Instituto Politécnico de Lisboa

Instituto Superior de Engenharia de Lisboa

Escola Superior de Tecnologia da Saúde de Lisboa

Development of an analysis pipeline for human microelectrode recordings in Parkinson's Disease

Sara Fernández Abalde

Thesis to obtain the Master Degree in

Biomedical Engineering

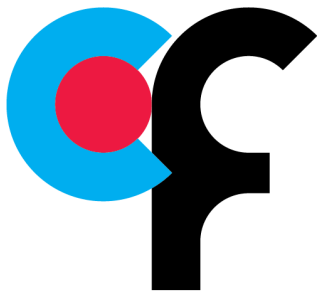Supervisors: Gonçalo Caetano Marques (ISEL)
Marcelo Mendonça (Fundação Champalimaud)

Examination Committee

Chairperson: João Costa, PhD
Members of the Committee: André Caetano, MD
João Sousa, PhD
Marcelo Mendonça, MD

December 2018

Champalimaud Foundation

Development of an analysis pipeline for human microelectrode recordings in Parkinson's Disease

Sara Fernández Abalde

Acknowledgments

Abstract

Background:

Deep brain stimulation is a common treatment for advanced Parkinson's Disease (PD). Intra-operative microelectrode recordings (MER) along preplanned trajectories are often used for accurate identification of subthalamic nucleus (STN), a common target for deep brain stimulation (DBS) in PD. However, this identification is performed manually and can be difficult in regions of transition. Misidentification may lead to suboptimal location of the DBS lead and inadequate clinical outcomes.

Methods:

A tool for unsupervised analysis and spike-sorting of human MER signals with feature extraction was developed. We also trained and tested a hybrid unsupervised/supervised machine learning approach that uses extracted MER time, frequency and noise properties for high-accuracy identification of STN. Lastly, we compared neurophysiological characteristics of different STN functional segments.

Results:

We obtained a classification accuracy of 96.28 ± 3.15 % (30 trajectories, 5 patients) for individual STN-DBS surgery MER using an approach of "leave one subject out" validation with support vector machine classifier, all features based on time and frequency domain and human expert labels.

The unsupervised sorting approach allowed us to sort a total of 357 STN neurons in 5 subjects. Dividing the STN in a dorsal, probably motor region, and a ventral, probably non-motor portion, we've found a higher burst rate (median (interquartile range) of 1.8 (1.5) vs 1.15 (0.05) bursts/s, $p=0.001$) and firing rate (median (interquartile range) of 21.4 (16.85) vs. 15.3 (14.33), $p=0.013$) of dorsal STN neurons among other features. Ongoing work will refine these results using anatomical gold standard through lead trajectory reconstruction, fused with an STN functional subdivision atlas.

Conclusions:

We've developed a tool for human MER analysis and extraction of related features, that

provided good preliminary results in STN classification. In line with the literature, we were able to find preliminary activity differences in functionally segregated STN segments. This tool is fast and generalizable for other brain regions. Ongoing work using patient's anatomy can further validate its' usefulness in optimizing electrode placement and research purposes.

Keywords

Microelectrode Recordings (MER); Deep Brain Stimulation (DBS); Parkinson's Disease (PD); Subthalamic Nucleus (STN); Support Vector Machines (SVM)

Resumo

A Estimulação Cerebral Profunda é um técnica utilizada para o tratamento dos sintomas motores da Doença de Parkinson com recurso a estimulação eléctrica intracraniana. Um dos alvos mais frequentemente utilizados para o tratamento é o núcleo subtalâmico e uma colocação precisa dos eléctrodos na região correcta é fulcral para o sucesso terapêutico e minimização de efeitos secundários.

Esta cirurgia faz-se com recurso a estereotaxia e monitorização intra-operatória. Previamente à cirurgia, as coordenadas do alvo são estabelecidas e é definido o núcleo subtalâmico utilizando imagens pré-operatórias de ressonância magnética e de tomografia computadorizada do doente. Intra-operatoriamente são utilizados registos de microeléctrodos a diferentes profundidades em trajetórias pré-definidas para verificar as coordenadas estabelecidas e identificar -pela actividade eléctrica- a localização dos eléctrodos no cérebro. No entanto, a identificação do núcleo subtalâmico com registos de microeléctrodos é frequentemente realizada por inspeção visual, e pode ser difícil em regiões de transição. Erros na identificação podem levar a um posicionamento subóptimo dos elétrodos e a um mau resultado clínico.

O correcto posicionamento do eléctrodo final de estimulação na porção sensoriomotora do núcleo subtalâmico está relacionado com os melhores resultados clínicos de acordo com a literatura. No entanto não é possível identificar claramente esta subdivisão com base a inspecção visual dos registos.

O objetivo do presente estudo é o desenvolvimento de ferramentas para análise de registos de microeléctrodos em diferentes áreas do cérebro e identificação do núcleo subtalâmico, e da porção sensoriomotora, em doentes com Doença de Parkinson submetidos a estimulação cerebral profunda.

Métodos

Foram desenvolvidas ferramentas para análise do sinal, *spike sorting* e extração de características por processamento não supervisionado de registos de microeléctrodos humanos.

As características relativas ao domínio do tempo e frequência foram obtidas para cada sinal através de medidas estatísticas (média, mediana e desvio padrão) dos registos fracionados em segmentos sobrepostos, para minimizar os efeitos do ruído. As características relacionadas com a

atividade neuronal foram também extraídas após deteção e identificação dos diferentes neurónios em cada registo de forma não supervisionada.

A localização dos diferentes registos ao longo do trajecto foi realizada por um especialista através de inspeção visual aleatória de todos os sinais. As características extraídas nas regiões do núcleo subtalâmico foram comparadas com as registadas fora de esta área.

Posteriormente, uma abordagem híbrida de classificação utilizando métodos de aprendizagem automática -*machine learning*- foi treinada e testada para identificação dos registos localizados no núcleo subtalâmico utilizando as características extraídas baseadas no domínio do tempo e frequência.

Para comparar as sub-regiões funcionais do núcleo subtalâmico, utilizaram-se os sinais identificados em esta região e dividiram-se estas profundidades etiquetadas pelo especialista como núcleo subtalâmico na porção mais dorsal (que terá maior probabilidade de ser motora) e outra ventral (que terá maior probabilidade de ser não motora). Utilizando esta subdivisão, compararam-se as características neurofisiológicas relativas às diferentes sub-regiões funcionais com testes estatísticos.

Conhecendo as limitações da classificação por inspecção visual, desenvolveu-se uma metodologia para refinamento das localizações dos diferentes registos baseada na reconstrução da trajetória dos eléctrodos implantados, utilizando imagens pre- e pos- operatórias. Esta reconstrução foi sobreposta a um atlas contendo as subdivisões funcionais do núcleo subtalâmico. Os resultados preliminares desta aplicação num sujeito, permitiram mostrar uma classificação mais realista destas sub-regiões na identificação da área motora vs. não motora (límbica e associativa).

Resultados

Foram identificadas várias características significativamente diferentes no domínio do tempo e frequência, entre os sinais classificados como pertencentes ao Núcleo Subtalámico e os sinais não-pertencentes. Estas diferenças permitiram o desenvolvimento de classificadores do tipo de máquinas de vectores de suporte, utilizando as localizações dos diferentes registos tendo como base a classificação baseada em inspecção visual por um especialista.

Utilizando os sinais de maior certeza de classificação, obteve-se uma precisão de classificação do núcleo subtalâmico de 96,18 $\pm$ 3,15 % (para 30 trajetórias, 5 pacientes) usando uma abordagem de validação cruzada de "deixar um sujeito fora"com máquina de vectores de suporte linear. Este modelo de validação garante que os dados relativos aos doentes utilizados para treinar o modelo não foram usados posteriormente para testar o mesmo. Utilizando todos os sinais considerados como núcleo subtalâmico independentemente do nível de certeza, obteve-se uma classificação de 84,32 $\pm$ 5,75 %.

A ferramenta não supervisionada para *spike sorting* permitiu a identificação de um total de 357 neurónios do núcleo subtalâmico identificado com a máxima certeza em 5 doentes (155 registos de microeléctrodos de 10 segundos de duração) e 521 neurónios foram extraídos de 420

sinais localizados fora do núcleo subtalâmico. Utilizando segmentos de 10 segundos livres de artefactos, identificamos 258 neurónios em 116 sinais do núcleo subtalâmico e 329 neurónios em 228 registos de não-Subtalámico. Foram encontradas diferenças significativas em em 13 das 15 características analisadas comparando STN com não-STN.

As subregiões funcionais do núcleo subtalâmico (dorsal vs. ventral) foram comparadas quanto às caracteristicas de tempo, frequência e actividade neuronal. Encontraram-se diferenças significativas em 6 características extraídas dos sinais sem artefactos, consistentes com a literatura e também em uma característica de frequência. Em linha com o descrito previamente, encontramos um *burts rate* mais elevado na subdivisão dorsal em relação à ventral (mediana (amplitude interquartil) of 1.8(1.5) vs 1.15(0.05) bursts/s, p=0.001) e um *firing rate* mais alto (mediana (amplitude interquartil) of 21.4(16.85) vs. 15.3(14.33), p=0.013) entre outras características.

Os resultados preliminares da análise de sinal refinado por imagem num sujeito, mostraram uma alta accuracy do expert na classificação de sinais STN/não-STN comparados com imagem (>85%). Apesar de neste sujeito não se obterem diferenças significativas entre regiões nas características extraídas, esta analise mostrou que divisão heurística dorsal/ventral era insuficiente para proceder à analise apropriada.

Conclusões

Foi desenvolvida uma ferramenta para análise de registo de microeléctrodos em humanos e para extração de características no domínio do tempo e frequência, que forneceu bons resultados na classificação do núcleo subtalâmico utilizando a identificação da localização de cada registo por um especialista.

A análise e extração de características relativas à atividade neuronal foi realizada por uma ferramenta não supervisionada, que foi desenvolvida combinando algoritmos existentes e que pode ser generalizável a registos em outras regiões ou outro tipo de registos. Esta ferramenta permitiu a identificação de características que permitem diferenciar os sinais colhidos no núcleo subtalâmico dos identificados fora de esta região.

De acordo com a literatura, fomos capazes de identificar diferenças em segmentos funcionalmente segregados do núcleo subtalâmico. O trabalho em curso, com refinação anatómica, vai-nos permitir avaliar a sua utilidade em optimizar o posicionamento dos elétrodos, podendo também ser utilizado para fins de investigação.

Palavras chave

Estimulação cerebral profunda; Doença de Parkinson; núcleo subtalâmico; registos de microeléctrodos intraoperatórios; máquinas de vector de suporte

x

Contents

List of Figures

List of Tables

Acronyms

AC anterior commisure

AI assimetry index

ANTs Advanced Normalization Tools

AP action potential

BA basal ganglia

BI burst index

BR burst rate

CF crest factor

CT computed tomography

CV cross validation

cvISI coefficient of variation of interspike interval

CWT continuous wavelet transform

DARTEL Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra

DA dopamine

DBS deep brain stimulation

DICOM Digital Imaging and Communications in Medicine

dSTN dorsal subthalamic nucleus

FF fano factor

FN false negative

FP false positive

FR firing rate

GMM gaussian mixture model

GPe external Globus Pallidus

GPi internal Globus Pallidus

ISI interspike interval

IS isolation score

LDA linear discriminant analysis

MAVS mean absolute value slope

MAV mean absolute value

MBI modified burst index

MC Mel-frequency cepstral coefficients

mdnBD median burst duration

mdnFRQ median frequency

mdnIBF median intraburst frequency

mdnIBI median interburst interval

mdnISI median interspike interval

MER microelectrode recording

mFRQ mean frequency

MNI Montreal Neurological Institute

MRI magnetic resonance imaging

PCA principal components analysis

PC posterior commisure

PCs principal components

PD Parkinson's Disease

PI pause index

PR pause ratio

PaCER Precise and Convenient Electrode Reconstruction for Deep Brain Stimulation

RMS root mean square

RSA Rank Surprise algorithm

SDN standard deviation of noise

SF spectral flux

SNc Substantia Nigra pars compacta

SPC spectral centroid

SPM12 Statistical Parametric Mapping 12

stdISI standard deviation of interspike interval

STN subthalamic nucleus

SVM support vector machine

SppA spike peak-to-peak amplitude

TE Teager energy

TN true negatives

TP true positives

VAR variance

vSTN ventral subthalamic nucleus

WL waveform of curve length

ZC zero crossings

Chapter 1

Introduction

Deep Brain Stimulation (DBS) is a common treatment for advanced Parkinson's Disease (PD). In this therapeutic technique, a small pair of electrodes is implanted surgically in a target area (subthalamic nucleus (STN) or the pars interna of Globus Pallidus (GPi)) to reverse motor symptoms through its stimulation.

During surgery, clinicians decide the final electrodes position based on different and complementary approaches: 1) target coordinates are defined prior to surgery based on patient's pre-operative magnetic resonance imaging (MRI) and computed tomography (CT) imaging. 2) During surgery, intra-operative microelectrode recordings (MER) along the trajectory to the target are visually inspected, and borders of STN identified. 3) Afterwards, final lead position is reviewed through therapeutic and side effects assessment by intra-operative stimulation.

In STN-DBS, a proper placement of the final electrode and particularly stimulation within sensorimotor STN is related with the best clinical results. However, identification of this region on regular MRI is impossible and MER-based features that can distinguish clearly this region remain to be clarified.

In this study, we contribute to the development of clinical-decision support tools for target identification in STN-DBS, through computational analysis using intraoperative MER. Objectives are explained in depth in chapter 2, but theoretical background regarding Parkinson's Disease and STN-DBS is presented in the following section in order to provide better understanding of the problem and purpose of our study.

1.1 Parkinson's Disease

Parkinson's Disease is the second most common neurodegenerative disease worldwide. It is a chronic and progressive disease and its annual incidence is estimated 15 per 100,000 people according to Tysnes and Storstein (2017). Its prevalence increases with age with 1 % of population

over 60 years old suffering from PD. Genetic factors are thought to be involved in all patients but 5–10% of the cases are thought to be monogenetic (caused by mutation in a single gene).

Although its cause is unknown in most cases, several environmental factors are related with increased risk of PD (Tysnes & Storstein, 2017). Previous studies also present higher incidence rate of PD in men than women (Wooten, Currie, Bovbjerg, Lee, & Patrie, 2004).

1.1.1 Symptomatology

Parkinson's Disease symptomatology is characterized by both motor and non-motor characteristics. These symptoms vary between patients and progression of the disease, which are usually mild at early stages (Hughes, Daniel, Kilford, & Lees, 1992; Gelb, Oliver, & Gilman, 1999).

Motor symptoms, also known as cardinal manifestations, are tremor, bradykinesia (slowness of movement), rigidity (muscle stiffness), akinesia (loss or impairment of the power of voluntary movement) and hypokinesia (decreased bodily movement) (Gelb et al., 1999). Other symptoms that may be present on PD patients are such as postural instability (trouble with balance and falls), dyskinesia (abnormal involuntary movements as jerking, fidgeting, twisting and turning movements), freezing of gait (temporary hesitation of walking, as being stuck in place), shuffling gait (dragging the feet), sialorrhea (excessive drooling, increased salivation and swallowing issues) or hypophonia (soft or muffled speech) (Hughes et al., 1992; Bartels & Leenders, 2009).

In addition to cardinal symptoms, also non-motor characteristics are common and diverse in PD. Depression, dementia, cognitive dysfunction and/or sleep disorders occur frequently. Fatigue, sexual problems, loss of sense of smell/taste or psychotic symptoms are more non-motor symptoms that could be present in some PD patients (Maiti, Manna, & Dunbar, 2017; Gelb et al., 1999).

1.1.2 Mechanism underlying PD symptoms

The basal ganglia (BA) are a group of interconnected subcortical structures that play an important function in motor control but also have non-motor roles in executive functions, behaviour or emotions. Subthalamic nucleus (STN), Substantia Nigra pars compacta (SNc), internal Globus Pallidus (GPi), external Globus Pallidus (GPe) and striatum (with both caudate nucleus and putamen) are the main nuclei in the basal ganglia (Fahn, Jankovic, & Hallett, 2011).

In order to control execution of movements, motor information is modulated by the combination of a net excitatory (direct) and inhibitory (indirect) pathway, through basal ganglia structures and cortex, as a closed circuit illustrated in figure 1.1 (left).

Parkinson's Disease is characterized by progressive degeneration of neurons in SNc, decreasing the secretion of dopamine (DA), an essential brain monoamine which regulates the excitability

of striatal neurons. Effects of loss of DA lead to alterations in neuronal activity resulting in difficulties for movement control and reflect both in indirect and direct pathways, as represented with thinner arrows in the Parkinson's circuitry on figure 1.1. Consequently, GPi is hyper-activated and highly increases inhibition signals in the thalamus, finally inhibiting the output to motor cortex and consequently reducing the control of voluntary movements (Maiti et al., 2017).
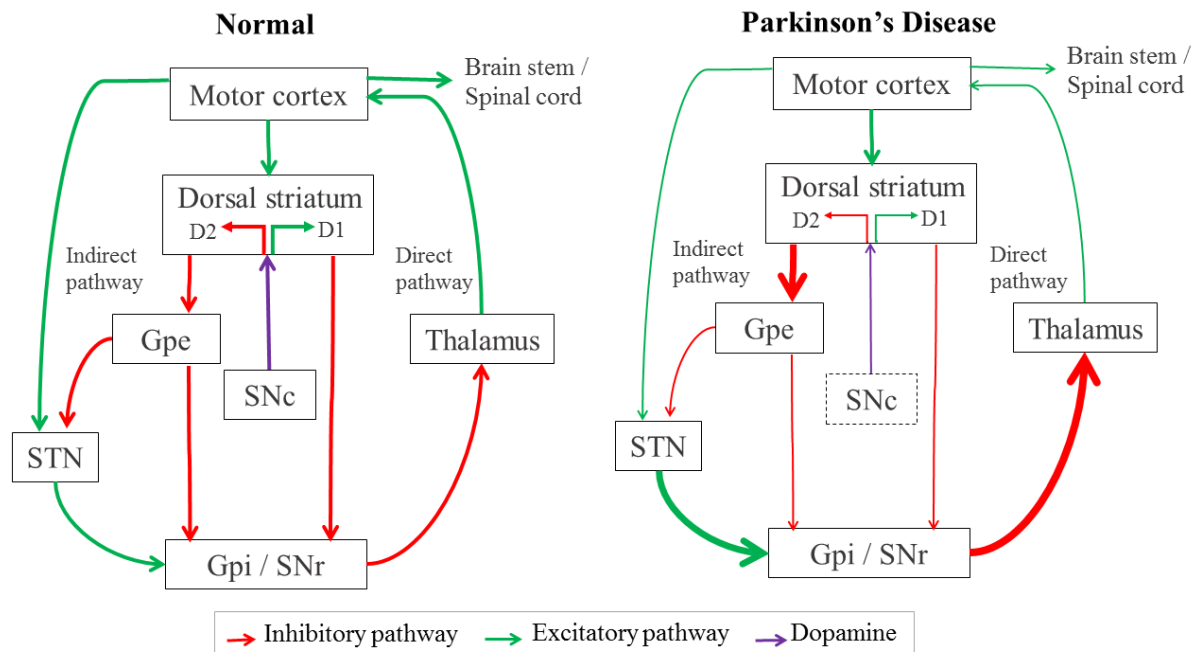
Figure 1.1: Circuity between basal ganglia-thalamo-cortical motor structures with simplified inhibitory and excitatory connections, in red and green respectively. The left panel indicates the "normal" state, and the right shows the overall changes in activity that have been associated with Parkinson's Disease. Dopamine from SNc is represented in purple. Thickness of each arrow is representative (directly) with its activity. For simplicity some connections have been omitted from this diagram.

1.1.3 Treatments for Parkinson's Disease

Different therapies are currently available for the treatment of motor symptoms in PD. At the present, both medication and surgery are the most frequently used therapies to treat PD symptoms, but new promising treatments are emerging and being studied such as stem cell transplantation or gene therapy (Maiti et al., 2017).

During early stages of PD treatment, medication typically is the best option to control motor symptoms. The most commonly used drug is Levodopa (L-dopa) which is a dopaminergic drug that replaces dopaminergic loss associated with SNc degeneration. Nevertheless, since increment of dopamine occurs in other brain regions beyond the basal ganglia, it may generate adverse effects, specially in long-term treatments, such as motor fluctuations or dyskinesias.

For patients with motor symptoms refractory to medication, different therapeutic approaches can be tried. Surgical approaches include deep brain stimulation and lesion surgery. Lesion surgeries for PD are known as pallidotomy and thalamotomy referring to the respective surgically lesioned part: globus pallidus or thalamus. Motor symptoms decrease with these procedures. It is proposed that the lesions reverse the elevated inhibitory neural activity of these structures. Nevertheless, since lesion surgery is irreversible and lesions in adjacent areas may occur, electrical neuro-stimulation surgery (Deep Brain Stimulation) is a good alternative and currently most commonly used for PD.

1.2 Deep Brain Stimulation

Deep Brain Stimulation (DBS) is a surgical treatment for refractory movement disorders, such as PD, dystonia or tremor. A target area of the brain is stimulated through surgically-implanted electrodes to reverse motor symptoms by inducing neuronal activity alterations.
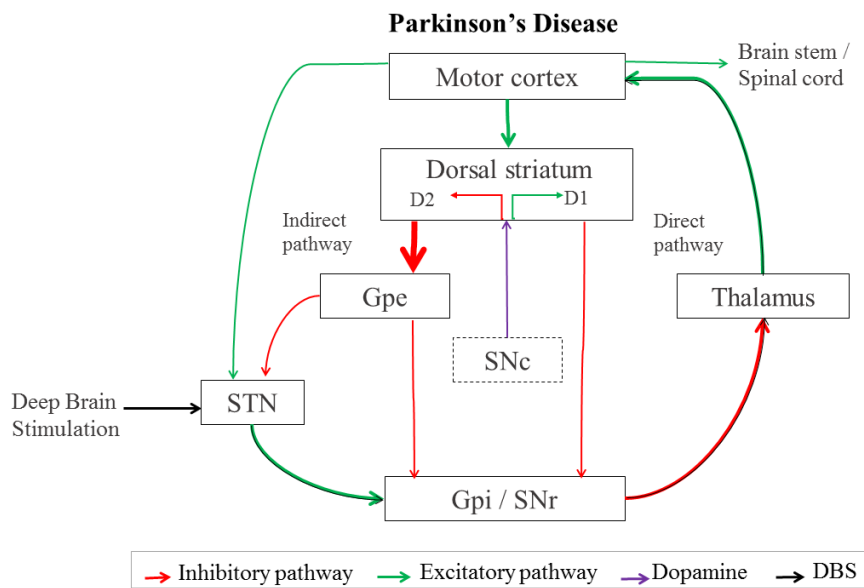
Figure 1.2: Basal ganglia-thalamo-cortical circuitry in Parkinson's Disease with alterations induced by STN-DBS in both inhibitory and excitatory pathways. Connecting pathways inhibitory (in red), excitatory (in green) and DA from SNc is represented in purple. Thickness of each arrow is representative (directly) with its activity.

Up to 20% of PD patients are considered suitable candidates for Deep Brain Stimulation (DBS), commonly patients in an advanced stage of the disease, with motor fluctuations, and preferentially without psychiatric illnesses (Wagle Shukla & Okun, 2014).

Different proposals on the mechanism underlying STN-DBS were set based on the known basal ganglia circuitry, figure 1.1. However, it still remains fully unclear what mechanisms underlie

STN-DBS. As illustrated in the basal ganglia neural network in figure 1.2, artificial stimulation in the hyper-activated STN may provoke an inhibition of its firing rate that activates the GPi, consequently inhibiting the thalamus and improving movement control from motor cortex (Maiti et al., 2017; Gradinaru, Mogri, Thompson, Henderson, & Deisseroth, 2009; Negida et al., 2018).

The subthalamic nucleus is divided into three functional regions: sensorimotor, limbic and associative, which are located in dorsolateral, anteroventral and medial territories respectively as shown in figure 1.3, and it is part of a bigger functional structure called the basal ganglia as mentioned before. Previous studies suggest that precise positioning of the electrode and stimulation within the sensorimotor part of the STN is of major importance for optimal results and to avoid side effects like mania (Castrioto, Lhommée, Moro, & Krack, 2014; Wodarg et al., 2012; Johnsen, Sunde, Mogensen, & Østergaard, 2010). However, identification of sensorimotor STN region based on 1.5 Tesla MRI images is impossible and still unclear based on MER characteristics.
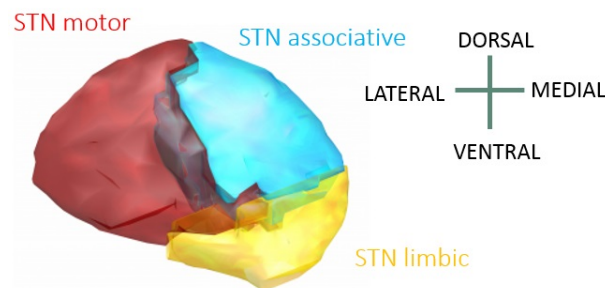
Figure 1.3: Subthalamic Nucleus with functional subdivisions: motor dorsolateral (red), associative medial (blue) and limbic ventro-lateral (yellow) STN. Adapted from Accolla's atlas.

1.2.1   DBS procedure

Previously to surgery and stereotactic frame implantation, pre-operative MRI is acquired in order to subsequently plan the lead trajectory based on the patient's anatomy. Pre-implantation contrast-enhanced computed tomography (CT) is performed to the patient, after installation of the stereotactic frame similar to the one in figure 1.4 (right), which displays a reference scale through its ferromagnetic materials.

Estimation of STN target coordinates is based on both CT and MRI pre-operative images co-registered and its STN localization in relation to its fiducial points, anterior commisure (AC) and posterior commisure (PC), as illustrated in 1.4 (left). For both hemispheres (one at a time), different trajectories and strategies are discussed and visualized to achieve the best trajectory based on merged images in both three views (sagital, coronal, axial) and also as *probe view*, which facilitates avoiding problematic trajectories. The best electrode's trajectory avoids white matter, critical cerebral tissue, veins or other conflict points.
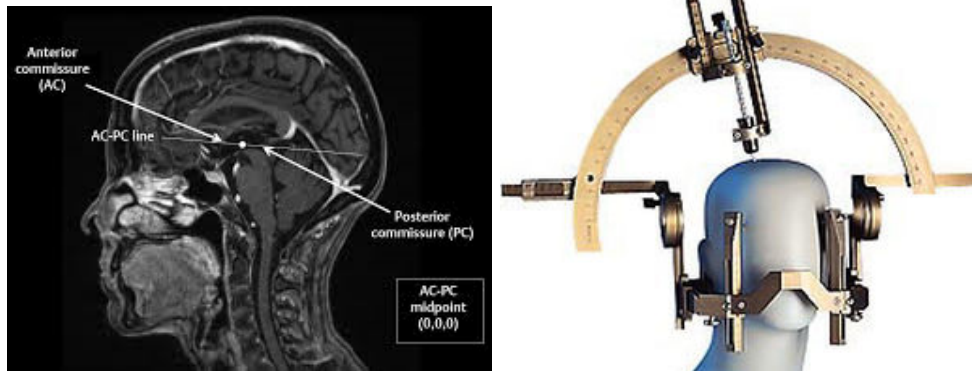
Figure 1.4: Fiducial points identified in coronal view anatomic image (left). Stereotactic frame. From Leksell, Elekta.

Once target coordinates are selected, surgery is performed one hemisphere at a time. Stereotactic frame is configured according to the target coordinates and after burr hole trepanations, electrodes are connected and implanted into the brain according to the fixed coordinates in the frame.

Registration of microelectrode recordings (MER) is subsequently performed, acquiring neuronal activity at different controlled depths along the trajectory to the planned target, as illustrated in figure 1.5. Differences in the signal are discernible between different regions of the recorded brain, as neurons present different patterns of spontaneous discharge (Starr, 2002). Recordings used in this study were acquired using 3 channels -central, lateral and anterior- as represented in figure 1.5A. Although up to 5 electrodes (concentrically distributed) are available to use for each lead, they are not commonly used to reduce the risk of complications.

Acquisitions of MER start when electrodes are positioned one centimeter above the planned target, as illustrated in figure 1.5C, and recordings are acquired at different depths simultaneously for all parallel channels. According to the acquired signals and based on intraoperative visual inspection, the region of the brain in which the recording electrode is located, is identified in a table for each recording point. After registration in the different depths and storage for off-line analysis, a reconstruction of the shape of the target is estimated using MER annotations and possible definitive target coordinates are determined/refined.

Intraoperative stimulation is then executed in order to refine the final localization of lead based on therapeutic and side effects assessment. Different estimated points of interest are stimulated through inserted macro-electrodes with gradual variations of both intensities and frequency. Once stimulation induced side effects, such as eye or muscle contractions, are assessed and the best position is determined, definitive electrode is implanted.

The second part of surgery includes the implementation of the neurostimulator with the battery in the patient and its connection to the electrodes as illustrated in figure 1.6 from Medtronic
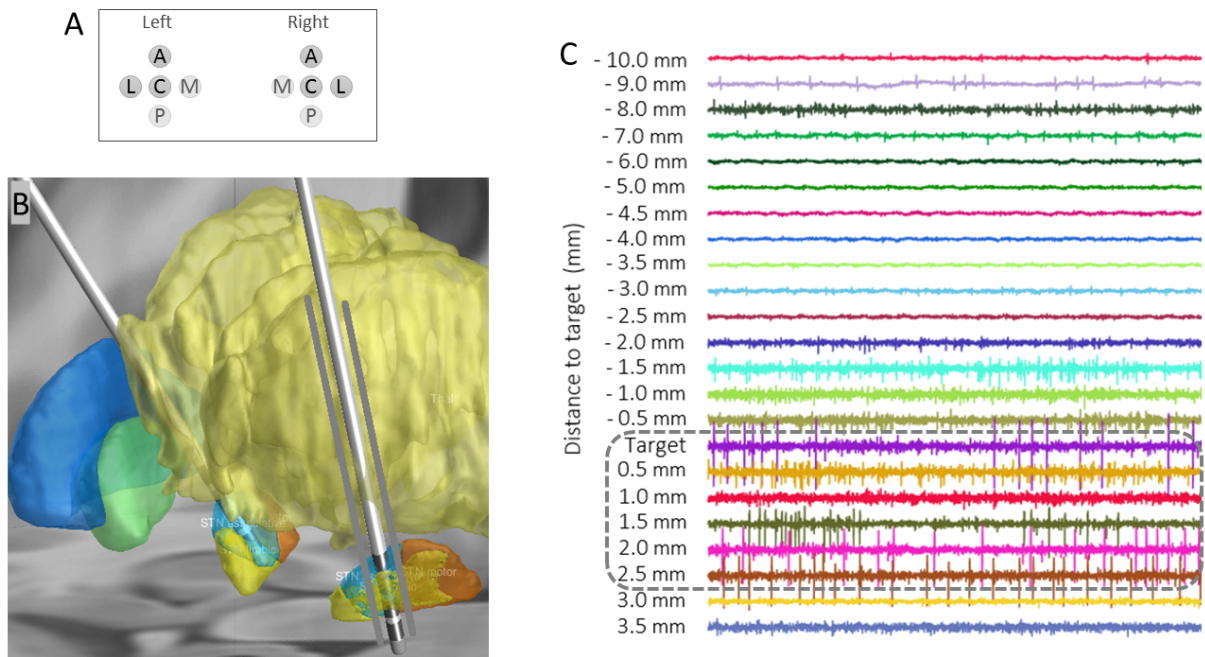
Figure 1.5: Position of the channels for intraoperative microelectrode recordings (A). Trajectory of the lead to STN with subdivisions and GPi in green, GPe in blue and thalamus in yellow. Lateral and anterior channels illustrated as grey lines (B). MER signals recorded through the trajectory with center, lateral and anterior electrodes. Recordings presumably from STN are identified within the grey box. Segments of MERs (500 ms duration) with central channel at each depth along the trajectory with central electrode, starting recordings from 10 mm above target to 3.5 mm under (C).

(2007). Both rechargeable and not rechargeable stimulators are available on the market for DBS. Rechargeable batteries may last between 9 and 25 years whereas non-rechargeable batteries may need to be replaced in 3 to 5 years depending on the patient and stimulation parameters (Hariz, 2017).
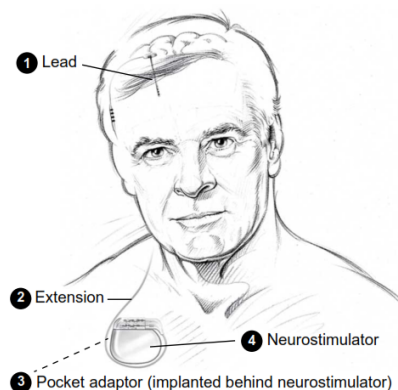
Figure 1.6: Illustration of basic implanted components of the DBS System Components: lead, extension, neurostimulator and pocket adaptor. From Medtronic.

Chapter 2

Objectives

Clinical-support tools for identification of the target area in DBS are necessary in the field. As mentioned before, positioning the final electrodes in the precise location has a vital role in obtaining the best therapeutic results. However, this decision depends on the expertise of the intra-operative evaluation performed by the neurologist, making it a time-consuming and subjective process.

In this study we intend to contribute to the development of computational tools that can assist target identification in STN-DBS. We split our objective into four main tasks:

1. Identification of features based on time and frequency domain to distinguish STN from non-STN intraoperative MER. Development of an unbiased unsupervised tool for extraction of these MER features through loading, preprocessing and analysis, but easily adaptable for other types of signals and regions.

2. Development of an automatic approach for spike sorting and signal analysis of human brain MER. Through this tool, features related with neuronal activity are studied to find relevant characteristics for STN identification. We aim to construct this tool generalizable to other brain regions and unsupervised through the adjustment of existing algorithms to our approach.

3. Development of a hybrid unsupervised/supervised machine learning classification approach that uses extracted MER features based on time and frequency domain for high-accuracy identification of STN.

4. Identification of features that distinguish sensorimotor subdivision of STN vs. limbic and associative using extracted features from both previously constructed tools. This approach can be optimized using individual patient lead trajectory localization reconstruction based on imagiology, fused with an STN functional subdivision atlas.

Chapter 3

Methods

In the present chapter, we describe the dataset used in the analysis. Then, procedures for labelling each MER regarding its position within the target area of the brain are detailed followed by the pre-processing approaches for MER to clean the signals. Methods for extraction of features, both directly from signal and from neuro-physiological information, are described in sections 3.5 and 3.4.

3.1   Microelectrode recordings (MER) dataset

Our database was collected from 5 PD patients undergoing bilateral STN-DBS. Three registry electrodes were used in each hemisphere and the lateral and anterior channels were positioned 2 mm from the central one (impedance of each microelectrode is 1 M$\Omega \pm 20\%$ at 1000 Hz). Recordings started 10 mm above the estimated target and ended approximately 3.5 mm after, thus signals were acquired at around 23 different depths in each channel, as shown in figure 1.5. Acquisitions were obtained through a MER system (Medtronic, Minnesota, MN, USA), sampled at 24 kHz per channel, filtered (band-pass filtering, 0.5-5 kHz) and amplified (with total gain of 1000), previously to their storage. Signal was recorded during 10 seconds for 3 patients, and 30 seconds for 2 patients and analysis were performed on 10 seconds duration MER segments.

3.2   Localization of MER position

Labelling each MER of our dataset regarding its position is important to effectively identify features distinguishing STN and non-STN MER and to construct the automatic classifier. Therefore, different approaches are explored for labelling MER regarding its location within the target, both further explained in the following sections.

Expert based classification was used in previous studies using 1) intraoperative annotations based on visual inspection during recording phase of DBS surgery, 2) off-line MER labelling or 3) other expert-based approaches (Cagnan et al., 2011; He, 2009; Bakštein et al., 2015). Nevertheless, expert classification as STN vs. non-STN based on visual inspection may be biased by previous geometric knowledge of the MER trajectories and/or estimated target area. Therefore, we approached this issue trying to reduce the bias with offline labelling of MER segments using random order.

Ongoing work is refining the labelling, making it based on lead trajectory localization reconstruction using individual patient imagiology, fused with an STN functional subdivision atlas. Gold-standard classification could be obtained through this approach and it allows us to compare physiological neurons properties in motor/non-motor STN and it is further explained in section 3.2.2.

3.2.1 Expert-based classification

Labels regarding each signal's position within the brain (STN/non-STN) were identified through visual inspection by an expert as previously referenced. In order to facilitate visualization and labelling of all signals and to preserve the classification unbiased to knowledge about MER depths, a customized script in Matlab was programmed.
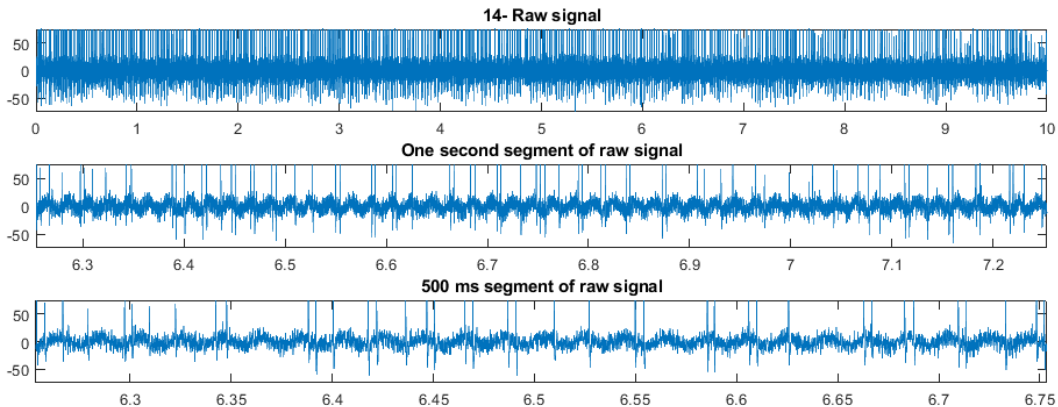
Figure 3.1: Segments of MERs (1 s and 500 ms duration) and raw signal shown for classification by the expert through visual inspection.

This script first shuffles all signals in the dataset and it plots in random order one signal at a time. The whole signal and two enlarged segments (1 second and 500 milliseconds duration, randomly chosen from the raw signal) are presented for classification, as shown in figure 3.1 for each signal in the dataset. Then, the expert is asked to identify the current signal as *STN (1)*, *not-STN (2)* or *try again (3)*. If any signal is classified as *try again* for the first time, it is shown again until the third attempt, when it is classified as *unclear (3)*. Additionally for each signal

classified as STN, the expert is asked to set its value of certainty of its localization from 1 (very unclear) to 5 (high certainty), which gives an idea of the confidence of each classification for further analysis.

However, labelling subterritories of STN based on visual inspection is not an optimal option since differences are not clearly known or visually recognizable on MER. Accordingly, to explore if preliminary differences may be present between STN subregions, this subdivision is manually defined on the consecutive previously labelled MER. Previously defined STN recordings were split in two parts: dorsal and ventral STN, for the shallower and deeper depths labelled as STN respectively as illustrated in figure 3.2. Additionally, MER classified as STN located between both deep and shallow regions are discarded, resulting in labels for both dorsal and ventral regions, with a higher probability of being motor and non-motor STN respectively.
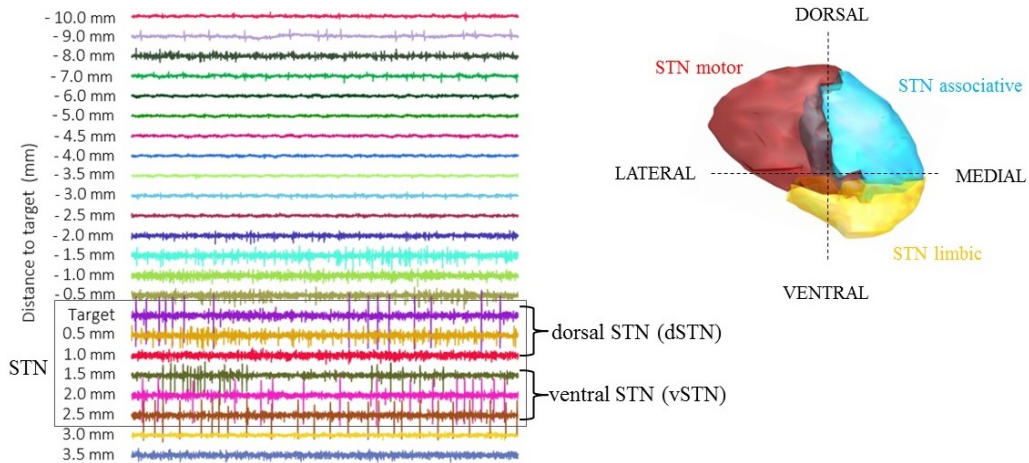
Figure 3.2: Segments of MER at different depths with STN expert identification and its subdivision on dorsal and ventral STN.

3.2.2 Anatomically-based classification

Reconstruction of trajectory localization for each lead presents a gold-standard approach, since it is based on both pre-operative and post-operative patient's imagiology fused to an atlas of STN functional subterritories. Ongoing analysis to localize MER based on the reconstructed electrode tip and its trajectory with patient's image processing was performed with LeadDBS, a Matlab toolbox [http://www.lead-dbs.org; (Horn & Kühn, 2015)]. Along with Slicer3D [https://www.slicer.org/], used to import all provided Digital Imaging and Communications in Medicine (DICOM) images into Nifti, which is the default file format in the used packages. Nifti stands for Neuroimaging Informatics Technology Initiative and it is a medical image file format created to simplify post-processing analysis (Larobina & Murino, 2014).

First, registration of all pre- and post-operative images is performed, which consists in the geometric alignment of these images between each other through transformations to allow their fusion. Therefore and through Lead DBS, MRI images are coregistered using Statistical Parametric Mapping 12 (SPM12) method [http://www.fil.ion.ucl.ac.uk/spm/software/spm12/], whereas post-operative contrast-enhanced CT image was coregistered to preoperative MRI images with Advanced Normalization Tools (ANTs) [http://stnava.github.io/ANTs/].
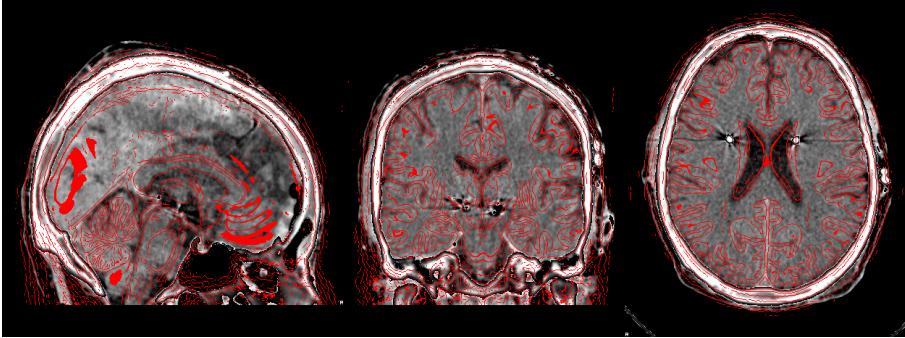
Figure 3.3: Normalization of post-operative CT to the MNI2009b space, previously registered to pre-operative MRIs.

Normalization is then performed to transform the previous volumes to a known space in which we will superpose an atlas. Consequently, all volumes are normalized to standard stereotactic MNI space (ICBM152 2009b non-linear asymmetric) using Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra (DARTEL) implemented in SPM12 (Friston, Ashburner, Kiebel, Nichols, & Penny, 2006).

After coregistration and normalization, subcortical refinement of previous transforms is performed to correct brainshift, which may happen during surgery when the skull is opened and air gets inside, moving the brain in respect to the skull. We applied this process through Lead-DBS using specific masks from Schönecker, Kupsch, Kühn, Schneider, and Hoffmann (2009) in 2 patients, to estimate a transform for refinement of the overlapping in subcortical regions, as illustrated in figure 3.4.

In order to obtain electrode's location and trajectory, reconstruction is manually supervised assessed with pre-reconstruction through Precise and Convenient Electrode Reconstruction for Deep Brain Stimulation (PaCER) method and Acolla's STN subterritories atlas (Accolla et al., 2014; Husch, Petersen, Gemmar, Goncalves, & Hertel, 2018), as illustrated in figure 3.5.

Once definitive coordinates of electrode are located, MER localizer tool in LeadDBS is used to determinate locations along the trajectory in the selected channels in relation to the STN-subdivisions atlas, as shown in figure 4.2 where each yellow ball represents the location of each introduced recording site. This step is currently ongoing, therefore only data from patient 4 (which presented the majority of STN located MER according to the expert labelling) is presented in the results for STN subterritories refinement, in section 4.5. However, all patient's trajectories
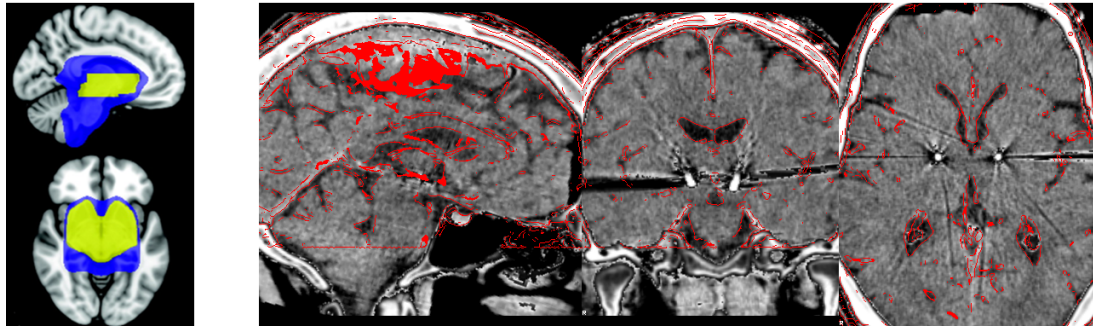
Figure 3.4: Application area for masks (Schönecker et al., 2009) (left) and results of subcortical refinement on normalized images (right).
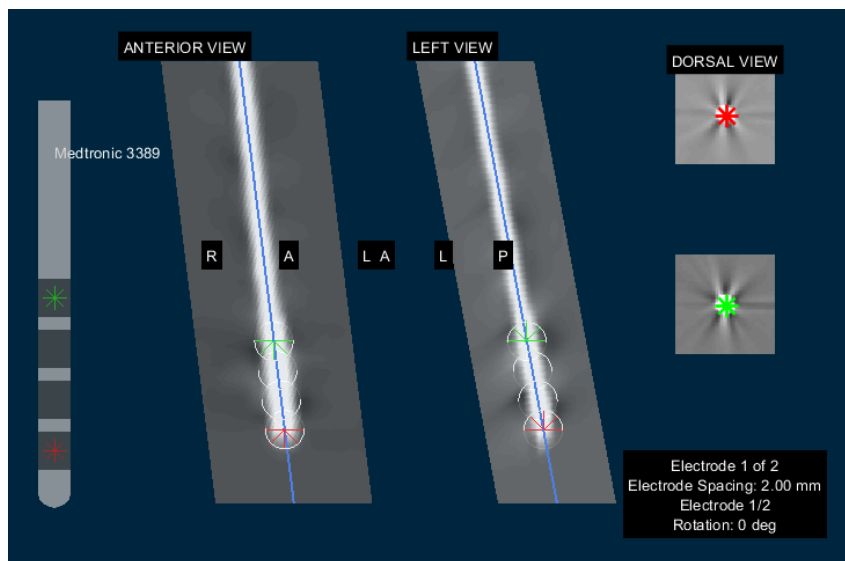
Figure 3.5: Lead reconstruction interface with Lead DBS using PaCER.

have been reconstructed, as illustrated in figure 3.6.
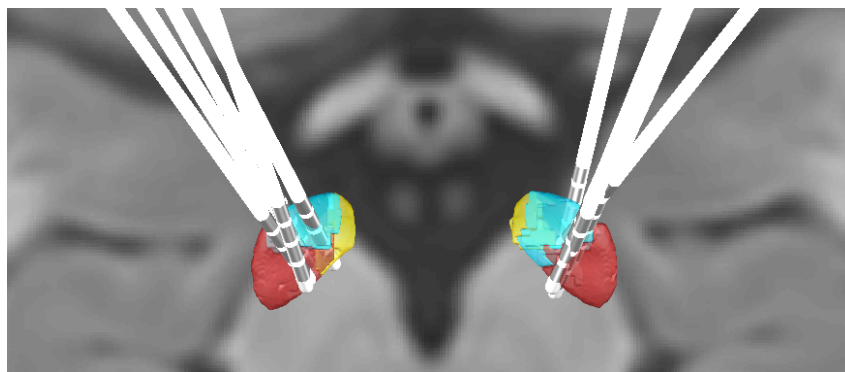
Figure 3.6: Axial view of STN subdivisions with reconstructed electrodes based on each patient's images through Lead DBS.

3.3 MER filtering and artifact detection

Noisy signals are present in MER dataset with different types of artifacts. According to Bakštein et al. (2017), noise in MER may affect more than 25% of the recording length and can arise due to a number of factors such as mechanical movement manifested by high power signal peaks, electromagnetic interference such as the 50 Hz interference from the power grid, low-frequency interferences (<50 Hz) and "irritated neurons" with high spiking activity. Consequently, pre-processing these signals is an important step to avoid errors in the extracted features and obtain better accuracy results in classification.
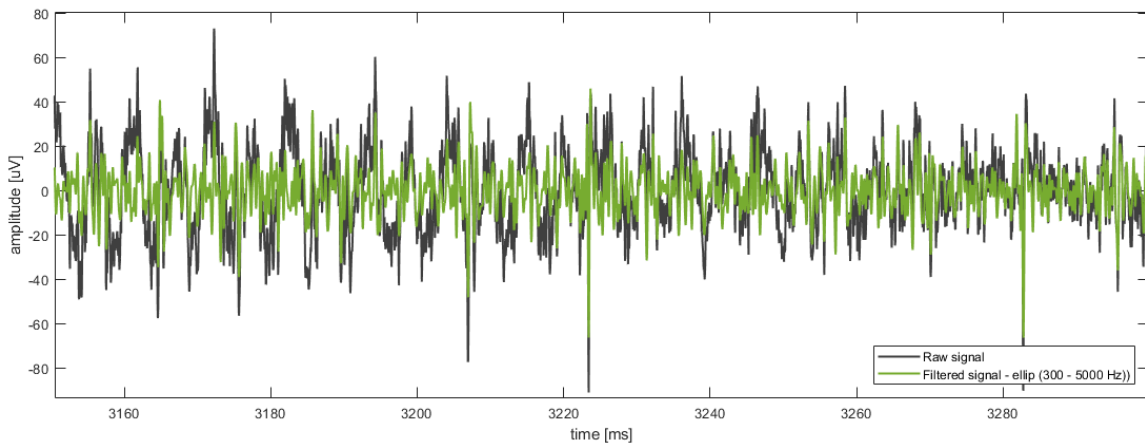
Figure 3.7: Filtered signal (150 ms) with elliptical band pass filter (green) overlapped to raw MER (grey) presenting low-frequency interference.

Filtering the signal is our first step in this pre-processing part, since it can smooth out high-frequency fluctuations and/or remove periodic trends of a specific frequency from data. Therefore, we filter the data using an elliptic band-pass filter between 300 and 5000 Hz. Elliptic filter was chosen after visual inspection of different types of filters (Butterworth and Chebyshev type II) and its effects on spike detection, since it keeps waveforms of firing neurons unaltered (Quiroga, 2009). Implemented filter of order 4 is set to 0.6 decibels (dB) of peak-to-peak ripple and 60 dB of attenuation and it is performed with zero-phase digital filtering in order to preserve time features. Nevertheless, all parameters can be easily modified in our tool and code for testing other types of filters is implemented.

In addition to the filtering, we reviewed previous approaches for MER artifact removal literature and visually inspected their performances on our dataset (O'Shea & Shenoy, 2017; Bakštein et al., 2017, 2015; Grubhoffer, 2016). Best observed approach for artifact detection was found using the autocorrelation-based method as implemented by Bakštein et al. (2015), but this algorithm was actually previously developed by Aboy and Falkenberg (2006). We adapted Bakštein et al.'s open-source code to extract the longest consecutive clean segment from signal used for

extraction of neuronal activity features, as illustrated in figure 3.8.
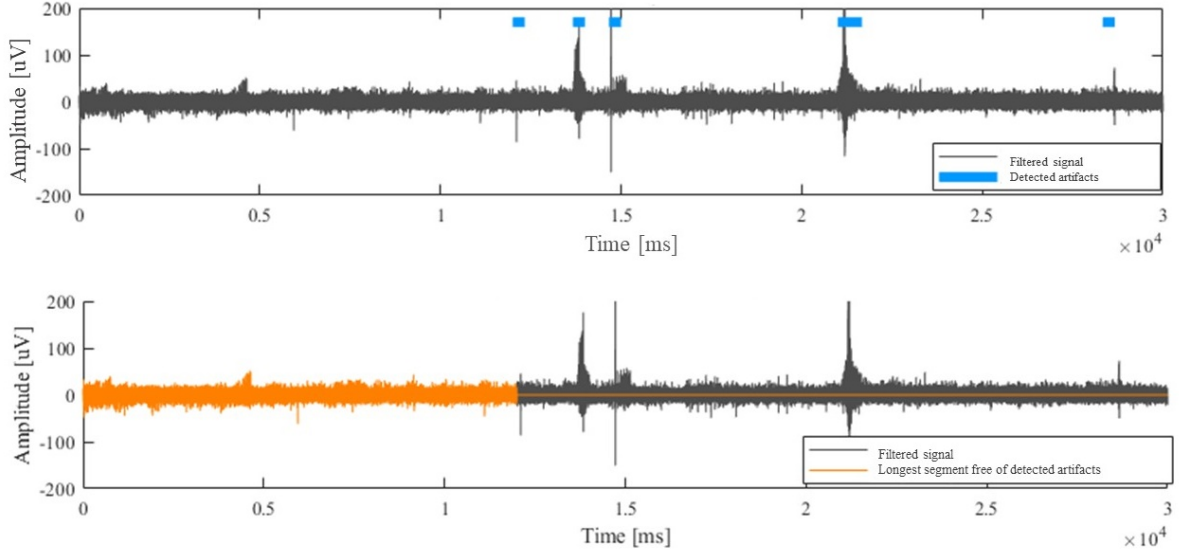
Figure 3.8: Artifacts detection. Filtered signal with detected noisy segments in blue (Top). Longest clean segment free of detected artifacts in orange (Bottom).

3.4   Extraction of features related with time, frequency and noise

As previously described, differences between raw signals along the trajectory to target are discernible through visual inspection. Therefore signal characteristics calculated through computational analysis might reflect these variations. In line with this, features related with time, frequency and noise are extracted from filtered MER to find features for distinguishing STN from non-STN recordings.

In order to identify STN, custom Matlab code was developed for MER analysis. After setting input parameters, the analysis was done in an unsupervised way, following steps presented in figure 3.9. Each raw signal is first filtered as explained in section 3.3 followed by artifact detection, in order to obtain segments free of artifacts.

Filtered signal is divided in smaller segments with Hanning windows of 2048 samples overlapped 50% (1024 samples), resulting in more than 230 segments for each used segment for analysis of 10 second (sampled at 24 kHz). Characteristics are computed for all frames independently and then statistics (mean, median and standard deviation) are stored for each MER using all segment's values. This division in frames should provide a more robust approach for signals with noisy segments or high variability within the same signal vs. extraction of characteristics from the whole signal.

Features related with time are then obtained directly from each frame: mean absolute
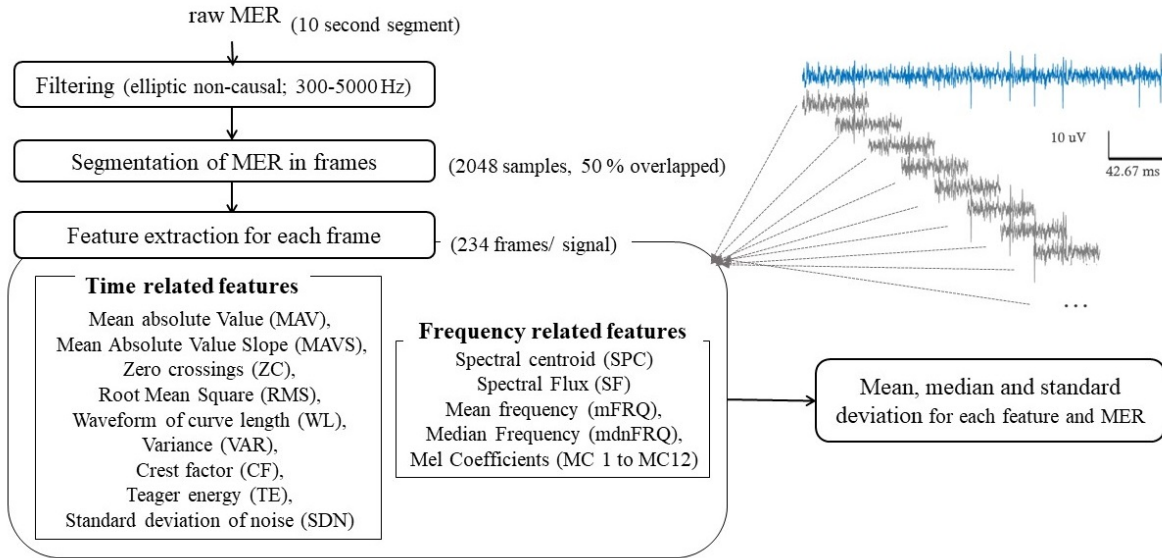
Figure 3.9: Workflow for MER analysis and extraction of time and frequency related features.

value (MAV), mean absolute value slope (MAVS), zero crossings (ZC), root mean square (RMS), waveform of curve length (WL), Teager energy (TE), variance (VAR) crest factor (CF). In relation to noise, an estimation of the standard deviation of noise (SDN) was obtained according to Dolan, Martens, Schuurman, and Bour (2009). Features related with frequency domain are: spectral centroid (SPC), spectral flux (SF), mean frequency (mFRQ), median frequency (mdnFRQ) and 12 Mel-frequency cepstral coefficients (MC). We selected this set of features related with time and frequency based on a preliminary study through the related literature for STN identification, MER analysis and typical features used in other signal processing fields such as sound processing (e.g. Mel Coefficients) and further definitions of these features are presented in appendix A.

3.5 Features related with neuronal activity

Microelectrodes register neuronal activity, part of which consists on action potentials (APs), spikes generated from neurons in the nearest region of the electrode tip. Because there are different structures with different functional and cellular density across the lead trajectory, MER of the brain show different neuronal activity. In addition, activity of more than one different neuron near the recording tip can be registered, but spikes from the same neuron present similar shapes and amplitudes in each MER. Consequently, prior to the spike-related features quantification, identification of the different neurons from each recording is important to obtain reliable features.

Action potentials of the different neurons are detected from the MER as spiking events by a process known as spike detection, further explained in section 3.5.1. After storage and

18

alignment of the detected spikes to their maximum, different neurons are separated in different clusters, process detailed in section 3.5.3 and known as spike sorting. Extraction of their features is implemented lastly, once neurons are sorted and detected from each MER. This sequence of steps for signal processing and neuronal characteristics is illustrated in the diagram in figure 3.10.
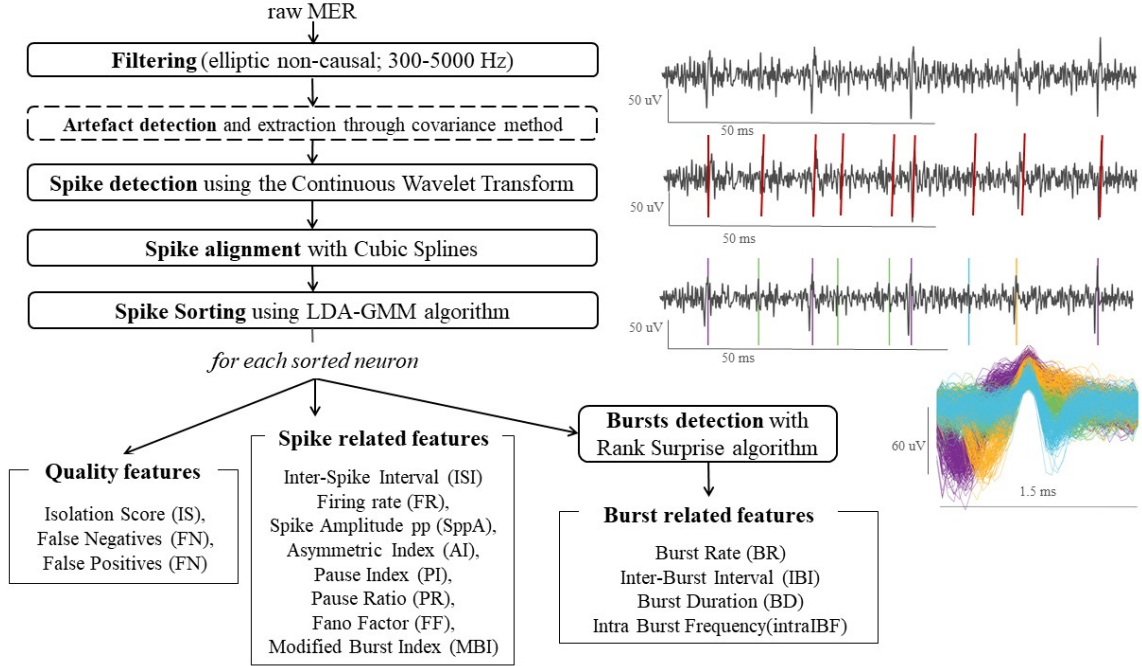
Figure 3.10: Workflow process for neuronal activity analysis and extraction of spike related features.

3.5.1 Spike detection

In this process, firing neurons are detected as spike events from the filtered signal. The chosen algorithm is Spike Detection with Continuous Wavelet Transform (Nenadic & Burdick, 2005). This algorithm uses the continuous wavelet transform (CWT) combined with information about the typical interval duration of APs to identify these spiking events. Biorthogonal family was used with this algorithm, though it can be used with other types of wavelets, but this family presents bi-phasic phase which is similar to the brain's action potentials and when translated and scaled, a bank of approximately matched filters for APs identification along the signal is formed.

Nenadic and Burdick's algorithm (2005) was chosen instead of threshold detection, the most commonly used method in related studies, since it presented better performance with an unsupervised approach on the number of detected spikes, particularly in noisy signals. This may be due to the fact that this method is less affected by the high variability of amplitudes and background noise between all signals and more stable for unsupervised analysis.
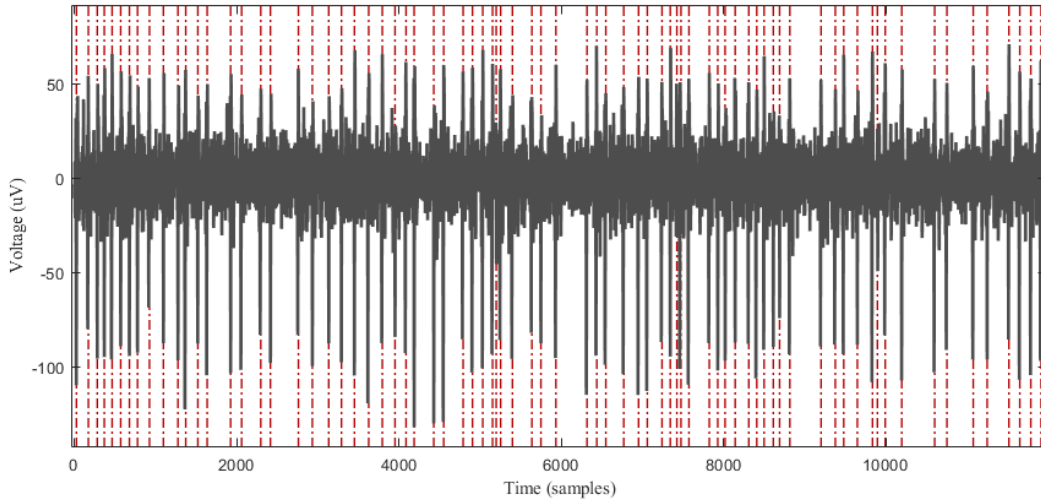
Figure 3.11: Spike train of detected spike events (red) on the clean filtered signal (grey).

3.5.2 Spike alignment

Once spikes are detected, further analysis is needed to store the spikes and sort them according to their neuron of origin, since as mentioned before detected spikes from one signal are probably provoked by more than one different neuron. Nevertheless, the maximum local value of the spike segment could appear on different samples and this miss-alignment may affect the spike sorting process. Therefore, previously to their sorting, a segment of 22 samples before the sample of the detected spike to 26 samples after is stored. Afterwards, cubic spline interpolation is performed to interpolate a curve with 10 times as many points which goes through the same original samples of the spike segment. Then, spikes segments are aligned to their local maximum and extrapolation is performed and the segment is cropped to 1.5 ms (16 samples before the maximum, 20 samples after) and stored for further analysis.

3.5.3 Spike sorting

Detected firing events within the same signal may be originated from more than one different neuron. Consequently, it is important to correctly identify their origin to avoid misclassifications and errors in the spike related features. Spike sorting consists in grouping the detected spikes into clusters based on the similarity of their waveforms.

Since we aim to develop an unsupervised tool for neurophysiological features extraction based on MER, we research the spike sorting literature to find a suitable algorithm to adapt to our analysis tool. Our spike sorting is performed with linear discriminant analysis (LDA), gaussian mixture model (GMM) and outlier detection, through an available algorithm and implemented code in Matlab from Keshtkaran and Yang (2017). This method performs clustering with subspace

learning using LDA, which is a linear transformation technique commonly used for dimensionality reduction, to extract most discriminative features from the spike waveforms. Automatic initial estimation of the number of the clusters is based on GMM, a clustering models that adapts to spherical clusters and facilitates implementation of outlier detection, as explained by the authors.
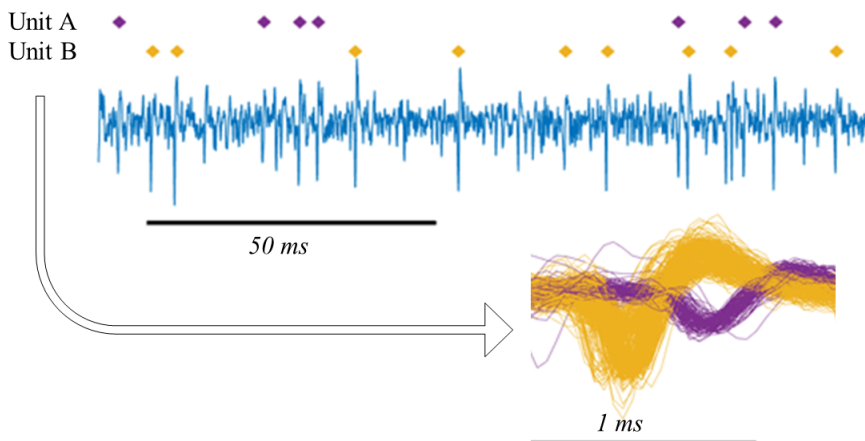
Figure 3.12: Detected spikes in filtered MER after spike detection and sorting (top) and their respective overlapped waveforms (bottom).

To address the quality of each cluster in an objective and standardized way, different related measures are extracted through existing algorithms. Regarding the isolation quality for each cluster, isolation score (IS) is quantified as explained in Joshua, Elias, Levine, and Bergman (2007). This score measures the overlap between the noise and the spike clusters. Also with the same algorithm, an estimation of the proportion of false positive (FP) and false negative (FN) classification errors is obtained.

3.5.4 Spike related features extraction

Different groups of characteristics were studied from the related literature on the field to select features that could provide relevant information for distinguishing STN and non-STN MER. We've extracted features related with the periodicity of each spike train, which is a binary vector with 1 where each neuron's spike event occurs. Features related with the bursting activity for each spike train were also calculated among other spike characteristics. In total, 15 spike related features were extracted through custom Matlab code and also 3 quality measures as explained before. Spike related features are presented following, but further definitions of these features are available in appendix A.

Quantification of the variability of the increments of a spike train is measured through the fano factor (FF), calculated according to the literature (Kuebler & Thivierge, 2014). Median values of spike peak-to-peak amplitude (SppA) are also stored as measurements of spikes events

amplitude, and firing rate (FR) is extracted as a measure of the number of spikes per unit.

A group of features is based on periodicity of the spike trains, specifically related with the interspike interval (ISI). These features are pause index (PI), pause ratio (PR), assimetry index (AI), median interspike interval (mdnISI), standard deviation of interspike interval (stdISI) and coefficient of variation of interspike interval (cvISI). Features for an estimation of bursting activity directly obtained from ISI are modified burst index (MBI) and burst index (BI) (Rajpurohit, Danish, Hargreaves, & Wong, 2015).

Nevertheless, characteristics related with bursting neurons are further calculated after bursts are previously detected through the Rank Surprise algorithm (RSA), which is again based on ISI for detecting bursts (Gourévitch & Eggermont, 2007). Extracted features from bursting neurons are: burst rate (BR), median interburst interval (mdnIBI), median burst duration (mdnBD), median intraburst frequency (mdnIBF).

3.6   Statistics

Statistical comparison of STN and non-STN regions based on MER features was computed through different statistical approaches. Considering independent signals, Student's T-test, a parametric test, was performed in samples with size equal or bigger than 30 or otherwise when following a normal distribution, according to Lilliefors test combined with visual inspection of histograms. Non-parametric Mann Whitney U test, also known as Wilcoxon rank sum test, was used in non-normal distributed features when the sample size is less than 30. Results were considered statistically significant for an α value equal to 0.05.

Random Permutations test (also known as randomization test) was also used to compare features across STN motor vs. non-motor in spike related features to verify the results obtained through the previous method. It is a non-parametric test which doesn't make assumptions about the data through the permutations of the values. Therefore, data from both samples was shuffled 100000 times, creating a reference distribution for the difference in means between the two samples, which is the observed value. Features are considered statistically different across compared regions if the observed value lies outside the confidence intervals of the reference distribution (again for $\alpha = 0.05$).

3.7   Classification with Machine Learning

Automatic identification of the signals recorded in the interest area was performed through classification with machine learning techniques. We've used previously extracted features related with time and frequency combined with the information of the electrode's localization along the trajectory. Classification consists in the categorization of an observation to a class or group

based on previous extracted features or characteristics and their respective labels through pattern recognition machine learning models.

This is a supervised approach since the algorithm needs to learn with a known set of labelled data in order to make predictions. Nevertheless, feature extraction and reduction is performed in an unsupervised manner.

Improvement in accuracy is found when the volume of training data is increased. Before training the model, the classification algorithm is projected and selected through exploration of prediction accuracy measurements. Once the selected algorithm is trained, test data is classified with the trained model and validation is performed to test the performance of our classifier and its capability of generalization.

Multiple models are available for classification of data such as decision trees, discriminant analysis, nearest neighbour classifiers, logistic regression classifiers or support vector machines among others. Since our data has two classes (STN and non-STN) and a set of features with no categorical values and due to its versatility through the use of different kernel functions, support vector machine (SVM) was chosen for the classification procedure as explained in next section, which uses a subset of points in the decision function, reducing memory usage, although further analysis are planned with different models as future work.

3.7.1 Support Vector Machine

Support vector machine classifies data into different classes by finding a hyperplane in an N-dimensional space (where N is the number of features) with the largest margin to separate both classes. Support vector are the data points that are closest to the hyperplane and influence its position and orientation, since the margin to maximize refers to the distance between the hyperplane and the support vectors (Cortes & Vapnik, 1995).

Data points from different classes may be separable linearly with the hyperplane, which is the SVM approach by definition. Nevertheless, if data is linearly non-separable, non-linear alternatives of SVM are available through the use of kernel functions. Data points are transformed to a higher dimension in order to separate classes using a hyperplane. Further information regarding SVM and kernels is available in literature (Hu, 2000; Cortes & Vapnik, 1995).

In the present study, 4 kernel functions are trained and tested along with the linear approach: quadratic, cubic, and 3 types of gaussian kernel (fine, medium and coarse), also known as radial basis function.

3.7.2 Cross validation

Standard approach for prediction of accuracy of the trained model is cross-validation and in particular k-fold cross validation is commonly used. This procedure consists on randomly splitting

our data in a number of folds (k), then for each fold, the model is trained using the out-of-fold observations and tested with the present fold. Evaluation of the classifier performance is measured through an estimated average test error with the results over all folds. Validation using 10-cross validation (CV) was performed through the Classification Learner App in Matlab, to explore algorithm's performance and visualize features in a fast and simple way.

Nevertheless, as explained in Little et al. (2017), the correct methodology for validation of a classifier with patient's observations is leave-one-subject out cross validation to ensure that features from patient for test has never seen the trained model. Leave-one-out cross validation is a specific case of k-fold validation where k is equal to the number of observations. Additionally, in our approach of leave-one-subject-out cross-validation, folds are not randomly chosen, but defined by patient identification. Therefore, for each iteration or validation, one patient is excluded from the training data and used for testing the trained model, as illustrated in figure 3.13. Since this approach is not available on the Classification Learner app, we trained, tested and validated our classifier through custom code for preparation of data, classification and error estimation.
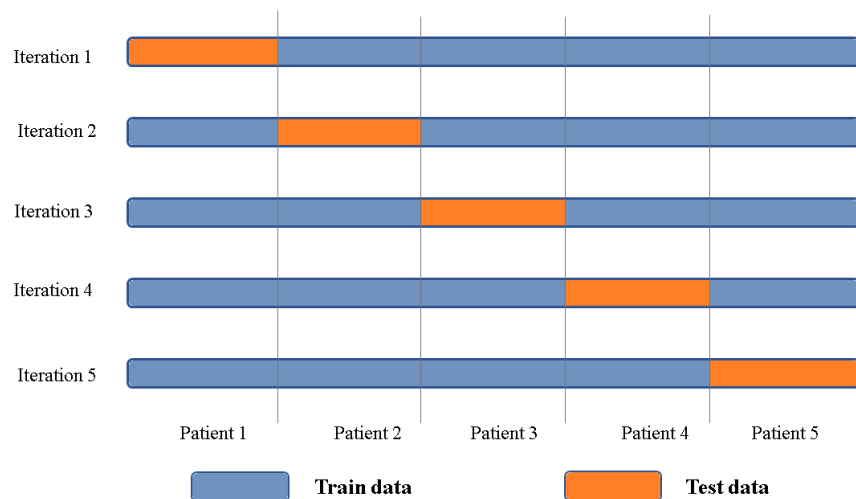
Figure 3.13: Cross-validation iterations for leave one subject out approach

3.7.3 Feature reduction

To reduce the dimension of the sample, feature reduction is performed through principal components analysis (PCA) combined with SVM as in previous studies (Guo et al., 2016; Sahak, Mansor, Lee, Yassin, & Zabidi, 2010). Original features are projected into a linear space of lower dimensionality, defined by the PCA principal components (PCs). Since this PCs are orthogonal to each other, redundant information and correlation from the original features is discarded.

3.7.4 Classification performance measurements

Evaluation of classification performance is assessed through four indicators: accuracy, specificity, sensitivity and precision. These indicators are calculated based on the prediction values from the confusion matrix for each model as indicated in figure 3.1, where true positives (TP) is the number of true positive predictions, false positive (FP) the incorrect positive prediction and true negatives (TN) and false negative (FN) are consequently the correct and incorrect negative predictions. Accuracy represents the proportion of the number of correct predictions, whereas sensitivity or recall and specificity represent the proportions of positives (STN) and negatives (not-STN) correctly classified, respectively. Additionally, precision is the portion of correctly classified as positives from all positives.

Table 3.1: Confusion matrix and its evaluation measures, where TP represents the number of true positives, FN as false negatives, FP is false positives and TN true negatives.

Confusion matrix		Target			
		Positive	**Negative**		
Model	**Positive**	TP	FN	**Sensitivity**	$TP/(TP+FN)$
	Negative	FP	TN	**Specificity**	$TN/(TN+FP)$

Precision

$TP/(TP+FP)$ **Accuracy** $(TP+TN)/(TP+FN+TN+FP)$

Chapter 4

Results

In the present chapter results are presented following the same order as the stated objectives, except for section 4.1 which first presents characteristics of our set of MER and their expert-based classification as STN or not.

4.1   MER dataset

Our set of MER is composed by 702 intraoperative MER during STN-DBS, from 5 patients, each with 2 hemispheres and 3 channels or electrodes recording simultaneously and in parallel through its trajectory, which changes depending on the patient as illustrated in table 4.1, between 21 to 27 depths for each hemisphere.

Table 4.1: Characteristics of subjects, recording duration and number of MER depths. UPDRS III represents the motor evaluation of Unified PD Rating Scale for each subject, scored through a structured neurological examination (range between 0 and 108). Gender indicated as female (F) and male (M).

Patient	Gender	Age at surgery (years)	Disease duration at surgery (years)	UPDRS III score	Duration (seconds)	Depths along trajectory to target	
						Right Hemisphere	Left Hemisphere
1	M	61	10	66	30	21	21
2	M	67	12	83	10	25	24
3	M	54	13	36	10	27	25
4	M	67	10	57	30	22	23
5	F	37	5	32	10	24	23

4.1.1   Expert labels and STN subdivision

Identification of MER localizations was performed based on visual inspection of the randomized dataset by an expert as explained in section 3.2.1. STN was identified in 285 recordings in

our dataset of 705 MER signals and detailed results and certainty of STN classification are illustrated in table 4.2. Number of STN vs. non-STN labels and certainty of MER classification as STN varies between patients, e.g. 38 recordings were labelled with the maximum certainty levels ($>95\%$) in patient 4, whereas in patient 3 no MER was identified as STN maximum certainty. This may be due to multiple reasons: differences in trajectories between patients and hemispheres, the electrode is actually not recording along the STN with the chosen trajectory or miss-classifications in the visual inspection labelling, which could be further explored with the gold-standard approach.

Table 4.2: Results of expert based classification of MER by visual inspection. Number and percentages of MER labelled as STN, not-STN and total per patient and for all patients. Number of recordings for each certainty level labelled as STN.

	Number of MER	non-STN Total	STN Total	Certainty $<25\%$	25-50%	50-75%	75-95%	$>95\%$
All patients	705	420 (59.6%)	285 (40.4%)	70	60	36	66	53
Patient 1	126 (17.9%)	58 (46.0%)	68 (54.0%)	10	14	16	19	9
Patient 2	147 (20.9%)	122 (83.0%)	25 (17.0%)	9	4	5	5	2
Patient 3	156 (22.1%)	113 (72.4%)	43 (27.6%)	14	25	3	1	0
Patient 4	135 (19.1%)	39 (28.9%)	96 (71.1%)	14	9	6	29	38
Patient 5	141 (20.0%)	88 (62.4%)	53 (37.6%)	23	8	6	12	4

The following results were computed considering STN-MER when the certainty level set is higher than 50%, and discarding signals with the lowest certainty levels. Therefore, 155 recordings are considered from the STN, which represents approximately 22% of the whole sample, almost 55% of the total recordings initially labelled as STN.

Regarding STN subdivisions (dorsal vs. ventral) manually defined on STN labelled for the results in section 4.5, we visualized these 155 STN-labelled recordings with maximum certainty and divide them in relation to its depth. Consequently 73 signals were considered as dorsal subthalamic nucleus (dSTN) and 74 as ventral subthalamic nucleus (vSTN), excluding 8 MER spatially located in the presumed frontier between dSTN and vSTN according to their depth.

4.2 Features related with time and frequency

Differences in extracted features were studied for distinguishing STN from non-STN MER from our dataset through our developed code.

To automatize the process of feature extraction with previous loading and preprocessing of MER, Matlab code was developed which can be generalizable for other types of recordings. Since our approach computes features from each signal based on small segments of the signal, we expected artifacts not to influence very significantly our results. Therefore, first analysis of extracted features for distinguishing STN and non-STN recordings using this tool was performed

calculating the median, mean and standard deviation of all frame values from 10 second segment filtered signal. Additionally, these results were extracted with features from 10 seconds segments free of artifacts. Following results for comparison of MER localization regarding to the STN were obtained through our custom code using parametric or non-parametric statistics, depending on the distribution of the sample as explained in section 3.6, and with the median values of all features and recordings.

4.2.1 Time domain features

Statistical analysis performed to find differences between STN and non-STN MER reported 8 time related features significantly different between STN/non-STN regions: MAV, ZC, RMS, VAR, WL, CF, TE and SDN. Only values of MAVS were not significant in our statistical analysis. Median values and interquartile range of features from 10 seconds filtered signal along with p-values for calculated features based on time domain are shown in table 4.3. Additionally, features values calculated from 10 seconds filtered segments free of artifacts are shown in table B.1, and significant features identified are the same using both datasets.

Table 4.3: Features based on time domain for (10 seconds) segments of filtered MER. Median results and interquartile ranges for time based features and p-values, obtained with T-test. Features with significant differences between STN and non-STN identified with *.

| Time related features | Filtered MER | | | | |
| | STN (n=155) | | non-STN (n=420) | | |
	mdn	*iqr*	*mdn*	*iqr*	p-value
MAV	10.311	4.396	4.084	1.216	<0.001*
MAVS	0.000	0.028	0.000	0.008	0.691
ZC	0.168	0.020	0.190	0.018	<0.001*
RMS	0.297	0.135	0.114	0.035	<0.001*
WL	9843.874	5115.080	4718.395	1206.946	<0.001*
VAR	180.719	171.272	26.573	16.878	<0.001*
CF	16.852	22.270	2.139	1.416	<0.001*
TE	68.537	82.086	13.718	7.741	<0.001*
SDN	13.227	5.105	5.566	1.556	<0.001*

4.2.2 Frequency domain features

A set of 16 features based on the frequency domain was extracted from 10 second segments of filtered raw MER. Both mFRQ, mdnFRQ and SPC show significant differences between STN and non-STN according to the statistics calculated. However, SF does not differ across both brain regions, along with some of the Mel Coefficients, although most of them present significant differences, as shown in table 4.4.

Table 4.4: Median and interquartile range results for features based on frequency domain. P-values obtained through T-test, with features showing significant differences between STN and non-STN indicated with *.

| Frequency related features | Filtered MER | | | | |
| | STN (n=155) | | non-STN (n=420) | | |
	mdn	*iqr*	*mdn*	*iqr*	p-value
SPC (Hz)	0.176	0.018	0.197	0.010	<0.001*
SF	0.631	0.129	0.627	0.139	0.256
mFRQ (Hz)	691.078	146.937	813.176	95.086	<0.001*
mdnFRQ (Hz)	113.000	35.750	141.000	33.000	<0.001*
MC1	18.764	0.888	16.838	0.613	<0.001*
MC2	62.416	3.470	56.207	1.884	<0.001*
MC3	-0.613	1.428	-1.654	0.738	<0.001*
MC4	-1.289	0.602	-0.575	0.598	<0.001*
MC5	-0.156	0.166	-0.120	0.161	<0.001*
MC6	-0.455	0.182	-0.261	0.125	<0.001*
MC7	-0.311	0.063	-0.289	0.073	<0.001*
MC8	-0.132	0.083	-0.054	0.084	0.142
MC9	-0.205	0.050	-0.183	0.056	0.008*
MC10	-0.034	0.052	0.005	0.056	<0.001*
MC11	-0.097	0.051	-0.091	0.054	0.069
MC12	0.007	0.061	0.019	0.049	0.460

Results from filtered clean segments after artifact detection were also statistically analyzed, as shown in table B.1. Features statistically significant for STN vs. non STN identification coincide, except for 3 MC coefficients. Differences across brain regions are statistically significant in MC7 and MC9 using the filtered signal but not when computing features from our set of segments free of artifacts. On the other hand, MC8 is significant with the clean MER although not significant with filtered recordings.

4.3   Neuronal activity MER analysis and related features

Comparison between STN and non-STN MER are presented in this section, based on its extracted features related with neuronal activity.

Our approach for spike analysis is based on existing algorithms for spike detection and sorting, as explained in section 3.5. We exhaustively inspected our results to refine and readjust input parameters or specific processes to fit our data set and extracted features. Regarding spike detection, we adapted input variables as recommend by the authors except for one factor. This factor trades off the sensibility of the detection, therefore detecting more spikes if it is increased. It represents an adapted rate between the cost of false alarms (probability of false detection) by the cost of omission (probability of not detecting spikes that truly are spiking events). As

implemented by the authors, this factor presents a range of ±0.2 and its recommended default value for unsupervised detection is 0. However, after comparison of different factor values, it was set to -0.15 to improve the detection in signals with low amplitude. In relation to spike sorting, it was found better clustering results by visual inspection of quality values and spike waveforms in our dataset when LDA subspace dimension was set to 3 instead of the default value (2).

Spike related features are then extracted from each sorted neuron and all results can be easily stored in tables for further analysis. Moreover, this tool is generalizable to other types of signals, adjusting a few input parameters such as sample frequency or duration among others.

4.3.1 Spike related features

Differences were found in the majority of the extracted features, except for the coefficient of variation of ISI as illustrated in table 4.5, using the filtered artifact free segments of 10 seconds segments from MER labelled with certainty level higher than 50%. Results are computed for each neuron. From 116 STN-MER were identified 258 neurons, whereas 329 neurons were found from 228 non-STN-MER. Results were computed from filtered-only MER, are shown in table B.2. Features identified as significantly different for STN identification are the same using both approaches although p-values and also feature values changed slightly.

Table 4.5: Spike related median results, interquartile range and p-values obtained through T-test from MER segments free of artifacts. Number of neurons in each sample represented as n. Features showing significant differences between STN and non-STN indicated with *.

| Spike related features | Filtered artifact-free MER | | | | |
| | STN (n=258) | | non-STN (n=329) | | |
	mdn	*iqr*	*mdn*	*iqr*	p-value
IS	0.240	0.331	0.051	0.160	<0.001*
FN	0.055	0.089	0.000	0.103	0.352
FP	0.422	0.458	0.125	0.600	<0.001*
mdnISI (ms)	35.344	31.792	114.938	175.396	<0.001*
stdISI (ms)	54.946	55.120	176.726	231.655	<0.001*
cvISI	0.990	0.149	0.970	0.186	0.113
FR (spks/s)	17.950	16.400	5.500	7.300	<0.001*
mdnSppA (uV)	79.029	34.133	26.922	9.478	<0.001*
AI	0.136	0.241	0.051	0.156	<0.001*
PI	0.615	0.810	2.773	5.083	<0.001*
PR	3.053	5.653	25.911	89.965	<0.001*
FF (ms)	1.637	0.578	1.189	0.266	<0.001*
BI	7.366	12.863	19.497	40.012	<0.001*
MBI (ms)	0.172	0.205	0.048	0.088	<0.001*
BR (burst/s)	1.500	1.300	0.400	0.700	<0.001*
mdnIBI (ms)	191.333	260.328	323.875	914.078	<0.001*
mdnBD (ms)	206.510	205.896	233.396	440.385	0.024*
mdnIBF (Hz)	$3.93 \ 10^{-5}$	$2.66 \ 10^{-5}$	$1.69 \ 10^{-5}$	$2.86 \ 10^{-5}$	<0.001*

Quality values were extracted regarding isolation score (IS) of the spike sorting and false

negatives (FN) and positives scores (FP), which quantify an estimation of the number of spikes missed and the number of spikes that are actually noise respectively. Higher values of IS were found in STN, as illustrated in figure 4.1, but more neurons are expected in STN than in non-STN regions and the majority of detected spikes outside STN present low values of isolation and high rate of false positives. In order to explore our features results improving the reliability of the spike sorting, minimum threshold of IS>0.5 was established to refine the rest of spike features, as shown in table B.3. However spike related features may be biased by this restriction in IS level specially on our dataset, since as illustrated bellow, points of features above 0.5 are mostly from the same patient and few MERs present higher values of IS.
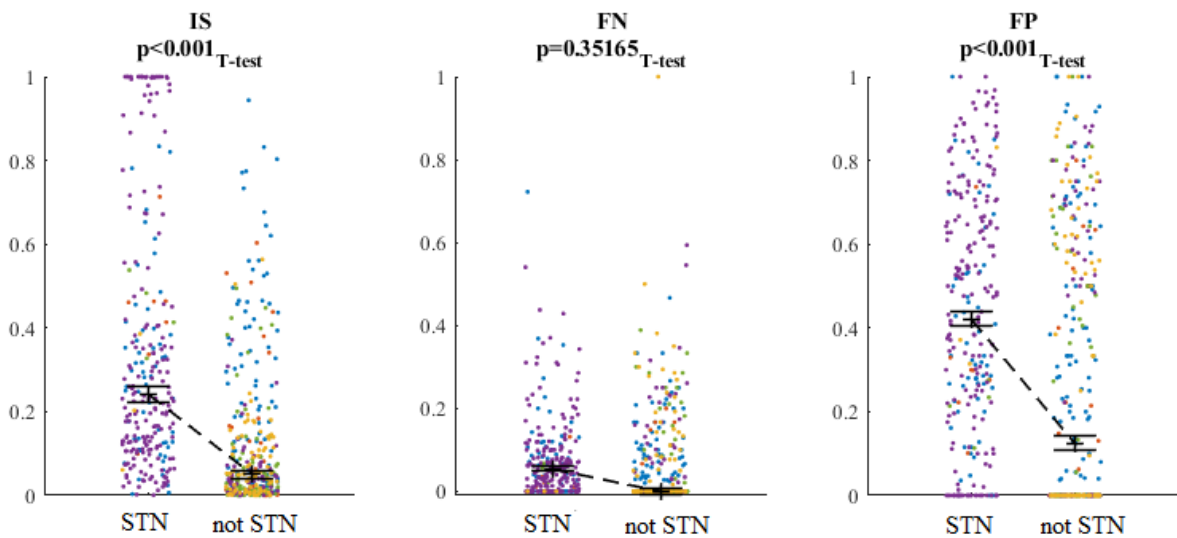
Figure 4.1: Quality measures obtained for all spikes in segments free of artifacts. Values of each sorted neuron are represented with a different color for each patient and median and standard errors are illustrated in black for both STN and non-STN samples.

4.4   Classification for STN identification

Classification was performed with Matlab using time and frequency related characteristics through our MER processing and expert-based labelling. Spike related features were not included in this classifier since extraction of these type of features is more time-consuming and less values are significantly different according to our statistic tests. Nevertheless, future analysis will be performed including both types of features to evaluate if they help the model and to differentiate STN subregions with gold standard labelling.

Performance measurements were extracted with six different SVM models using a set of extracted features related with time and frequency (25 median and 25 standard deviation). Classification using STN labelled with maximum certainty (>50 %) results in accuracy values

higher than 95% with 10-cross validation for all classification models except Fine Gaussian, as illustrated in table 4.6. Best performance of all implemented models is achieved with Linear SVM, presenting an accuracy of 96.7%.

Table 4.6: Classification performance results using all set of features (50) from STN labelled with maximum certainty (155) vs. non STN (420) and 10-cross validation to 6 different SVM models.

SVM model	10-CV using maximum certainty STN labels					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Confusion matrix	
Linear	96.70	90.97	98.81	96.58	141	14
					5	415
Quadratic	96.52	90.32	98.81	96.55	140	15
					5	415
Cubic	95.30	87.74	98.10	94.44	136	19
					8	412
Fine Gaussian	82.96	37.42	99.76	98.31	58	97
					1	419
Medium Gaussian	95.65	86.45	99.05	97.10	134	21
					4	416
Coarse Gaussian	95.13	84.52	99.05	97.04	131	24
					4	416

Classification was also implemented using all MER labelled by the expert as STN, as shown in table 4.7. As expected, overall performance measures decreased in comparison to the classification using refined STN labels. Nevertheless, obtained accuracy values are higher than 85% for all models except for Fine Gaussian, which again presents the worst classification results (79.72%).

Table 4.7: Classification performance results using all set of features (50) and STN labelled (285) vs. non STN (420) and 10-cross validation to 6 different SVM models.

SVM model	10-CV using all STN labels					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Confusion matrix	
Linear	89.79	81.05	95.71	92.77	231	54
					18	402
Quadratic	88.37	82.11	92.62	88.30	234	51
					31	389
Cubic	85.53	81.75	88.10	82.33	233	52
					50	370
Fine Gaussian	79.72	84.91	76.19	70.76	242	43
					100	320
Medium Gaussian	88.37	78.25	95.24	91.77	223	62
					20	400
Coarse Gaussian	86.38	71.23	96.67	93.55	203	82
					14	406

4.4.1 Feature reduction using PCA

Classification was performed also with feature reduction through PCA as explained in methods, which reduces dimensionality by projecting the data into its PCs. In order to explain 95% of variance, classification was performed keeping the first 2 components, which explain 94.5% and 4.9% of the total variance respectively for both approaches. Nevertheless, performance results were worst in general than without PCA for all models except for fine gaussian SVM, as illustrated in figure 4.8, which presented the worst accuracy result using all features. Accuracy values decreased specially for cubic SVM, from 95.3% using all features to 48.17%.

Table 4.8: Classification performance results using 2 PCs, STN labelled with maximum certainty (155) vs. non STN (420) and 10-cross validation to 6 different SVM models.

SVM model	10-CV using 2 PCs and max. certainty STN labels				
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Confusion matrix
Linear	95.65	90.32	97.62	93.33	140 15 10 410
Quadratic	92.35	78.06	97.62	92.37	121 34 10 410
Cubic	48.17	99.35	29.29	34.15	154 1 297 123
Fine Gaussian	96.00	89.68	98.33	95.21	139 16 7 413
Medium Gaussian	96.00	89.03	98.57	95.83	138 17 6 414
Coarse Gaussian	94.79	84.62	98.57	95.65	132 24 6 414

Classification was then performed adding principal components 3 and 4, which explained 0.4% and 0.1% of the total variance respectively. As shown in table 4.9, this classification approach improved accuracy results for some models in relation to the use of 2 PCs, such as quadratic and cubic SVM, or to the whole dataset (50 features), as fine gaussian.

Feature reduction was also performed to 2 and 4 PCs using all STN labels and both tables are available in appendix C. Although all values were accordingly lower in relation to the classification of refined STN, calculated performance measures presented similar behaviours. For instance, cubic SVM provides an accuracy of 36.7 %, presenting again the worst performance value using 2 PCs, as illustrated in C.1. Fine gaussian model also increased its accuracy using PCs in relation to the whole dataset and it presents the best performance measures using 2 PCs, same as the classification of refined STN. Additionally, best performance model using 4 PCs is the quadratic SVM for both all STN labels (table C.2) and maximum certainty.

Table 4.9: Classification performance results using 4 PCs, STN labelled with maximum certainty (155) vs. non STN (420) and 10-cross validation to 6 different SVM models.

SVM model	10-CV using 4 PCs and max. certainty STN labels					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Confusion matrix	
Linear	95.48	88.39	98.10	94.48	137	18
					8	412
Quadratic	96.00	90.97	97.86	94.00	141	14
					9	411
Cubic	95.30	89.68	97.38	92.67	139	16
					11	409
Fine Gaussian	94.61	94.19	94.76	86.90	146	9
					22	398
Medium Gaussian	95.48	88.39	98.10	94.48	137	18
					8	412
Coarse Gaussian	94.43	82.58	98.81	96.24	128	27
					5	415

4.4.2 Leave-one-subject out validation

As mentioned in section 3.7.2, evaluation of performance in classifiers with data from different subjects may be achieved through testing the model using data from each patient which wasn't used for training.

Therefore, cross validation through leave-one-subject out was performed using all features (50) in the 6 SVM models. All accuracy values varied in relation to 10-cross validation, but the best performances are still achieved using linear and quadratic SVM algorithm. As illustrated in table 4.10 and 4.11, accuracy results for each iteration presents quite different results due to inter-patient's variability, different number of STN-MER and also since the samples are not the same size, which might be a problem to obtain reliable results.

Table 4.10: Estimated accuracy percentage for leave-one-subject-out validation using STN maximum certainty.

SVM model	Accuracy of leave-one-subject-out CV (%)						
	Total		Iter. 1	Iter. 2	Iter. 3	Iter. 4	Iter. 5
	mean	std					
Linear	96.28	3.15	91.18	95.52	97.44	99.11	98.18
Quadratic	96.48	2.79	95.10	92.54	96.58	99.11	99.09
Cubic	93.81	2.97	93.14	91.04	97.44	91.07	96.36
Fine Gaussian	94.06	4.02	88.24	91.79	96.58	98.21	95.45
Medium Gaussian	94.98	4.28	89.22	91.79	96.58	98.21	99.09
Coarse Gaussian	91.56	9.21	75.49	94.78	96.58	98.21	92.73

Table 4.11: Estimated accuracy percentage for leave-one-subject-out validation using all STN-MER.

SVM model	Total mean	std	Iter. 1	Iter. 2	Iter. 3	Iter. 4	Iter. 5
	Accuracy of leave-one-subject-out CV (%)						
Linear	84.32	5.76	74.60	87.07	85.90	89.63	84.40
Quadratic	83.42	7.42	71.43	88.44	84.62	90.37	82.27
Cubic	69.28	28.17	79.37	90.48	89.10	22.22	65.25
Fine Gaussian	85.22	4.63	88.89	85.03	75.00	93.33	82.27
Medium Gaussian	85.06	7.05	89.68	85.03	75.00	93.33	82.27
Coarse Gaussian	80.34	7.84	73.02	85.71	72.44	90.37	80.14

4.5 Features for STN subdivision identification

Subterritories of STN were identified on MER expert-based labels as explained in the section 3.2.1 and therefore features extracted through the previously developed tools were compared to identify differences across both regions. Regarding time and frequency features, significant differences were not found, using 73 and 74 MER respectively as dorsal and ventral, (8 signals were considered as frontier or undefined due to its depth and consequently excluded from these analysis), as illustrated in table D.1. Statistics after artifact detection only in 10 seconds clean segments were also extracted and consequently the number of MER in the dorsal and ventral subdivision decreased to 51 and 54 respectively, but only one significant feature based on time and frequency domain was found.

As presented in table 4.12, spike related features from artifact free segments showed significant differences in 6 features in 113 neurons from 54 dorsal STN-MER vs. 129 from 51 MER considered as dorsal subdivision: mdnISI, stdISI, FR, PI, FF, MBI and BR. Also spike related features from filtered segments (without artifact extraction) are illustrated in D.2.

Random Permutation test on extracted features from both filtered and artifact-free filtered MER verified these significant features, as illustrated in appendix for spike related features in figures D.1 and D.2.

4.5.1 Refinement of MER labelling through images

More reliable definition of MER labels regarding STN subterritories is based on patient's anatomical images. Therefore this approach is explored in the present section using data from patient 4, since it presents the higher number of MER labelled as STN by the expert.

Labels of STN recordings based on the image approach are illustrated in table 4.13, which corresponds to the reconstruction illustrated in figure 4.2. Additionally, table 4.13 presents the certainty levels of STN expert-based identification showing the dorsal subdivision colored in grey, considering those not colored MER with certainty level equal or higher than 3 as ventral labelled.

Table 4.12: Statistic results of spike related features for STN subdivisions (based on expert labelling) in filtered free of artifact segments for STN subdivisions.

Spike related features	Artifact-free filtered MER				
	dorsal STN (n=113)		ventral STN (n=129)		
	mdn	*iqr*	*mdn*	*iqr*	p-value
IS	0.232	0.334	0.238	0.339	0.708
FN	0.058	0.089	0.048	0.102	0.397
FP	0.478	0.522	0.381	0.429	0.298
mdnISI (ms)	29.708	22.120	41.979	33.568	0.040*
stdISI (ms)	47.651	46.728	63.468	58.509	0.020*
cvISI	0.997	0.155	0.987	0.152	0.706
FR (spks/s)	21.400	16.850	15.300	14.325	0.013*
mdnSppA (uA)	77.661	32.762	79.095	34.160	0.543
AI	0.126	0.238	0.159	0.248	0.483
PI	0.447	0.597	0.786	0.970	0.037*
PR	2.170	3.972	3.983	6.432	0.151
FF (ms)	1.727	0.429	1.519	0.482	0.001*
BI	7.922	11.797	6.286	13.184	0.614
MBI (ms)	0.215	0.184	0.135	0.165	0.004*
BR (burst/s)	1.800	1.500	1.200	1.025	0.001*
mdnIBI (ms)	170.229	232.260	272.865	289.417	0.579
mdnBD (ms)	190.958	176.464	229.958	245.427	0.103
mdnIBF (Hz)	$4.32\ 10^{-5}$	$2.20\ 10^{-5}$	$3.34\ 10^{-5}$	$2.65\ 10^{-5}$	0.051
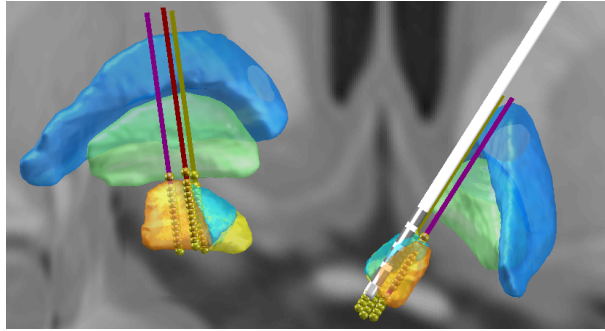
Figure 4.2: Trajectory of central, lateral and anterior leads through STN subterritories (motor in orange, associative in cyan and limbic in yellow). GPi and GPe also shown in green and blue respectively. Intraoperative MER locations represented by yellow spheres and definitive right reconstructed lead. Corresponding colors to each channel indicated in table 4.13.

The gold standard anatomical MER classification for this patient was compared to the STN identification by the expert through a confusion matrix in order to quantify its' accuracy in this region. Therefore, an accuracy of 86.81% was found using STN identified with certainty higher than 50%, as illustrated in table 4.14. On the other hand, using all STN-MER, the accuracy increased to 77.51% since some MER identified by the expert as STN actually they are not, as presented in 4.15. Nevertheless, its sensitivity increased to the 100% although its specificity was

Table 4.13: MER sites regarding STN subdivisions through image approach (gold standard) and expert certainty levels (in columns Cert.) for one patient. Grey cells represent MER positions labelled as motor or dorsal subterritories. Depths where the electrode is located on the frontier of this structures, both subterritories (or npn-STN if outside the STN) are indicated. Final implemented electrode's depth is presented in bold.

	Central		Lateral		Anterior	
	Gold Standard	*Cert.*	*Gold Standard*	*Cert.*	*Gold Standard*	*Cert.*
	Left hemisphere					
10	nSTN	0	nSTN	0	nSTN	3
11	nSTN	1	nSTN/motor	0	nSTN	3
12	nSTN	0	motor	2	nSTN/associative	5
13	nSTN/motor	5	motor	5	associative	5
14	nSTN/motor	4	motor	4	associative	4
15	nSTN/motor	5	motor	4	associative	5
16	nSTN/motor	4	motor	5	associative	5
17	nSTN/motor	4	motor	5	associative	4
18	nSTN/motor	5	motor	4	motor/asociative	5
19	motor	5	motor	5	motor/asociative	4
20	motor	4	motor	5	motor/asociative	4
21	motor	4	motor/nSTNmotor	5	motor/asociative	4
22	motor	4	motor/nSTNmotor	1	motor	5
23	nSTN	5	motor/nSTNmotor	5	motor/nSTN	5
	Right hemisphere					
8	nSTN	3	nSTN	4	nSTN	0
9	nSTN	2	nSTN	2	nSTN	2
10	nSTN	2	nSTN	5	nSTN	4
11	nSTN	4	nSTN	5	nSTN	4
12	nSTN	5	nSTN	5	associative/nSTN	4
13	associative	5	associative	5	associative	4
14	associative	4	motor/associative	5	associative	2
15	associative	5	motor	5	associative	4
16	associative	5	motor	5	limbic	4
17	associative	5	motor	5	limbic	5
18	motor	5	motor	5	limbic	4
19	motor	5	motor	4	limbic	5
20	motor	3	motor	4	limbic	4
21	motor	5	motor	3	limbic	4
22	motor	4	motor	2	nSTN	3

reduced to 25%.

Table 4.14: Confusion matrix and evaluation measures for STN identification based on expert labels with certainties higher than 50% vs. on lead reconstruction with patient's anatomy, as gold standard approach.

		Expert based			
		STN	non-STN		
Gold standard	STN	62	3	**Sensitivity**	*95.59 %*
	non-STN	9	17	**Specificity**	*65.38 %*
		Precision			
		87.32 %		**Accuracy**	*86.81 %*

Table 4.15: Confusion matrix and evaluation measures for STN identification based on expert labels (all certainties) vs. on lead reconstruction with patient's anatomy, as gold standard approach.

		Expert based			
		STN	non-STN		
Gold standard	STN	65	0	**Sensitivity**	*100 %*
	non-STN	21	7	**Specificity**	*25 %*
		Precision			
		75.58 %		**Accuracy**	*77.51 %*

On the other hand, regarding STN subdivisions according to the image based approach, 38 MER were labelled as motor subdivisions whereas 21 from the non-motor part, since depths located in frontiers between any STN subdivision and non-STN were considered from the consequent subregion, but boundaries between motor and non-motor subdivisions were excluded from analysis. Identification of MER sites using subdivisions of the expert STN identification in relation to the gold standard approach differed, since some dorsal MER were actually not located on the motor part of the STN and vice versa.

Features values were compared for motor and non-motor regions based on the image approach using time, frequency and spike related features, and all results are available in tables D.3 and D.4. Although some features were significantly different, obtained results were not consistent with results in table 4.12 neither the literature, which may be due to the small size of the sample and only from one patient.

Regarding features of time and frequency, SF was significantly different only from filtered 10 seconds segment whereas CF and MC10 were the only significant features across the defined subdivisions of STN. In relation to spike related features, surprisingly, no significantly different features were found both in filtered and artifact free signals except for FP.

Chapter 5

Discussion

5.1 Developed tools for MER processing and feature extraction

As intended for this project, tools for MER processing were developed to analyze and extract characteristics from raw MER related with time and frequency. As mentioned in results, the algorithm for time and frequency extraction is quite robust to noise due to extraction of features through statistics of smaller frames from the signal.

Multiple tools for spike sorting exist such as Wave Clus 3 (Quiroga, Nadasdy, & Ben-Shaul, 2004), which is a widely known hybrid approach which requires temperature adjustment after the automatic sorting was performed, or Mountain Sort (Chung et al., 2017), a recent approach fully automatic which might be a better alternative to our spike sorting, but it is implemented in Linux and therefore it difficulties integration with Matlab for feature extraction. As explained before, our proposed approach combines existing algorithms for spike detection and sorting in order to extract action potentials from each MER and then export its related features all in an unsupervised way in order to be able to analyze a big number of MER.

Different approaches for spike sorting provide heterogeneous results depending on the spike identification algorithms, and classification methods as it is illustrated in figure 5.1, which shows the same detected spikes using CWT with our approach vs. Wave Clus 3 (Quiroga et al., 2004). In this particular case, sorted neurons are different using both methods and the supervised approach was unable to correctly sort as many spikes in unsupervised clustering, resulting in incorrectly classified outliers since supervision is performed through limited temperature adjustment.

Additionally, extracted measures of sorting quality provided quite poor results in terms of isolation of the sorting, false positives and negatives. This may be due to over-clustering, which was observed in some cases using our unsupervised sorting or due to the fact that spike detection using CWT detects spikes with amplitudes lower than the background noise level . Also in some cases, falsely detected spikes that are actually noise are sometimes sorted as neurons, which
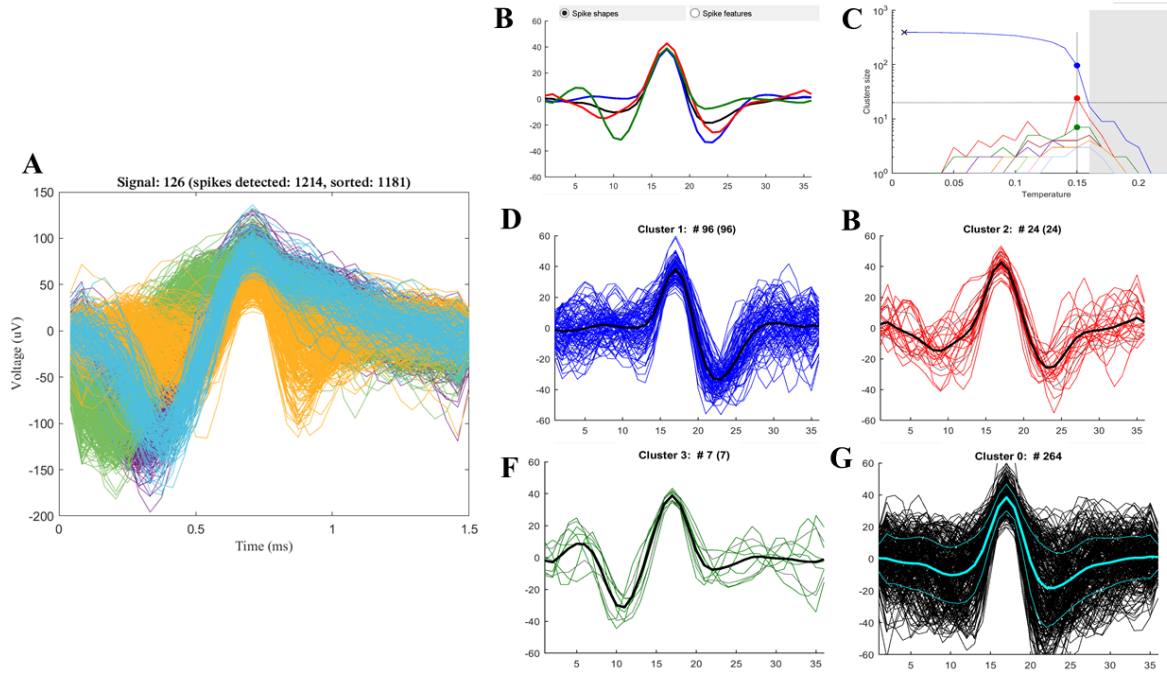
Figure 5.1: Spikes waveforms with cluster identification through different color with unsupervised sorting. (A) Clusters in same signal extracted through supervised sorting (Wave Clus) (B shows mean waveforms from D, E, F, G) and temperature graph for supervision (right). Detected 264 outliers with Wave Clus (vs. 33 in unsupervised approach) (G)

decreases the IS.

5.2 Features for STN identification

Time and frequency analysis showed significant differences between expert-labelled STN and non-STN. Significant differences in features identified using artifact-free filtered MER were similar as using only filtered MER except for some MC, which represent contribution to different frequency bands and therefore are more susceptible to artifacts.

Regarding features related with neuronal activity, significant differences were found across neurons detected in STN and non STN regions. Since quality values were quite low, results for neurons with a minimum threshold value of 0.5 were extracted, resulting in a lower number of significant features. Nevertheless for both approaches, features regarding bursting activity such as median burst index (BI), modified burst index (MBI) and burst rate (BR) are significantly higher for neurons labelled as STN. Also median spike amplitude (mdnSppA) showed significantly higher values for STN regions along with the Fano Factor, on the contrary to the assimetric index (AI) showing higher values for non-STN regions. This is in line with previous results collected from STN recordings in PD.

Although extracted features were used to compare STN and non-STN regions, our developed

tool could be used to compare other brain regions.

5.3   Classification of STN-MER

Classification results of STN regions evaluated in 6 SVM models using all features and expert labelling presented quite good results in most of the trained and tested models. As expected, when using STN labelled with the maximum certainty and using all features, the results were particularly good. Through PCA, feature dimensionality was reduced to 2 and 4 principal components and although this approach improved performance results in some models, best performance was achieved with the linear model without feature reduction, using all features.

Lastly, classification accuracy was obtained using leave one subject out validation. Overall results decreased, as expected, and accuracy for the different iterations highly varies, but mean results over 90 % were obtained with all models using maximum certainty STN.

5.4   Features for motor STN identification

Differences across STN subdivisions (dorsal and ventral) were not discernible based on time and frequency domain (only on one spectral centroid) but on features related with action potential. Features extracted from 10 seconds artifact-free segments presented statistically significant differences in 6 features related with the firing rate and bursting activity, although only mdnBD was significant using filtered-only segments.

Obtained results using clean segments are consistent with the previous literature since according to Lourens et al. (2013), higher burst activity and firing rate was found in the ventral subdivisions of STN. Measures directly related with ISI showed no significant differences as previously reported in Seifried et al. (2012).

Comparison of STN subregions identification based on gold standard approach (lead reconstruction) vs. expert subdivision through the confusion matrix showed that the classification through subdivisions of expert labelled STN is not actually accurate, since superpositions of the different STN subterritories are not taken into account.

Regarding features obtained using the gold standard labelling, using only one subject, no significant differences were identified. This may be due to the small proof-of-concept analysis. Further analysis using more patients is aimed in order to contrast the results obtained with the expert-based labels.

Chapter 6

Conclusion

In summary, differences between STN and non-STN regions have been found based on features from PD-DBS intraoperative recordings. Features extraction was performed through unsupervised MER analysis tools on Matlab and time, frequency and also neuronal activity features were calculated. This tool is easily generalizables for other types of recordings or brain regions and all developed code is open-source[1].

A classification model for STN-MER was developed using time and frequency related features. This classifier could support clinicians in distinguishing STN from non-STN. Although different models were studied, one of the best performances was achieved with linear SVM classifier, which presents a mean accuracy of 96.28 ± 3.15 %, validated with leave-one-subject-out approach and using STN labelled with certainty level higher than 50%.

Differences on the extracted features based on neuronal activity were found comparing STN-MER with recordings outside the STN area, but these features were not included in the classifier.

Furthermore, since the area within the STN with major clinical interest for STN-DBS is the motor subdivision, features were studied using the developed tools for MER analysis. Preliminary analysis have found differences between the probably motor functional subdivision of the STN from the associative and limbic subregions, consistent with the previous literature.

Nevertheless, these results may be biased by errors in expert based labels or the reduced size of the sample and may alter classification and analysis results, since our database is composed by 5 patients with high inter-patient-variability.

---

[1]Available in online repository: `https://bitbucket.org/sarafernaandez/thesis_development-of-pipeline-for-mer-analysis`

6.1 Future work

Different lines remain open for future work in this study. This includes: image and signal processing and classification with machine learning techniques.

More accurate labelling of the different MER locations in relation to the STN and subterritories (motor, limbic and associative) can be achieved through electrode reconstruction based on patient's imagiology. The methodology described in section 3.2.2 could be scaled up to more patients and segregated into a study of the three subregions. This gold-standard labels could be used to refine the classification classes, which may improve the performance of the algorithms.

Additionally, ongoing work in the classification field is using SVM recursive feature selection (Lin et al., 2018) to decrease the number of features used in the classification and improving its performance. Moreover, other classification models such as deep neural networks, regression trees, naive bayes or k-means classifiers could be explored for our features and labels.

References

Aboy, M., & Falkenberg, J. H. (2006). An automatic algorithm for stationary segmentation of extracellular microelectrode recordings. *Med. Biol. Eng. Comput.*, *44*(6), 511–515. doi: 10.1007/s11517-006-0052-2

Accolla, E. A., Dukart, J., Helms, G., Weiskopf, N., Kherif, F., Lutti, A., ... Draganski, B. (2014). Brain tissue properties differentiate between motor and limbic basal ganglia circuits. *Hum. Brain Mapp.*, *35*(10), 5083–5092. doi: 10.1002/hbm.22533

Bakštein, E., Schneider, J., Sieger, T., Novak, D., Wild, J., & Jech, R. (2015). Supervised segmentation of microelectrode recording artifacts using power spectral density. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, *2015-Novem*, 1524–1527. doi: 10 .1109/EMBC.2015.7318661

Bakštein, E., Sieger, T., Wild, J., Novák, D., Schneider, J., Vostatek, P., ... Jech, R. (2017). Methods for automatic detection of artifacts in microelectrode recordings. *J. Neurosci. Methods*, *290*, 39–51. doi: 10.1016/j.jneumeth.2017.07.012

Bartels, A. L., & Leenders, K. L. (2009). Parkinson's disease: The syndrome, the pathogenesis and pathophysiology. *Cortex*, *45*(8), 915 - 921. doi: 10.1016/j.cortex.2008.11.010

Cagnan, H., Dolan, K., He, X., Contarino, M. F., Schuurman, R., Van Den Munckhof, P., ... Martens, H. C. (2011). Automatic subthalamic nucleus detection from microelectrode recordings based on noise level and neuronal activity. *J. Neural Eng.*, *8*(4). doi: 10.1088/ 1741-2560/8/4/046006

Castrioto, A., Lhommée, E., Moro, E., & Krack, P. (2014). Mood and behavioural effects of subthalamic stimulation in Parkinson's disease. *Lancet Neurol.*, *13*(3), 287–305. doi: 10.1016/S1474-4422(13)70294-1

Chung, J. E., Magland, J. F., Barnett, A. H., Tolosa, V. M., Tooker, A. C., Lee, K. Y., ... Greengard, L. F. (2017). A Fully Automated Approach to Spike Sorting. *Neuron*, *95*(6),

1381–1394.e6. doi: 10.1016/j.neuron.2017.08.030

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi: 10.1007/BF00994018

Dolan, K., Martens, H. C., Schuurman, P. R., & Bour, L. J. (2009). Automatic noise-level detection for extra-cellular micro-electrode recordings. *Med. Biol. Eng. Comput.*, *47*(7), 791–800. doi: 10.1007/s11517-009-0494-4

Fahn, S., Jankovic, J., & Hallett, M. (2011). Functional neuroanatomy of the basal ganglia. *Princ. Pract. Mov. Disord.*, 55–65. doi: 10.1016/B978-1-4377-2369-4.00003-2

Friston, K., Ashburner, J., Kiebel, S., Nichols, T., & Penny, W. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press. Retrieved from `http://store.elsevier.com/product.jsp?isbn=9780123725608`

Gelb, D., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for parkinson's disease. *Archives of Neurology*, *56*(1), 33-39. doi: 10.1001/archneur.56.1.33

Gourévitch, B., & Eggermont, J. J. (2007). A nonparametric approach for detection of bursts in spike trains. *J. Neurosci. Methods*, *160*(2), 349–358. doi: 10.1016/j.jneumeth.2006.09.024

Gradinaru, V., Mogri, M., Thompson, K. R., Henderson, J. M., & Deisseroth, K. (2009). Optical deconstruction of parkinsonian neural circuitry. *Science*, *324*(5925), 354–359. doi: 10.1126/science.1167093

Grubhoffer, B. T. (2016). *Automatic artifact detection in micro-EEG signals* (Unpublished master's thesis). Czech Technical University in Prague.

Guo, H., Huang, Y., Lin, C., Chien, J., Haraikawa, K., & Shieh, J. (2016, Oct). Heart rate variability signal features for emotion recognition by using principal component analysis and support vectors machine. In *2016 IEEE 16th International Conference on Bioinformatics and Bioengineering* (p. 274-277). doi: 10.1109/BIBE.2016.40

Hariz, M. (2017). My 25 stimulating years with DBS in Parkinson's disease. *J. Parkinsons. Dis.*, *7*, S33–S41. doi: 10.3233/JPD-179007

He, X. (2009). *Neural Signal Processing of Microelectrode Recordings for Deep Brain Stimulation* (Unpublished master's thesis). Chalmers University of Technology, Sweden.

Horn, A., & Kühn, A. A. (2015). Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *Neuroimage*, *107*, 127–135. doi: 10.1016/j.neuroimage.2014.12.002

Hu, Y. H. (2000). *Handbook of neural network signal processing* (1st ed.; J.-N. Hwang & J.-N. Hwang, Eds.). Boca Raton, FL, USA: CRC Press, Inc.

Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry*, *55*(3), 181–184. doi: 10.1136/jnnp.55.3.181

Husch, A., Petersen, M. V., Gemmar, P., Goncalves, J., & Hertel, F. (2018). PaCER - A fully automated method for electrode trajectory and contact reconstruction in deep brain stimulation. *NeuroImage Clin.*, *17*, 80–89. doi: 10.1016/j.nicl.2017.10.004

Hutchison, W. D., Allan, R. J., Opitz, H., Levy, S. R., Dostrovsky, J., Lang, a. E., & Lozano, a. M. (1998). Neurophysiological Identification of the Subthalamic Nucleus in Surgery for Parkinson's Disease. *Ann. Neurol.*, 622–628.

Johnsen, E. L., Sunde, N., Mogensen, P. H., & Østergaard, K. (2010). MRI verified STN stimulation site - Gait improvement and clinical outcome. *Eur. J. Neurol.*, *17*(5), 746–753. doi: 10.1111/j.1468-1331.2010.02962.x

Joshua, M., Elias, S., Levine, O., & Bergman, H. (2007). Quantifying the isolation quality of extracellularly recorded action potentials. *J. Neurosci. Methods*, *163*(2), 267–282. doi: 10.1016/j.jneumeth.2007.03.012

Keshtkaran, M. R., & Yang, Z. (2017). Noise-robust unsupervised spike sorting based on discriminative subspace learning with outlier handling. *J. Neural Eng.*, *14*(3), 036003. doi: 10.1088/1741-2552/aa6089

Kuebler, E. S., & Thivierge, J.-P. (2014). Spiking variability: Theory, measures and implementation in matlab. *The Quantitative Methods for Psychology*, *10*(2), 131-142. doi: 10.20982/tqmp.10.2.p131

Larobina, M., & Murino, L. (2014). Medical image file formats. *J. Digit. Imaging*, *27*(2), 200–206. doi: 10.1007/s10278-013-9657-9

Lee, S., Kim, J., & Lee, I. (2012). Speech/audio signal classification using spectral flux pattern recognition. In *2012 ieee workshop on signal processing systems* (p. 232-236). doi: 10.1109/SiPS.2012.36

Lin, X., Li, C., Zhang, Y., Su, B., Fan, M., & Wei, H. (2018). Selecting feature subsets based on SVM-RFE and the overlapping ratio with applications in bioinformatics. *Molecules*, *23*(1). doi: 10.3390/molecules23010052

Little, M. A., Varoquaux, G., Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., & Kording, K. P. (2017). Using and understanding cross-validation strategies. Perspectives on Saeb et al. *Gigascience*, *6*(5), 1–6. doi: 10.1093/gigascience/gix020

Lourens, M., Meijer, H., Contarino, M., van den Munckhof, P., Schuurman, P., van Gils, S., & Bour, L. (2013). Functional neuronal activity and connectivity within the subthalamic nucleus in Parkinson's disease. *Clin. Neurophysiol.*, *124*(5), 967–981. doi: 10.1016/j.clinph .2012.10.018

Maiti, P., Manna, J., & Dunbar, G. L. (2017). Current understanding of the molecular mechanisms in Parkinson's disease: Targets for potential treatments. *Transl. Neurodegener.*, *6*(1), 28. doi: 10.1186/s40035-017-0099-z

Medtronic. (2007). Mri guidelines for medtronic deep brain stimulation systems. *English*. doi: 10.1016/B978-0-12-374976-5.00014-1

Negida, A., Elminawy, M., El Ashal, G., Essam, A., Eysa, A., & Abd Elalem Aziz, M. (2018). Subthalamic and Pallidal Deep Brain Stimulation for Parkinson's Disease. *Cureus*, *10*(2). doi: 10.7759/cureus.2232

Nenadic, Z., & Burdick, J. W. (2005). Spike detection using the continuous wavelet transform. *IEEE Transactions on Biomedical Engineering*, *52*(1), 74-87. doi: 10.1109/TBME.2004 .839800

On, C. K., Pandiyan, P. M., Yaacob, S., & Saudi, A. (2006, June). Mel-frequency cepstral coefficient analysis in speech recognition. In *2006 international conference on computing informatics* (p. 1-5). doi: 10.1109/ICOCI.2006.5276486

O'Shea, D. J., & Shenoy, K. V. (2017). ERAASR: An algorithm for removing electrical stimulation artifacts from multielectrode array recordings. *bioRxiv*. doi: 10.1101/185850

Oskoei, M. A., & Hu, H. (2006, Dec). Ga-based feature subset selection for myoelectric classification. In *2006 ieee international conference on robotics and biomimetics* (p. 1465-1470). doi: 10.1109/ROBIO.2006.340145

Quiroga, R. Q. (2009). What is the real shape of extracellular spikes? *J. Neurosci. Methods*, *177*(1), 194–198. doi: 10.1016/j.jneumeth.2008.09.033

Quiroga, R. Q., Nadasdy, Z., & Ben-Shaul, Y. (2004). Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering. *Neural Comput.*, *16*(8), 1661–1687. doi: 10.1162/089976604774201631

Rajpurohit, V., Danish, S. F., Hargreaves, E. L., & Wong, S. (2015). Optimizing computational feature sets for subthalamic nucleus localization in DBS surgery with feature selection. *Clin. Neurophysiol.*, *126*(5), 975–982. doi: 10.1016/j.clinph.2014.05.039

Sahak, R., Mansor, W., Lee, Y. K., Yassin, A. I. M., & Zabidi, A. (2010). Performance of Combined Support Vector Machine and Principal Component Analysis in recognizing infant cry with asphyxia. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (p. 6292-6295). doi: 10.1109/IEMBS.2010.5628084

Schönecker, T., Kupsch, A., Kühn, A., Schneider, G.-H., & Hoffmann, K.-T. (2009). Automated optimization of subcortical cerebral mr imaging-atlas coregistration for improved postoperative electrode localization in deep brain stimulation. *American Journal of Neuroradiology*, *30*(10), 1914–1921. doi: 10.3174/ajnr.A1741

Seifried, C., Weise, L., Hartmann, R., Gasser, T., Baudrexel, S., Szelényi, A., . . . Hilker, R. (2012). Intraoperative microelectrode recording for the delineation of subthalamic nucleus topography in Parkinson's disease. *Brain Stimul.*, *5*(3), 378–387. doi: 10.1016/j.brs.2011.06.002

Starr, P. A. (2002). Placement of deep brain stimulators into the subthalamic nucleus or globus pallidus internus: Technical approach. *Stereotact. Funct. Neurosurg.*, *79*(3-4), 118–145. doi: 10.1159/000070828

Tysnes, O. B., & Storstein, A. (2017). Epidemiology of Parkinson's disease. *J. Neural Transm.*, *124*(8), 901–905. doi: 10.1007/s00702-017-1686-y

Wagle Shukla, A., & Okun, M. S. (2014). Surgical Treatment of Parkinson's Disease: Patients, Targets, Devices, and Approaches. *Neurotherapeutics*, *11*(1), 47–59. doi: 10.1007/s13311-013-0235-0

Wodarg, F., Herzog, J., Reese, R., Falk, D., Pinsker, M. O., Steigerwald, F., . . . Volkmann, J. (2012). Stimulation site within the MRI-defined STN predicts postoperative motor outcome. *Mov. Disord.*, *27*(7), 874–879. doi: 10.1002/mds.25006

Wooten, G. F., Currie, L. J., Bovbjerg, V. E., Lee, J. K., & Patrie, J. (2004). Are men at greater risk for parkinson's disease than women? *J. Neurol. Neurosurg. Psychiatry*, *75*(4), 637–639. doi: 10.1136/jnnp.2003.020982

Appendix A

Definition of features

A.1 Time related features

1. **Mean absolute value (MAV)**

$$MAV = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{A.1}$$

   where N is the number of samples in data x.

2. **Mean absolute value slope (MAVS)**

$$MAVS = MAV_{i+1} - MAV_i \tag{A.2}$$

3. **Zero crossings (ZC)**

   It quantifies the number of times that the signal crosses the horizontal axis, changing from positive amplitudes to negatives or vice versa.

4. **Root mean square (RMS)**

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2} \tag{A.3}$$

   where N is the number of samples in data x.

5. **Waveform of curve length (WL)**

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{A.4}$$

where N is the number of samples in data x.

6. **Variance (VAR)**

$$VAR = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{A.5}$$

where N is the number of samples in data x.

7. **Crest factor (CF)**

$$CF = \frac{1}{2}\frac{x_{max} - x_{min}}{RMS} \tag{A.6}$$

where RMS is the root mean square of signal x.

8. **Teager energy (TE)**

According to Rajpurohit et al. (2015)

$$TE = \frac{1}{(N-2)}\sum_{i=2}^{N-1} x_i^2 - x_{i-1}x_{i+1} \tag{A.7}$$

where x_i are each components of the data vector with length N.

9. **Standard deviation of noise (SDN)**

Estimation of noise is performed as explained in Dolan et al. (2009), as the scale parameter of the Hilbert transform of the segment.

A.2 Frequency related features

1. **Spectral centroid (SPC)**

$$SPC = \frac{\sum_{k=0}^{N-1} kx(k)}{\sum_{k=0}^{N-1} x(k)} \tag{A.8}$$

where x(k) is the amplitude corresponding to the bin k in Discrete Fourier Transform spectrum.

2. **Spectral flux (SF)**

According to Lee, Kim, and Lee (2012), spectral flux measures the change in the spectrum between two adjacent frames and it is obtained as

$$SF = \sum_{m=0}^{N} [|\hat{X}[m]| - |\hat{X}_{preFrame}[m]|]^2 \tag{A.9}$$

where X[m] denotes the results of the Fast Fourier Transform (FFT) for the input signal x[n]; N is the frame length, $\hat{X}[m] = \frac{X[m]}{argmax[X[m]]}$ and $\hat{X}_{preFrame}$ is the spectral flux from the previous frame.

3. **Mean frequency (mFRQ)**

$$mFRQ = \frac{\sum_{i=1}^{N} f_i PSD_i}{\sum_{i=1}^{N} PSD_i} \tag{A.10}$$

where PSD_i is the ith line of the Power Spectral Density, calculated as the discrete Fourier Transform, as defined in Oskoei and Hu (2006).

4. **Median frequency (mdnFRQ)**

$$mdnFRQ = \sum_{i=1} f_{med} PSD_i = \sum_{i=f_{med}}^{N} PSD_i = \frac{1}{2} \sum_{i=1}^{N} NPSD_i \tag{A.11}$$

according to Oskoei and Hu (2006), where PSD_i is the ith line of the Power Spectral Density, calculated as the discrete Fourier Transform.

5. **Mel Cepstrum coefficients (MC)**

Coefficients derived from the mel-frequency cepstrum, commonly used in speech signal processing as described in On, Pandiyan, Yaacob, and Saudi (2006). The power of spectrum (of the Fast Fourier Transform) is obtained from the framed signal onto the Mel scale. Then, the Mel spectrum is extracted and lastly the Mel cepstrum is obtained by taking the logarithms of the powers of the spectrum at each mel frequency and the Discrete Cosine Transform of the mel logarithmic powers for each cepstral coefficient (1 to 12).

A.3 Spike related features

1. **Inter spike interval (ISI)**

It represents the time (in ms) between subsequent spikes (action potentials) of a neuron. Since it is a vector, statistical measures extracted are:

- **Median Inter Spike Interval (mdnISI)**
- **Standard deviation of ISI (stdISI)**
- **Coefficient of variation of ISI (cvISI)**: ratio between standard deviation and mean of ISI.

2. **Firing rate (FR)**

$$FR = \frac{n_{spk}}{T} \tag{A.12}$$

where n_{burst} is the number of spikes in interval T (10 seconds).

3. **Median spike amplitude peak to peak (mdnSppA)**

Amplitude (in μV) for each spike segment from lowest peak (minimum value) to the highest peak (maximum value) is obtained. Since it results in a vector, its median value of all the spike segments for each neuron is stored.

4. **Asymmetry index (AI)**

$$AI = \frac{mode(ISI)}{mean(ISI)} \tag{A.13}$$

where ISI is the Inter spike interval.

5. **Pause index (PI)**

$$PI = \frac{\#(ISI > 50ms)}{\#(ISI < 50ms)} \tag{A.14}$$

where $\#(ISI < 50ms)$ and $\#(ISI > 50ms)$ are the number of ISI with shorter and longer than 10 ms of duration respectively.

6. **Pause ratio (PR)**

$$PR = \frac{\sum(ISI > 50ms)}{\sum(ISI < 50ms)} \tag{A.15}$$

where $\sum(ISI < 50ms)$ and $\sum(ISI > 50ms)$ are the cumulative sum of ISI with shorter and longer than 50 ms of duration respectively.

7. **Fano factor (FF)**

This factor presents a measure of the variability in spike rate distributions, as defined in Kuebler and Thivierge (2014):

$$FF = \frac{\sqrt{\frac{1}{J-1} \sum_{j=1}^{J} (R_j - \overline{R})^2}}{\overline{R}} \tag{A.16}$$

where J is the total number of spike rates (R_j) and $\overline{R}$ is the mean spike rate.

8. **Burst index (BI)**

According to Hutchison et al. (1998), burst index is obtained as:

$$BI = \frac{mean(ISI)}{mode(ISI)} \tag{A.17}$$

where ISI is the inter spike interval and the cell is considered bursty if $BI > 10$.

9. **Modified burst index (MBI)**

$$MBI = \frac{\#(ISI < 10ms)}{\#(ISI < 10ms)} \tag{A.18}$$

where $\#(ISI < 10ms)$ and $\#(ISI > 10ms)$ are the number of inter spike intervals with duration shorter and longer than 10 ms respectively.

10. **Burst rate (BR)**

$$BR = \frac{n_{burst}}{T} \tag{A.19}$$

where n_{burst} is the number of bursts in interval T (10 seconds).

11. **Median inter burst interval (mdnIBI)**

It represents the median value of the vector with time intervals between the end of one burst and the beginning of the next consecutive burst.

12. **Median burst duration (mdnBD)**

It is obtained as the median value of all bursts durations (ms) for each sorted neuron.

13. **Median intra-burst frequency (mdnIBF)**

It quantifies the median number of spikes per second inside each burst detected for each spike train.

Appendix B

Features for STN identification.

Table B.1: Features based on time and frequency domain for (10 seconds) segments of free-of artifact filtered MER. Median results and interquartile ranges for each feature and p-values, obtained with T-test. Features with significant differences between STN and non-STN in bold.

Time and frequency related features	Free of artifact filtered MER				
	STN (n=116)		non-STN (n=228)		
	mdn	*iqr*	*mdn*	*iqr*	p-value
MAV	11.179	4.649	4.259	1.244	<0.001*
MAVS	0.000	0.021	0.000	0.007	0.589
ZC	0.170	0.016	0.191	0.017	<0.001*
RMS	0.328	0.156	0.118	0.036	<0.001*
WL	1.20E+04	5.18E+03	5.01E+03	1.44E+03	<0.001*
VAR	220.917	205.780	28.747	17.197	<0.001*
CF	19.894	27.290	2.236	1.443	<0.001*
TE	97.161	96.123	15.415	9.184	<0.001*
SDN	14.359	5.505	5.817	1.589	<0.001*
SPC (Hz)	0.180	0.017	0.198	0.010	<0.001*
SF	0.661	0.097	0.652	0.102	0.405
mFRQ (Hz)	720.299	130.738	823.057	86.831	<0.001*
mdnFRQ (Hz)	124.500	31.000	145.000	36.000	<0.001*
MC1	18.925	0.988	16.918	0.614	<0.001*
MC2	63.371	3.725	56.351	1.807	<0.001*
MC3	-0.913	1.285	-1.772	0.747	<0.001*
MC4	-1.469	0.470	-0.752	0.581	<0.001*
MC5	-0.157	0.147	-0.153	0.152	0.374
MC6	-0.492	0.151	-0.290	0.101	<0.001*
MC7	-0.315	0.071	-0.306	0.058	0.001*
MC8	-0.139	0.070	-0.070	0.073	0.030*
MC9	-0.206	0.040	-0.191	0.051	0.294
MC10	-0.046	0.047	-0.003	0.043	<0.001*
MC11	-0.095	0.041	-0.093	0.049	0.846
MC12	0.006	0.046	0.020	0.054	0.582

Table B.2: Statistical results of features based on action potentials in filtered MER to compare STN and non STN regions. Number of neurons for each sample represented by n.

Spike related features	Filtered MER				
	STN (n=357)		non-STN (n=521)		
	mdn	iqr	mdn	iqr	p-value
IS	0.311	0.463	0.074	0.167	<0.001*
FN	0.047	0.086	0.000	0.119	0.707
FP	0.320	0.516	0.286	0.602	0.035*
mdnISI (ms)	35.604	31.089	95.396	143.984	<0.001*
stdISI (ms)	57.599	59.953	163.753	208.091	<0.001*
cvISI	0.993	0.187	1.016	0.233	0.923
FR (spks/s)	17.800	15.975	6.400	7.325	<0.001*
mdnSppA (uA)	79.933	38.248	27.328	11.527	<0.001*
AI	0.125	0.220	0.041	0.097	<0.001*
PI	0.583	0.819	2.238	3.889	<0.001*
PR	3.221	6.192	20.171	66.970	<0.001*
FF (ms)	1.633	0.568	1.237	0.325	<0.001*
BI	7.969	18.589	24.611	40.364	<0.001*
MBI (ms)	0.169	0.195	0.059	0.095	<0.001*
BR (burst/s)	1.400	1.300	0.300	0.700	<0.001*
mdnIBI (ms)	207.500	271.422	386.042	738.042	<0.001*
mdnBD (ms)	207.500	205.036	208.792	460.250	0.001*
mdnIBF (Hz)	$4.00\ 10^{-5}$	$2.89\ 10^{-5}$	$2.03\ 10^{-5}$	$3.16\ 10^{-5}$	<0.001*

Table B.3: Spike features with IS > 0.5 from filtered and artifact free filtered 10 seconds MER, both with and without artifact detection. Median and interquartile range and number of neurons (n) in each sample. Statistical test used are T-test [1] or Wilcoxon Rank Sum test [2].

| Spike related features | Filtered MER with IS>0.5 | | | | | | Artifact-free filtered MER with IS>0.5 | | | | | |
| | STN (n=113) | | non-STN (n=21) | | | | STN (n=55) | | non-STN (n=20) | | | |
	mdn	iqr	mdn	iqr	p-value	test	mdn	iqr	mdn	iqr	p-value	test
IS	0.875	0.362	0.640	0.218	0.002*	[1]	0.942	0.306	0.603	0.229	<0.001*	[2]
FN	0.021	0.053	0.014	0.060	0.005*	[2]	0.011	0.034	0.000	0.039	<0.001*	[2]
FP	0.032	0.115	0.083	0.128	0.071	[1]	0.013	0.109	0.048	0.108	<0.001*	[2]
mdnISI (ms)	32.688	24.188	31.208	30.839	0.005*	[2]	33.375	33.448	55.021	39.719	<0.001*	[2]
stdISI (ms)	61.946	62.808	62.589	92.182	0.005*	[2]	57.425	54.248	73.369	37.466	<0.001*	[2]
cvISI	1.027	0.334	1.043	0.741	0.005*	[2]	0.987	0.206	0.931	0.064	0.011*	[1]
FR (spks/s)	17.300	13.725	14.800	16.375	0.005*	[2]	16.900	18.000	12.500	6.150	0.001*	[1]
mdnSppA (uA)	89.243	40.822	33.100	15.937	<0.001*	[1]	87.963	38.383	30.847	6.002	<0.001*	[1]
AI	0.135	0.237	0.231	0.226	0.005*	[2]	0.152	0.279	0.251	0.128	<0.001*	[2]
PI	0.577	0.653	0.554	0.843	0.005*	[2]	0.588	0.956	1.102	0.832	<0.001*	[2]
PR	3.230	4.691	3.856	4.954	0.005*	[2]	3.128	5.744	5.701	6.099	<0.001*	[2]
FF (ms)	1.667	0.588	1.927	1.053	0.005*	[2]	1.650	0.703	1.349	0.164	<0.001*	[1]
BI	7.430	14.863	4.338	11.561	0.005*	[2]	6.601	8.694	3.979	2.191	<0.001*	[2]
MBI (ms)	0.157	0.213	0.095	0.328	0.006*	[2]	0.129	0.260	0.017	0.078	<0.001*	[2]
BR (burst/s)	1.200	1.200	0.800	0.925	0.012*	[1]	1.200	1.575	1.200	1.050	0.150	[1]
mdnIBI (ms)	233.708	274.042	339.875	447.859	0.006*	[2]	232.500	258.354	239.958	161.385	<0.001*	[2]
mdnBD (ms)	235.063	336.073	217.583	361.573	0.329	[1]	251.708	482.500	219.958	269.938	<0.001*	[2]
mdnIBF (Hz)	$3.81\ 10^{-5}$	$2.94\ 10^{-5}$	$3.75\ 10^{-5}$	$2.05\ 10^{-5}$	0.005*	[2]	$3.89\ 10^{-5}$	$3.86\ 10^{-5}$	$2.78\ 10^{-5}$	$1.87\ 10^{-5}$	0.002*	[1]

Appendix C

Classification performance measures

Table C.1: Classification performance results using 2 PCs, all MER labelled as STN (285) vs. non STN (420) and 10-cross validation to 6 different SVM models.

SVM model	10-CV using 2 PCs and STN labels					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Confusion matrix	
Linear	83.55	72.28	91.19	84.77	206	79
					37	383
Quadratic	50.21	68.77	37.62	42.79	196	89
					262	158
Cubic	36.74	57.54	22.62	33.54	164	121
					325	95
Fine Gaussian	85.25	73.68	93.10	87.87	210	75
					29	391
Medium Gaussian	83.97	70.53	93.10	87.39	201	84
					29	391
Coarse Gaussian	82.84	69.12	92.14	85.65	197	88
					33	387

Table C.2: Classification performance results using 4 PCs, all MER labelled as STN (285) vs. non STN (420) and 10-cross validation to 6 different SVM models.

| SVM model	10-CV using 4 PCs and STN labels					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Confusion matrix	
Linear	85.39	75.44	92.14	86.69	215	70
					33	387
Quadratic	86.52	74.04	95.00	90.95	211	74
					21	399
Cubic	69.08	78.95	62.38	58.75	225	60
					158	262
Fine Gaussian	86.52	81.40	90.00	84.67	232	53
					42	378
Medium Gaussian	86.24	73.68	94.76	90.52	210	75
					22	398
Coarse Gaussian	84.96	67.72	96.67	93.24	193	92
					14	406

Appendix D

Features in STN subdivisions

Table D.1: Results of time and frequency related features for STN subdivisions based on expert labels. P-value obtained through T-test for 10 second segments both filtered and free of artifacts. N is the number of cells in each sample, mdn represents median values and iqr interquartile range for each feature.

| Time and frequency features | Filtered MER | | | | | Artifact-free filtered MER | | | | |
| | dorsal STN (n=73) | | ventral STN (n=74) | | | dorsal STN (n=54) | | ventral STN (n=51) | | |
	mdn	*iqr*	*mdn*	*iqr*	p-value	*mdn*	*iqr*	*mdn*	*iqr*	p-value
MAV	9.491	3.377	10.630	4.677	0.191	10.689	4.550	11.643	4.445	0.362
MAVS	0.000	0.028	-3.57E-05	0.026	0.414	0.000	0.026	-5.23E-05	0.021	0.650
ZC	0.170	0.019	0.167	0.020	0.245	0.172	0.021	0.169	0.013	0.061
RMS	0.272	0.106	0.305	0.142	0.204	0.314	0.138	0.334	0.152	0.356
WL	9668.139	4728.871	10028.708	5360.854	0.271	11580.966	4566.076	12130.117	5597.096	0.623
VAR	151.246	126.984	190.278	181.592	0.198	201.701	172.175	229.285	199.914	0.368
CF	15.814	18.104	17.088	21.977	0.268	18.397	26.089	19.894	33.265	0.385
TE	61.409	71.506	64.783	92.096	0.249	92.513	76.670	101.380	99.720	0.514
SDN	12.211	4.162	13.748	5.188	0.135	13.548	5.746	14.910	5.170	0.319
SPC (Hz)	0.177	0.018	0.174	0.019	0.262	0.180	0.015	0.178	0.020	0.045*
SF	0.631	0.114	0.628	0.152	0.396	0.652	0.082	0.672	0.122	0.721
mFRQ (Hz)	696.634	125.652	677.251	155.897	0.356	726.980	120.888	704.532	146.195	0.063
mdnFRQ (Hz)	118.000	36.750	110.500	36.000	0.306	126.000	34.000	121.500	35.000	0.073
MC1	18.566	0.733	18.785	0.943	0.247	18.863	0.940	19.002	0.968	0.358
MC2	62.041	3.066	62.547	3.594	0.318	63.190	3.494	63.638	3.571	0.447
MC3	-0.640	1.431	-0.427	1.414	0.347	-1.004	1.240	-0.768	1.486	0.059
MC4	-1.287	0.521	-1.277	0.668	0.738	-1.434	0.415	-1.491	0.565	0.808
MC5	-0.166	0.175	-0.146	0.151	0.361	-0.165	0.154	-0.152	0.123	0.664
MC6	-0.434	0.164	-0.455	0.239	0.933	-0.488	0.148	-0.506	0.149	0.567
MC7	-0.315	0.062	-0.300	0.066	0.238	-0.323	0.071	-0.308	0.074	0.315
MC8	-0.124	0.067	-0.131	0.109	0.994	-0.138	0.077	-0.137	0.069	0.814
MC9	-0.207	0.051	-0.200	0.051	0.301	-0.209	0.039	-0.204	0.039	0.403
MC10	-0.028	0.057	-0.039	0.055	0.474	-0.044	0.045	-0.049	0.052	0.628
MC11	-0.101	0.050	-0.093	0.060	0.712	-0.100	0.052	-0.095	0.033	0.934
MC12	0.011	0.064	0.004	0.067	0.776	0.012	0.050	0.005	0.043	0.546

Table D.2: Statistic results of spike related features for STN subdivisions (based on expert labelling) in filtered free of artifact segments.

Spike related features	Filtered MER				
	dorsal STN (n=168)		ventral STN (n=171)		
	mdn	*iqr*	*mdn*	*iqr*	p-value
IS	0.331	0.496	0.312	0.445	0.490
FN	0.048	0.078	0.052	0.098	0.263
FP	0.299	0.489	0.333	0.543	0.485
mdnISI (ms)	30.917	33.938	37.188	26.120	0.439
stdISI (ms)	54.962	63.677	60.420	59.494	0.977
cvISI	1.007	0.233	0.993	0.174	0.765
FR (spks/s)	19.750	16.800	16.500	15.125	0.178
mdnSppA (uA)	75.648	39.976	81.071	37.596	0.168
AI	0.118	0.185	0.143	0.247	0.852
PI	0.520	0.922	0.658	0.726	0.908
PR	2.816	6.716	3.482	5.358	0.390
FF (ms)	1.684	0.590	1.614	0.552	0.897
BI	8.460	16.429	6.973	20.386	0.904
MBI (ms)	0.187	0.185	0.153	0.204	0.208
BR (burst/s)	1.400	1.450	1.300	1.200	0.201
mdnIBI (ms)	196.063	279.292	209.833	272.271	0.363
mdnBD (ms)	194.010	195.677	235.625	250.094	0.205
mdnIBF (Hz)	$4.25\ 10^{-5}$	$2.84\ 10^{-5}$	$3.60\ 10^{-5}$	$2.63\ 10^{-5}$	0.407
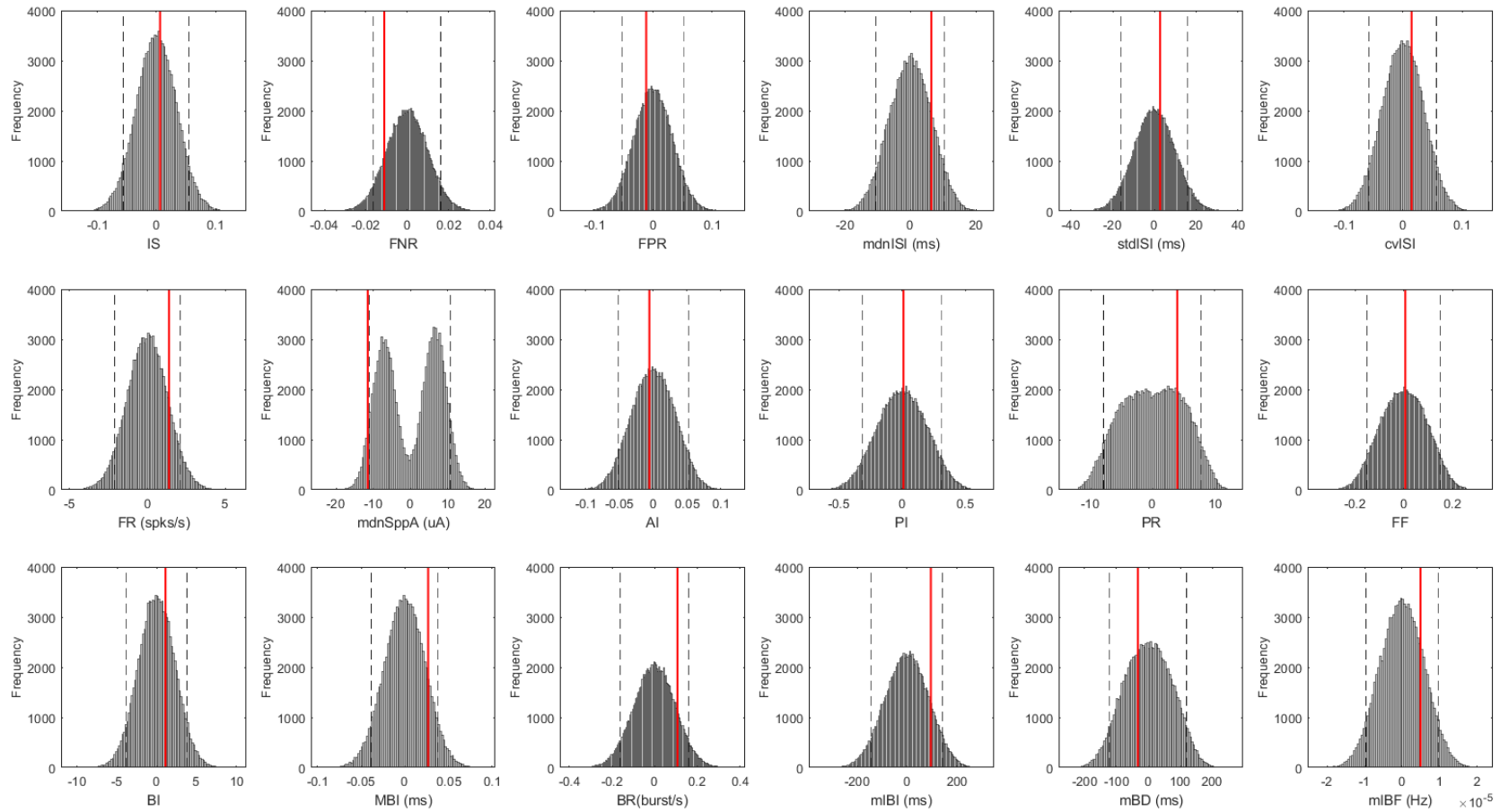
Figure D.1: Random Permutation tests between STN subdivisions for spike related features (based on expert labels) using 10 second segments filtered.
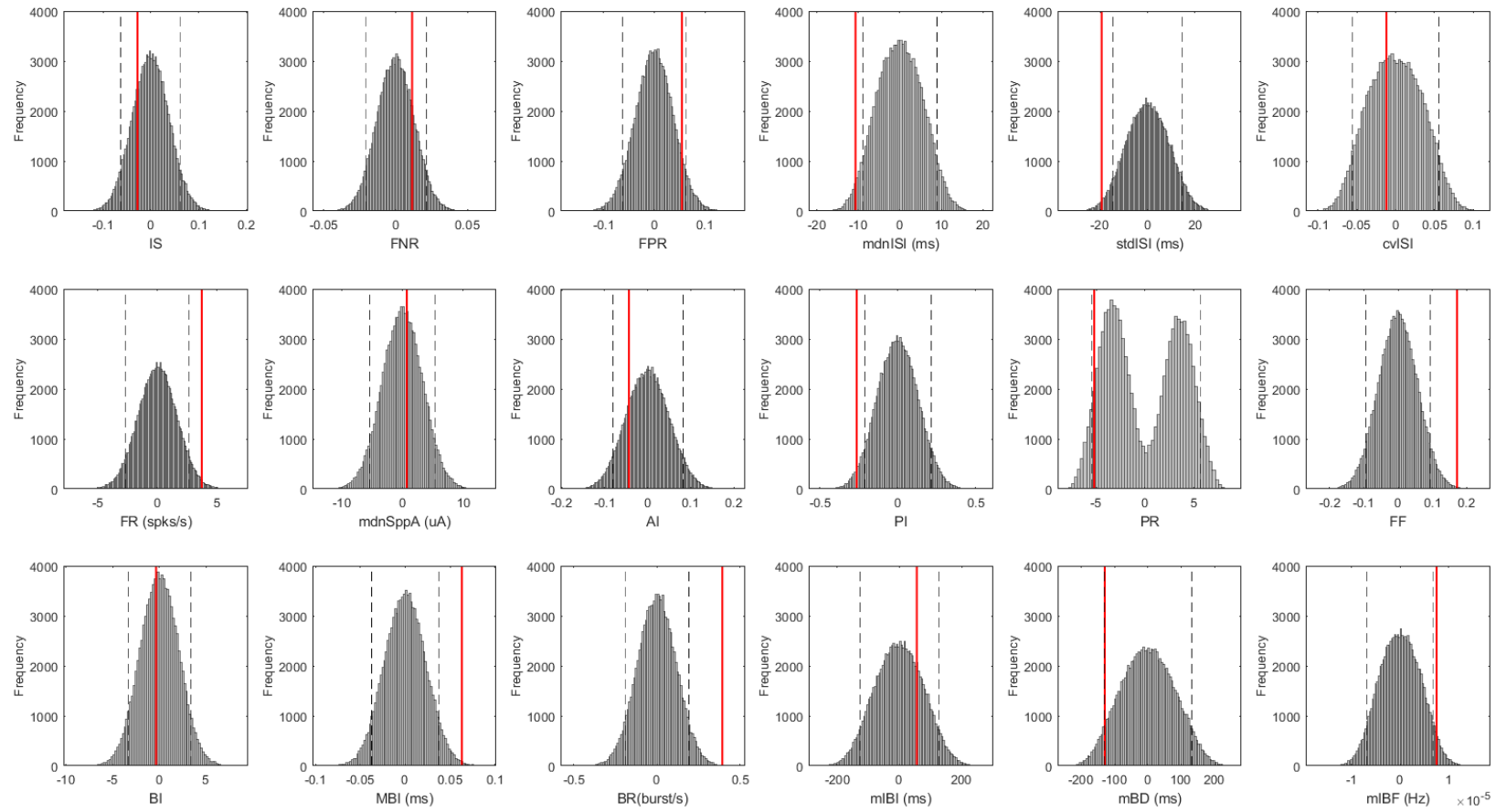
Figure D.2: Random Permutation tests between STN subdivisions for spike related features (based on expert labels) using 10 second segments free of artifacts.

Table D.3: Statistical results of time and frequency related features for STN subdivisions refined with lead reconstruction, both with and without artifact detection in 10 second segments. Median and interquartile range and number of neurons (n) in each sample. Statistical test used are T-test [1] or Wilcoxon Rank Sum test [2]; not shown for artifact-free MER since after testing normality, T-test was implemented for all features.

| Time and frequency features | Filtered MER | | | | | | Artifact-free filtered MER | | | | |
| | motor STN (n=38) | | non-motor STN (n=25) | | | | motor STN (n=36) | | non-motor STN (n=25) | | |
	mdn	iqr	mdn	iqr	p-value	test	mdn	iqr	mdn	iqr	p-value
MAV	12.706	3.557	12.768	2.984	0.449	[1]	12.979	2.649	12.510	3.504	0.850
MAVS	0.000	0.028	0.003	0.028	0.897	[2]	0.002	0.049	0.000	0.026	0.460
ZC	0.168	0.010	0.170	0.007	0.104	[1]	0.168	0.009	0.171	0.007	0.091
RMS	0.386	0.136	0.372	0.088	0.911	[1]	0.407	0.152	0.369	0.105	0.264
WL	13421.669	2618.334	13515.693	2470.441	0.103	[1]	13706.386	1749.290	12932.094	2869.635	0.661
VAR	305.938	216.660	284.131	138.007	0.643	[1]	339.924	257.226	279.061	161.329	0.196
CF	31.274	40.532	32.376	20.609	0.199	[1]	35.848	46.131	30.450	22.372	0.043*
TE	135.785	78.999	125.563	60.255	0.755	[1]	139.803	65.516	116.153	65.402	0.447
SPN	16.003	2.421	16.345	3.549	0.143	[1]	16.359	2.247	16.066	3.586	0.690
SPCENT	0.180	0.017	0.181	0.008	0.098	[1]	0.178	0.016	0.182	0.007	0.056
SF	0.666	0.070	0.680	0.029	0.039*	[1]	0.655	0.054	0.679	0.033	0.589
mFRQ	704.946	115.304	720.874	55.555	0.091	[1]	699.760	111.876	730.019	64.120	0.066
mdnFRQ	125.500	22.000	129.000	13.000	0.092	[1]	125.500	24.000	130.500	18.000	0.074
MC1	19.275	0.731	19.217	0.581	0.744	[1]	19.377	0.746	19.193	0.595	0.435
MC2	64.573	2.527	64.382	2.030	0.631	[1]	64.900	2.315	64.274	2.050	0.503
MC3	-0.941	1.260	-1.090	0.573	0.071	[1]	-0.840	1.156	-1.164	0.649	0.058
MC4	-1.562	0.426	-1.608	0.106	0.204	[1]	-1.616	0.356	-1.591	0.159	0.879
MC5	-0.113	0.092	-0.137	0.046	0.402	[1]	-0.137	0.097	-0.133	0.070	0.419
MC6	-0.522	0.062	-0.537	0.059	0.897	[2]	-0.513	0.073	-0.537	0.063	0.583
MC7	-0.307	0.060	-0.300	0.051	0.694	[1]	-0.303	0.044	-0.306	0.056	0.635
MC8	-0.170	0.059	-0.171	0.064	0.610	[1]	-0.167	0.049	-0.166	0.042	0.803
MC9	-0.205	0.033	-0.191	0.029	0.264	[1]	-0.201	0.037	-0.185	0.033	0.428
MC10	-0.054	0.035	-0.043	0.037	0.482	[1]	-0.054	0.022	-0.029	0.025	0.048*
MC11	-0.096	0.037	-0.083	0.034	0.335	[1]	-0.091	0.034	-0.088	0.032	0.754
MC12	-0.015	0.042	-0.005	0.037	0.130	[1]	-0.014	0.030	-0.008	0.027	0.863

Table D.4: Statistical results of spike related features for STN subdivisions refined with lead reconstruction, both with and without artifact detection in 10 second segments. Median and interquartile range and number of neurons (n) in each sample.

| Spike related features | Filtered MER | | | | | Artifact-free filtered MER | | | | |
| | motor STN (n=101) | | non-motor STN (n=64) | | | motor STN (n=98) | | non-motor STN (n=49) | | |
	mdn	*iqr*	*mdn*	*iqr*	p-value	*mdn*	*iqr*	*mdn*	*iqr*	p-value
IS	0.219	0.588	0.215	0.479	0.405	0.228	0.436	0.194	0.257	0.274
FN	0.051	0.092	0.055	0.106	0.422	0.049	0.089	0.060	0.092	0.778
FP	0.458	0.567	0.535	0.714	0.424	0.429	0.476	0.582	0.496	0.018*
mdnISI (ms)	37.188	33.828	34.958	53.500	0.261	35.146	32.313	29.667	23.953	0.308
stdISI (ms)	60.185	48.050	51.142	67.699	0.367	56.267	53.091	41.874	42.294	0.169
cvISI	0.976	0.150	0.962	0.119	0.527	0.998	0.143	0.991	0.156	0.418
FR (spks/s)	16.600	14.550	19.350	23.000	0.224	17.050	16.000	23.100	16.550	0.300
mdnSppA (uA)	92.702	30.994	91.620	37.297	0.947	87.570	33.659	92.935	28.528	0.472
AI	0.112	0.185	0.112	0.153	0.075	0.120	0.241	0.149	0.192	0.874
PI	0.660	0.936	0.523	1.418	0.357	0.580	0.957	0.429	0.645	0.143
PR	3.221	5.959	2.600	9.412	0.150	3.138	5.579	1.894	4.012	0.247
FF	1.559	0.580	1.640	0.637	0.467	1.611	0.638	1.740	0.582	0.364
BI	8.926	17.483	8.914	14.932	0.791	8.307	12.917	6.719	11.158	0.573
MBI	0.151	0.196	0.175	0.229	0.340	0.168	0.203	0.218	0.223	0.384
BR (bursts/s)	1.300	1.300	1.700	1.700	0.581	1.450	1.300	1.700	1.500	0.440
mdnIBI (ms)	234.646	286.281	165.458	269.833	0.968	171.042	224.786	173.792	212.729	0.825
mdnBD (ms)	207.500	262.760	208.000	161.500	0.160	225.458	205.208	201.875	157.271	0.294
mdnIBF (Hz)	$3.61\ 10^{-5}$	$2.88\ 10^{-5}$	$4.42\ 10^{-5}$	$2.99\ 10^{-5}$	0.840	$3.67\ 10^{-5}$	$2.93\ 10^{-5}$	$4.30\ 10^{-5}$	$2.72\ 10^{-5}$	0.294