second stage, classify as physiological or pathological using extra features, e.g. amplitude, number of cycles. The main issue concerning this classification is that no clear feature has been defined yet. Therefore this benchmark only assesses the performance of the detectors on their capability to detect HFOs, whether pathological or physiological, and does not presume to bind the performance of a detector with the surgical outcome of a patient.

III.4. Conclusion

The present paper describes a framework of simulations which shows the possibility of reproducing SEEG signal realistically in a controlled manner. By extracting elements of interest from their original BKG using DWT and by simulating their surrounding BKG using AR coefficients, we were able to reproduce signals from different brain areas while controlling the SNR of each added elements of interest separately. This control enabled us to study the performance and robustness of the detectors to several variations, i.e BKG variability, modification in amplitude and type of the elements.

Our simulation study has highlighted weaknesses and assets of several detectors that were never mentioned in previous studies as far as we know. They have pointed out:

1. the instability of the threshold to distinguish HFOs from background due to the unstable approach to estimate the background activity (section Robust estimation of the background activity).
2. the underestimated impact of the $1/\mathrm{f}$ spectrum (section Accounting for the power spectrum) which causes a difference in the class sensitivity,
3. the fuzzy definition of HFOs which leads to a decrease in precision (section Clear definition of HFOs).

To solve these issues we propose (1) to estimate the BKG activity by either detecting sections of baseline/BKG or by extracting its characteristics from the whole signal, (2) to take into account the shape of the spectrum by, for example, whitening the signal and finally (3) to use a strict definition of HFOs to avoid detecting events which look alike but originate from different entities.

We encouraged other groups to take part in the construction of a wider simulation set by either providing data of other brain areas and modalities or by generating their own simulation. This benchmark will be uploaded in open-access on the AnyWave (`http://meg.univ-amu.fr/wiki/AnyWave:Plugin_Simul)` website as well as the simulated SEEG data and the detection markers of the detectors. We believe that, by building such a broad and cross-modality simulation set, we will answer some unsolved questions concerning HFOs.

III.A. Supporting Information

Appendix A Let $x, \tilde{x}$ and Γ be the signal, the Fourier transform of x and the autocorrelation function of x, respectively. The power spectrum density (PSD) of x is defined as

$$\Sigma = |\tilde{x}|^2 . \tag{III.3}$$

The first stage of the SLL detector is a first-order backward differencing filter. We showed in [Roehri et al., 2016] that this process, which corresponds to a pre-whitening stage aiming at flattening the $1/\mathrm{f}$ spectrum, does not completely flatten the spectrum but does not over-express the high frequency. In fact, the algorithm of the SLL detector can be rewritten to highlight the fact that, indeed, the short line length metric correspond to a second differencing filter. Let x, x_1, x_2 and n be the original signal, the differentiated signal, the second order differentiated signal and the sample.

$$\begin{aligned} x_1[n] =& x[n] - x[n-1], \\ x_2[n] =& (x[n] - x[n-1]) - (x[n-1] - x[n-2]), \\ x_2[n] =& x_1[n] - x_1[n-1]. \end{aligned} \tag{III.4}$$

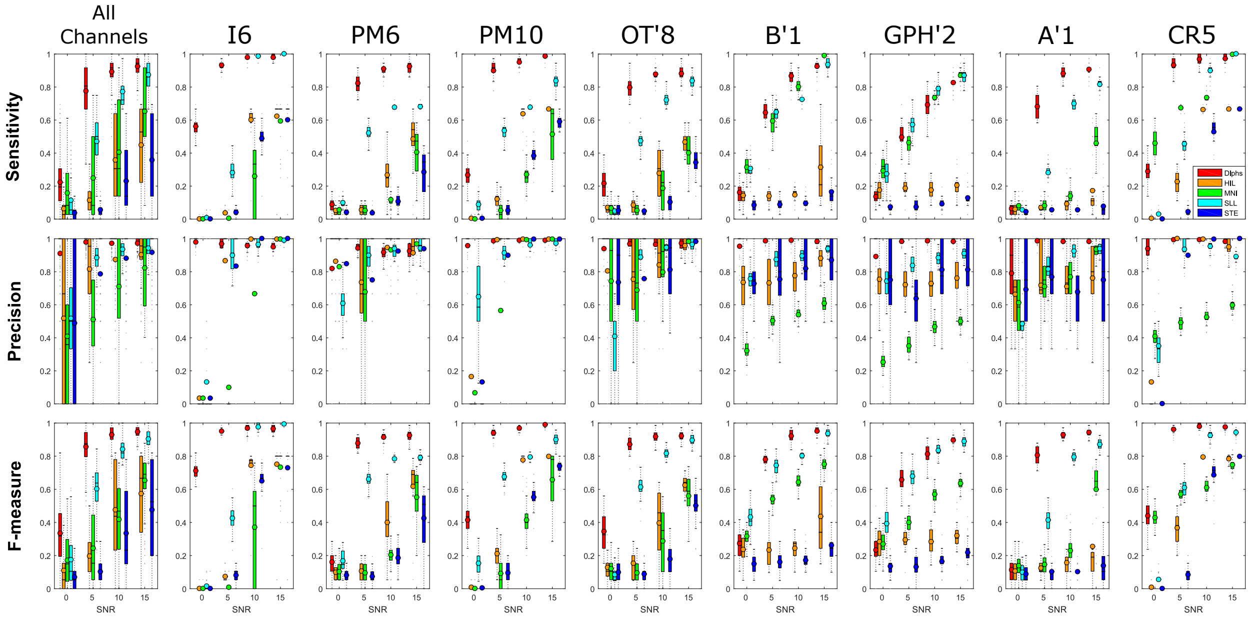


Figure III.11. – Boxplots of the sensitivity, precision and F-measure of the detectors ran over this benchmark for different SNRs. The first column is identical to Fig III.6 and shows the results for every channels and the other columns represent the result for each channel. The sensitivity of all detectors increases for each channels with the SNR. The STE and HIL has different behaviors for the group of channels I6,PM6, PM10, OT'8 and CR5 compared to B'1, GPH'2 and A'1. The MNI detectors switches into the "no baseline" mode for B'1, GPH'2 and CR5. Delphos and the SLL detectors have consistent behavior across channels.

The SLL energy expression for a time window W is

$$E_{SLL}[n] = \sum_{k=n-W+2}^{n} |x_1[k] - x_1[k-1]| \,, \tag{III.5}$$

and using III.4

$$E_{SLL}[n] = \sum_{k=n-W+2}^{n} |x_2[k]| \,. \tag{III.6}$$

We therefore derive that a first order differencing stage followed by a SLL energy calculation is proportional to a rectified moving average of a second order differencing signal. By applying a second order differencing, higher frequencies are more enhanced and the power spectrum almost inverses its slope. This explains the very high sensitivity in the fast-ripple band and poor sensitivity in the ripple band. Furthermore, the differencing method is dependent on the sampling frequency which appears clearly when we study the discrete Fourier transform

$$\mathrm{DTFT}\left\{x_2[\,.\,]\right\} = \left(1 - e^{-j\omega}\right)^2 \tilde{x}\left(e^{j\omega}\right), \tag{III.7}$$

where $\omega = 2\pi f / f_s$ is the normalized frequency and f_s the sampling frequency. Its PSD thus is

$$\begin{aligned} \left|\mathrm{DTFT}\left\{x_2[\ .\ ]\right\}\right|^2 &= \left|1 - e^{-j\omega}\right|^4 \Sigma\left(\omega\right) \\ &= 2(3 + \cos(2\omega) - 4\cos(\omega))\Sigma\left(\omega\right) \end{aligned} \tag{III.8}$$

The baseline detector is based on wavelet entropy (WE). The Continuous Wavelet Transform (CWT) of a signal y is defined as

$$T_y(b,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} y(t)\overline{\psi}\left(\frac{t-b}{a}\right) dt, \tag{III.9}$$

with ψ the wavelet function, a the scaling factor, and b the shifting factor. The wavelet entropy is defined as

$$S(b) = -\sum_{k \in A} P(k,b)\log_{10}(P(k,b)), \tag{III.10}$$

where A is the set of scales and P is the normalized wavelet power which is similar to a probability and corresponds to

$$P(a,b) = \frac{\left|T_y(a,b)\right|^2}{\sum\limits_{k \in A} \left|T_y(k,b)\right|^2}, \tag{III.11}$$

Let us show how the WE is dependent on the PSD and thus the profile of the $1/$f spectrum. Applying III.9 to the auto-correlation Γ as it is done in the MNI detector, we have

$$T_\Gamma(b,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \Gamma(t)\overline{\psi\left(\frac{t-b}{a}\right)} dt, \tag{III.12}$$

then using the Parseval theorem [Carmona et al., 1998] it yields

$$T_\Gamma(b,a) = \frac{1}{2\pi}\sqrt{a} \int_{-\infty}^{+\infty} \tilde{\Gamma}(\xi)\overline{\tilde{\psi}\left(a\xi\right)}e^{i\xi b} d\xi. \tag{III.13}$$

Finally, the Wiener-Khinchin theorem states that the Fourier transform of the auto-correlation function of a process x equal the PSD of the process x. Consequently,

$$T_\Gamma(b,a) = \frac{1}{2\pi}\sqrt{a} \int_{-\infty}^{+\infty} \Sigma(\xi)\overline{\tilde{\psi}\left(a\xi\right)}e^{i\xi b} d\xi. \tag{III.14}$$

Acknowledgments

This work has been carried out within the FHU EPINEXT with the support of the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the “Investissements d’Avenir” French Government program managed by the French National Research Agency (ANR). This study was also supported by the ANR projet FORCE (ANR-13-TECS-0013).

Closing remarks

In this article we have designed a method to extract spikes and HFOs which are further used to produce realistic simulations. This interface (Appendix B.2 Simulation Maker) is available online and could benefit other groups working on other modalities (e.g. ECoG, scalp-EEG, MEG). Delphos was shown to be very robust to many parameters. Most importantly, it is its constant precision (Eq. III.2) which ensures that its detections are reliable and could be used to map interictal activities.

Furthermore, the tested detectors (except Delphos) were designed to detect HFOs in the whole HFO band, i.e. 80-500 Hz. We thus decided not to examine the accuracy in classifying an HFO as ripple or fast-ripple. Here are however the results of Delphos in classifying the detection as spike, ripple and fast ripple (Fig III.12). *Note that the sensitivity of the spikes are constant across SNR because the channels were classed by the SNR of the ripples and the fast ripples and not the SNR of the spikes.* The sensitivity in classifying the ripples and the fast ripples increases and tends to one when the SNR of the oscillations augments. This is not surprising and is in line with the content of this chapter.

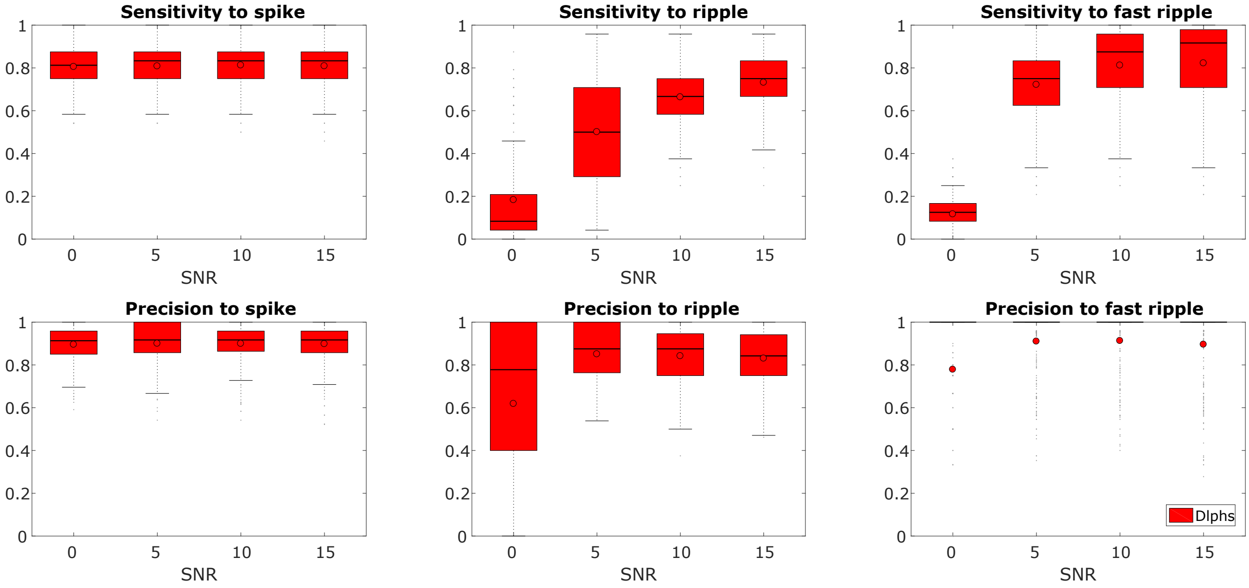


Figure III.12. – Performance of Delphos in classifying the different detections. *Note that the sensitivity of the spikes are constant across SNR because the channels were classed by the SNR of the ripples and the fast ripples and not the SNR of the spikes*

The median precision of each class is around 90%. On the one hand, the precision of classifying a fast ripple is almost always one for all SNRs. On the other hand, the precision of classifying a ripple is around 90%. This decrease in precision may be explained by the local interference of the wavelet coefficients (the cross-terms) between the spikes and the background or the spikes and the fast ripples. This decrease is not alarming since the precision is remains high. Similarly, cross-terms between co-occurring ripples and fast-ripples may generate ‘blobs’ with a large frequency spread. This probably causes the median precision in classifying a spike to be around 90%. Again, nothing alarming. Delphos thus seems suitable to map interictal activities.

This benchmark does not consider muscle artifacts nor spontaneous bursts of noise. Therefore, before any application and after the detection procedure, one has to verify that such artifacts were not present in the recording. In Delphos, there is a built-in function which discards any previously marked artifact (Appendix B.1.1 Delphos).

IV. Clinical Application

Roehri N. *et al.*, Ann. Neurol. Accepted

Preamble

We have previously designed and validated Delphos. The next step is to map the interictal activities and compare them to the EZ. In the next section, the ROC paradigm is used to quantify how well the interictal markers can predict the EZ. We also put particular effort in designing a patient-based statistical analysis which would reflect and benefit clinical needs. Such an analysis is of tremendous importance as clinicians face individual cases. Please note that we provide the SEEG location template with the reference name of each electrode in Appendix A Electrode Location Template.

High-frequency oscillations are not better biomarkers of epileptogenic tissues than spikes

Nicolas Roehri[1], Francesca Pizzo[1,2], Stanislas Lagarde[1,2], Isabelle Lambert[1,2], Anca Nica[3], Aileen McGonigal[1,2], Bernard Giusiano[1,4], Fabrice Bartolomei[1,2], and Christian-George Bénar[1]

[1] Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France
[2] APHM, Timone hospital, Clinical Neurophysiology, Marseille, France
[3] CHU Rennes, Neurology, Rennes, France
[4] APHM, Public Health Department, Marseille, France

Objective*: High-frequency oscillations (HFOs) in intracerebral EEG (stereoencephalography, SEEG) are considered as better biomarkers of epileptogenic tissues than spikes. How this can be applied at the patient level remains poorly understood. We investigated how well the HFOs and the spikes can predict epileptogenic regions with a large spatial sampling at the patient level. *Methods*: We analyzed non-REM sleep SEEG recordings sampled at 2048 Hz of thirty patients. Ripples (R, 80-250 Hz), fast ripples (FR, 250-500 Hz) and spikes were automatically detected. Rates of these markers and several combinations – spikes co-occurring with HFOs or FRs and cross rate (Spk$\otimes$HFO) – were compared to a quantified measure of the seizure onset zone (SOZ) by performing a receiver operating characteristic analysis for each patient individually. We used a Wilcoxon sign rank test corrected for false-discovery rate to assess whether a marker was better than the others for predicting the SOZ. *Results*: A total of 2930 channels was analyzed (median of 100 channels per patient). The HFOs or any of its variants were not statistically better than spikes. Only one feature, the cross-rate was better than all the other markers. Moreover, fast ripples, even though very specific, did not delineate all epileptogenic tissues. *Interpretation*: At the patient level, the performance of the HFOs is weakened by the presence of strong physiological HFO generators. Fast ripples are not sensitive enough to be the unique biomarker of epileptogenicity. Nevertheless, combining HFOs and spikes using our proposed measure –the cross rate– is a better strategy than using only one marker.

IV.1. Introduction

During presurgical examination of patients with drug-resistant epilepsies, clinicians face the difficult task of defining the epileptogenic zone (EZ), i.e. the subset of brain regions involved in generating seizures [Bancaud et al., 1965, Talairach and Bancaud, 1966, Kahane et al., 2006]. The ictal period is the privileged moment to delineate the EZ based on the regions involved at seizure onset. However, the non-negligible rate of failure in epilepsy surgery [West et al., 2015] has led to search for other electrophysiological criteria to delineate the EZ. On the one hand, methods have been proposed to better quantify cerebral activity at the onset of the seizure [Bartolomei et al., 2008, David et al., 2011]. On the other hand, renewed interest has been given to the interictal period in order to find putative biomarkers of epileptogenicity.

For the past decade, research on electrophysiological biomarkers has been fueled by the discovery of interictal high-frequency oscillations (HFOs, 80-500 Hz) [Urrestarazu et al., 2007, Staba et al., 2002]. These brief and small oscillations visible on intracranial EEG (iEEG) are considered strongly bound to the seizure onset zone (SOZ) [Jacobs et al., 2008, Crépon et al., 2010] and to correlate with surgical outcome [Haegelen et al., 2013, Wu et al., 2010, van't Klooster et al., 2015]. They are regarded as being more focal and specific than classical epileptic spikes (Spk) [Jacobs et al., 2008, Haegelen et al., 2013, Jacobs et al., 2009]. Indeed, interictal spikes are only partially concordant with the EZ, often found outside the EZ and thus lack specificity [Bartolomei et al., 2016]. However, not all HFOs are pathological. Even though fast ripples (FR, 250-500 Hz) seem to be always pathological [Menendez de la Prida et al., 2015], ripples (R, 80-250 Hz) are involved in physiological processes such as memory consolidation in the hippocampus [Girardeau et al., 2009, Alkawadri et al., 2014].

Most previous studies have drawn conclusions about HFOs at the group level. Indeed, the HFO rate is higher inside than outside the EZ when analyzing an entire group of patients, but this cannot be readily verified at the patient level due to the heterogeneity of patients (e.g. type of epilepsy, implantation sites, age) and to the presence of physiological HFO generators. It is therefore relevant to confront HFOs and spikes not only in epileptogenic areas but also in remote areas to better describe the extent of the zones defined by each marker, at the patient level. Because visual marking of HFOs is tedious, time-consuming, and results in a low inter-rater agreement [Menendez de la Prida et al., 2015, Spring et al., 2017], most articles comparing HFOs and spikes have studied few channels per patient. With the emergence of automated HFO detectors, the number of studied channels per patient increased but, since existing detectors were not designed to detect both HFOs and spikes, information about spikes were lost [Crépon et al., 2010, Burnos et al., 2014, Fedele et al., 2016].

We designed the present study to investigate the spatial extent of each marker at the individual patient level. We prospectively analyzed thirty patients with large spatial sampling from intracerebral SEEG electrodes (median number of contacts studied by patient was 100.5) and various type of focal epilepsies. We estimated the EZ using a quantitative method (Epileptogenicity Index [Bartolomei et al., 2008]), and applied an automatic detection technique for both HFOs and spikes [Roehri et al., 2016, Roehri et al., 2017]) on all channels. We computed several interictal markers (spikes; gamma oscillations (40-80 Hz), HFOs (80-500 Hz), Rs (80-250 Hz), FRs (250-500 Hz); HFOs co-occurring with spikes, FRs co-occurring with spikes, Spk $\otimes$ HFO). We then compared the performance of these interictal markers to delineate the EZ at the individual patient level.

IV.2. Methods

IV.2.1. Patient Selection

We studied forty all consecutive patients who had undergone stereoelectroencephalography (SEEG) exploration at 2048 Hz. Recordings were performed between May 2014 and January 2017, during pre-surgical evaluation of drug-resistant epilepsy. Before SEEG, a detailed evaluation was performed for each patient, including medical history, neurologic examination, neuropsychological assessment, scalp EEG recording, cerebral magnetic resonance imaging (MRI) and positron emission tomography (PET). As part of patients' usual clinical care, SEEG was carried out after the noninvasive phase. Recordings were performed using intracerebral macroelectrodes (10–15 contacts, length of the contact: 2 mm, diameter: 0.8 mm, 1.5 mm apart) placed according to Talairach's stereotactic method.

Since we aimed at comparing the fast oscillations and the spikes to the EZ, we selected patients fulfilling the following criteria: spontaneous seizures recorded with well-defined seizure onset[1], and without major artifact on the interictal recording (i.e. without high frequency noise nor saturation). Ten patients did not fulfill these criteria and were discarded; thirty patients were thus included for the rest of the study. From these patients, ten were operated upon, six are awaiting surgery or surgical decision, and the others (fourteen) were contraindicated (because of EZ overlapping with functional areas, too wide or multifocal). Therefore, the surgical outcome is available for only ten patients. This relatively low number of patients operated is in line with a recent study [Cloppenborg et al., 2016]. Table IV.1 provides clinical information about the selected patients. Brain surface meshes were made with Freesurfer [Fischl, 2012]. The electrode segmentation was done on a home-made software GARDEL [Medina et al., 2017].
The institutional review board of the French Institute of Health (IRB15226) approved this study and patient's consent was obtained.

IV.2.2. Recording methods

Signals were acquired on a 128 channel Deltamed™system. Ictal periods sampling rates ranged from 256 Hz up to 2048 Hz. The interictal sections were recorded during the night at a sampling frequency of 2048 Hz, at least 48 h after the day of implantation. All recordings were low-pass filtered at one third of their sampling frequency. We selected 5 minutes of non-REM sleep at stage N2 or N3 for detecting HFOs and spikes. N2 and N3 stages were defined by two neurologists (FP, IL) as sleep periods with clear presence of

---

[1] i.e. where a discharge can been seen and not only a slow pattern confirming that the SOZ was not missed.

Patient Index	Sex	Age at SEEG	Epilepsy onset (year)	Epilepsy duration (year)	MRI	Lateralisation	Epilepsy localization	SEEG SOZ	Surgery	Follow up (months)	Engel class
1	M	5	1	4	MCD	L	Insular	Insular	L insular cortectomy (second surgery, enlargment of the first surgery)	12 (2nd surgery)	IA
2	M	14	4	10		L	TLE	Temporal plus	Contraindication (extended network and functional key region)		
3	F	14	10	4	MCD	R	TLE	Ltle	Lesionectomy + R temporo-lateral cortectomy	10	IA
4	F	3	0,003	3	MCD	R	Posterior	Posterior	R temporo-parieto-occipital deconnection	10	IA
5	F	61	24	37	Scar	R	TLE	Mtle	Possibly mesial temporal lobe cortectomy		
6	M	54	36	18	Scar	L	TLE	Mltle	Contraindication (extended network)		
7	M	45	17	28		L	TLE	Mtle	Contraindication (extended network)		
8	M	34	17	17	MCD	R	TLE	Mtle	R anterio-temporal lobectomy	3	IA
9	M	27	15	12		R	FLE	Premotor	Contraindication (extended network and functional key region)		
10	M	39	0	39		L	TLE	Temporal plus	Contraindication (extended network)		
11	M	16	12	4	NDT	L	Posterior	Posterior	Lesionectomy (14/11/16)	6	IA
12	F	17	17	0	HS	Bilateral	TLE	Mtle	Contraindication (bilateral)		
13	M	32	1	31		R	FLE	Premotor	Contraindication (extended network)		
14	M	19	14	5		Bilateral	Multifocal	Multifocal	Contraindication (extended network)		
15	F	15	7	8	MCD	L	TLE	Ltle	Containdication (language deficency)		
16	F	29	19	10	MCD	Bilateral	TLE	Mtle	Contraindication (bilateral)		
17	M	17	1	16		R	Insular	Insulo-opercular	Multiple surgery and GK in the insulo-opercular region	110	IV
18	F	21	3	18	MCD	L	FLE	Premotor	Lesionectomy	4	IV
19	M	6	0,25	6	MCD	R	FLE	Premotor	R frontal lobectomy and callostomy (2nd surgery)	11	III
20	F	16	3	13	MCD	R	Insular	Insular	Contraindication (extended network 2nd SEEG in program for thermocoagulation)		
21	M	47	36	11	MCD	L	FLE	Frontal	Lesionectomy+cortectomy (awaiting surgery)		
22	M	19	2	17	MCD	L	FLE	Orbitofrontal	L ventro orbito frontal cortectomy	3	IA
23	M	13	3	10		L	FLE	Premotor	Contraindication (functional deficit)		
24	F	34	32	2		R	TLE	Mtle	Indication to anterior temporal lobectomy (awaiting surgery)		
25	M	14	10	4	NDT	L	TLE	Mtle	L anterio-temporal lobectomy	3	IA
26	M	59	44	15		L	FLE	Bi-frontal	Contraindication (bilateral bifrontal)		
27	F	46	38	8		L	TLE	Temporal plus	Awaiting surgical decision		
28	M	21	0	21	MCD	L	Posterior	Parietal epilepsy	Cortectomy (awaiting surgery)		
29	F	15	12	3		R	Insular	Insular	Contraindication (extended network)		
30	M	43	26	17		L	TLE	Mtle	Awaiting surgical decision		

TLE: Temporal Lobe Epilepsy, FLE: Frontal Lobe Epilepsy, LTLE: Lateral TLE, MTLE: Medial TLE; Mesial-Lateral TLE
MCD: malformation of cortical development, NDT: neurodevelopmental tumors, HS: hippocampus sclerosis.

Table IV.1. – Patients' characteristics spindles or slow waves. This 5 minutes duration was shown to be sufficient to study HFO even though these rates may vary over time [Zelmann et al., 2009]. The analysis was performed in a bipolar montage.

IV.2.3. Mesial and lesional contact localization

Contacts in each selected brain area were labeled by two neurologists (FP, SL). We particularly focused on regions in the mesial temporal lobe (amygdala (Am), hippocampus (Hip) and mesial temporal pole (MTP)). The classification was done using a computed tomography (CT) scan obtained directly after SEEG implantation and an MRI obtained before implantation. Images were co-registered to the same space. The same procedure was used to label contacts in the lesion in patients with malformation of cortical development (MCD), neurodevelopmental tumors (NDT), hippocampus sclerosis (HS) or scar.

IV.2.4. Quantification of the SOZ

Several seizures were studied for each patient. The SOZ was visually defined by two expert neurophysiologists (FB, FP) for each seizure. It was defined as the group of channels that were involved at seizure onset for at least one seizure. We also calculated on each seizure the Epileptogenicity Index (EI) [Bartolomei et al., 2008], a semi-automatic procedure that aims at quantifying the SOZ. The EI ranks brain structures according to the tonicity of the fast-discharge and the delay of involvement of the structure at seizure onset (Fig IV.1 A). Its values range from 0 to 1, with 1 corresponding to the most epileptogenic region. The EI method was shown to be relevant to quantify the epileptogenicity of brain structures [Aubert et al., 2009, Bartolomei et al., 2017b]. To compute the EI, we used the plugin designed for the open-source software AnyWave [Colombet et al., 2015], available at http://meg.univ-amu.fr. To have one EI value for each channel, we took the maximum EI of a given channel across seizures. Channels exhibiting EI values above 0.2-0.3 are said to be epileptogenic [Bartolomei et al., 2008, Bartolomei et al., 2016, Gollwitzer et al., 2016, Marchi et al., 2016, Bartolomei et al., 2010]. We set the cutoff at 0.25 in order to have a median correspondence per patients between the EIZ and the SOZ of approximately 50%. The binarized EI gives the Epileptogenicity Index Zone (EIZ), which is an estimation of the extent of the EZ. We ensured that the EIZ was included in the visually defined SOZ.

The definition of the EZ used in the current study is different from the one of Lüders: “the minimum amount of cortex that must be resected (inactivated or completely disconnected) to produce seizure freedom” [Lüders et al., 2006]. We used the Bancaud and Talairach view on the EZ which is the site of the beginning and of the primary organization of the seizure [Bancaud et al., 1965, Talairach and Bancaud, 1966, Kahane et al., 2006] (independently of the potential surgery). Notably, both definitions are theoretical since the minimum amount is not known for the former and the outcome is not available for the latter (the localization of the EZ is less certain but it permits however to study contraindicated patients).

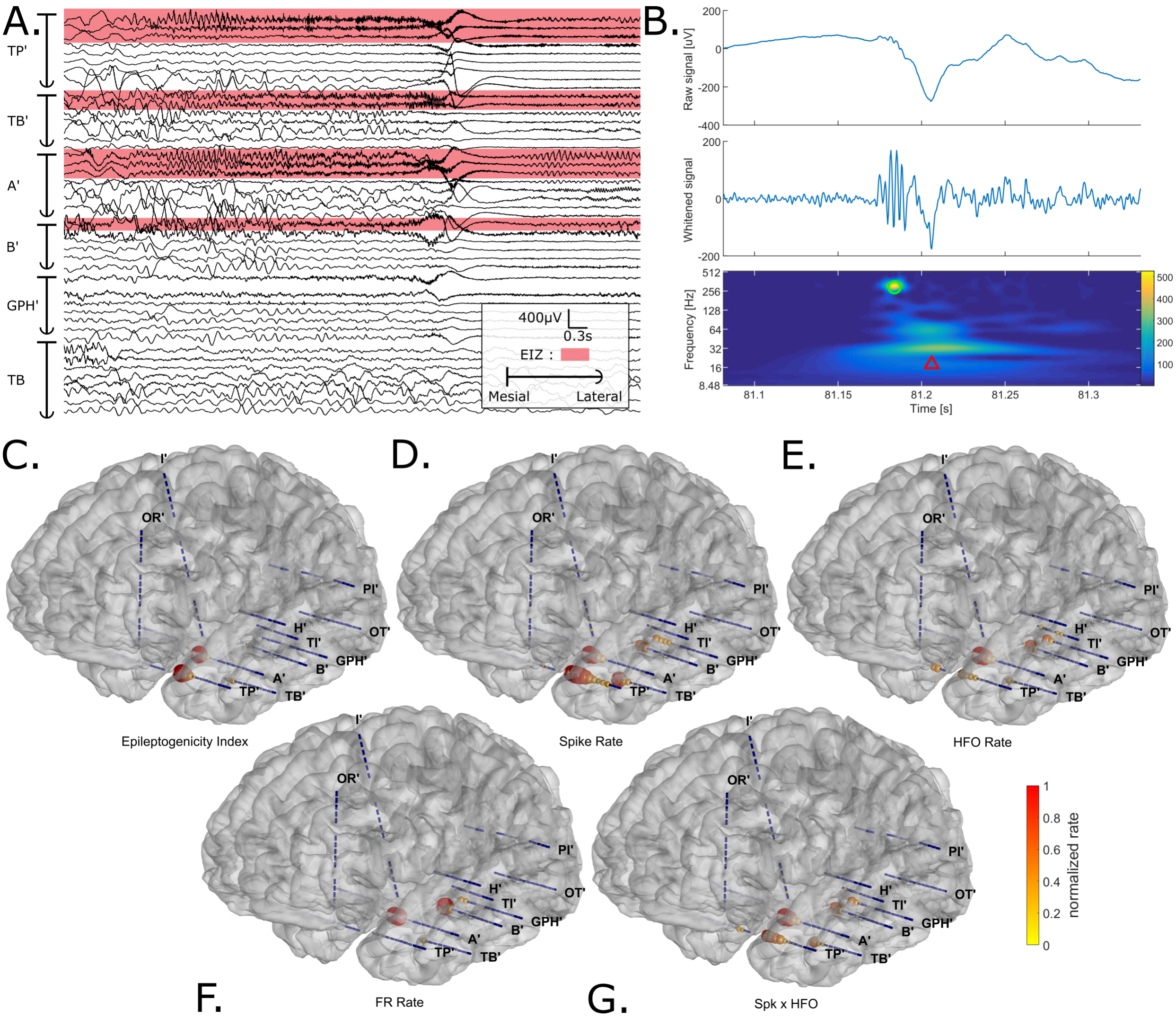


Figure IV.1. – Illustration of the analysis pipeline for Patient 30. (A) SEEG recording of a seizure in bipolar montage. Electrode TP' explored the temporal pole; TB' was placed in the temporo-basal area (lateral contacts) and in the entorhinal cortex (mesial contacts); electrodes A', B' and, GPH' targeted the amygdala, the hippocampus, and the parahippocampal gyrus respectively (mesial contacts) and recorded the middle temporal gyrus as well (lateral contacts). The apostrophe indicates the left hemisphere. The channels highlighted in red correspond to the channels exhibiting an EI value above 0.25. These channels form the Epileptogenicity Index Zone (EIZ) of the patient, which is an estimation of the Epileptogenic Zone. (B) Example of an interictal fast ripple riding an epileptic spike detected by Delphos. From top to bottom, the panels display the raw signal, the whitened signal and the normalized time-frequency image. The green circle represents a detected oscillation (here, a fast ripple) and the red triangle a detected spike. (C-G) Mapping of the different biomarkers in the patient's brain surface. The sphere color and diameter correspond to either the Epileptogenicity Index or the normalized rate of the marker (ranging from 0 to 1). One can see that each marker seems to delineate the EIZ. The irritative zone is broad but the channels with the highest spike rate correspond well to the EIZ (TP', TB', A', and B' inner contacts). The HFO rate and FR rate seem to indicate more the amygdala and the hippocampus but have a lower rate in TP' and TB'. Spk $\otimes$ HFO seems to be the best compromise in term of extent and precision.

IV.2.5. Automated detection

HFOs and spikes were automatically detected in AnyWave using Delphos (Detector of ElectroPhysiological Oscillations and Spikes) [Roehri et al., 2016]. Delphos was designed to solve the "false-ripple" issue due to filtering artifacts [Bénar et al., 2010] and to optimize the signal to noise ratio at each frequency. It does so by detecting events of interest above a threshold in the Z_{H_0}-normalized time-frequency (TF) image [Roehri et al., 2016] and by measuring the time width and frequency spread of the detected islands (Fig IV.1 B). The algorithm classifies the detections as "oscillation" or "spike" according to the two aforementioned measures. Moreover, this method allows us to detect events occurring simultaneously and to give a frequency label to the detected oscillations. This detector was tested against a benchmark of realistic simulations [Roehri et al., 2017]. These simulations were obtained by inserting visually marked HFOs and spikes from three patients (from the current pool) into the simulated background of their respective brain region (for more details see Methods of [Roehri et al., 2017]). This allowed us to study the performance of the detection knowing the true occurrence of the HFOs and test the variability across different brain areas. Delphos showed constant high precision, i.e. the detections were almost always correct, and the highest sensitivity compared to the 4 other detectors from the Ripplelab Toolbox [Navarrete et al., 2016]. Delphos showed less imbalance toward a class of HFO compared to the other detectors, i.e. it detects similarly well Rs and FRs, and robustness across brain regions and across variations in brain activities.

These characteristics enabled us to study oscillations in the Gamma (γ), HFO, R and FR bands as well as spikes. Moreover, we could study HFO or FR superimposed on spikes. Two events were defined as co-occurring if their detection times were separated by less than 100 ms. For each channel, we computed the rate per minute for every marker. We studied 8 markers: spikes, gamma oscillations, HFO, R, FR, spike co-occurring with HFOs (Spk-HFO), or with FRs (Spk-FR), and a combination of spikes and HFOs (Spk $\otimes$ HFO). Spk $\otimes$ HFO corresponds to the geometric mean of the spike and HFO rate obtained by calculating the square root of the product of the two rates; $\text{Spk} \times \text{HFO} = \sqrt{\text{spike rate} \times \text{HFO rate}}$. This measure permits us to explore another approach of combining the two markers, less restrictive than the co-occurrence. Before and after the detector was run, we visually verified for muscle artifact and high frequency noise and removed these sections if necessary.

IV.2.6. Statistical Analysis

Statistical analyses were applied to assess how well the rates of the different markers can delineate the EIZ for each patient. The Receiver Operating Characteristic (ROC) framework [Fawcett, 2006] aims at computing the performance of binary classifiers ("EIZ" vs "no EIZ" based on event rates) across a range of thresholds. The ROC curve is obtained by plotting the sensitivity (Sens), i.e. the proportion of correctly labeled epileptic channels, as a function of the specificity (Spe), i.e. the proportion of correctly labeled non-epileptic channels. To generate the ROC curve, we compared the EIZ to the rates of the different markers at several thresholds.

Firstly, we calculated the ROC curve on the entire group of patients to represent the global behavior of the markers compared to epileptic regions. Plotting this curve reveals whether the markers span the whole ROC space, or if there is a plateau indicating a lack of specificity or sensitivity. Drawing further conclusions about their performance using this curve would be misleading [Fawcett, 2006]; a measure of variance (based on the patient) is needed to compare the markers.

We studied their performance for each patient individually to determine whether a given marker was in general better than the others. We chose to study the ROC curve between 85% and 100% of specificity. As a global performance measure, the partial Area Under the Curve of this ROC curve (pAUC) was used. The pAUC is especially useful when one can assume that the curves could cross [Ma et al., 2015] (spikes are said to be more sensitive than specific and inversely for the HFOs) and because the proportion of epileptic channels in our study is quite low compared to "normal" channels (298/2930) [Fawcett, 2006]. This is also justified in a clinical perspective, because it is more important to minimize the number of wrongly labeled normal channels (false positives). We divided the partial AUC pAUC value by the maximum possible value area ($\text{Sens}_{\text{max}} \times (1 - \text{Spe}_{\text{min}}) = 1 \times 0.15 = 0.15$) to have an index ranging from 0 to 1. Normally, the AUC of chance level is 0.5; here, it is $0.15^2/2 \times 0.15 = 0.075$ after normalization (considering the partial ROCcurve). Figure IV.2 illustrates the definition and transform of the pAUC. This thus yield one pAUC for each patient and each marker.

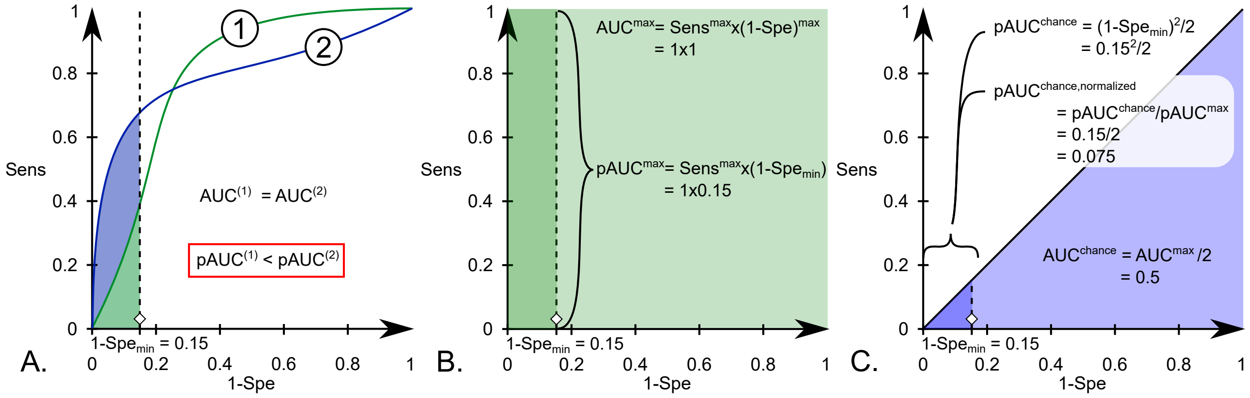


Figure IV.2. – A. strength of the pAUC in the current context for crossing ROC curves; B. calculation of the maximal AUC for the full and partial AUC; C. calculation of the AUC of chance level for the full and partial AUC.

We applied a Wilcoxon signed rank test (which is a paired test) to the pAUC of all pairs of markers across patients and corrected for multiple comparisons (False Discovery Rate [Benjamini and Hochberg, 1995] below 0.05 corrected for 28 comparisons). As these statistics approximates normal distribution for number of samples above 15, the p-values were transformed into z-values to capture the two-tailed test. In other words, a test is significant if $|z| \geq 2.63$ after correction.

Secondly, we investigated the rates of some markers in mesial structures and lesions at the group and patient level. We studied channels inside the hippocampus (Hip), the amygdala (Am) and the lesions (Les) (all three are known to exhibit high rate of HFOs and spikes), and compared them to the rest of the channels, which corresponds to all the contacts which were not inside the aforementioned regions (Oth). For the group level analysis, the statistical significance was assessed by computing a Bonferroni corrected Mann-Whitney U test (p below 0.05 corrected for 24 comparisons).

IV.3. Results

IV.3.1. Patients' characteristics

Thirty patients were studied; 11 women and 19 men. The median duration of epilepsy was 10.5 years [interquartile interval, IQ: 4.25-17 years]. There were 14 patients with TLE (8 MTLE, 3 temporal plus, 2 LTLE, 1 MLTLE), 8 with FLE, 4 insular epilepsies and 3 posterior epilepsies. The etiologies of the epilepsy were structural (12 MCDs, 2 NDT, 2 Scars (post-traumatic and post-encephalitis), 1 HS) and unknown (13 MRI negative patients).

The total number of analyzed channels was 2930 with a median number per patient of 100.5 [IQ 89-109]. From these channels, 298 were inside the EIZ. The median percentage per patient of channels inside both the EIZ and the SOZ was 48.5% [32.1%-60%]. The median percentage per patient of channels with spikes outside the EIZ was 77.9% [69.4-83.3%]; for the HFOs it was 63.4% [48.8-74.3%]; for the FRs it was 13.7% [9.78-18.5%] and for channels outside the EIZ with HFO and Spk it was 60.7% [43.1-69.4%].

The first question was to determine whether the rates of the markers differed between inside and outside the Epileptogenic Index Zone (EIZ and NEIZ). The rates of spikes, ripples and fast ripples were all significantly higher (***) inside the EIZ than outside when taking all channels (Fig IV.3). Similarly, the rates were higher (**) in the hippocampus inside the EIZ than in the hippocampus outside the EIZ. (NEIZ). The rest of channels (Oth), which were not in Hip, Am nor Les, also showed higher rates inside than outside the EIZ (***). This was also the case for the fast ripple rate in the Am (*) and the Les (**) but not for the spike and ripple rates.

The second test aims at showing whether the rates in the Hip, Am and Les outside the EIZ are different from the rates inside the EIZ in the other regions (Oth). Only two tests were significant (¤¤¤), i.e. the ripple rate in the amygdalae outside the EIZ and the spike rates in the lesion outside the EIZ (here the median was higher in the NEIZ lesion than in the rest of the epileptogenic regions). The fact that the ripple rate can be statistically significant between the Am (NEIZ) and the Oth (EIZ) but not between the Am (NEIZ) and the Am (EIZ) is likely because the variance of the rate in the Oth (EIZ) is smaller since there are many more channels in the Oth (EIZ) than in the Am (EIZ).

IV.3.2. Markers vs. EIZ at group level

In this section, we evaluated how the markers globally behave in term of sensitivity and specificity when taking the whole group of patients (Fig IV.4). The combinations of spikes

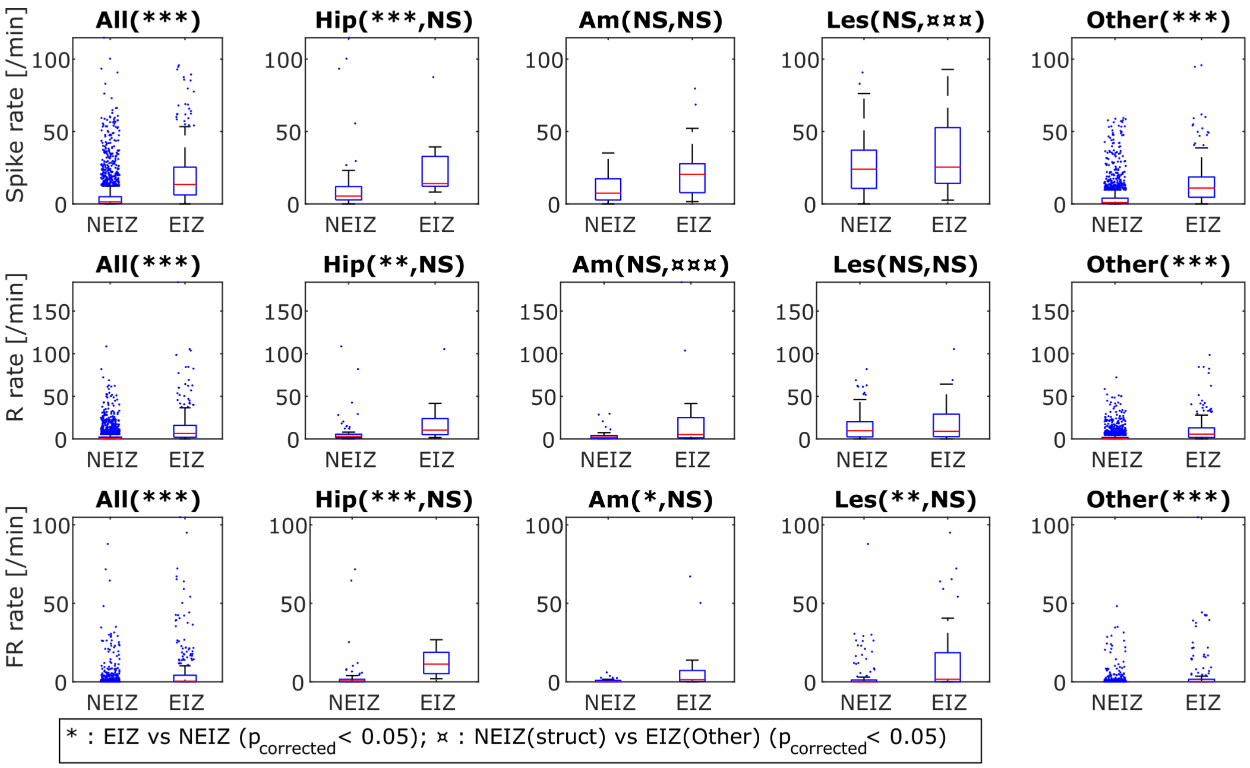


Figure IV.3. – Boxplots of the rate inside the EIZ or outside (NEIZ) in the whole brain and in specific structures. The spike rate, ripple (R) rate and fast ripple (FR) rate are given in the first, second and last line respectively. The first to the fifth column correspond to the rate in the whole brain (All), the hippocampi (Hip), the amygdalae (Am), the lesions (Les) and in the rest of the brain excluding the previous regions (Oth). Two statistical tests were applied. The first one tested the difference in the rate inside the EIZ compared to the NEIZ within each group (e.g. All, Hip, Am) for each marker. The asterisk (*) shows the differences which were statistically significant after correction for 24 comparisons (*: corrected p-value ≤ 0.05, **: $p \leq 0.01$, ***: $p \leq 0.001$, NS: not significant). The second test investigates the differences between the rate outside the EIZ in the Hip, Am and Les compared to the rate inside the EIZ in the other regions (Oth). This symbol ¤ shows the differences which were statistically significant after correction for 24 comparisons (¤: corrected p-value ≤ 0.05, ¤¤: $p \leq 0.01$, ¤¤¤: $p \leq 0.001$, NS: not significant). The rate in the EIZ is higher than in the NEIZ in the whole brain and within each group for every marker except for the ripple and spike rate in the Am and Les. There are almost no statistical differences between the rate in the NEIZ of the Hip, Am, and Les and the rate inside the EIZ in the rest of the brain. The spike rate is higher in the NEIZ of the Les compared to the spike rate in the EIZ in Oth and the ripple rate is lower in the NEIZ of the Am compared to the ripple rate in the EIZ in Oth.

and HFOs have higher sensitivity for specificity below 90% and perform better globally (higher AUC, Fig IV.4 A,B,D). FR and Spk-FR have higher sensitivity at higher specificity (95%, Fig IV.4 C). The ROC curve (Fig IV.4 E) clearly shows a plateau for the FRs and Spk-FRs; this means that these events were not present in ≈35% and ≈45% of the epileptic channels respectively. A similar plateau can be observed for the Spk-HFO at ≈90% of sensitivity.

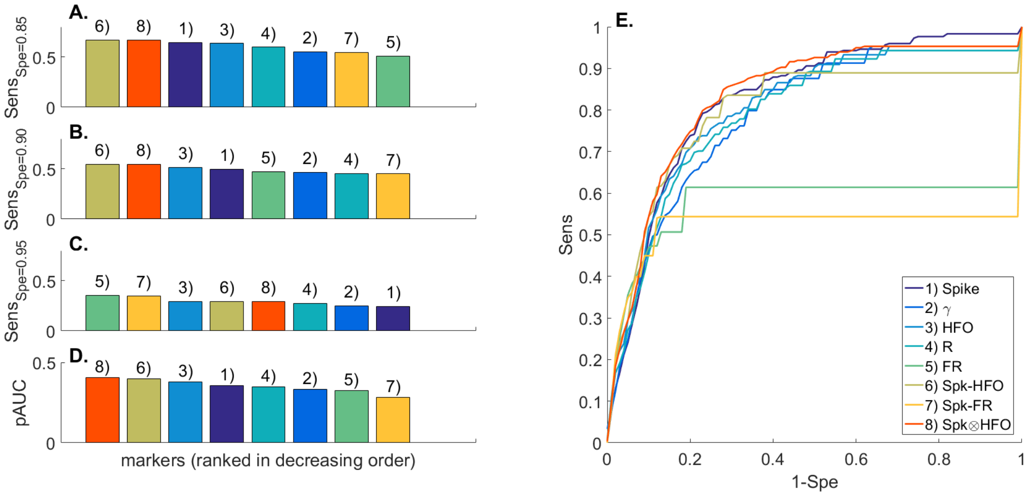


Figure IV.4. – ROC curve for the whole group of patients. (A-C) These panels give the sensitivity of each marker to discriminate EIZ from NEIZ channels at given specificities (0.85, 0.90, and 0.95, respectively) for the whole group of patients. (D) The area under the curve calculated between 0.85 and 1 of specificity is drawn here (pAUC). (E) This panel plots the ROC curves for every marker. Note the plateau for the fast ripples and spikes with fast ripples.

IV.3.3. Markers vs. EIZ at patient level

Figure IV.5 shows examples of different results. Patient 22 is a concordant case in whom each marker points to the orbitofrontal gyrus, in agreement with the EI values. For patient 11, the interictal markers only partially describe the EIZ. The spikes miss the inner contacts of PFG' but are present in both the hippocampus (B') and the temporo-basal electrode (TB'). The ripple rate is highest in the regions of the occipital and parietal lobe (CU' and PA' respectively). The fast-ripples are only present in the amygdala (A') and are not seen in the electrode exploring the lesion (L'). Spk $\otimes$ HFO is low in the PFG' electrode and still keeps a relative high value in the CU' but is negligible in TB' and PA'. For patient 3, spikes and ripples globally map the EIZ but are present in the hippocampus (B) electrode and the amygdala (A) respectively. Fast ripples are only detected in the dysplasia (DYS). Spk $\otimes$ HFO delineates well the EIZ by combining the ripple and spike map.

Here, the investigation concerns the comparison of the markers within each patient. Partial ROC curves were drawn for each patient individually between 85% and 100% of specificity and we calculated the pAUC (Fig IV.6 A). The box corresponding to the FRs and Spk-FR are the lowest. Only Spk $\otimes$ HFO stands out from other measures. The result of the statistical paired test (Fig IV.6 B) indicates that there is no statistical evidence showing that spikes, gamma oscillations, HFOs, Rs and Spk-HFO are generally different from one another. All markers are however better than the FRs and Spk-FR except gamma oscillations. One marker is better than all the others: Spk $\otimes$ HFO (illustrated by the complete yellow line with asterisks, p below 0.05 after FDR correction).

Figure IV.7 gives a visual description of the results of the statistical test. Note that the pAUC at chance level is 0.075. One can see that the performance of Spk $\otimes$ HFO is glob-

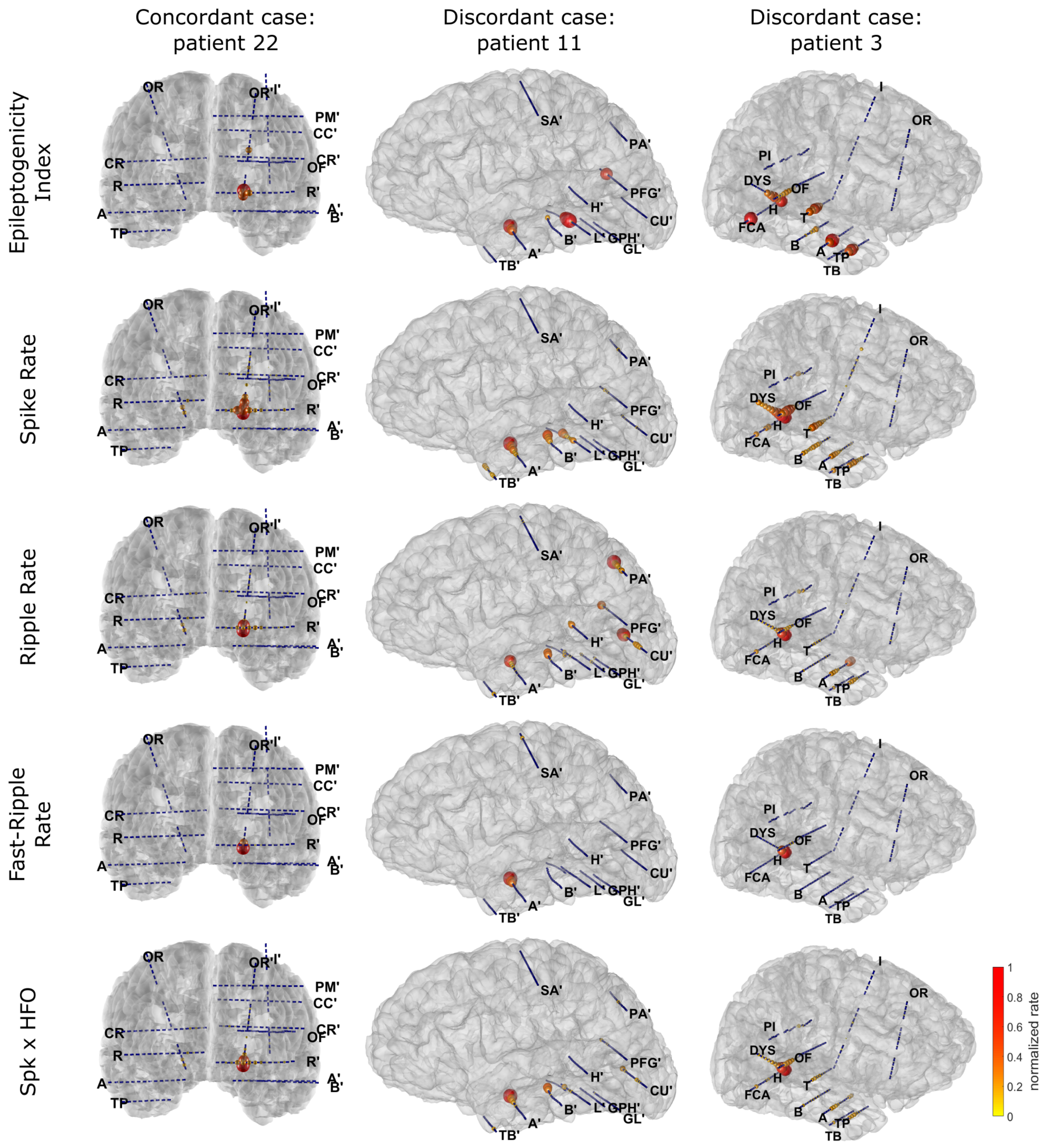


Figure IV.5. – Illustration of the mapping of 4 interictal markers and the EI values for 3 patients. Each column illustrates one of the three patients and each line the following markers: EI, spike, ripple, fast-ripple, Spk $\otimes$ HFO. To avoid being blinded by region exhibiting large amount of a marker (e.g. dysplasia (DYS)), we decided to represent the normalized value of the square root of the rate divided by the maximum value for each patient and each marker. For the first case (patient 22), every marker points toward the orbitofrontal gyrus (OR'). The markers are discordant for the two other cases. [...]

For the patient 11, the spikes are frequent in the hippocampus (B') and in the temporo-basal area (TB') which are not in the EIZ. They are however few in PFG' belonging to the EIZ. Similarly, the ripple rate is high in the inner contacts of CU' and PA' which are outside the EIZ. The fast ripples are only present in the amygdala. The Spk ⊗ HFO values are similar to the one of the spikes but do not manage to completely remove the impact of the CU' electrode. The effects of PA' and TB' are however reduced. In that case, Spk ⊗ HFO is better than the HFOs but slightly less good than the spikes (Fig IV.7). For patient 3, the ripple and spike values are similar to the EI values. The spikes are however present in the electrode B but not the ripples; the ripples are present in the amygdala whereas the spikes are absent from this structure. Spk ⊗ HFO is better than both the ripples and spikes (Fig IV.7) thanks to the complementarity of the HFOs and spikes. The fast-ripples are mainly in the dysplasia (DYS).

ally above the performances of the other markers for each patient. When Spk ⊗ HFO does not perform best for a patient, it is at least in the 3 first best markers. Note the variable performances of spikes and HFOs; spikes perform better than HFOs for some patients (1,3,4,6-9,11,16,18,21-28,30) and, for some others, HFOs perform better (2,5,10,12-15,17,19,20,29). All these results were similar when changing the threshold of the EIZ (0.2-0.3) and the minimum specificity (80-90%) (Supplementary Figure IV.8).

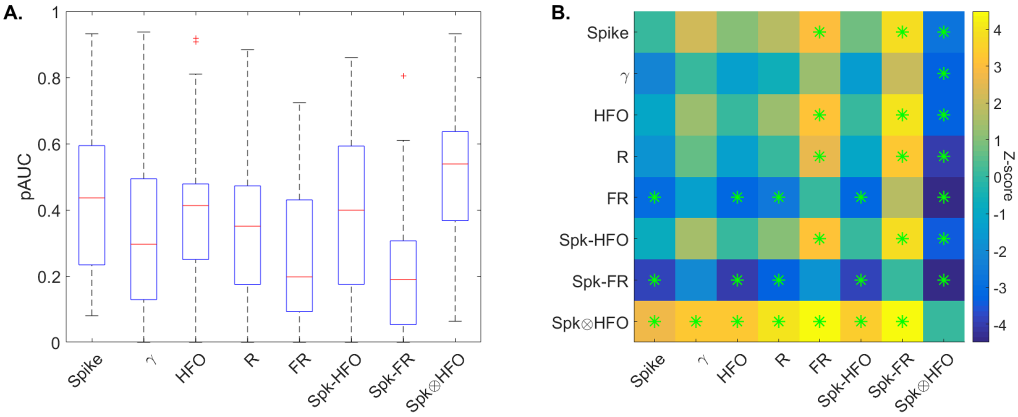


Figure IV.6. – Boxplot of the partial AUC of each marker beside its significance test. (A) The results of the partial AUC calculated for each patient separately is given as boxplots for each marker. The variance is high for each marker due to patient variability but Spk ⊗ HFO is the only measure to stands out of the others. (B) The matrix shows the z-values of the Wilcoxon signed rank test for each pair of markers. It reads from line to column. For instance, the tile of the first line and the second column means that spikes are better than gamma oscillations with Z≈2. The green asterisks show statistical significance below 0.05 after FDR correction (28 comparisons). There is no statistical evidence showing that HFOs or any subgroup are better predictors than the spikes. The yellow line of the Spk ⊗ HFO means that this feature is in general better than the other marker for each patient.

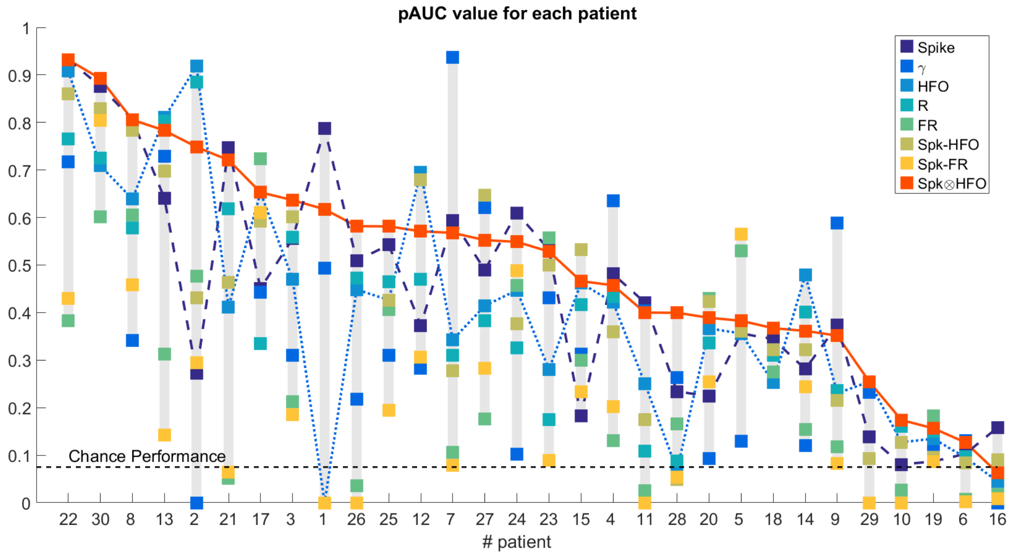


Figure IV.7. – Partial AUC for each patient, ranked in decreasing order of the Spk ⊗ HFO performance. The solid, dotted, and dashed line correspond to the performance of the Spk ⊗ HFO, HFOs, and the spikes respectively for each patient. The Spk ⊗ HFO is almost always higher than the other markers as it was shown in Fig IV.6. Note the alternating performance of the spikes and HFOs. The performance at chance level is 0.075 due to partial AUC analysis and normalization (see Methods).

IV.4. Discussion

IV.4.1. HFOs are not better than spikes

The main result of this study is the absence of statistical evidence showing that HFOs or its variants (R, FR, Spk-HFO, Spk-FR) are globally better than epileptic spikes to delineate the EIZ at for every patient. Regarding figure IV.7, for some patients, HFOs are better than spikes, whereas spikes are better than HFOs for others. This explains the absence of statistical evidence. In previous articles studying HFOs, little spatial information was available about the extent of the HFO zone [Crépon et al., 2010]. In this study, we emphasized the spatial investigation of the zone mapped by each marker by studying every channel available without a priori selection. The variable performance of HFOs and spikes is most likely due to the detection of physiological ripples in some regions and to the large distribution of the spikes (Fig IV.3 and IV.5) respectively. It was possible to differentiate epileptogenic areas from non-epileptogenic ones at the group level; this is even possible at the sublobar level for hippocampi and amygdalae (Fig IV.3). It is however more difficult, already at the group level, to separate epileptogenic structures in other regions (Oth) from non-epileptogenic temporo-mesial structures and lesions (Fig IV.3).

We focused on the ripple band to investigate the impact of physiological HFOs. We defined ripples outside the EIZ as physiological ripples. This type of ripples occurred more frequently in some regions, such as the occipital lobes or the mesial temporal regions. Four patients (4,8,11,28) were explored in the occipital lobe which was diagnosed as non-epileptogenic after SEEG. For these patients, spikes were better predictors than ripples. Alkawadri and colleagues [Alkawadri et al., 2014] showed that the rate of physiological ripples was high in the occipital region. In our study, the threshold had to be high enough to overcome the high rate of physiological ripples in order to achieve high specificity, which resulted in lower sensitivity. Our patients however did not have many spikes within occipital lobes (Fig IV.5). We also found high rates of ripples in hippocampi which were not epileptogenic. The hippocampus is known to produce physiological ripples [Girardeau et al., 2009, Csicsvari et al., 1999]. Addressing the issue of physiological and pathological ripples in order to identify the epileptogenic zone cannot be solved by looking at the overall rate of ripples, especially when regions being known as generators of physiological ripples such as the occipital region and the hippocampus are being recorded. Spikes also exhibit a high rate in the hippocampi, amygdalae and lesions, and therefore the same conclusion could be made. We did not find any relationship with the localization of the SOZ (Table IV.1) explaining the variable performance of spikes and HFOs. When ripples perform better than spikes for a given patient in terms of predicting the EIZ, it is because the spike rate in the NEIZ is too high compared to the rate in the EIZ, whereas in these channels the ripple rate is not necessarily higher in the NEIZ than in the EIZ, and inversely for spikes overcoming ripples (Fig IV.5, IV.6 and IV.7).

A rule that could be drawn with these results is that tissues are unlikely to belong to the EIZ if they produce a high rate of only one of the two markers. This translates to Spk $\otimes$ HFO and not Spk+HFO. By multiplying the two rates, we highlight the channels exhibiting high rate of both markers and attenuate those with few or none of at least one marker. Spk-HFO (HFOs occurring together with spikes) is also a combination of the two markers, but it is more constrained since both events have to occur together within a short time window and is therefore less sensitive. We acknowledge that detecting HFOs superimposed on spikes is challenging especially when these events overlap both in time and frequency [Amiri et al., 2015], which could require further developments such as spike removal [Jmail et al., 2016]. We showed in a previous study [Roehri et al., 2017] that Delphos was sensitive enough to have a good estimation of the rate of spikes co-occurring with HFOs. We thus believe that the low performance cannot only be explained by this lower detection sensitivity but more by the fact that spikes without HFOs and HFOs without spikes can be pathological. Concerning the gamma oscillations, they were the only marker which was not statistically better than the FR. Note however that the rate of gamma was the best predictor in three patients (4,7,9). Combining the three markers may be a possible perspective.

IV.4.2. Low sensitivity of FRs and Spk-FR

We found that fast ripples lack sensitivity in delineating the EIZ. In our study, FRs and Spk-FR were very specific markers, as they were found only in about 14% of channels outside the EIZ whereas the HFOs and spikes were found in more than 50%. The median rate of the FR and Spk-FR outside the EIZ is null (Fig IV.3). They however do not perform well under the pAUC criterion. This poor performance seems to go against previous publications [Jacobs et al., 2008, van't Klooster et al., 2015]. In [Jacobs et al., 2008], FR and Spk-FR had the best sensitivity compared to spikes, R and spikes co-occuring with ripples at 95% of specificity at the group level (12 patients). Interestingly, this is also what we found (Fig IV.4 C). The fact that this occurs at exactly the same specificity (95%) may be a coincidence because the ratio between epileptogenic and non-epileptogenic channels is different in the two studies. This conclusion highly depends on the value of studied specificities (Fig IV.4 A-B). If 90% or 85% specificity had rather been chosen, the Spk-HFO and Spk $\otimes$ HFO would have had higher sensitivity than FRs and Spk-FR. This strengthens our choice of using the partial area under the curve as the performance criterion instead of sensitivity for a given specificity.

The plateau seen on Fig IV.4 E (when analyzing the whole group of patients) could be explained by the fact that FRs rarely occur in neocortical epilepsies [Jacobs et al., 2008, Crépon et al., 2010]. Jacobs and colleagues found that at 95% specificity, spikes were better predictors for neocortical epilepsies (for 2/3 of the patients). This, however, does not explain the poor performance at the patient level because we would expect the FR class to achieve high performances for some patients and poor for some others (Fig IV.6 and IV.7). We did not find any statistical evidence when comparing the performance of FRs in neocortical against mesio-temporal epilepsies. Since we studied only the performance at high specificities, the low boxplots show the poor global sensitivity. FRs thus describe only a subset of the EIZ. In a recent study [van't Klooster et al., 2015], the lack of sensitivity of FRs was assumed to be related to the insufficient spatial sampling of deep structures due to the recording modality (electrocorticography). We believe that FRs do not occur in all epileptogenic tissues, in line with another SEEG study [von Ellenrieder et al., 2016a] where they did not find FRs in some epileptogenic channels. For instance, if we focus in the mesial area, the rate of the FR inside the EIZ is much lower in the mesial temporal pole 0.2 min^{-1} [0-2 min^{-1}] than in the other structures (Am: 1.4 min^{-1}[0.25-7.2 min^{-1}]; Hip: 11.3 min^{-1} [5.2-18.8 min^{-1}]).

This is of clinical interest since the temporal pole can generate seizures along with other mesiotemporal structures [Bartolomei et al., 2010, Kahane and Bartolomei, 2010, Chabardès et al., 2005]. Fig IV.1 illustrates this issue with patient 30 where the FR rate is very high in the amygdala and the hippocampus but not in the mesial temporal pole even though each structure had high EI values. Patient 11 and 3 in Fig IV.5 also describe the low sensitivity of the FRs which are only seen in one part of the EIZ. We would like to emphasize that such conclusions can only be made if the detection procedure is not biased by the variation in activity across brain structures and if it detects equally precisely Rs and FRs. We showed that most of the detectors were biased because the impact of the power spectrum was neglected or because the estimation of the background activity was not robust [Roehri et al., 2017]. Delphos was the only method to overcome both issues whithin our proposed benchmark. Therefore the present study shows that it is highly likely that FRs do not occur in all epileptogenic regions. This highlights the limitation of focusing only on FRs when including pathological Rs should be more appropriate. This however begs the question of how to distinguish physiological from pathological ripples.

IV.4.3. Physiological HFOs

To better predict the EZ, some teams have proposed to distinguish physiological from pathological HFOs by studying their relationship with spikes [Jacobs et al., 2008, Wang et al., 2013] with slow waves [von Ellenrieder et al., 2016a] or with spindles [Bruder et al., 2016]. Although promising, none of these approaches has so far succeeded in fully separating them. Other teams have suggested to separate them by clustering HFOs based on features such as frequency, duration, amplitude [Matsumoto et al., 2013, Malinowska et al., 2015]. The former study managed to differentiate pathological HFOs from normal task-induced HFOs. This segregation remains to be validated on spontaneous HFOs such as those encountered in the current study. The latter found that some features were higher inside than outside the SOZ (i.e. amplitude and duration) but the difference was small. In the current study, we investigated the potential of a multi-marker analysis rather than directly separating physiological from pathological HFOs. The question whether pathological HFOs are better than spikes remains open.

IV.4.4. Limitations

In the analyzed cohort, only ten patients underwent resective surgery. This is certainly a limitation of our study and more patients with outcome should be investigated to validate our observations. However, studying only patients who underwent surgery may lead to erroneous conclusions. Indeed, the resected area may not be exclusively composed of the epileptic zone. The outcome solely defines the sensitivity of the resection and not its specificity. Moreover, surgery is not proposed to every implanted patient, and this would entail a clear selection bias. Future studies should apply statistical analysis that consider outcomes and resective zones along with the putative EZ to include patients who did not undergo surgery. To estimate the EZ, we used the EI; our results are relative to the EIZ and may not generalize to other means of estimating the EZ. We set the threshold at 0.25 in order to have a correspondence of approximately 50% between the EIZ and the SOZ. The classical threshold as defined in the original study [Bartolomei et al., 2008] is 0.3 but this was determined in TLE patients; the heterogeneity of our patients may explain this slight decrease. However, the supplementary table shows that neither the change in the EI threshold nor the modification of the minimum specificity affect the main findings and the performance of the Spk $\otimes$ HFO. It is unknown whether the 5-min duration chosen here may be suboptimal to map the interictal activities; this is still an active area of research. Although a given duration might be advantageous for one marker, it might on the other hand be detrimental for another because of non-stationarity of the signal. The bipolar montage could possibly hinder HFO visibility [Menendez de la Prida et al., 2015]. This montage is however the standard one for studying seizures in a clinical setting. It seems unlikely that two macro-contacts record the same small HFO generator yielding a destructive interference. Nevertheless, macro-electrodes have a wider spatial resolution compared to micro-electrodes; the resulting background activity could cover HFOs. Only hybrid micro-macro electrodes, recording the same activity at two different scales, could help resolve this question. The contacts near the gray-white matter interface may be an issue as they may be more sensitive to ictal discharge and spikes than HFOs and thus bias the performance of the markers. This is a complex question which needs further investigation.

IV.4.5. Conclusion and future perspectives

This study investigates the reliability of different interictal biomarkers in predicting the epileptogenic zone. HFOs do not seem to delineate epileptogenic areas better than epileptic spikes both at the group and at the patient level. Ripples suffer from physiological counterparts and or at least spikes from less pathological counterparts (possibly propagated spikes). FRs seem to be specific to the epileptogenic zone but do not fully describe it; they cannot be considered as a unique biomarker of epileptogenicity, they are only the tip of the iceberg. A promising path is to combine biomarkers. Combining spikes with HFOs by multiplying the two rates (Spk $\otimes$ HFO) improves the prediction of the epileptogenic zone with respect to biomarkers considered separately. Tissues are unlikely to be epileptogenic if they produce a high rate of either spikes or HFOs. The performance of Spk $\otimes$ HFO is however not optimal for every patient. This means that combining markers may only be a first step in improving the delineation of the epileptogenic zone. Future studies should focus on how to better combine markers (e.g. with the use of machine learning techniques) and on other markers (e.g. slow wave, low-frequency oscillations). Considering the low sensitivity of the FRs, future works should also investigate whether this is related to factors such as brain region, pattern of seizure discharge, epilepsy etiology, recording techniques (macro vs. micro-electrodes), montage (bipolar vs referential) [Menendez de la Prida et al., 2015] or suboptimal recording duration.

Acknowledgments

We would like to thank Didier Scavarda and Romain Carron for the stereotactic placement of the electrodes, Philippe Kahane for leading the FORCE project and Samuel Medina Villalon for the 3D visualization. This work has been carried out within the FHU EPINEXT with the support of the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the “Investissements d’Avenir” French government program managed by the French National Research Agency (ANR). This study was also supported by the ANR project FORCE (ANR-13-TECS-0013).

Author Contributions

NR, FP, SL, IL, AN, AM, FB and CGB took part in the conception and design of the study. FP, AN, AM and FB contributed in acquiring the data, and NR, FP, SL, IL, AN, AM, BG, FB and CGB analyzed the data. NR drafted the manuscript and the figures.

Conflicts of Interest

NR, CGB and FB have a patent pending (PCT/EP2017/054574) about the detection procedure.

IV.A. Supplementary Material

IV.A.1. Meta-matrix of performance

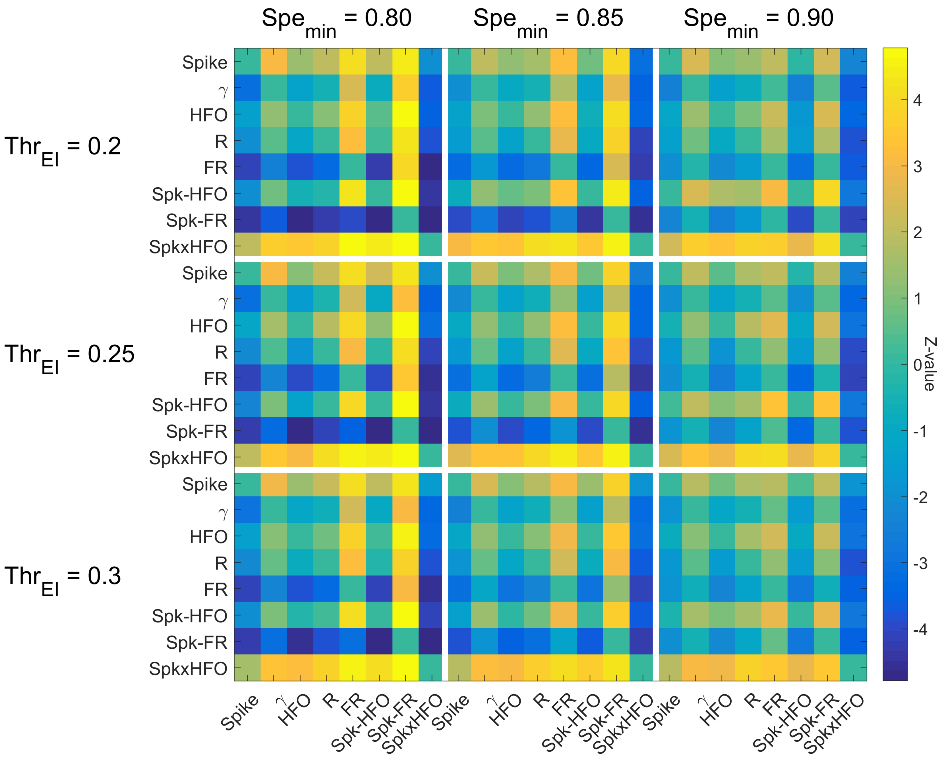


Figure IV.8. – Meta-matrix of performance for different settings of EI threshold and minimum specificity. The matrix is composed of 9 matrices obtained by calculating the Wilcoxon signed rank test for each pair of markers for different value of EI threshold (0.2; 0.25; 0.3) and minimum specificity (0.80; 0.85; 0.90). The settings used for our study is EI threshold = 0.25 and minimum specificity = 0.85. It corresponds to the matrix in the center; the same as in Fig IV.6. The results are similar across settings. Note that Spk $\otimes$ HFO is still better than the other markers whatever the setting (yellow line).

IV.A.2. Variability of spike and HFO rate across structures and patients

Here is a figure IV.9 which was removed from the above manuscript because it was too complicated. I however think that what it illustrates is interesting and worth mentioning for someone willing to decode it.

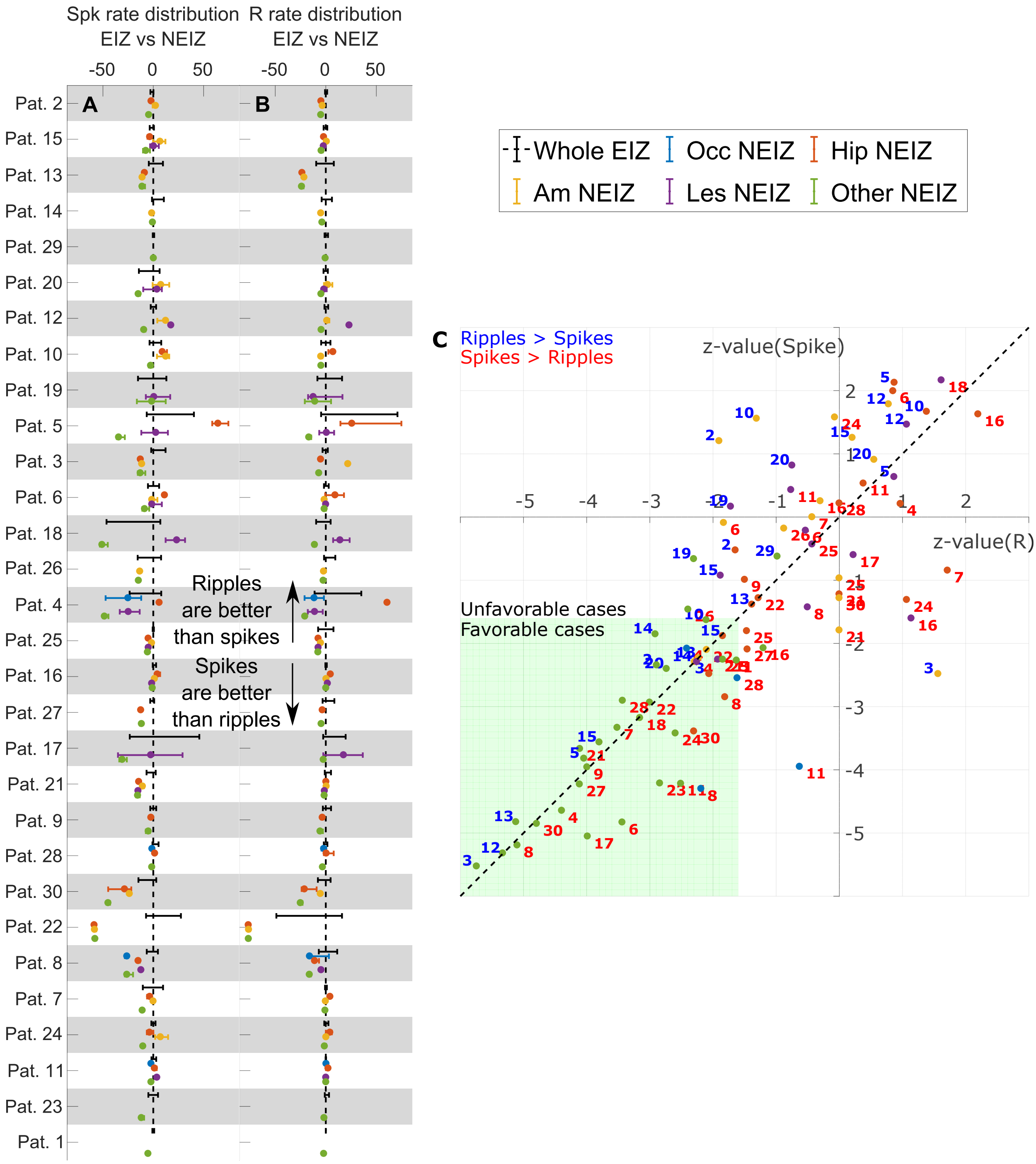


Figure IV.9. – Spike and ripple rate in the EIZ compared to their rate in specific structures outside the EIZ. (A and B) shows the median and interquartile range of the rate of the spike (A) and ripple (B) for different areas for every patient. Each rate interval is aligned on the median of the rate inside the EIZ and patients are ranked according to the difference in performance between the ripple and the spike. (C) represents the z-values of the Mann-Whitney U test comparing the rates in the NEIZ structures of each patient independently to the rate inside the EIZ. The dots relative to the structures are color-coded in a similar way as in A and B (relative to specific structures). Numbers correspond to the patient numbers. [...]

These numbers are colored in blue (red) when the ripples (spikes) performed better than the spikes (ripples) in term of pAUC. The green space delineates the favorable cases (when there is not much overlap); the rest corresponds to the unfavorable cases (when there is an overlap or when the rate in the NEIZ is higher than in the EIZ). When a dot lies above the dashed line (y=x), it means the difference in the rates in the brain area is more unfavorable to the spikes than to the ripples. Inversely, when a dot lies below the dashed line it is unfavorable towards the ripples. One can observe that cases favorable using the ripples - left side of A and B - have less overlap in the ripple rates than in the spike rates. Conversely, the overlap in the ripple rates is more stressed than in the spike rates on the right side of this figure. As an example of the former, one could look at the NEIZ amygdala of patient 15, which exhibits spike rates above the rate in the EIZ but only overlaps with the EIZ in the ripple rate. Therefore, the coordinate of the amygdala of patient 15 is above the dashed line and has a z-score for the ripple close to zero. To illustrate the latter, the hippocampus of patient 24 has a spike rate below the rate in the EIZ while its ripple rate is above the median of the EIZ. The relative dot thus lies in the positive z for the ripple but in the negative z for the spike. Note that almost all the blue numbers – when ripples perform better than spikes – are above the dashed line and inversely the red are mainly below this line. Note the blue dot of patient 11 correspond to the electrode CU' of Fig IV.5. The dot is below the dashed line, i.e. it is more difficult to differentiate the EIZ from the NEIZ occipital lobe by measuring the ripple rate than the spike rate.

Closing remarks

This work concludes that no marker taken separately is better than another, but fusing the information of the spikes and HFOs may be a better approach. This conclusion could have a real impact on the clinical environment as both spikes and HFOs should be handled with caution. Since epilepsy is a complex disease, it seems unwise to seek a single biomarker. Merging information from different biomarkers appears to be a better approach. Already in clinical practice, information coming from different sources (e.g. MRI, semiology, electrophysioloy) are fused together. It thus seems appropriate to take advantage of all interictal markers.

In the following chapter, I discuss the results obtained in this thesis and give some suggestions regarding other possible metrics or methods that could be used to refine this analysis.

V. Discussion and future perspectives

V.1. Discussion

V.1.1. Few surgical outcome: a limitation?

In the pool of thirty patients we ran the analysis onto (IV Clinical Application), only ten underwent resective surgery. Six are awaiting surgery or surgical decision. This is certainly a limitation of our last study and more patients should be investigated to validate our observations. We however believe that studying only patients who underwent surgery or focusing only on the outcome may lead to erroneous conclusions. By definition, an Engel I class patient had its epileptogenic zone removed. Nevertheless, the resected area may not be exclusively composed of the epileptic zone. By the outcome, one can solely define the sensitivity of the resection and not its specificity nor precision (I.6 Binary Classification). The amount of cortex which was mistakenly removed cannot be comprehended. When patients have recurrent seizures after surgery, it could not only be because the epileptic focus was not entirely resected but also because the EZ was not focal (which is the first criterion for resective surgery [Kahane et al., 2006]) but rather an extended network. Contrariwise, if an extended network was suspected, surgery may not have been carried out. The specificity can only be obtained when one wants to estimate the EZ prior to the operation. The delineation of the EZ is less certain; the estimation is noisier. It is however the only option to estimate the specificity. One obvious bias is that the resected area depends on the estimated EZ. The surgeon may however not exactly remove the proposed EZ because of anatomical constrains and blood vessels. It is therefore worth combining the outcome with the proposed EZ.

Moreover, surgery is not proposed to every implanted patient. There is here a clear selection bias; the cohort of patient is not representative of the population intended to be analyzed. According to statistics, one cannot compare biomarkers using only outcome and generalize the results to focal refractory epilepsy since EZs corresponding to extended networks or overlapping with eloquent areas were not studied. Furthermore, large exploration is more likely when extended networks are suspected. With such exploration it is more likely to measure physiological HFOs. But again, this is not considered when focusing on surgery. Our study may be biased because we could not take into account the resection; we could however study every patient regardless of their type of EZ. Future studies should apply a statistical analysis that consider outcomes and resective zones along with the putative EZ to include patients which did not undergo surgery. For these patients, resection is not an option but disconnection of well investigated network could be [Jirsa et al., 2016, Bartolomei et al., 2017a].

We believe that our work, by investigating every patient, may alert the clinicians concerning physiological HFOs. Our work suggests that physiological HFOs may greatly impact the delineation of the EZ based on HFOs. HFOs have to be handled carefully because they may arise from physiological HFO generators. This is especially important since some teams have started to investigate HFOs non-invasively [Pizzo et al., 2015, van Klink et al., 2015, Nissen et al., 2016b, von Ellenrieder et al., 2016b]. Non-invasive recordings measure the activity of the whole brain and may more likely record cortical HFO generators. A recent publication [Mooij et al., 2017] showed that ripples were detected in scalp-EEG in epileptic and non-epileptic children. Above all, these ripples were very similar in epileptic and normal children. Separating physiological from pathological ripples thus seems an intricate but essential step toward the clinical application of HFOs. Our proposition of combining spikes and HFOs is an indirect way of distinguish the two HFO types. We could consider the spikes to be a biased/smooth estimator because they are mainly present in the EZ but more or less present outside the EZ. The HFOs are rather a noisy estimator since their rate can be very high in the EZ but also outside due to putative physiological HFO generators. In V.2.4 Spike and HFO multi-layer network, we propose a method which goes a step further in combining the two markers. This could distinguish the two HFO types and deserves further investigations.

V.1.2. Fast ripple: the hidden Grail?

In IV Clinical Application, we have found that fast ripples were specific to the EIZ (because there are rarely found outside the EIZ) but lack in sensitivity. We hypothesized that fast ripples only appear in a subset of the EZ; other interpretations could explain this results. I discuss these proposals below.

Increasing number of publications [Menendez de la Prida et al., 2015, Shimamoto et al., 2017, Waldman et al., 2017] have raised the question of the relevance of the bipolar montage for HFO investigations. Calculating the difference of potential between two adjacent channels may cancel out HFOs and thus hinder their visibility. In a recent review [Menendez de la Prida et al., 2015], an example showed that a bipolar montage distorted an HFO which was recorded on two different contacts. The distortion was such that the resulting signal did not resemble an HFO anymore. This example however was obtained using micro-electrodes. It seems incorrect to extrapolate this example to clinical macro-electrode because they are not merely up-scaled micro-electrode; the size of the generator is likely constant but the inter-contact distance is different as well as the integrated field. Supposing that this assumption is true, i.e. the bipolar montage distorts HFO up to a point that there do not appear as HFOs anymore, could the lack of sensitivity of fast ripples be explained by the bipolar montage?

In such case, one would expect a general decrease in the number of fast ripples. Fig IV.5 shows that the fast ripples are only located in a subset of the estimated EZ (EIZ) for patients #3 and #11. Both of these patients had a lesionectomy plus a temporal cortectomy and are Engel I. Patient #11 had only fast ripples in the amygdala, whereas ripples were found in the hippocampus and it is known that fast ripple can be recorded in the hippocampus. It thus seems unlikely that this lack of sensitivity is explained by the montage. Could it be related to noise level?

Fedele and colleagues [Fedele et al., 2017b] have developed a custom-made low-noise amplifier to improve the signal-to-noise ratio of intra-operative recordings. They showed that this system increased the detectability of fast ripples compared to a commercial amplifier. One has to take into account that the operative room is much noisier that clinical shielded rooms. While this low-noise amplifier may be useful for intra-operative recordings, it may not drastically improve fast ripple detectability in chronic long-term monitoring. Moreover, this amplifier only targets electrical noise. Another type of "noise", which is likely to cover fast ripples, is the background activity.

As background activity differs in different brain regions and in lesion types and is influenced by the recording techniques, this could be a reason why in certain cases one cannot record fast ripples (or more generaly HFOs). Biophysical models which take into account both the mechanism producing HFOs and the recording means [Shamas et al., 2016] could provide pieces of explanation. Hybrid micro-macro electrodes may also bring answers to this question by recording simultaneously at two different scales the same activity. If this hypothesis is proven, the application of the hybrid micro-macro electrodes in a clinical setting may however be hampered due to the practical and safety issues of having micro-wires everywhere in the brain... Even the best whitening method would not improve the detection since it does not permit to see through the background activity at the same frequency... The only viable solution is to utilize multivariate analysis such as independent component analysis that could improve the signal-to-background-activity ratio. Such technique would have to be improved because they are attracted by high power sources and may miss small focal activity.

V.1.3. Future of HFO detectors

Characterizing HFO detectors by calculating their sensitivity and specificity is unsatisfactory. In III Detection Validation (and in [Roehri et al., 2017]), we have shown that other parameters have to be taken into account such as the amplitude of the oscillations relative to the background level, the variability of the background and the content of the signal. Moreover, we have stressed out that sensitivity should be separately investigated for ripples and fast-ripples to avoid misleading conclusions. In I.6 Binary Classification, we have illustrated why detectors have to be compared on the same dataset. In chapter III, we could not test every detector because most of them were not available online or complicated to re-implement (especially those with multiple stages and settings [Burnos et al., 2014, Burnos et al., 2016, Liu et al., 2016]). This is why our benchmark is available online along with the detections of the tested detectors.

Furthermore, the tested detectors (except Delphos) were designed to detect HFOs in the whole HFO band, i.e. 80-500 Hz. We thus decided not to examine the accuracy in classifying an HFO as ripple or fast-ripple (we however tested the sensitivity in the different classes). Such classification is however of utmost importance when attempting to map interictal activities. We have not published but have reported in III.4 Closing remarks quantitative measures concerning Delphos' performance in classifying the oscillations (Fig III.12). The precision of each class was around 90%. For other detectors which rely on filter and spectrum based classifiers, several issues may be predicted. Classifying HFOs by finding the peak frequency in the power spectrum of the detection was shown to be insufficiently robust [Gliske et al., 2016]. Assigning a class to the detection based on whether the detection occurred in the ripple or fast ripple bandpass filtered signal is also not robust. Indeed, bandpass filters are not perfect filters. Some events with a frequency content near the limit of the band pass are still visible in the filtered signal. Furthermore, the oscillations we are dealing with do not have a punctual frequency support (Fig I.10). For instance, a ripple at 230 Hz lasting four periods is likely captured in both filter settings. Fedele and colleagues refered to this phenomenon as unimodal fast-ripple and ripple [Fedele et al., 2017a]. Time-frequency based classifier (like Delphos or [Burnos et al., 2014, Liu et al., 2016]) thus seem the best available option.

If this sort of benchmarks does not make its way to undeniable detector testing procedure, one could fear an overload of pseudo-novel HFO detectors and this field of research may be blocked in an infinite loop. Obviously, other type of benchmarks which would target other aspects such as the effect of background non-stationarity is encouraged [Birot et al., 2013, Jrad et al., 2016]. We generally think that benchmarking is an inevitable path the HFO research has to take. Another bridge HFO detectors have to cross is the transfer from research algorithm to embedded software. This is why, one aspect of my thesis was to propose a prototype of Delphos to Micromed™(company manufacturing electro-medical devices). Delphos is already an AnyWave plugin which does not need a MATLAB™ license and is already installed in most of the computers in the Epilepsy unit of *la Timone* Hospital. The remaining steps are to adapt the input and output to the Micromed™ Software. Delphos will also be available for clinical research on the AnyWave website[1] once the legal agreement concerning the patent will be handled.

---

[1] `http://meg.univ-amu.fr/wiki/AnyWave`

V.2. Future perspectives

In this section, I present some possible future applications of this thesis. The first two concern other properties of the Z_{H_0}-normalization: Z_{H_0}-suppression and Ridge-Riding Denoising. The third section describes the utility of Delphos in mapping oscillations at each frequency. The final section details a multi-layer network approach which goes one step further than the Spk $\otimes$ HFO in merging the information of the two markers.

V.2.1. Z_{H_0}-suppression

Kobayashi and colleagues have developed a method to detect changes of high-frequency activity in epileptic spikes [Kobayashi et al., 2009]. In brief, two types of spikes were studied, those preceding a slow wave and those being alone. The Gabor transform was applied to spikes and baseline sections. A t-test was used to highlight the significant increases and decreases in power during the spikes compared to the baselines. Notably, there was a significant decrease during slow waves after spikes and this effect was stronger in SOZ channels [Jacobs et al., 2011]. The drawback of this analysis is that it needs many baselines and the result is not given for a single spike but for a group of spikes.

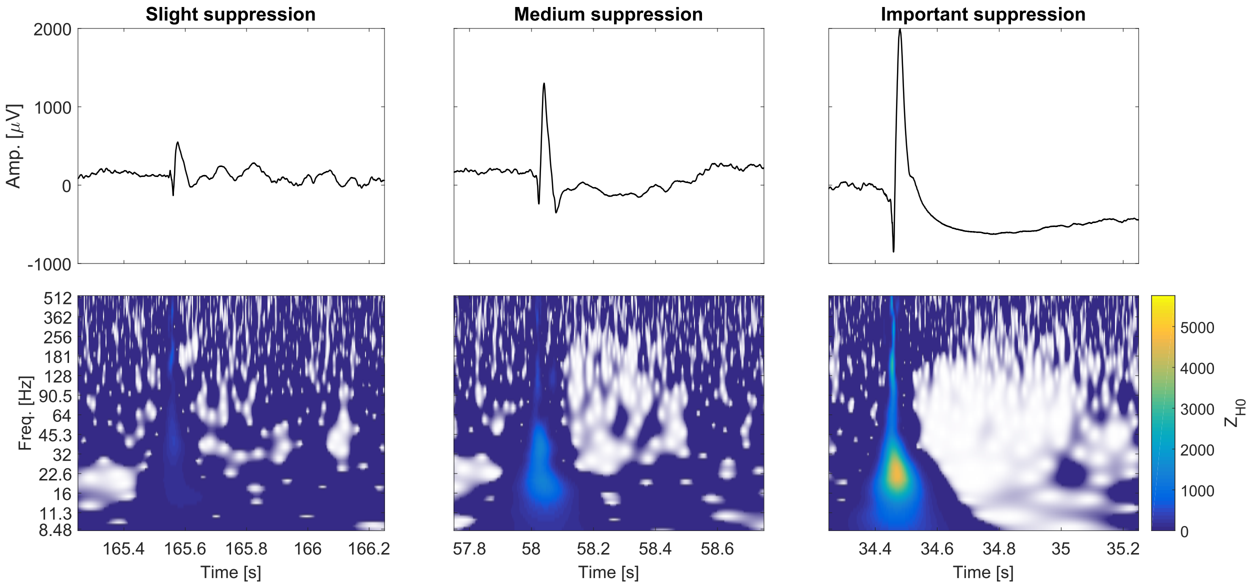


Figure V.1. – Illustration of post-spike suppression in Z_{H_0}. Values below 1 are represented in white. The whiter they are, the closer they are to zero and thus the more suppression there is.

Since the Z_{H_0}-normalization does not need a baseline and has a statistical meaning, one can easily capture such decrease or post-spike suppression by thresholding the TF image. Indeed, values below 1 – i.e. below one standard deviation of the estimated background activity (Fig II.3) – are sparsely present in the whole dataset but produce a broader spot in case of such suppression. Fig V.1 shows Z_{H_0}-suppression. The last spike clearly exhibits an important suppression during the slow wave with a large white spot. These spots could be characterized in terms of perimeter or area and used as new features for the spikes. One could also characterize each spot to extract a global suppression metric (at the channel level) and compare the chosen metric to the one of the white noise. One could use the distance between the two distributions which could be higher in EZ channels.

V.2.2. Ridge-Riding Denoising

A second aspect of the Z_{H_0} method which could be interesting is the ridge properties [Carmona et al., 1998, Mallat, 2008]. The wavelet ridges are defined as the local maxima (according to the frequency axis) of the TF plane. Instantaneous frequency and amplitude can be measured from the ridge. Moreover, using the ridge as a mask, one can reconstruct the underlying oscillation with Eq. II.16. This can be used as a denoising technique [Carmona et al., 1998].

In terms of HFOs, Delphos, after detecting the oscillations contained in the signal, could, from the detected local maxima in the Z_{H_0}-plane, ride down the ridge of each detected oscillation (hence the name I gave to this technique) and reconstruct them – possibly separately – without their surrounding background and potential spikes (Fig V.2). It is possible to reconstruct the oscillation in the original time-domain and in the whitened version. From this sparse representation, one could then apply connectivity measures or even cross-frequency measures (e.g. between spindles and ripples [Bruder et al., 2016]). Synchrony based on phase can be directly applied in the TF plane using the ridge mask [Zerouali et al., 2013] (which would be more precise than estimating the phase by applying the Hilbert transform on a band-pass signal). Moreover, while one filters the data and calculates any connectivity metric to epileptic channels, one cannot know whether the connectivity arises from synchrony in the background activity, in spikes or in oscillations... This technique should reduce the spurious connectivity due to the $1/f$-spectrum [Achard, 2006, Achard and Bullmore, 2007] and spikes, and represent a better estimation of oscillation-based networks.

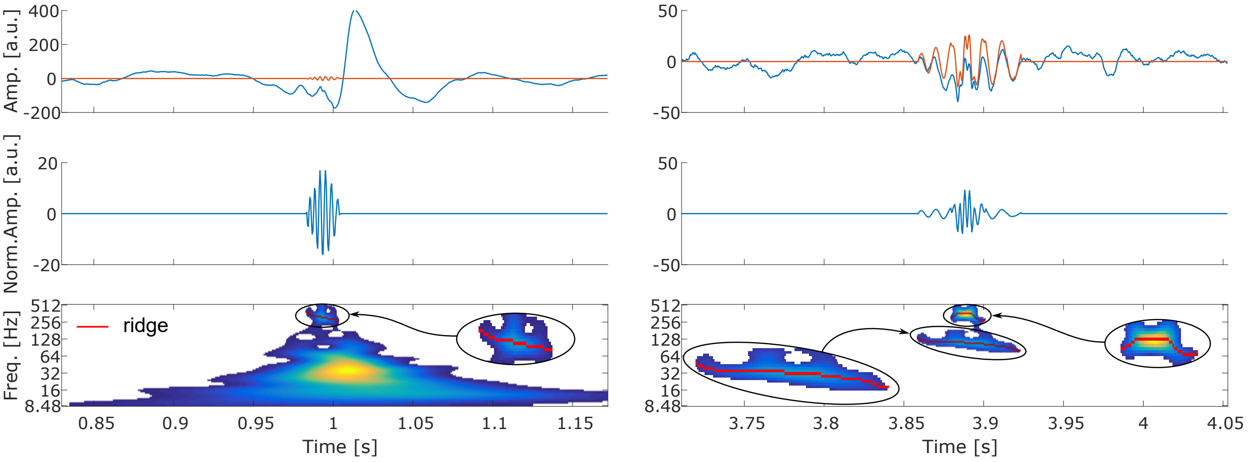


Figure V.2. – Examples of ridge-riding denoising in both the original and whitened signal.

V.2.3. Entire frequency range oscillation mapping

As its name suggests, Delphos was not designed to detect only HFOs but oscillations at any frequency. In the analysis undertaken in IV Clinical Application, only the HFO and gamma rates were analyzed but oscillations at lower frequencies were detected as well. Fig V.3 represents the histogram of the average number of oscillations at a given frequency in the EIZ and NEIZ for each patient (as defined in the previous study, IV Clinical Application). When applying the Wilcoxon signed-rank test (nonparametric paired test) for each frequency bin across patients and correcting for multiple comparisons (FDR [Benjamini and Hochberg, 1995] below 0.05), two frequency bands appear significant. One corresponds to a high frequency band $[176, 352]$ Hz and another to a low frequency band $[8.7, 15.5]$ Hz.

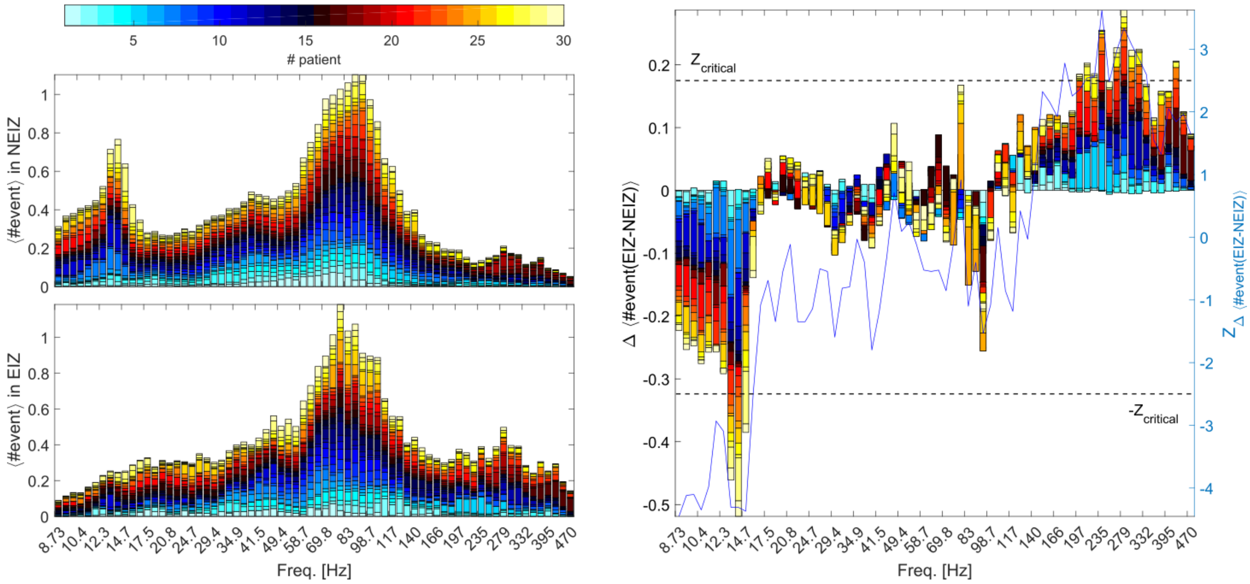


Figure V.3. – Complete oscillation mapping histogram in EIZ and NEIZ

The high frequency band is not surprising because it was shown that the number of HFOs is higher inside than outside the SOZ [Jacobs et al., 2008, Crépon et al., 2010, Malinowska et al., 2015] and we found similar results with the EIZ and NEIZ in chapter IV. The fact that this band is only a subset of the HFO band could be explained by physiological ripples occurring preferentially in the range of $[80, 176)$ Hz and by the presence of fast ripples in only a subset of the EIZ (see IV.4.2 Low sensitivity of FRs and Spk-FR).

Interestingly, this analysis further suggests that the oscillations in the identified low frequency band are suppressed in the EIZ. Especially the negative peak in $[12, 14]$ Hz, i.e. in the frequency range of the sleep spindle, shows a drastic decrease in the number of spindles in the EIZ. This is in line with previous studies [Tezer et al., 2014, Frauscher et al., 2015]. It indicates that information about epileptogenicity is not restricted to a specific frequency range. Delphos may be an interesting tool to map oscillations at all frequencies since it is, to my knowledge, the only detector capable of broadband oscillation mapping.

Does that mean that one should calculate the rate of the oscillations in the low frequency and do an analysis similar to the one applied to HFOs? Probably not, since not all regions generate spindles. This approach should be incorporated into an atlas-based framework where the number of oscillations in a given structure at a given frequency should be compared to the number obtained in the same healthy structure at the same frequency. The assumption that structures without spikes and outside the SOZ are healthy is debatable but it is the best approximation of 'control' one could think of. Such a framework would necessitate numerous patients to have a reliable estimation of the 'normal' and 'epileptic' regimes for each region. This highlights the necessity of a multi-center initiative to build such a colossal dataset.

V.2.4. Spike and HFO multi-layer network

Several studies have attempted to improve the delineation of the EZ by separating physiological from pathological HFOs (issue discussed above IV.4.3 Physiological HFOs). On the one hand, teams have tried to cluster the two ripple types according to their characteristics (e.g. frequency, duration, amplitude). It seems to work well for normal task-induced HFOs [Matsumoto et al., 2013] (which could actually be HFAs: high-frequency activities [Lachaux et al., 2012]) but the difference is weak when studying spontaneous HFOs [Malinowska et al., 2015]. There is a strong underlying hypothesis when trying to separate HFOs based on features. It supposes that pathological HFOs share the same characteristics across brain regions. This is however not verified. This weak difference between the two classes could be explained by this biased hypothesis or poorly-chosen features.

On the other hand, teams have tried to separate them by studying their occurrence regarding other physiological activities (e.g. spikes, slow waves and spindles [Jacobs et al., 2008, Wang et al., 2013, von Ellenrieder et al., 2016a, Bruder et al., 2016]). This seems a captivating idea especially at a mechanistic level. At a detection theory level, it seems less optimal. Indeed, such co-occurring complexes necessitate to detect both the HFOs and the other activities. The maximum sensitivity of detecting such complexes are limited by the lowest sensitivity of the elements detected separately. This means that, in reality, the sensitivity would most likely be lower.

A strategy which, as far as I know, has not been studied yet is graph theory. It seems relevant to think that physiological HFOs as produced by physiological processes may arise from different networks than pathological HFOs. As mentioned above in V.2.2 Ridge-Riding Denoising, it seems inappropriate to use conventional connectivity metrics to study HFO networks. Using the work performed on neuron-spiking activities [Grün et al., 2002a, Grün et al., 2002b], one represents the occurrence of HFOs as a raster plot. To do so, one discretizes the time-domain into bins of size δt and binarizes the data as 0 (when no events happened in this time bin) and 1 otherwise. Using information theory, the mutual information of such "codes" can be used to capture the synchrony between different brain regions. With such connectivity matrix, one can calculate diverse metrics such as the degre of nodes – here channels – (i.e. the number of connections of a given node), and the nodal centrality (i.e. the importance of a given node within a network) [Rubinov and Sporns, 2010]. As the metrics are assigned to a channel, one can apply the same ROC paradigm used with the rates.

Another possibility could be to find the different modules or communities of a network. A community is defined as a group of channels which has more connections within their community than outside. Again, it appears conceivable that physiological and pathological HFOs emerge from different communities. Some methods even permit to distinguish overlapping communities, i.e. communities which have some channels in common. This seems particularly interesting since the hippocampus is known to generate both HFO types [Csicsvari et al., 1999, Alkawadri et al., 2014]. One could study whether the EZ is composed of one or more HFO communities and if this number may predict the surgical outcome. One could also try to predict the EZ by applying the ROC paradigm on features calculated on communities instead of nodes.

Following the idea of merging information of spikes and HFOs (Spk $\otimes$ HFO), one could use a multi-layer network approach (Fig V.4). For example, let us take the example of the urban transport system [Gallotti and Barthelemy, 2015]. Here, the nodes are the different train, bus or metro stations and the edges or links the different available train, bus or metro lines. One could see these transport means as separate networks. However, to go from one location to another, one can successively take the train, the bus and finally the metro. Therefore, these networks are interconnected. Each transport mean has its own network (in its own layer) and there are networks which connect each of the layers. This approach is still under mathematical development [Kivelä et al., 2014] and starts to be used in electrophysiology [Brookes et al., 2016, Guillon et al., 2016].

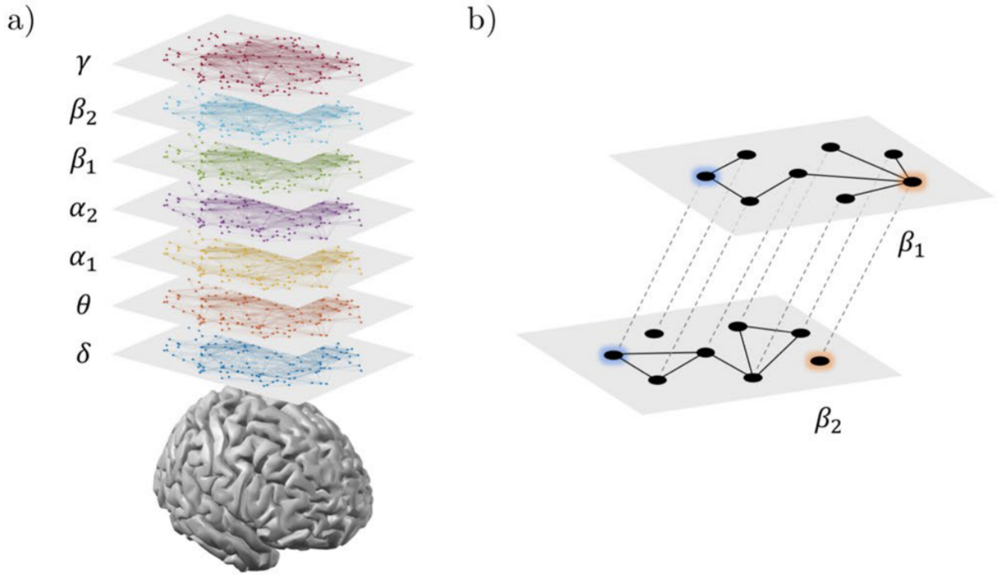


Figure V.4. – Example of multi-layer network in electrophysiology (MEG). (a) Brain networks for seven frequency bands. (b) Illustration of inter-layer connections in two bands. (adapted from [Guillon et al., 2016])

Fig V.5 represents the supra-adjacency matrix (spike and HFO multi-layer network) of two patients from the pool used in IV Clinical Application. These matrices were obtained with the technique described above (raster plot of spikes and HFOs and mutual information with a threshold based on surrogates). These supra-adjacency matrices are composed of four block matrices. The ones on the diagonal are the adjacency matrices or the networks within a layer and the others are the connections between the layers. In this case, the diagonal shows the spike network (top) and HFO network (bottom) and the others the spike to HFO and HFO to spike networks. To detect the different communities, I used the algorithm proposed by Ahn and colleagues [Ahn et al., 2010] (implemented in the Brain Connectivity Toolbox [Rubinov and Sporns, 2010]). The communities were ranked according to the overlap between the HFO and spike communities. It is interesting to see how well the first community corresponds to the EIZ in both examples. In patient 13, the SOZ outside the EIZ seems to form a different module (#2) possibly related to ictal but also interictal propagation. The community #5 in the two matrices show isolated HFO modules which could possibly be physiological HFO networks.

In these two examples I presented a proof of concept for a possible future application of Delphos which goes a step further in merging spikes and HFOs. Obviously, one has to verify whether the community hypothesis can be generalized in other patients and if isolated HFO modules correspond to physiological HFO networks. Moreover, the mutual information metric used here may not be optimal since it is influenced by the number of detected events in each class (there are more spikes than HFO detections). This approach has to be further enhanced but seems a fruitful alternative.

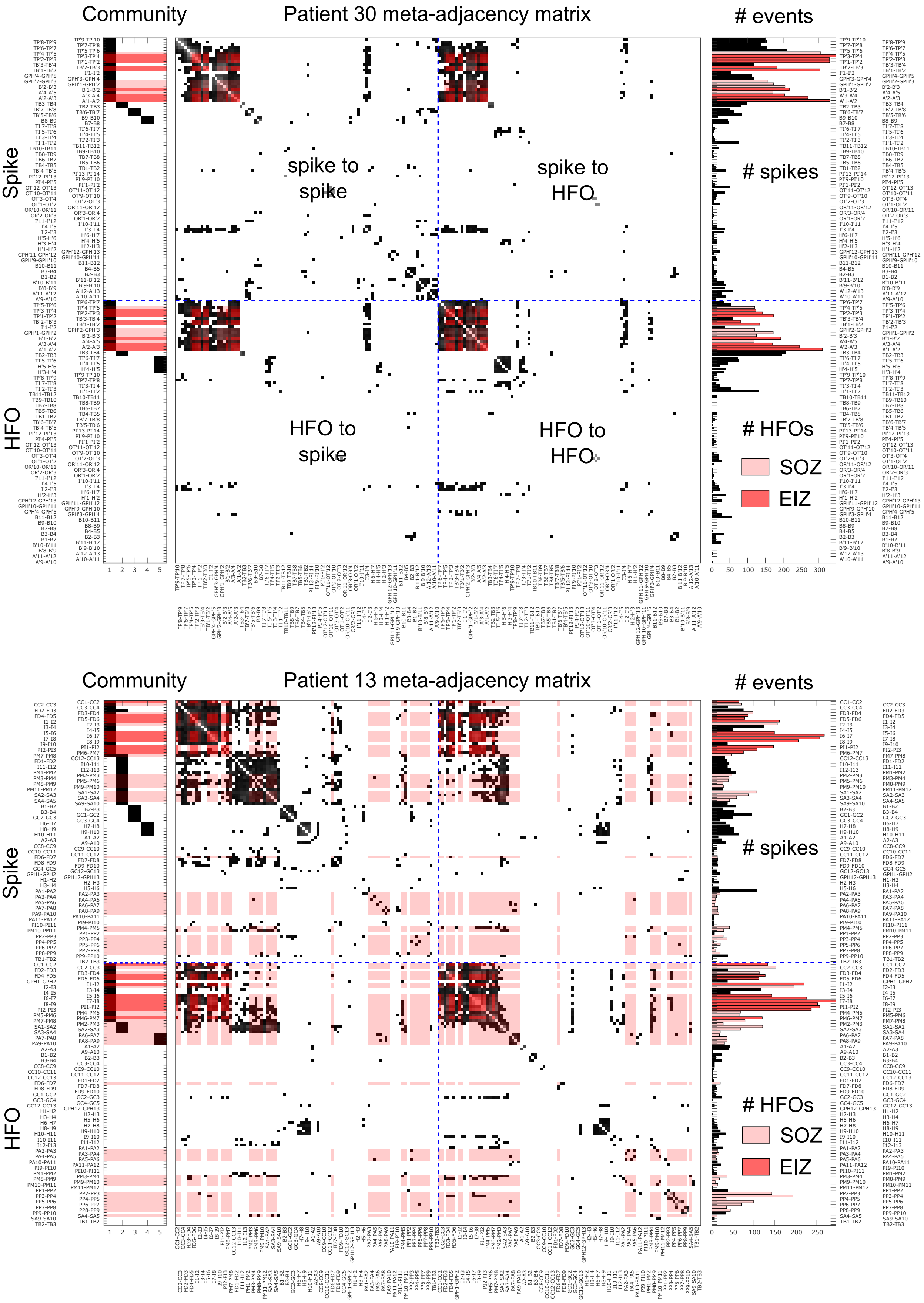


Figure V.5. – Supra-adjacency matrix of patient 30 and 13. Along the supra-adjacency matrix are the detected communities (on the left side) and the number of detected spikes and HFOs. The EIZ and SOZ are defined as in IV Clinical Application.

VI. General conclusion

This thesis resulted in three publications and one pending international patent. In the first chapter, I have described the novel and robust Z_{H_0}-normalization procedure. This method has been shown to be equivalent to baseline-based normalization. The Z_{H_0}-representation has been further used as a basis for an innovative detector: *Delphos*.

To validate the detector, we have designed a new framework to thoroughly characterize HFO detectors. Based on realistic simulations, we have been able to highlight unreported issues concerning some standard detectors, i.e. the lack of robust estimation of the background activity, the underestimated impact of the $1/f$-spectrum of physiological data, and the inadequate criteria defining an HFO. Thanks to the robustness of the Z_{H_0} technique and to the properties of the wavelet transform, Delphos overcame these issues and showed constant precision, higher sensitivity and less skewness toward any HFO classes.

Finally, Delphos has been applied to thirty patients with refractory focal epilepsies to map the interictal activities. These interictal maps were then compared to an estimation of the epileptogenic zone. Interestingly, neither the HFOs nor their variations (ripple, fast ripples or co-occurring spikes) were better predictors of the epileptogenic zone than the spikes. Is it the moment to renamed the HFOs as: Highly Fantasized Oscillations? This is fortunately not the case because the spike and HFO cross rate was the only marker surpassing the others. Furthermore, the separation between physiological and pathological HFOs remains an open question. This raises the idea that a better strategy to predict the epileptogenic zone would be to merge information from several markers.

In the future, Delphos could be a milestone in the approach of merging data of different markers since it is the only available detector which can detect spikes and oscillations at any frequency. It could open the way to new multi-frequency and multi-marker analysis to better understand this complex disease which is epilepsy.

A. Electrode Location Template

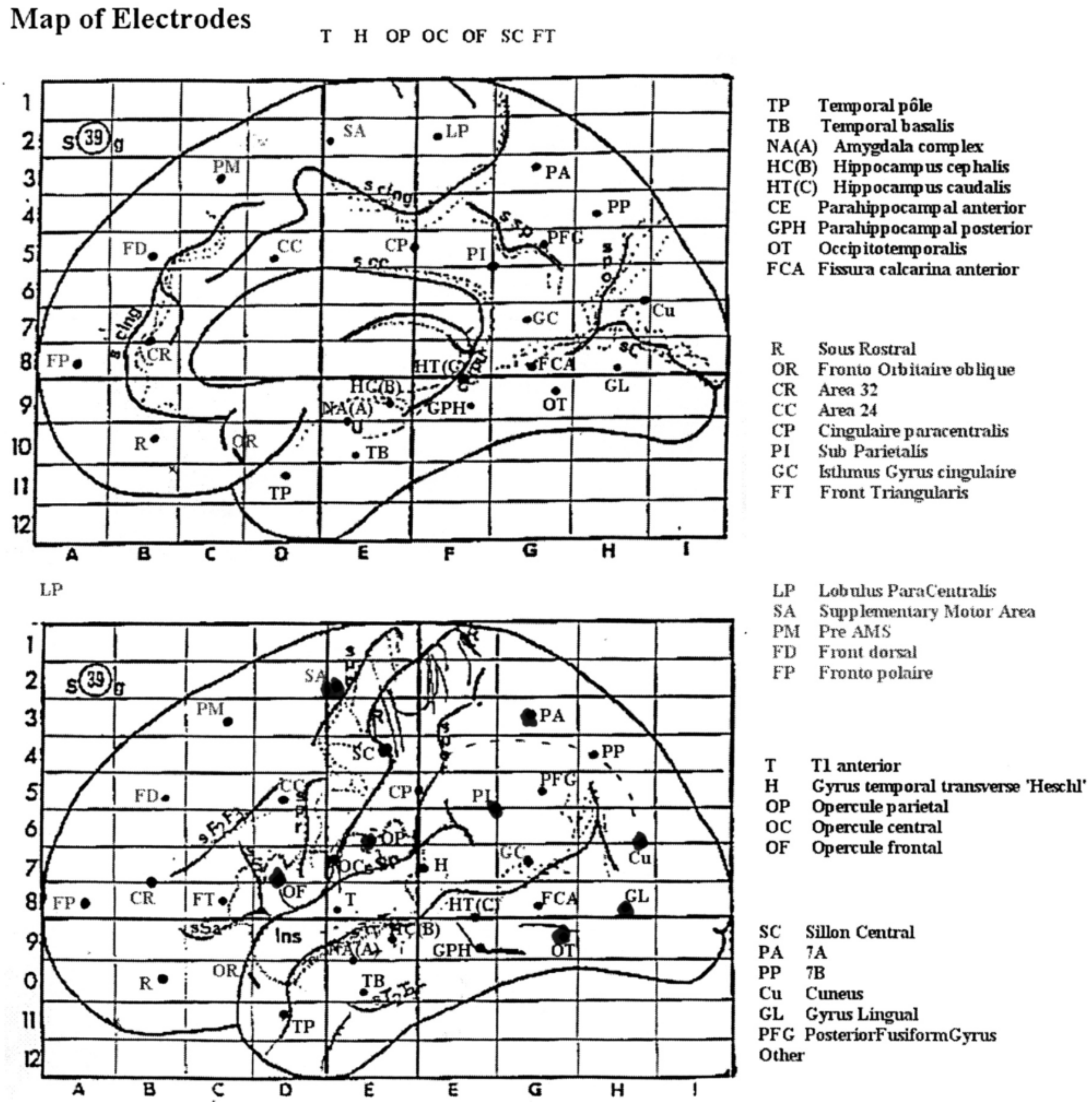


Figure A.1. – SEEG electrode location template

B. Developed Plugins

In this section, I present the plugins that I developed during this thesis. These plugins are modules of the AnyWave open-source software [Colombet et al., 2015], which was developed in the DynaMap team at INS. AnyWave was made in such a way that one can design algorithms or interfaces in Matlab™ or Python, which can then communicate with AnyWave. For instance, the plugin can receive SEEG recording data from AnyWave and extract some metrics/values and send back markers in the main interface. For an easy use and distribution of the plugins, they can be compiled and run without Matlab™. In the following paragraph I briefly describe the five plugins that I developed (B.1.1 Delphos, B.1.2 Delphos Viewer, B.2 Simulation Maker, B.3 Epileptogenicity Index, and B.4 Co-occurrence Graph, Fig B.1). I also mention the GARDEL software B.5 which aims at segmenting the SEEG electrodes automatically. I did not designed the interface but the image processing part which segments the electrodes.

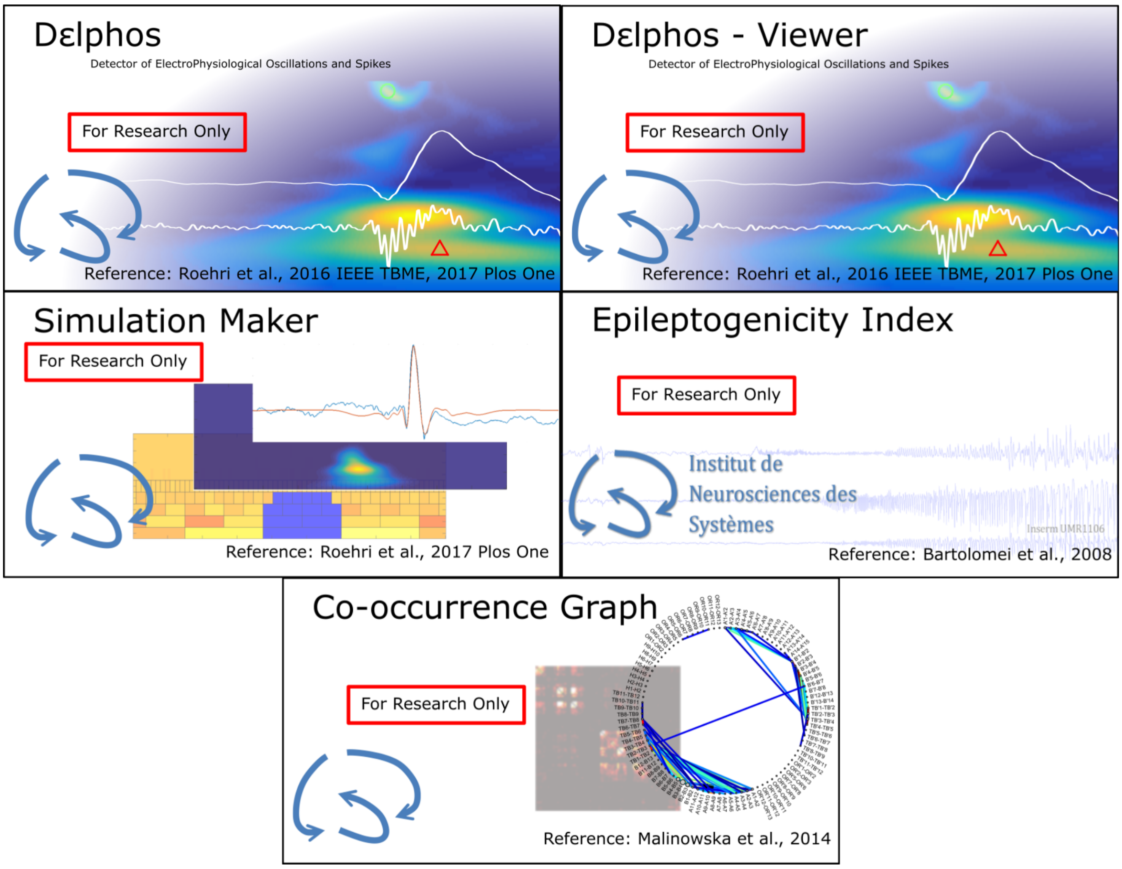


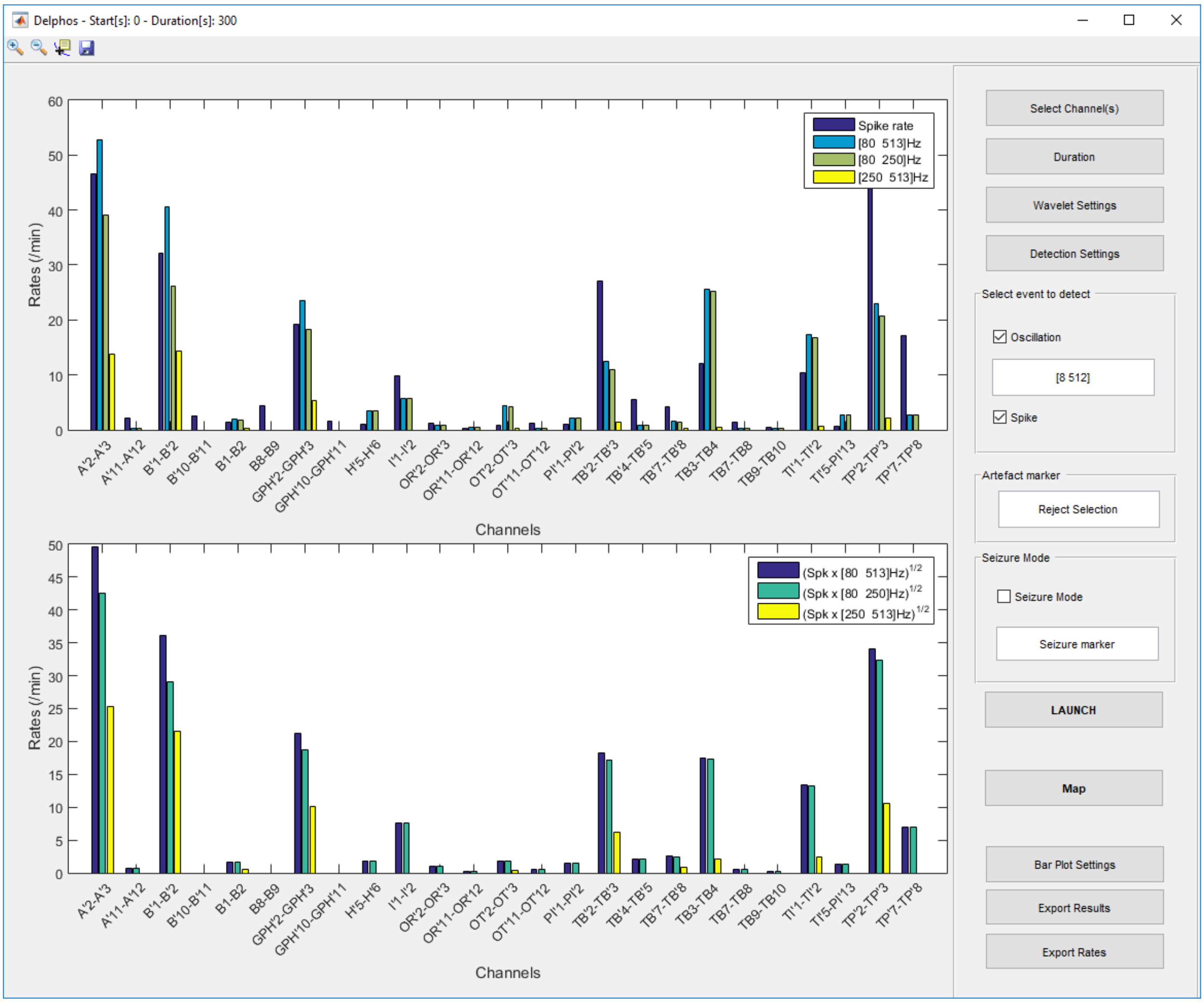
Figure B.1. – Splash screens of the five developed AnyWave plugins

B.1. Delphos: Detector of ElectroPhysiological Oscillations and Spikes

Delphos, standing for Detector of ElectroPhysiological Oscillations and Spikes, is a detector based on the Z_{H_0}-normalization of the wavelet transform described above (II, III). Two separate interfaces were developed: *Delphos* and *Delphos Viewer*. The former is designed to run the detectors while the latter permits to review the detections. These plugins were used in chapter III and IV and also in two published papers [Roehri et al., 2017, Pizzo et al., 2017].

B.1.1. Delphos

Once the plugin starts, the user selects the channels and the duration of the recording to be analyzed as well as the type of events to be detected (oscillations given a certain frequency band and spikes) as illustrated in Fig B.2. He or she can also modify the wavelet


Delphos - Start[s]: 0 - Duration[s]: 300

Select event to detect
☑ Oscillation
[8 512]
☑ Spike

Seizure Mode
☐ Seizure Mode

Figure B.2. – Delphos interface

setting and the detection parameters. If artifacts were previously marked in AnyWave, the user enters the name of the marker and Delphos discards the concerned section. At the end of the detection, the main frame displays the rates per minutes of the detections on all selected channels as a bar plot.

This plugin was designed to be applied on many channels to give a global map of the interictal activities. On a computer with 32 GB of RAM and 12 cores, it takes Delphos 10 min to analyze 100 to 120 channels of 5 min recording. To reach this speed, we parallelize the detection procedure according to the channels. In other words, each core launches the procedure on different channels. To avoid exceeding the available amount of RAM, Delphos cuts the data into several chunks of the same duration.

Every detection is automatically sent to AnyWave as color-coded markers, each color corresponding to different classes of oscillations or spikes. These markers can be used to review the detection in AnyWave or in Delphos viewer. The results can be exported in an Excel or Matlab™ file.

There are also two experimental modules. One is made for the detection during the ictal period estimating the background preceding the seizure. The second outputs a kind of raster plot of the detections.

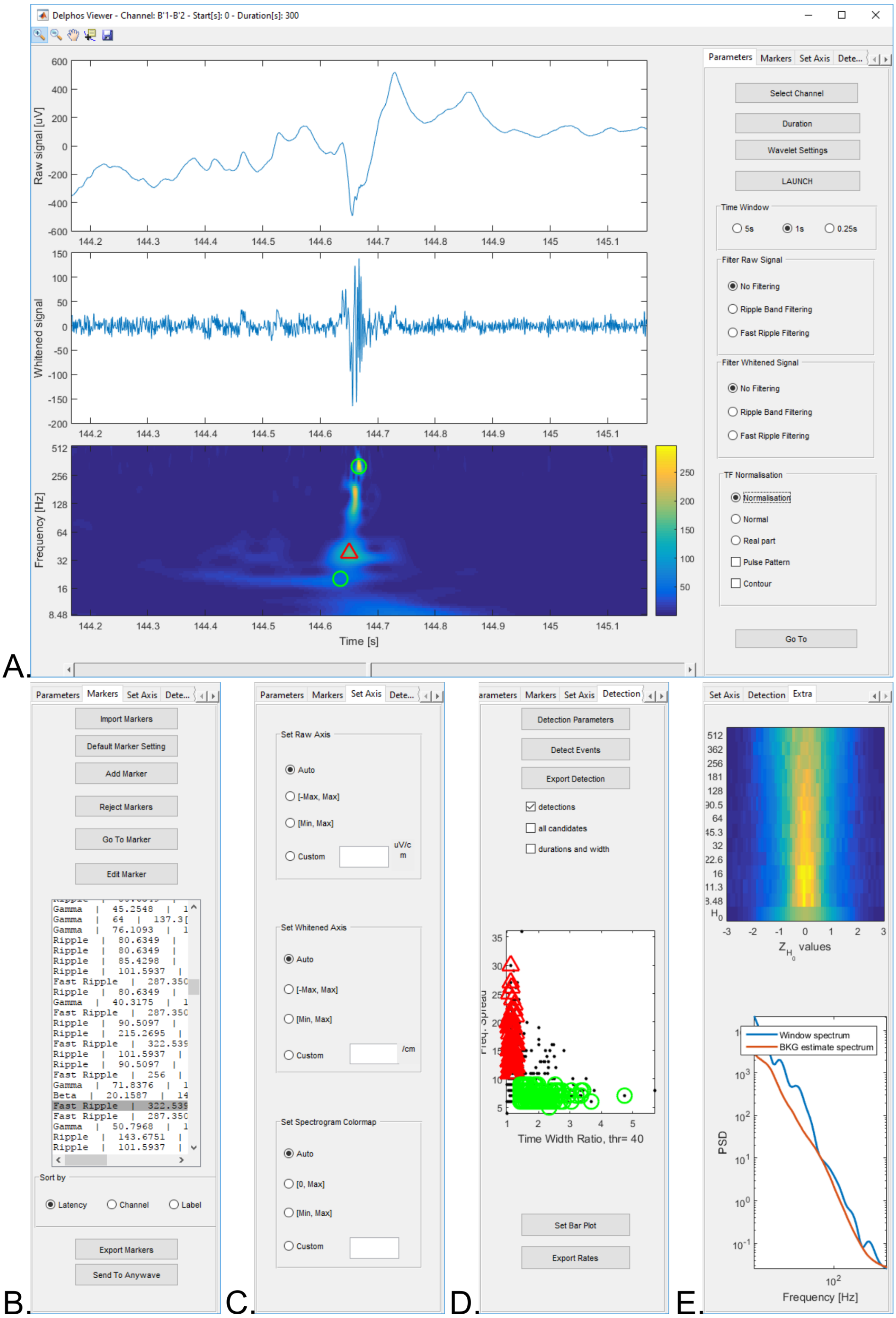
B.1.2. Delphos Viewer

Delphos Viewer was designed to display the raw signal of a channel along its whitened time-domain and time-frequency versions (Fig B.3).

One can visualize the traces in different time-windows and band-pass filtering (Parameters panel in Fig B.3 A.). Interestingly, one can import markers from AnyWave, review them to reject or validate them, and even add new markers (Markers panel in Fig B.3 B.). It is also possible for the user to modify the axis of the traces and the color axis of the TF image (Axis panel in Fig B.3 C.).

In the Detection panel (Fig B.3 D.), the user can launch the detection procedure on the selected channel. The detections are represented in the TF axis as green circles for oscillations and red triangles for spikes. The marker list in the Marker panel is updated as soon as the detection stage finishes. By double-clicking on the marker, both the main AnyWave view and the viewer are centered on this detection.

The last panel (Extra panel in Fig B.3 E.) allows verifying that the normalization process worked out correctly. The top image corresponds to the histogram of the real part of the wavelet coefficients after normalization (see Fig II.3). Each line should resemble the H_0-line. The bottom plot shows the estimation of the power spectrum of the background activity and the power spectrum of the current time-window.



Time Window
○ 5s ● 1s ○ 0.25s

Filter Raw Signal
● No Filtering
○ Ripple Band Filtering
○ Fast Ripple Filtering

Filter Whitened Signal
● No Filtering
○ Ripple Band Filtering
○ Fast Ripple Filtering

TF Normalisation
● Normalisation
○ Normal
○ Real part
☐ Pulse Pattern
☐ Contour

```
Gamma  |  45.2548  |  1
Gamma  |  64  |  137.3[
Gamma  |  76.1093  |  1
Ripple  |  80.6349  |
Ripple  |  80.6349  |
Ripple  |  85.4298  |
Ripple  |  101.5937  |
Fast Ripple  |  287.350
Ripple  |  80.6349  |
Gamma  |  40.3175  |  1
Fast Ripple  |  287.350
Ripple  |  90.5097  |
Ripple  |  215.2695  |
Fast Ripple  |  322.539
Ripple  |  101.5937  |
Ripple  |  90.5097  |
Fast Ripple  |  256  |
Gamma  |  71.8376  |  1
Beta  |  20.1587  |  14
Fast Ripple  |  322.539
Fast Ripple  |  287.350
Gamma  |  50.7968  |  1
Ripple  |  143.6751  |
Ripple  |  101.5937  |
```

Sort by
● Latency ○ Channel ○ Label

Set Raw Axis
● Auto
○ [-Max, Max]
○ [Min, Max]
○ Custom uV/cm

Set Whitened Axis
● Auto
○ [-Max, Max]
○ [Min, Max]
○ Custom /cm

Set Spectrogram Colormap
● Auto
○ [0, Max]
○ [Min, Max]
○ Custom

☑ detections
☐ all candidates
☐ durations and width

Figure B.3. – Delphos Viewer interface and the different panels. A. Main frame which displays the traces in original and whitened time domains above the whitened time-frequency image. the associated panel handles the parameters such as the channel selection or the filtering settings. B. Marker panel. C. Axis Setting panel. D. Detection panel. E. Extra panel for sanity check.

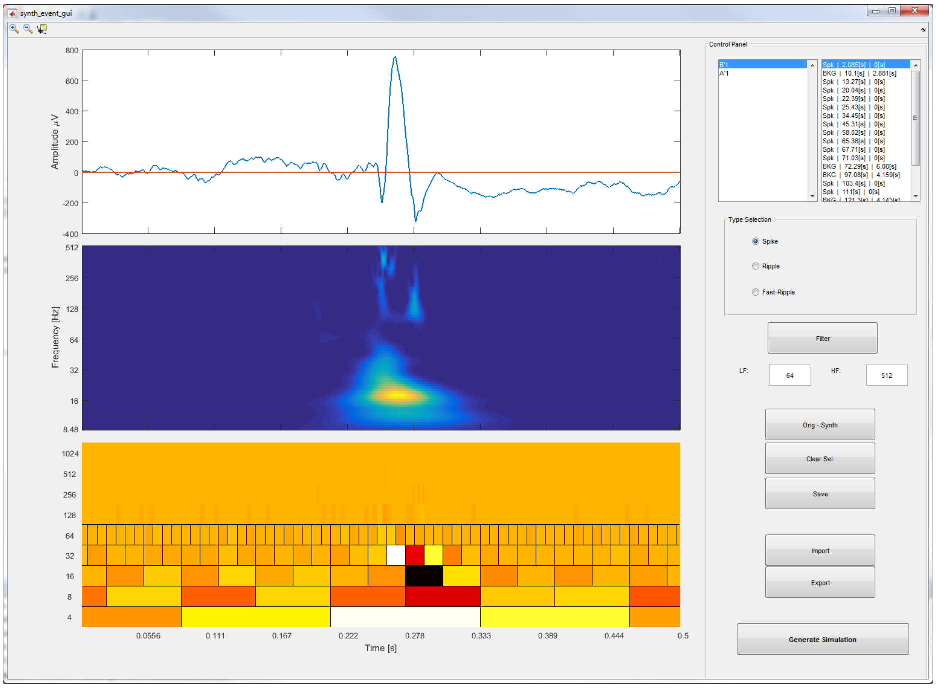
B.2. Simulation Maker

This plugin aims at extracting events of interest from real data and use these extracted events to build realistic simulations. This plugin was used in chapter III and in the related publication [Roehri et al., 2017].

Fig B.4 and Fig B.5 illustrate the process of extracting a spike and a FR, respectively[1]. Each figure is composed, from top to bottom, of the original signal (blue) with the synthesized signal (orange), the normalized continuous TF image and the discrete TF image. The user selects the coefficients of interest in the discrete TF image by left-clicking the appropriate tile. Once a tile is selected, it turns blue and the inverse transform is applied to the corresponding coefficients in order to reconstruct the signal. At the beginning of the process, the continuous TF represents the TF of the original data and during the process it shows the TF of the extracted signal.

---

[1] These figures are animated and may not work in a reader other than Adobe Acrobat Reader™



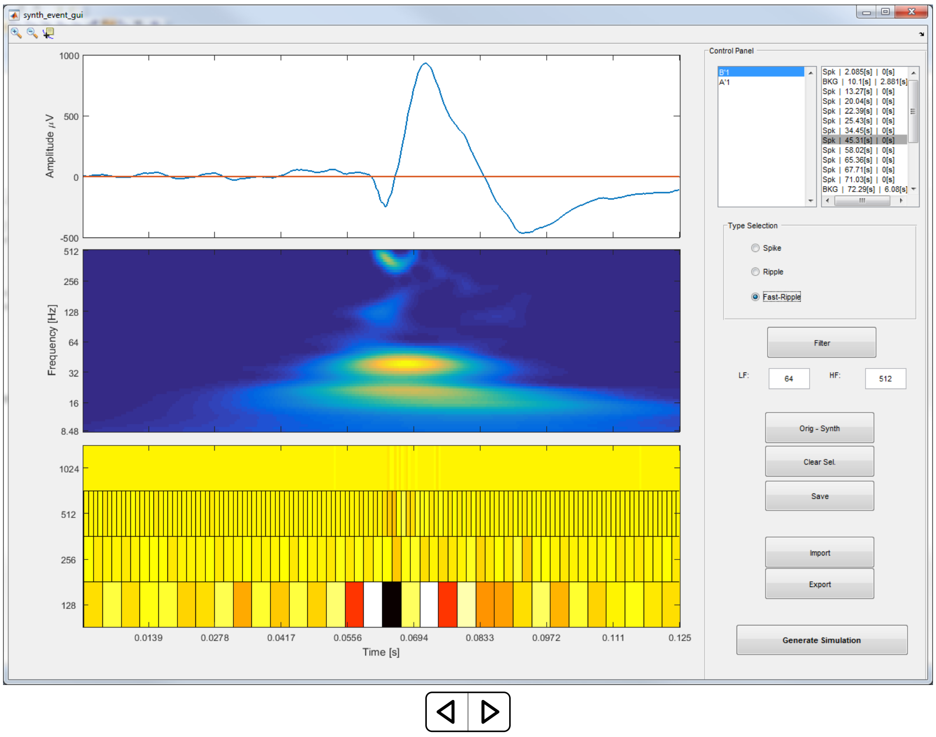
```
B'1
A'1
```

```
Spk | 2.085[s] | 0[s]
BKG | 10.1[s] | 2.881[s]
Spk | 13.27[s] | 0[s]
Spk | 20.04[s] | 0[s]
Spk | 22.39[s] | 0[s]
Spk | 25.43[s] | 0[s]
Spk | 34.45[s] | 0[s]
Spk | 45.31[s] | 0[s]
Spk | 58.02[s] | 0[s]
Spk | 65.36[s] | 0[s]
Spk | 67.71[s] | 0[s]
Spk | 71.03[s] | 0[s]
BKG | 72.29[s] | 6.08[s]
BKG | 97.08[s] | 4.159[s]
Spk | 103.4[s] | 0[s]
Spk | 111[s] | 0[s]
BKG | 121.3[s] | 4.143[s]
```

Type Selection
● Spike
○ Ripple
○ Fast-Ripple

LF: 64 HF: 512

Figure B.4. – Example of the GUI with a Spike



```
B'1
A'1
```

```
Spk | 2.085[s] | 0[s]
BKG | 10.1[s] | 2.881[s]
Spk | 13.27[s] | 0[s]
Spk | 20.04[s] | 0[s]
Spk | 22.39[s] | 0[s]
Spk | 25.43[s] | 0[s]
Spk | 34.45[s] | 0[s]
Spk | 45.31[s] | 0[s]
Spk | 58.02[s] | 0[s]
Spk | 65.36[s] | 0[s]
Spk | 67.71[s] | 0[s]
Spk | 71.03[s] | 0[s]
BKG | 72.29[s] | 6.08[s]
```

Type Selection
○ Spike
○ Ripple
● Fast-Ripple

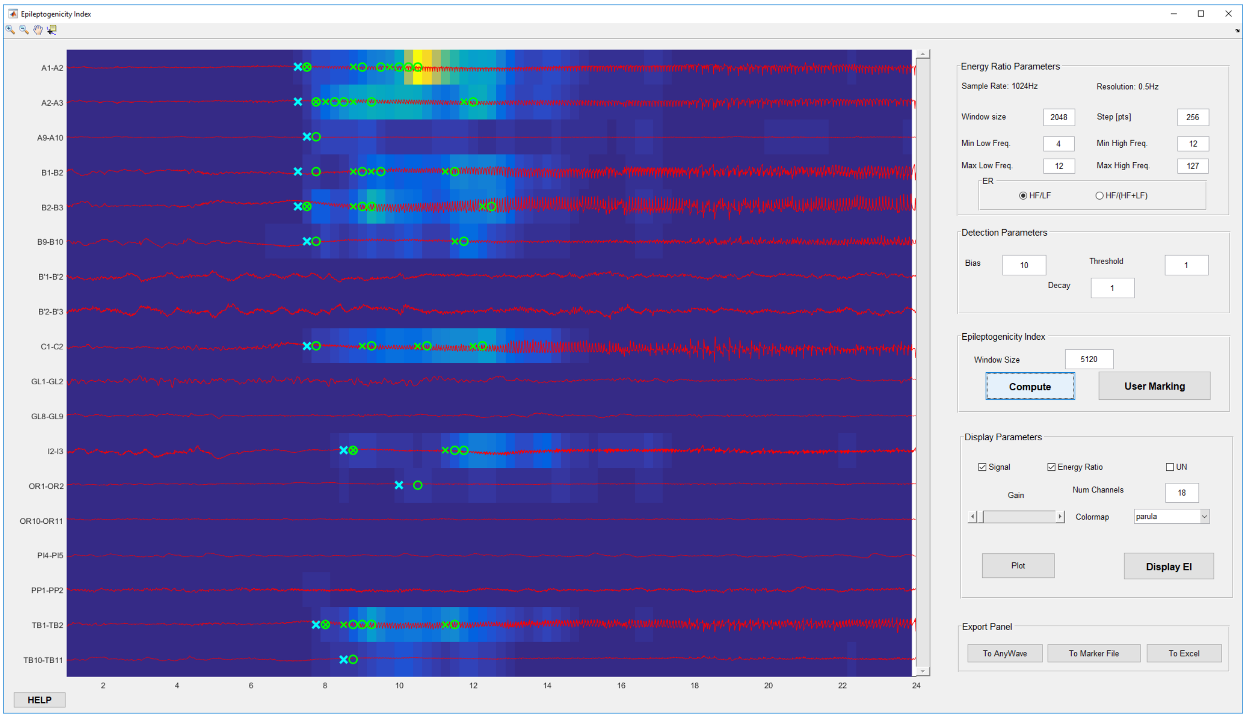
LF: 64 HF: 512

Figure B.5. – Example of the GUI with a fast ripple

B.3. Epileptogenicity Index

The Epileptogenicity Index plugin quantifies the degree of involvement of brain structures during a seizure according to their frequency content and their onset [Bartolomei et al., 2008]. This plugin was used in chapter IV and in [Pizzo et al., 2017].

The main frame plots the energy ratio (the energy of the high frequency over the energy the low frequency) for each previously selected channel and their respective trace. The user can tune each parameter of the algorithm to better fit the fast discharge. The results are shown in a bar plot and can be exported as an Excel file. The detected onsets can be sent to AnyWave as markers. Importantly, the EI algorithm does not aim to automatically detect the onset of the seizure and quantify the epileptogenicity of each structure. It rather quantifies the visual analysis of the clinician by constraining some aspect of the analysis such as the frequency bands or the criteria for the onset.



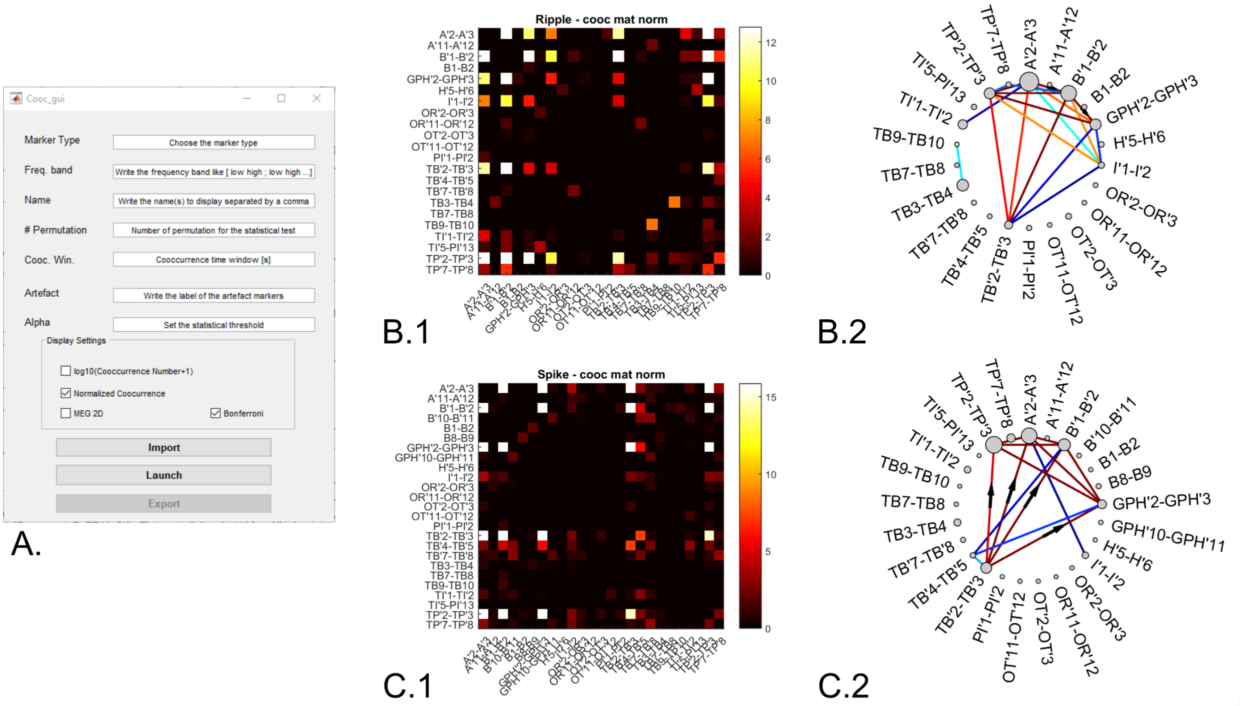
Sample Rate: 1024Hz
Resolution: 0.5Hz
Window size 2048
Step [pts] 256
Min Low Freq. 4
Min High Freq. 12
Max Low Freq. 12
Max High Freq. 127
ER ● HF/LF ○ HF/(HF+LF)
Bias 10
Threshold 1
Decay 1
Window Size 5120
☑ Signal ☑ Energy Ratio ☐ UN
Gain
Num Channels 18
Colormap parula

Figure B.6. – Interface of the Epileptogenicity Index plugin.

B.4. Co-occurrence Graph

Based on the algorithm described in [Malinowska et al., 2014], this plugin aims at finding networks using the co-occurrence of a given event in different channels (Fig B.7). Similar to the Hebbian theory, neural populations producing events together wire together. This plugin was used in a published article [Lambert et al., 2017].

The results are shown in form of matrices (Fig B.7 B.1 and C.1) and circular networks (Fig B.7 B.2 and C.2).



Cooc_gui

Marker Type: Choose the marker type
Freq. band: Write the frequency band like [low high ; low high ...]
Name: Write the name(s) to display separated by a comma
Permutation: Number of permutation for the statistical test
Cooc. Win.: Cooccurrence time window [s]
Artefact: Write the label of the artefact markers
Alpha: Set the statistical threshold

Display Settings
☐ log10(Cooccurrence Number+1)
☑ Normalized Coocurrence
☐ MEG 2D
☑ Bonferroni

Figure B.7. – Co-occurrence graph interface and examples of results. A. Main interface which allows choosing the marker types, frequency band, and statistical test parameters. B.1 Normalized co-occurrence matrix of ripples beside B.2 its representation in a circular network. C.1 Normalized co-occurrence matrix of spikes beside C.2 its representation in a circular network.

B.5. GARDEL

GARDEL is one of the softwares of the EpiTool suite [Medina et al., 2017]. Standing for *GUI for Automatic Registration and Depth Electrodes Localization*, GARDEL was designed to automatically segment depth SEEG electrodes using a post-implantation CT scan. Originally based on the work of Rodrigo Paz during for his Master thesis, I designed a new segmentation algorithm. The former algorithm was based on models of electrodes. It was a good first approximation but was not precise enough to be used to label the contacts according to an atlas. The interface is the result of the common work of Rodrigo Paz and Samuel Medina.

After removing the skull and thresholding the CT scan, the binary image stack is processed in two stages (Fig B.8 A.1). The first stage aims at defining each electrode and the second stage aims at detecting each contact of a given electrode. The binary image is dilated to bind the contacts of the same electrode together. Each connected component is obtained and corresponds to an electrode (Fig B.8 A.2). This results in masks of each individual electrode. These electrode masks are iteratively applied to the undilated, thresholded CT to obtain a binary image of the contacts. A distance transform is applied to the binary image prior to a watershed segmentation [Meyer, 1994]. The watershed technique may oversegment some contacts, i.e. identify several contacts instead of one, or miss some contacts because the contacts were too small or removed after thresholding. A robust intrinsic template of the electrodes is estimated and outliers are removed. Based on the inter-contact distance and main axis of the electrode, the missing contacts are reconstructed (Fig B.8 A.3). The final results are shown on three maximal intensity projected images in Fig B.8 B.
The sensitivity and precision of the segmentation were evaluated on thirty patients and both had a median value of 100%. The interface offers a manual reconstruction scheme for wrongly segmented or missed electrodes.

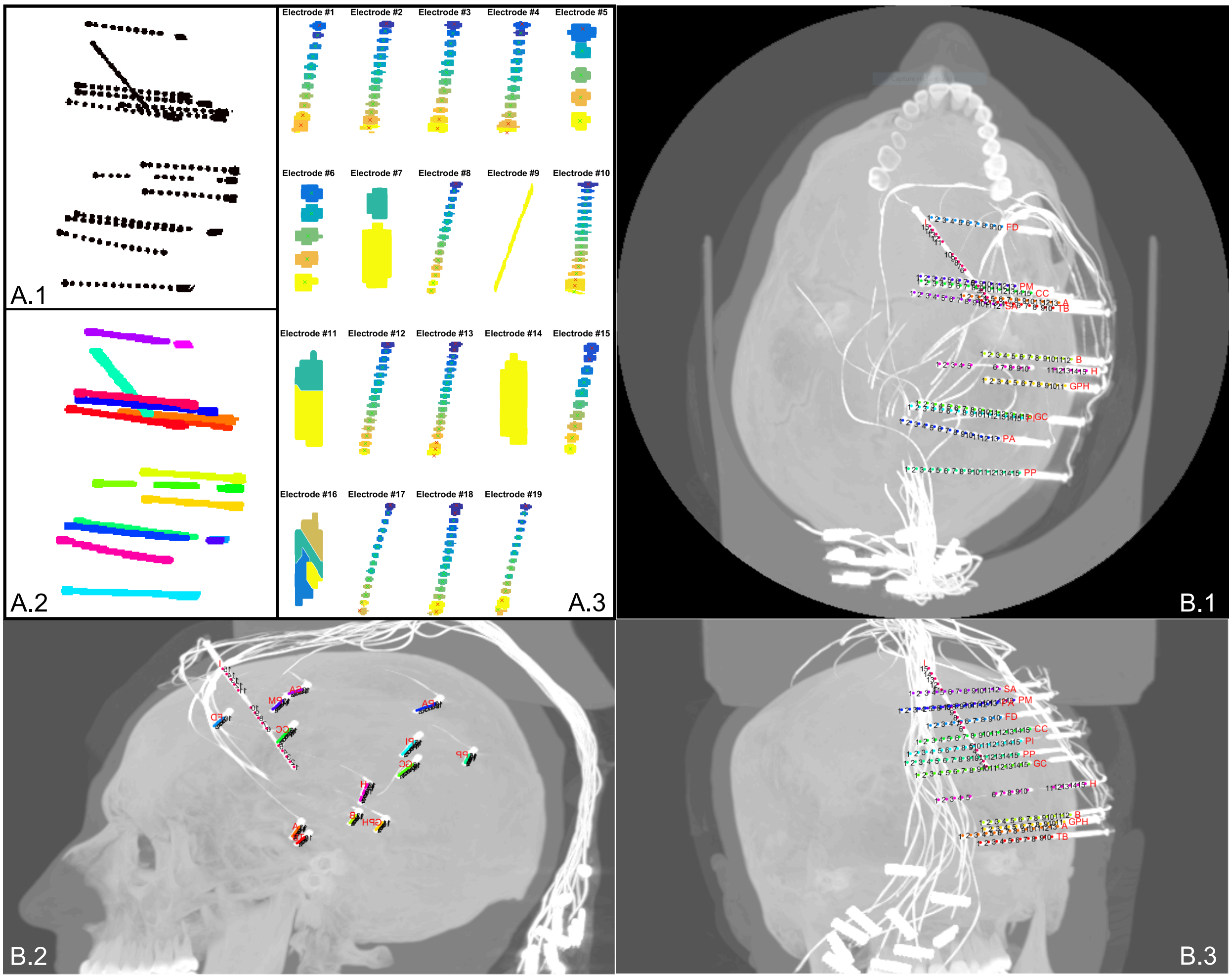


Figure B.8. – Electrode segmentation in GARDEL. A.1 projection of the thresholded CT scan, A.2 projection of the dilated and labeled electrodes, and A.3 segmentation of each contact within each electrode. B. projection of the results in the original CT scan onto three different axes.

References

[Achard, 2006] Achard, S. (2006). A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs. *Journal of Neuroscience*, 26(1):63–72.

[Achard and Bullmore, 2007] Achard, S. and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Computational Biology*, 3(2):0174–0183.

[Achard et al., 2012] Achard, S., Delon-Martin, C., Vértes, P. E., Renard, F., Schenck, M., Schneider, F., Heinrich, C., Kremer, S., and Bullmore, E. T. (2012). Hubs of brain functional networks are radically reorganized in comatose patients. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50):20608–13.

[Adams, 1868] Adams, C. D. (1868). *The Genuine Works of Hippocrates*. Dover, New York.

[Ahn et al., 2010] Ahn, Y.-Y. Y., Bagrow, J. P., and Lehmann, S. (2010). Link communities reveal multiscale complexity in networks. *Nature*, 466(7307):761–764.

[Akiyama et al., 2011] Akiyama, T., McCoy, B., Go, C. Y., Ochi, A., Elliott, I. M., Akiyama, M., Donner, E. J., Weiss, S. K., Snead, O. C., Rutka, J. T., Drake, J. M., and Otsubo, H. (2011). Focal resection of fast ripples on extraoperative intracranial EEG improves seizure outcome in pediatric epilepsy. *Epilepsia*, 52(10):1802–1811.

[Alkawadri et al., 2014] Alkawadri, R., Gaspard, N., Goncharova, I. I., Spencer, D. D., Gerrard, J. L., Zaveri, H., Duckrow, R. B., Blumenfeld, H., and Hirsch, L. J. (2014). The spatial and signal characteristics of physiologic high frequency oscillations. *Epilepsia*, 55(12):1986–1995.

[Amiri et al., 2015] Amiri, M., Lina, J.-M., Pizzo, F., and Gotman, J. (2015). High Frequency Oscillations and spikes: Separating real HFOs from false oscillations. *Clinical Neurophysiology*.

[Andrzejak et al., 2015] Andrzejak, R. G., David, O., Gnatkovsky, V., Wendling, F., Bartolomei, F., Francione, S., Kahane, P., Schindler, K., and de Curtis, M. (2015). Localization of Epileptogenic Zone on Pre-surgical Intracranial EEG Recordings: Toward a Validation of Quantitative Signal Analysis Approaches. *Brain Topography*, 28(6):832–837.

[Aubert et al., 2009] Aubert, S., Wendling, F., Regis, J., McGonigal, A., Figarella-Branger, D., Peragut, J. C., Girard, N., Chauvel, P., and Bartolomei, F. (2009). Local and remote epileptogenicity in focal cortical dysplasias and neurodevelopmental tumours. *Brain*, 132(11):3072–3086.

[Bagshaw et al., 2009] Bagshaw, A. P., Jacobs, J., LeVan, P., Dubeau, F., and Gotman, J. (2009). Effect of sleep stage on interictal high-frequency oscillations recorded from depth macroelectrodes in patients with focal epilepsy. *Epilepsia*, 50(4):617–628.

[Bancaud et al., 1965] Bancaud, J., Talairach, J., and Bonis, A. (1965). *La Stéréoencéphalographie dans l'épilepsie.* Masson, Paris.

[Bartolomei et al., 2008] Bartolomei, F., Chauvel, P., and Wendling, F. (2008). Epileptogenicity of brain structures in human temporal lobe epilepsy : a quantified study from intracerebral EEG. *Brain : a journal of neurology*, pages 1818–1830.

[Bartolomei et al., 2010] Bartolomei, F., Cosandier-Rimele, D., McGonigal, A., Aubert, S., Régis, J., Gavaret, M., Wendling, F., and Chauvel, P. (2010). From mesial temporal lobe to temporoperisylvian seizures: A quantified study of temporal lobe seizure networks. *Epilepsia*, 51(10):2147–2158.

[Bartolomei et al., 2017a] Bartolomei, F., Guye, M., and Jirsa, V. K. (2017a). Individual brain structure and modelling predict seizure propagation. *Brain : a journal of neurology*, 140(3):641–654.

[Bartolomei et al., 2017b] Bartolomei, F., Lagarde, S., Wendling, F., Mcgonigal, A., Jirsa, V., Guye, M., and Bénar, C. (2017b). Defining epileptogenic networks : Contribution of SEEG and signal analysis. *Epilepsia*, pages 1–17.

[Bartolomei et al., 2016] Bartolomei, F., Trébuchon, A., Bonini, F., Lambert, I., Gavaret, M., Woodman, M., Giusiano, B., Wendling, F., and Bénar, C. (2016). What is the concordance between the seizure onset zone and the irritative zone? A SEEG quantified study. *Clinical Neurophysiology*, 127(2):1157–1162.

[Beghi, 2017] Beghi, E. (2017). Epilepsy: New classification of seizures and epilepsies — an advance? *Nature Reviews Neurology*.

[Bénar et al., 2010] Bénar, C.-G., Chauvière, L., Bartolomei, F., and Wendling, F. (2010). Pitfalls of high-pass filtering for detecting epileptic oscillations: A technical note on "false" ripples. *Clinical Neurophysiology*, 121(3):301–310.

[Benjamini and Hochberg, 1995] Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.

[Besson et al., 2017] Besson, P., Bandt, S. K., Proix, T. E., Lagarde, S., Jirsa, V. K., Ranjeva, J.-P., Bartolomei, F., and Guye, M. (2017). Anatomic consistencies across epilepsies: a stereotactic-EEG informed high-resolution structural connectivity study. *Brain : a journal of neurology*, 140(10):1–14.

[Birot et al., 2013] Birot, G., Kachenoura, A., Albera, L., Bénar, C., and Wendling, F. (2013). Automatic detection of fast ripples. *Journal of Neuroscience Methods*, 213(2):236–249.

[Blanco et al., 2010] Blanco, J. A., Stead, M., Krieger, A., Viventi, J., Marsh, W. R., Lee, K. H., Worrell, G. A., and Litt, B. (2010). Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients. *Journal of neurophysiology*, 104(5):2900–2912.

[Bragin et al., 1999a] Bragin, A., Engel, J., Wilson, C. L., Fried, I., and Buzsáki, G. (1999a). High-frequency oscillations in human brain. *Hippocampus*, 9(2):137–142.

[Bragin et al., 1999b] Bragin, A., Engel, J., Wilson, C. L., Fried, I., and Mathern, G. W. (1999b). Hippocampal and entorhinal cortex high-frequency oscillations (100-500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures. *Epilepsia*, 40(2):127–137.

[Bragin et al., 2002] Bragin, A., Mody, I., Wilson, C. L., and Jr, J. E. (2002). Local Generation of Fast Ripples in Epileptic Brain. *The Journal of neuroscience*, 22(5):2012–2021.

[Brookes et al., 2016] Brookes, M. J., Tewarie, P. K., Hunt, B. A. E., Robson, S. E., Gascoyne, L. E., Liddle, E. B., Liddle, P. F., and Morris, P. G. (2016). A multi-layer network approach to MEG connectivity analysis. *NeuroImage*, 132:425–438.

[Bruder et al., 2016] Bruder, J. C., Dümpelmann, M., Piza, D. L., Mader, M., Schulze - Bonhage, A., and Jacobs - Le Van, J. (2016). Physiological Ripples Associated with Sleep Spindles Differ in Waveform Morphology from Epileptic Ripples. *International Journal of Neural Systems*, 27(0):1750011.

[Burnos et al., 2016] Burnos, S., Frauscher, B., Zelmann, R., Haegelen, C., Sarnthein, J., and Gotman, J. (2016). The morphology of high frequency oscillations (HFO) does not improve delineating the epileptogenic zone. *Clinical Neurophysiology*, pages 1–9.

[Burnos et al., 2014] Burnos, S., Hilfiker, P., Sürücü, O., Scholkmann, F., Krayenbühl, N., Grunwald, T., and Sarnthein, J. (2014). Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis. *PLoS ONE*, 9(4).

[Buzsáki et al., 2012] Buzsáki, G., Anastassiou, C. a., and Koch, C. (2012). The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. *Nature reviews. Neuroscience*, 13(6):407–20.

[Buzsaki et al., 1992] Buzsaki, G., Horvath, Z., Urioste, R., Hetke, J., and Wise, K. (1992). High-frequency network oscillation in the hippocampus. *Science*, 256(5059):1025–1027.

[Cardinale et al., 2013] Cardinale, F., Cossu, M., Castana, L., Casaceli, G., Schiariti, M. P., Miserocchi, A., Fuschillo, D., Moscato, A., Caborni, C., Arnulfo, G., and Lo Russo, G. (2013). Stereoelectroencephalography: Surgical methodology, safety, and stereotactic application accuracy in 500 procedures. *Neurosurgery*, 72(3):353–366.

[Cardinale et al., 2016] Cardinale, F., González-Martínez, J., and Lo Russo, G. (2016). SEEG, Happy Anniversary! *World Neurosurgery*, 85:1–2.

[Carmona et al., 1998] Carmona, R., Hwang, W.-L., and Torresani, B. (1998). *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms, with an Implementation in S*. Academic Press.

[Chabardès et al., 2005] Chabardès, S., Kahane, P., Minotti, L., Tassi, L., Grand, S., Hoffmann, D., and Benabid, A. L. (2005). The temporopolar cortex plays a pivotal role in temporal lobe seizures. *Brain*, 128(8):1818–1831.

[Chaibi et al., 2014] Chaibi, S., Lajnef, T., Sakka, Z., Samet, M., and Kachouri, A. (2014). A reliable approach to distinguish between transient with and without HFOs using TQWT and MCA. *Journal of Neuroscience Methods*, 232:36–46.

[Chaitanya et al., 2015] Chaitanya, G., Sinha, S., Narayanan, M., and Satishchandra, P. (2015). Scalp high frequency oscillations (HFOs) in absence epilepsy: An independent component analysis (ICA) based approach. *Epilepsy Research*, 115:133–140.

[Chander, 2007] Chander, R. (2007). *Algorithms to Detect High Frequency Oscillations in Human Intracerebral EEG.* Meng thesis, McGill University, Montreal.

[Chrobak and Buzsáki, 1996] Chrobak, J. J. and Buzsáki, G. (1996). High-frequency oscillations in the output networks of the hippocampal-entorhinal axis of the freely behaving rat. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(9):3056–3066.

[Cloppenborg et al., 2016] Cloppenborg, T., May, T. W., Blümcke, I., Grewe, P., Hopf, L. J., Kalbhenn, T., Pfäfflin, M., Polster, T., Schulz, R., Woermann, F. G., and Bien, C. G. (2016). Trends in epilepsy surgery: stable surgical numbers despite increasing presurgical volumes. *Journal of Neurology, Neurosurgery & Psychiatry*, pages jnnp–2016–313831.

[Colombet et al., 2015] Colombet, B., Woodman, M., Badier, J., and Bénar, C.-G. (2015). AnyWave: A cross-platform and modular software for visualizing and processing electrophysiological signals. *Journal of Neuroscience Methods*, 242:118–126.

[Crépon et al., 2010] Crépon, B., Navarro, V., Hasboun, D., Clemenceau, S., Martinerie, J., Baulac, M., Adam, C., and Le Van Quyen, M. (2010). Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*, 133(1):33–45.

[Csicsvari et al., 1999] Csicsvari, J., Hirase, H., Czurkó, a., Mamiya, a., and Buzsáki, G. (1999). Fast network oscillations in the hippocampal CA1 region of the behaving rat. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19(16):RC20.

[David et al., 2011] David, O., Blauwblomme, T., Job, A. S., Chabards, S., Hoffmann, D., Minotti, L., and Kahane, P. (2011). Imaging the seizure onset zone with stereo-electroencephalography. *Brain*, 134(10):2898–2911.

[Davis and Goadrich, 2006] Davis, J. and Goadrich, M. (2006). The Relationship Between Precision-Recall and ROC Curves. *Proceedings of the 23rd International Conference on Machine learning – ICML'06*, pages 233–240.

[Demont-Guignard et al., 2012] Demont-Guignard, S., Benquet, P., Gerber, U., Biraben, A., Martin, B., and Wendling, F. (2012). Distinct hyperexcitability mechanisms underlie fast ripples and epileptic spikes. *Annals of Neurology*, 71(3):342–352.

[Efron, 2005] Efron, B. (2005). Local false discovery rates. *Discovery*, pages 63–79.

[Fawcett, 2006] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letter*, pages 861–874.

[Fedele et al., 2017a] Fedele, T., Burnos, S., Boran, E., Krayenbuhl, N., Hilfiker, P., Grunwald, T., and Sarnthein, J. (2017a). Resection of high frequency oscillations predicts seizure outcome in the individual patient. *Scientific Reports*, 7(1).

[Fedele et al., 2017b] Fedele, T., Ramantani, G., Burnos, S., Hilfiker, P., Curio, G., Grunwald, T., Krayenbühl, N., and Sarnthein, J. (2017b). Prediction of seizure outcome improved by fast ripples detected in low-noise intraoperative corticogram. *Clinical Neurophysiology*, 128(7):–.

[Fedele et al., 2016] Fedele, T., Van, M., Burnos, S., Zweiphenning, W., van't Klooster, M., Burnos, S., Zweiphenning, W., van Klink, N., Leijten, F., Zijlmans, M., and Sarnthein, J. (2016). Automatic detection of high frequency oscillations during epilepsy surgery predicts seizure outcome. *Clinical Neurophysiology*, 127(9):3066–3074.

[Fink et al., 2015] Fink, C. G., Gliske, S., Catoni, N., and Stacey, W. C. (2015). Network mechanisms generating abnormal and normal hippocampal High Frequency Oscillations: A computational analysis. *eNeuro*, 2(3):ENEURO.0024–15.2015.

[Fischl, 2012] Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2):774–781.

[Fisher et al., 2014] Fisher, R. S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J. H., Elger, C. E., Engel, J., Forsgren, L., French, J. A., Glynn, M., Hesdorffer, D. C., Lee, B. I., Mathern, G. W., Moshé, S. L., Perucca, E., Scheffer, I. E., Tomson, T., Watanabe, M., and Wiebe, S. (2014). ILAE Official Report: A practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482.

[Fisher et al., 2017] Fisher, R. S., Cross, J. H., D'Souza, C., French, J. A., Haut, S. R., Higurashi, N., Hirsch, E., Jansen, F. E., Lagae, L., Mosh??, S. L., Peltola, J., Roulet Perez, E., Scheffer, I. E., Schulze-Bonhage, A., Somerville, E., Sperling, M., Yacubian, E. M., and Zuberi, S. M. (2017). Instruction manual for the ILAE 2017 operational classification of seizure types. *Epilepsia*, 58(4):531–542.

[Fisher et al., 2005] Fisher, R. S., Van Emde Boas, W., Blume, W., Elger, C., Genton, P., Lee, P., and Engel, J. (2005). Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4):470–472.

[Frauscher et al., 2015] Frauscher, B., Bernasconi, N., Caldairou, B., von Ellenrieder, N., Bernasconi, A., Gotman, J., and Dubeau, F. (2015). Interictal Hippocampal Spiking Influences the Occurrence of Hippocampal Sleep Spindles. *Sleep*, 38(12):1927–1933.

[Frauscher et al., 2016] Frauscher, B., Von Ellenrieder, N., Dubeau, F., and Gotman, J. (2016). EEG desynchronization during phasic REM sleep suppresses interictal epileptic activity in humans. *Epilepsia*, 57(6):879–888.

[Gallotti and Barthelemy, 2015] Gallotti, R. and Barthelemy, M. (2015). The multilayer temporal network of public transport in Great Britain. *Scientific data*, 2(6):140056.

[Gardner et al., 2007] Gardner, A. B., Worrell, G. a., Marsh, E., Dlugos, D., and Litt, B. (2007). Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings. *Clinical Neurophysiology*, 118(5):1134–1143.

[Girardeau et al., 2009] Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., and Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature neuroscience*, 12(10):1222–1223.

[Gliske et al., 2016] Gliske, S. V., Irwin, Z. T., Chestek, C., Stacey, W. C., Gliske, S. V., Irwin, Z. T., Chestek, C., Stacey, W. C., and Stacey, W. C. (2016). Effect of sampling rate and filter settings on High Frequency Oscillation detections. *Clinical Neurophysiology*, 127(9):3042–3050.

[Gnatkovsky et al., 2011] Gnatkovsky, V., Francione, S., Cardinale, F., Mai, R., Tassi, L., Lo Russo, G., and De Curtis, M. (2011). Identification of reproducible ictal patterns based on quantified frequency analysis of intracranial EEG signals. *Epilepsia*, 52(3):477–488.

[Gollwitzer et al., 2016] Gollwitzer, S., Valente, I., Rodionov, R., Scott, C., Ritter, L. M., Wehner, T., Hamer, H. M., Bartolomei, F., and Diehl, B. (2016). Visual and semiautomated evaluation of epileptogenicity in focal cortical dysplasias - An intracranial EEG study. *Epilepsy and Behavior*, 58:69–75.

[Grandchamp and Delorme, 2011] Grandchamp, R. and Delorme, A. (2011). Single-trial normalization for event-related spectral decomposition reduces sensitivity to noisy trials. *Frontiers in Psychology*, 2(SEP):1–14.

[Grün et al., 2002a] Grün, S., Diesmann, M., and Aertsen, A. (2002a). Unitary events in multiple single-neuron spiking activity: I. Detection and Significance. *Neural computation*, 14(1):81–119.

[Grün et al., 2002b] Grün, S., Diesmann, M., and Aertsen, A. (2002b). Unitary events in multiple single-neuron spiking activity: II. Nonstationary data. *Neural computation*, 14(1):81–119.

[Guillon et al., 2016] Guillon, J., Attal, Y., Colliot, O., La Corte, V., Dubois, B., Schwartz, D., Chavez, M., and Fallani, F. D. V. (2016). Loss of brain inter-frequency hubs in Alzheimer's disease. *Scientific Reports*, pages 1–13.

[Haegelen et al., 2013] Haegelen, C., Perucca, P., Châtillon, C. E., Andrade-Valença, L., Zelmann, R., Jacobs, J., Collins, D. L., Dubeau, F., Olivier, A., and Gotman, J. (2013). High-frequency oscillations, extent of surgical resection, and surgical outcome in drug-resistant focal epilepsy. *Epilepsia*, 54(5):848–857.

[He, 2014] He, B. J. (2014). Scale-free brain activity: Past, present, and future. *Trends in Cognitive Sciences*, 18(9):480–487.

[Höller et al., 2015] Höller, Y., Kutil, R., Klaffenböck, L., Thomschewski, A., Höller, P. M., Bathke, A. C., Jacobs, J., Taylor, A. C., Nardone, R., and Trinka, E. (2015). High-frequency oscillations in epilepsy and surgical outcome. A meta-analysis. *Frontiers in Human Neuroscience*, 9(October):1–14.

[Huang et al., 1998] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Yen, N.-c., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454(1971):995, 903.

[Iida and Otsubo, 2017] Iida, K. and Otsubo, H. (2017). Stereoelectroencephalography: Indication and Efficacy. *Neurologia medico-chirurgica*, pages 1–11.

[Jacobs et al., 2011] Jacobs, J., Kobayashi, K., and Gotman, J. (2011). High-frequency changes during interictal spikes detected by time-frequency analysis. *Clinical Neurophysiology*, 122(1):32–42.

[Jacobs et al., 2008] Jacobs, J., LeVan, P., Chander, R., Hall, J., Dubeau, F., and Gotman, J. (2008). Interictal high-frequency oscillations (80-500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain. *Epilepsia*, 49(11):1893–1907.

[Jacobs et al., 2009] Jacobs, J., Levan, P., Châtillon, C. E., Olivier, A., Dubeau, F., and Gotman, J. (2009). High frequency oscillations in intracranial EEGs mark epileptogenicity rather than lesion type. *Brain*, 132(4):1022–1037.

[Jacobs et al., 2012] Jacobs, J., Staba, R., Asano, E., Otsubo, H., Wu, J. Y., Zijlmans, M., and Mohamed, I. (2012). Progress in Neurobiology High-frequency oscillations (HFOs) in clinical epilepsy. *Progress in Neurobiology*, 98:302–315.

[Jacobs et al., 2010] Jacobs, J., Zijlmans, M., Zelmann, R., Chatillon, C. É., Hall, J., Olivier, A., Dubeau, F., and Gotman, J. (2010). High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery. *Annals of Neurology*, 67(2):209–220.

[Jiménez-Jiménez et al., 2015] Jiménez-Jiménez, D., Nekkare, R., Flores, L., Chatzidimou, K., Bodi, I., Honavar, M., Mullatti, N., Elwes, R. D. C., Selway, R. P., Valentín, A., and Alarcón, G. (2015). Prognostic value of intracranial seizure onset patterns for surgical outcome of the treatment of epilepsy. *Clinical Neurophysiology*, 126(2):257–267.

[Jirsa et al., 2016] Jirsa, V., Proix, T., Perdikis, D., Woodman, M., Wang, H., Gonzalez-Martinez, J., Bernard, C., Bénar, C.-G., Guye, M., Chauvel, P., and Bartolomei, F. (2016). The Virtual Epileptic Patient: Individualized whole-brain models of epilepsy spread. *NeuroImage*.

[Jmail et al., 2016] Jmail, N., Gavaret, M., Bartolomei, F., and Bénar, C.-G. (2016). Despiking SEEG signals reveals dynamics of gamma band preictal activity. *Physiological Measurement*, in press.

[Jmail et al., 2011] Jmail, N., Gavaret, M., Wendling, F., and Kachouri, A. (2011). A comparison of methods for separation of transient and oscillatory signals in EEG. *Journal of Neuroscience Methods*, 199(2):273–289.

[Jobst et al., 2015] Jobst, B. C., Cascino, G. D., and R, K. (2015). Resective Epilepsy Surgery for Drug-Resistant Focal Epilepsy. *Jama*, 313(3):285.

[Jrad et al., 2016] Jrad, N., Kachenoura, A., Merlet, I., Bartolomei, F., Nica, A., Biraben, A., and Wendling, F. (2016). Automatic detection and classification of High Frequency Oscillations in depth-EEG signals. *IEEE Transactions on Biomedical Engineering*, PP(99):1–1.

[Kahane and Bartolomei, 2010] Kahane, P. and Bartolomei, F. (2010). Temporal lobe epilepsy and hippocampal sclerosis: Lessons from depth EEG recordings. *Epilepsia*, 51(SUPPL. 1):59–62.

[Kahane et al., 2006] Kahane, P., Landré, E., Minotti, L., Francione, S., and Ryvlin, P. (2006). The Bancaud and Talairach view on the epileptogenic zone: A working hypothesis. *Epileptic Disorders*, 8(SUPPL. 2):16–26.

[Kamath, 2013] Kamath, C. (2013). A New Approach to Detect Epileptic Seizures in Electroencephalograms Using Teager Energy. *ISRN Biomedical Engineering*, pages 1–14.

[Kamwa et al., 2011] Kamwa, I., Pradhan, A. K., Member, S., and Joós, G. (2011). Robust Detection and Analysis of Power System Oscillations Using the Teager-Kaiser Energy Operator. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 26(1):323–333.

[Kerber et al., 2014] Kerber, K., Dümpelmann, M., Schelter, B., Van, P. L., Korinthenberg, R., Schulze-bonhage, A., and Jacobs, J. (2014). Differentiation of specific ripple patterns helps to identify epileptogenic areas for surgical procedures. *Clinical Neurophysiology*, 125(7):1339–1345.

[Khalilov et al., 2005] Khalilov, I., Le Van Quyen, M., Gozlan, H., and Ben-Ari, Y. (2005). Epileptogenic actions of GABA and fast oscillations in the developing hippocampus. *Neuron*, 48(5):787–796.

[Kivelä et al., 2014] Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, 2(3):203–271.

[Kobayashi et al., 2009] Kobayashi, K., Jacobs, J., and Gotman, J. (2009). Detection of changes of high-frequency activity by statistical time-frequency analysis in epileptic spikes. *Clinical Neurophysiology*, 120(6):1070–1077.

[Kobayashi et al., 2010] Kobayashi, K., Watanabe, Y., Inoue, T., Oka, M., Yoshinaga, H., and Ohtsuka, Y. (2010). Scalp-recorded high-frequency oscillations in childhood sleep-induced electrical status epilepticus. *Epilepsia*, 51(10):2190–2194.

[Krieg et al., 2011] Krieg, J., Trébuchon-Da Fonseca, A., Martínez-Montes, E., Marquis, P., Liégeois-Chauvel, C., and Bénar, C.-G. (2011). A comparison of methods for assessing alpha phase resetting in electrophysiology, with application to intracerebral EEG in visual areas. *NeuroImage*, 55(1):67–86.

[Lachaux et al., 2012] Lachaux, J.-P., Axmacher, N., Mormann, F., Halgren, E., and Crone, N. E. (2012). High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research. *Progress in Neurobiology*, 98(3):279–301.

[Lagarde et al., 2016] Lagarde, S., Bonini, F., McGonigal, A., Chauvel, P., Gavaret, M., Scavarda, D., Carron, R., Régis, J., Aubert, S., Villeneuve, N., Giusiano, B., Figarella-Branger, D., Trebuchon, A., and Bartolomei, F. (2016). Seizure-onset patterns in focal cortical dysplasia and neurodevelopmental tumors: Relationship with surgical prognosis and neuropathologic subtypes. *Epilepsia*, pages 1–10.

[Lambert et al., 2017] Lambert, I., Roehri, N., Giusiano, B., Carron, R., Wendling, F., Bénar, C.-G., and Bartolomei, F. (2017). Brain Regions and Epileptogenicity Influence Epileptic Interictal Spikes Production and Propagation during NREM Sleep in comparison with Wakefulness. *Epilepsia*, in press.

[Lee et al., 2000] Lee, S.-A., Spencer, D. D., and Spencer, S. S. (2000). Intracranial EEG Seizure-Onset Patterns in Neocortical Epilepsy. *Epilepsia*, 41(3):297–307.

[Lilly and Olhede, 2009] Lilly, J. M. and Olhede, S. C. (2009). Higher-order properties of analytic wavelets. *IEEE Transactions on Signal Processing*, 57(1):146–160.

[Linnet, 1988] Linnet, K. (1988). A Review on the Methodologyfor Assessing Diagnostic Tests. *Clinical Chemistry*, 34(7):1379–1386.

[Liu et al., 2016] Liu, S., Sha, Z., Sencer, A., Aydoseli, A., Bebek, N., Abosch, A., Henry, T., Gurses, C., and Ince, N. F. (2016). Exploring the time-frequency content of high frequency oscillations for automated identification of seizure onset zone in epilepsy. *Journal of neural engineering*, 13(2):026026.

[Ljung, 1987] Ljung, L. (1987). System Identification Theory for User.pdf.

[Lüders et al., 2006] Lüders, H. O., Najm, I., Nair, D., Widdess-Walsh, P., and Bingman, W. (2006). The epileptogenic zone: General principles. *Epileptic Disorders*, 8(SUPPL. 2):1–9.

[Ma et al., 2015] Ma, H., Bandos, A. I., and Gur, D. (2015). On the use of partial area under the ROC curve for comparison of two diagnostic tests. *Biometrical Journal*, 57(2):304–320.

[Malinowska et al., 2014] Malinowska, U., Badier, J.-m., Gavaret, M., Bartolomei, F., Chauvel, P., and Bénar, C. G. (2014). Interictal Networks in Magnetoencephalography. *Human Brain Mapping*, 35(1):2789–2805.

[Malinowska et al., 2015] Malinowska, U., Bergey, G. K., Harezlak, J., and Jouny, C. C. (2015). Identification of seizure onset zone and preictal state based on characteristics of high frequency oscillations. *Clinical Neurophysiology*, 126(8):1505–1513.

[Mallat, 2008] Mallat, S. (2008). *A Wavelet Tour of Signal Processing : The Sparse way*. Elsevier.

[Marchi et al., 2016] Marchi, A., Bonini, F., Lagarde, S., McGonigal, A., Gavaret, M., Scavarda, D., Carron, R., Aubert, S., Villeneuve, N., Médina Villalon, S., Bénar, C., Trebuchon, A., and Bartolomei, F. (2016). Occipital and occipital "plus" epilepsies: A study of involved epileptogenic networks through SEEG quantification. *Epilepsy & Behavior*, 62:104–114.

[Mari et al., 2012] Mari, F., Zelmann, R., Andrade-Valenca, L., Dubeau, F., and Gotman, J. (2012). Continuous high-frequency activity in mesial temporal lobe structures. *Epilepsia*, 53(5):797–806.

[Matsumoto et al., 2013] Matsumoto, J. Y., Stead, M., Kucewicz, M. T., Matsumoto, A. J., Peters, P. A., Brinkmann, B. H., Danstrom, J. C., Goerss, S. J., Marsh, W. R., Meyer, F. B., and Worrell, G. A. (2013). Network oscillations modulate interictal epileptiform spike rate during human memory. *Brain : a journal of neurology*, pages 2444–2456.

[McGonigal et al., 2007] McGonigal, A., Bartolomei, F., Régis, J., Guye, M., Gavaret, M., Fonseca, A. T. D., Dufour, H., Figarella-Branger, D., Girard, N., Péragut, J. C., and Chauvel, P. (2007). Stereoelectroencephalography in presurgical assessment of MRI-negative epilepsy. *Brain*, 130(12):3169–3183.

[McNeil et al., 1975] McNeil, B. J., Keeler, E., and Adelstein, S. J. (1975). Primer On Certain Elements of Medical Decision Making. *The New England Journal of Medicine*, 293(5):211–215.

[Medina et al., 2017] Medina, S., Paz, R., Roehri, N., Carron, R., Brovelli, A., Dubarry, A.-S., Trébuchon, A., Coulon, O., Bartolomei, F., and Bénar, C.-G. (2017). EpiTools: A software suite for presurgical brain mapping in epilepsy (under review).

[Melani et al., 2013] Melani, F., Zelmann, R., Mari, F., and Gotman, J. (2013). Continuous High Frequency Activity: A peculiar SEEG pattern related to specific brain regions. *Clinical Neurophysiology*, 124(8):1507–1516.

[Menendez de la Prida et al., 2015] Menendez de la Prida, L., Staba, R. J., and Dian, J. A. (2015). Conundrums of high-frequency oscillations (80-800 Hz) in the epileptic brain. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, 32(3):207–19.

[Meyer, 1994] Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing*, 38(1):113–125.

[Mooij et al., 2017] Mooij, A. H., Raijmann, R. C. M. A., Jansen, F. E., Braun, K. P. J., and Zijlmans, M. (2017). Physiological Ripples (± 100 Hz) in Spike-Free Scalp EEGs of Children With and Without Epilepsy. *Brain Topography*, 0(0):0.

[Navarrete et al., 2016] Navarrete, M., Alvarado-Rojas, C., Le Van Quyen, M., and Valderrama, M. (2016). RIPPLELAB: A comprehensive application for the detection, analysis and classification of high frequency oscillations in electroencephalographic signals. *PloS one*, page In press.

[Nissen et al., 2016a] Nissen, I. A., Stam, C. J., Reijneveld, J. C., van Straaten, I. E., Hendriks, E. J., Baayen, J. C., De Witt Hamer, P. C., Idema, S., and Hillebrand, A. (2016a). Identifying the epileptogenic zone in interictal resting-state MEG source-space networks. *Epilepsia*, pages 1–12.

[Nissen et al., 2016b] Nissen, I. A., van Klink, N. E., Zijlmans, M., Stam, C. J., and Hillebrand, A. (2016b). Brain areas with epileptic high frequency oscillations are functionally isolated in MEG virtual electrode networks. *Clinical Neurophysiology*.

[Pizzo et al., 2015] Pizzo, F., Ferrari-Marinho, T., Amiri, M., Frauscher, B., Dubeau, F., and Gotman, J. (2015). When spikes are symmetric, ripples are not: Bilateral spike and wave above 80Hz in focal and generalized epilepsy. *Clinical Neurophysiology*, 127(3):1794–1802.

[Pizzo et al., 2017] Pizzo, F., Roehri, N., Catenoix, H., Medina, S., McGonigal, A., Giusiano, B., Carron, R., Scavarda, D., Ostrowsky, K., Lepine, A., Boulogne, S., Scholly, J., Hirsch, E., Rheims, S., Bénar, C.-G., and Bartolomei, F. (2017). Epileptogenic networks in nodular heterotopia: A stereoelectroencephalography study. *Epilepsia*, pages 1–12.

[Pugliatti et al., 2007] Pugliatti, M., Beghi, E., Forsgren, L., Ekman, M., and Sobocki, P. (2007). Estimating the cost of epilepsy in Europe: A review with economic modeling. *Epilepsia*, 48(12):2224–2233.

[Roehri et al., 2016] Roehri, N., Lina, J.-M., Mosher, J. C., Bartolomei, F., and Bénar, C.-G. (2016). Time-frequency strategies for increasing high frequency oscillation detectability in intracerebral. *IEEE Transactions on Biomedical Engineering*, 63(12):2595–2606.

[Roehri et al., 2017] Roehri, N., Pizzo, F., Bartolomei, F., Wendling, F., and Bénar, C.-G. (2017). What are the assets and weaknesses of HFO detectors? A benchmark framework based on realistic simulations. *PLoS ONE*, 12(4):e0174702.

[Rossi et al., 1984] Rossi, G., Colocchio, G., and Pola, P. (1984). Interictal Epileptic Activity During Sleep: A Stereo-EEG Study In Patients With Partial Epilepsy. *Electroemephalography and clintcal Neurophrsiologv*, 58:97 106.

[Rubinov and Sporns, 2010] Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069.

[Rubinov and Sporns, 2011] Rubinov, M. and Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage*, 56(4):2068–2079.

[Sammaritano et al., 1991] Sammaritano, M., Gigli, G. L., and Gotman, J. (1991). Interictal spiking during wakefulness and sleep and the localization of foci in temporal lobe epilepsy. *Neurology*, 41(2, Part 1):290–290.

[Schuele, 2016] Schuele, S. U. (2016). Stereoelectroencephalography. *Journal of Clinical Neurophysiology*, 33(6):477.

[Shamas et al., 2016] Shamas, M., Benquet, P., Merlet, I., El Falou, W., Khalil, M., and Wendling, F. (2016). Computational modeling of high frequency oscillations recorded with clinical intracranial macroelectrodes. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016-Octob:1014–1017.

[Shimamoto et al., 2017] Shimamoto, S., Waldman, Z., Orosz, I., Song, I., Bragin, A., Fried, I., Engel, J., Staba, R., Sharan, A., Wu, C., Sperling, M. R., and Weiss, S. A. (2017). Utilization of independent component analysis for accurate pathological ripple detection in intracranial EEG recordings recorded extra- and intra-operatively. *Clinical Neurophysiology*.

[Spring et al., 2017] Spring, A. M., Pittman, D. J., Aghakhani, Y., Jirsch, J., Pillay, N., Bello-Espinosa, L. E., Josephson, C., and Federico, P. (2017). Interrater reliability of visually evaluated high frequency oscillations. *Clinical Neurophysiology*, 128(3):433–441.

[Staba et al., 2002] Staba, R. J., Wilson, C. L., Bragin, A., Fried, I., and Engel, J. (2002). Quantitative analysis of high-frequency oscillations (80-500 Hz) recorded in human epileptic hippocampus and entorhinal cortex. *Journal of neurophysiology*, 88(4):1743–1752.

[Stefan et al., 2008] Stefan, H., Hopfengärtner, R., Kreiselmeyer, G., Weigel, D., Rampp, S., Kerling, F., Blümcke, I., and Buchfelder, M. (2008). Interictal triple ECoG characteristics of temporal lobe epilepsies: An intraoperative ECoG analysis correlated with surgical outcome. *Clinical Neurophysiology*, 119(3):642–652.

[Talairach and Bancaud, 1966] Talairach, J. and Bancaud, J. (1966). Lesion, "Irritative" Zone and Epileptogenic Focus. *Confin. neurol*, 27(1):91–94.

[Tezer et al., 2014] Tezer, F. I., Rémi, J., Erbil, N., Noachtar, S., and Saygi, S. (2014). A reduction of sleep spindles heralds seizures in focal epilepsy. *Clinical Neurophysiology*, 125(11):2207–2211.

[Urrestarazu et al., 2007] Urrestarazu, E., Chander, R., Dubeau, F., and Gotman, J. (2007). Interictal high-frequency oscillations (100-500 Hz) in the intracerebral EEG of epileptic patients. *Brain*, 130(9):2354–2366.

[van Klink et al., 2015] van Klink, N., Arjan Hillebrand, and Maeike Zijlmans (2015). Identification of epileptic high frequency oscillations in the time domain by using MEG beamformer-based virtual sensors. *Clinical Neurophysiology*.

[van't Klooster et al., 2015] van't Klooster, M. A., Van Klink, N. E. C., Leijten, F. S. S., Zelmann, R., Gebbink, T. A., Gosselaar, P. H., Braun, K. P. J., Huiskamp, G. J. M., and Zijlmans, M. (2015). Residual fast ripples in the intraoperative corticogram predict epilepsy surgery outcome. *Neurology*, 85(2):120–128.

[van't Klooster et al., 2011] van't Klooster, M. A., Zijlmans, M., Leijten, F. S. S., Ferrier, C. H., Van Putten, M. J. A. M., and Huiskamp, G. J. M. (2011). Time-frequency analysis of single pulse electrical stimulation to assist delineation of epileptogenic cortex. *Brain*, 134(10):2855–2866.

[Ventosa et al., 2008] Ventosa, S., Simon, C., Schimmel, M., Danobeitia, J. J., and Manuel, A. (2008). The S-transform from a wavelet point of view. *IEEE Transactions on Signal Processing*, 56(7 I):2771–2780.

[von Ellenrieder et al., 2016a] von Ellenrieder, N., Frauscher, B., Dubeau, F., and Gotman, J. (2016a). Interaction with slow waves during sleep improves discrimination of physiologic and pathologic high-frequency oscillations (80-500 Hz). *Epilepsia*, pages 869–878.

[von Ellenrieder et al., 2016b] von Ellenrieder, N., Pellegrino, G., Hedrich, T., Gotman, J., Lina, J.-M., Grova, C., and Kobayashi, E. (2016b). Detection and Magnetic Source Imaging of Fast Oscillations (40–160 Hz) Recorded with Magnetoencephalography in Focal Epilepsy Patients. *Brain Topography*.

[Waldman et al., 2017] Waldman, Z., Shimamoto, S., Song, I., Orosz, I., Bragin, A., Fried, I., Engel, J., Staba, R., Sperling, M. R., and Weiss, S. A. (2017). A Method for the Topographical Identification and Quantification of High Frequency Oscillations in Intracranial Electroencephalography Recordings. *Clinical Neurophysiology*.

[Wang et al., 2013] Wang, S., Wang, I. Z., Bulacio, J. C., Mosher, J. C., Gonzalez-Martinez, J., Alexopoulos, A. V., Najm, I. M., and So, N. K. (2013). Ripple classification helps to localize the seizure-onset zone in neocortical epilepsy. *Epilepsia*, 54(2):370–376.

[Weber, 2005] Weber, M. (2005). Épilepsie : la maladie aux mille noms. *Épilepsies*, 17(3):172–175.

[West et al., 2015] West, S., Sj, N., Cotton, J., Gandhi, S., Weston, J., Ramirez, R., and Newton, R. (2015). Surgery for epilepsy. *Cochrane library*, (7).

[Worrell et al., 2012] Worrell, G. A., Jerbi, K., Kobayashi, K., Lina, J. M., Zelmann, R., and Quyen, M. L. V. (2012). Progress in Neurobiology Recording and analysis techniques for high-frequency oscillations. *Progress in Neurobiology*, 98(3):265–278.

[Wu et al., 2010] Wu, J. Y., Sankar, R., Lerner, J. T., Matsumoto, J. H., Vinters, H. V., and Mathern, G. W. (2010). Removing interictal fast ripples on electrocorticography linked with seizure freedom in children. *Neurology*, 75(19):1686–1694.

[Zelmann et al., 2014] Zelmann, R., Lina, J. M., Schulze-Bonhage, a., Gotman, J., and Jacobs, J. (2014). Scalp EEG is not a blur: it can see high frequency oscillations although their generators are small. *Brain topography*, 27(5):683–704.

[Zelmann et al., 2010] Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Dubeau, F., and Gotman, J. (2010). Automatic detector of High Frequency Oscillations for human recordings with macroelectrodes. *Conf Proc IEEE Eng Med Biol Soc*, 123(3):106–116.

[Zelmann et al., 2012] Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Dubeau, F., and Gotman, J. (2012). A comparison between detectors of high frequency oscillations. *Clinical Neurophysiology*, 123(1):106–116.

[Zelmann et al., 2009] Zelmann, R., Zijlmans, M., Jacobs, J., Châtillon, C. E., and Gotman, J. (2009). Improving the identification of High Frequency Oscillations. *Clinical Neurophysiology*, 120(8):1457–1464.

[Zerouali et al., 2013] Zerouali, Y., Herry, C. L., Jemel, B., and Lina, J. M. (2013). Localization of synchronous cortical neural sources. *IEEE Transactions on Biomedical Engineering*, 60(3):770–780.

[Zijlmans et al., 2009] Zijlmans, M., Jacobs, J., Zelmann, R., Dubeau, F., and Gotman, J. (2009). High-frequency oscillations mirror disease activity in patients with epilepsy. *Neurology*, 72(11):979–986.

[Zijlmans et al., 2017] Zijlmans, M., Worrell, G. A., Dümpelmann, M., Stieglitz, T., Barborica, A., Heers, M., Ikeda, A., Usui, N., and Le Van Quyen, M. (2017). How to record high-frequency oscillations in epilepsy: A practical guideline. *Epilepsia*, pages 1–11.

[Zweig and Campbell, 1993] Zweig, M. H. and Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39(4):561–577.