

Géométrie des signatures temps-fréquence de HFO dans les EEG

Nemo Fournier

Stage réalisé sous l'encadrement de Paulo Gonçalves et Patrick Flandrin

Juin et Juillet 2018

Résumé

On s'intéresse dans ce travail à l'analyse de signaux issus d'enregistrements encéphalographiques à forte résolution fréquentielle à la recherche d'évènements qualifiés dans la littérature de *High Frequency Oscillations* (HFO), phénomènes oscillatoires de fréquence élevée (dans la bande 80Hz-500Hz) qui pourraient constituer des marqueurs des zones épileptogènes du cerveau. On introduit des méthodes temps-fréquence, judicieuses vis à vis des caractéristiques connues des HFO et des difficultés posées par leur détection, et applique une méthode, originale dans ce cadre, fournissant un nouveau point de vue géométrique sur les évènements analysés.

Une preuve de concept du détecteur développé lors de ce stage est disponible à l'adresse <https://github.com/Nemh/hfo-zeroes>.

Table des matières

1	Introduction, motivations	2
2	<i>High Frequency Oscillations</i>, de quoi parle-t-on là?	2
3	Temps-Fréquence [5]	5
4	Détection de HFO avec des outils temps-fréquence	8
	4.1 Approches énergétiques	9
	4.2 Approche géométrique	10
5	Conclusion	15
6	Références	17
A	Inégalité d'Heisenberg-Gabor	18
B	Contexte institutionnel et social du stage	19

1 Introduction, motivations

L'intérêt porté à l'encéphalographie s'est longtemps focalisé sur la bande dite de Berger, des signaux de fréquence comprise entre 0Hz et 70Hz. Les progrès actuels dans les méthodes d'acquisition de l'activité cérébrale ont ouvert la voie à l'étude de signaux de fréquences plus élevées. Ainsi, en épileptologie, des phénomènes oscillatoires se déployant dans une bande de 80Hz à 500Hz, dits *High Frequency Oscillations* (HFO), ont pu être étudiés, récemment, comme de possibles marqueurs de zones du cerveau propices à l'apparition de crises épileptiques [8]. Il est espéré qu'une bonne connaissance de tels événements pourra améliorer le traitement chirurgical de l'épilepsie et les résultats de telles pratiques pour les patients dont la pathologie est résistante aux médicaments (soit près de 30% des patients). Les avancées susnommées dans la résolution fréquentielle des mesures se sont accompagnées de progrès dans la résolution spatiale des acquisitions, avec l'émergence de techniques telles que l'électroencéphalographie intracrânienne stéréotaxique. De telles méthodes permettant une couverture plus fine des zones du cerveau étudiées et de dresser une correspondance précise entre chaque mesure et la zone du cerveau à l'origine de l'activité mesurée. Cependant cette amélioration de la résolution spatiale s'accompagne d'une explosion de la quantité de données, et rend compliquée, sinon au moins chronophage l'analyse manuelle des mesures : à titre d'exemple, une exploration visuelle d'un enregistrement de 10 minutes sur 10 canaux spatiaux, dans le but de localiser de potentielles HFO peut mobiliser jusqu'à une dizaine d'heures de praticien [3]. Il est alors judicieux de se pencher sur la détection automatique de telles oscillations. De nombreuses méthodes ont été proposées. Cependant le manque de définition rigoureuse — en tout cas universellement acceptée et utilisée comme standard — des HFO, associée à la part de subjectivité des classifications manuelles par les praticiens et à la variabilité du matériel utilisé et des jeux de données, peinent encore à faire émerger une méthodologie dont la fiabilité serait assurée et quantifiée.

Dans ce travail, nous dresserons d'abord en section 2 un portrait un peu plus précis de ces oscillations haute fréquence, leurs intérêts, et caractéristiques connues dans la littérature, avant de nous intéresser plus avant à la question de leur détection automatique, se concentrant sur les approches temps-fréquence (que l'on introduit en section 3), puis d'explorer un nouveau point de vue géométrique sur l'analyse temps-fréquence des HFO pour esquisser une méthode de détection et d'extraction de caractéristiques, en section 4.2.

2 *High Frequency Oscillations*, de quoi parle-t-on là?

Les oscillations haute fréquence sont caractérisées, dans la littérature, comme étant des augmentations soudaines, courtes et de faible amplitude de l'énergie du signal cérébral dans une bande de fréquence limitée, comprise entre 80Hz et 500Hz voire 600Hz, la définition exacte de la bande considérée fluctuant selon les articles. Ces brusques augmentations résulteraient d'une éphémère synchronisation d'une population neuronale localisée [9]. Il a été observé que les zones du cerveau où les HFO étaient les plus fréquentes pourraient coïncider avec les *Seizure Onset Zones* (SOZ), soit des zones dans lesquelles se concentrent les tissus épileptogènes et donc des zones d'intérêt particulier pour les approches chirurgicales de la maladie. Cette surreprésentation dans les zones épileptogènes est particulièrement vérifiée pour les HFO évoluant dans la bande de

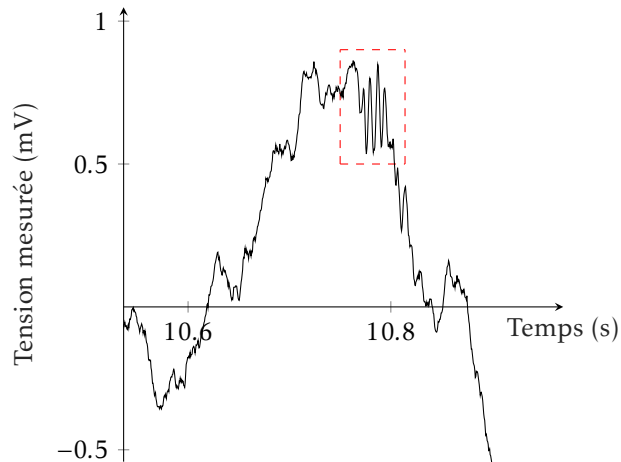


FIGURE 1 – HFO (de type *ripple* ici) visible sur un électroencéphalogramme, encadrée en rouge, d’une durée d’une trentaine de millisecondes.

250Hz à 500Hz, dite des *fast ripples* (FR), par opposition aux *ripples* (R) qui sont des HFO de fréquence plus basse, comprise entre 80Hz et 250Hz [10].

L’intérêt (toujours étudié) des HFO comme miroirs de l’activité pathogène au sein des zones d’intérêt, leur localisation tenue au sein de ces zones, associé à la haute précision spatiale atteinte par les moyens modernes d’acquisition en font un sujet d’étude de choix pour la préparation des opérations de résection, consistant à retirer ou rendre inactives les SOZ, touchant ainsi à la source des crises épileptiques. Pouvoir identifier avec précision ces HFO parmi la masse de données obtenues pourrait donc permettre un ciblage plus précis des zones à opérer. On trouve également au delà de l’intérêt pré-chirurgical pour les patients dont la pathologie est résistante aux traitements médicamenteux un potentiel intérêt pour les autres patients, certains liens en effet observés entre médication et taux d’occurrence des HFO [8] suggèrent la possibilité d’utiliser la surveillance des HFO comme moyen de suivi de l’action des médicaments administrés.

La figure 1 montre une HFO observable directement, facilement repérable, sur un électroencéphalogramme. Cependant ce cas ne constitue qu’un cas de figure parmi les occurrences de HFO. En effet, l’oscillation en figure 1 survient isolément de toute activité autre que celle cérébrale de fond. Les HFO peuvent pourtant se produire conjointement à d’autres phénomènes, et on peut même rencontrer plusieurs HFO distinctes survenant simultanément, dans des bandes de fréquence différentes. La figure 2 montre plusieurs de ces cas, tant de superposition avec des pics d’activité physiologique que de plusieurs HFO distinctes (une de type *ripple*, l’autre *fast-ripple*).

On s’aperçoit vite qu’un simple examen des séries temporelles brutes ne permet pas facilement d’identifier ce qui est à l’œuvre dans chacune des situations. Pourtant ces cas sont relativement fréquents : l’article [13], étudiant ces phénomènes d’occurrences simultanées montre que 64% des *ripples* et 86% des *fast-ripples* se produisent en même temps que des pics. Ces problèmes d’observation et d’identification mal aisée sur les séries temporelles brutes justifient le besoin de s’intéresser à des méthodes pour traiter ce signal et caractériser plus précisément ces phénomènes.

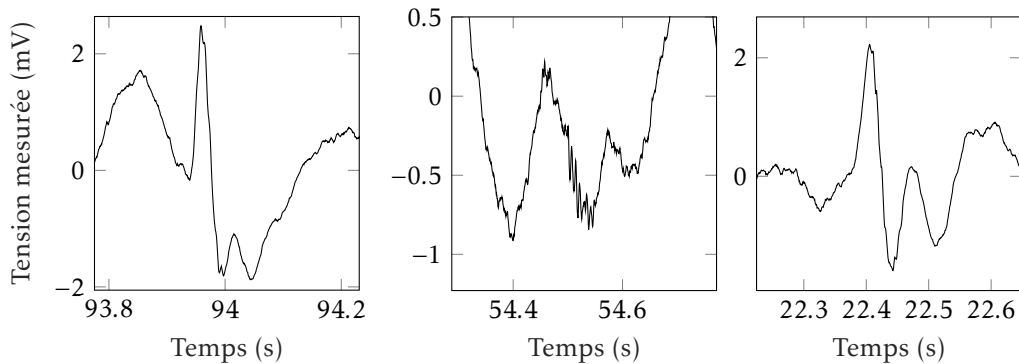


FIGURE 2 – Séries temporelles associées à des HFO marqués. À gauche, un *ripple* se produisant simultanément avec un pic, au centre un *ripple* et un *fast-ripple* simultanés, et à droite un pic, un *ripple* et un *fast-ripple*. Ces courbes sont issues de données simulées provenant de [12]), le marquage des divers évènements est alors fiable.

La première étape dans la plupart des méthodes automatiques de détection des HFO déjà proposées consiste usuellement en une phase de filtrage de la série temporelle afin d’isoler les composantes dans les bandes de fréquence qui nous intéressent. Les évènements de type HFO pourront alors apparaître et ressortir comme des évènements d’énergie élevée dans les bandes de fréquence leur correspondant (80-250Hz pour les *ripples*, 250-500Hz pour les *fast-ripples*). Seulement, à cet égard, la présence dans le signal de pics pose problème : ceux-ci étant des évènements « étroits » (*sharp*), en raison de phénomènes dits de Gibbs ils se déploient dans une large bande de fréquence. C’est encore plus vrai pour des évènements de type *artefacts*, non physiologiques mais résultant des aléas de l’acquisition, plus localisés encore. En particulier une part de l’énergie du signal dans les bandes d’intérêt peut alors être expliquée par des pics et des artefacts. Un simple filtrage et une sélection des évènements sur le critère de l’énergie contenue dans la bande des HFO est alors susceptible d’engendrer des faux-positifs. La figure 3 illustre ce phénomène avec l’effet d’un filtrage sur un artefact et un HFO.

Compte tenu de la relative difficulté de l’identification des HFO, accompagnée de la subjectivité sous-jacente à l’évaluation manuelle des signaux (icelle évaluation souffrant d’une faible fiabilité inter-évaluateur [12]), il est encore difficile d’évaluer et comparer les différents détecteurs mis au point. Pour progresser sur cet aspect, l’article [12] décrit un cadre permettant de simuler des électroencéphalogrammes avec leur lot de HFO, et pics, dans nombre de situations de simultanéité, en contrôlant le rapport signal / bruit (SNR), et les particularités propres aux différentes zones du cerveau desquelles on tire des mesures. Une archive de données simulées et marquées est fournie par les auteur-e-s : ce sont les données que j’ai utilisées lors du stage. Toutes les figures de ce rapport impliquant des électroencéphalogrammes sont d’ailleurs issues du traitement de ces données.

On va par la suite s’intéresser à des approches temps-fréquence, techniques de traitement du signal particulièrement adaptées à l’étude des évènements transitoires.

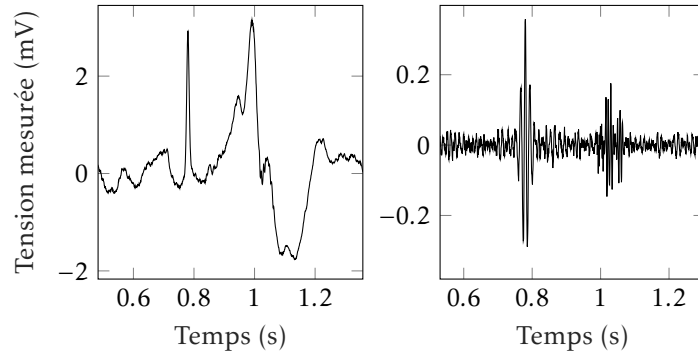


FIGURE 3 – Insuffisance du filtrage, phénomène de Gibbs à ses œuvres... À gauche un artefact (à $t \simeq 0,8s$) suivi de deux HFO de type *ripple* et *fast-ripples* simultanés (pour $t \simeq 1s$), un pic en plus. À droite la filtrée, dans la bande des *ripples* (ie. 80Hz-250Hz), qui fait ressortir de l'énergie tant pour l'artefact que pour la véritable oscillation.

3 Temps-Fréquence [5]

Une approche naturelle et historique pour l'analyse spectrale des signaux est l'analyse au travers de la *transformée de Fourier*. L'idée sous-jacente de cette approche est de projeter, un signal, sur des atomes de la forme $t \mapsto e^{2i\pi\nu t}$, pour $\nu \in \mathbf{R}$ jouant un rôle de fréquence, et de tirer de ces projections des informations sur la structure harmonique du signal étudié.

Ainsi, la valeur fréquentielle à la fréquence ν d'un signal réel s au sens de la transformée de Fourier traditionnelle est :

$$\hat{s}(\nu) = (s \mid t \mapsto e^{2i\pi\nu t}) = \int_{-\infty}^{+\infty} s(t)e^{-2i\pi\nu t} dt$$

Il apparaît alors que la connaissance de cette valeur fréquentielle nécessite la connaissance de l'histoire complète du signal, d'un point de vue temporel (en témoignent les bornes d'intégration). De plus, si les atomes considérés sont des atomes parfaitement localisés en fréquence (puisque correspondant à une onde monochromatique de fréquence ν donnée), il ne le sont absolument pas en temps : leurs oscillations sont infinies. Ces deux aspects font sentir que l'analyse de Fourier n'est pas particulièrement bien adaptée à l'étude de signaux transitoires, l'interprétation des spectres obtenus étant rendue difficile par l'absence de réalité physique derrière les valeurs des projections... Ainsi, par exemple, une nullité temporaire du signal s se traduira par une annulation locale *via* des phénomènes d'interférence entre les atomes sur lesquels on projette, et non par des nullités d'amplitude dans les projections. On ne peut se contenter de s'intéresser à l'amplitude de ces projections : leur phase est l'information complémentaire qui permet de caractériser et comprendre complètement le signal. Mais toutes complémentaires qu'elles soient, ces deux informations peuvent se révéler délicates à interpréter simultanément...

De la même manière, on peut penser l'échantillonnage temporel d'un signal comme l'approche duale de la transformation de Fourier : on projette cette fois-ci sur des atomes

de Dirac (l'échantillonnage étant l'opération de projection sur un maillage en *peigne de Dirac*), qui, bien que parfaitement localisés en temps, ne le sont absolument pas en fréquence (le spectre d'un pic de Dirac occupant l'ensemble des fréquences réelles...).

On peut alors essayer d'aller au delà du cadre offert par la transformation de Fourier, et considérer conjointement les aspects temps et fréquence. Pour ce faire, on va introduire une localisation temporelle sur les atomes monochromatiques considérés jusque là, en les multipliant par une fenêtre temporelle w centrée sur un temps τ .

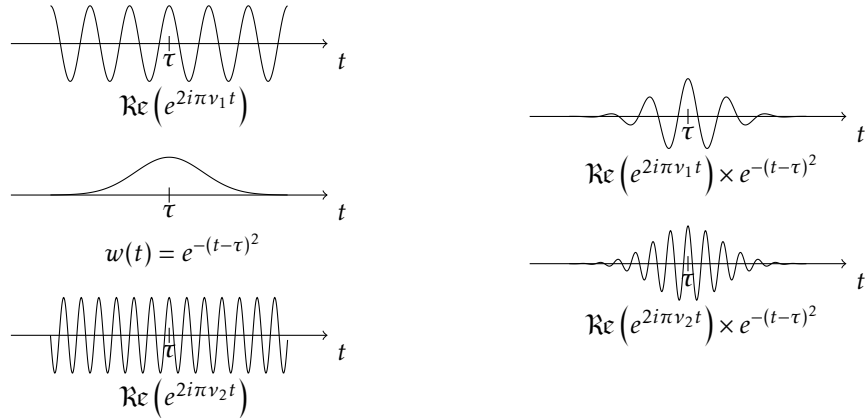


FIGURE 4 – Construction de deux atomes, à droite, localisés temporellement autour de τ , fréquentiellement respectivement autour de ν_1 et ν_2 ; ici en utilisant une fenêtre w gaussienne.

On peut visualiser ces atomes, leurs localisations tant en fréquence qu'en temps, en les représentant sur un plan *temps-fréquence*. Sur la figure 5, un tel plan, et la surface couverte sur celui-ci par les atomes construit plus tôt en figure 4.

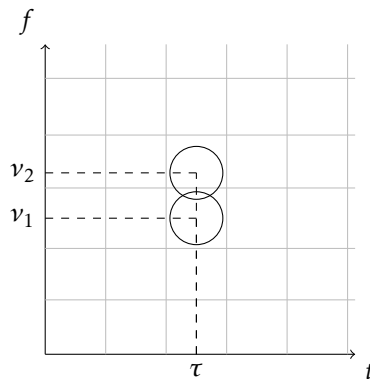


FIGURE 5 – Plan temps-fréquence, et localisation de deux atomes construit en figure 4.

On obtient là le formalisme de la *transformée de Fourier à court terme*. Les projections que l'on calcule alors sont les suivantes :

$$S(\tau, \nu) = \left(s \mid t \mapsto w(t - \tau)e^{2i\pi\nu t} \right) = \int_{-\infty}^{+\infty} s(t)\overline{w(t - \tau)}e^{-2i\pi\nu t} dt$$

On constate que si on s'approche là d'une solution convenable au problème motivant le dépassement de la décomposition de Fourier, c'est à dire ayant obtenu un atome localisé à la fois autour d'une fréquence donnée et d'un temps donné, cette localisation n'est pas « parfaite », en ce sens qu'un étalement autour de ses fréquence et temps centraux subsiste (sur la figure 5 : une surface non nulle occupée par l'atome). On se heurte là à des limites théoriques, données notamment par l'inégalité d'Heisenberg-Gabor, qui fournit une borne inférieure strictement positive à la surface occupée dans le plan temps-fréquence par tout atome. Borne qui est atteinte pour des atomes construits comme ci-haut avec un fenêtre gaussienne, appelés *atomes de Gabor*.

Plus formellement, en supposant notre atome a d'énergie E_a finie, et centré temporellement comme fréquentiellement (quitte à se ramener à ce dernier cas), on peut définir des étalements respectivement temporel et fréquentiel comme suit :

$$\Delta t = \sqrt{\frac{1}{E_a} \int_{-\infty}^{+\infty} t^2 |a(t)|^2 dt} \quad \Delta \nu = \sqrt{\frac{1}{E_a} \int_{-\infty}^{+\infty} \nu^2 |\hat{a}(\nu)|^2 d\nu}$$

Il s'agit là sous les racines des expressions de moments d'ordre 2, les grandeurs Δt et $\Delta \nu$ s'assimilant à l'écart type des distributions temporelles et fréquentielles, et traduisent alors un certain sens d'« étalement », dans ces domaines respectifs. On peut alors montrer la relation suivante, inégalité d'Heisenberg-Gabor :

$$\Delta t \Delta \nu \geq \frac{1}{4\pi}$$

On fournit une preuve dans l'annexe A. Ce résultat scelle les ambitions de localisation arbitrairement simultanée des atomes en fréquence et temps, mais ouvre en un même mouvement le champ des compromis... Ainsi, par exemple, les fenêtres gaussiennes évoquée plus tôt peuvent par parfois être avantageusement remplacées par d'autres fenêtres (de Hamming, de Blackman, rectangulaire, etc) selon les spécificités du problème rencontré, chaque choix de fenêtre permettant de réaliser différent compromis sur les résolutions fréquentielles, temporelles ou sur la manière dont le signal est affecté par la transformation.

Une fois la transformée de Fourier à court terme calculée (ayant choisi un schéma de discrétisation prenant usuellement la forme d'un maillage rectangulaire), on représente généralement la surface obtenue en prenant le carré du module de la transformée de Fourier à court terme et on obtient alors le *spectrogramme* du signal. On représente en figure 6 le spectrogramme obtenu à partir d'un *chirp* linéaire, signal qui peut se penser comme une oscillation dont la fréquence augmente linéairement au cours du temps.

Dans un premier temps j'ai implémenté un calcul de transformée à Fourier à court terme en Python, pour être sûr de bien saisir les diverses questions de discrétisation et calcul en jeu, ceci me permettant d'ailleurs de me rendre compte pour de bon de l'intérêt pratique de l'algorithme de transformation rapide de Fourier (on peut voir le calcul effectif de transformée de Fourier à court terme comme l'application successive de transformations rapide de Fourier au signal que l'on fenêtré autour de chaque temps), avant de me rabattre pour la suite du stage sur la boîte à outils TFTB développée pour

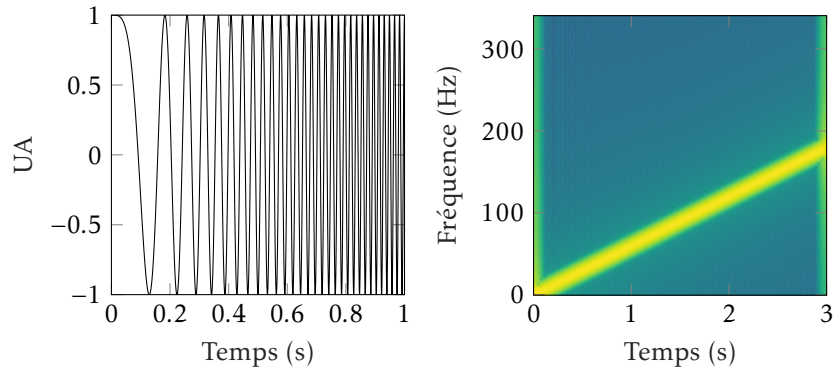


FIGURE 6 – Représentation temporelle d’un *chirp* linéaire (ici $\cos(2\pi f_0 t^2)$, pour $f_0 = 10\text{Hz}$) à gauche. À droite son spectrogramme. Avec l’échelle de couleurs utilisée ici pour la représentation de la surface, les parties jaune clair représentent les zones d’énergie élevée, et les zones virant vers le bleu plus foncé représentent celles d’énergie plus faible. Le spectrogramme permet alors de donner un sens à cette notion de « fréquence évoluant au cours du temps », et laisse apparaître la relation linéaire entre le temps et la fréquence du signal.

Matlab et Octave [7], permettant de travailler avec ces approches temps-fréquence de manière bien plus souple et d’utiliser les fonctionnalités de Matlab / Octave adaptées à ces questions de traitement de signal.

4 Détection de HFO avec des outils temps-fréquence

De nombreuses approches, variées, ont pu être proposées pour l’étude des HFO, on peut en trouver un large panorama dans [9]. On se concentre ici sur les stratégies usant de l’approche temps-fréquence.

Comme détaillé plus tôt, les HFO que l’on cherche à détecter sont caractérisées comme étant des phénomènes dont le déploiement est à la fois limité en temps et en fréquence. Leur signature dans le plan temps-fréquence sera porteuse de cette information structurale : les évènements apparaîtront sous la forme d’îlots isolés. La figure 7 illustre ceci, mettant côte à côte les signatures de HFO et la signature d’un artefact. Cette description est déjà utilisée pour faciliter la détection visuelle des HFO, et le rejet des faux HFO, par exemple dans [4].

Bien que l’on ait plus haut décrit le formalisme de la transformée de Fourier à court terme, ce n’est pas la seule stratégie temps-fréquence que l’on peut envisager, la majorité des articles étudiés optant d’ailleurs pour une approche en transformation d’ondelette ; si la théorie régissant cette approche est différente de celle liée à Fourier à court terme, les idées restent les mêmes, en particulier la lecture des HFO comme îlots sur la carte temps-fréquence associée.

Dans la suite, on s’intéresse à des formalisations de cette remarque sur la structure des HFO et leur signature dans le plan temps-fréquence pour la mise au point de détecteurs automatiques.

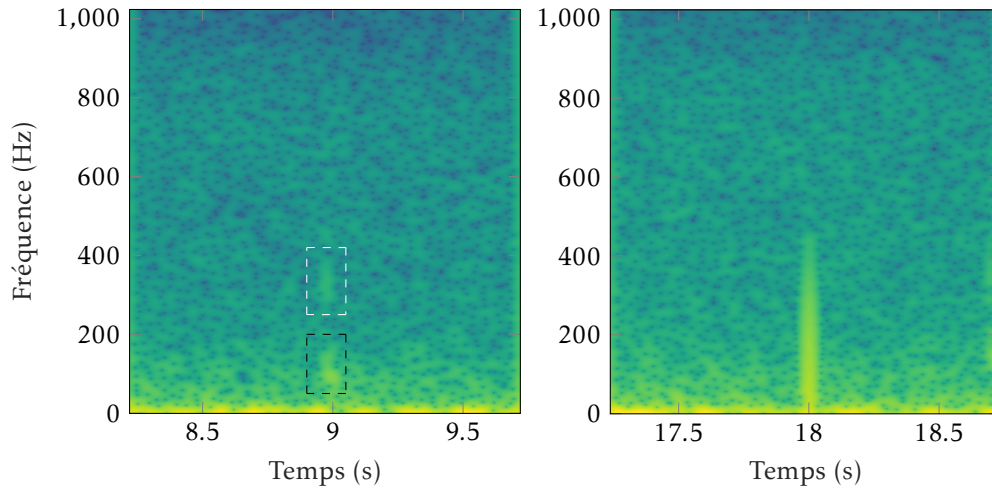


FIGURE 7 – Spectrogrammes. À gauche de deux HFO, de type *ripple* (signature encadrée en noir) et *fast-ripple* (signature encadrée en blanc) simultanées, à droite d'un artefact, qui se déploie dans la bande de fréquence des HFO.

4.1 Approches énergétiques

Une approche assez naturelle vis à vis de la remarque précédente, pour discriminer HFO et pics / artefacts est de s'intéresser à cette notion, jusqu'alors juste intuitive, d'îlot visible sur la carte temps-fréquence du point de vue de l'énergie du signal révélée par le spectrogramme. En effet, un îlot qui ressort à l'œil est une zone localisée du plan temps-fréquence où l'énergie est plus élevée qu'en tous ses abords. Ainsi, une fois repéré un évènement d'intérêt, on peut s'intéresser à son profil d'énergie spectrale instantanée. Une signature correspondant à une HFO serait alors de la forme suivante : un maximum local d'énergie dans la bande de fréquence des HFO, entourée de minima locaux suffisamment « profonds » pour correspondre à une signature de type îlot dans le plan temps fréquence. C'est cette approche qu'adopte l'article [2], mettant en place la méthode suivante.

On commence (comme une majorité de méthodes) par filtrer le signal dans la bande qui nous intéresse, celle où se déploient les HFO. On calcule ensuite l'enveloppe (de Hilbert) du signal filtré. Sont alors dits d'intérêt les instants où cette enveloppe dépasse un seuil, défini là comme 3σ où σ est l'écart type de l'amplitude de l'enveloppe. On fusionne les instants proches détectés temporellement ainsi (qui correspondent au même évènement). Cette première étape, commune à bien des détecteurs (à la variation seule du seuil choisi), isole des évènements intéressants, mais ne permet pas seule d'assurer une bonne *spécificité*. On trouve *a priori* des faux-positifs dus à des phénomènes de Gibbs décrits plus haut. On passe alors à l'analyse temps-fréquence.

Pour chaque évènement d'intérêt détecté à la première étape, on va sélectionner trois fréquences : après avoir calculé le spectre instantané de l'évènement (dans l'article *via* une transformée de Stockwell, méthode temps-fréquence hybride entre la transformée de Fourier à court terme et la transformée en ondelette), on localise la fréquence

f_{max} d'énergie maximum dans la bande 60Hz-500Hz, la fréquence f_{min} d'énergie minimum entre 40Hz et f_{max} , et enfin f_{pic} , première fréquence d'énergie localement maximum et inférieure à f_{min} .

Pour chaque temps étudié, pour être marqué comme HFO, l'article exige d'avoir $P(f_{min})/P(f_{max}) < 0,8$ et $P(f_{max})/P(f_{pic}) > 0,5$ pour chaque temps considéré (avec $P(f)$ l'énergie à la fréquence f dans le spectre instantané de l'évènement). C'est à dire qu'on veut ici respectivement un « trou » en f_{min} suffisamment profond, et un évènement en f_{max} contenant une énergie significative par rapport à l'activité basse fréquence du signal. Ces seuils ont été déterminés en utilisant des données de référence, pour maximiser les performances du détecteur.

La méthode est dans l'article testée avec pour *gold-standard* (c'est à dire critère pour décider de la validité ou non des résultats) le repérage des *Seizure Onset Zones* (SOZ). On utilise pour chaque patient une cartographie de référence de ces zones, obtenue par l'étude de crises antérieures. Les HFO validés par l'algorithme précédent permettent de faire ressortir les canaux d'acquisition (associés chacun à une zone du cerveau) où l'activité des HFO est particulièrement élevée, et on considèrera que l'algorithme a produit un vrai positif si un des canaux plus actif correspond à une SOZ déjà connue, et un faux positif sinon. Si ce choix est pertinent d'un point de vue de l'intérêt médical des HFO (aider à étudier ces SOZ), il ne fournit pas d'information très précise sur la capacité d'une telle méthode à catégoriser individuellement les évènements. Cela, et le faible jeu de données utilisées (un seul patient pour l'entraînement et détermination des seuils, quatre patients pour tester la méthode) rendent difficile une véritable évaluation de la pertinence.

De plus les auteurs de [2] n'abordent pas de distinction entre *ripples* et *fast-ripples*, là où celle-ci semble pourtant pouvoir être judicieuse; et plus généralement peu d'information est extraite de chaque évènement. La méthode repose également sur l'énergie contenue dans le spectrogramme, avec un certain *a priori* sur sa répartition lors de l'observation de HFO.

Des études ont d'ailleurs été réalisées s'intéressant à la répartition de cette énergie dans le plan temps-fréquence et l'influence de celle-ci sur la détection de HFO : l'article [11] étudie divers procédés de normalisation de l'énergie dans les spectrogrammes, dans le but premier de faciliter la détection visuelle des HFO, mais propose d'utiliser ces procédés de normalisation pour améliorer les résultats des détecteurs. En particulier ils suggèrent qu'il pourrait être envisageable (et même souhaitable) de s'affranchir de la première phase de détection consistant en un filtrage et un seuillage sur l'enveloppe du signal filtré, pour la remplacer par un seuillage sur l'énergie d'un spectrogramme normalisé, détecter ainsi des évènements d'intérêt et ensuite extraire des informations sur l'évènement en regardant la répartition de l'énergie autour de l'évènement, en particulier les étalements spectraux et temporels, qui permettent de formaliser l'idée « d'ilot ».

4.2 Approche géométrique

On s'intéresse dans cette partie à une approche géométrique, originale dans le domaine de la détection des HFO, dont les idées ont été introduites dans [6]. L'idée est de s'intéresser au signal vu comme une perturbation de son bruit de fond, dont la signature temps-fréquence comporte, sous certaines hypothèses, quelque régularité géométrique.

En effet, se plaçant dans le cadre de la transformée de Fourier à court terme fenêtrée

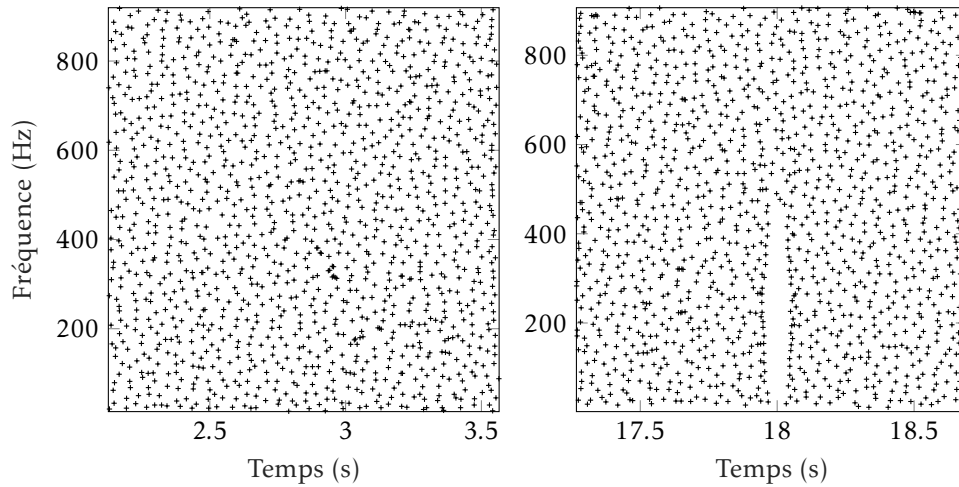


FIGURE 8 – Champ des zéros dans le spectrogramme d’un électroencéphalogramme, à gauche sans évènements d’intérêt observé, à droite avec la présence de l’artefact visible en figure 7.

gaussienne — soit le cadre introduit dans la section 3 — et s’intéressant aux propriétés du spectrogramme d’un bruit blanc gaussien, on peut montrer ([6]) que ce dernier se comporte tel un champ homogène d’ordre deux. Cette vertu d’homogénéité du spectrogramme se transmet à ses points particuliers, en particuliers à ses zéros, dans leur répartition. Dans ce même cadre, on peut également montrer que le spectrogramme d’un signal est entièrement caractérisé par la distribution de ses zéros : il en existe une expression faisant intervenir uniquement ces derniers.

De ces deux résultats, l’article [6] déduit un procédé de débruitage sous hypothèse de bruit blanc gaussien, sans connaissance *a priori* des caractéristiques des signaux. De fait, si les zéros du spectrogramme caractérisent entièrement ce dernier, il est raisonnable de s’attendre à ce que leur distribution fasse ressortir les propriétés du signal d’intérêt. C’est effectivement ce que l’on observe, vis à vis de l’interaction entre le champ des zéros induit par le bruit, et l’énergie du signal d’intérêt : celle-ci va alors « pousser » les zéros autour d’elle, perturber l’homogénéité du champ de fond. C’est ce que l’on peut observer sur la figure 8, comparant la distribution des zéros sur le spectrogramme d’un épisode sans évènement d’intérêt et d’un épisode avec.

On va chercher alors à formaliser cette notion de répartition « anormale » des zéros qui émerge en présence de signal dans du bruit. Étant établi que l’on peut se contenter de la répartition des zéros pour décrire le spectrogramme, il est assez naturel d’en chercher une tessellation et d’obtenir alors une représentation purement géométrique du spectrogramme : on choisit dans le procédé d’en calculer une triangulation, de Delaunay, qui fournit alors une description plus clairsemée du spectrogramme.

Du fait de l’homogénéité du champ des zéros hors du signal, et du signal « poussant » ces dits zéros autour de lui, des triangles vont ressortir de la triangulation : ceux ayant au moins un de leurs côtés particulièrement long comparé au reste de la tessellation. Il est raisonnable de penser alors ces triangle comme issus de perturbations par le

ment du signal et de manipulation de séries temporelles, et leur mise en œuvre dans un script Matlab / Octave est alors aisée, la plupart s’implémentant en quelques lignes. Le calcul du spectrogramme, et en particulier la configuration de sa fenêtre (pour respecter le cadre théorique assurant des résultats décrits plus tôt) fait partie intégrante de ce que permet la bibliothèque TFTB[7] évoquée dans la section 3. Je me suis alors penché sur le traitement du spectrogramme pour réaliser les opérations qui suivent.

Pour des raisons de performance, j’ai choisi d’interfacer le script Matlab / Octave avec du code écrit en C++. En effet, Matlab et Octave sont particulièrement adaptés à la production de code *vectorisé* (c’est à dire vu comme applications d’opérateurs et fonctions à des tableaux et matrices), les performances se dégradant rapidement lorsque l’on adopte un style plus impératif, avec boucles, parcours de tableau, etc, style qu’il me paraissait délicat d’éviter pour la suite des opérations.

De par les discrétisations en jeu, et questions d’arrondis, la recherche de zéros dans la surface qu’est le spectrogramme peut s’avérer hasardeuse si elle s’établit simplement à partir de seuils numériques. On peut montrer que celle-ci se ramène en fait à la recherche de minimum locaux, ceci découlant de propriétés de régularité du spectrogramme. Il s’agit alors simplement d’un parcours de la surface à la recherche de points (de pixels dans notre cas discret) d’énergie plus faible qu’en leur voisinage. La complexité est moindre que celle du calcul du spectrogramme.

Matlab comme Octave possèdent une implémentation du calcul d’une triangulation de Delaunay d’un nuage de point (qui se résume pour tous deux en un appel à la bibliothèque qhull) : il suffit de l’appliquer aux zéros décelés précédemment pour en obtenir la tessellation. Ici les performances ne sont pas critiques, le nombre de zéro est usuellement assez faible avec les fenêtres temporelle et fréquentielle utilisées (de l’ordre du millier), et les algorithmes connus et déjà implémentés atteignant la borne de $O(n \log n)$ pour n le nombre de points. Le calcul est alors peu coûteux.

On peut alors conserver les triangles à l’un des côtés anormalement long (parmi les 1% les plus longs), opération que Matlab / Octave permet bien de faire sans trop de heurts. C’est à ce stade que l’on peut par exemple produire la figure 9.

Construire le graphe d’adjacence de tels triangles et en calculer les composantes connexes permet de grouper les triangles anormalement grands ayant un côté mitoyen pour former des composantes de triangles. Là encore cette opération est plus intéressante à réaliser en C++, non trop par soucis de performance (à ce moment là, il ne reste qu’assez peu de données à traiter, la plupart des triangles ayant été éliminés), mais plus par praticité, l’accès à des structures de données telles que les piles rendant ces opérations bien plus agréables.

Le résultat produit est un code qui analyse le signal et fait défiler les évènements d’intérêt avec les composantes capturées par la méthode ci-dessus; quelques exemples de sorties sont visibles sur la figure 10.

Cette approche ne prétend pas encore discriminer les vraies HFO des fausses. À son terme on obtient en effet seulement une description géométrique de chaque évènement candidat, sous la forme de composantes de triangles dans le plan temps-fréquence. Description qui permet d’accéder très simplement à des caractéristiques qui semblent, au vu de la littérature, pertinentes pour l’étude des HFO : notamment les étalement temporels et fréquentiels de chacune des composantes (il suffit de regarder les coordonnées extrémales parmi les sommets des triangles formant la composante). C’est ici une différence avec les méthodes précédemment élaborées : ces dernières extraient ces informations

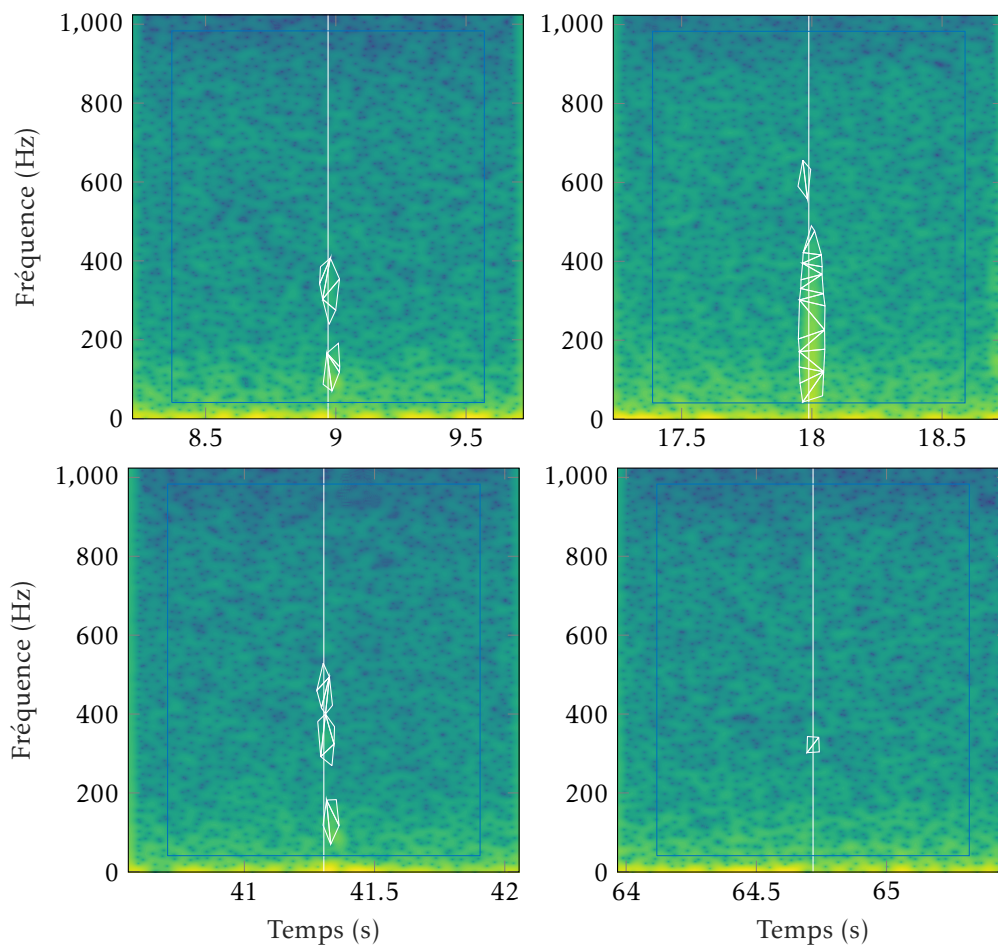


FIGURE 10 – Quelques résultats de l’algorithme de détection et d’extraction de composantes, appliqué sur la bande des *fast-ripples*. Sur la ligne du haut on retrouve les deux évènements de la figure 7. En bas : à gauche une autre occurrence simultanée de *ripple* et *fast-ripple*. À droite un faux-positif issu de la première phase : bien qu’il n’y ait pas de HFO à cet instant, on trouve de l’énergie dans la bande des *fast-ripples*, suffisamment pour passer le seuil de $4\sigma_h$ choisi, mais la seconde étape capture seulement peu de structure derrière cet évènement. Le trait blanc vertical représente le temps détecté lors de la première étape sur l’enveloppe de la filtrée, le cadre léger bleu symbolise la fenêtre dans laquelle on s’intéresse aux triangles, pour éviter des effets de bord.

en s'appuyant sur l'énergie du signal dans le spectrogramme, éventuellement après divers traitements ([11]), utilisant par exemple les largeurs (temporelles et fréquentielles) à mi-hauteur autour du maximum d'énergie de l'évènement. Notre approche est plus structurelle. On se concentre sur les zéros du spectrogramme, soit des points singuliers du spectrogramme, plutôt que des valeurs d'énergie particulières autour d'un maximum local, par exemple.

Notre méthode permet également une extraction différenciée des composantes associées par exemple à deux HFO de type *ripple* et *fast-ripple* simultanées.

On peut envisager utiliser notre méthode d'extraction pour analyser des séries de HFO déjà marquées. On peut par exemple s'intéresser, de manière plus géométrique que le seul déploiement temporel ou fréquentiel, à l'aire occupée par les composantes, également accessible aisément de par la nature de la description (il suffit de sommer l'aire des triangles contenus dans la composante). Ainsi, sur une série de l'ordre de 200 HFO marqués, on peut observer un certain confinement de ces aires, comme l'indique l'histogramme de répartition de celles-ci (figure 11). Ce critère permettrait sur nos exemples de discriminer, en rejetant les composantes d'aire trop petite ou grande par rapport à ce qu'on peut observer sur les HFO de référence, entre évènements de type artefact et HFO (aire trop importante pour les artefacts), ainsi qu'entre HFO et évènements consistant en de l'énergie contenue « par hasard » dans les bandes de fréquence des HFO, mais ne présentant pas suffisamment de structure pour être considérée comme d'intérêt, le critère ici étant une aire trop petite de la composante (c'est par exemple l'exemple en bas à droite de la figure 10).

On peut imaginer s'intéresser à d'autres caractéristiques : nombre de triangles par composantes, aire moyenne par triangle, nombre de composante (certains *fast-ripples* présenteraient plusieurs composantes, comme en figure 10, en bas à gauche, on peut se demander s'il y a une pertinence particulière à y trouver).

Malheureusement, six semaines passant vite, je n'ai pas eu le temps d'aller plus loin que le port des idées à l'état de « preuve de faisabilité ». Il faudrait soumettre ceci à des tests plus systématiques, en variant la provenance des signaux (l'archive fournie par [12] permet de parcourir quelques zones diverses du cerveau), s'intéresser au comportement de la méthode à SNR plus faible ou plus fort, les images et essais ayant été produits et conduits avec des signaux d'un SNR de l'ordre de 10dB. De plus, il me semble que le côté simulé des données que j'ai utilisées conduit à manipuler des signaux peut-être « trop propres » : parmi mes explorations visuelles, certes à l'œil non exercé, je n'ai pas trouvé d'évènement qui rentrerait dans la case *artefact*. Ceux visibles sur les figures plus hauts et lors de mes essais ont été ajoutés de mon fait, et bien qu'inspiré par l'article [1] pour rester dans un cadre réaliste, il serait sans doute bon de confronter la méthode à de vrais aléas d'acquisition.

Il serait peut-être aussi intéressant, suivant la direction de [11], de voir s'il est possible d'adapter la méthode pour se passer de la première étape et se contenter d'une extraction à partir du spectrogramme brut.

5 Conclusion

Ainsi, on a introduit une approche nouvelle d'extraction de caractéristiques d'évènements, en se concentrant sur la structure géométrique du signal dans le plan temps-

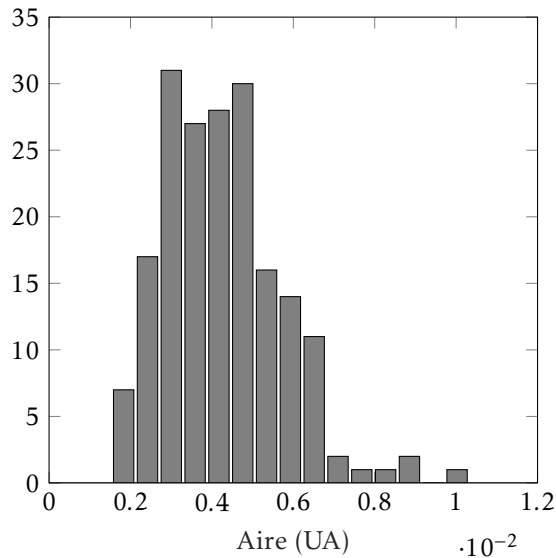


FIGURE 11 – Histogramme des aires des composantes détectées par notre méthode pour des HFO marquées, de type *fast-ripple*, à un SNR de 10dB

fréquence, ce qui constitue une approche originale par rapport aux détecteurs utilisant des seuils sur l'énergie des spectrogrammes. L'étude conduite ici n'a cependant pas été assez approfondie pour permettre de mettre au point un véritable détecteur, bien autonome et qui ait été testé, avec des performances quantifiées ou au moins comparées à d'autres détecteurs. Reste que la description géométrique de chaque évènement, à la fois simple à manipuler et intuitive, fournie par notre approche la rend sûrement propice à l'exhibition des critères qui pourraient constituer la base d'un détecteur intéressant. Dans tous les cas, sans aller jusqu'à un détecteur complet et autonome, on pourrait aussi chercher à utiliser notre approche en l'adjoignant à une méthode déjà existante pour en augmenter la robustesse.

Dans une perspective plus large, il serait également intéressant de reprendre les idées lancées dans [6], évoquées dans la section 4.2, déjà appliquées sur les données à ma disposition à la recherche de HFO, et de proposer par exemple une bibliothèque applicable, plus souplement qu'en son état actuel, à d'autres contextes que la détection de HFO, implémentant ces approches géométriques.

6 Références

- [1] Christian-G Bénar, Laetitia Chauvière, Fabrice Bartolomei, and Fabrice Wendling. Pitfalls of high-pass filtering for detecting epileptic oscillations : a technical note on “false” ripples. *Clinical Neurophysiology*, 121(3) :301–310, 2010.
- [2] Sergey Burnos, Peter Hilfiker, Oguzkan Sürücü, Felix Scholkmann, Niklaus Kraysenbühl, Thomas Grunwald, and Johannes Sarnthein. Human intracranial high frequency oscillations (hfos) detected by automatic time-frequency analysis. *PLoS One*, 9(4) :e94381, 2014.
- [3] Sahbi Chaibi, Tarek Lajnef, Zied Sakka, Mounir Samet, and Abdennaceur Kachouri. A reliable approach to distinguish between transient with and without hfos using tqwt and mca. *Journal of neuroscience methods*, 232 :36–46, 2014.
- [4] Jounghong Ryan Cho, Eun Yeon Joo, Dae Lim Koo, Seung Chyul Hong, and Seung Bong Hong. Clinical utility of interictal high-frequency oscillations recorded with subdural macroelectrodes in partial epilepsy. *Journal of Clinical Neurology*, 8(1) :22–34, 2012.
- [5] Patrick Flandrin. *Temps-fréquence*. Number BOOK. Hermes, 1993.
- [6] Patrick Flandrin. Time-frequency filtering based on spectrogram zeros. *IEEE Signal Processing Letters*, 22(11) :2137–2141, 2015.
- [7] Auger François, Lemoine Olivier, Gonçalves Paulo, and Flandrin Patrick. Time-frequency toolbox (tftb). <http://tftb.nongnu.org/>. [En ligne, consulté le 7 août 2018].
- [8] Birgit Frauscher, Fabrice Bartolomei, Katsuhiko Kobayashi, Jan Cimbalnik, Maryse A van ‘t Klooster, Stefan Rampp, Hiroshi Otsubo, Yvonne Höller, Joyce Y Wu, Eishi Asano, et al. High-frequency oscillations : The state of clinical research. *Epilepsia*, 58(8) :1316–1329, 2017.
- [9] Miguel Navarrete, Jan Pyrzowski, Juliana Corlier, Mario Valderrama, and Michel Le Van Quyen. Automated detection of high-frequency oscillations in electrophysiological signals : Methodological advances. *Journal of Physiology-Paris*, 110(4) :316–326, 2016.
- [10] Martin Pail, Pavel Řehulka, Jan Cimbalník, Irena Doležalová, Jan Chrastina, and Milan Brázdil. Frequency-independent characteristics of high-frequency oscillations in epileptic and non-epileptic regions. *Clinical Neurophysiology*, 128(1) :106 – 114, 2017.
- [11] 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 eeg. *IEEE Transactions on Biomedical Engineering*, 63(12) :2595–2606, 2016.
- [12] 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, 2017.
- [13] Elena Urrestarazu, Rahul Chander, François Dubeau, and Jean Gotman. Interictal high-frequency oscillations (100–500 hz) in the intracerebral eeg of epileptic patients. *Brain*, 130(9) :2354–2366, 2007.

A Inégalité d'Heisenberg-Gabor

L'énergie E_s supposée finie du signal s s'écrit, par théorème de Parseval,

$$E_s = \int_{-\infty}^{+\infty} |s(t)|^2 dt = \int_{-\infty}^{+\infty} |\hat{s}(\nu)|^2 d\nu$$

Or, en notant I la quantité

$$I = \int_{-\infty}^{+\infty} t \overline{s(t)} s'(t) dt$$

On a la suite d'inégalité suivante :

$$\begin{aligned} |I|^2 &\leq \int_{-\infty}^{+\infty} t^2 |s(t)|^2 dt \times \int_{-\infty}^{+\infty} |s'(t)|^2 dt && \text{(Cauchy-Schwarz)} \\ &\leq E_s \Delta t^2 \times \int_{-\infty}^{+\infty} |\widehat{s'}(\nu)|^2 d\nu && \text{(Définition de } \Delta t \text{ et identité de Parseval)} \\ &\leq E_s \Delta t^2 \times \int_{-\infty}^{+\infty} 4\pi^2 \nu^2 |\hat{s}(\nu)|^2 d\nu && \text{(Propriété de la transformée de Fourier d'une dérivée)} \\ &\leq 4\pi^2 E_s^2 \Delta t^2 \Delta \nu^2 && (1) \end{aligned}$$

De plus, en intégrant I par partie (en supposant que le signal est suffisamment évanescant à l'infini pour avoir $\lim_{t \rightarrow \pm\infty} t |s(t)|^2 = 0$) :

$$I = \int_{-\infty}^{+\infty} t \overline{s(t)} s'(t) dt = \underbrace{\left[t |s(t)|^2 \right]_{-\infty}^{+\infty}}_{=0} - \int_{-\infty}^{+\infty} |s(t)|^2 dt - \int_{-\infty}^{+\infty} t \overline{s'(t)} s(t) dt = -E_s - \bar{I}$$

D'où

$$\Re(I) = -\frac{E_s}{2} \quad (2)$$

On déduit des équations (1), (2) et du fait que $\Re(I)^2 \leq |I|^2$ l'inégalité suivante, qui est l'inégalité souhaitée :

$$\frac{1}{4\pi} \leq \Delta t \Delta \nu$$

De plus on peut retrouver le résultat sur la saturation de l'inégalité par les fonctions gaussiennes, évoqué en section 3 en s'intéressant au cas d'égalité dans l'inégalité de Cauchy-Schwarz : celui-ci est donné pour des fonctions proportionnelles, soit ici des solutions de l'équation différentielle suivante, solutions données par les fonctions gaussiennes :

$$ts(t) = ks'(t)$$

B Contexte institutionnel et social du stage

J'ai été pour ce stage accueilli par l'équipe DANTE, hébergée au sein de l'IXXI, institut rhônalpin des systèmes complexes, dont les laboratoires sont attenants à l'école. Je remercie encore Paulo Gonçalves de l'équipe DANTE et Patrick Flandrin de l'équipe Sisyphe de m'avoir proposé ce stage sur un sujet qui toujours a su susciter mon intérêt au long des — bien courtes finalement — six semaines qu'a duré ce stage, et d'avoir été disponible au long de celui-ci.

J'ai partagé un bureau avec deux doctorants de l'équipe, et puis mes repas avec bien d'autres de l'IXXI encore. J'y ai pu d'une conversation l'autre découvrir le large spectre des recherches menées à l'IXXI, à la croisée souvent entre informatique, physique, dont les applications portent plus loin parfois : de la sociologie aux télécommunications, passant médecine et biologie, profiter des parcours scolairement comme géographiquement divers des personnes y travaillant, me forcer à me saisir de mon anglais pour y communiquer et raviver quelques bouts d'espagnol. C'était également l'occasion de m'imprégner d'une vie de laboratoire, à suivre les odysées en conférence des uns ou en école d'été des autres comme face aux deadlines des derniers, venant ponctuer et rythmer ces jours à me plonger dans mon sujet... Assister aussi, auditeur juste, à des séminaires de l'équipe et des doctorants, tout cela participant de cette découverte des recherches s'établissant autour de moi.